

# Towards Accurate Cuffless Blood Pressure Estimation with Reduced Calibration Requirements

by

Shan He

Thesis submitted to the University of Ottawa  
in partial fulfillment of the requirements for the  
Doctor of Philosophy in Electrical and Computer Engineering

School of Electrical Engineering and Computer Science  
Faculty of Engineering  
University of Ottawa

© Shan He, Ottawa, Canada, 2025

## **Examining Committee**

The following served on the Examining Committee for this thesis.

External Member: Mukkamala Ramakrishna  
Professor, Swanson School of Engineering  
University of Pittsburgh

Carleton Member: Sreeraman Rajan  
Professor, Department of Systems and Computer Engineering  
Carleton Univeristy

Internal Member(s): Voicu Groza  
Professor, School of Electrical Engineering & Computer Science  
University of Ottawa

Internal Member(s): Changjian Shui  
Assistant Professor, School of Electrical Engineering & Computer Science  
University of Ottawa

Supervisor(s): Miodrag Bolic  
Professor, School of Electrical Engineering & Computer Science  
University of Ottawa

## **Declaration of Authorship**

I hereby certify that this thesis is entirely my own original work except where otherwise indicated. I am aware of the University of Ottawa regulations concerning plagiarism, including those regarding consequent disciplinary actions. Any use of the works of any other author, in any form, is properly acknowledged at their point of use.

## Copyright Compliance Declaration

This thesis incorporates material from four research articles authored by Shan He in collaboration with Hilmi R. Dajani and Miodrag Bolic. Written permissions from all co-authors have been obtained for the inclusion of these works. The use of each article complies with the copyright and licensing policies of the respective publishers, as outlined below:

1. He, S., Dajani, H. R., & Bolic, M. (2021). "Novel Cuffless Blood Pressure Estimation Method Using a Bayesian Hierarchical Model." *Proceedings of the 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*.
  - This material is included in the thesis in accordance with IEEE's self-archiving policy. Only the accepted manuscript (author's version) is reproduced, not the final published PDF. The IEEE copyright notice is included.
2. He, S. & Bolic, M. (2024). "Novel Metrics for Tracking Blood Pressure Changes in Continuous Cuffless Blood Pressure Estimations." *Scientific Reports (Nature Portfolio)*.
  - This article is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0). The publisher's version may be reproduced in full, provided that proper attribution is given.
3. He, S. & Bolic, M. (2025). "Analyzing and Processing the Underrated Effect of Respiration for Cuffless Blood Pressure Estimation."
  - This manuscript is original work by the authors. As it has not yet been published, the authors retain full copyright and grant permission for its inclusion in this thesis.
4. He, S. & Bolic, M. (2025). "Improving Generalization and Personalization in Cuffless Blood Pressure Estimation via Trend-Aware Time-Series Learning and Few-Shot Domain Adaptation."
  - This manuscript is original work by the authors. As it has not yet been published, the authors retain full copyright and grant permission for its inclusion in this thesis.

All co-authors (Shan He, Hilmi R. Dajani, and Miodrag Bolic) have reviewed and approved the inclusion of these works in this thesis.



## Abstract

Cuffless Blood Pressure (BP) estimation using Electrocardiogram (ECG) and Photoplethysmogram (PPG) features is a promising alternative to cuff-based devices for continuous BP monitoring. However, its practical deployment faces several critical challenges: Respiration Modulation (RM) introduces confounding fluctuations into both physiological features and BPs, degrading estimation accuracy and trend tracking; existing models often fail to accurately track BP trends, which is crucial for long-term monitoring; poor model generalization to new subjects leads to performance degradation; personalization strategies are inefficient, typically requiring a large number of calibration measurements; and evaluation metrics are inadequate, primarily focusing on pointwise accuracy while neglecting trend tracking.

To address these limitations, this thesis proposes a comprehensive framework. First, the often-overlooked impact of RM on features and BP is mitigated using Variational Mode Decomposition (VMD) for pre-processing, enhancing accuracy and trend tracking. Second, a novel lightweight estimation model, TCN-SENet-KAN, coupled with a customized shape-oriented Soft Dynamic Time Warping (sDTW) loss function, is proposed to improve BP trend tracking and model generalization. Third, a robust three-step personalization framework is developed, leveraging Domain Adversarial Neural Network (DANN) pre-training for subject-invariant features, followed by sDTW-guided fine-tuning and Procrustes analysis for post-hoc alignment. Finally, a composite similarity metric is introduced to jointly assess both distance and trend accuracy for continuous monitoring scenarios.

The proposed framework achieves Mean Absolute Error (MAE) of 4.34 mmHg for SBP and 2.70 mmHg for DBP using only two calibrations, meeting the IEEE 1708a-2019 standard for continuous BP monitoring and qualifying as “Grade A” performance. Furthermore, the method demonstrates robust BP trend tracking with median Pearson correlation coefficients of 0.781 (SBP) and 0.603 (DBP) across 32 subjects from the BCG BP dataset. Overall, this comprehensive framework significantly advances the practical viability of cuffless BP estimation for real-world applications.

## **Acknowledgements**

I want to give my deepest thanks to my supervisor, Dr. Miodrag Bolic. His guidance, support, and belief in me were essential throughout my PhD journey. I truly appreciate all his help. My thanks also go to Dr. Hilmi Dajani for his helpful feedback and time.

To my wonderful wife – thank you with all my heart. Your constant love and support kept me going through everything. I honestly couldn't have finished this without you by my side. And to my parents and family – thank you for your endless love and encouragement. Your support and sacrifices truly made this PhD possible. I'm so grateful for all of you.

# Table of Contents

|                                                                                            |            |
|--------------------------------------------------------------------------------------------|------------|
| <b>List of Tables</b>                                                                      | <b>xiv</b> |
| <b>List of Figures</b>                                                                     | <b>xvi</b> |
| <b>Abbreviations</b>                                                                       | <b>xxv</b> |
| <b>1 Introduction</b>                                                                      | <b>1</b>   |
| 1.1 Motivations . . . . .                                                                  | 2          |
| 1.1.1 The Overlooked Role of Respiration in Feature-Based Cuffless BP Estimation . . . . . | 6          |
| 1.1.2 Unsatisfactory Trend Tracking Performance . . . . .                                  | 7          |
| 1.1.3 Poor Model Generalization and Inefficient Personalization . . . . .                  | 7          |
| 1.1.4 Inadequate Evaluation Metrics . . . . .                                              | 8          |
| 1.2 Objectives . . . . .                                                                   | 8          |
| 1.3 Contributions . . . . .                                                                | 9          |
| 1.4 Outline of the Thesis . . . . .                                                        | 10         |
| <b>2 Literature Review</b>                                                                 | <b>11</b>  |
| 2.1 Review of Current BP Measurement Techniques and Devices . . . . .                      | 11         |
| 2.1.1 Invasive BP Measurement Techniques . . . . .                                         | 11         |
| 2.1.2 Noninvasive BP Measurement Techniques . . . . .                                      | 11         |

|         |                                                                     |    |
|---------|---------------------------------------------------------------------|----|
| 2.1.2.1 | Auscultation . . . . .                                              | 12 |
| 2.1.2.2 | Oscillometry . . . . .                                              | 13 |
| 2.1.2.3 | Volume Clamping . . . . .                                           | 13 |
| 2.1.2.4 | Tonometry . . . . .                                                 | 15 |
| 2.1.3   | New Wearable Cuff-based BP Devices . . . . .                        | 15 |
| 2.1.4   | Calibration for Conventional BP Devices . . . . .                   | 16 |
| 2.1.5   | Limitations of Conventional Methods . . . . .                       | 17 |
| 2.2     | Cuffless BP Estimation Techniques . . . . .                         | 17 |
| 2.2.1   | Ultrasound-Based BP Estimation . . . . .                            | 18 |
| 2.2.2   | Bioimpedance-Based BP Estimation . . . . .                          | 18 |
| 2.2.3   | Computer Vision-Based BP Estimation (rPPG) . . . . .                | 18 |
| 2.2.4   | Ballistocardiogram (BCG)-Based BP Estimation . . . . .              | 19 |
| 2.2.5   | Physiological Signal (ECG and/or PPG)-Based BP Estimation . . . . . | 19 |
| 2.2.6   | Discussion . . . . .                                                | 21 |
| 2.3     | Review of ECG and/or PPG Based Studies . . . . .                    | 21 |
| 2.3.1   | Predetermined Feature-Based Methods (feat2bp) . . . . .             | 21 |
| 2.3.1.1 | PTT/PAT Based Method . . . . .                                      | 21 |
| 2.3.1.2 | Pulse Wave Analysis (PWA) Based Methods . . . . .                   | 23 |
| 2.3.2   | Non-predetermined Feature-Based Methods (sig2bp) . . . . .          | 25 |
| 2.3.3   | ABP Waveform Prediction (sig2sig) . . . . .                         | 27 |
| 2.3.4   | Calibration for Cuffless BP Estimation . . . . .                    | 27 |
| 2.3.5   | Current Cuffless BP Devices . . . . .                               | 29 |
| 2.4     | Limitations of Current Cuffless BP Studies . . . . .                | 29 |
| 2.4.1   | Generalization and Personalization . . . . .                        | 29 |
| 2.4.1.1 | Inadequate Model Generalization . . . . .                           | 30 |
| 2.4.1.2 | Few-Shot Personalization . . . . .                                  | 34 |
| 2.4.2   | Unsatisfactory BP Trend Tracking Performance . . . . .              | 34 |
| 2.5     | Conclusion and Discussion . . . . .                                 | 37 |

|          |                                                                                                       |           |
|----------|-------------------------------------------------------------------------------------------------------|-----------|
| <b>3</b> | <b>Respiratory Effect on Cuffless BP Estimation</b>                                                   | <b>41</b> |
| 3.1      | Prior Insights on Respiratory Effects in Cuffless BP Estimation . . . . .                             | 41        |
| 3.2      | Data, Pre-Processing and Feature Extraction . . . . .                                                 | 43        |
| 3.2.1    | Experiment Data, Signal Processing and Feature Extraction . . . . .                                   | 43        |
| 3.3      | Respiration Modulation . . . . .                                                                      | 44        |
| 3.3.1    | Frequency Bands and Spectral Analysis . . . . .                                                       | 44        |
| 3.3.1.1  | Spectral bands from 0 to 0.04 Hz . . . . .                                                            | 45        |
| 3.3.1.2  | Spectral Band from 0.04 to 0.15 Hz . . . . .                                                          | 45        |
| 3.3.1.3  | Spectral Band from 0.15 to 0.4 Hz . . . . .                                                           | 45        |
| 3.3.2    | Mechanisms and Previous Studies of Respiratory Modulation . . . . .                                   | 46        |
| 3.3.3    | The Complexity of Respiration Modulation . . . . .                                                    | 48        |
| 3.3.3.1  | RM Fluctuations Are Enhanced by Upright Posture . . . . .                                             | 48        |
| 3.3.3.2  | The Amplitude of RM Fluctuations Increases With De-<br>creasing RR . . . . .                          | 49        |
| 3.3.3.3  | The Frequency of RM Fluctuations Might Be a Function of<br>RR . . . . .                               | 49        |
| 3.3.3.4  | The Relation Between the RM Fluctuations and the Res-<br>piratory Phase is a Function of RR . . . . . | 49        |
| 3.3.4    | The General Existence of RM on a Wide Variety of Physiological<br>Features and BPs . . . . .          | 50        |
| 3.3.5    | Separating RM from Physiological Features and BPs . . . . .                                           | 51        |
| 3.3.6    | Respiration Modulation Ratio (RMR) . . . . .                                                          | 53        |
| 3.4      | Effect of Respiration on Cuffless BP Estimation . . . . .                                             | 55        |
| 3.4.1    | Impact of RM on Feature Selection . . . . .                                                           | 55        |
| 3.4.2    | Impact of RM for Cuffless BP Estimation Accuracy . . . . .                                            | 55        |
| 3.5      | Conclusion and Discussion . . . . .                                                                   | 57        |

|           |                                                                                                            |           |
|-----------|------------------------------------------------------------------------------------------------------------|-----------|
| <b>4</b>  | <b>Novel Metrics for Tracking Blood Pressure Changes in Continuous Cuffless Blood Pressure Estimations</b> | <b>60</b> |
| 4.1       | Conventional Performance Metrics for Cuffless BP Evaluation . . . . .                                      | 60        |
| 4.1.1     | Distance Metrics . . . . .                                                                                 | 60        |
| 4.1.2     | Correlation Coefficient . . . . .                                                                          | 61        |
| 4.1.3     | A Composite Score to Represent Similarity in Mean, Variance and Correlation Coefficient . . . . .          | 63        |
| 4.2       | A New Comprehensive Performance Evaluation for Cuffless BP Estimation                                      | 63        |
| 4.2.1     | Pre-processing . . . . .                                                                                   | 64        |
| 4.2.2     | Segment Similarity Measures . . . . .                                                                      | 66        |
| 4.2.2.1   | Segment Distance Similarity . . . . .                                                                      | 66        |
| 4.2.2.2   | Segment Trend Similarity . . . . .                                                                         | 67        |
| 4.2.2.3   | Segment Composite Similarity . . . . .                                                                     | 68        |
| 4.2.3     | Combining the Segment Similarity Measures (global evaluation) . .                                          | 69        |
| 4.2.4     | Interpretation of the Proposed Metrics . . . . .                                                           | 70        |
| 4.3       | Experiments and Results . . . . .                                                                          | 71        |
| 4.3.1     | Applying the Proposed New Metrics to Simulated Data . . . . .                                              | 71        |
| 4.3.2     | Applying the Proposed New Metrics to Results of Two SOTA Models                                            | 73        |
| 4.4       | Conclusion and Discussion . . . . .                                                                        | 75        |
| <b>5</b>  | <b>Methods to Improve Generalization of BP Estimation Models</b>                                           | <b>77</b> |
| 5.1       | Trend-Related Features for BP Estimation . . . . .                                                         | 77        |
| 5.2       | Cuffless Blood Pressure Estimation Models . . . . .                                                        | 84        |
| 5.2.1     | A TCN-SENet-KAN Model . . . . .                                                                            | 84        |
| 5.2.1.1   | Temporal Convolutional Network (TCN) Block . . . . .                                                       | 84        |
| 5.2.1.1.1 | Causal Convolutions . . . . .                                                                              | 85        |
| 5.2.1.1.2 | Dilated Convolutions . . . . .                                                                             | 85        |
| 5.2.1.1.3 | Residual Connections . . . . .                                                                             | 85        |

|           |                                                                |     |
|-----------|----------------------------------------------------------------|-----|
| 5.2.1.2   | Squeeze-and-Excitation Network (SENet) Block . . . . .         | 85  |
| 5.2.1.2.1 | Squeeze Operation . . . . .                                    | 86  |
| 5.2.1.2.2 | Excitation Operation . . . . .                                 | 87  |
| 5.2.1.3   | Recalibration . . . . .                                        | 87  |
| 5.2.2     | Kolmogorov-Arnold Network (KAN) Linear Layer . . . . .         | 87  |
| 5.2.3     | A Parallel Transformer-BiGRU Model with Attention Mechanisms . | 89  |
| 5.2.3.1   | Transformer Encoder Branch . . . . .                           | 89  |
| 5.2.3.1.1 | Multi-Head Self-Attention (MHSA): . . . . .                    | 90  |
| 5.2.3.1.2 | Position-wise Feed-Forward Network (FFN): . . . . .            | 90  |
| 5.2.3.2   | Bidirectional GRU (BiGRU) Branch with Global Attention         | 90  |
| 5.2.3.2.1 | Bidirectional GRU (BiGRU): . . . . .                           | 90  |
| 5.2.3.2.2 | Global Attention Mechanism: . . . . .                          | 91  |
| 5.2.3.3   | Cross-Attention Fusion . . . . .                               | 91  |
| 5.3       | Loss Function . . . . .                                        | 92  |
| 5.3.1     | Pointwise Distance-Based Loss Function . . . . .               | 92  |
| 5.3.1.1   | Mean Squared Error (MSE) . . . . .                             | 92  |
| 5.3.1.2   | Mean Absolute Error (MAE) . . . . .                            | 92  |
| 5.3.2     | A Shape-Oriented Loss Function . . . . .                       | 93  |
| 5.3.2.1   | Dynamic Time Warping (DTW) . . . . .                           | 93  |
| 5.3.2.2   | Soft Dynamic Time Warping (sDTW) . . . . .                     | 94  |
| 5.3.2.3   | Customized sDTW for Accurate BP Trend Tracking . . . . .       | 96  |
| 5.4       | Experiments . . . . .                                          | 96  |
| 5.4.1     | Bed-Based Ballistocardiography Dataset . . . . .               | 96  |
| 5.4.2     | Pre-processing . . . . .                                       | 98  |
| 5.4.3     | Fiducial Points Detection . . . . .                            | 98  |
| 5.4.3.1   | ECG Fiducial Point . . . . .                                   | 99  |
| 5.4.3.2   | PPG Fiducial Points . . . . .                                  | 100 |

|          |                                                                         |            |
|----------|-------------------------------------------------------------------------|------------|
| 5.4.3.3  | ABP Fiducial Points . . . . .                                           | 101        |
| 5.4.4    | Feature Extraction and Processing . . . . .                             | 101        |
| 5.4.5    | Implementation Details . . . . .                                        | 101        |
| 5.4.6    | Leave-One-Subject-Out Cross Validation . . . . .                        | 103        |
| 5.5      | Results . . . . .                                                       | 104        |
| 5.6      | Conclusion and Discussion . . . . .                                     | 105        |
| <b>6</b> | <b>Accurate BP Estimation with Few Calibrations</b>                     | <b>110</b> |
| 6.1      | A Bayesian Hierarchical Regression Model for Few-Shot BP Estimation . . | 111        |
| 6.1.1    | Theoretical Introduction . . . . .                                      | 111        |
| 6.1.2    | Implementation and Results . . . . .                                    | 113        |
| 6.1.3    | Limitations . . . . .                                                   | 113        |
| 6.2      | Problem Formalization for Few-Shot Adaptation . . . . .                 | 114        |
| 6.3      | Model Fine-Tuning for Subject-Specific Adaptation . . . . .             | 115        |
| 6.3.1    | Adaptation via Fine-Tuning . . . . .                                    | 115        |
| 6.4      | Meta-Learning for Subject-Specific Adaptation . . . . .                 | 116        |
| 6.4.1    | Inner Loop: Task-Specific Adaptation . . . . .                          | 116        |
| 6.4.2    | Outer Loop: Meta Optimization . . . . .                                 | 117        |
| 6.4.3    | Meta-Testing Stage . . . . .                                            | 118        |
| 6.5      | Domain Adaptation for Subject-Specific Adaptation . . . . .             | 118        |
| 6.5.1    | Adversarial Training . . . . .                                          | 119        |
| 6.5.2    | K-shot Adaptation for the Test Subject . . . . .                        | 121        |
| 6.6      | Statistical Post-Hoc Methods for Subject-Specific Adaptation . . . . .  | 122        |
| 6.6.1    | Affine Transformation . . . . .                                         | 122        |
| 6.6.2    | Procrustes Analysis . . . . .                                           | 123        |
| 6.7      | Experiments . . . . .                                                   | 124        |
| 6.8      | Results . . . . .                                                       | 125        |
| 6.9      | Conclusion and Discussion . . . . .                                     | 129        |



|          |                                                                                                            |            |
|----------|------------------------------------------------------------------------------------------------------------|------------|
| <b>7</b> | <b>Conclusions and Future Work</b>                                                                         | <b>133</b> |
| 7.1      | Conclusions . . . . .                                                                                      | 133        |
| 7.2      | Future Work . . . . .                                                                                      | 135        |
|          | <b>References</b>                                                                                          | <b>137</b> |
|          | <b>APPENDICES</b>                                                                                          | <b>160</b> |
| <b>A</b> | <b>Database, Signal Preprocessing and Feature Extraction</b>                                               | <b>161</b> |
| A.1      | Research Database . . . . .                                                                                | 161        |
| A.2      | Features Employed for Cuffless BP Estimation . . . . .                                                     | 162        |
| A.2.1    | PPG Features . . . . .                                                                                     | 162        |
| A.3      | Examples of Respiratory Effects on Extracted Features . . . . .                                            | 165        |
| A.4      | Supplementary Material for Novel Evaluation Metrics . . . . .                                              | 165        |
| A.4.1    | MLP Model for Example Shown in Figure 2.9(a) . . . . .                                                     | 165        |
| A.4.2    | Pearson Correlation Coefficient (PCC) and Concordance Correlation Coefficient (CCC) . . . . .              | 165        |
| A.4.3    | Re-implementation of SOTA Models . . . . .                                                                 | 168        |
| A.4.4    | Features Used for DeepRNN-4L . . . . .                                                                     | 173        |
| A.4.5    | Benchmark and Diagnose Estimation Algorithms . . . . .                                                     | 174        |
| A.4.6    | Detailed Interpretation of the Proposed metrics . . . . .                                                  | 179        |
| A.4.7    | Composite Similarity of Examples Shown in Figure 4.6 with Different $w_1$ and $w_2$ Combinations . . . . . | 182        |
| A.5      | Impact of Sampling Frequency . . . . .                                                                     | 182        |
| A.6      | Impact of Differential Filter Orders . . . . .                                                             | 184        |
| A.7      | Few-Shot Fine Tuning Strategies . . . . .                                                                  | 186        |

# List of Tables

|     |                                                                                                                                                                                                                                     |    |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1.1 | Statistics of hypertension in Canada (ages 20-79). . . . .                                                                                                                                                                          | 4  |
| 2.1 | Summary and comparison of the conventional BP measurement methods. .                                                                                                                                                                | 17 |
| 2.2 | Summary and comparison of the cuffless BP estimation methods. . . . .                                                                                                                                                               | 20 |
| 2.3 | Different mathematical models to calculate BP from PAT and RMSE for SBP and DBP estimations (unit: mmHg) . . . . .                                                                                                                  | 23 |
| 2.4 | Evaluation of BP estimation in the MIMIC-II UCI-ML dataset with regard to MD, STD and MAE (adapted from [110]). . . . .                                                                                                             | 32 |
| 2.5 | Mean and standard deviation of absolute error between estimated BPs and reference BPs using 4 different types of SOTA models with UCI dataset. .                                                                                    | 33 |
| 2.6 | Key aspects of validation protocols developed for cuff-based and cuffless BP devices. . . . .                                                                                                                                       | 35 |
| 2.7 | A summary of existing state-of-art cuffless BP estimation studies. . . . .                                                                                                                                                          | 39 |
| 3.1 | Mean and standard deviation of absolute error, and Pearson correlation coefficient ( $r$ ) between estimated BPs and reference BPs using two estimation models with originally extracted features and RM-suppressed features. . . . | 57 |
| 4.1 | Definition of distance metrics and correlation coefficients widely-used for cuffless BP estimation evaluation. . . . .                                                                                                              | 62 |
| 4.2 | Equation for calculating GCSM and definitions of each component. . . . .                                                                                                                                                            | 63 |
| 4.3 | Segment trend similarity for different scenarios . . . . .                                                                                                                                                                          | 68 |
| 4.4 | Equations of calculating temporally-normalized similarity metrics. . . . .                                                                                                                                                          | 69 |

|     |                                                                                                                                                                                                                                                                              |     |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 4.5 | Segment similarity measures of the example shown in Figure 4.4 proposed by existing works and this study and their temporally-normalized results for whole sequences comparison. . . . .                                                                                     | 70  |
| 4.6 | Conventional distance metrics, correlation coefficients, composite similarity measure GCSM and the proposed similarity metrics calculated between simulated BP estimations and reference BP. . . . .                                                                         | 72  |
| 4.7 | Conventional distance metrics, correlation coefficients, composite similarity measure GCSM and the proposed similarity metrics calculated between the estimated BPs and reference BPs using SOTA models from DailA BP dataset. . . . .                                       | 75  |
| 5.1 | Definition of PPG fiducial points, and steps to detect these fiducial points. . . . .                                                                                                                                                                                        | 80  |
| 5.2 | Definition of 69 features extracted from ECG, PPG and its derivatives with correlation coefficients ( $r$ ) to SBP and DBP with RM and without RM (results represented in median (iqr) among 32 subjects). . . . .                                                           | 81  |
| 5.3 | The mean and standard deviation of absolute error between reference and estimated BPs for all subjects, and the median (IQR) of the Pearson correlation coefficient between each subject's reference and estimated BPs using different models and loss functions. . . . .    | 104 |
| 6.1 | The mean and standard deviation of the absolute error between estimated SBP and reference SBP using different approaches(Note: it is impossible to conduct OLS regression one a single measurement). . . . .                                                                 | 113 |
| 6.2 | Mean and standard deviation of AE between the reference and estimated BPs for all 32 subjects using different subject-specific adaptation methods with different number of calibration samples . . . . .                                                                     | 128 |
| 6.3 | Median and interquartile range (IQR) of Pearson correlation coefficient ( $r$ ) between the reference and estimated BPs for each of the 32 subjects using different subject-specific adaptation methods with different number of calibration samples . . . . .               | 128 |
| 6.4 | Median and interquartile range (IQR) of the proposed distance, trend and composite similarities between the reference and estimated BPs for each of the 32 subjects using different adaptation method with different number of adaptation samples ( $k = 1, 2, 4$ ). . . . . | 128 |
| A.1 | Four widely used open access databases for cuffless BP estimation . . . . .                                                                                                                                                                                                  | 162 |

# List of Figures

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                |    |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1.1 | An example of peripheral BP waveform with indication of SBP, DBP and MBP. . . . .                                                                                                                                                                                                                                                                                                                                                                              | 2  |
| 1.2 | (a) Deaths caused by hypertension and other diseases per year globally [14] (data from "Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016" in The Lancet and WHO), (b) the amount of people with hypertension have it controlled globally [14] (data from "Global disparities of hypertension prevalence and control" in Circulation). . . . . | 3  |
| 1.3 | The overview of this thesis: the pipeline of the ECG and PPG feature based BP estimation model, four critical research problems identified by this thesis with their effects on practical application, along with the implications and contributions. . . . .                                                                                                                                                                                                  | 5  |
| 2.1 | An example of the auscultatory BP device (left, image adapted from [30]) and the methodology of auscultatory BP measurement (right, image adapted from [31]). . . . .                                                                                                                                                                                                                                                                                          | 12 |
| 2.2 | Demonstrations of oscillometry, volume-clamp and tonometry BP measurement methods (figure adapted from [39]). . . . .                                                                                                                                                                                                                                                                                                                                          | 14 |
| 2.3 | Commercialized wearable continuous BP monitors, (a) Omron HeartGuide; (b) Huawei Watch D; (c) Biobeat Tech BB-613WP; (d) Samsung Galaxy Active2. . . . .                                                                                                                                                                                                                                                                                                       | 16 |
| 2.4 | A demonstration of fiducial points of ECG, PPG, the first derivative of PPG (VPG) and the second derivative of PPG (APG), as well as PATs which can be extracted as the time intervals between ECG R peaks and PPG onsets, max_slope points and systolic peaks (x axis: time (s), y axis: amplitude). . . . .                                                                                                                                                  | 24 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |    |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2.5  | (a) an example of a regular PPG pulse with distinct dicrotic notch and diastolic peak; (b) an example of PPG pulse without distinct dicrotic notch and diastolic peak. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 25 |
| 2.6  | Examples of PPG and ABP morphology similarity analysis. $r$ refers to the morphology correlation between PPG and ABP waveforms (adapted from [94], y axis: normalized amplitude). . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                | 27 |
| 2.7  | An example of the BP waveform modelling using PPG waveform demonstrated in [95]. (a). the PPG waveform used as input of the model; (b). the predicted ABP waveform using <b>approximation network</b> (red), the predicted waveform using <b>refinement network</b> (blue) and the ground truth ABP waveform (green). . . . .                                                                                                                                                                                                                                                                                                            | 28 |
| 2.8  | Four SOTA models with different input-output structures for cuffless BP estimation. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 33 |
| 2.9  | (a) An example of the proposed BP model unable to track BP changes but the overall MAD and SDAD is low. This segment is 543 seconds, $MAD \pm SDAD$ for SBP is $(3.77 \pm 2.51)$ mmHg. (b) 24-hour ambulatory brachial SBP pattern, stratified by age. Solid lines are mean values and dash lines are 95% confidence intervals (reproduction of the middle figure of Figure 2 demonstrated in work [121], author authorized the use of this figure in this study). . . . .                                                                                                                                                               | 37 |
| 2.10 | (a) BP waveform when hand grips were performed by squeezing a rolled towel at maximum voluntary contraction (0-60s) and at rest (after 60 s), (b) BP waveform when the subject was walking for 1 minute, (c) BP waveform when Valsalva maneuver was conducted, the BP reached maximal during Phase 1, (d) an example of BP changes extracted from <b>DailA BP</b> when a subject changed posture from sitting to standing, the lower track is the orientation of subject's knee (unit: degrees), $0^\circ$ indicates thighs are parallel to ground (sitting) $90^\circ$ indicates thighs are perpendicular to ground (standing). . . . . | 38 |
| 3.1  | (a) A demonstration of baseline wandering (BW) caused by respiration on PPG waveform, (b) a 4th order Chebyshev type II bandpass filter with low cut-off frequency of 0.5 Hz and high cut-off frequency of 12 Hz applied to PPG waveform to remove BW. . . . .                                                                                                                                                                                                                                                                                                                                                                           | 42 |
| 3.2  | (a) respiration can cause SBP fluctuations up to 3 to 6 mmHg with normal breathing depth and rate, (b) respiration can cause up to 15 to 20 mmHg with shallow and deep breathing. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                | 47 |

|      |                                                                                                                                                                                                                                                                     |    |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.3  | A demonstration of RSA, and power spectrum of HR which exhibits two prominent peaks of LF and HF bands. . . . .                                                                                                                                                     | 47 |
| 3.4  | Demonstration of amplitude of RM fluctuations may vary with RR, the RM components extracted from PATs of both examples (left and right) are demonstrated in a same amplitude scale. . . . .                                                                         | 49 |
| 3.5  | Demonstration of phase of RM fluctuations may vary with RR (y axis: amplitude). . . . .                                                                                                                                                                             | 50 |
| 3.6  | Twelve originally extracted features of different types, SBP, DBP, MBP and PP with the simultaneous reference respiration waveforms and their power spectrums. . . . .                                                                                              | 51 |
| 3.7  | An example of phase difference RM components between four features–PAT_ms, pulse amplitude, systolic time and IPA, and SBP, the time delay between features and SBP are calculated using cross-correlation based method. . . . .                                    | 52 |
| 3.8  | A demonstration of originally extracted PAT_ms with RM and decompose the original PAT_ms into 4 IMFs with different frequency components using VMD, and the RM-suppressed PAT_ms is obtained by aggregating IMFs with peak frequency $f_{pk} \leq 0.12$ Hz. . . . . | 54 |
| 3.9  | An example of a 400-seconds segment of originally extracted SBP, DBP, PAT_ms and PIR and their RM-suppressed versions, along with the scatter plots and Pearson correlation coefficients between PAT_ms, PIR and BPs. . . . .                                       | 55 |
| 3.10 | Examples of two subjects’ estimated and reference SBP using Model 1 with both respiration modulated features and RM-free features. . . . .                                                                                                                          | 58 |
| 3.11 | Examples of two subjects’ estimated and reference SBP using Model 2 with both respiration modulated features and RM-free features. . . . .                                                                                                                          | 59 |
| 4.1  | The framework of the proposed comprehensive performance evaluation method for cuffless BP estimation, including sequence pre-processing, segment similarity evaluation and sequence similarity evaluation. . . . .                                                  | 64 |
| 4.2  | A demonstration of the proposed segment distance similarity which is a customized logistic function according to IEEE Standard 1708a-2019 evaluation criteria [25]. . . . .                                                                                         | 67 |
| 4.3  | A demonstration of the proposed angular difference based trend similarity measures of approximated linear segments in the same direction and in opposite directions, with the corresponding calculations of $\theta$ and $\theta_{\max}$ . . . . .                  | 68 |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |    |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.4 | A demonstration of (a) segmenting the compared reference and estimated BPs using BEAST method, the reference BP is divided into 8 segments while the estimated BP is divided into 9 segments ( $t = 4$ is a change point for both reference and estimation); (b) the equal segment number (ESN) will re-process the change points and re-divide the compared sequences such that the compared sequences have the same number of segments (e.g., the reference and estimated BPs are re-processed to 10 segments), and the least squares are used to approximate the linear trend for each segment. . . . . | 71 |
| 4.5 | Waveforms of the reference SBP measurement ( $BP_{ref}$ ) and five simulated SBP estimations ( $BP_{est1}$ to $BP_{est5}$ ), and relations between the simulated SBPs and reference SBP. . . . .                                                                                                                                                                                                                                                                                                                                                                                                           | 72 |
| 4.6 | An approximate 30 minutes' example of reference SBPs and estimated SBPs using SOTA models (BP variations might be caused by daily activities, such as standing up and walking), the distance, trend and composite similarities between estimated SBPs using ApproximateNet and reference SBPs are 0.37, 0.23 and 0.32 respectively, and 0.81, 0.81 and 0.81 between estimated BPs using DeepRNN-4L and reference BPs respectively (the composite similarities with different combinations of $w_1$ and $w_2$ are demonstrated in Appendix A.4.7). . . . .                                                  | 74 |
| 5.1 | Demonstration of fiducial points of a typical PPG pulse wave, its first derivative PPG' and second derivative PPG'' which are also termed as VPG and APG respectively. Their fiducial points which are widely used to extract features. The waveform can be separated into anacrotic and catacrotic phases, which are dominated by systolic ejection and wave reflections from periphery respectively. . . . .                                                                                                                                                                                             | 79 |
| 5.2 | The framework of the proposed TCN-SENet-KAN model for BP estimation via time series regression method. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 84 |
| 5.3 | The structures of (a) TCN block, (b) SENet block employed for the proposed BP estimation model. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 86 |
| 5.4 | Architectures of the KAN and MLP. The diagram highlights the layer composition and activation function differences between the KAN and MLP models (figure adapted from [195]). . . . .                                                                                                                                                                                                                                                                                                                                                                                                                     | 88 |
| 5.5 | The structure of a parallel transformer and BiGRU with attention mechanisms for BP estimation. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 90 |
| 5.6 | Demonstrations of (a) dynamic time warping (DTW) and (b) soft dynamic time warping (sDTW) of $\gamma = 1.0$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 95 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |     |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 5.7  | Distribution of BP readings of selected 32 subjects in BCG BP dataset. . .                                                                                                                                                                                                                                                                                                                                                                                                                 | 97  |
| 5.8  | Demonstration of (a) a good quality PPG waveform segment, and (b) abnormal PPG waveform which has multiple ripples in diastolic phase is excluded from further analysis. . . . .                                                                                                                                                                                                                                                                                                           | 97  |
| 5.9  | A demonstration of applying both 4-th order Butterworth bandpass filter and Chebyshev Type II bandpass filter with low cut-off frequency of 0.5 Hz and high cut-off frequency of 12 Hz to a 5-second PPG segment. Chebyshev bandpass filter can make onset and dicrotic notch of PPG more prominent. . . . .                                                                                                                                                                               | 98  |
| 5.10 | Examples of ECG pulses, (a) ECG waveform with a distinct QRS complex, (b) ECG waveform with a minor R wave. . . . .                                                                                                                                                                                                                                                                                                                                                                        | 99  |
| 5.11 | The demonstration of the proposed ECG R peak detection method. (a) an original ECG waveform with minor QRS complex, (b) the high frequency components (10-25 Hz) of the ECG signal (blue curve) and its upper envelope (orange curve), (c) the upper envelope (orange curve) with its peaks (black dots) and the corrected ECG R peaks (red dots). . . . .                                                                                                                                 | 100 |
| 5.12 | Examples of detecting fiducial points from PPG waveform and the second derivative of PPG (APG) using methods proposed by <i>PulseAnalyse</i> , <i>PPGFeat</i> and <i>pyPPG</i> , resultinng in different detection results (data extracted from <b>BCG BP</b> dataset, sampling frequency is 1 KHz). . . . .                                                                                                                                                                               | 106 |
| 5.13 | Data preparation strategies for model training, (a) strategy for time series regression models, TCN-SENet-KAN and a parallel transformer-BiGRU with attention mechanisms which takes 4 seconds' of physiological features to predict BPs, (b) strategy for sig2bp model-CNN-LSTM which uses 2.048 seconds' waveform segments to predict BPs, (c) strategy for a hybrid model-HTCTNet which uses 4.096 seconds' waveform segments as well as physiological features to predict BPs. . . . . | 107 |
| 5.14 | Reference and estimated SBP of Subject 1 using (a) SOTA models and TCN-SENet-KAN with sDTW loss function, (b) UNet model, (c) CNN-LSTM model. . . . .                                                                                                                                                                                                                                                                                                                                      | 107 |
| 5.15 | Reference and estimated SBP of Subject 1 using (a) proposed Transformer-BiGRU attention, TCN-SENet-Linear model using MAE loss function, (b) proposed TCN-SENet-Linear and TCN-SENet-KAN model with MAE and sDTW loss functions. . . . .                                                                                                                                                                                                                                                   | 108 |



|      |                                                                                                                                                                                                                                                                                                                                                                                                                            |     |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 5.16 | Reference and estimated SBP of Subject 3 using (a) SOTA models and TCN-SENet-KAN with sDTW loss function, (b) UNet model, (c) CNN-LSTM model. . . . .                                                                                                                                                                                                                                                                      | 108 |
| 5.17 | Reference and estimated SBP of Subject 3 using (a) proposed Transformer-BiGRU attention, TCN-SENet-Linear model using MAE loss function, (b) proposed TCN-SENet-Linear and TCN-SENet-KAN model with MAE and sDTW loss functions. . . . .                                                                                                                                                                                   | 108 |
| 5.18 | Reference and estimated SBP of Subject 5 using (a) SOTA models and TCN-SENet-KAN with sDTW loss function, (b) UNet model, (c) CNN-LSTM model. . . . .                                                                                                                                                                                                                                                                      | 109 |
| 5.19 | Reference and estimated SBP of Subject 5 using (a) proposed Transformer-BiGRU attention, TCN-SENet-Linear model using MAE loss function, (b) proposed TCN-SENet-Linear and TCN-SENet-KAN model with MAE and sDTW loss functions. . . . .                                                                                                                                                                                   | 109 |
| 6.1  | (a) The Kruschke diagram of the proposed hierarchical model and the selection of the hyperpriors. (b) A demonstration of SBP estimation for one subject using the proposed hierarchical regression (red), OLS (green) with only 3 calibrations, while the true regression is the black line. (c) The estimated SBP trend (red) using the hierarchical regression and the reference SBP trend (adapted from [206]). . . . . | 112 |
| 6.2  | A demonstration of key process of domain adversarial neural network (DANN) (figure adapted and modified based on [212]). . . . .                                                                                                                                                                                                                                                                                           | 121 |
| 6.3  | Demonstration of adaptation window, evaluation set and calibration samples selection method applied to each test subject of LOSO cross validation for this study. . . . .                                                                                                                                                                                                                                                  | 125 |
| 6.4  | Box plot of MAE between reference and estimated BPs using different adaptation methods with 4 calibration measures. . . . .                                                                                                                                                                                                                                                                                                | 129 |
| 6.5  | Box plot of MAE between reference and estimated BPs using different adaptation methods with 2 calibration measures. . . . .                                                                                                                                                                                                                                                                                                | 130 |
| 6.6  | SBP estimation for three selected subjects using different adaptation methods with only 2 calibration measures. . . . .                                                                                                                                                                                                                                                                                                    | 130 |
| 6.7  | DBP estimation for three selected subjects using different adaptation methods with only 2 calibration measures. . . . .                                                                                                                                                                                                                                                                                                    | 131 |

|      |                                                                                                                                                                                                                                          |     |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| A.1  | (a) a demonstration of unknown time delay during data collection which causes unreliable PAT values (MIMIC database); (b) regular PAT extracted from ECG and PPG signals without extra time delay. . . . .                               | 163 |
| A.2  | The demonstration of PPG feature extraction . . . . .                                                                                                                                                                                    | 164 |
| A.3  | The demonstration of extracted features without respiratory effect removal (feature1-feature12, exemplary on subject1 from BCG dataset). . . . .                                                                                         | 166 |
| A.4  | The demonstration of extracted features without respiratory effect removal (feature13-feature24, exemplary on subject1 from BCG dataset). . . . .                                                                                        | 167 |
| A.5  | The demonstration of extracted features without respiratory effect removal (feature25-feature36, exemplary on subject1 from BCG dataset). . . . .                                                                                        | 168 |
| A.6  | The demonstration of extracted features without respiratory effect removal (feature37-feature48, exemplary on subject1 from BCG dataset). . . . .                                                                                        | 169 |
| A.7  | The demonstration of extracted features without respiratory effect removal (feature49-feature60, exemplary on subject1 from BCG dataset). . . . .                                                                                        | 170 |
| A.8  | The demonstration of extracted features, SBP and DBP without respiratory effect removal (feature61-feature69, SBP and DBP, exemplary on subject1 from BCG dataset). . . . .                                                              | 171 |
| A.9  | The demonstration of extracted features without respiratory effect removal (feature1-feature12, exemplary on subject8 from BCG dataset). . . . .                                                                                         | 172 |
| A.10 | The demonstration of extracted features without respiratory effect removal (feature13-feature24, exemplary on subject8 from BCG dataset). . . . .                                                                                        | 173 |
| A.11 | The demonstration of extracted features without respiratory effect removal (feature25-feature36, exemplary on subject8 from BCG dataset). . . . .                                                                                        | 174 |
| A.12 | The demonstration of extracted features without respiratory effect removal (feature37-feature48, exemplary on subject8 from BCG dataset). . . . .                                                                                        | 175 |
| A.13 | The demonstration of extracted features without respiratory effect removal (feature49-feature60, exemplary on subject8 from BCG dataset). . . . .                                                                                        | 176 |
| A.14 | The demonstration of extracted features, SBP and DBP without respiratory effect removal (feature61-feature69, SBP and DBP, exemplary on subject8 from BCG dataset). . . . .                                                              | 177 |
| A.15 | Demonstration of concordance correlation coefficient ( $\rho_c$ ) and Pearson correlation coefficient ( $\rho$ ) for different scenarios, (a) location shift, (b) scale shift, (c) location and scale shift, (d) no concordance. . . . . | 178 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| A.16 (a) Diagram of ApproximateNet [117], which is a deeply supervised one-dimensional U-Net to approximate ABP waveforms from preprocessed PPG waveforms; (b) Diagram of DeepRNN-4L [84], each rectangle box is an LSTM cell (unit = 128), and the sequence length of each training sample is set to 32. The bottom green dashed block is a bidirectional LSTM layer (unit = 128); there is a dropout layer (rate = 0.2) applied after each bidirectional and forward LSTM layer; residual connections are also applied between LSTM layers. The last layer (unit = 2, activation=linear) was added for SBP and DBP prediction. . . . . | 178 |
| A.17 Illustration of 7 ECG and PPG features used for SBP and DBP estimation proposed in [84]. $tf_i$ , $tp_i$ , $tn_i$ and $td_i$ indicate the time index of foot point, systolic peak, diastolic notch and diastolic peak of $i^{th}$ PPG pulse, respectively.                                                                                                                                                                                                                                                                                                                                                                          | 179 |
| A.18 Percentage of segments with different MAD levels for SBP and DBP estimation using ApproximateNet and DeepRNN-4L. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 179 |
| A.19 Percentage of segments with correct BP trend tracking for SBP and DBP using ApproximateNet and DeepRNN-4L. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 180 |
| A.20 Percentage of segments with $MAD \leq 5$ mmHg and correct BP trend tracking for SBP and DBP using ApproximateNet and DeepRNN-4L. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 180 |
| A.21 Reference and estimated BPs with different distance and trend similarities, and composite measure of both distance and trend similarities with different weights combination $w_1$ and $w_2$ (assume $t_{iR} - t_{iL} = 10$ , $\max(x_L, y'_L) - \min(x_L - y'_L) = 10$ mmHg for simulation1-10 s and $\max(x_L, y'_L) - \min(x_L - y'_L) = 20$ mmHg for simulation11 (refer to Table4 and Figure4). . . . .                                                                                                                                                                                                                        | 181 |
| A.22 the composite similarity metric (TN_CS) between estimations of DeepRNN-4L, ApproximateNet model and reference shown in Figure7 with different combinations of $w_1$ and $w_2$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                             | 182 |
| A.23 Impact of sampling frequency on PAT precision and its correlation with SBP for an example subject. (a) PAT extracted from signals sampled at 1 kHz (blue), downsampled to 500 Hz (green), and downsampled to 125 Hz (magenta), showing increased quantization at lower sampling rate. (b) Scatter plot of SBP versus PAT for the three sampling frequencies, illustrating a wider range of SBP values for a given PAT measurement at lower sampling rate and decreased correlation. . . . .                                                                                                                                         | 183 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| A.24 Comparison of PAT calculated using different filter configurations for Subject 1. Time series plot of PATs, with blue line representing PAT extracted from the thesis configuration and red line representing PAT extracted from the control configuration, showing high temporal agreement. The scatter plot (bottom left) confirms near-perfect linear correlation ( $r > 0.99$ ). The Bland-Altman plot (bottom right) illustrates the small magnitude and near-constant nature of the difference between the two filter configurations. . . . | 185 |
| A.25 Comparison of PAT calculated using different filter configurations for Subject 2. Time series plot of PATs, with blue line representing PAT extracted from the thesis configuration and red line representing PAT extracted from the control configuration, showing high temporal agreement. The scatter plot (bottom left) confirms near-perfect linear correlation ( $r > 0.99$ ). The Bland-Altman plot (bottom right) illustrates the small magnitude and near-constant nature of the difference between the two filter configurations. . . . | 186 |

# Abbreviations

**AAMI** Association for the Advancement of Medical Instrumentation [13](#), [26](#)

**ABP** Arterial Blood Pressure [21](#), [43](#), [48](#)

**AE** Absolute Error [104](#), [113](#)

**ANN** Artificial Neural Network [23](#)

**BCG** Ballistocardiogram [43](#)

**BHS** British Hypertension Society [26](#)

**BiGRU** Bidirectional Gated Recurrent Unit [89](#), [90](#)

**BP** Blood Pressure [v](#), [1](#)

**BW** Baseline Wandering [41](#)

**CNN** Convolutional Neural Network [26](#)

**DANN** Domain Adversarial Neural Network [v](#), [9](#), [118](#), [134](#)

**DBP** Diastolic Blood Pressure [1](#), [12](#)

**DTW** Dynamic Time Warping [93](#)

**ECG** Electrocardiogram [v](#), [29](#), [43](#)

**FDA** Food and Drug Administration [15](#), [29](#)

**GRU** Gated Recurrent Unit [26](#), [90](#)

**HRV** Heart Rate Variability [6](#), [43](#)

**KAN** Kolmogorov-Arnold Network [84](#)

**LOSO** Leave-One-Subject-Out [7](#), [114](#)

**LRCN** Long-term Recurrent Convolutional Network [26](#)

**LSTM** Long Short-Term Memory [26](#)

**MAE** Mean Absolute Error [v](#), [8](#), [24](#), [56](#), [92](#), [129](#)

**MAML** Model-Agnostic Meta-Learning [116](#)

**MBP** Mean Blood Pressure [1](#), [46](#)

**MCMC** Markov Chain Monte Carlo [113](#)

**ME** Mean Error [8](#)

**MFDS** Ministry of Food and Drug Safety [29](#)

**MSE** Mean Squared Error [92](#), [116](#)

**OLS** Ordinary Least Squares [111](#)

**PAT** Pulse Arrival Time [6](#), [21](#), [22](#)

**PP** Pulse Pressure [56](#)

**PPG** Photoplethysmogram [v](#), [13](#), [29](#), [43](#)

**PTT** Pulse Transit Time [19](#), [21](#), [29](#)

**PWA** Pulse Wave Analysis [23](#)

**PWV** Pulse Wave Velocity [18](#), [21](#)

**RM** Respiration Modulation [v](#), [8](#), [43](#), [47](#)

**RMSE** Root Mean Squared Error [25](#)

**RNN** Recurrent Neural Network [24](#), [26](#), [84](#), [90](#)

**RR** Respiration Rate [42](#), [46](#), [52](#)

**SBP** Systolic Blood Pressure [1](#), [12](#)

**sDTW** Soft Dynamic Time Warping [v](#), [92](#), [93](#), [134](#)

**SENet** Squeeze-and-Excitation Network [84](#)

**SOTA** state-of-the-art [4](#), [21](#)

**TCN** Temporal Convolutional Network [84](#)

**VMD** Variational Mode Decomposition [v](#), [43](#), [51](#)

# Chapter 1

## Introduction

The cardiovascular system, consisting of the heart, arteries, veins, and capillaries, circulates blood to supply oxygen and nutrients to every cell and to remove metabolic waste products [1]. Proper tissue perfusion requires adequate arterial pressure: blood must be driven through capillary beds, but pressures that are too high can damage delicate vessel walls [2]. BP is the mechanical force exerted by circulating blood on vessel walls, and it is one of the vital signs – together with respiratory rate, heart rate, oxygen saturation, and body temperature – that can evaluate an individual's health [3]. BP depends on cardiac output (heart rate  $\times$  stroke volume) and total peripheral resistance (vascular tone) [4]. A BP reading consists of two numbers (i.e., 120/80 mmHg) and it is usually measured in millimeters of mercury (mmHg) above the surrounding atmospheric pressure. The first (top) number is called the **Systolic Blood Pressure (SBP)**, which refers to the amount of pressure in the arteries during the heart muscle contraction. The second (bottom) number is called the **Diastolic Blood Pressure (DBP)**, which refers to the minimum pressure between heart beats. **Mean Blood Pressure (MBP)** therefore can be approximated as  $MBP \approx DBP + (SBP - DBP)/3$  [5]. An example of BP waveform of one cardiac cycle is shown in Figure 1.1. In an normal adult, normal resting BP is approximately 120 mmHg systolic over 80 mmHg diastolic, denoted as "120/80 mmHg". Globally, the average BP, age standardized, has remained about the same since 1975, at approximately 127/79 mmHg for males and 122/77 mmHg for females, although these average data mask quite large divergent regional trends [6].

As a vital sign, BP is a key indicator of cardiovascular health and abnormal BP has significant consequences. The most common BP disorders are high blood pressure (hypertension) and low blood pressure (hypotension). Hypertension is generally considered as a BP reading higher than 140 mmHg for SBP and 90 mmHg for DBP. Over time, even mild



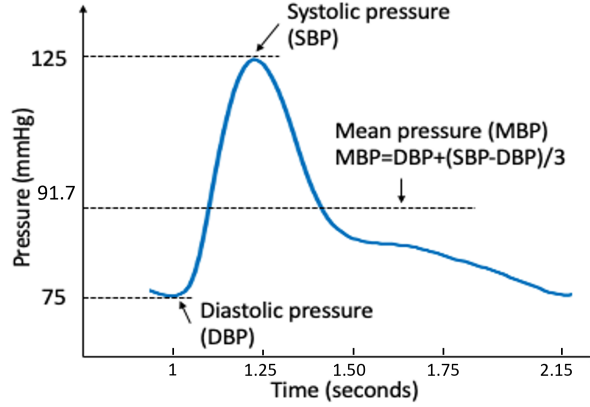


Figure 1.1: An example of peripheral BP waveform with indication of SBP, DBP and MBP.

hypertension can increase the risk of atherosclerosis, stroke, aneurysm, heart failure, heart attack and kidney damage [7–10]. Left untreated, chronic high BP can cause progressive damage to one or more vital organs [11]. Hypotension is generally considered as a BP reading lower than 90 mmHg for the SBP and 60 mmHg for the DBP. The symptoms of hypotension may include blurred or fading vision, dizziness or lightheadedness.

BP has been mostly measured by cuff-based devices. However, these conventional methods are often uncomfortable, cumbersome, and most critically, unsuitable for continuous and long-term monitoring. Cuffless BP estimation, which aims to address the limitations of conventional methods, has emerged as a promising and popular research topic recently. Moreover, the feasibility of integrating this technique into wearable devices, such as smartwatches, making cuffless BP estimation technique highly attractive for both clinical and everyday health monitoring applications.

## 1.1 Motivations

BP-related health issues have become a major global concern. A recent study [12] indicates that hypertension has been the world’s leading cause of death and killed more people, including young adults, than COVID-19 (see Figure 1.2(a)). According to World Health Organization (WHO), 1.28 billion adults aged 30-79 years worldwide are estimated to have hypertension [13]. However, an estimation of 54% adults with hypertension are aware that they have the condition and less than half of adults (42%) with hypertension are diagnosed and treated. Only approximately 1 in 7 people (~14%) with hypertension have

it under control as demonstrated in Figure 1.2(b). The statistics about the hypertension in Canada among different age groups from 2008 to 2019 are illustrated in Table 1.1. There are approximately 22.4% Canadian (age 20-79) were hypertensive from 2016 to 2019, but nearly 23% of people were unaware of this health problem.

There are two main reasons contributing to the unawareness or neglect to hypertension:

1. Hypertension is called a "silent killer". Most people with hypertension are unaware of the problem because it may have no warning signs or symptoms [13].
2. Currently, BP is still mostly measured by the traditional auscultatory and oscillometric method, which utilize an inflatable cuff which makes the measurement cumbersome and uncomfortable. Moreover, these methods cannot be used for continuous measurement.

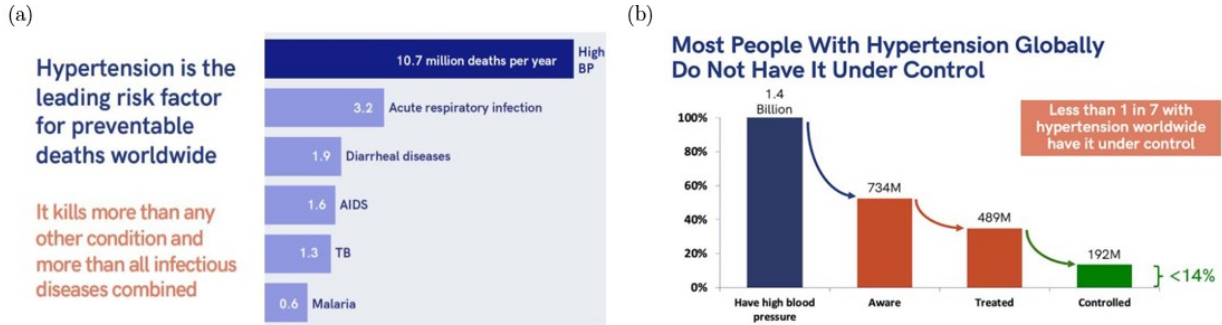


Figure 1.2: (a) Deaths caused by hypertension and other diseases per year globally [14] (data from "Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016" in The Lancet and WHO), (b) the amount of people with hypertension have it controlled globally [14] (data from "Global disparities of hypertension prevalence and control" in Circulation).

Due to the global burden of hypertension and the substantial proportion of individuals who remain undiagnosed or inadequately managed, there is an urgent demand for more accessible, convenient, and effective BP monitoring solutions. Cuffless BP estimation presents a promising alternative to traditional cuff-based methods by enabling continuous and long-term monitoring in a non-intrusive manner. This capability is particularly valuable for facilitating early diagnosis, optimizing antihypertensive therapy, and enhancing

Table 1.1: Statistics of hypertension in Canada (ages 20-79).

| Measures                         | Age group     | 2008 to 2011 |                   | 2012 to 2015 |                   | 2016 to 2019 |                   |
|----------------------------------|---------------|--------------|-------------------|--------------|-------------------|--------------|-------------------|
|                                  |               | Percent      | Number of persons | Percent      | Number of persons | Percent      | Number of persons |
| <b>Hypertension</b> <sup>1</sup> | Ages 20 to 79 | 22.5         | 5,378,000         | 23.2         | 5,865,800         | 22.4         | 5,902,200         |
|                                  | Ages 20 to 39 | 3.2*         | 285,500*          | 3.0*         | 280,600*          | 3.1*         | 293,900*          |
|                                  | Ages 40 to 59 | 21.4         | 2,108,900         | 23.2         | 2,303,800         | 21.0         | 2,056,000         |
|                                  | Ages 60 to 79 | 57.7         | 2,984,700         | 23.2         | 3,281,300         | 51.2         | 3,552,300         |
| <b>Awareness</b> <sup>2</sup>    | Ages 20 to 79 | 84.7         | 4,557,600         | 84.2         | 4,941,600         | 77.8         | 4,591,900         |
|                                  | Ages 20 to 39 | 57.7         | 164,700           | 64.6*        | 181,300*          | 49.6*        | 145,900*          |
|                                  | Ages 40 to 59 | 82.5         | 1,740,300         | 79.7         | 1,835,600         | 73.7         | 1,515,500         |
|                                  | Ages 60 to 79 | 88.9         | 2,652,500         | 89.1         | 2,924,600         | 82.5         | 2,930,500         |
| <b>Treatment</b> <sup>3</sup>    | Ages 20 to 79 | 80.7         | 4,340,300         | 79.5         | 4,665,600         | 74.2         | 4,378,100         |
|                                  | Ages 20 to 39 | 45.6*        | 130,100*          | 55.4*        | 155,600*          | 45.8*        | 134,600*          |
|                                  | Ages 40 to 59 | 75.3         | 1,588,000         | 72.5         | 1,670,400         | 66.4         | 1,365,100         |
|                                  | Ages 60 to 79 | 87.9         | 2,622,200         | 79.5         | 2,839,600         | 81.0         | 2,878,500         |
| <b>Control</b> <sup>4</sup>      | Ages 20 to 79 | 67.3         | 3,619,500         | 66.0         | 3,870,500         | 61.0         | 3,600,000         |
|                                  | Ages 20 to 39 | 44.3*        | 126,500*          | 51.2*        | 143,500*          | 42.4*        | 124,500*          |
|                                  | Ages 40 to 59 | 65.5         | 1,382,100         | 59.5         | 1,370,000         | 54.0         | 1,110,500         |
|                                  | Ages 60 to 79 | 70.7         | 2,110,800         | 71.8         | 2,356,900         | 66.6         | 2,365,100         |

Data source: Statistics Canada, released on 28/06/2021.

<sup>1</sup> Hypertension: systolic blood pressure  $\geq 140$  mmHg or diastolic blood pressure  $\geq 90$  mmHg or antihypertensive medication used.

<sup>2</sup> Awareness: diagnosed hypertension or antihypertensive medication used.

<sup>3</sup> Treatment: a hypertensive respondent's report of antihypertensive medication use in the past month.

<sup>4</sup> Controlled: a hypertensive respondent used antihypertensive medication and blood pressure is normal.

\* Use with caution.

patient adherence to treatment. Consequently, the widespread adoption of cuffless BP monitoring is expected to play a significant role in mitigating the prevalence and severity of hypertension and improving overall cardiovascular health outcomes.

After a comprehensive review of [state-of-the-art \(SOTA\)](#) studies and the current progress in cuffless BP estimation, this thesis identifies and addresses four critical research limitations that have hindered the practical implementation of this technique. These limitations, observed across numerous prior works, are considered fundamental barriers to achieving accurate BP estimation using pre-defined morphological features extracted from ECG and/or PPG signals with deep learning models in real-world scenarios. By systematically tackling these limitations, this thesis aims to enhance estimation performance and significantly improve the feasibility of deploying cuffless BP monitoring in everyday and clinical applications, thereby bring this technique closer to widespread real-world adoption. The overview of the research problems identified and solutions proposed by this thesis, as well as the contributions are demonstrated in Figure 1.3.

## Critical Limitations in Cuffless BP Estimation Studies

Among various different cuffless BP estimation methods, such as ultrasound-based methods, bioimpedance-based methods and physiological signal-based methods (see Section 2.2

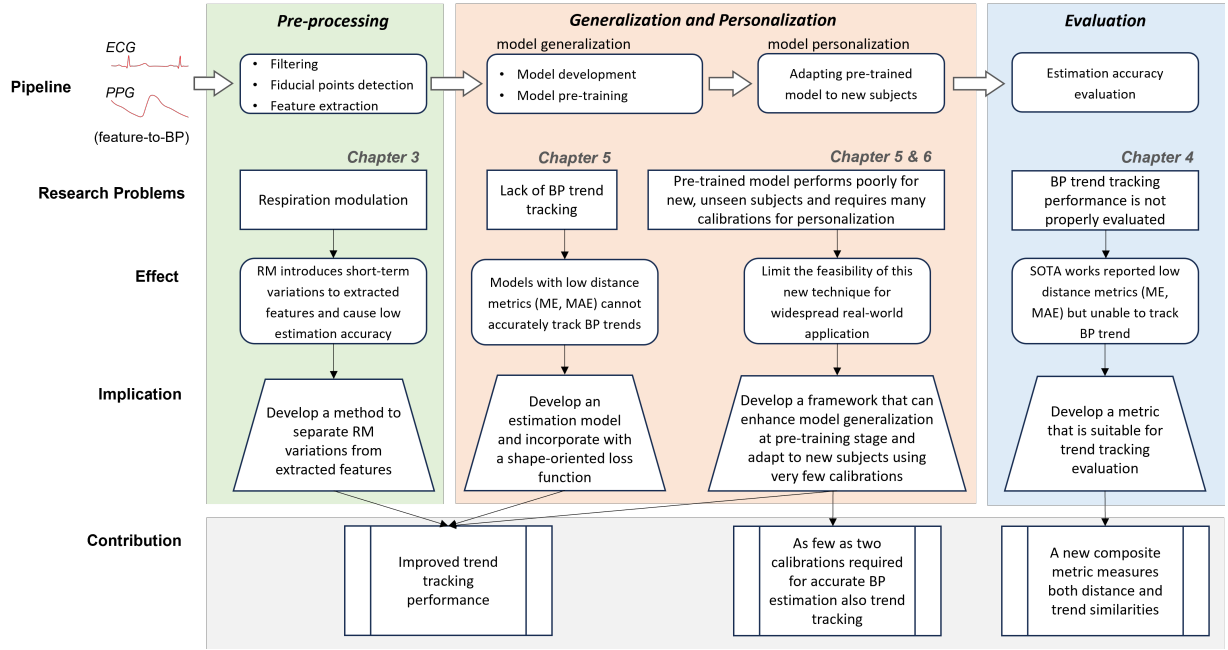


Figure 1.3: The overview of this thesis: the pipeline of the ECG and PPG feature based BP estimation model, four critical research problems identified by this thesis with their effects on practical application, along with the implications and contributions.

for details), this thesis focuses on the physiological signal-based method, specifically estimating BPs using pre-defined morphological features extracted from ECG and PPG waveforms and deep learning models. A typical pipeline of this method comprises several key stages, as illustrated in Figure 1.3. The pre-processing stage involves denoising the raw ECG and PPG signals, detecting fiducial points and extracting relevant morphological features. The model generalization stage focuses on developing a robust deep learning based estimation model and pre-train it using data from a large and diverse subject group to enhance its generalization capability. The model personalization stage aims to adapt the pre-trained model to completely new and unseen subject using a minimal number of calibration measures. Finally, the model evaluation stage is designed to assess the model’s performance using appropriate metrics, evaluating both the accuracy of BP estimation and model’s ability to track BP trends effectively.

After thoroughly reviewing existing SOTA studies and analyzing the findings obtained from preliminary experiments conducted in this thesis, four critical research problems which are situated at different stages of the cuffless BP estimation pipeline (see Figure 1.3) have been identified. These challenges that limit the practical deployment of this new technique are: **respiration-induced variations in morphological features and BPs, unsatisfactory trend tracking performance, poor generalization and inefficient personalization and inadequate evaluation metrics.**

### 1.1.1 The Overlooked Role of Respiration in Feature-Based Cuffless BP Estimation

Respiration has been shown to introduce variations in several physiological signals and measures. For instance, it can modulate the R–R intervals between ECG R-peaks, thereby affecting [Heart Rate Variability \(HRV\)](#) analysis [15]. Previous studies have also observed that respiration induces fluctuations in BP [16, 17], as well as in morphological features extracted from PPG and ECG waveforms, such as [Pulse Arrival Time \(PAT\)](#) [18], systolic time interval (STI), and diastolic time interval (DTI) [19]. However, the influence of respiration on a wide range of morphological features used for cuffless BP estimation, and its overall impact on estimation accuracy, has never been systematically examined or addressed.

Throughout the experiments conducted in this thesis, respiration has been found to negatively impact cuffless BP estimation performance. Specifically, respiration can introduce rapid variations across various morphological features, each with differing magnitudes and phase shifts. This adds considerable difficulty for deep learning–based models to cap-

ture the underlying long-term BP trend when these respiration-induced fluctuations are not effectively separated from the input features and target BP values during training. As a result, the estimation outputs may become noisy and inconsistent.

### 1.1.2 Unsatisfactory Trend Tracking Performance

Unlike conventional cuff-based methods, cuffless BP technique is intended for continuous, long-term monitoring. Capturing the temporal dynamics and trends of BP is also essential for real-world application. For example, the "non-dipping" or "reverse dipping" of nocturnal BP patterns are strong predictors of cardiovascular risk [20–22]. Moreover, monitoring long-term BP trends enable more precise hypertension management, supports medication adjustment and helps clinicians assess treatment effectiveness over time. However, current studies commonly overlook BP trend tracking and mainly focus on pointwise accuracy. This neglect results in unsatisfactory BP trend tracking performance which makes this technique less practical for real-world long-term monitoring.

### 1.1.3 Poor Model Generalization and Inefficient Personalization

Although recent studies in cuffless BP estimation have shown promising results, many rely on non-rigorous training and evaluation strategies that can lead to overly optimistic performance assessments. These approaches often fail to reveal fundamental issues related to poor model generalization, particularly under more stringent evaluation protocols such as [Leave-One-Subject-Out \(LOSO\)](#) cross-validation (see Section 2.4.1 for details). In real-world scenarios where models are applied to entirely unseen subjects, this lack of generalization becomes a critical barrier, as the estimation performance tends to degrade significantly due to inter-subject variability.

Moreover, although personalization aims to improve model performance for new subjects, current strategies remain limited and often impractical. Existing methods typically require a large amount of subject-specific calibration data to adapt the pre-trained models, which hinders their feasibility in real-world applications. Even with extensive calibration efforts, some models still fail to meet clinically acceptable accuracy standards, underscoring the inefficiency and ineffectiveness of current approaches. These limitations point to an urgent need for more robust generalization techniques and data-efficient personalization strategies capable of maintaining high estimation accuracy with minimal subject-specific data.

### 1.1.4 Inadequate Evaluation Metrics

Most existing cuffless BP estimation studies continue to rely on evaluation standards originally developed for traditional cuff-based devices, which are primarily intended for intermittent and discrete measurements [23–26]. These standards typically focus on pointwise error metrics, such as [Mean Error \(ME\)](#) and [MAE](#), which assess accuracy at individual time points. However, such metrics are inadequate for evaluating the performance of cuffless BP systems designed for continuous, long-term monitoring. In particular, pointwise distance-based metrics fail to capture a model’s ability to accurately track the trends and dynamic fluctuations of BP over time which is an essential requirement for clinical relevance in continuous monitoring applications. Therefore, there is a pressing need to develop and adopt new evaluation frameworks that better reflect both BP value estimation accuracy and trend tracking performance of cuffless BP estimation models.

## 1.2 Objectives

The objectives of this thesis are defined and proposed to address the aforementioned four critical research problems identified from the existing SOTA works.

- Develop a preprocessing method to separate [RM](#) from extracted morphological features and BPs in order to mitigate short-term variations caused by respiration and improve BP estimation accuracy.
- Improve BP trend tracking performance and model generalization by developing an estimation model that effectively captures informative patterns from morphological features and incorporates a shape-oriented loss function to emphasize temporal dynamics.
- Establish an efficient personalization strategy that adapts the pre-trained model to new subjects using very few calibration samples, thereby enhancing subject-specific estimation performance while minimizing calibration requirements.
- Propose a comprehensive evaluation metric that simultaneously captures both distance-based accuracy and temporal trend similarity, addressing the limitations of existing metrics and enabling more meaningful assessment of cuffless BP models.

## 1.3 Contributions

This study presents several novel contributions to the field of cuffless BP estimation, including:

- This thesis presents a comprehensive approach to enhance BP trend tracking performance which is commonly overlooked in prior works, by jointly leveraging respiration-resilient preprocessing, a temporal-aware model architecture, and a shape-oriented loss function. This integrated framework enables more accurate BP trend tracking performance and improves the feasibility for continuous long-term BP monitoring.
- This thesis proposes a novel framework that enhances both model generalization and personalization for cuffless BP estimation. **DANN** is employed during the pre-training stage to encourage the learning of subject-invariant features, while subject-specific adaptation is achieved through lightweight model fine-tuning and Procrustes analysis. The proposed method achieves IEEE evaluation standard "Grade A" performance for both SBP and DBP using only two calibration samples, substantially reducing the calibration burden compared to prior SOTA approaches while simultaneously improving estimation accuracy in terms of both absolute BP values and trend tracking.
- This thesis is the first to reveal the limitations of current evaluation protocols for BP trend tracking and systematically assess widely used metrics in this context. A composite evaluation metric is proposed for cuffless BP estimation, combining both distance and trend similarities measures. These normalized metrics (ranging from -1 to 1) are interpretable, generalizable, and applicable not only to BP monitoring but also to broader healthcare time-series evaluation tasks.

The following papers were submitted and published based on the research leading to this thesis:

- Shan He, Hilmi R. Dajani, and Miodrag Bolic, "Novel Cuffless Blood Pressure Estimation Method using a Bayesian Hierarchical Model," 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pages 898–901. IEEE, 2021.
- Shan He and Miodrag Bolic, "Novel Metrics for Tracking Blood Pressure Changes in Continuous Cuffless Blood Pressure Estimations," Scientific Reports, 14(1):27478, 2024.



- Shan He and Miodrag Bolic, "Analyzing and Processing the Underrated Effect of Respiration for Cuffless Blood Pressure Estimation," submitted to IEEE Transaction in Biomedical Engineering.

## 1.4 Outline of the Thesis

The thesis is organized as follows:

1. Chapter 2 reviews conventional BP measurement techniques and recent developments in cuffless BP estimation. It also highlights critical limitations in existing studies.
2. Chapter 3 investigates the respiration variations observed in physiological features and BPs, which can degrade estimation both estimation accuracy and trend tracking. A VMD-based method is proposed to separate respiration modulation before model training.
3. Chapter 4 proposes a novel evaluation metric that combines distance accuracy and trend similarity, addressing a key gap in existing evaluation standards that primarily focus on ME and MAE without assessing trend tracking which is crucial for long-term monitoring.
4. Chapter 5 systematically analyzes the relationships between a wide range of morphological features to BPs, proposes two BP estimation models which predict BPs in time-series regression and aim to improve BP trend tracking performance and model generalization with a shape-oriented loss function, soft-DTW.
5. Chapter 6 evaluates several adaptation strategies including fine-tuning, MAML, DANN, and Procrustes transformation. The combination of DANN-based pre-training, fine-tuning with sDTW loss, and Procrustes alignment enables effective subject-specific BP estimation using only 2–4 calibration measures.
6. Chapter 7 summarizes the key contributions of this thesis in improving model generalization, personalization, and trend tracking in cuffless BP estimation. It also discusses future directions, including model robustness, further reduction of calibration needs, and deployment in real-world scenarios.

# Chapter 2

## Literature Review

This chapter first introduces the conventional techniques for BP measurement. Then, the recent advances in cuffless BP estimation using physiological signals such as ECG and PPG are reviewed. Finally, the major limitations of the current studies are critically examined.

### 2.1 Review of Current BP Measurement Techniques and Devices

#### 2.1.1 Invasive BP Measurement Techniques

**Catheterization** is the gold standard method of blood pressure measurements [27]. This method measures instantaneous BP by placing a strain gauge in fluid contact with blood at any arterial site (e.g., the radial artery or the aorta). This method is typically used for perioperative monitoring during major surgical procedures and in critically ill patients. A big advantage of the arterial catheterization is that this approach is the most accurate BP measurement method, and it is also able to measure blood pressure as well as many other hemodynamic parameters to evaluate the health of the cardiovascular system at different measurement sites.

#### 2.1.2 Noninvasive BP Measurement Techniques

**Auscultation**, **oscillometry**, and **volume clamping** are noninvasive methods and these methods employ an inflatable cuff. The auscultation and oscillometry methods are widely

used in hospital and home healthcare.

### 2.1.2.1 Auscultation

The auscultatory method is the standard clinical method [28] which uses a stethoscope and a sphygmomanometer. This comprises an inflatable cuff placed around the upper arm at approximately the same vertical height as the heart, attached to a mercury or aneroid manometer. The mercury manometer measures the height of a column of mercury, giving an absolute result without need for subject-specific calibration and, consequently, not subject to the errors and drift of calibration which affect other methods. The use of mercury manometers is often required in clinical trials and for the clinical measurement of hypertension in high-risk patients. To conduct the measurement, a cuff of the appropriate size is fitted smoothly and snugly around the upper arm [29], and then it is inflated manually by repeatedly squeezing a rubber bulb until the artery is completely occluded. By listening with the stethoscope to the brachial artery at the antecubital area of the elbow, the examiner slowly releases the pressure in the cuff. When blood just starts to flow in the artery, the turbulent flow creates a “whooshing” or pounding (first Korotkoff sound). The pressure at which this sound is first heard is the SBP. The cuff pressure is further released until no sound can be heard (fifth Korotkoff sound), at the DBP. The auscultatory BP device and the procedures of auscultatory BP measurement are shown in Figure 2.1.

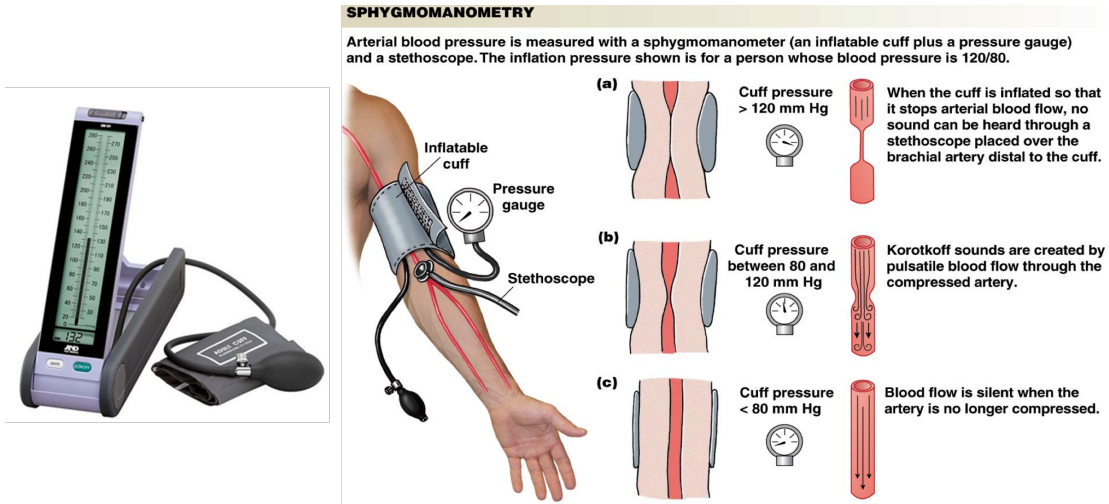


Figure 2.1: An example of the auscultatory BP device (left, image adapted from [30]) and the methodology of auscultatory BP measurement (right, image adapted from [31]).

### 2.1.2.2 Oscillometry

The oscillometric method is the most popular noninvasive, automatic method [32, 33]. It was first demonstrated in 1876 [34] and involves the observations in the sphygmomanometer cuff pressure which are caused by the oscillation of blood flow, i.e., the pulse. The electronic version of this method is widely used for home BP measurements nowadays. It uses a sphygmomanometer cuff, an electronic pressure sensor to observe cuff pressure oscillations, electronics to automatically interpret them and automatically inflate and deflate the cuff. The device should be calibrated periodically to maintain accuracy [35]. For a regular oscillometric BP measurement, the cuff is inflated to a pressure initially in excess of the systolic arterial pressure and then reduced to below diastolic pressure over a period about 30 seconds. When blood flow is nil (cuff pressure exceeding systolic pressure) or unimpeded (cuff pressure below diastolic pressure), cuff pressure will be essentially constant. When blood flow is present, but restricted, the cuff pressure, which is monitored by the pressure sensor, will vary periodically in synchrony with the cyclic expansion and contraction of the brachial artery. Over the deflation period, the recorded pressure waveform forms a signal known as the cuff deflation curve. A bandpass filter is utilized to extract the oscillometric pulses from the cuff deflation curve. Over the deflation period, the extracted oscillometric pulses form a signal known as the oscillometric waveform. The amplitude of the oscillometric pulses increases to a maximum and then decreases with further deflation. A variety of analysis algorithms can be employed in order to estimate the systolic, diastolic, and mean arterial pressure [36, 37]. The calculation of systolic, diastolic and mean BP using oscillometric method is shown in Figure 2.2. However, oscillometry is unreliable in subjects with certain conditions such as arterial fibrillation, stiff arteries, and pre-eclampsia [38].

### 2.1.2.3 Volume Clamping

Volume clamping is a noninvasive, automatic method used in research [40, 41]. This method measures instantaneous (finger) BP by using a cuff and a PPG sensor to measure the blood volume. The blood volume at zero transmural pressure is estimated via oscillometry. The cuff pressure is then continually varied to maintain this blood volume throughout the cardiac cycle via a fast servo-control system. The applied cuff pressure may thus equal BP. Volume clamping devices also achieve BP errors within Association for the Advancement of Medical Instrumentation (AAMI) limits when evaluated against auscultation and near AAMI limits when evaluated against radial artery catheterization [42]. A demonstration of the volume-clamp method can be found in Figure 2.2. Monitoring systems, such as Finapres NOVA [43] (Finapres Medical System), CNAP Monitor 500 [44] (BIOPAC Systems.

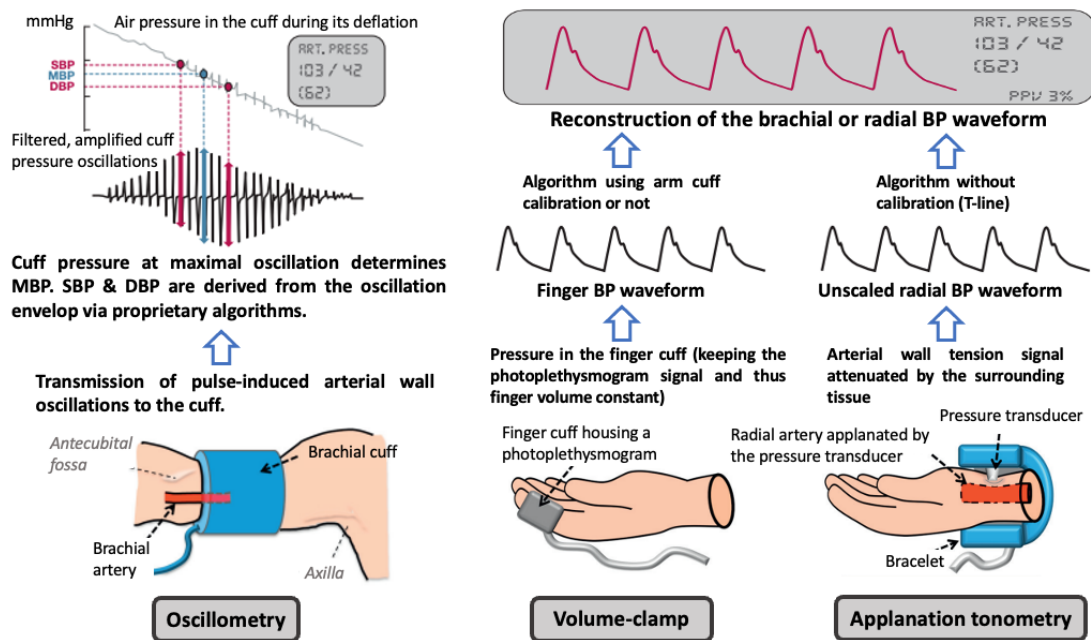


Figure 2.2: Demonstrations of oscillometry, volume-clamp and tonometry BP measurement methods (figure adapted from [39]).

INC), ClearSight System [45] (Edwards Lifesciences), utilize the volume clamp method to provide continuous, non-invasive, beat-to-beat BP monitoring from a finger cuff. These devices have been widely used by cuffless BP estimation studies to record reference brachial ABP waveforms.

#### 2.1.2.4 Tonometry

Tonometry is another non-invasive method used in research that, in theory, does not require an inflatable cuff [46, 47]. Arterial applanation tonometry consists of placing, over the skin, a pressure transducer that gently compresses the underlying artery. This action allows the reconstruction of the BP waveform, using a proprietary algorithm taking into account the soft tissue-related signal loss. Hence, a tonometer quantifies the arterial pulse that physicians otherwise subjectively assess through radial palpation by estimating the arterial wall tension. Among several different devices of this kind, the T-Line device is user-friendly and free of user bias because the sensor is housed by a wrist bracelet rather than handheld by a health-care provider. The sensor automatically moves over the radial artery until maximal pulse pressure is recorded within the bracelet. And it is reported that this device does not require calibrations [39, 48, 49]. However, it is still recommended that the devices of this kind should be routinely calibrated with cuff BP whenever a BP change is anticipated [50]. A brief demonstration of tonometry method is shown in Figure 2.2.

### 2.1.3 New Wearable Cuff-based BP Devices

Omron HeartGuide (see Figure 2.3 (a)) is the first clinically-accurate, wearable BP monitor designed in the innovative form of a wristwatch, and it was approved as a medical device by the U.S. Food and Drug Administration (FDA) in November 2018 [51]. The device's BP reading accuracy is  $\pm 3$  mmHg [52]. Huawei Watch D (see Figure 2.3 (b)), which integrates an inflatable cuff into the watch band, is another commercialized cuff-based wearable BP monitor. The Huawei Watch D has been cleared as a Class II medical device by China's National Medical Products Administration and the BP measurement accuracy has been reported as  $\pm 3$  mmHg [53]. The oscillometry method is used by both the Omron HeartGuide and the Huawei Watch D.



Figure 2.3: Commercialized wearable continuous BP monitors, (a) Omron HeartGuide; (b) Huawei Watch D; (c) Biobeat Tech BB-613WP; (d) Samsung Galaxy Active2.

#### 2.1.4 Calibration for Conventional BP Devices

The aforementioned conventional BP devices, such as sphygmomanometers and volume clamp based devices (e.g., Finapres) require **calibration** to maintain measurement accuracy.

For manual auscultatory or automated oscillometric devices, they are manufactured to meet specific international standards for accuracy (e.g., ISO 81060 series). "**Calibration**" in this sense refers to ensuring the pressure transducer within the device accurately measures and displays pressure (in mmHg). This is typically verified against a reference manometer during manufacturing and quality control. Also, for clinical and reliable personal use, these devices require periodic checks to maintain the measurement accuracy. This involves comparing the device's readings against a recently calibrated reference device, such as a mercury sphygmomanometer or a validated digital reference.

For Finapres and similar devices which utilize the volume clamp method to provide continuous, non-invasive, beat-to-beat BP monitoring from a finger cuff. Their calibration involves several aspects. When the device is first applied to a new user, an initial "physiological calibration" or setup procedure is performed. This involves the device automatically determining the correct set-point for the PPG and this set-point corresponds to the finger artery's unloaded diameter or a target transmural pressure where the PPG signal is most sensitive to arterial pressure changes. Once the set-point is established, the device continuously adjusts the finger cuff pressure on a beat-to-beat basis to maintain the constant arterial volume detected by the PPG during measurement.

### 2.1.5 Limitations of Conventional Methods

The advantages and disadvantages of the aforementioned techniques are summarized in Table 2.1. The major drawbacks of the aforementioned BP measurement methods are that they are cumbersome and uncomfortable to use, calibrations required, and most importantly, they are incapable of continuous monitoring. Due to the limitations of conventional methods and the significant advancements in computer vision and sensing technologies, new approaches have been developed to monitor BP without the use of a cuff. These cuffless BP monitoring techniques, which are well-suited for continuous and long-term tracking, have become a prominent research focus and have demonstrated promising results.

Table 2.1: Summary and comparison of the conventional BP measurement methods.

| Methods                              | Advantages                                                                                                                                                              | Disadvantages                                                                                                                                                                                           |
|--------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Auscultatory</b><br>(noninvasive) | <ul style="list-style-type: none"> <li>• Accurate</li> <li>• Low cost</li> <li>• No calibration required</li> </ul>                                                     | <ul style="list-style-type: none"> <li>• Professional experience required</li> <li>• Not for continuous measurements</li> <li>• Cuff-based, cumbersome and uncomfortable</li> </ul>                     |
| <b>Oscillometry</b><br>(noninvasive) | <ul style="list-style-type: none"> <li>• Automatic</li> <li>• Easy to use</li> <li>• Low cost</li> <li>• For Home healthcare</li> </ul>                                 | <ul style="list-style-type: none"> <li>• Not accurate enough</li> <li>• Calibrations required</li> <li>• Not for continuous measurements</li> <li>• Cuff-based, cumbersome and uncomfortable</li> </ul> |
| <b>Tonometry</b><br>(noninvasive)    | <ul style="list-style-type: none"> <li>• No cuff required</li> </ul>                                                                                                    | <ul style="list-style-type: none"> <li>• Expensive</li> <li>• Professional experience required</li> <li>• Not for continuous measurements</li> <li>• Not accurate enough</li> </ul>                     |
| <b>Volume-clamp</b><br>(noninvasive) | <ul style="list-style-type: none"> <li>• Continuous measurements</li> <li>• Beat-to-beat measurements</li> </ul>                                                        | <ul style="list-style-type: none"> <li>• Calibrations required</li> <li>• Expensive</li> <li>• Cuff-based, cumbersome and uncomfortable</li> </ul>                                                      |
| <b>Catheterization</b><br>(invasive) | <ul style="list-style-type: none"> <li>• The most accurate</li> <li>• Measure at different measurement sites</li> <li>• Measures cardiovascular hemodynamics</li> </ul> | <ul style="list-style-type: none"> <li>• Professional experience required</li> <li>• Only conducted in hospitals</li> <li>• Procedures are risky</li> </ul>                                             |

## 2.2 Cuffless BP Estimation Techniques

This section will discuss several key methodologies: ultrasound-based, bioimpedance-based, ballistocardiogram (BCG)-based, computer vision-based (e.g., remote photoplethysmography - rPPG), and physiological signal-based (e.g., ECG and PPG), detailing their prin-



ciples, SOTA works, performance, and a comparative summary of their strengths and limitations.

### 2.2.1 Ultrasound-Based BP Estimation

Ultrasound-based BP estimation methods typically determine BP by measuring changes in arterial dimensions, such as diameter, using ultrasonic waves that reflect off vessel walls [54]. The arterial diameter fluctuates with pressure according to the vessel’s mechanical properties, and this relationship can be used to calculate [Pulse Wave Velocity \(PWV\)](#) or directly correlate diameter changes to BP [55]. Recent advancements have seen the development of wearable ultrasound patches which can track blood vessel diameter for continuous BP monitoring and have been validated on 117 subjects, showing results comparable to clinical standards [56]. Another recent study [55] focused on lead-free KNN-based wearable ultrasonic patches, demonstrating effective signal transmission and vascular diameter detection on artificial vessels. The progression from lab concepts to clinically validated wearable patches indicates a strong drive to translate this high-fidelity sensing into practical continuous monitoring.

### 2.2.2 Bioimpedance-Based BP Estimation

Bioimpedance techniques measure the electrical impedance of a body segment, which changes as blood volume within arteries alters due to the pulse wave. BP is commonly estimated by deriving PWV from the transit time of the bioimpedance pulse waveform between two arterial sites, relating the change in arterial cross-sectional area to the change in measured impedance [57]. Sel *et al.* developed wearable ring sensors and achieved correlations of 0.81 between estimation and reference and estimation errors around  $0.11 \pm 5$  mmHg [58]. Huynh *et al.* proposed a bioimpedance sensor for smartwatches and reported PWV-BP correlations around 0.81–0.84 and RMSEs of approximately 5-7.5 mmHg [57]. More recently, Hsiao *et al.* developed a multi-modal wearable with two PPG and two bioimpedance sensors and achieved MAEs < 3.3 mmHg between estimated and reference BPs using random forest regression [59].

### 2.2.3 Computer Vision-Based BP Estimation (rPPG)

Remote photoplethysmography (rPPG) can approximate the PPG signal by analyzing subtle skin color changes from video recordings of a subject, typically the face, using standard

cameras [60]. BP is then estimated from the rPPG signal, often using **Pulse Transit Time (PTT)** principles or morphological analysis with machine learning. Dasari *et al.* extracted rPPG from face and hand ROIs (regions of interest), incorporating demographic data, achieved SBP MAE of 17.1 mmHg and DBP MAE of 13.2 mmHg [61]. Hwang & Lee proposed a dual-network approach (DRP-Net, BBP-Net) with phase-shifted rPPG to estimate BPs and the MAE between estimation and reference is 10.19 mmHg and 7.09 mmHg for SBP and DBP respectively [62].

#### 2.2.4 Ballistocardiogram (BCG)-Based BP Estimation

BCG is the measure of the mechanical recoil of the body caused by cardiac ejection of blood. Features from BCG waveform, such as the time interval and amplitude of I, J and K waves can be correlated with BP, or PTT can be derived using BCG as a proximal reference with a distal sensor like impedance plethysmography (IPG) or PPG [63]. Liu *et al.* measured BCG and IPG from a modified weight scale and estimated BPs using an XGBoost model, the correlation coefficients between the estimations and references are  $0.953 \pm 0.007$  for SBP and  $0.935 \pm 0.007$  for DBP, the reported MAE is  $3.35 \pm 0.28$  mmHg for SBP and  $2.67 \pm 0.15$  mmHg for DBP estimation respectively [63]. In 2018, Kim *et al.* demonstrated the BCG I-J interval was comparable to PTT for DBP monitoring [64].

#### 2.2.5 Physiological Signal (ECG and/or PPG)-Based BP Estimation

Methods based on ECG and/or PPG signals are central to cuffless BP research, utilizing features like PAT, or pulse wave analysis (PWA) from waveform morphology [65]. The PPG waveform’s morphology such as fiducial points of systolic and diastolic peaks, dicrotic notch, and its derivatives (VPG and APG) provide rich cardiovascular information and are widely employed for cuffless BP estimation [66]. In 2023, a deep RNN model which consists of 2 bi-directional Long Short-term Memory (Bi-LSTM) layers and 6 LSTM layers was developed to estimate BPs from 14 features extracted from ECG and PPG waveforms, the reported RMSEs are 3.223 mmHg and 1.781 mmHg for SBP and DBP estimation respectively [67]. In 2025, Zheng *et al.* included transformer mechanisms to UNet model (UTransBPNet) to estimate ABP waveforms from ECG, PPG and PPG derivatives, and achieved MAE of 4.38 mmHg and 2.25 mmHg for SBP and DBP estimations on MIMIC dataset [68].

Table 2.2: Summary and comparison of the cuffless BP estimation methods.

| Methods                         | Advantages                                                                                                                                                                                     | Disadvantages                                                                                                                                                                                                                                                  |
|---------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Ultrasound</b>               | <ul style="list-style-type: none"> <li>• Direct vascular measurement</li> <li>• High potential precision; SOTA clinically validated</li> <li>• Continuous measure by wearable patch</li> </ul> | <ul style="list-style-type: none"> <li>• Highly motion artifact sensitive</li> <li>• Good acoustic alignment/coupling required</li> <li>• Higher system complexity and cost</li> </ul>                                                                         |
| <b>Bioimpedance</b>             | <ul style="list-style-type: none"> <li>• Compact wearable integration</li> <li>• Generally skin tone insensitive</li> </ul>                                                                    | <ul style="list-style-type: none"> <li>• Electrode-skin contact required</li> <li>• Susceptible to motion/electrodermal interference</li> <li>• Calibration required</li> </ul>                                                                                |
| <b>Ballistocardiogram (BCG)</b> | <ul style="list-style-type: none"> <li>• Unobtrusive; integration into everyday objects (e.g., weight scale)</li> <li>• Potential for unconstrained monitoring</li> </ul>                      | <ul style="list-style-type: none"> <li>• Highly motion/vibration sensitive</li> <li>• Generally low signal-to-noise ratio of BCG signal</li> <li>• Often requires secondary distal sensor (e.g., IPG or PPG)</li> <li>• Calibration required</li> </ul>        |
| <b>Computer Vision (rPPG)</b>   | <ul style="list-style-type: none"> <li>• Completely non-contact and remote</li> <li>• Potential for opportunistic screening</li> <li>• Scalable for population monitoring</li> </ul>           | <ul style="list-style-type: none"> <li>• Highly sensitive to light and motion</li> <li>• Significant privacy and data security concerns</li> <li>• Accuracy affected by distance/angle/skin tone</li> <li>• Calibration required</li> </ul>                    |
| <b>ECG and/or PPG</b>           | <ul style="list-style-type: none"> <li>• Accessible sensors in consumer wearables</li> <li>• Continuous, long-term monitoring</li> </ul>                                                       | <ul style="list-style-type: none"> <li>• Calibration often required (periodic/individual)</li> <li>• PPG signal susceptible to motion/light/pressure</li> <li>• Accuracy model-dependent; generalization challenges</li> <li>• Calibration required</li> </ul> |

## 2.2.6 Discussion

Cuffless BP monitoring is a dynamic and rapidly evolving field, with diverse methodologies demonstrating strengths and limitations. As compared in Table 2.2, method based on physiological signals, particular ECG and/or PPG, exhibits exceptional promise. Their distinct advantages include low cost and widespread availability of sensors in existing consumer electronics, such as smartwatches and fitness trackers, which pave the way for cost-effective and highly accessible BP monitoring solutions. The ability of these sensors to facilitate real continuous beat-to-beat and long-term monitoring provides data for proactive hypertension management and understanding BP variability. Therefore, this thesis will focus on ECG and/or PPG based cuffless BP estimation methods, analyze the primary limitations which prevent this technique for widespread real-world adoption and propose solutions to address these limitations.

## 2.3 Review of ECG and/or PPG Based Studies

Cuffless BP estimation which uses ECG and/or PPG signals with machine learning models to estimate BPs, is fundamentally a supervised regression problem. Current research within this paradigm can be categorized by the model’s input-output architecture: feature-to-BP (denoted as **feat2bp**) approaches estimate BP from extracted physiological features; signal-to-BP (denoted as **sig2bp**) methods directly predict BP from raw physiological signals; and signal-to-signal (denoted as **sig2sig**) techniques first reconstruct an [Arterial Blood Pressure \(ABP\)](#) waveform from input signals, from which BP values are then derived.

In this section, a selection of current [SOTA](#) ECG and/or PPG based cuffless BP measurement studies is reviewed. The subsequent subsections, along with Table 2.7, will detail these studies and include a discussion of their advantages and limitations.

### 2.3.1 Predetermined Feature-Based Methods (feat2bp)

#### 2.3.1.1 PTT/PAT Based Method

A number of continuous cuffless BP measurement methods have been developed over the last decade and most of them employed [PTT](#) and [PAT](#). PTT refers to the time it takes an arterial pulse wave to travel between two arterial sites (usually from the heart to the periphery) and it is inversely related to [PWV](#) which is further related to BP. PAT which can

be approximated by the time interval from the R-peak on an ECG to the peripheral arrival of the pulse wave, for example, a PPG waveform (as demonstrated in Figure 2.4). PAT has been used as a surrogate of PTT for noninvasive BP estimation [69]. The theoretical relation between PTT and BP can be expressed by Moens-Kortweg (M-K) equation [70] which models the relationship between PWV and the incremental elastic modulus of the arterial wall or its distensibility.

$$\text{PWV} = \frac{L}{\text{PTT}} = \sqrt{\frac{E \cdot h}{2r\rho}} \quad (2.1)$$

where,  $L$  is the pulse transit distance, PTT is the pulse transit time,  $\rho$  is the blood density,  $r$  is the inner radius of the vessel,  $h$  is the vessel wall thickness, and  $E$  is the elastic modulus of the vascular wall. By using different assumptions and different approximations, different relationships between PTT and BP can be drawn from Equation (2.1) [71]. For example, the relationship between the elastic modulus  $E$  and BP can be expressed by applying an empirical exponential relation [72].

$$E = E_0 \cdot e^{\kappa \cdot (\text{BP} - \text{BP}_0)} \quad (2.2)$$

where,  $\kappa$  is a constant,  $E_0$  and  $\text{BP}_0$  are nominal values of Young's modulus and BP respectively. By setting different assumptions and approximations of the variables in Equation (2.1) and Equation (2.2), different relations between PTT and BP can be derived. For example, by assuming all other variables are constant, a logarithmic dependency between BP and PTT can be extracted from Equation (2.1) and Equation (2.2). Thus, the relationship between PTT and BP can be expressed by:

$$\text{BP} = \alpha + \beta \cdot \ln(\text{PTT}) \quad (2.3)$$

The coefficients  $\alpha$  and  $\beta$  are optimized based on the acquired simultaneous BP and PTT. As aforementioned, PAT has been reported can be used as the surrogate of PTT in BP estimation. Therefore, the non-linear relation Equation (2.3) can be adopted by replacing PTT with PAT so that  $\text{BP} = \alpha + \beta \cdot \ln(\text{PAT})$ .

Some reported different mathematical models derived on a basis of M-K equation with different assumptions can be found in Table 2.3.

Table 2.3: Different mathematical models to calculate BP from PAT and RMSE for SBP and DBP estimations (unit: mmHg)

| Models      | BP model                                                            | Dataset                                    | RMSE       |
|-------------|---------------------------------------------------------------------|--------------------------------------------|------------|
| Model1 [73] | SBP: $SBP_0 - (2/(\gamma * PAT_0)) * (PAT - PAT_0)$                 | 12 subjects                                | SBP: 6.09  |
|             | DBP: $SBP - PP_0 * (PAT_0/PAT)^2$                                   | (11 males, 1 female)                       | DBP: 3.27  |
| Model2 [74] | SBP: $DBP + PP_0 * (PAT_0/PAT)^2$                                   | 24 subjects                                | SBP: 3.43  |
|             | DBP: $MBP_0 + (2/\gamma) \ln(PAT_0/PAT) - (PP_0/3) * (PAT_0/PAT)^2$ | (15 elderly hypertensive, 9 young healthy) | DBP: 3.32  |
| Model3 [75] | SBP: $a_3 + \sqrt{b_3 + c_3 * (1/PAT_3)}$                           | 32 healthy subjects                        | SBP: 5.77  |
|             | DBP: $a_3' + \sqrt{b_3' + c_3' * (1/PAT^2)}$                        | (24 males, 8 females)                      | DBP: 4.22  |
| Model4 [72] | SBP: $a_2 + b_2 * PAT$                                              | 20 patients                                | SBP: 21.63 |
|             | DBP: $a_2' + b_2' * PAT$                                            | (14 males, 6 females)                      | DBP: 10.21 |

**Note:**  $\gamma$  denotes a vascular information parameter which might be altered with age and the development of CVD. For healthy subjects, it is  $0.031 mm \cdot Hg^{-1}$  [74].  $PP_0 = SBP_0 - DBP_0$ ,  $MBP_0 = (1/3)SBP_0 + (2/3)DBP_0$ .  $SBP_0$ ,  $DBP_0$ ,  $PP_0$  are the initial measurements conducted by cuff-based devices. The RMSE is calculated for each subject [71] (personalized model)

### 2.3.1.2 Pulse Wave Analysis (PWA) Based Methods

To address the limitations and inconvenience of PAT/PTT-based BP estimation, an alternative approach known as [Pulse Wave Analysis \(PWA\)](#) has gained prominence. PWA methods employ machine learning models that utilize a broader array of physiological features extracted from waveforms like PPG and/or ECG, rather than relying solely on PAT/PTT, and have demonstrated promising results in cuffless BP estimation.

A number of features extracted from both the time domain and frequency domain of PPG signals and even features extracted from first and second-order derivatives of the PPG waveform (common fiducial points of PPG and its derivatives are demonstrated in Figure 2.4), have been reported to correlate with BP changes [76–79]. In this section, various studies which extract predetermined features to be used as inputs to machine learning models for BP estimation will be reviewed.

Several researchers have utilized machine learning algorithms to predict BP based on features other than PTT and PAT extracted from physiological signals. In 2009, Suzuki and Oguri [80] suggested a pioneering method for feature-based BP estimate using AdaBoost model trained with a large amount of data (160 males and 208 females), and the mean error for SBP estimation was  $-1.2 \pm 11.7$  mmHg. In 2013, Kurylyak *et al.* [81] employed [Artificial Neural Network \(ANN\)](#) to predict BP using 21 predetermined PPG features. Miao *et al.* [82] in 2017 retrieved 14 features from simultaneous ECG and PPG data for beat-to-beat BP estimation, with a genetic algorithm applied for feature selection as BP indicators for each subject. The BP estimation model was developed using multivariate linear regression and support vector regression. The accuracy of BP prediction as demonstrated by this

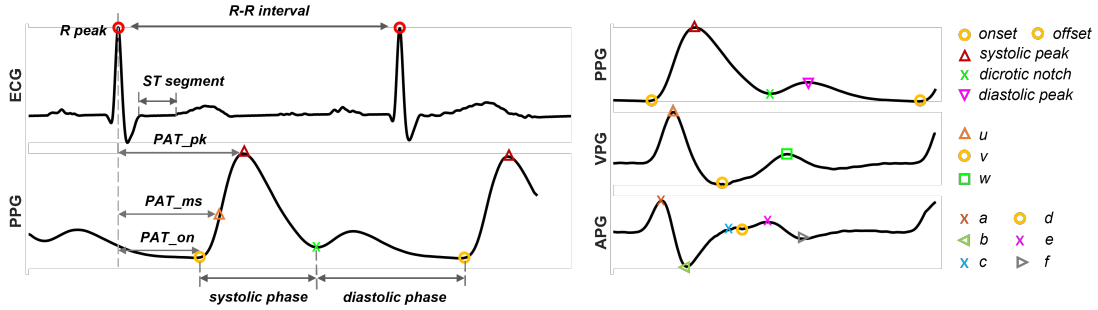


Figure 2.4: A demonstration of fiducial points of ECG, PPG, the first derivative of PPG (VPG) and the second derivative of PPG (APG), as well as PATs which can be extracted as the time intervals between ECG R peaks and PPG onsets, max\_slope points and systolic peaks (x axis: time (s), y axis: amplitude).

study is relatively satisfactory in terms of standard deviation, this is likely due to the fact that the features are carefully selected for physiological reasons. Khalid *et al.* [83] in 2018 only utilized five pre-determined PPG features, such as pulse area and pulse rising time, for BP estimation using a regression tree model. The MAE was reported to be  $4.82 \pm 4.31$  mmHg for SBP and  $3.25 \pm 4.17$  mmHg for DBP, which makes this work one of the first to demonstrate satisfactory outcomes using a single PPG sensor for estimating BP. In 2018, Su *et al.* [84] presented a novel deep Recurrent Neural Network (RNN) consisting of a bidirectional long short-term memory (BiLSTM) to access larger temporal scale context information of input sequence and model the temporal dependencies in BP dynamics to overcome accuracy decay over time. Seven predetermined features are used as BiLSTM inputs. The study improves the standard deviations of the estimation error of long-term BP prediction to within 5.81 mmHg for SBP and 5.21 mmHg for DBP.

In 2020, Khalid *et al.* [85] proposed a two-step algorithm using a regression tree with predetermined features to classify subjects into groups of hypotensives, normotensives and hypertensives, the mean error is  $0.07 \pm 7.1$  mmHg for SBP and  $-0.08 \pm 6.0$  mmHg for DBP. Song *et al.* [86] in 2020 developed a two-stage system of stacked deep neural networks to estimate BPs. The first stage of the system was to extract features such as PTTs from collected ECG and PPG signals and the second stage was to estimate BPs. This study results in the mean error of SBP  $-3.0 \pm 6.0$  mmHg and DBP  $-1.5 \pm 6.0$  mmHg. Li *et al.* [77] in 2020 utilized the BiLSTM and a residual connection inside for BP estimation. However there are only seven predetermined features used as the model inputs and the concluded accuracy in terms of standard deviations is not small, 14.51 mmHg and 6.44 mmHg for SBP and DBP respectively. In 2020, Hsu *et al.* [76] dedicated effort to extracting

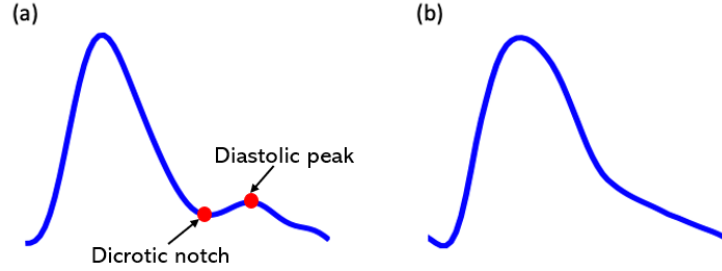


Figure 2.5: (a) an example of a regular PPG pulse with distinct dicrotic notch and diastolic peak; (b) an example of PPG pulse without distinct dicrotic notch and diastolic peak.

59 predetermined features from raw PPG and its first derivative, along with 32 features selected as the input of the fully-connected DNN for BP estimation. This research includes data of 2,176,188 heart pulses to calibrate and validate the proposed model. With [Root Mean Squared Error \(RMSE\)](#) of SBP and DBP of 4.63 mmHg and 3.0 mmHg, respectively, the resulting accuracy is promising. El-Hajj and Kyriacou [87] in 2021 proposed a deep learning model of bi-directional gated recurrent units (Bi-GRU) with attention mechanism and 22 pre-determined features are extracted from PPG waveform as model inputs. The resulted estimation accuracy in terms of standard deviations of 4.22 mmHg and 2.07 mmHg for SBP and DBP, respectively.

Despite the fact that the predetermined feature based cuffless BP studies have yielded promising results, there are some drawbacks. Firstly, the PPG waveforms are highly dependent on the cardiovascular characteristics for each individual. For example, the dicrotic notch as seen in Figure 2.5 may be nonexistent for the people in older ages and/or with chronic cardiovascular disease, and specific algorithm designed to extract many predetermined features could end up with erroneous features and thus cause unreliable BP estimations for these people. In addition, the optical properties of skin, tissue, fat and bones around and above the blood vessels to detect its pulsation among different subjects also affect the PPG waveform, not to mention possible mis-positioning of measuring location, motion artifacts and the varied optical performances among differently-designed PPG devices.

### 2.3.2 Non-predetermined Feature-Based Methods (sig2bp)

In recent years, some works have utilized machine learning methods to address the inadequacy of predetermined features in BP estimation and the difficulty of extracting predeter-



mined temporal and spectral features from physiological signals (ECG, PPG, etc) due to the difference of cardiovascular characteristics between individuals. These studies aimed to combine both feature extraction and regression for BP prediction into a single model.

A hierarchical artificial neural network with long short-term memory (ANN-LSTM) was proposed in 2019 for BP estimation [84]. The model consists of two hierarchy levels, with the lower hierarchy level employing ANNs to extract a number of non-predetermined morphological features from ECG and PPG waveforms and the top hierarchy level employing [Long Short-Term Memory \(LSTM\)](#) layers to account for the time domain variation of the features extracted by the lower level. Thus, the efficiency of extracting characteristics from raw PPG/ECG data can be significantly enhanced. Based on the data from 39 participants, this study yields estimation errors of  $1.10 \pm 1.56$  mmHg for SBP and  $0.58 \pm 0.85$  for DBP. Slapnicar *et al.* [88] reported in 2019 a spectro-temporal deep neural network using ResNet with residual connections, designed to extract non-predetermined features and subsequently estimate BP. This study also compared the accuracy of BP predictions using the proposed model and a standard regression with 12 predefined temporal and frequency features. The MAEs of 9.43 mmHg and 6.9 mmHg for SBP and DBP respectively which is better than the traditional regression method, however, these results are still relatively large and do not meet [AAMI](#) or [British Hypertension Society \(BHS\)](#) standards. In 2020, Landry *et al.* [89] developed a novel nonlinear dynamic autoregressive model with exogenous input by employing an artificial neural network to model the BP waveforms using ECG and/or PPG, extracting non-predetermined features, and then predicting BPs. The final mean error of BP estimation for this study was  $-1.2 \pm 17.8$  mmHg and  $-0.6 \pm 8.3$  mmHg for SBP and DBP, respectively, with a much larger standard deviation for SBP than other studies indicated in Table. In 2020, Panwar *et al.* [90] proposed a model comprised of [Convolutional Neural Network \(CNN\)](#) and LSTM, named as a [Long-term Recurrent Convolutional Network \(LRCN\)](#). CNN-LSTM is conceptually similar to Tanveer and Hasan’s ANN-LSTM [91], where CNN is designed to extract features and LSTM to estimate BP. The LRCN is expected to be both effective and lightweight. To calibrate and test the proposed LRCN, the 1557 subjects’ data from MIMIC II were evaluated. The reported standard deviations in terms of accuracy for SBP and DBP are 5.65 mmHg and 5.41 mmHg, respectively. To estimate BP, Aguirre *et al.* [92] developed an [RNN](#) encoder-decoder architecture comprised of [Gated Recurrent Unit \(GRU\)](#) and an attention network in 2021, the estimation errors in terms of the standard deviation are relatively large with 15.67 mmHg and 7.32 mmHg for SBP and DBP. In 2021, a combined deep CNN-LSTM multitasking network was developed for estimating BPs from ECG and PPG waveforms and attained Mean Estimation Errors (MEs) of  $0.0 \pm 1.6$  mmHg for SBP and  $0.2 \pm 1.3$  mmHg for DBP [93].

### 2.3.3 ABP Waveform Prediction (sig2sig)

The morphology of the PPG and BP waveforms has been proved to be highly correlated in [94] (overall correlation coefficient  $r > 0.9$ , see Figure 2.6) which indicates the feasibility of modeling BP waveforms directly from simultaneous PPG waveform and BP values can be then extracted from predicted ABP waveforms. In [95], a U-Net-based deep learning model was developed to reconstruct BP waveforms from PPG waveforms and achieved satisfactory results, an example of PPG-BP waveform mapping is shown in Figure 2.7.

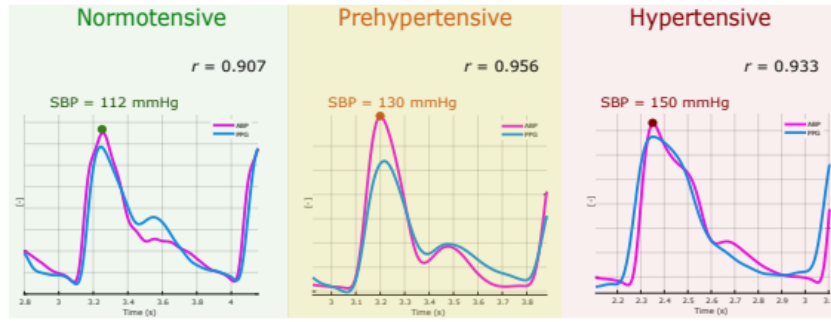


Figure 2.6: Examples of PPG and ABP morphology similarity analysis.  $r$  refers to the morphology correlation between PPG and ABP waveforms (adapted from [94], y axis: normalized amplitude).

### 2.3.4 Calibration for Cuffless BP Estimation

Calibration is critically required for new subjects because substantial inter-subject physiological variability means that generalized cuffless models, derived from ECG and PPG features, often fail to accurately map these features to an individual's true BP without subject-specific adaptation. This personalization is essential as the relationship between physiological parameters and BP is significantly influenced by individual-specific factors such as arterial stiffness, vascular tone, and anatomical differences, which generic models cannot adequately capture.

The "calibration" for cuffless BP estimation differs significantly from the calibration process conducted for conventional BP devices described in Section 2.1.4. In the domain of cuffless BP estimation utilizing physiological signals such as the ECG and/or PPG, **calibration** is defined as the process of personalizing a pre-trained BP estimation model to a new subject. This procedure is critical to account for the inherent inter-subject

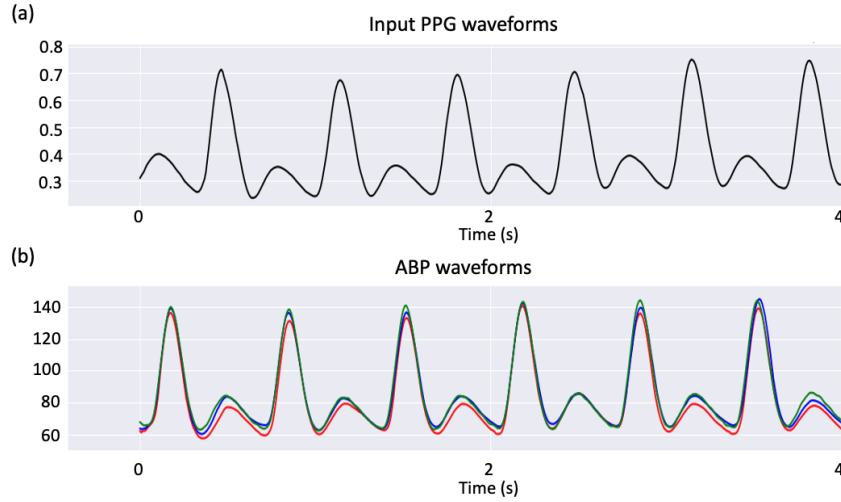


Figure 2.7: An example of the BP waveform modelling using PPG waveform demonstrated in [95]. (a). the PPG waveform used as input of the model; (b). the predicted ABP waveform using **approximation network** (red), the predicted waveform using **refinement network** (blue) and the ground truth ABP waveform (green).

physiological variability, which significantly influences the relationship between measured physiological features and actual BP values. The primary purpose of calibration is to enhance the accuracy and reliability of BP estimates for a specific individual by adapting the model to their unique cardiovascular characteristics. The calibration process typically involves the following sequential steps:

1. Concurrent data acquisition: simultaneous recording of the subject's ECG and PPG waveforms. Contemporaneous measurement of subject's reference BP values using a validated, gold-standard cuff-based device (e.g., an oscillometric or auscultatory sphygmomanometer). This is usually performed at one or multiple discrete time points.
2. Model Personalization and Adjustment: Extraction of relevant physiological features from the recorded ECG and PPG signals (e.g., PAT, morphological features of the waveforms). Then utilization of these extracted features along with the corresponding reference BP readings to adjust, fine-tune or derive the parameters of the cuffless BP estimation model specifically for that new subject.

Common calibration schemes employed by previous studies:

- Initial calibration: this is the most prevalent approach, where a new user provides one or more initial cuff-based BP readings during simultaneous ECG and PPG recording. These paired data points are then used to optimize the subject-specific model parameters. This can be a single point calibration or involve multiple readings to create a more robust personalized model.
- Periodic recalibration: due to physiological changes over time (e.g., vascular aging, lifestyle changes, medication effects), the accuracy of the initial calibration can degrade. Consequently, periodic recalibration with a reference cuff device (e.g., daily, weekly, monthly, and annually) is often recommended to maintain the model’s performance. However, the optimal frequency for recalibration is still an open problem.

### 2.3.5 Current Cuffless BP Devices

A number of wearable cuffless BP monitoring devices have been commercialized. For example, the Biobeat Tech BB-613WP (see Figure 2.3 (c)) device was cleared as cuffless BP monitoring device by the U.S. FDA in August 2019. According to the report [96], this device predicts BP using the traditional PTT method based on both ECG and PPG signals. This device, however, requires calibration every three months for each user. Neither the techniques of signal preprocessing used for retaining quality PPG waveforms nor how the features of PTT are extracted in real time were disclosed to the public. Samsung revealed that the software for estimating BP for Galaxy Watch Active2 (see Figure 2.3 (d)) and newer models have been approved by South Korea’s Ministry of Food and Drug Safety (MFDS), but there is no technical validation report available but a summary [97].

Although these products are already available to consumers, their unknown monitoring accuracy and reliability, frequent calibrations required prevent them from being widely used for clinical diagnosis purpose.

## 2.4 Limitations of Current Cuffless BP Studies

### 2.4.1 Generalization and Personalization

Despite the promising results shown in current studies, cuffless BP estimation remains far from real-world deployment, primarily due to challenges in model generalization and the lack of effective personalization strategies.

#### 2.4.1.1 Inadequate Model Generalization

Cuffless BP estimation, which aims to predict BP values from physiological features or directly from signals such as PPG and ECG, is fundamentally a supervised regression task. In this context, the concept of model generalization is paramount. Model generalization refers to a machine learning model's capacity to maintain performance on new, previously unseen data that originates from the same underlying distribution as data used for training. The significance of robust model generalization is amplified in healthcare applications like cuffless BP monitoring, where the primary goal is to develop reliable devices for continuous and unobtrusive patient monitoring. Moreover, physiological signals are known to exhibit significant inter-subject variability, meaning that individuals can have distinct signal characteristics (then cause completely different features extracted) even under similar physiological conditions. Inaccurate predictions on new individuals can lead to severe consequences, such as delayed recognition of hypertension.

However, a large number of studies are not cautious about data used to train, validate and test their proposed model and cause inflated model performance due to potential subject-level data usage. Some examples of statements regarding data usage for model training and evaluation from existing studies are shown as follows:

"After processing of the dataset, 4,254 records were obtained. **The new dataset is then randomly partitioned into three sets: 60% as training, 20% as validation and the remaining as test samples.**" [98]

"The data were *shuffled and randomly selected* to train the deep learning model. In total, **60% of the data was used for training, 20% was used for validation, and 20% was used for testing.**" [99].

"**80% of the total dataset** was used to train the model and **20%** to evaluate the model performance. Furthermore, **10% of the training data** was used for validation to prevent the model from overfitting the training dataset." [93].

"**Seventy percent of the dataset** was used for training, **10%** for validation, and **20%** for testing." [100].

"Before the training of our proposed deep neural network predictor, the selected feature set is split into three parts, and part contains **70%, 20% and 10%** of the data, which serve as training, testing and validation datasets, respectively." [76].

"The dataset were divided into **70%** train, **15%** validation and **15%** test sets." [101].

"In this experiment, **60%** of the data is used as a training set, **20%** as a test set, and **20%** as a validation set." [102].

"We used **70%** of the total data for training our model, **15%** for validation, and the remaining **15%** for testing." [103].

"In total, there are 58,795 valid intervals of PPG signal (subject number is 72) and corresponding BP values for different people and different time instances. In order to avoid overfitting, we use **70%** of them for network training, **15%** of them for validation, and **15%** of them for testing." [104].

"In each of the four non-linear regression models, estimation of each of the BP values (DBP, MAP, SBP) is done by separating training and testing data with the 10 fold cross validation algorithm." [105].

These studies applied naive random split or k-fold cross validation may place data from the same subject in both training and testing sets. This causes subject-level data leakage: the model can learn the subject's baseline pulse morphology, distribution of features or calibration offset from the training data, and thus trivially predict that subject's BP in the test set. As Sánchez *et al.* explain in their study that randomly shuffling a dataset with multiple samples per subject "leads to examples of the same subject in different folds, i.e. information leakage". In this case, the model relies "mainly on subject-specific characteristics" and achieves artificially low errors [106]. Related studies also validated that random splits can overestimate model performance. In 2024, LaBerge found that random splitting improved SBP MAE by an average of  $\sim 13.8$  mmHg and DBP MAE by  $\sim 7.3$  mmHg compared to a patient-wise split. A number of BP estimation models proposed previously were re-implemented [106], and most methods exhibited their SBP error drop by approximately one half. For example, the deep ResNet reported MAE of SBP approximately 21.0 mmHg under subject-wise evaluation, but only  $\sim 6.3$  mmHg under the random split evaluation [107]. In a recent benchmark study (2025 Feb), a deep learning model XResNet1d101 was trained and evaluated on the PulseDB [108] dataset with different strategies. When using a subject-specific calibration approach (where part of data from the test subject are used in training phase), the MAEs were 9.4 mmHg for SBP and 6.0 mmHg for DBP. However, when the same model was evaluated in a calibration-free manner (where the data from the test subject was never used for training), the MAEs increased significantly to 14.0 mmHg for SBP and 8.5 mmHg for DBP. The performance degradation was even more significant when these calibration-free models were tested on a different, out-of-distribution datasets, where SBP MAEs ranged from 15.0 mmHg to as high as 25.1 mmHg [109].

These examples consistently underscore a critical reality: models trained and evaluated without strict subject separation often learn subject-specific characteristics rather than generalizable physiological features indicative of BP. This will lead to a substantial per-

formance drop when these models are tested on completely new, unseen individuals. For example, several SOTA models were re-implemented and evaluated using two strategies: an intra-subject approach, where a separate model is trained for each individual, and an inter-subject approach, where a model is trained on a group of subjects and tested on entirely unseen subjects [110]. The results from these two evaluation strategies showed substantial differences, as demonstrated in Table 2.4.

Table 2.4: Evaluation of BP estimation in the MIMIC-II UCI-ML dataset with regard to MD, STD and MAE (adapted from [110]).

| Algorithm      |     | Intra-subject metrics (mmHg) |       |       | Inter-subject metrics (mmHg) |       |       |
|----------------|-----|------------------------------|-------|-------|------------------------------|-------|-------|
|                |     | MD                           | STD   | MAE   | MD                           | STD   | MAE   |
| XGBoost        | SBP | 0.07                         | 7.48  | 4.86  | 4.09                         | 20.49 | 16.67 |
|                | DBP | 0.02                         | 3.89  | 2.29  | 0.87                         | 9.60  | 7.14  |
| LightGBM       | SBP | 0.01                         | 11.00 | 8.00  | 0.26                         | 20.45 | 16.49 |
|                | DBP | 0.00                         | 5.65  | 3.77  | 0.10                         | 9.59  | 7.21  |
| CatBoost       | SBP | 0.01                         | 8.58  | 5.94  | 0.42                         | 20.29 | 16.43 |
|                | DBP | 0.01                         | 4.59  | 2.94  | 0.12                         | 9.77  | 7.38  |
| Residual U-Net | SBP | 0.41                         | 5.22  | 3.17  | -0.65                        | 24.08 | 19.10 |
|                | DBP | 0.07                         | 2.89  | 1.54  | -0.20                        | 11.09 | 8.18  |
| ResNet18       | SBP | 0.30                         | 9.94  | 7.23  | 2.04                         | 24.72 | 19.85 |
|                | DBP | -0.05                        | 5.75  | 4.05  | -0.34                        | 11.25 | 8.42  |
| ResNet-LSTM    | SBP | 0.75                         | 14.76 | 11.19 | 0.15                         | 22.30 | 17.71 |
|                | DBP | 0.02                         | 7.94  | 5.74  | -0.76                        | 10.26 | 7.96  |

In this thesis, experiments were also conducted to demonstrate the model generalization problem for cuffless BP estimation studies. Four SOTA models, MLP-based model [76], CNN-LSTM [93], HGCTNet [111] and UNet [95] which estimate BP using different input-output architectures were re-implemented (see Figure 2.8 for model structures). 680 subsets of the UCI Cuffless BP Estimation dataset were extracted and prepared in 10-fold leave-k-subject-out cross validation (k=68). In each fold, 68 subsets were held-out as testing set, the remaining 598 subsets were used as training set and 20% of each training subset are used for validation. The structures of 4 SOTA models are shown in Figure 2.8 and the inter-subject MAE between the reference and estimated BPs for test subsets are demonstrated in Table 2.5.

As demonstrated in Table 2.4 and Table 2.5, the MAE > 15 mmHg for SBP estimation and MAE > 7 mmHg for DBP estimation using different types of models, indicating the challenge of model generalization and also highlights the requirement of effective personalization strategies for accurate estimation for new subjects.

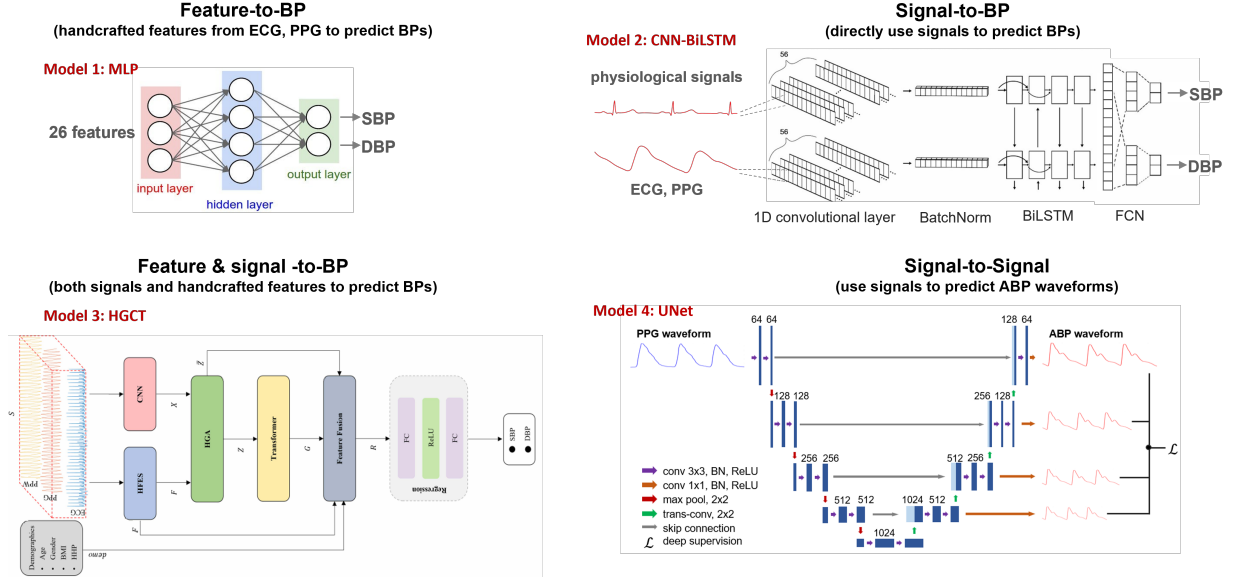


Figure 2.8: Four SOTA models with different input-output structures for cuffless BP estimation.

Table 2.5: Mean and standard deviation of absolute error between estimated BPs and reference BPs using 4 different types of SOTA models with UCI dataset.

|                              | SBP (AE) |       | DBP (AE) |      |
|------------------------------|----------|-------|----------|------|
|                              | mean     | std   | mean     | std  |
| <b>MLP</b> (feat2bp)         | 17.61    | 14.96 | 8.91     | 8.83 |
| <b>CNN-LSTM</b> (sig2bp)     | 17.58    | 12.83 | 8.63     | 7.69 |
| <b>HGCTNet</b> (sig&feat2bp) | 18.61    | 12.17 | 8.29     | 7.78 |
| <b>UNet</b> (sig2sig)        | 15.46    | 13.05 | 7.37     | 6.97 |



#### 2.4.1.2 Few-Shot Personalization

Although new cuffless BP estimation models continue to emerge, research specifically focused on the personalization of deep learning-based approaches remains limited. Slapničar *et al.* [112] proposed a personalization approach by incrementally training the entire pre-trained model using an additional 20% of data from a specific subject, which reduced the SBP/DBP MAE from 15.41 mmHg / 12.38 mmHg to 9.43 mmHg / 6.88 mmHg. In 2022, Leitner *et al.* [113] introduced BP-CRNN-Transfer, a recent approach to personalizing cuffless BP estimation models. Their method involves freezing the lower layers of the network and fine-tuning the last convolutional and fully connected layers, which reduces the SBP/DBP MAE from 4.59/2.27 mmHg to 3.52/2.20 mmHg. To mimic domain adaptation, the dataset is partitioned strictly by subjects to define source and target domains. However, this method requires a substantial amount of data—up to 3600 samples for optimal performance and at least 50 samples for fine-tuning—making it less practical for low-calibration scenarios. In 2024, a few-shot transfer learning method comprised of a modularized multi-modal network architecture called MSTNN and a personalization adapter (PA) was proposed to address the limitation of the previous studies that a large number of calibrations are required for personalization [114]. This study demonstrated the SBP MAE can be improved from 8.85 mmHg to 6.92 mmHg and DBP MAE can be improved from 4.82 mmHg to 3.73 mmHg by using 5 calibration measures for model personalization. However, the results barely satisfies IEEE 1708a-2019 evaluation standard [25] ( $MAE \leq 7$  mmHg) and the SBP estimation can only be considered as "Grade C".

#### 2.4.2 Unsatisfactory BP Trend Tracking Performance

Cuffless BP estimation technologies are inherently intended for continuous and longitudinal BP monitoring. However, the majority of current research predominantly concentrates on the accuracy of discrete BP value estimation, often neglecting a critical assessment of model performance in continuous BP trend tracking.

Although the current cuffless BP studies have shown promising results, existing performance validation standards are inadequate for this new research problem. In the last three decades, several organizations, such as the British Hypertension Society (BHS) [115], Association for Advancement of Medical Instrumentation (AAMI) [23], and the European Society of Hypertension (ESH) [116] have proposed their individual protocols of clinical validation for BP measurement devices. In 2018, a single universal standard (AAMI/ESH/ISO Universal Standard - ISO 81060-2:2018 [24]) was developed by AAMI, ESH and ISO experts

for global use of BP device validation and it was expected to replace all other previous standards/protocols. Although these validation protocols were designed for cuff-based devices and for intermittent assessments, the majority of current cuffless BP studies [76, 87, 117] still employed these protocols to evaluate the performance of their proposed approaches, which was considered inadequate and misleading for continuous and cuffless BP measurement [118]. Therefore, new standards were adjusted and adapted based on aforementioned cuff-based protocols for cuffless BP devices specifically. In 2014, the IEEE (Institute of Electrical and Electronics Engineers) published the first standard for validating the accuracy of cuffless wearable BP devices (IEEE 1708-2014) [119] and an amendment of this protocol was updated in 2019 (IEEE 1708a-2019) [25]. ISO (International Organization for Standardization) published their protocol ISO 81060-3:2022 [26] which is specific for continuous noninvasive sphygmomanometers in 2022. In 2023, EHS also proposed their recommendations for the validation of cuffless BP measuring devices [120]. A summary of key aspects of validation protocols from several aforementioned standards are listed in Table 2.6.

Table 2.6: Key aspects of validation protocols developed for cuff-based and cuffless BP devices.

|                                | ANSI/AAMI<br>SP10:2002 [23]        | ISO Standard<br>81060-2:2018 [24]                  | IEEE Standard<br>1708a-2019 [25] | ISO Standard<br>81060-3:2022 [26]  | ESH<br>recommendation 2023 [120]   |
|--------------------------------|------------------------------------|----------------------------------------------------|----------------------------------|------------------------------------|------------------------------------|
| <b>Device type</b>             | cuff-based; intermittent           | cuff-based; intermittent                           | cuffless; intermittent           | cuffless; continuous               | cuffless; continuous               |
| <b>Reference method</b>        | Ausculatory;<br>intra-arterial     | mercury sphygmomanometers;<br>accurate non-mercury | ausculatory                      | intra-arterial                     | ausculatory<br>24-h oscillometric  |
| <b>Main accuracy metrics</b>   | MD;<br>SDD                         | MD;<br>SDD                                         | MAD                              | MD;<br>SDD                         | MD;<br>SDD                         |
| <b>Pass requirement (mmHg)</b> | $-5 \leq MD \leq 5$<br>$SD \leq 8$ | $-5 \leq MD \leq 5$<br>$SD \leq 8$                 | $MAD \leq 7$                     | $-5 \leq MD \leq 5$<br>$SD \leq 8$ | $-5 \leq MD \leq 5$<br>$SD \leq 8$ |

MD: mean estimation difference; SDD: standard deviation of estimation difference. See Table 4.1.

MAD: mean absolute estimation difference. See Table 4.1.

In Table 2.6, several similarities can be observed among different protocols for both cuff-based and cuffless devices. The pairwise distances between the estimations and references are used as the primary metrics to validate the performance of the proposed device against the reference measurements. However, the results obtained from experiments indicate that sometimes an estimation model may meet the distance criteria, even though it may not accurately monitor BP changes. For example, the  $MAD \pm SDAD$  of a 543 seconds' testing subset shown in Figure 2.9(a) is  $(3.77 \pm 2.51)$  mmHg for SBP which is considered as "Grade A" of IEEE Standard 1708a-2019 [25]. However, the model is unable to track one acute BP change occurred between 325 and 375 seconds. We believe this is a general but underestimated problem of the existing studies. It is because the existing validation standards

do not well define the metrics to evaluate the ability of the proposed methods/devices in tracking BP changes.

Recently, the increasing number of studies revealed the BP pattern is related to different diseases and it is also important to accurately track BP patterns for both short-term and long-term. BP pattern refers to the morphology of sequential BP readings from a certain period of time which consists a number of monotonous trends. For example, several studies [121–123] observed that BP has a daily pattern and a global experiment [121] of 144,509 brachial SBP measurements from 2,423 healthy subjects is shown in Figure 2.9(b). It can be observed that BP starts to rise a few hours before a person wakes up and continuous to increase during the day, peaking in midday. Then BP typically decreases in the late afternoon and evening, and even lower at night while sleeping. An accurate measure of BP pattern can assess BP changes and BP variability. Based on the time window BP is assessed, BP pattern can be categorized and measured in different ranges, including very short-term (beat-to-beat, minute-to-minute), short-term (within 24h, hour-to-hour, day-to-night), long-term (day-to-day, months and years) [124, 125].

Abnormal BP patterns may indicate specific forms of hypertension (e.g. nocturnal hypertension) or may be associated with concomitant illness (e.g. sleep apnea) and recognition of these abnormal BP patterns can be helpful in diagnosis and management of hypertension [126]. Cardiovascular analysis indicates that obstructive sleep apnea (OSA) will cause acute hemodynamic changes [127]; both BP and heart rate (HR) initially decrease and then progressively increase by approximately 20 mmHg and 15 bpm, respectively [128]. Orthostatic hypotension, which indicates impaired physical performance and cerebral autoregulation, also causes rapid BP changes, SBP declines at least 20 mmHg and/or DBP declines at least 10 mmHg within 3 minutes (most declines occur within the first minute [129]) after standing up [130]. Activities can also cause BP variations, as shown in Figure 2.10. Figure 2.10(a) illustrates the BP waveform of a subject performing multiple hand grippings for one minute and then recovered. the BP decreases when the force applied to the hand grip increases, and there is a rapid BP drop when hand griping is stopped (e.g., a 32 mmHg decrease of SBP from 60s to 65s) and then BP gradually recovers. Figure 2.10(b) illustrates the BP steadily increases when a participant walks for one minute (SBP increases from 128 mmHg to 152 mmHg), study [131] also reveals that SBP increases significantly and proportionally to workload increase during exercise test in healthy adults. Figure 2.10(c) illustrates BP changes responding to Valsalva Maneuver (VM), VM can be divided into four phases based on hemodynamic changes. The BP will transiently increase in Phase 1 and then steadily decrease in Phase 2. In Phase 3, BP will further decreases and increases and overshoot in Phase 4, and gradually recover to normal. Comparing to exercises which steadily change BP, VM can significantly change BP within a very short

time, the mean and standard deviation of SBP changes ( $\max(\text{SBP}_{\text{VM}}) - \min(\text{SBP}_{\text{VM}})$ , e.g.,  $174-85=89\text{mmHg}$  in Figure 2.10(c)) is  $115 \pm 21 \text{ mmHg}$  concluded from 5 subjects' 19 VM segments in DailA BP dataset (see Section 4.3.2 for details). Figure 2.10(d) illustrates BP decreases when a subject changes posture from sitting to standing and then recovers to normal. The work [132] indicates 24-hour BP values are linearly related to stroke risk and the Dublin Outcome study concluded that nocturnal BP was overall the best predictor of cardiovascular risk, with a 10 mmHg increase in mean nocturnal SBP being associated with a 21% increase in cardiovascular mortality [133]. It was also studied that sleep phases, such as rapid eye movement and non-rapid eye movement, have different BP patterns [134]. Therefore, it is vital for future studies to evaluate the performance of BP pattern tracking with their estimation models, and eventually facilitating the adoption of continuous cuffless BP estimation in real-world clinical applications.

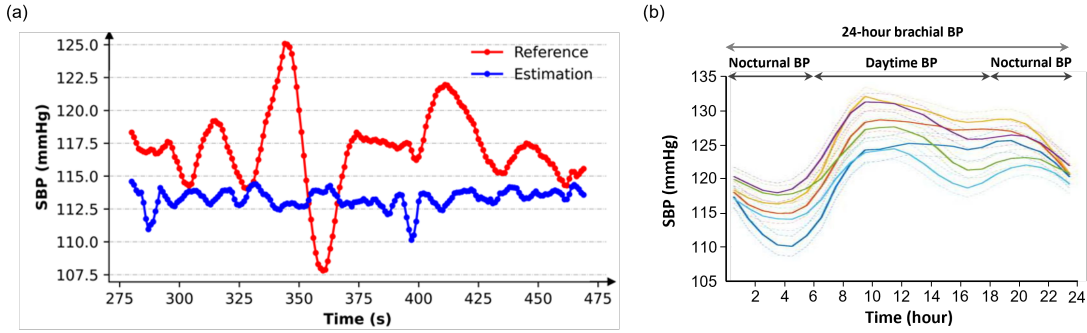


Figure 2.9: (a) An example of the proposed BP model unable to track BP changes but the overall MAD and SDAD is low. This segment is 543 seconds,  $\text{MAD} \pm \text{SDAD}$  for SBP is  $(3.77 \pm 2.51) \text{ mmHg}$ . (b) 24-hour ambulatory brachial SBP pattern, stratified by age. Solid lines are mean values and dash lines are 95% confidence intervals (reproduction of the middle figure of Figure 2 demonstrated in work [121], author authorized the use of this figure in this study).

## 2.5 Conclusion and Discussion

Despite the fact that the existing SOTA cuffless BP estimation research have yielded encouraging results. This new BP monitoring method remains far from achieving widespread real-world application due to aforementioned limitations.

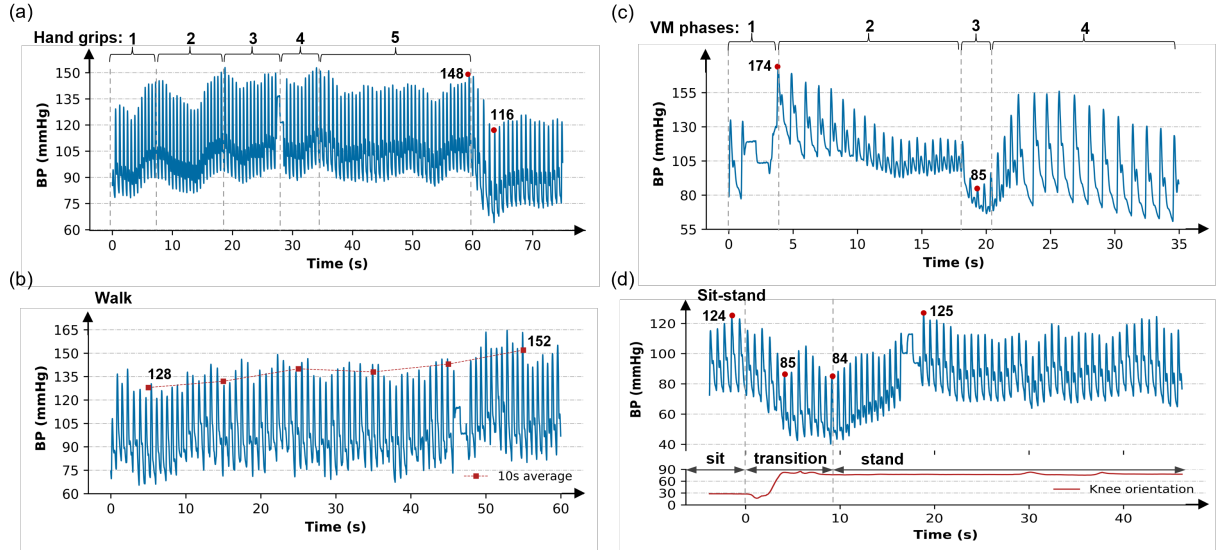


Figure 2.10: (a) BP waveform when hand grips were performed by squeezing a rolled towel at maximum voluntary contraction (0-60s) and at rest (after 60 s), (b) BP waveform when the subject was walking for 1 minute, (c) BP waveform when Valsalva maneuver was conducted, the BP reached maximal during Phase 1, (d) an example of BP changes extracted from **DailA BP** when a subject changed posture from sitting to standing, the lower track is the orientation of subject's knee (unit: degrees),  $0^\circ$  indicates thighs are parallel to ground (sitting)  $90^\circ$  indicates thighs are perpendicular to ground (standing).

Consequently, this study will analyze the limitations of the current methods and focus on several key problems to improve the generalization of BP estimation models and accuracy of BP trend tracking.

Table 2.7: A summary of existing state-of-art cuffless BP estimation studies.

| Work (year)   | Sensors (databases)             | Model inputs                                                     | Models                      | Training strategy | Accuracy                                            |
|---------------|---------------------------------|------------------------------------------------------------------|-----------------------------|-------------------|-----------------------------------------------------|
| [83]<br>2018  | PPG<br>(Queensland)             | 5 features<br>(pre-determined)                                   | Regression tree             | G*                | SBP $-0.1 \pm 6.5^*$<br>DBP $-0.6 \pm 5.2^*$        |
| [104]<br>2018 | PPG<br>(MIMIC)                  | 22 features<br>(pre-determined)                                  | ANN                         | P                 | SBP $4.02 \pm 2.8^{**}$<br>DBP $2.27 \pm 1.8^{**}$  |
| [84]<br>2018  | PPG ECG<br>(MIMIC)              | 7 features<br>(pre-determined)                                   | Deep RNN,<br>LSTM           | P                 | SBP $N/A \pm 5.8^{**}$<br>DBP $N/A \pm 5.2^{**}$    |
| [135]<br>2018 | PPG<br>(MIMIC)                  | 13 features of PPG/dPPG<br>(pre-determined)                      | Linear regression           | P                 | SBP $4.06 \pm 9.2^{**}$<br>DBP $3.99 \pm 5.2^{**}$  |
| [136]<br>2018 | PPG<br>(MIMIC)                  | PPG features<br>(pre-determined)                                 | Regression tree             | P                 | SBP $7.87 \pm 7.5^{**}$<br>DBP $3.84 \pm 3.6^{**}$  |
| [137]<br>2018 | PPG<br>(MIMIC)                  | features from time<br>and frequency domain<br>(pre-determined)   | ELM                         | G                 | SBP $4.25 \pm 5.2^{**}$<br>DBP $3.95 \pm 4.9^{**}$  |
| [105]<br>2019 | PPG<br>(MIMIC)                  | features from PPG/SDPPG<br>(pre-determined)                      | AdaBoost                    | G*                | SBP $0.19 \pm 4.18^*$<br>DBP $-0.05 \pm 8.9^*$      |
| [85]<br>2020  | PPG<br>(MIMIC<br>Queensland)    | 16 features<br>(pre-determined)                                  | Regression tree             | G*                | SBP $0.07 \pm 7.1^*$<br>DBP $-0.08 \pm 6.0^*$       |
| [79]<br>2020  | PPG<br>(222 subjects)           | selected from 101 time and frequency domain<br>(pre-determined)  | Gaussian process regression | G                 | SBP $3.02 \pm 9.3^{**}$<br>DBP $1.74 \pm 5.5^{**}$  |
| [86]<br>2020  | PPG ECG<br>(110 subjects)       | features from<br>ECG and PPG                                     | DNN                         | P                 | SBP $-4.80 \pm 4.7^{**}$<br>DBP $4.8 \pm 3.9^{**}$  |
| [77]<br>2020  | PPG ECG<br>(MIMIC II)           | 7 features from ECG/PPG<br>(pre-determined)                      | Bi-LSTM                     | G*                | SBP $4.64 \pm 14.5^{**}$<br>DBP $3.16 \pm 6.4^{**}$ |
| [76]<br>2020  | PPG ECG<br>(MIMIC II)           | 59 features from ECG/PPG<br>(pre-determined)                     | Fully-connected<br>DNN      | G                 | SBP $3.21 \pm 4.6^{**}$<br>DBP $2.23 \pm 3.2^{**}$  |
| [87]<br>2021  | PPG<br>(MIMIC II)               | 59 PPG features<br>(pre-determined)                              | Bi-GRU,<br>Attention        | G                 | SBP $-1.26 \pm 1.6^{**}$<br>DBP $2.58 \pm 3.4^{**}$ |
| [99]<br>2021  | ECG/PPG<br>(18 subjects)        | 7 features from ECG and PPG<br>(pre-determined)                  | Bi-LSTM                     | G*                | SBP $5.82 \pm 6.8^{**}$<br>DBP $5.24 \pm 6.1^{**}$  |
| [91]<br>2019  | PPG ECG<br>(MIMIC, 39 subjects) | 2 second segments,<br>features from ECG/PPG (not pre-determined) | ANN-LSTM                    | P                 | SBP $1.10 \pm 1.6^{**}$<br>DBP $0.58 \pm 0.9^{**}$  |

Continued on next page

Table 2.7 – continued from previous page

| Work<br>(year)                                                                                                                                   | Sensors<br>(databases)           | Model inputs                                                                                | Models              | Training<br>strategy | Accuracy                                           |
|--------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------|---------------------------------------------------------------------------------------------|---------------------|----------------------|----------------------------------------------------|
| [138]<br>2019                                                                                                                                    | PPG ECG<br>(MIMIC, 604 subjects) | 10 second segments,<br>continuous data sequence input<br>(not pre-determined)               | CNN                 | G*                   | SBP $3.09 \pm 2.8^{**}$<br>DBP $2.11 \pm 2.0^{**}$ |
| [112]<br>2019                                                                                                                                    | PPG<br>(MIMIC, 510 subjects)     | pulse segments of PPG, DPPG and SDPPG,<br>frequency domain features<br>(not pre-determined) | ResNet,<br>and LOSO | G                    | SBP $9.43 \pm N/A^{**}$<br>DBP $6.88 \pm N/A^{**}$ |
| [139]<br>2019                                                                                                                                    | PPG ECG<br>(MIMIC)               | Segments of the original,<br>FFT and derivatives<br>of signals<br>(not pre-determined)      | CNN                 | G*                   | SBP $5.32 \pm 5.5^{**}$<br>DBP $3.38 \pm 3.8^{**}$ |
| [100]<br>2020                                                                                                                                    | ECG, PPG, BCG<br>(15 subjects)   | 8 second segments,<br>CNN and Bi-BRU to extract<br>128 features                             | Bi-GRU<br>Attention | G*                   | SBP $4.06 \pm 4.0^{**}$<br>DBP $3.33 \pm 3.4^{**}$ |
| [89]<br>2020                                                                                                                                     | PPG, ECG<br>(MIMIC)              | ECG and PPG data<br>(not pre-determined)                                                    | ANN                 | P                    | SBP $-1.2 \pm 17.8^*$<br>DBP $-0.6 \pm 8.4^*$      |
| [90]<br>2020                                                                                                                                     | PPG<br>(MIMIC II, 1557 subjects) | 8 second segments<br>(not pre-determined)                                                   | LSTM                | G*                   | SBP $-1.25 \pm 5.7^*$<br>DBP $1.55 \pm 5.4^*$      |
| [103]<br>2021                                                                                                                                    | PPG<br>(MIMIC III)               | segments of 350 samples<br>(not pre-determined)                                             | U-Net               | G*                   | SBP $3.68 \pm 4.4^{**}$<br>DBP $1.97 \pm 2.9^{**}$ |
| <b>Training strategy:</b> P–Personalized model, G–Generalized model, G*–Partial generalized model(part of test subjects' data used for training) |                                  |                                                                                             |                     |                      |                                                    |
| <b>Accuracy:</b> *–mean error (ME), **–mean absolute error (MAE)                                                                                 |                                  |                                                                                             |                     |                      |                                                    |

## Chapter 3

# Respiratory Effect on Cuffless BP Estimation

In this chapter, the respiratory effect for cuffless BP estimation is comprehensively revealed. Then a VMD-based method is proposed to separate variations caused by respiration from physiological features and BPs before training an estimation model.

A version of the method described in this Chapter was prepared and submitted to IEEE Transaction on Biomedical Engineering.

### 3.1 Prior Insights on Respiratory Effects in Cuffless BP Estimation

Despite the advances of current cuffless BP estimation studies, this technique remains far from achieving widespread real-world application, and we hypothesize that addressing an ostensibly ordinary, yet overlooked, factor—**respiration**—could potentially enhance its performance. Respiration, which is the physiological process of gas exchange, has been considered by most of the previous studies as a main contributor to [Baseline Wandering \(BW\)](#) of physiological signals, such as ECG and PPG which are widely employed for cuffless BP estimation. As shown in Figure 3.1(a), the baseline level of the PPG waveform shows slow and low-frequency fluctuations over time which is in phase with simultaneous respiration waveform. BW can interfere with the analysis of signals, making it difficult to accurately measure the pulsatile component of the signal. For example, BW can cause inaccurate detection of ST segment (demonstrated in Figure 2.4) from ECG waveform



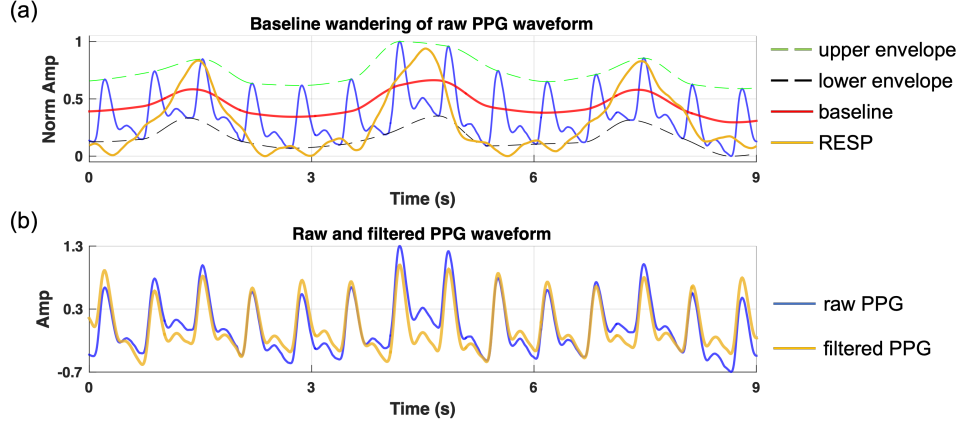


Figure 3.1: (a) A demonstration of baseline wandering (BW) caused by respiration on PPG waveform, (b) a 4th order Chebyshev type II bandpass filter with low cut-off frequency of 0.5 Hz and high cut-off frequency of 12 Hz applied to PPG waveform to remove BW.

and lead to misdiagnosis if not properly addressed [140]. Similarly, this low-frequency noise introduced by BW may distort PPG waveform, making it difficult to accurately identify fiducial points and potentially impact feature extraction for cuffless BP estimation. Therefore, existing cuffless BP estimation studies commonly employed filtering methods to remove BW before fiducial points detection and further analysis. It is generally accepted that normal [Respiration Rate \(RR\)](#) for an adult person ranges from 12 to 20 breaths per minute (0.2 Hz to 0.33 Hz) at rest [141]. Therefore, a bandpass filter, specifically a 4th order Chebyshev bandpass filter which can emphasize the difference between the systolic and diastolic waves of PPG and resulting more salient dicrotic notches [142, 143] (see Figure 5.9), is recommended to conserve the frequency content of the PPG pulse waves while filtering out lower-frequency content (e.g. BW caused by respiration) and higher-frequency content (e.g. muscle noise or power interference) [66, 79, 144]. Therefore the 0.5-12 Hz frequency band was recommended for filtering purpose [145]. A demonstration of a 9-second original PPG waveform and filtered PPG waveform using a 4th order Chebyshev type II bandpass filter with low cut-off frequency of 0.5 Hz and high cut-off frequency of 12 Hz is shown in Figure 3.1(b).

Besides the BW observed in physiological waveforms which is caused by respiration. Charleston *et al.* [146] further revealed respiration can also effects PPG waveform in terms of both amplitude modulation (AM) and frequency modulation (FM), and morphological features extracted from PPG waveform also illustrates respiratory variation which can be further applied for RR estimation. There have also been studies discussing the effects of

respiration on BP, [HRV](#) analysis [15, 17]. A previous related study [18] also discovered respiration can add variations to PAT which is a widely used feature for cuffless BP estimation, indicating [RM](#) is probably the most important reason for the inaccuracy of BP estimation with only the PAT and the requirement of intermittent calibration to maintain accuracy. Although these relevant findings revealed the overlooked role of respiration in various physiological measurements, it has never been comprehensively discussed and assessed for cuffless BP estimation. Therefore, this study is dedicated to validate the existence of RM in physiological waveforms, features extracted from these waveforms and eventually, its impact on cuffless BP estimation accuracy. Towards this objective, the rest of this chapter will be organized as follows. The mechanisms of RM are introduced, and its general existence in physiological features used for cuffless BP estimation are demonstrated. A [VMD](#) based method is proposed to separate RM component and a new metric–respiration modulation ratio (RMR)–is proposed to quantify the strength of RM for a wide variety of physiological features. Finally, the correlation coefficients between extracted features and BPs are evaluated before and after removing RM, which can provide a new guideline for feature selection. Additionally, the difference between cuffless BP estimation performance using physiological features with RM and without RM are demonstrated and compared.

## 3.2 Data, Pre-Processing and Feature Extraction

### 3.2.1 Experiment Data, Signal Processing and Feature Extraction

To validate the existence of RM on physiological features for BP estimation and study the effect of RM on cuffless BP estimation accuracy. The Bed-based [Ballistocardiogram](#) (BCG) Dataset [147] (from now on referred as **BCG BP**) which collects three-lead [ECG](#), finger [PPG](#), simultaneous reference [ABP](#) and respiration waveforms at the sampling frequency of 1 KHz which enhances the resolution of extracted features and can potentially improve BP estimation accuracy. This dataset includes 40 participants (17 males) whose ages range from 18 to 65 years and their body mass indices (BMIs) range from 18 to 48 kg/m<sup>2</sup>, and four participants indicated that they had some form of past or current cardiovascular-related conditions. Each subject was asked to recline in a customized bed with BCG sensors and an average of 6.97 minutes’ data were collected. To ensure fiducial points of physiological signals and morphological features are accurately detected and extracted for RM evaluation and cuffless BP estimation, 8 subjects whose physiological waveforms are in poor quality were excluded from this study.

A 4th order Chebyshev type II bandpass filter with low cut-off frequency of 0.5 Hz

and high cut-off frequency of 12 Hz was applied to remove low frequency BW and high frequency noises from PPG waveform. A lowpass filter with cut-off frequency of 12 Hz was applied to ABP waveform. The fiducial points of ECG, PPG and its first and second derivatives (VPG and APG respectively), as well as ABP waveforms were detected to extract physiological features for RM assessment and cuffless BP estimation. The R peaks of ECG was detected using the Pan-Tompkins algorithm [148]. The PPG fiducial points (onset, systolic peak, dicrotic notch and diastolic peak), fiducial points of the first derivative of PPG ( $u$ ,  $v$  and  $w$  points) and fiducial points of the second derivative of PPG ( $a$ ,  $b$ ,  $c$ ,  $d$ ,  $e$  and  $f$  points) were detected using an adjusted method based on *pyPPG* algorithm [145]. Then 69 physiological features and simultaneous SBP and DBP values are extracted from 32 subjects of **BCG BP** dataset. These physiological features can be categorized into 8 types: 3 PATs (from ECG and PPG), 17 interval-based features (from PPG and its derivatives), 4 amplitude-based features (from PPG and its derivatives), 3 area-based features (from PPG), 16 ratios between interval based features (from PPG and its derivatives), 17 ratios between amplitude based features (from PPG and its derivatives), 1 ratio between area based feature (from PPG) and 8 ratios between combined features (from PPG and its derivatives). A demonstration of fiducial points of ECG, PPG and PPG derivatives, as well as three PATs is shown in Figure 2.4. The detailed descriptions of 69 features as well as their relevance to BP changes can be found in Section 5.1.

These extracted beat-to-beat features and simultaneous BPs are then resampled to a fixed sampling rate of 2 Hz to form uniformly sampled time series, and outliers caused by inaccurate fiducial point detection especially for features extracted from PPG first and second derivatives were detected using the three-sigma rule and then replaced by piecewise cubic hermite interpolating polynomial ("pchip") interpolation method.

## 3.3 Respiration Modulation

### 3.3.1 Frequency Bands and Spectral Analysis

Frequency domain analysis is commonly applied for HRV analysis, as distinct spectral ranges provide insights into underlying physiological mechanisms [149]. Techniques such as Fast Fourier Transformation (FFT) and autoregressive (AR) modeling can be used to separate HRV into its component ultra-low-frequency (ULF), very-low-frequency (VLF), low-frequency (LF), high-frequency (HF) rhythms that operate within different frequency ranges.

### 3.3.1.1 Spectral bands from 0 to 0.04 Hz

The ULF band which ranges between 0 to 0.003 Hz, and VLF band which ranges from 0.003 to 0.04 Hz, representing slow variations over relatively long time scales. The ULF may be primarily driven by circadian rhythms and other contributors operating over long time periods include core body temperature, metabolism and renin-angiotensin system [150]. The exact meaning of VLF is not entirely clear [151], it could correspond to slower influencing rhythms such as hormonal rhythms, thermoregulation and renin-angiotensin system [150, 151].

### 3.3.1.2 Spectral Band from 0.04 to 0.15 Hz

The LF range is typically defined between 0.04 to 0.15 Hz. However, the origin of LF power is complex and unsettled. Some studies indicate that LF power may be produced by both the parasympathetic nervous system (PNS) and sympathetic nervous system (SNS), and BP regulation via baroreceptors [150, 151], though some suggest primarily by the PNS [150]. During resting conditions, the LF band is stated to reflect baroreflex activity and not cardiac sympathetic innervation [150]. However, sympathetic modulation of vasomotor tone, which occurs at approximately 0.1-0.15 Hz, is suggested to contribute to the LF band [18, 152], and Mayer waves which are slow spontaneous BP oscillations and are driven by sympathetic mediation are reflected in the HRV spectrum within this range [151]. Furthermore, during periods of slow RRs (below 8.5 bpm or 7-second periods), vagal activity can easily generate oscillations in heart rhythms that cross over into the LF band, meaning respiratory-related efferent vagally mediated influences can be particularly present in the LF band [150, 151].

### 3.3.1.3 Spectral Band from 0.15 to 0.4 Hz

The HF band in frequency-domain analysis is typically defined as the range from 0.15 to 0.4 Hz. The HF band is widely accepted as an indicator of parasympathetic activity [150, 151]. It is also called the respiratory band because it corresponds to the HR variations related to respiratory cycle. The HF band in frequency-domain analysis is typically defined as the range from 0.15 to 0.4 Hz. The HF band is widely accepted as an indicator of parasympathetic activity [150, 151, 153]. It is also called the respiratory band because it corresponds to the HR variations related to the respiratory cycle [150].

As demonstrated in Figure 3.6, some widely employed physiological features for cuffless BP estimation exhibit cyclical variations caused by respiration. Features, such as

pulse arrival time (PAT), time interval of PPG systolic phase (T\_sys) and systolic area of PPG pulses (AUC\_sys) illustrate dominant HF bands. Furthermore, throughout our experiments, we carefully analyzed different frequency bands of the extracted physiological features and their effects on cuffless BP estimation, the HF band which is primarily driven by respiratory activities is found to impact the BP estimation results the most. Therefore, the following sections of this chapter will focus on analyzing the HF band and proposing methods to suppress its effect for cuffless BP estimation.

### 3.3.2 Mechanisms and Previous Studies of Respiratory Modulation

Respiration has been shown to modulate cardiovascular dynamics, particularly influencing BP regulation and the spectral characteristics of HRV. In 1733, Stephen Hales first observed fluctuations in arterial pressure associated with respiration [16]. Subsequent studies, such as one conducted in 1952, have demonstrated that respiratory variations in BP are affected by factors including breathing patterns, RR, and body posture [17]. For example, breathing at a normal depth and rate can cause SBP fluctuations of up to 3 to 6 mmHg (e.g. shown in Figure 3.2(a)), and slow and deep breathing has been found to cause large fluctuations in SBP of up to 15 to 20 mmHg (e.g. shown in Figure 3.2(b)) [154]. Empirical results obtained from our experiments indicate that the fluctuations caused by respiration is generally more prominent in SBP and pulse pressure (PP), and less prominent in DBP and MBP, like shown in Figure 3.2 and Figure 3.6. This result is in line with a relevant previous study [18]. Respiration introduces fluctuations in BP through both mechanical and neural mechanisms. Mechanically, inspiration lowers intrathoracic pressure, which enhances the filling of the right heart and consequently increases right ventricular stroke volume [17]. This leads to a rise in effective pressure in the pulmonary veins, promoting increased filling of the left heart and a subsequent rise in left ventricular stroke volume. The mechanical effects are modulated by respiratory rate, with stroke volume modulation decreasing as respiratory rate increases. Neurally, respiratory activity directly influences sympathetic motoneurons; inspiratory activity inhibits sympathetic motoneurons, while expiratory activity facilitates it. This is further influenced by afferent pulmonary and thoracic stretch receptors, which inhibit sympathetic activity during inspiration. Spontaneous changes in BP or reductions in arterial baroreceptor afferent input can also modify respiratory modulation of muscle sympathetic outflow [149].

In the mid-19th century, Carl Ludwig described Respiratory Sinus Arrhythmia (RSA), initiating systematic investigations into the influence of respiration on HR [15]. RSA

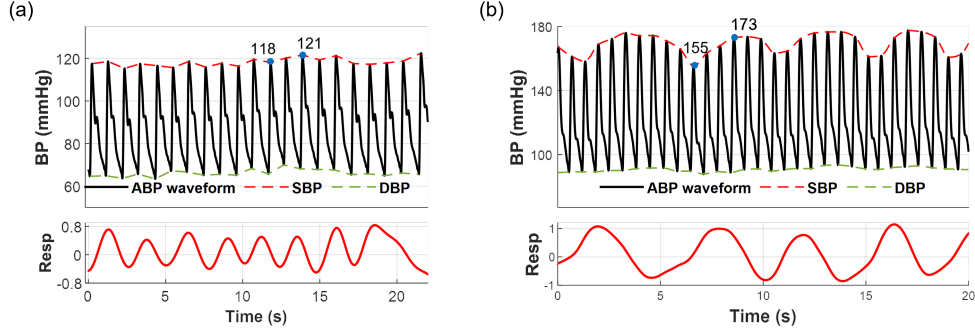


Figure 3.2: (a) respiration can cause SBP fluctuations up to 3 to 6 mmHg with normal breathing depth and rate, (b) respiration can cause up to 15 to 20 mmHg with shallow and deep breathing.

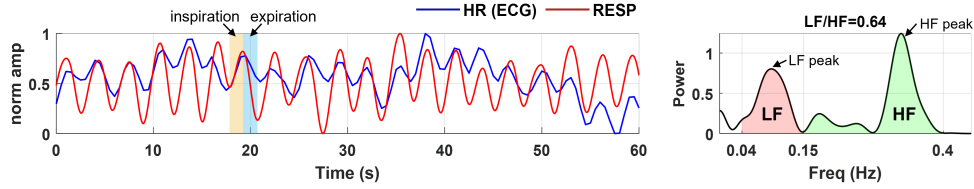


Figure 3.3: A demonstration of RSA, and power spectrum of HR which exhibits two prominent peaks of LF and HF bands.

illustrates that HR increases during inspiration and decreases during expiration, as shown in Figure 3.3. This modulation stems from the direct impact of the central respiratory drive on cardiac vagal preganglionic neurons, coupled with the inhibition of cardiac vagal efferent activity through lung inflation [155]. RSA enhances pulmonary gas exchange by synchronizing the heartbeat with the respiratory rhythm [155].

RM, such as RSA, manifests within the HF range, reflecting the changes in HR that synchronize with the respiratory cycle [146, 152]. For example, analyzing the HR shown in Figure 3.3 in frequency domain, it demonstrates both prominent LF and HF ranges and the peak of the HF range can exactly indicate the RR ( $f_{\text{resp}} \approx 0.32$  Hz). Several methods were proposed to separate RM from HR for more accurate HRV analysis. One approach involves using an adaptive noise canceller (ANC) with an FIR filter and the least mean squares (LMS) algorithm to reduce the influence of respiration on the HRV signal by comparing HRV with simultaneously recorded respiration signal [156]. Another method uses orthogonal subspace projections (OSP) to decompose the HRV signal into a respiratory component and a residual component, thereby separating the linear respiratory

influences from the HR to improve the estimation of RSA and sympathovagal balance [153]. Additionally, spectral filtering techniques like notch filters or raised cosine filters can be applied to the R-R intervals to remove the spectral component of a constant respiratory influence, aiming to isolate the underlying HRV parameters without respiratory artifacts [157].

Additionally, previous studies discovered respiration can also cause variations to physiological features which can be applied for cuffless BP estimation. Respiration has been proven to heavily modulate PAT which is the time interval between ECG R peak and PPG systolic peak, consequentially cause unsatisfied BP estimation results. Therefore a new feature—PPG intensity ratio (PIR, see Figure 3.6)—which demonstrates less HF components was applied to improve BP estimation accuracy [18]. A previous study [158] found that PAT exhibits more significant cyclical fluctuations than PTT, which indicates respiration modulates more heavily on pre-ejection period (PEP), as PAT is defined as the summation of both PTT and PEP. Other physiological features, such as systolic time interval (STI) and diastolic time interval (DTI) were also reported can be modulated by respiration [19].

However, the influence of RM on a variety of features for BP estimation, as well as its effect on cuffless BP estimation accuracy, has rarely been previously examined or addressed. This is probably due to the fact that most of the existing cuffless BP estimation studies employed MIMIC databases (e.g., MIMIC-II [159] and MIMIC-III [160] waveform databases, and UCI cuff-less blood pressure estimation dataset [161]) where the physiological waveforms, such as ECG, PPG and ABP are sampled at 125 Hz, resulting in a low temporal resolution of 0.008 s of time interval based features, such as PAT, pulse interval and systolic time of PPG pulses. However, this low sampling rate may not be sufficient to detect RM, as it was recommended that a minimum sampling frequency of 500 Hz for reliable RSA assessment [150], especially for patients with heart failure requires higher sampling rates for adequate temporal resolution [162].

### 3.3.3 The Complexity of Respiration Modulation

#### 3.3.3.1 RM Fluctuations Are Enhanced by Upright Posture

The BP fluctuations caused by respiration are found more exaggerated in the erect posture, comparing to subjects in the supine posture [17]. The power of HF band extracted from PAT is significantly higher when subjects are in the tilt position than subjects are in the supine position, with the average power of HF bands calculated from 14 subjects of  $4.59 \times 10^{-3} \text{ s}^2$  in the tile position, comparing to  $\sim 2.68 \times 10^{-3} \text{ s}^2$  in the supine posture [158].



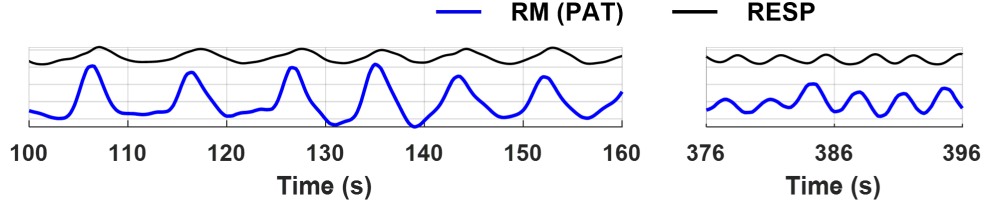


Figure 3.4: Demonstration of amplitude of RM fluctuations may vary with RR, the RM components extracted from PATs of both examples (left and right) are demonstrated in a same amplitude scale.

### 3.3.3.2 The Amplitude of RM Fluctuations Increases With Decreasing RR

As shown in Figure 3.4, the amplitude of RM component extracted from PAT is greater at a slower RR comparing to a faster RR.

### 3.3.3.3 The Frequency of RM Fluctuations Might Be a Function of RR

The BP may exhibit fluctuations at a submultiple of the respiration rate. An example shown in [17] the BP fluctuation of a subject occurs at one third of the respiration rate. The RM fluctuations occur approximately at the rate of 6 per minute when the subject actually breaths at the rate between 15 breaths to 20 breaths per minute (see Figure 6 of [17]). Even though this study mentioned this is a quite common phenomenon, it has not been observed from our experiments.

Study [163] explored the effects of OM chanting—an ancient Indian mantra practiced at a breathing rate of approximately 3 breaths per minute (0.05 Hz)—on the synchronization between ECG R-R intervals, respiration, and SBP. Interestingly, the researchers observed unexpected oscillations around 0.1 Hz in both SBP and RR intervals during breathing at 0.05 Hz. This phenomenon of frequency doubling may be attributed to the interaction between the breathing rhythm and intrinsic Mayer wave activity.

### 3.3.3.4 The Relation Between the RM Fluctuations and the Respiratory Phase is a Function of RR

As demonstrated in Figure 3.5 (a), the RM variation extracted from SBP decreases during inspiration and increases during expiration at slow respiration rate (approx. 7 breaths per minute). However, the RM variation extracted from SBP increases during inspiration and



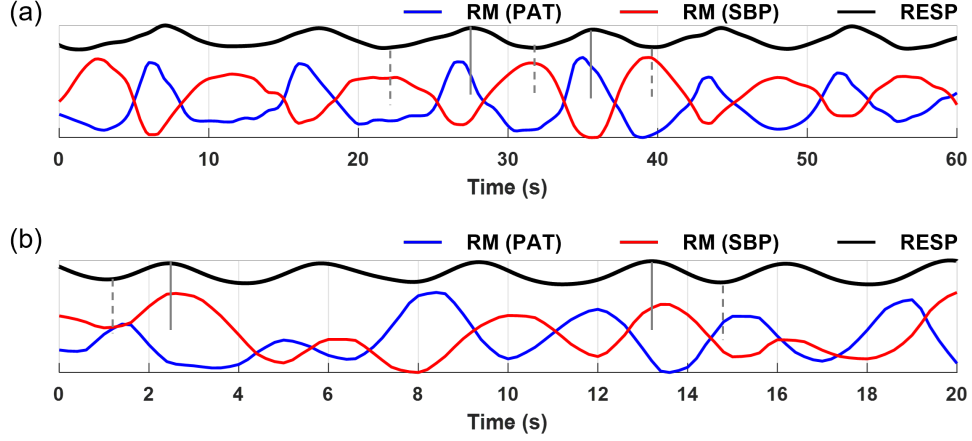


Figure 3.5: Demonstration of phase of RM fluctuations may vary with RR (y axis: amplitude).

decreases during expiration at a fast respiration rate (approx. 18 breaths per minute), as shown in Figure 3.5 (b). The phase difference between RM fluctuations of SBP and PAT remain the same regardless different RRs.

### 3.3.4 The General Existence of RM on a Wide Variety of Physiological Features and BPs

Our experiment results indicate that all 69 features extracted from 32 subjects from **BCG BP** dataset are affected by RM, although the degree of effect varies across different types of features. In general, respiration modulates interval-based features and interval-related features more heavily, like the example shown in Figure 3.6 which illustrates twelve features selected from different types and BPs, as well as their power spectrum. Moreover, the RM components derived from the features exhibit different phase characteristics. As an example shown in Figure 3.7, the extracted RM components of PPG pulse amplitude ( $A_{sp}$ ), PPG systolic time ( $T_{sys}$ ) and SBP almost synchronize in phase. However, the extracted RM components of  $PAT_{ms}$ , IPA and SBP are exactly out of phase. RM affects physiological features differently, making it significantly challenging for an estimation model predict BP accurately (examples of effects on 69 extracted features and BPs for 2 subjects are demonstrated in Figure A.3 to Figure A.14).

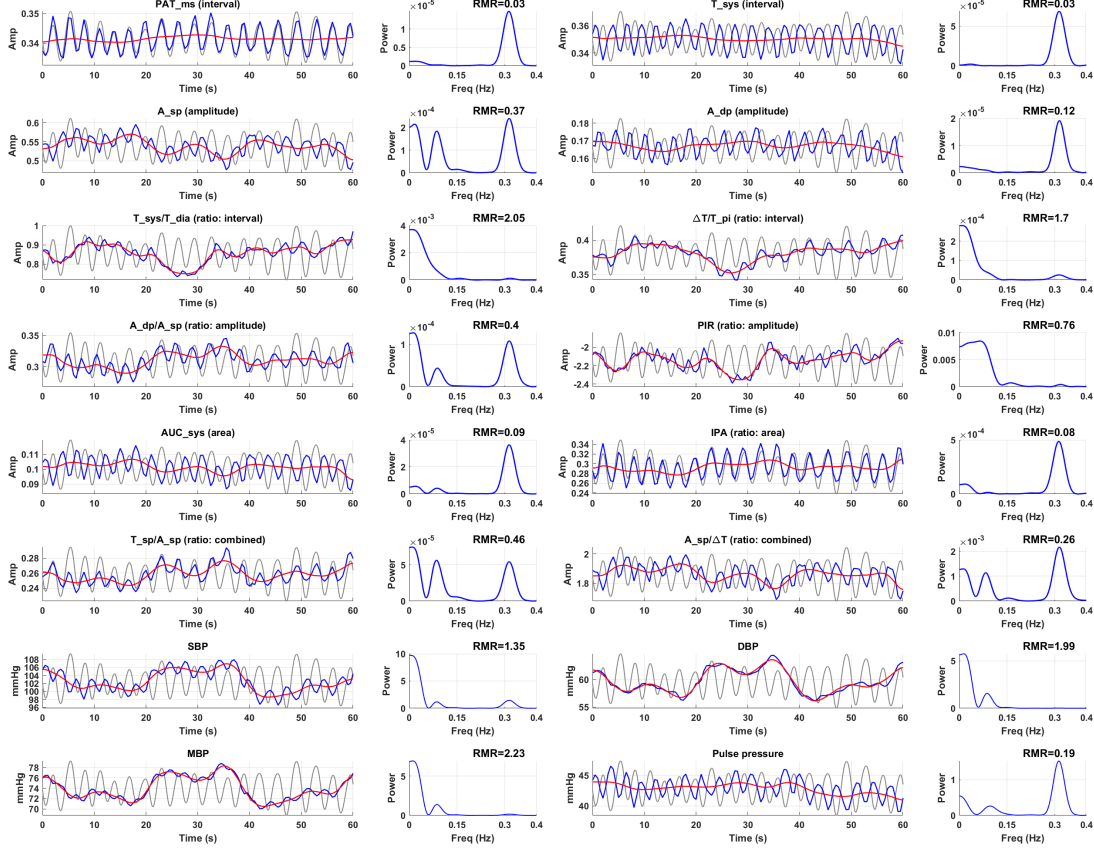


Figure 3.6: Twelve originally extracted features of different types, SBP, DBP, MBP and PP with the simultaneous reference respiration waveforms and their power spectrums.

### 3.3.5 Separating RM from Physiological Features and BPs

The [VMD](#) method is a robust and adaptive signal decomposition technique that decomposes a signal into an ensemble of band-limited intrinsic mode functions (IMFs). Unlike the conventional Empirical Mode Decomposition (EMD), which is based on empirical sifting processes and is prone to issues such as mode mixing and sensitivity to noise, VMD formulates the decomposition as a well-defined variational problem [164]. This formulation

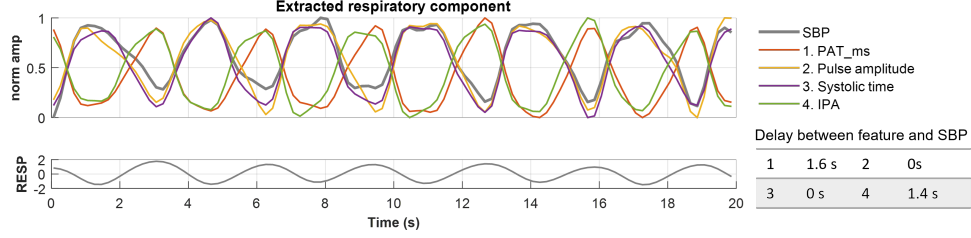


Figure 3.7: An example of phase difference RM components between four features—PAT\_ms, pulse amplitude, systolic time and IPA, and SBP, the time delay between features and SBP are calculated using cross-correlation based method.

allows for a simultaneous extraction of modes by minimizing the sum of their estimated bandwidths under the constraint that their collective sum reconstructs the original signal. The reason to use VMD based method is based on two considerations. First, the results obtained using VMD-based method doesn't exhibit edge distortion as conventional low-pass filter. Second, as VMD can decompose the originally extracted features into different frequency components which can be further used for applications such as RR estimation and respiration waveform monitoring.

VMD is particularly effective for amplitude and frequency modulated signals because it inherently constrains each mode to be compact around a central frequency, ensuring that each decomposed component is nearly monochromatic [165]. This property facilitates the accurate separation of distinct oscillatory components, making it highly suitable for analyzing signals with complex modulation patterns.

VMD decomposes the originally extracted feature or BP  $F_i(t)$  into  $K$  IMFs:

$$F_i(t) = \sum_{k=1}^K u_k(t) + r(t) \quad (3.1)$$

where  $u_k(t)$  is the  $k$ -th IMF and  $r(t)$  the residual. The number of modes  $K$  is determined by spectral analysis.

For each IMF  $u_k(t)$ , compute its peak frequency  $f_k$  via Fourier transform:

$$f_k = \arg \max_f |\mathcal{F}\{u_k(t)\}| \quad (3.2)$$

where  $\mathcal{F}$  denotes the Fourier transform.

Reconstruct the RM-suppressed features or BPs  $\tilde{F}_i(t)$  by aggregating IMFs with peak frequencies within 0 to 0.12 Hz which is the combination of ULF, VLF and LF bands:

$$\tilde{F}_i(t) = \sum_{k \in \mathcal{S}} u_k(t), \quad \mathcal{S} = \{k | f_k < 0.12 \text{ Hz}\} \quad (3.3)$$

In this study, the lower limit of HF band is adjusted to 0.12 Hz to accommodate slow respiration ( $< 7.2$  breaths/min).

As shown in Figure 3.8, VMD decomposes a 90 seconds' PAT\_ms of a subject from **BCG BP** dataset into 4 IMFs with different frequencies. The IMF1 and IMF2 corresponds to the RM caused by faster breathing rates occur at the first 10 seconds, and from 30 to 70 seconds. The IMF3 corresponds to the RM caused by slower breathing rates occur between 10 to 30 seconds, and the last 20 seconds. The IMF4 with its peak frequency of 0.02 Hz indicates the long-term trend of PAT\_ms. According to Equation 3.3, the IMF4 is used as the RM-free version of PAT\_ms.

### 3.3.6 Respiration Modulation Ratio (RMR)

A new metric—respiration modulation ratio (RMR)—is proposed to quantify the strength of the RM component of each physiological feature and BP, and the RMR is defined as:

$$\text{RMR} = \frac{P_{\text{signal}}}{P_{\text{RM}}} = \frac{\frac{1}{N} \sum_{n=1}^N (x[n] - \mu_x)^2}{\frac{1}{N} \sum_{n=1}^N (x_{\text{RM}}[n] - \mu_{\text{RM}})^2} \quad (3.4)$$

where  $P_{\text{signal}}$  is the power of the originally extracted physiological features or BPs, and  $P_{\text{RM}}$  is the power of the extracted RM component using the proposed method.  $x[n]$  is the original feature sequence which has  $n$  samples,  $\mu_x$  is the mean of the original feature,  $x_{\text{RM}}[n]$  is the extracted RM component and  $\mu_{\text{RM}}$  is the mean of the RM component. The interpretation of RMR is similar to signal-to-noise-ratio (SNR) which are described as follows:

- $\text{RMR} > 1$ : This indicates that the power of the original feature is greater than the power of RM component. In this case, the feature is less affected by respiration, suggesting that it predominantly reflects the underlying physiological signal rather than respiratory influence.
- $\text{RMR} < 1$ : This indicates that RM contributes more power than the original feature itself. A low RMR suggests that the feature is highly influenced by respiration,

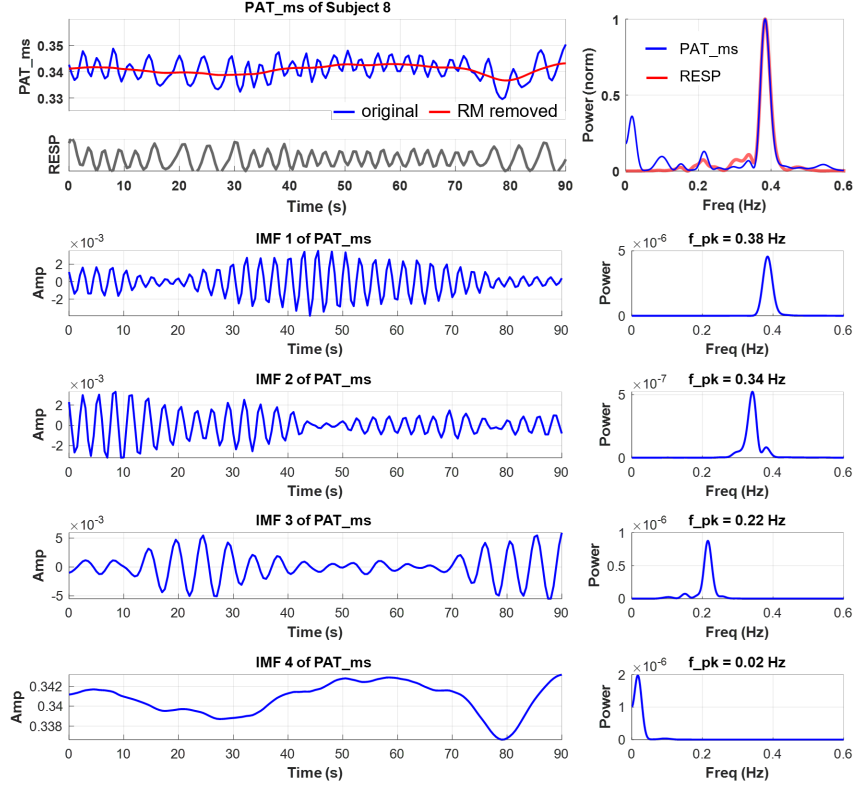


Figure 3.8: A demonstration of originally extracted PAT\_ms with RM and decompose the original PAT\_ms into 4 IMFs with different frequency components using VMD, and the RM-suppressed PAT\_ms is obtained by aggregating IMFs with peak frequency  $f_{pk} \leq 0.12$  Hz.

which may introduce unwanted variability and compromise the feature's reliability for downstream analysis.

- $RMR = 1$ : This implies that the power of the original feature and RM component are equal, indicating that respiration-induced variations and the intrinsic feature variations are of comparable magnitude. This scenario suggests a moderate level of respiration interference, where respiration-related fluctuations should be carefully considered in the analysis.

The RMR of twelve selected features, as well as the corresponding power spectrums are shown in Figure 3.6.

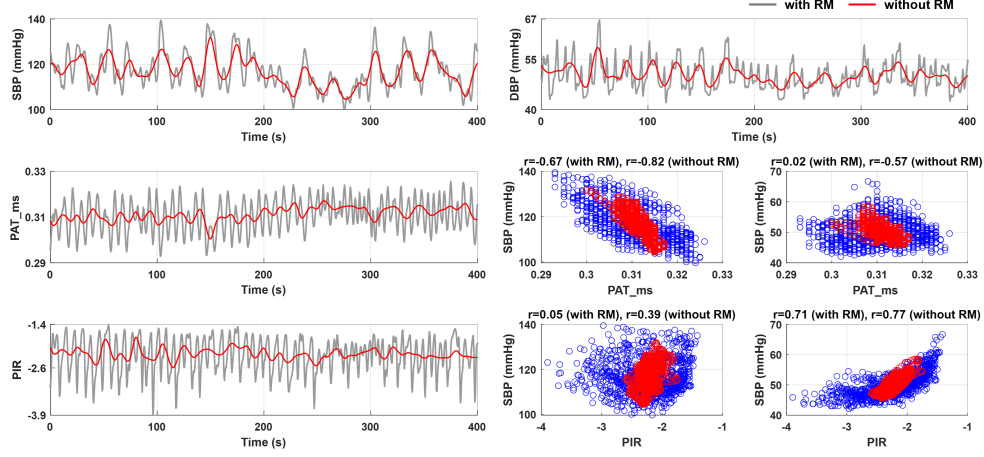


Figure 3.9: An example of a 400-seconds segment of originally extracted SBP, DBP, PAT\_ms and PIR and their RM-suppressed versions, along with the scatter plots and Pearson correlation coefficients between PAT\_ms, PIR and BPs.

### 3.4 Effect of Respiration on Cuffless BP Estimation

#### 3.4.1 Impact of RM on Feature Selection

RM may influence the result of feature selection which is an important process for cuffless BP estimation. After removing RM components from physiological features and BPs, the relationship or relevance between physiological features and BPs may be different and can potentially improve cuffless BP estimation accuracy. For example, the originally extracted SBP, DBP, PAT\_ms and PIR, as well as their RM-free versions are shown in Figure 3.9. It can be observed that both PAT\_ms and PIR are significantly modulated by respiration and their correlation coefficients with both SBP and DBP improve significantly. The correlation coefficients between 69 physiological features extracted from ECG, PPG and PPG derivatives and BPs with RM and without RM are shown in Appendix A.3, which can be used as a new guideline of feature selection for cuffless BP estimation.

#### 3.4.2 Impact of RM for Cuffless BP Estimation Accuracy

Two models are employed to assess the impact of RM for cuffless BP estimation accuracy.

- **BP estimation with PAT and PIR (Model 1):** the SBP and DBP can be

estimated using PAT and PIR based on Moens-Korteweg equation [18]:

$$\text{DBP} = \text{DBP}_0 \cdot \frac{\text{PIR}_0}{\text{PIR}} \quad (3.5)$$

$$\text{SBP} = \text{DBP}_0 \cdot \frac{\text{PIR}_0}{\text{PIR}} + \text{PP}_0 \cdot \left( \frac{\text{PAT}_0}{\text{PAT}} \right)^2 \quad (3.6)$$

where  $\text{DBP}_0$ ,  $\text{PAT}_0$  and  $\text{PIR}_0$  are the initial measurements of DBP, PAT and PIR respectively.  $\text{PP}_0$  is the initial measurement of [Pulse Pressure \(PP\)](#) which can be calculated by  $\text{PP}_0 = \text{SBP}_0 - \text{DBP}_0$  and  $\text{SBP}_0$  is the initial measurement of SBP.

The mean values of the first 5 samples of SBP, DBP, PIR and PAT are calculated as the initial measurements for each subject. Each subject's SBP and DBP are then estimated based on Equation 3.5 and Equation 3.6.

- **BP estimation with 69 physiological features and CNN-BiLSTM model (Model 2):** a CNN-BiLSTM model is developed which consists of three 1D convolutional layers (channels = 32,32,64) with kernel size of 3 and ReLU activation functions to gradually transform the 69 physiological features into a richer representation. Then two bidirectional long short-term memory (BiLSTM) layers (hidden size = 64, 128) further refine the feature representation obtained from CNN layers to capture temporal dependencies. Finally, a linear layer maps the output of the last BiLSTM to predict SBP or DBP respectively.

32 subjects are divided into 8 folds. For each fold, a leave-4-subjects-out cross-validation strategy is implemented, whereby the proposed CNN-BiLSTM model is initially trained on the features and reference BP measurements of 28 training subjects. Subsequently, for each of the four testing subjects, the first five samples of features and their corresponding BP values are utilized to fine-tune the model for personalized adaptation. Finally, the fine-tuned model is employed to predict BP for the remaining samples of each testing subject.

For each estimation method, both originally extracted features and BPs with RM, and RM-suppressed features and BPs are employed to evaluate the impact of RM on accuracy of cuffless BP estimation. [MAE](#), the standard deviation of absolute error (std), and Pearson correlation coefficient ( $r$ ) are employed to quantify the accuracy and similarity between the estimations and references. The results of BP estimations using two models with originally extracted features and RM-suppressed features are illustrated in Table 3.1. It can be observed that after removing RM from physiological features and BPs, the BP estimation

Table 3.1: Mean and standard deviation of absolute error, and Pearson correlation coefficient ( $r$ ) between estimated BPs and reference BPs using two estimation models with originally extracted features and RM-suppressed features.

|         |                  | MAE (std)   | $r$  |
|---------|------------------|-------------|------|
| Model 1 | SBP (with RM)    | 7.78 (7.15) | 0.38 |
|         | DBP (with RM)    | 7.29 (6.86) | 0.39 |
|         | SBP (without RM) | 5.57 (5.26) | 0.48 |
|         | DBP (without RM) | 5.16 (5.14) | 0.41 |
| Model 2 | SBP (with RM)    | 7.11 (6.77) | 0.30 |
|         | DBP (with RM)    | 4.25 (3.52) | 0.14 |
|         | SBP (without RM) | 6.73 (6.10) | 0.36 |
|         | DBP (without RM) | 3.66 (3.52) | 0.25 |

accuracy and trend similarity which can be implicitly quantified by Pearson correlation improved for both estimation models. For Model 1, the MAE between estimated SBP and reference SBP improves to 5.57 mmHg and Pearson correlation coefficient improves from 0.38 to 0.48 by using RM-free features. And for Model 2, the MAE between estimated SBP and reference SBP improves from 7.11 mmHg to 6.73 mmHg and Pearson correlation coefficient improves from 0.30 to 0.36 by using RM-free features.

Examples of the reference and estimated SBP of two subjects using Model 1 and Model 2 with both respiration modulated features and BPs, and RM-suppressed features and BPs are shown in Figure 3.10 and Figure 3.11. It can be observed that the estimations using RM-free features exhibit less noises and more accurate long-term trend tracking.

### 3.5 Conclusion and Discussion

In this section, we conducted a comprehensive analysis on the impact of respiration on cuffless BP estimation. Our experiments revealed that respiratory modulation introduces short-term fluctuations across a broad spectrum of physiological features used for BP estimation, with the magnitude and phase of these variations differing significantly among feature types. To mitigate these challenges, we propose a frequency analysis method based on VMD to separate and remove RM from the extracted physiological features and BPs before training the estimation models. Alongside this, we introduce a novel metric—the respiration modulation ratio (RMR), analogous to the signal-to-noise ratio (SNR)—to quantify the intensity of RM.

Our results indicate that after suppressing RM, the physiological features exhibit a stronger correlation with BPs, particularly in terms of interval-based features. This find-



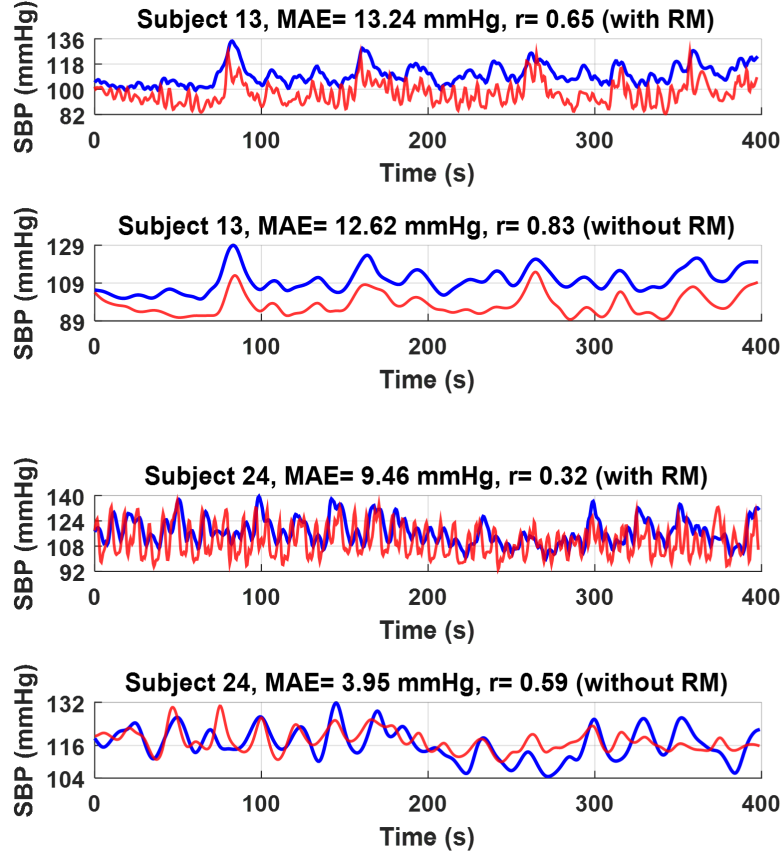


Figure 3.10: Examples of two subjects' estimated and reference SBP using Model 1 with both respiration modulated features and RM-free features.

ing provides a new guideline for signal processing and feature selection aimed at enhancing cuffless BP estimation accuracy. Comparative analyses using two different BP estimation models for 32 subjects demonstrate that RM-suppressed features yield improved BP prediction performance, as reflected by reduced MAE and enhanced Pearson correlation coefficients, thereby indicating a closer alignment of long-term BP trends with reference measurements. Consequently, we strongly recommend that future research pay more attention to the effect of respiration—an often overlooked factor—in advancing the accuracy of cuffless BP estimation.

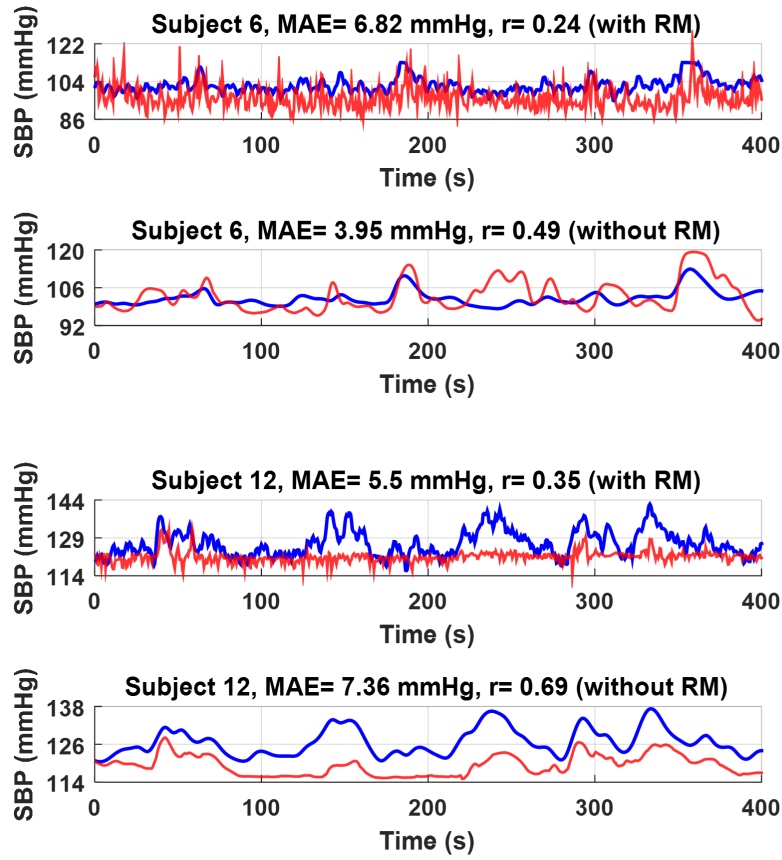


Figure 3.11: Examples of two subjects' estimated and reference SBP using Model 2 with both respiration modulated features and RM-free features.

## Chapter 4

# Novel Metrics for Tracking Blood Pressure Changes in Continuous Cuffless Blood Pressure Estimations

This chapter proposes a comprehensive evaluation framework for cuffless BP estimation which compares the similarities between the references and estimations in terms of both distance and trend. A composite metric is also proposed to combine both distance and trend similarity metrics. The proposed similarity metrics are normalized metrics which range from -1 to 1 and can be easily interpreted. The new proposed similarity metrics are generic and can be applied to other healthcare monitoring studies and even general time series similarity comparison problems.

**Publication:** a version of the method described in this chapter was previously published by Nature Scientific Reports [\[166\]](#).

### 4.1 Conventional Performance Metrics for Cuffless BP Evaluation

#### 4.1.1 Distance Metrics

The pairwise distance measures between the estimated and reference BPs are the most common metrics employed by the majority of current studies, as well as the aforementioned

validation standards. For SBP and DBP, the mean and standard deviation of estimation difference (MD and SDD), the mean and standard deviation of absolute difference (MAD and SDAD) and the root mean squared difference (RMSD) of the  $n$  individual paired measurements of the estimation and of the reference measurements for all subjects are calculated using equations shown in Table 4.1. Although these pairwise Euclidean distance based metrics are straightforward, intuitive and less computationally complex, they can not properly evaluate the accuracy of BP pattern tracking. For instance, the values of these aforementioned distance metrics between a set of simulated BP estimations and the reference (see Figure 4.5 and Section 4.3.1 for details) are illustrated in Table 4.6. It can be observed that simulated BP estimations  $BP_{est1-4}$  satisfy the passing requirement of IEEE Standard 1708a-2019 ( $MAD \leq 7$  mmHg), but  $BP_{est2-4}$  were unable to accurately track BP patterns. Though MAD between  $BP_{est5}$  and  $BP_{ref}$  is 8.14 mmHg, but  $BP_{est5}$  almost has the identical morphology to  $BP_{ref}$  which performs better than  $BP_{est2-4}$  in BP pattern tracking. It can also be observed that MAD and RMSD can better indicate the accuracy of the estimation model in BP change tracking than MD, for example, comparing  $BP_{est2}$  and  $BP_{est3}$  against  $BP_{ref}$ . Furthermore, the contributions of individual error values to the final results are different for MAD and RMSD, since the errors are squared before being averaged, the RMSD assigns a higher weight to large errors than MAD. For example, the RMSD between  $BP_{est4}$  and  $BP_{ref}$  shown in Figure 4.5 is 6.24 mmHg whereas the MAD is 4.29 mmHg.

There are also some distance metrics developed to measure curve similarity, such as **Dynamic time warping** (DTW) [167], **Fréchet** distance [168] and **Hausdorff** distance [169]. These metrics address the key limitations of pairwise distance based metrics in terms of time scaling and shifting and time series in different lengths, but they are more computationally complex than pairwise distance metrics. As the length of estimation and reference are identical and the time delay (may be caused by time series estimation models, such as RNN and LSTM) should be composited before comparison. These metrics are not necessary for cuffless BP evaluation.

Therefore, we suggest to use MAD or RMSD as the primary distance metric for future studies.

### 4.1.2 Correlation Coefficient

A popular metric to evaluate the agreement between estimations and reference measurements is the correlation coefficient. **Pearson correlation coefficient** (PCC, denoted as  $\rho$ ), which measures the linear correlation between two variables, is commonly employed

Table 4.1: Definition of distance metrics and correlation coefficients widely-used for cuffless BP estimation evaluation.

| Distance metrics             |                                                                                  |                                                                                            | Correlation coefficients |                                                                             |
|------------------------------|----------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|--------------------------|-----------------------------------------------------------------------------|
| Estimation difference        | Mean                                                                             | Standard deviation                                                                         | Pearson correlation      | $\rho = \frac{\sigma_{12}}{\sigma_1\sigma_2}$                               |
|                              | $MD = \frac{1}{n} \sum_{i=1}^n (BP_{\text{est}i} - BP_{\text{ref}i})$            | $SDD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (BP_{\text{est}i} - BP_{\text{ref}i} - MD)^2}$     |                          |                                                                             |
| Absolute difference          | Mean                                                                             | Standard deviation                                                                         | Concordance correlation  | $\rho_c = \frac{2\sigma_{12}}{(\mu_1 - \mu_2)^2 + \sigma_1^2 + \sigma_2^2}$ |
|                              | $MAD = \frac{1}{n} \sum_{i=1}^n  BP_{\text{est}i} - BP_{\text{ref}i} $           | $SDAD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n ( BP_{\text{est}i} - BP_{\text{ref}i}  - MAD)^2}$ |                          |                                                                             |
| Root mean squared difference | $RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n (BP_{\text{est}i} - BP_{\text{ref}i})^2}$ |                                                                                            |                          |                                                                             |

$BP_{\text{est}}$ : BP estimated by the test device/method;  $BP_{\text{ref}}$ : BP measured by the reference device/method.  $i$ : index of the individual element;  $n$ : the number of measurements.  $\sigma_{12}$ : covariance of  $BP_{\text{est}}$  and  $BP_{\text{ref}}$ ;  $\sigma_1$  and  $\sigma_2$ : standard deviation of  $BP_{\text{est}}$  and  $BP_{\text{ref}}$ ;  $\mu_1$  and  $\mu_2$ : mean of  $BP_{\text{est}}$  and  $BP_{\text{ref}}$ .

by the majority of existing studies [76, 79]. The mathematical definition of PPC is shown in Table 4.1. However,  $\rho$  has one major shortcoming when assessing the reproducibility of measurements: it is invariant to additive or/and multiplicative shifts by a constant value, which is also an issue for the Spearman correlation coefficient. Therefore, using  $\rho$  to compare the similarity of the estimated BPs and the reference BPs may be misleading.

The **concordance correlation coefficient** (CCC, denoted as  $\rho_c$ ) was originally introduced by Lawrence Lin to address the aforementioned limitations of PCC in 1989 [170], and the formulation of CCC is shown in Table 4.1. The relationship between  $\rho$  and  $\rho_c$  can be further derived as  $\rho_c = \rho \cdot C_b$  (see Appendix for details) which means CCC ( $\rho_c$ ) contains the measurements of both  $C_b$  (a measure of accuracy) and PCC (a measure of precision). A demonstration of limitations of PCC and the robustness of CCC in detecting scale and location shifts comparing to PCC is shown in Appendix A.4.2.

The previous works [171, 172] translate  $\rho$  into descriptors like "weak," "moderate," or "strong", and McBride [173] suggested that  $\rho_c > 0.99$  for "almost perfect,"  $0.95 < \rho_c \leq 0.99$  for "substantial,"  $0.90 < \rho_c \leq 0.95$  for "moderate" and  $\rho_c \leq 0.90$  for "poor." However, there is still no clear-cut agreement how to interpret both PCC and CCC for cuffless BP evaluation.

Though the results of PCC and CCC shown in Table 4.6 can indicate the morphology similarities of a simulated dataset. They are not suitable for evaluating BP pattern tracking of noisy and long-term data collected from real subjects, as described in Section 4.2.3 and Section 4.3.2.

### 4.1.3 A Composite Score to Represent Similarity in Mean, Variance and Correlation Coefficient

A composite score, generic composite similarity measure (GCSM) [174], which was originally proposed to compare similarities of spatial or temporal patterns from images for ecological studies can be used to address the limitation that a single distance metric can not measure the pattern similarity for two compared sequences (as described in Section 4.1.1). GCSM composes the similarity measures of mean, standard deviation and correlation coefficient of two compared sequences ( $x$  and  $y$ ) to a single score, and its formulation is shown in Table 4.2.

Table 4.2: Equation for calculating GCSM and definitions of each component.

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |  |  |  |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|--|--|
| <b>GCSM</b> ( $x, y$ ) = $s_1(x, y) \cdot s_2(x, y) \cdot s_3(x, y)$                                                                                                                                                                                                                                                                                                                                                                                             |  |  |  |
| <b>Components:</b>                                                                                                                                                                                                                                                                                                                                                                                                                                               |  |  |  |
| $s_1(x, y) = \begin{cases} 1 - \frac{ \mu_x - \mu_y }{\max_{x,y} - \min_{x,y}} & s_3 \geq 0 \\ 1 - \frac{ \min_{x,y} + \max_{x,y} - \mu_x - \mu_y }{\max_{x,y} - \min_{x,y}} & s_3 < 0 \end{cases} \quad s_2(x, y) = 1 - \frac{ \sigma_x - \sigma_y }{(\max_{x,y} - \min_{x,y})/2} \quad s_3(x, y) = \begin{cases} 1 & \sigma_x = \sigma_y = 0 \\ 0 & \sigma_x \text{ or } \sigma_y = 0 \\ \frac{\sigma_{xy}}{\sigma_x \sigma_y} & \text{otherwise} \end{cases}$ |  |  |  |
| $\mu_x$ and $\mu_y$ : mean values of the compared sequences $x$ and $y$ ; $\max_{x,y}$ and $\min_{x,y}$ are the global maximum and minimum of sequences $x$ and $y$ ; $\sigma_x$ and $\sigma_y$ : standard deviation of two sequences $x$ and $y$ ; $\sigma_{xy}$ is the covariance of sequences $x$ and $y$ .                                                                                                                                                   |  |  |  |

It can be inferred from the formulations that  $s_1(x, y) \in [0, 1]$ ,  $s_2(x, y) \in [0, 1]$  and  $s_3(x, y) \in [-1, 1]$  and therefore  $\text{GCSM}(x, y) \in [-1, 1]$ . The mean, standard deviation and PCC are more similar for two compared sequences when  $\text{GCSM}(x, y)$  runs to 1. Although the GCSM consists of three different metrics to measure the overall similarity between sequences, it still has drawbacks.  $s_1(x, y)$  and  $s_2(x, y)$  are sensitive to outliers if  $|\mu_x - \mu_y| \ll \max_{x,y} - \min_{x,y}$  or  $|\sigma_x - \sigma_y| \ll (\max_{x,y} - \min_{x,y})/2$ , the  $s_1(x, y)$  or  $s_2(x, y)$  can still be close to 1 which indicates two sequences have similar means and standard deviations, though the actual differences (the numerators of  $s_1(x, y)$  and  $s_2(x, y)$ ) can be significantly large and cannot be ignored.

## 4.2 A New Comprehensive Performance Evaluation for Cuffless BP Estimation

In this section, a new comprehensive performance evaluation method which quantifies the similarity of **distance**, **trend** and the composite of them, between the reference and estimated BP measurements will be introduced. This new method addresses the problems

or limitations of the existing evaluation metrics and methods employed for cuffless BP estimation (as described in Section 4.1). The workflow of this new method is demonstrated in Figure 4.1, and the pseudocode of the proposed method is shown in Algorithm 4.1. The pre-processing is done with piecewise linear representation (PLR) which divides the compared estimated and reference BP sequences into segments according to trend changes and approximates segment linear trends. For each segment, new proposed distance and trend similarity metrics will be evaluated, and a composite similarity metric which combines both distance and trend similarities using weighted sum is then calculated. Finally, the temporal normalization method will combine the segment similarity results for entire sequence evaluation and result interpretation.

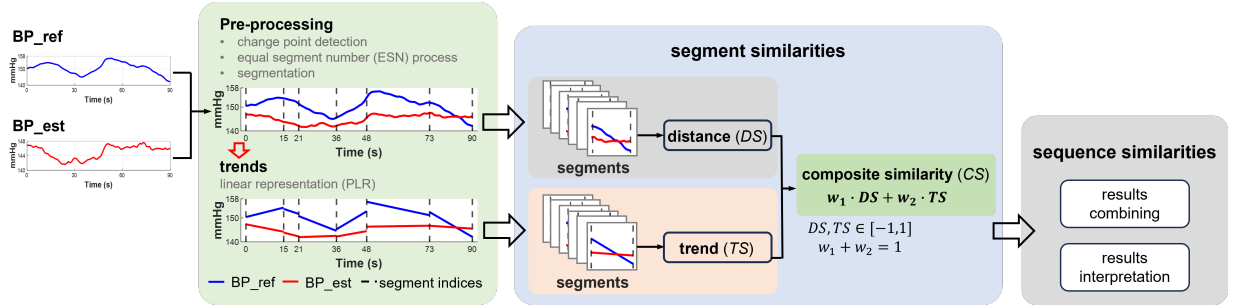


Figure 4.1: The framework of the proposed comprehensive performance evaluation method for cuffless BP estimation, including sequence pre-processing, segment similarity evaluation and sequence similarity evaluation.

### 4.2.1 Pre-processing

The similarity of two raw time series is difficult to measure as time series data are typically noisy and of high dimensions. As a result, it is common and efficient to work with data representations that are lighter than the raw data by reducing the dimensions of the data while maintaining their main properties, a number of techniques can be applied for this purpose (see study [175] for methods introduction and comparison). In this study, piecewise linear representation (PLR) [176] which can effectively extract trend information from the original time series data by dividing the original data into a number of segments according to trend changes and approximating the linear trends is employed. The number of segments that one time series should be divided into can be decided using several approaches, such as sliding window [177], top-down [178] and bottom-up [179] algorithms (see [180] for details).

---

**Algorithm 4.1** Segment distance, trend and composite similarity metrics and temporally normalized metrics

---

**Input:** two sequences for comparison  $x$  and  $y$ ,  $\text{length}(x) = \text{length}(y)$   
**Output:**  $\text{TNCSM}(x, y) \in [-1, 1]$

```

1: changepoint_x = segmentation(x)                                ▷ segment x and extract change points, according to tendency
2: changepoint_y = segmentation(y)                                ▷ segment y and extract change points, according to tendency
3: changepoint_xy = sort([0, unique(changepoint_x ∪ changepoint_y), length(x) - 1])    ▷ ESN process
4: changepoint_process(changepoint_xy)    ▷ process changepoints such that each segment has enough samples for PLR
5: T_D = length(x)                                                ▷ duration of the sequence
6: TN_DS = 0, TN_TS = 0, TN_CS = 0
7: for i ← 1 to length(changepoint_xy) - 1 do
8:   t_i = changepoint_xy(i) + 1                                ▷ for the first segment, t_i = changepoint_xy(i)
9:   t_{i+1} = changepoint_xy(i + 1)
10:  x_i = x(t_i : t_{i+1})
11:  y_i = y(t_i : t_{i+1})
12:  x_{L_i} = least_squares(x_i)
13:  y_{L_i} = least_squares(y_i)                                ▷ approximate a straight line for each segment using least squares method
14:  DS(x_i, y_i) = distance_similarity(x_i, y_i)                ▷ evaluate segment distance similarity
15:  TS(x_i, y_i) = trend_similarity(x_{L_i}, y_{L_i})            ▷ evaluate segment trend similarity
16:  CS(x_i, y_i) = w_1 · DS(x_i, y_i) + w_2 · TS(x_i, y_i)    ▷ calculate composite similarity using weighted sum
17:  t_{d_i} = t_{i+1} - t_i                                    ▷ duration of i-th segment
18:  TN_DS +=  $\frac{1}{T_D} \cdot \text{DS}(x_i, y_i) \cdot t_{d_i}$             ▷ compute sequence distance similarity with temporal normalization
19:  TN_TS +=  $\frac{1}{T_D} \cdot \text{TS}(x_i, y_i) \cdot t_{d_i}$             ▷ compute sequence trend similarity with temporal normalization
20:  TN_CS +=  $\frac{1}{T_D} \cdot \text{CS}(x_i, y_i) \cdot t_{d_i}$             ▷ compute sequence composite similarity with temporal normalization
21: end for

```

---

However, these methods are parametric which means parameters such as the number of segments to be divided into or the approximation error threshold must be pre-defined, and these parameters will impact the segmentation results and therefore affect similarity comparison between sequences (e.g., the results demonstrated in Table 2 of study [181]). In this study, a non-parametric Bayesian method, Bayesian estimator of abrupt change, seasonality, and trend (BEAST) [182], is employed to pre-process the original compared sequences. A demonstration of detecting change points and segmenting the reference and estimated BPs using BEAST is shown in Figure 4.4(a). However, two time series may have different number of segments (e.g., the reference BP is divided into 8 segments and estimated BP is divided into 9 segments as shown in Figure 4.4(a)), then a equal segment number (ESN) process which re-divides both time series into the same number of segments while also ensures a segment has enough samples ( $m \geq 2$ ) for linear trend approximation is developed and applied to the compared sequences.

After segmentation and ESN process, the least squares method is applied to approximate the original segment data to a straight line which represents the segment trend (pre-processing steps are described in Algorithm 4.1 line1 to line4). A demonstration of ESN process and linear trend approximation is shown in Figure 4.4(b).

To better describe the proposed similarity metrics in the following sections, a set of



denotations are formally introduced. The original reference sequence (reference BPs) is denoted as  $x$  and the original estimated sequence (estimated BPs) is denoted as  $y$  respectively. The  $i$ -th segment of the original sequences  $x$  and  $y$  are denoted as  $x_i$  and  $y_i$  respectively. The approximated linear representations of the original sequences  $x$  and  $y$  using aforementioned PLR methods are denoted as  $x_L$  and  $y_L$  respectively. Finally, the  $i$ -th segment of the linear trends of the compared sequences are denoted as  $x_{L_i}$  and  $y_{L_i}$  respectively.

## 4.2.2 Segment Similarity Measures

After pre-processing two compared sequences using PLR, the segment similarity measures of **distance**, **trend** and the composite of them will be evaluated.

### 4.2.2.1 Segment Distance Similarity

The new proposed segment distance similarity evaluates the MAD between the estimated and reference BPs, and maps the MAD to a value which is between -1 to 1. The distance similarity metric is a logistic function customized according to IEEE Standard 1708a-2019 [25] evaluation protocol as shown in Figure 4.2. For each segment, the MAD between two compared sequences is calculated and then the segment distance similarity  $DS(x_i, y_i)$  is defined as:

$$DS(x_i, y_i) = -\frac{2}{1 + e^{-2.2 \cdot (MAD_i - 5.5)}} + 1 \quad (4.1)$$

where  $MAD_i$  is the MAD of the  $i$ -th segment of compared sequences after change point detection and segmentation,  $MAD_i = \frac{\sum_{j=1}^m |x_{ij} - y_{ij}|}{m}$ ,  $x_{ij}$  and  $y_{ij}$  are the  $j$ -th measurement values of the  $i$ -th segment of compared sequences  $x$  and  $y$  respectively, and  $m$  is number of samples of the  $i$ -th segment ( $m \geq 2$ ). It can be inferred from Equation 4.1 and observed from Figure 4.2 that  $DS(x_i, y_i) \in [-1, 1]$ .  $DS(x_i, y_i)$  is close to 1 when  $MAD_i \leq 4$  mmHg, corresponding to "Grade A" of IEEE Standard and "pass" grade of ANSI/AAMI SP10 [23] requirement.  $DS(x_i, y_i) \in [0.5, 0.93)$  when  $4\text{mmHg} < MAD_i \leq 5\text{mmHg}$ , corresponding to "Grade B" of IEEE Standard and "mostly pass" grade of ANSI/AAMI SP10 standard respectively.  $DS(x_i, y_i) \in [-0.5, 0.5)$  when  $5\text{mmHg} < MAD_i \leq 6\text{mmHg}$  ( $DS(x_i, y_i) = 0$  when  $MAD_i = 5.5$  mmHg), corresponding to "Grade C" of IEEE Standard and "pass or fail" grade of ANSI/AAMI SP10 standard respectively.  $DS(x_i, y_i) \in [-0.93, -0.5)$  when  $6\text{mmHg} < MAD_i \leq 7\text{mmHg}$ , corresponding to "Grade D" of IEEE Standard and "mostly fail" grade of ANSI/AAMI SP10 standard respectively.  $DS(x_i, y_i)$  is close to -1 when

$MAD_i \geq 7\text{mmHg}$ , corresponding to "Grade D (unacceptable)" of IEEE Standard and "fail" grade of ANSI/AAMI SP10 standard respectively.

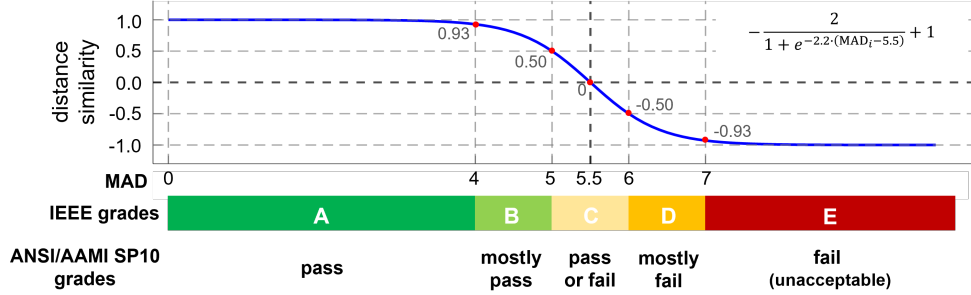


Figure 4.2: A demonstration of the proposed segment distance similarity which is a customized logistic function according to IEEE Standard 1708a-2019 evaluation criteria [25].

#### 4.2.2.2 Segment Trend Similarity

The new proposed segment trend similarity metric evaluates the accuracy of the estimation method in tracking BP trends. Firstly, the approximated linear representation of the estimated BP  $y_L$  is processed such that it has the same mean to the approximated linear representation of the reference BP  $x_L$  ( $y'_L = y_L - (\bar{y}_L - \bar{x}_L)$ ,  $\bar{x}_L$  and  $\bar{y}_L$  are the mean values of the approximated linear representations for reference BP  $x_L$  and estimated BP  $y_L$  respectively). For each segment, the slopes of the approximated linear trends are calculated to determine if the compared segments are in the same direction or not. If two compared segment trends are in the same direction, as showing in Figure 4.3(a) and (b), both of the segments are up or down, or at least one segment is unchanged ( $\text{slope}(x_{L_i}) \times \text{slope}(y'_{L_i}) \geq 0$ ). Therefore, the range of angular difference  $\theta$  between the segment linear trends  $x_{L_i}$  and  $y'_{L_i}$  is  $[0, \theta_{\max}]$  and the segment trend similarity is defined as  $TS(x_i, y_i) = 1 - \frac{\theta}{\theta_{\max}}$ . If two compared segment trends are in different directions, as showing in Figure 4.3 (c) and (d), one segment is up and the other segment is down ( $\text{slope}(x_{L_i}) \times \text{slope}(y'_{L_i}) < 0$ ). Therefore, the range of angular difference  $\theta$  between segment linear trends  $x_{L_i}$  and  $y'_{L_i}$  is  $(0, \theta_{\max}]$  and the segment trend similarity is defined as  $TS(x_i, y_i) = -\frac{\theta}{\theta_{\max}}$ . The formulations of  $\theta$  and  $\theta_{\max}$  for different scenarios are illustrated in Table 4.3 and Figure 4.3.

As shown in Figure 4.3, the angular difference  $\theta$  is determined by the starting values  $v_{xL}$  and  $v_{yL}$ , and the ending values  $v_{xR}$  and  $v_{yR}$  of linear trends of the compared segments, as well as the segment duration  $t_i$  ( $t_i = t_{iR} - t_{iL}$ ). The maximum angular difference

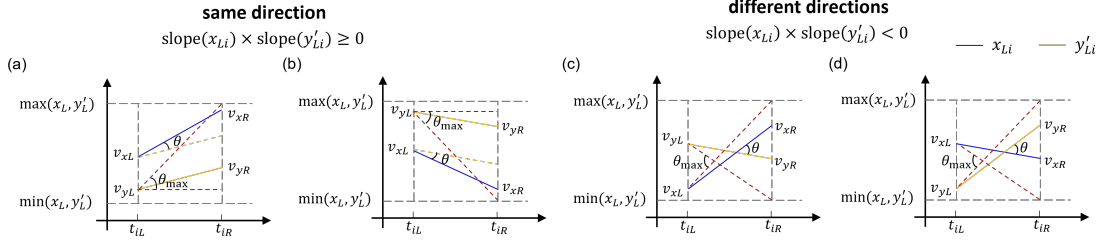


Figure 4.3: A demonstration of the proposed angular difference based trend similarity measures of approximated linear segments in the same direction and in opposite directions, with the corresponding calculations of  $\theta$  and  $\theta_{\max}$ .

Table 4.3: Segment trend similarity for different scenarios

| $x_{Li}$ and $y_{Li}$ in the same direction<br>$\text{slope}(x_{Li}) \times \text{slope}(y_{Li}') \geq 0$ |                                                                                                                       | $x_{Li}$ and $y_{Li}'$ in different directions<br>$\text{slope}(x_{Li}) \times \text{slope}(y_{Li}') < 0$ |                                                                                                                                         |
|-----------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| $\text{slope}(x_{Li}) \geq 0, \text{slope}(y_{Li}') \geq 0$ (Figure 4.3(a))                               | $\text{slope}(x_{Li}) \leq 0, \text{slope}(y_{Li}') \leq 0$ (Figure 4.3(b))                                           | $\text{slope}(x_{Li}) > 0, \text{slope}(y_{Li}') < 0$ (Figure 4.3(c))                                     | $\text{slope}(x_{Li}) < 0, \text{slope}(y_{Li}') > 0$ (Figure 4.3(d))                                                                   |
| $1 - \frac{\theta}{\theta_{\max}}$                                                                        |                                                                                                                       | $-\frac{\theta}{\theta_{\max}}$                                                                           |                                                                                                                                         |
| $TS(x_i, y_i)$                                                                                            |                                                                                                                       |                                                                                                           |                                                                                                                                         |
| where                                                                                                     |                                                                                                                       |                                                                                                           |                                                                                                                                         |
| $\theta$                                                                                                  | $ \arctan\left(\frac{v_{yR}-v_{yL}}{t_{iR}-t_{iL}}\right) - \arctan\left(\frac{v_{xR}-v_{xL}}{t_{iR}-t_{iL}}\right) $ |                                                                                                           | $\arctan\left(\frac{v_{yR}-v_{yL}}{t_{iR}-t_{iL}}\right) + \arctan\left(\frac{v_{xR}-v_{xL}}{t_{iR}-t_{iL}}\right)$                     |
| $\theta_{\max}$                                                                                           | $\arctan\left(\frac{\max(x_L, y_L) - \min(v_{xL}, v_{xL}')}{t_{iR}-t_{iL}}\right)$                                    | $\arctan\left(\frac{\max(v_{xL}, v_{xL}') - \min(x_L, y_L)}{t_{iR}-t_{iL}}\right)$                        | $\arctan\left(\frac{\max(x_L, y_L) - v_{xL}}{t_{iR}-t_{iL}}\right) + \arctan\left(\frac{v_{yL} - \min(x_L, y_L)}{t_{iR}-t_{iL}}\right)$ |
|                                                                                                           |                                                                                                                       |                                                                                                           | $\arctan\left(\frac{\max(x_L, y_L) - v_{yL}}{t_{iR}-t_{iL}}\right) + \arctan\left(\frac{v_{xL} - \min(x_L, y_L)}{t_{iR}-t_{iL}}\right)$ |

$\theta_{\max}$  is determined by the maximum and minimum values of the approximated PLRs of the compared sequences  $x_L$  and  $y'_L$ , the starting values of the compared linear trends  $v_{xL}$  and  $v_{yL}$  as well as the segment duration  $t_i$ . It can be inferred that the segment trend similarity  $TS(x_i, y_i) \in [-1, 1]$  ( $TS(x_i, y_i) \in [0, 1]$  for segment trends in a same direction,  $TS(x_i, y_i) \in [-1, 0]$  for segment trends in different directions). The segment trends are more similar when  $TS(x_i, y_i)$  is close to 1 and they are more reverse when  $TS(x_i, y_i)$  is close to -1.  $TS(x_i, y_i) > 0$  indicates the compared segment trends are in the same direction while  $TS(x_i, y_i) < 0$  indicates the segment trends are in different directions.

#### 4.2.2.3 Segment Composite Similarity

The segment composite similarity  $CS(x_i, y_i)$  combines both segment **distance** and **trend** similarities using the weighted sum method which assigns different weights according to the importance of each metric, as shown in Equation 4.2.

$$CS(x_i, y_i) = w_1 \cdot DS(x_i, y_i) + w_2 \cdot TS(x_i, y_i) \quad (4.2)$$

The existing evaluation standards and studies primarily evaluate distance metrics (e.g.,

MAD, RMSD) between the estimations and references. In this study, we emphasize the importance of trend tracking in cuffless BP monitoring. Therefore, we assign the weight for distance similarity  $w_1 = 0.6$  and the weight for trend similarity  $w_2 = 0.4$ . And these weights can be adjusted according to different evaluation priorities.

### 4.2.3 Combining the Segment Similarity Measures (global evaluation)

Inspired by the method used in [181], a temporal-normalization method which combines the segment similarity measures based on the proportion of the segment duration was applied for this study. The temporally normalized distance, trend and composite similarity of the compared reference and estimation of BPs are therefore defined and calculated as shown in Table 4.4.

Table 4.4: Equations of calculating temporally-normalized similarity metrics.

|                                                                                 |                                                                    |
|---------------------------------------------------------------------------------|--------------------------------------------------------------------|
| <b>Temporally-normalized similarity of distance (TN_DS)</b>                     | $TN\_DS(x, y) = \frac{1}{t_D} \sum_{i=1}^K DS(x_i, y_i) \cdot t_i$ |
| <b>Temporally-normalized similarity of trend (TN_TS)</b>                        | $TN\_TS(x, y) = \frac{1}{t_D} \sum_{i=1}^K TS(x_i, y_i) \cdot t_i$ |
| <b>Temporally-normalized composite similarity of distance and trend (TN_CS)</b> | $TN\_CS(x, y) = \frac{1}{t_D} \sum_{i=1}^K CS(x_i, y_i) \cdot t_i$ |

As shown in Table 4.4,  $DS(x_i, y_i)$ ,  $TS(x_i, y_i)$  and  $CS(x_i, y_i)$  are the segment distance, trend and composite similarity respectively.  $t_i$  is the duration of the  $i$ -th segment,  $t_D$  is the duration of the compared sequences,  $K$  is the number of segments. Therefore,  $\sum_{i=1}^K \frac{t_i}{t_D} = 1$  which means the temporal normalization method is similar to the weighted sum method and the weight of each segment result contributes to the final result is determined by the segment duration, indicating the longer the segment is, the more ratios it contributes to the final result. Also, it can be inferred that  $TN\_DS(x, y) \in [-1, 1]$ ,  $TN\_TS(x, y) \in [-1, 1]$  and  $TN\_CS(x, y) \in [-1, 1]$ . The segment distance, trend and composite similarity metrics, as well as their temporally normalized results for whole sequence comparison of the example shown in Figure 4.4 are illustrated in Table 4.5.

The results illustrated in Table 4.5 indicate PCC and CCC can not properly evaluate trend similarity between the reference and estimated BPs. For example, the trends of the 5<sup>th</sup> segment are almost parallel and the MAD between the estimations and references is only 0.44 mmHg which are accurately evaluated by new proposed distance and trend similarity metrics, whereas the PCC and CCC are 0.601 and 0.096. It can also be observed

Table 4.5: Segment similarity measures of the example shown in Figure 4.4 proposed by existing works and this study and their temporally-normalized results for whole sequences comparison.

|                             | segment (local) |            |             |              |              |              |             |             |              |              | whole sequence (global) |
|-----------------------------|-----------------|------------|-------------|--------------|--------------|--------------|-------------|-------------|--------------|--------------|-------------------------|
|                             | 1               | 2          | 3           | 4            | 5            | 6            | 7           | 8           | 9            | 10           |                         |
| <b>duration</b> [start,end] | 4s [0, 4)       | 9s [4, 13) | 4s [13, 17) | 10s [17, 27) | 15s [27, 42) | 11s [42, 53) | 4s [53, 57) | 6s [57, 63) | 17s [63, 80) | 10s [80, 90] | 77s [0, 77]             |
| <b>MAD</b> (mmHg)           | 12.87           | 5.28       | 9.00        | 4.13         | 0.44         | 1.51         | 3.96        | 1.39        | 1.99         | 0.75         | 2.96                    |
| $PCC(x_i, y_i)$             | -0.998          | -0.998     | -0.012      | -0.082       | 0.601        | -0.532       | 0.788       | -0.465      | 0.949        | -0.126       | -0.016                  |
| $CCC(x_i, y_i)$             | -0.007          | -0.625     | 0           | -0.003       | 0.096        | -0.113       | 0.019       | -0.049      | 0.432        | -0.030       | -0.015                  |
| $DS(x_i, y_i)$              | -1.000          | 0.240      | -0.999      | 0.907        | 1.000        | 0.999        | 0.935       | 0.999       | 0.999        | 1            | 0.73                    |
| $TS(x_i, y_i)$              | -0.14           | -0.68      | 0.95        | -0.36        | 0.98         | -0.18        | 0.93        | -0.19       | 0.94         | -0.14        | 0.26                    |
| $CS(x_i, y_i)$              | -0.66           | -0.13      | -0.22       | 0.40         | 0.99         | 0.53         | 0.93        | 0.53        | 0.98         | 0.54         | 0.54                    |

in Figure 4.4 that the trends of the 2<sup>nd</sup> segment are more distinct than the 1<sup>st</sup> segment with the proposed trend similarity of -0.68 and -0.14 respectively. However, the PCC values for these two segments are identical.

#### 4.2.4 Interpretation of the Proposed Metrics

The proposed segment distance, trend similarity metrics and their composite metric, as well as their temporally-normalized versions for whole sequence evaluation, are normalized metrics (defined in [183]) and they have a same range of  $[-1, 1]$ .

The estimation is more reliable when the distance similarity  $DS(x, y) > 0$  (also applicable to segment distance similarity  $DS(x_i, y_i)$ ) and it is more accurate when  $DS(x, y)$  is close to 1. The estimation is more unreliable when  $DS(x, y) < 0$  and it is more inaccurate when  $DS(x, y)$  is close to -1. The estimated and the reference segment BP trends are in the same direction when the segment trend similarity  $TS(x_i, y_i) > 0$  and they are more similar when  $TS(x_i, y_i)$  is close to 1. The estimated and the reference segment BP trends are in opposite directions when the segment trend similarity  $TS(x_i, y_i) < 0$  and they are more reverse when  $TS(x_i, y_i)$  is close to -1. The estimated BP pattern are more similar to the reference BP pattern when  $TS(x, y)$  is close to 1 and the majority of the estimated and the reference BP trends are in the same direction when  $TS(x, y) > 0$  and vice versa.

Since the composite similarity metric  $CS(x, y)$  combines distance  $DS(x, y)$  and trend  $TS(x, y)$  similarity metrics with weighted sum, the estimation is more accurate in terms of both distance and trend when  $CS(x, y)$  is close to 1, the estimation is more different to the reference in terms of distance and/or trend when  $CS(x, y)$  is close to -1. Therefore, a BP estimation model is expected to have the composite similarity metric  $CS(x, y)$  that is as close to 1 as possible and we suggest that a BP estimation method/device can be

considered capable for tracking BP changes if the proposed metrics  $TN\_DS \geq 0$ ,  $TN\_TS \geq 0$  and  $TN\_CS \geq 0$ . Detailed interpretation of the proposed metrics with a set of simulations can be found in Appendix A.4.6.

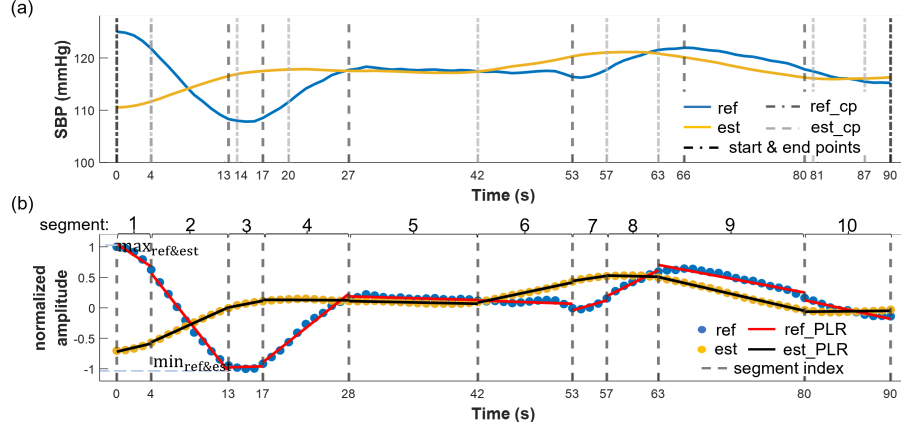


Figure 4.4: A demonstration of (a) segmenting the compared reference and estimated BPs using BEAST method, the reference BP is divided into 8 segments while the estimated BP is divided into 9 segments ( $t = 4$  is a change point for both reference and estimation); (b) the equal segment number (ESN) will re-process the change points and re-divide the compared sequences such that the compared sequences have the same number of segments (e.g., the reference and estimated BPs are re-processed to 10 segments), and the least squares are used to approximate the linear trend for each segment.

## 4.3 Experiments and Results

To demonstrate the performance of the proposed distance, trend and composite similarity metrics, these metrics are firstly applied to compare a set of different simulated BP estimations to the reference BP for feasibility validation. Then, these new metrics are applied to compare the estimated BPs using two SOTA models to the reference BPs of a real world dataset.

### 4.3.1 Applying the Proposed New Metrics to Simulated Data

In this section, a synthetic dataset spanning a 3-minute interval was generated, simulating BP estimations using reference experimental data characterized by substantial variations

within that time frame. Consequently, the dataset includes waveforms with varying degrees of fidelity such that various conventional metrics and the proposed metrics can be assessed in tracking BP changes. The waveforms of the reference BP and simulated BP estimations and their relations are illustrated in Figure 4.5.  $BP_{est1}$  and  $BP_{est5}$  have almost identical pattern to  $BP_{ref}$  but with different distances.  $BP_{est2}$  simulates a random estimation of  $BP_{ref}$ .  $BP_{est3}$  simulates an estimation model that is unable to track any BP change and just estimates the average BP level, and  $BP_{est4}$  simulates a reverse estimation of  $BP_{ref}$ .

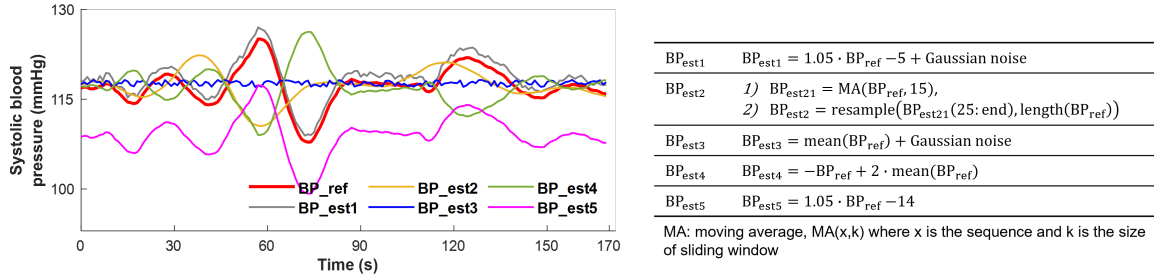


Figure 4.5: Waveforms of the reference SBP measurement ( $BP_{ref}$ ) and five simulated SBP estimations ( $BP_{est1}$  to  $BP_{est5}$ ), and relations between the simulated SBPs and reference SBP.

The conventional distance metrics (MD, MAD, RMSD), correlation coefficients (PCC and CCC), a composite similarity measure GCSM [174], and the proposed similarity metrics between different simulated BP estimations and the reference BP are shown in Table 4.6.

Table 4.6: Conventional distance metrics, correlation coefficients, composite similarity measure GCSM and the proposed similarity metrics calculated between simulated BP estimations and reference BP.

|                         | conventional metrics |                 |       |       |        | composite measure | temporally normalized metrics (ours) |       |       |
|-------------------------|----------------------|-----------------|-------|-------|--------|-------------------|--------------------------------------|-------|-------|
|                         | MD $\pm$ SDD*        | MAD $\pm$ SDAD* | RMSD* | PCC   | CCC    |                   | TN_DS                                | TN_TS | TN_CS |
| $BP_{est1} \& BP_{ref}$ | $1.36 \pm 0.34$      | $1.36 \pm 0.34$ | 1.40  | 0.996 | 0.914  | 0.91              | 0.999                                | 0.98  | 0.99  |
| $BP_{est2} \& BP_{ref}$ | $0.08 \pm 4.45$      | $2.85 \pm 3.41$ | 4.43  | -0.23 | -0.22  | -0.19             | 0.56                                 | 0.17  | 0.40  |
| $BP_{est3} \& BP_{ref}$ | $0.51 \pm 3.16$      | $2.27 \pm 2.26$ | 3.20  | -0.02 | -0.004 | -0.012            | 0.81                                 | 0.34  | 0.62  |
| $BP_{est4} \& BP_{ref}$ | $-0.28 \pm 6.28$     | $4.30 \pm 4.57$ | 6.27  | -1    | -0.996 | -1                | 0.34                                 | -0.46 | 0.02  |
| $BP_{est5} \& BP_{ref}$ | $-8.14 \pm 0.16$     | $8.14 \pm 0.16$ | 8.14  | 1     | 0.24   | 0.68              | -0.994                               | 0.98  | -0.20 |

\*unit: mmHg

It can be observed from Table 4.6 that conventional distance metrics can only measure the dissimilarity of pairwise values for the compared sequences. Though PCC can indicate the pattern similarities between the simulated BP estimations and the references.

For example,  $BP_{est5}$  and  $BP_{est1}$  have almost identical morphology to  $BP_{ref}$  and therefore  $PCC=1$ ,  $PCC=-1$  for  $BP_{est4}$  indicates the estimated BP has a reverse morphology to  $BP_{ref}$  and  $PCC \approx 0$  for  $BP_{est3}$  indicates the estimation model does not track any BP changes. The GCSM, which composes similarities of mean, standard deviation and PCC, demonstrates its potential in assessing both distance and trend similarities for most of the scenarios, except for  $BP_{est5}$ . As previously stated, the sensitivity of  $s_1(x, y)$  and  $s_2(x, y)$  of GCSM (see Section 4.1.3) is significantly affected by outliers or instances where the disparity between the global maximum and global minimum values ( $\max_{x,y} - \min_{x,y}$  in Table 4.2) is substantial. Consequently,  $s_1(x, y)$  and  $s_2(x, y)$  when comparing  $BP_{est5}$  and  $BP_{ref}$  will still approximate to 1 even though the MAD is 8.14 mmHg, which is a considerable large error for cuffless BP estimation. CCC and GCSM will exhibit more limitations in cases when the compared time series are long and noisy which can be observed from results obtained from DailA BP in Section 4.3.2. The new similarity metrics are proposed to address the drawbacks of the aforementioned metrics, and the results from this simulated dataset validated the proposed methods can accurately quantify both distance and trend similarity for cuffless BP studies.

### 4.3.2 Applying the Proposed New Metrics to Results of Two SOTA Models

The wearable physiological and BP measurements during activities of daily living dataset [184] (from now on referred as DailA BP) will be used to train and evaluate two SOTA cuffless BP estimation models with proposed metrics. DailA BP includes approximately 6.5 hours' physiological waveforms (ECG, PPG and BP) from 5 healthy subjects (4 males, age  $(28 \pm 6.6)$  yrs) performing daily activities and several activities such as Valsalva Maneuver, walking and hand gripping were performed to induce BP changes (see Section 2.4.2 and Figure 2.10).

Two SOTA models, which differ in BP estimation errors and trend tracking, were selected to assess the feasibility of the proposed metrics. ApproximateNet [117] which is a deeply supervised one-dimensional U-Net model, was proposed to approximate BP waveform from a simultaneous PPG waveform. It was originally trained and evaluated with 127,260 subsets of simultaneous PPG and BP waveforms from UCI BP [185]. The reported  $MAD \pm SDAD$  for SBP and DBP estimation is  $(6.42 \pm 9.52)$  mmHg and  $(3.58 \pm 5.80)$  mmHg respectively [117]. DeepRNN-4L [84], which consists of a bidirectional LSTM layer and four forward LSTM layers with residual connections, was proposed to estimate SBP and DBP using 7 features extracted from ECG and PPG in a time series analysis fashion.



DeepRNN-4L was originally trained and evaluated on the static continuous BP dataset which includes simultaneous ECG, PPG and BP data from 84 healthy subjects (51 males and 33 females) while stationary, and the reported RMSD of SBP and DBP is 3.75 and 2.43 mmHg respectively [84]. Different training strategies were applied for ApproximateNet and DeepRNN-4L models (details of implementation, model parameters and training processes are described in Appendix A.4.3), also due to the fact that DeepRNN-4L model considers temporal dependency. Therefore, these two models generate different estimations on the same data in terms of both distance and pattern tracking accuracy. A proportion (approximate 30 minutes) of reference SBPs and estimated SBPs using two SOTA models are shown in Figure 4.6.

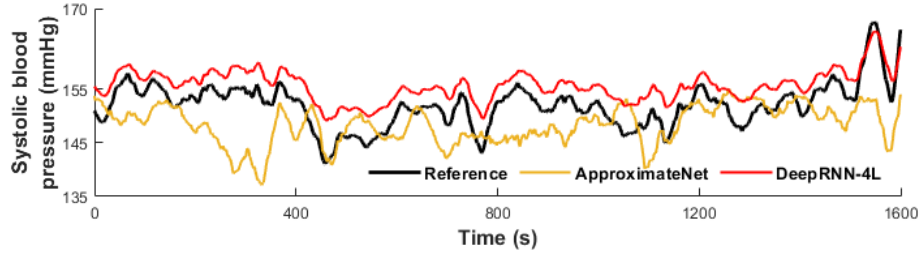


Figure 4.6: An approximate 30 minutes’ example of reference SBPs and estimated SBPs using SOTA models (BP variations might be caused by daily activities, such as standing up and walking), the distance, trend and composite similarities between estimated SBPs using ApproximateNet and reference SBPs are 0.37, 0.23 and 0.32 respectively, and 0.81, 0.81 and 0.81 between estimated BPs using DeepRNN-4L and reference BPs respectively (the composite similarities with different combinations of  $w_1$  and  $w_2$  are demonstrated in Appendix A.4.7).

The conventional distance metrics (MD, MAD, RMSD), correlation coefficients (PCC and CCC), a composite similarity measure GCSM, and the proposed similarity metrics between BPs estimated by two SOTA models and reference BPs from DailA BP dataset are shown in Table 4.7.

As shown in Table 4.7, DeepRNN-4L outperforms ApproximateNet in estimating BP using DailA BP dataset by comparing distance metrics and correlation coefficients. However, these conventional metrics can not indicate the performance of these estimation models in BP change tracking. The GCSM results almost align with PCC, this is because the denominators of the  $s_1(x, y)$  and  $s_2(x, y)$  can be much larger than their numerators ( $\max_{x,y} - \min_{x,y} \gg |\mu_x - \mu_y|$  or  $(\max_{x,y} - \min_{x,y})/2 \gg |\sigma_x - \sigma_y|$ ) and therefore  $s_1(x, y)$  or  $s_2(x, y) \approx 1$  for long and noisy sequences. Eventually, GCSM is mostly deter-

Table 4.7: Conventional distance metrics, correlation coefficients, composite similarity measure GCSM and the proposed similarity metrics calculated between the estimated BPs and reference BPs using SOTA models from DailA BP dataset.

|                       |            | conventional metrics |                 |       |      |      | composite measure | temporally normalized metrics (ours) |       |       |
|-----------------------|------------|----------------------|-----------------|-------|------|------|-------------------|--------------------------------------|-------|-------|
|                       |            | MD $\pm$ SDD*        | MAD $\pm$ SDAD* | RMSD* | PCC  | CCC  | GCSM              | TN_DS                                | TN_TS | TN_CS |
| <b>ApproximateNet</b> | <b>SBP</b> | $-5.82 \pm 11.77$    | $9.88 \pm 8.65$ | 13.13 | 0.68 | 0.62 | 0.64              | -0.22                                | 0.36  | 0.015 |
|                       | <b>DBP</b> | $-1.70 \pm 8.04$     | $6.50 \pm 5.02$ | 8.22  | 0.65 | 0.62 | 0.61              | -0.001                               | 0.34  | 0.13  |
| <b>DeepRNN-4L</b>     | <b>SBP</b> | $0.21 \pm 5.28$      | $4.13 \pm 3.29$ | 5.29  | 0.94 | 0.93 | 0.92              | 0.87                                 | 0.70  | 0.80  |
|                       | <b>DBP</b> | $0.77 \pm 5.25$      | $4.15 \pm 3.30$ | 5.30  | 0.75 | 0.73 | 0.72              | 0.74                                 | 0.55  | 0.66  |

\*unit: mmHg

mined by  $s_3(x, y)$  which is PCC of the compared sequences. The aforementioned issues are tackled by the proposed methods through the utilization of PLR, which simplifies noisy and high-dimensional time series with a number of line segments. This approach leads to improved accuracy in quantifying trend similarity. The utilization of temporal normalization also effectively tackles the problem associated with the evaluation strategy by providing varying weights to the final outcome based on the duration of each segment. The distance similarity TN\_DS of ApproximateNet is -0.22 and -0.001 for SBP and DBP, respectively, indicating the estimations of this method more likely dissatisfy IEEE Standard. DeepRNN-4L outperforms ApproximateNet in terms of both distance and trend tracking, which is also indicated by the composite similarity TN\_CS.

## 4.4 Conclusion and Discussion

Continuous cuffless BP tracking has become a popular research problem recently and has shown promising results. However, current evaluation standards, which primarily use pairwise distance metrics, are not robust and comprehensive. In this chapter, we first reviewed several commonly used BP evaluation standards and then introduced BP patterns and the importance of accurate tracking of BP changes. The conventional distance metrics, correlation coefficients which are commonly used by existing studies, a composite similarity metric GCSM were reviewed and their drawbacks in terms of BP pattern tracking were described and demonstrated with examples. Finally, we proposed a comprehensive framework with new distance, trend and composite similarity metrics with temporal normalization, for continuous cuffless BP evaluation.

The results from the simulated dataset and the re-implementation of two SOTA models using a realistic dataset DailA BP verified the feasibility of the proposed similarity metrics

for cuffless BP evaluation. Our results demonstrated that the composite similarity metric has advantages in evaluating both distance and trend similarities for long and noisy time series, compared to conventional metrics and GCSM. Additionally, the proposed metrics are normalized metrics and range from -1 to 1, making them intuitively interpretable, similar to well-known correlation coefficients. The proposed methods and metrics can also be used to benchmark different algorithms for cuffless BP estimation and diagnose and improve the estimation accuracy by analyzing proposed metrics (see Appendix [A.4.5](#)), facilitating the adoption of this new technique in real-world applications. Future work will adapt the proposed similarity metrics as a customizable loss function to improve the performance of estimation models in tracking BP patterns.

# Chapter 5

## Methods to Improve Generalization of BP Estimation Models

Generalization in machine learning refers to the ability of a trained model to accurately make predictions on new, unseen data. The purpose of generalization is to force the model to understand the patterns and relationships within its training data and apply them to previously unseen examples from the same distribution or even a different distribution as the training set [186]. Generalization is crucial to the cuffless BP estimation models because it allows them to make accurate and reliable predictions in diverse real-world scenarios.

This chapter focuses on improving model generalization by shifting attention from isolated BP value prediction to accurate BP trend tracking. Most existing studies prioritize point-wise BP estimation, but this overlooks the importance of capturing how BP changes over time. To address this, a new framework is explored: first, several overlooked morphological features that show a strong correlation with BP trends are identified. Then, two-time series regression models are proposed to better capture temporal dependencies in the data. Finally, a shape-oriented loss function is introduced to guide the model in learning features that relate to BP trend patterns, thereby enhancing generalization for real-world, continuous BP monitoring.

### 5.1 Trend-Related Features for BP Estimation

As introduced and reviewed in Section 2.3, current cuffless BP estimation studies can be categorized by the model’s input-output type. **feat2bp** approaches estimate BPs from

pre-determined features from physiological signals, such as most commonly employed ECG and/or PPG waveforms. **sig2bp** methods directly predict BP from physiological waveforms and **sig2sig** methods first reconstruct ABP waveform from input physiological signals, from which BP values are then derived. Even though, there has been a surge of interest in both **sig2bp** and **sig2sig** methods recently, mainly facilitated by the adoption of computer vision (CV) architectures (e.g., CNNs) for time-series feature extraction and the growing application of large language models (LLMs) to numerical data prediction using a transformer and various attention mechanisms. We still observe some pre-determined morphological features that are highly related to BP changes and can outperform **sig2bp** and **sig2sig** in terms of model generalization and trend tracking performance.

In this section, 69 morphological features are extracted from ECG, PPG and derivatives of PPG signals based on fiducial points identified across 32 subjects. The relationships between these features and BPs are then systematically evaluated. These fiducial points include R peak of ECG; onset, systolic peak, dicrotic notch and diastolic peak of PPG;  $u$ ,  $v$  and  $w$  points of the first derivative of PPG (VPG); and  $a$ ,  $b$ ,  $c$ ,  $d$ ,  $e$  and  $f$  points of the second derivative of PPG (APG). The demonstration and definition of fiducial points of PPG and its derivatives can be found in Figure 5.1 and Table 5.1 (see Section 5.4.3 for details of fiducial points detection methods). These physiological features can be categorized into eight types: 3 PATs (from ECG and PPG), 17 interval-based features (from PPG and its derivatives), four amplitude-based features (from PPG and its derivatives), three area-based features (from PPG), 16 ratios between interval-based features (from PPG and its derivatives), 17 ratios between amplitude-based features (from PPG and its derivatives), one ratio between area-based features (from PPG) and eight ratios between combined features (from PPG and derivatives). The definitions of these morphological features and their correlation coefficients to simultaneous SBP and DBP before and after removing RM (respiration modulation, see Chapter 3 for details) are demonstrated in Table 5.2.

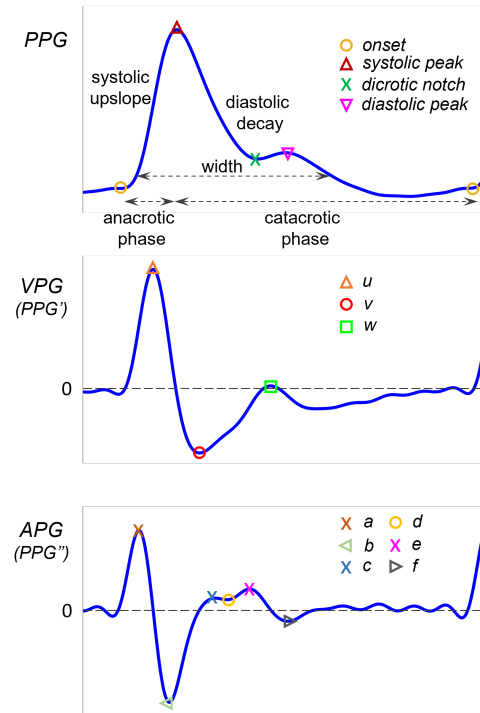


Figure 5.1: Demonstration of fiducial points of a typical PPG pulse wave, its first derivative PPG' and second derivative PPG'' which are also termed as VPG and APG respectively. Their fiducial points which are widely used to extract features. The waveform can be separated into anacrotic and catacrotic phases, which are dominated by systolic ejection and wave reflections from periphery respectively.

Table 5.1: Definition of PPG fiducial points, and steps to detect these fiducial points.

| Step | Fiducial pt.    | Fiducial point definitions                                                                                                                                                                                                                                                                           |
|------|-----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1    | <i>sp</i> (PPG) | <i>systolic peak</i> . The point with the highest amplitude of each PPG pulse.                                                                                                                                                                                                                       |
| 2    | <i>u</i> (VPG)  | <i>max slope point</i> . The point with highest amplitude of VPG pulses 0.1 pulse interval before systolic peaks (pulse interval calculated from systolic peaks).                                                                                                                                    |
| 3    | <i>a</i> (APG)  | The point with highest amplitude of APG pulse 0.1 pulse interval before point <i>u</i>                                                                                                                                                                                                               |
| 4    | <i>b</i> (APG)  | The minimum point APG pulse between point <i>a</i> and 0.2 pulse interval after point <i>a</i>                                                                                                                                                                                                       |
| 5    | <i>on</i> (PPG) | <i>Pulse onset</i> . The last trough between 0.15 pulse interval before point <i>u</i> and point <i>u</i> , if no troughs detected within this range, then select the point with minimum amplitude of APG pulse within this range                                                                    |
| 6    | <i>e</i> (APG)  | The peak point with the maximum amplitude between point <i>b</i> and 0.7 pulse amplitude after pulse onset                                                                                                                                                                                           |
| 7    | <i>c</i> (APG)  | The local maximum with the highest amplitude between the point <i>b</i> and <i>e</i> , or if no local maximum is present, then the inflection point between <i>b</i> and <i>e</i> on APG                                                                                                             |
| 8    | <i>d</i> (APG)  | The local minimum with the lowest amplitude between the point <i>c</i> and <i>e</i> , or if no local minimum is present, then use the same location of point <i>c</i>                                                                                                                                |
| 9    | <i>f</i> (APG)  | The trough point with minimum amplitude between point <i>e</i> to 0.7 pulse interval after pulse onset of APG pulse                                                                                                                                                                                  |
| 10   | <i>dp</i> (PPG) | <i>Diastolic peak</i> The first local maximum of the PPG pulse after systolic peak <i>sp</i> and 0.7 pulse interval after pulse onset <i>on</i> ; if there is no maxima, then use the first local maximum of the PPG pulse between point <i>e</i> and 0.8 pulse interval after pulse onset <i>on</i> |
| 11   | <i>dn</i> (PPG) | <i>Dicrotic notch</i> If a diastolic peak is present, then it is the local minimum between systolic peak <i>sp</i> and diastolic peak <i>dp</i> . If there is no diastolic peak, then it is the inflection point between the systolic peak and point <i>f</i>                                        |
| 12   | <i>v</i> (VPG)  | The local minimum of VPG pulse between systolic peak <i>sp</i> and diastolic peak <i>dp</i>                                                                                                                                                                                                          |
| 13   | <i>w</i> (VPG)  | The first peak of VPG pulse between dicrotic notch <i>dn</i> and 0.6 pulse interval after systolic peak <i>sp</i>                                                                                                                                                                                    |

Table 5.2: Definition of 69 features extracted from ECG, PPG and its derivatives with correlation coefficients ( $r$ ) to SBP and DBP with RM and without RM (results represented in median (iqr) among 32 subjects).

| Notation                         | Definition        | $r$ (with RM)                                                                               |              | $r$ (without RM) |              |              |
|----------------------------------|-------------------|---------------------------------------------------------------------------------------------|--------------|------------------|--------------|--------------|
|                                  |                   | SBP                                                                                         | DBP          | SBP              | DBP          |              |
| Interval based features          |                   |                                                                                             |              |                  |              |              |
| 1                                | $PAT_{on}$        | time interval between ECG R peak and PPG onset                                              | -0.32 (0.41) | -0.07 (0.33)     | -0.36 (0.71) | -0.20 (0.51) |
| 2                                | $PAT_{ms}$        | time interval between ECG R peak and PPG max_slope point (VPG $u$ point)                    | -0.54 (0.26) | -0.23 (0.38)     | -0.51 (0.34) | -0.43 (0.41) |
| 3                                | $PAT_{sys}$       | time interval between ECG R peak and PPG systolic peak                                      | -0.41 (0.47) | -0.14 (0.38)     | -0.35 (0.68) | -0.32 (0.40) |
| 4                                | $T_{pi}$          | pulse interval, the interval between pulse onset and pulse offset                           | 0.23 (0.36)  | -0.10 (0.43)     | 0.06 (0.54)  | -0.04 (0.55) |
| 5                                | $T_{sys}$         | systolic time, the interval between pulse onset and dicrotic notch                          | -0.03 (0.25) | -0.13 (0.24)     | -0.22 (0.49) | -0.28 (0.37) |
| 6                                | $T_{dia}$         | diastolic time, the interval between dicrotic notch and pulse offset                        | 0.20 (0.27)  | -0.08 (0.40)     | 0.02 (0.55)  | -0.03 (0.46) |
| 7                                | $T_{sp}$          | systolic peak time, the interval between pulse onset and systolic peak                      | -0.12 (0.59) | -0.06 (0.42)     | -0.13 (0.92) | -0.09 (0.60) |
| 8                                | $T_{dp}$          | diastolic peak time, the interval between pulse onset and diastolic peak                    | -0.20 (0.48) | -0.31 (0.55)     | -0.24 (0.61) | -0.45 (0.57) |
| 9                                | $\Delta T$        | time delay, the interval between systolic peak and diastolic peak                           | -0.15 (0.63) | -0.27 (0.43)     | -0.19 (0.88) | -0.32 (0.66) |
| 10                               | $T_u$             | u-point time, the interval between pulse onset and u-point                                  | -0.20 (0.54) | -0.16 (0.44)     | -0.35 (0.83) | -0.20 (0.67) |
| 11                               | $T_v$             | v-point time, the interval between pulse onset and v-point                                  | -0.02 (0.46) | -0.09 (0.59)     | -0.35 (0.68) | -0.21 (0.86) |
| 12                               | $T_w$             | w-point time, the interval between pulse onset and w-point                                  | -0.10 (0.30) | -0.21 (0.38)     | -0.41 (0.53) | -0.32 (0.58) |
| 13                               | $T_a$             | a-point time, the interval between pulse onset and a-point                                  | -0.19 (0.47) | -0.16 (0.44)     | -0.39 (0.72) | -0.27 (0.70) |
| 14                               | $T_b$             | b-point time, the interval between pulse onset and b-point                                  | -0.19 (0.52) | -0.20 (0.47)     | -0.41 (0.77) | -0.22 (0.66) |
| 15                               | $T_c$             | c-point time, the interval between pulse onset and c-point                                  | -0.28 (0.36) | -0.35 (0.26)     | -0.48 (0.44) | -0.55 (0.34) |
| 16                               | $T_d$             | d-point time, the interval between pulse onset and d-point                                  | 0.09 (0.26)  | -0.18 (0.35)     | -0.15 (0.48) | -0.27 (0.54) |
| 17                               | $T_e$             | e-point time, the interval between pulse onset and e-point                                  | -0.01 (0.38) | -0.22 (0.38)     | -0.32 (0.50) | -0.33 (0.58) |
| 18                               | $T_f$             | f-point time, the interval between pulse onset and f-point                                  | -0.06 (0.20) | -0.15 (0.24)     | -0.28 (0.53) | -0.34 (0.54) |
| 19                               | $T_{b-c}$         | b-c time, the interval between b-point and c-point                                          | -0.18 (0.27) | -0.27 (0.23)     | -0.39 (0.44) | -0.44 (0.35) |
| 20                               | $T_{b-d}$         | b-d time, the interval between b-point and d-point                                          | 0.19 (0.26)  | -0.08 (0.31)     | 0.13 (0.52)  | -0.05 (0.44) |
| Amplitude based features         |                   |                                                                                             |              |                  |              |              |
| 21                               | $A_{sp}$          | systolic peak amplitude, the different in amplitude between pulse onset and systolic peak   | -0.65 (0.50) | -0.48 (0.69)     | -0.75 (0.42) | -0.47 (0.68) |
| 22                               | $A_{dn}$          | dicrotic notch amplitude, the different in amplitude between pulse onset and dicrotic notch | -0.10 (0.46) | 0.02 (0.52)      | -0.41 (0.76) | -0.09 (0.69) |
| 23                               | $A_{dp}$          | diastolic peak amplitude, the different in amplitude between pulse onset and diastolic peak | -0.38 (0.36) | -0.21 (0.43)     | -0.55 (0.51) | -0.27 (0.48) |
| 24                               | $A_{off}$         | pulse onset amplitude, the different in amplitude between pulse onset and pulse offset      | 0.11 (0.15)  | 0.13 (0.19)      | 0.11 (0.24)  | 0.26 (0.30)  |
| Area based features              |                   |                                                                                             |              |                  |              |              |
| 25                               | $AUC_{pi}$        | pulse area, the area under the pulse wave between pulse onset and pulse offset              | -0.47 (0.47) | -0.44 (0.61)     | -0.67 (0.36) | -0.45 (0.63) |
| 26                               | $AUC_{sys}$       | systolic area, the area under the pulse wave between pulse onset and dicrotic notch         | -0.56 (0.46) | -0.44 (0.58)     | -0.72 (0.37) | -0.47 (0.61) |
| 27                               | $AUC_{dia}$       | diastolic area, the area under the pulse wave between dicrotic notch and pulse offset       | -0.22 (0.52) | -0.20 (0.43)     | -0.37 (0.73) | -0.23 (0.50) |
| Ratio of interval based features |                   |                                                                                             |              |                  |              |              |
| 28                               | $IPR$             | instantaneous pulse rate, $60/T_{pi}$                                                       | -0.23 (0.32) | 0.11 (0.44)      | -0.03 (0.50) | 0.02 (0.56)  |
| 29                               | $T_{sys}/T_{dia}$ | ratio of the systolic time versus diastolic time                                            | -0.19 (0.22) | 0.04 (0.41)      | -0.06 (0.48) | 0.01 (0.38)  |
| 30                               | $T_{sp}/T_{pi}$   | ratio of the systolic peak time versus pulse interval                                       | -0.23 (0.27) | 0.12 (0.36)      | -0.15 (0.64) | 0.08 (0.60)  |
| 31                               | $T_{dp}/T_{pi}$   | ratio of the diastolic peak time versus pulse interval                                      | -0.24 (0.35) | 0.01 (0.43)      | -0.10 (0.47) | -0.05 (0.52) |
| 32                               | $\Delta T/T_{pi}$ | ratio of the delay time versus pulse interval                                               | -0.26 (0.44) | -0.06 (0.43)     | -0.16 (0.64) | -0.11 (0.50) |
| 33                               | $T_u/T_{pi}$      | ratio of the u-point versus the pulse interval                                              | -0.26 (0.28) | 0.04 (0.28)      | -0.22 (0.61) | 0.02 (0.50)  |
| 34                               | $T_v/T_{pi}$      | ratio of the v-point time versus pulse interval                                             | -0.10 (0.53) | 0.10 (0.56)      | 0.04 (0.54)  | 0.01 (0.61)  |
| 35                               | $T_w/T_{pi}$      | ratio of the w-point time versus pulse interval                                             | -0.15 (0.36) | 0.01 (0.36)      | -0.22 (0.55) | -0.04 (0.47) |

Continued on next page



Table 5.2 – continued from previous page

|                                          | Notation                   | Definition                                                                                                                  | $r$ (with RM) |              | $r$ (without RM) |              |
|------------------------------------------|----------------------------|-----------------------------------------------------------------------------------------------------------------------------|---------------|--------------|------------------|--------------|
|                                          |                            |                                                                                                                             | SBP           | DBP          | SBP              | DBP          |
| 36                                       | $T_a/T_{pi}$               | ratio of the a-point time versus pulse interval                                                                             | -0.30 (0.40)  | -0.06 (0.30) | -0.37 (0.69)     | -0.14 (0.58) |
| 37                                       | $T_b/T_{pi}$               | ratio of the b-point time versus pulse interval                                                                             | -0.33 (0.30)  | 0.02 (0.28)  | -0.19 (0.54)     | 0.05 (0.48)  |
| 38                                       | $T_c/T_{pi}$               | ratio of the c-point time versus pulse interval                                                                             | -0.32 (0.18)  | -0.07 (0.31) | -0.29 (0.34)     | -0.22 (0.48) |
| 39                                       | $T_d/T_{pi}$               | ratio of the d-point time versus pulse interval                                                                             | -0.18 (0.35)  | 0.05 (0.30)  | -0.03 (0.45)     | 0.01 (0.45)  |
| 40                                       | $T_e/T_{pi}$               | ratio of the e-point time versus pulse interval                                                                             | -0.21 (0.21)  | 0.03 (0.37)  | -0.10 (0.37)     | -0.03 (0.44) |
| 41                                       | $T_f/T_{pi}$               | ratio of the f-point time versus pulse interval                                                                             | -0.19 (0.28)  | 0.02 (0.38)  | -0.12 (0.42)     | -0.01 (0.50) |
| 42                                       | $(T_u - T_a)/T_{pi}$       | ratio of the difference between the u-point-time and a-point-time versus pulse interval                                     | -0.15 (0.30)  | 0.13 (0.31)  | 0.04 (0.46)      | 0.12 (0.55)  |
| 43                                       | $(T_v - T_b)/T_{pi}$       | ratio of the difference between the v-point-time and b-point-time versus pulse interval                                     | -0.05 (0.38)  | 0.15 (0.52)  | 0.04 (0.59)      | 0.14 (0.74)  |
| <b>Ratio of amplitude based features</b> |                            |                                                                                                                             |               |              |                  |              |
| 44                                       | $A_{sp}/A_{off}$           | ratio of the systolic peak amplitude versus pulse offset amplitude                                                          | 0.06 (0.12)   | 0.05 (0.12)  | 0.04 (0.24)      | 0.08 (0.17)  |
| 45                                       | $A_{dp}/A_{sp}$            | reflection index, ratio of the diastolic peak amplitude versus systolic peak amplitude                                      | 0.42 (0.73)   | 0.45 (0.62)  | 0.66 (0.70)      | 0.47 (0.60)  |
| 46                                       | $(A_{sp} - A_{dp})/A_{sp}$ | ratio of the difference between systolic peak amplitude and diastolic peak amplitude versus systolic peak amplitude         | -0.42 (0.73)  | -0.45 (0.62) | -0.66 (0.70)     | -0.47 (0.60) |
| 47                                       | $A_{dn}/A_{sp}$            | ratio of the dicrotic notch amplitude versus systolic peak amplitude                                                        | 0.57 (0.63)   | 0.39 (0.51)  | 0.59 (0.60)      | 0.49 (0.43)  |
| 48                                       | $PIR$                      | pulse intensity ratio, $A_{sp}/A_{on}$ , ratio of the systolic peak amplitude versus the pulse onset amplitude ( $A_{on}$ ) | 0.32 (0.42)   | 0.48 (0.42)  | 0.56 (0.37)      | 0.59 (0.43)  |
| 49                                       | $A_u/A_{sp}$               | ratio of the u-point amplitude versus the systolic peak amplitude                                                           | 0.02 (0.08)   | 0.00 (0.07)  | 0.04 (0.22)      | 0.01 (0.21)  |
| 50                                       | $A_v/A_u$                  | ratio of the v-point amplitude versus the u-point amplitude                                                                 | -0.01 (0.07)  | 0.00 (0.08)  | -0.02 (0.25)     | 0.02 (0.24)  |
| 51                                       | $A_w/A_u$                  | ratio of the w-point amplitude versus the u-point amplitude                                                                 | 0.00 (0.07)   | 0.00 (0.08)  | -0.01 (0.25)     | -0.01 (0.23) |
| 52                                       | $A_b/A_a$                  | ratio of the b-point amplitude versus the a-point amplitude                                                                 | -0.01 (0.06)  | 0.01 (0.05)  | 0.02 (0.26)      | 0.02 (0.24)  |
| 53                                       | $A_c/A_a$                  | ratio of the c-point amplitude versus the a-point amplitude                                                                 | 0.01 (0.08)   | -0.01 (0.07) | 0.03 (0.20)      | 0.01 (0.18)  |
| 54                                       | $A_d/A_a$                  | ratio of the d-point amplitude versus the a-point amplitude                                                                 | 0.00 (0.10)   | -0.01 (0.08) | -0.01 (0.25)     | -0.01 (0.13) |
| 55                                       | $A_e/A_a$                  | ratio of the e-point amplitude versus the a-point amplitude                                                                 | -0.01 (0.11)  | 0.00 (0.09)  | -0.01 (0.25)     | -0.02 (0.21) |
| 56                                       | $A_f/A_a$                  | ratio of the f-point amplitude versus the a-point amplitude                                                                 | 0.00 (0.07)   | -0.01 (0.06) | 0.00 (0.22)      | 0.01 (0.28)  |
| 57                                       | $(A_c - A_b)/A_a$          | ratio of the difference between the b-point amplitude and c-point amplitude versus the a-point amplitude                    | 0.00 (0.08)   | 0.00 (0.06)  | -0.02 (0.19)     | -0.02 (0.17) |
| 58                                       | $(A_d - A_b)/A_a$          | ratio of the difference between the b-point amplitude and d-point amplitude versus the a-point amplitude                    | 0.01 (0.08)   | 0.00 (0.06)  | 0.00 (0.20)      | 0.01 (0.20)  |
| 59                                       | $AGI$                      | aging index, $(A_b - A_c - A_d - A_e)/A_a$                                                                                  | -0.01 (0.09)  | 0.00 (0.05)  | -0.03 (0.23)     | 0.00 (0.22)  |
| 60                                       | $AGI_{inf}$                | informal aging index, $(A_b - A_e)/A_a$                                                                                     | 0.00 (0.09)   | 0.01 (0.08)  | 0.01 (0.28)      | 0.03 (0.26)  |
| <b>Ratio of area based features</b>      |                            |                                                                                                                             |               |              |                  |              |
| 61                                       | $IPA$                      | inflection point area, $AUC_{dia}/AUC_{sys}$                                                                                | 0.35 (0.49)   | 0.27 (0.50)  | 0.58 (0.43)      | 0.37 (0.55)  |
| <b>Ratio of combined features</b>        |                            |                                                                                                                             |               |              |                  |              |
| 62                                       | $T_{sp}/A_{sp}$            | ratio of the systolic peak time versus the systolic peak amplitude                                                          | 0.56 (0.63)   | 0.42 (0.68)  | 0.67 (0.49)      | 0.42 (0.61)  |
| 63                                       | $A_{sp}/\Delta T$          | stiffness index, ratio of the systolic peak amplitude versus the time delay                                                 | -0.60 (0.47)  | -0.42 (0.64) | -0.71 (0.39)     | -0.43 (0.67) |
| 64                                       | $A_{sp}/(T_{pi} - T_{sp})$ | ratio of the systolic peak amplitude versus the difference between the pulse interval and systolic peak time                | -0.68 (0.49)  | -0.30 (0.66) | -0.70 (0.42)     | -0.37 (0.64) |
| 65                                       | $T_{sys}/A_{dn}$           | ratio of the systolic time versus the dicrotic notch amplitude                                                              | 0.10 (0.57)   | -0.03 (0.41) | 0.31 (0.74)      | -0.01 (0.57) |
| 66                                       | $T_{dp}/A_{dp}$            | ratio of the diastolic peak time versus the diastolic peak amplitude                                                        | 0.31 (0.38)   | 0.12 (0.42)  | 0.45 (0.54)      | 0.19 (0.49)  |
| 67                                       | WN                         | Womersley number: $A_{on} \cdot \sqrt{1060 \cdot IPR/A_{sp}}$ , $A_{on}$ is the pulse onset amplitude                       | 0.31 (0.32)   | -0.02 (0.50) | 0.39 (0.51)      | 0.14 (0.43)  |
| 68                                       | SC                         | spring constant, $APG(T_{sp})/((A_{sp} - A_u)/(A_{sp}))$                                                                    | -0.02 (0.09)  | -0.02 (0.07) | -0.04 (0.34)     | -0.07 (0.24) |
| 69                                       | IPAD                       | inflection point area normalized d-point amplitude, $AUC_{dia}/AUC_{sys} + A_d/A_a$                                         | 0.03 (0.12)   | 0.01 (0.10)  | 0.07 (0.24)      | 0.06 (0.20)  |

Previously, PATs which is considered as the surrogate to PTT which relates to BP changes via Moens-Korteweg equation (as described in Section 2.3.1.1) have been long and widely for cuffless BP estimation, a number of other features extracted from PPG and its derivatives have been reported strongly relate to BPs, such as PIR [187], and Womersley number [78], and features extracted from the second derivative of PPG (APG) which can reflect age-related arterial stiffness, like  $e/a$ ,  $c/a$  were found to have median ranking coefficients of 0.89, 0.84 to BPs [66]. Our experiments showed that some rarely discussed features have strong correlations to BP changes.

The PPG pulse amplitude (feature 21, denoted as  $A_{sp}$  in Table 5.2) which is an indicator of the pulsatile changes in blood volume caused by arterial blood flow around the measurement site [188, 189] has the highest median correlation coefficient of -0.75 to SBP after removing respiration modulation using the method proposed in Chapter 3, some  $A_{sp}$  related features also demonstrated strong correlation to SBP, for example, the median correlation coefficients between  $A_{sp}/\Delta T$ ,  $A_{sp}/(T_{pi} - T_{sp})$  and  $T_{sp}/A_{sp}$  (feature 63, 64 and 62 in Table 5.2), and SBP are -0.71, -0.70 and 0.67 respectively among 32 subjects from **BCG BP** dataset. Additionally, the median correlation coefficients between  $(A_{sp} - A_{dp})/A_{sp}$ ,  $A_{dp}/A_{sp}$  and  $A_{dn}/A_{sp}$  (feature 46, 45 and 47 in Table 5.2), and SBP are -0.66, 0.66 and 0.59 respectively which are all greater than widely used and generally agreed important features,  $PAT_{ms}$  and  $PIR$ , with the median correlation coefficients to SBP of -0.51 and 0.56. These findings indicate that the PPG pulse amplitude  $A_{sp}$  is strongly inversely proportional to SBP and can be potentially used for BP trend tracking. Previous study [189] also demonstrates the similar results. The correlation coefficient between reference and estimated SBP using PPG pulse amplitude is 0.70, comparing to the correlation coefficient of 0.43 when using PAT and subject is in supine posture with spontaneous breathing. Another study [190] found that PPG pulse amplitude was more effective than PAT for estimating SBP during sleep.

It is also worth noting that the relationship between physiological features and BPs may vary depending on the subject's posture. For example, the correlation coefficient between reference and estimated SBP when subject is standing with spontaneous breathing using PAT is 0.64, comparing to 0.34 of using PPG pulse amplitude. And this study also indicates that BP estimations are more accurate when using PAT and PPG pulse amplitude together for both supine and standing postures [189]. Since the results shown in Table 5.2 are obtained from the **BCG BP** dataset where physiological signals such as PPG, ECG and ABP were collected when subjects were asked to recline in a customized bed, which may indicate different sets of important features should be evaluated and extracted for more accurate BP estimation. This hypothesis needs to be studied and verified by future studies.

## 5.2 Cuffless Blood Pressure Estimation Models

In this section, two cuffless BP estimation models are proposed to predict BPs using trend-related features via time series regression.

### 5.2.1 A TCN-SENet-KAN Model

The proposed model combines [Temporal Convolutional Network \(TCN\)](#), [Squeeze-and-Excitation Network \(SENet\)](#) and a [Kolmogorov-Arnold Network \(KAN\)](#). TCNs aim to capture temporal dependencies within the input physiological features, SENet is employed to adaptively recalibrate channel-wise feature importance, and KAN inspired linear layers to enhance modeling non-linear relationships. The structure of this proposed TCN-SENet-KAN model is demonstrated in Figure 5.3.

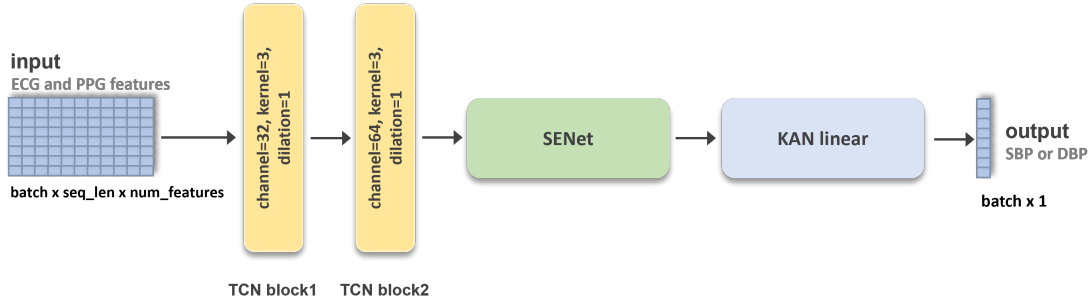


Figure 5.2: The framework of the proposed TCN-SENet-KAN model for BP estimation via time series regression method.

#### 5.2.1.1 Temporal Convolutional Network (TCN) Block

[TCNs](#) [191] are a class of neural network architectures specifically designed for sequence modeling tasks. TCNs adapt a convolutional approach to processing time series data, providing an alternative to [RNN](#). TCNs improve the ability to capture long-range temporal dependencies through a combination of causal convolutions, dilated convolutions, and residual connections. The TCN block employed for this model is demonstrated in Figure 5.3 (a).

**5.2.1.1.1 Causal Convolutions** A fundamental property of TCNs is causality, meaning that the output at time step  $t$  can only depend on inputs at time  $t$  and previous time steps  $(t-1, t-2, \dots)$ . This is achieved by using one-dimensional (1D) convolutional layers where the kernel is constrained to only capture past and current information. For an input sequence  $\mathbf{x} = (x_0, x_1, \dots, x_T)$  and a filter  $\mathbf{f} = (f_0, f_1, \dots, f_{k-1})$  of kernel size  $k$ , the causal convolution output  $y_t$  at time  $t$  is given by:

$$y_t = (\mathbf{x} * \mathbf{f})(t) = \sum_{i=0}^{k-1} f_i \cdot x_{t-i} \quad (5.1)$$

**5.2.1.1.2 Dilated Convolutions** TCNs employ dilated convolution to efficiently capture long-range dependencies without requiring a very deep network or excessively large kernels. A dilated convolution introduces gaps between kernel taps, effectively expanding the receptive field. For a 1D input sequence  $\mathbf{x}$ , a kernel  $\mathbf{f}$  of size  $k$ , and a dilation factor  $d$ , the dilated convolution operation at time  $t$  is defined as:

$$y_t = \sum_{i=0}^{k-1} f_i \cdot x_{t-d \cdot i} \quad (5.2)$$

By stacking layers of dilated convolutions with exponentially increasing dilation factors (e.g.,  $d = 1, 2, 4, 8, \dots$ ), TCNs can achieve very large receptive fields with a modest number of layers and parameters.

**5.2.1.1.3 Residual Connections** TCN blocks typically incorporate residual connections [192]. As shown in Figure 5.3 (a), the designed TCN block consists of two dilated causal convolutions, Chomp modules to remove paddings and a non-linear activation function (e.g., ReLU). The input to the block  $\mathbf{x}_{in}$  is added to the output of these convolutional layers  $F(\mathbf{x}_{in})$  via residual connection. If the input and output have different dimensions, a 1x1 convolution is used on  $\mathbf{x}_{in}$  to match the dimensions. The output of a residual block is:

$$\mathbf{o} = \text{Activation}(\mathbf{x}_{in} + F(\mathbf{x}_{in})) \quad (5.3)$$

## 5.2.1.2 Squeeze-and-Excitation Network (SENet) Block

The Squeeze-and-Excitation Network (SENet) [193] is an architectural unit designed to improve the representation of a network by enabling it to perform dynamic channel-wise

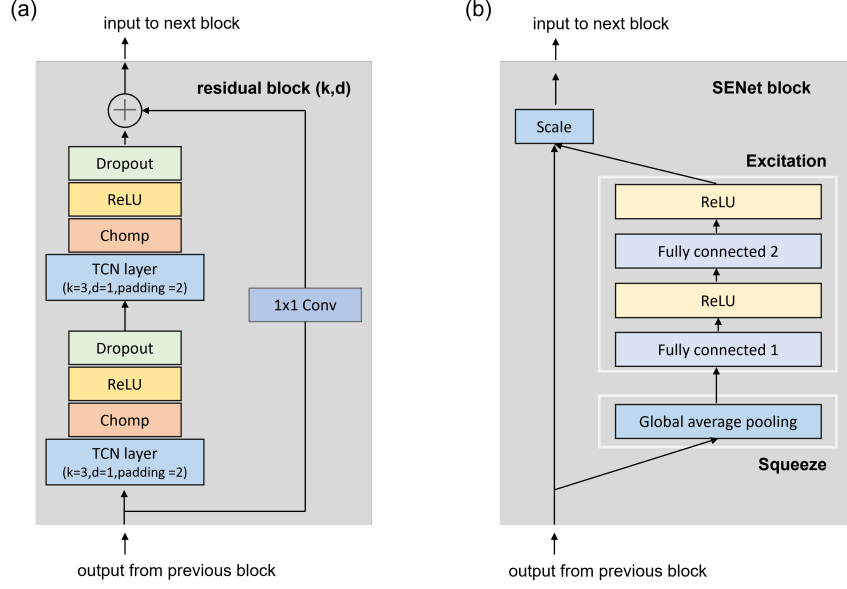


Figure 5.3: The structures of (a) TCN block, (b) SENet block employed for the proposed BP estimation model.

feature recalibration. Essentially, SENet allows the network to explicitly model interdependencies between channels and adaptively re-weight feature maps, emphasizing informative features while suppressing less useful ones. The SENet block operates in two main steps: Squeeze and Excitation.

**5.2.1.2.1 Squeeze Operation** The Squeeze operation aggregates global spatial or temporal for 1D signal information into a channel descriptor. This is typically achieved using global average pooling across the spatial/temporal dimensions of each feature map. For an input feature map  $\mathbf{U} \in \mathbb{R}^{C \times L}$  (where  $C$  is the number of channels and  $L$  is the length of the sequence), the  $c$ -th element of the squeezed vector  $\mathbf{z} \in \mathbb{R}^C$  is calculated as:

$$z_c = F_{sq}(\mathbf{u}_c) = \frac{1}{L} \sum_{i=1}^L u_c(i) \quad (5.4)$$

This descriptor  $z_c$  embeds the global receptive field information for channel  $c$ .

**5.2.1.2.2 Excitation Operation** The Excitation operation aims to learn a non-linear, channel-wise gating mechanism based on the information aggregated by the Squeeze step. This is typically implemented using two fully connected (FC) layers with a dimensionality reduction and subsequent increase, followed by a sigmoid activation function. The channel-wise statistics  $\mathbf{z}$  are passed through these layers to produce channel weights  $\mathbf{s} \in \mathbb{R}^C$ :

$$\mathbf{s} = F_{ex}(\mathbf{z}, \mathbf{W}) = \sigma(\mathbf{W}_2 \delta(\mathbf{W}_1 \mathbf{z})) \quad (5.5)$$

where  $\delta$  is the ReLU activation function,  $\sigma$  is the sigmoid activation function,  $\mathbf{W}_1 \in \mathbb{R}^{\frac{C}{r} \times C}$  is the weight matrix for the first FC layer which is used to reduce dimensionality by a ratio  $r$ , and  $\mathbf{W}_2 \in \mathbb{R}^{C \times \frac{C}{r}}$  is the weight matrix for the second FC layer which is employed to restore dimensionality. The learned weights  $s_c$  which range from 0 to 1 represent the importance of each channel.

### 5.2.1.3 Recalibration

The final output of the SE block is obtained by rescaling the original input feature map  $\mathbf{U}$  with the learned channel activations  $\mathbf{s}$ . The  $c$ -th channel of the output feature map  $\tilde{\mathbf{X}}_c$  is:

$$\tilde{\mathbf{x}}_c = F_{scale}(\mathbf{u}_c, s_c) = s_c \cdot \mathbf{u}_c \quad (5.6)$$

This operation allows the network to selectively amplify useful feature channels and diminish less relevant ones. The structure of the SENet employed for the proposed BP estimation model is demonstrated in Figure 5.3 (b).

## 5.2.2 Kolmogorov-Arnold Network (KAN) Linear Layer

Kolmogorov-Arnold Networks (KANs) [194] present an alternative to traditional Multi-Layer Perceptrons (MLPs). Inspired by the Kolmogorov-Arnold representation theorem, KANs place learnable activation functions on the edges (weights) of the network, rather than having fixed activation functions at the nodes like conventional MLP. These learnable activation functions are typically parameterized as splines. This architectural design allows KANs to learn complex non-linear relationships with potentially greater accuracy and parameter efficiency compared to MLPs, which rely on the composition of linear transformations and fixed non-linear activation functions. The KAN model and its difference to MLP model are demonstrated in Figure 5.4. The KAN linear layer, as adapted in this model, aims to capture intricate non-linearities. A general KAN layer transforms an input

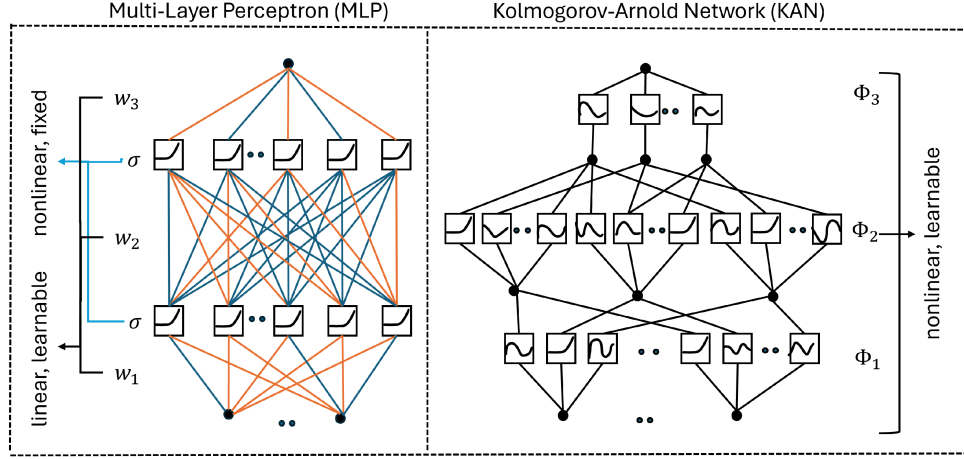


Figure 5.4: Architectures of the KAN and MLP. The diagram highlights the layer composition and activation function differences between the KAN and MLP models (figure adapted from [195]).

vector  $\mathbf{x} = (x_1, \dots, x_{n_{in}})$  to an output vector  $\mathbf{y} = (y_1, \dots, y_{n_{out}})$  where the  $j$ -th output  $y_j$  is given by:

$$y_j = \sum_{i=1}^{n_{in}} \phi_{j,i}(x_i) \quad (5.7)$$

Here, each  $\phi_{j,i} : \mathbb{R} \rightarrow \mathbb{R}$  is a learnable 1D function which is an activation function on the edge connecting input  $x_i$  to the computation of  $y_j$ , parameterized as a spline.

This study employs an implementation of KAN linear layer [196] which combines a traditional linear transformation with a spline-based transformation to predict SBP or DBP with the last step of the extracted feature representations from SENet. The function  $\phi_{j,i}(x_i)$  can be conceptualized as consisting of two parts: a base component and a spline component.

$$\phi_{j,i}(x_i) = w_{base_{j,i}} \cdot \psi(x_i) + \sum_{k=1}^M c_{j,i,k} B_k(x_i) \quad (5.8)$$

where:

- $\psi(x_i)$  is a fixed base activation function (e.g., SiLU) applied to the input  $x_i$ .
- $w_{base_{j,i}}$  is a learnable weight for the base activated component of input  $x_i$  contributing to output  $y_j$ .

- $B_k(x_i)$  are B-spline basis functions evaluated at  $x_i$ . B-splines are piecewise polynomials that provide a flexible way to approximate functions. The  $k$ -th order B-spline basis function,  $B_{m,k}(x)$ , defined over a knot vector  $t_0, t_1, \dots, t_{m+k+1}$ , can be computed recursively using the Cox-de Boor formula:

$$B_{m,0}(x) = \begin{cases} 1 & \text{if } t_m \leq x < t_{m+1} \\ 0 & \text{otherwise} \end{cases} \quad (5.9)$$

$$B_{m,k}(x) = \frac{x - t_m}{t_{m+k} - t_m} B_{m,k-1}(x) + \frac{t_{m+k+1} - x}{t_{m+k+1} - t_{m+1}} B_{m+1,k-1}(x) \quad (5.10)$$

For  $t_{m+k} = t_m$  or  $t_{m+k+1} = t_{m+1}$ , the respective term is taken as zero.

- $c_{j,i,k}$  are learnable coefficients for the B-spline basis functions.
- $M$  is the number of B-spline basis functions used.

The KAN linear layer's output is thus a sum of a linear transformation of base-activated inputs and a linear combination of B-spline transformations of inputs, allowing it to model complex non-linear functions effectively. The splines' control points (coefficients  $c_{j,i,k}$ ) and the grid on which they are defined can be learned during training, providing greater adaptability than fixed activation functions. Regularization terms can be employed to encourage smoother or sparser spline functions.

### 5.2.3 A Parallel Transformer-BiGRU Model with Attention Mechanisms

A parallel model which consists a transformer encoder branch to capture global relationships and long-range dependencies from model input, a [Bidirectional Gated Recurrent Unit \(BiGRU\)](#) with global attention mechanism branch is employed to capture sequential context and temporal order, and a cross-attention block to fuse feature representations of both branches for BP estimation. The structure of this model is shown in [Figure 5.5](#)

#### 5.2.3.1 Transformer Encoder Branch

This branch utilizes a Transformer Encoder architecture introduced by Vaswani et al. [197]. Each layer in the Transformer Encoder primarily consists of two sub-layers: Multi-Head Self-Attention (MHSA) and a position-wise Feed-Forward Network (FFN), with residual connections and layer normalization applied to each.



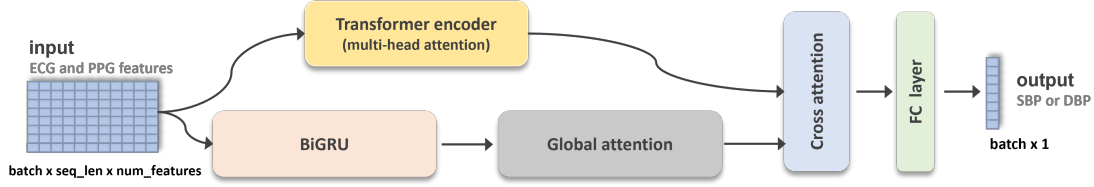


Figure 5.5: The structure of a parallel transformer and BiGRU with attention mechanisms for BP estimation.

**5.2.3.1.1 Multi-Head Self-Attention (MHSA):** The MHSA allows Transformer block to evaluate the significance of different parts of the input sequence. For each position in the sequence, self-attention computes a weighted sum of all other positions, where the weights (attention scores) indicate relevance. This allows the model to capture context by directly relating different segments of the sequence, regardless of their distance. The "multi-head" aspect allows this process to be performed multiple times in parallel with different learned linear projections (queries  $\mathbf{Q}$ , keys  $\mathbf{K}$ , and values  $\mathbf{V}$ ), capturing various types of relationships. The fundamental scaled dot-product attention is:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right) \mathbf{V} \quad (5.11)$$

The outputs from multiple heads are then concatenated and linearly projected.

**5.2.3.1.2 Position-wise Feed-Forward Network (FFN):** Following the MHSA module, each position's representation is independently processed by a fully connected feed-forward network which consists of two linear transformations with a ReLU activation in between, allowing further non-linear transformation of the attended features.

### 5.2.3.2 Bidirectional GRU (BiGRU) Branch with Global Attention

This parallel branch focuses on capturing sequential context and temporal order using [BiGRUs](#), enhanced by a subsequent global attention mechanism.

**5.2.3.2.1 Bidirectional GRU (BiGRU):** [GRUs](#) [198] are a type of [RNN](#) designed to better capture dependencies in sequences and mitigate issues like the vanishing gradient problem. GRUs employ gating mechanisms which consists of an update gate and a reset gate. These gates collaboratively control the amount of the past information (previous

hidden state) to retain and the amount of the new input information should be integrated to form the current hidden state. A BiGRU processes the input sequence in two directions, a forward GRU reads the sequence from start to end, and a backward GRU reads it oppositely. The hidden states from both directions at each time step are combined, providing a representation that incorporates both past and future context for every point in the sequence.

**5.2.3.2 Global Attention Mechanism:** Following the BiGRU layers, a global attention mechanism [199] is employed to compute a single, fixed-length context vector that summarizes the entire output sequence from the BiGRU ( $\mathbf{H}_{gru}$ ). This is achieved by calculating attention weights for each BiGRU hidden states which can indicate their relevance with respect to a query (often derived from the final BiGRU hidden state or a separate learnable query vector). The context vector  $\mathbf{c}$  is then a weighted average of all BiGRU hidden states:

$$\mathbf{c} = \sum_{t=1}^L \alpha_t \mathbf{h}_t^{gru} \quad (5.12)$$

where  $\alpha_t$  are the attention weights. This allows the model to dynamically focus on the most salient parts of the BiGRU-processed sequence when forming its summary representation.

### 5.2.3.3 Cross-Attention Fusion

To integrate the distinct feature representations learned by the Transformer and the BiGRU-GlobalAttention branches, a cross-attention mechanism is employed. This allows one branch's representation to selectively draw information from the other, facilitating a synergistic fusion of global and sequential features. In this setup, one set of features (e.g., the sequence output from the Transformer branch,  $\mathbf{H}_{trans}$ ) can act as the "query" sequence. The other set of features (e.g., the context vector or sequence output from the BiGRU-GlobalAttention branch,  $\mathbf{H}_{bgru}$ ) provides the "keys" and "values." The cross-attention mechanism then computes attention scores to indicate the relevance between the key/value sequence ( $\mathbf{H}_{bgru}$ ) and the query sequence ( $\mathbf{H}_{trans}$ ). The output is a weighted sum of the values, where weights are determined by query-key similarities calculated by scaled dot-product attention:

$$\text{CrossAttentionOutput}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax} \left( \frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}} \right) \mathbf{V} \quad (5.13)$$

Here,  $\mathbf{Q}$  would be derived from  $\mathbf{H}_{trans}$ , and  $\mathbf{K}, \mathbf{V}$  from  $\mathbf{H}_{bgru}$ . This results in a refined representation of the query sequence, enriched with context from the other branch. A fully connected (FC) layer is then added to predict SBP and DBP from the output of cross attention module.

## 5.3 Loss Function

As described in the last section, both proposed TCN-SENet-KAN and a parallel Transformer-BiGRU-Attention aim to capture the temporal dependencies and extract important features from the model inputs to enhance BP trend tracking via time series regression. In this section, a shape-oriented loss function—sDTW is introduced and customized to further improve the performance and generalization of BP estimation model for trend tracking.

### 5.3.1 Pointwise Distance-Based Loss Function

Conventionally, pointwise distance metrics, such as MAE and Mean Squared Error (MSE) are commonly used as the loss function for regression problems using deep learning models.

#### 5.3.1.1 Mean Squared Error (MSE)

The MSE loss is one of the most widely used loss functions for regression tasks. Formally, for a dataset containing  $N$  samples with predictions  $\hat{y}_i$  and true targets  $y_i$ , the MSE is defined as:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2 \quad (5.14)$$

MSE penalizes large errors more severely than small errors due to the square term.

#### 5.3.1.2 Mean Absolute Error (MAE)

The MAE loss computes the average absolute difference between predictions and targets:

$$\mathcal{L}_{\text{MAE}} = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (5.15)$$

Unlike MSE, MAE is less sensitive to outliers because errors are not squared. However, MAE is not differentiable at zero, although in practice sub-gradient methods or smooth approximations are used to train deep networks. MAE typically leads to a more robust estimation when the training data contains significant noise or outliers.

### 5.3.2 A Shape-Oriented Loss Function

[Dynamic Time Warping \(DTW\)](#) is a well-established algorithm used to measure the similarity between two temporal sequences, which may vary in speed, phase, or length [200]. [sDTW](#) [201] is a differentiable surrogate for the classic DTW algorithm, meaning it can be used as the loss function for deep learning models.

#### 5.3.2.1 Dynamic Time Warping (DTW)

Unlike pointwise distance measures (e.g., MSE and MAE mentioned in Section 5.3.1) that compare the distance the  $i$ -th points of two compared sequences, DTW finds an optimal non-linear alignment between the sequences by "warping" the time axis. This process makes it particularly effective for comparing time series that are qualitatively similar but locally out of phase.

Let two time series for comparison be  $\mathbf{X} = (x_1, x_2, \dots, x_N)$  of length  $N$  and  $\mathbf{Y} = (y_1, y_2, \dots, y_M)$  of length  $M$ . The core idea of DTW is to construct a cost matrix,  $\mathbf{C} \in \mathbb{R}^{N \times M}$ , where each element  $C(i, j)$  represents the distance between point  $x_i$  from sequence  $\mathbf{X}$  and point  $y_j$  from sequence  $\mathbf{Y}$ . This local distance  $d(x_i, y_j)$  is calculated by Euclidean distance, such as squared error:

$$d(x_i, y_j) = (x_i - y_j)^2 \quad (5.16)$$

or the absolute difference  $d(x_i, y_j) = |x_i - y_j|$ .

The goal is to find a warping path  $\mathbf{W} = (w_1, w_2, \dots, w_K)$  where each  $w_k = (i_k, j_k)$  is a pair of indices from  $\mathbf{X}$  and  $\mathbf{Y}$  respectively, mapping  $x_{i_k}$  to  $y_{j_k}$ . This path starts at  $w_1 = (1, 1)$  and ends at  $w_K = (N, M)$ , and must satisfy certain conditions like monotonicity (indices are non-decreasing) and continuity (steps are to adjacent cells). The optimal warping path is the one that minimizes the total cumulative cost of aligned points. A demonstration of locating the optimal warping path for two compared sequences using DTW is shown in Figure 5.6 (a).

This optimal path is found using dynamic programming. An accumulated cost matrix  $\mathbf{D} \in \mathbb{R}^{N \times M}$  is constructed, where  $D(i, j)$  is the minimum cumulative cost of any warping path from  $(1, 1)$  to  $(i, j)$ . The recurrence relation to compute  $D(i, j)$  is:

$$D(i, j) = d(x_i, y_j) + \min \begin{cases} D(i-1, j) \\ D(i, j-1) \\ D(i-1, j-1) \end{cases} \quad (5.17)$$

with initial conditions  $D(1, 1) = d(x_1, y_1)$ , and  $D(i, 1) = \sum_{k=1}^i d(x_k, y_1)$ ,  $D(1, j) = \sum_{k=1}^j d(x_1, y_k)$  (or by setting costs for out-of-bounds indices to infinity).

The DTW distance between sequences  $\mathbf{X}$  and  $\mathbf{Y}$  is then the final accumulated cost  $D(N, M)$ . A lower DTW distance indicates higher similarity between the shapes of the two sequences, irrespective of minor shifts or scaling in the time dimension.

### 5.3.2.2 Soft Dynamic Time Warping (sDTW)

The standard DTW algorithm's reliance on the non-differentiable min operator in its dynamic programming formulation (Equation 5.17) prevents its direct application as a loss function in gradient-based optimization, which is prevalent in deep learning.

The key innovation of sDTW is the substitution of the min operator with a differentiable "soft-minimum" operator, denoted as  $\min^\gamma$ . Given a smoothing parameter  $\gamma > 0$ , the soft-minimum of a set of values  $\{a_1, a_2, \dots, a_k\}$  is defined as:

$$\min^\gamma(a_1, \dots, a_k) = -\gamma \log \sum_{s=1}^k \exp\left(-\frac{a_s}{\gamma}\right) \quad (5.18)$$

This operator effectively computes a smoothed, differentiable approximation of the minimum value. Using this soft-minimum, the recurrence relation for computing the sDTW accumulated cost, denoted as  $D^\gamma(i, j)$ , becomes:

$$D^\gamma(i, j) = d(x_i, y_j) + \min^\gamma \begin{cases} D^\gamma(i-1, j) \\ D^\gamma(i, j-1) \\ D^\gamma(i-1, j-1) \end{cases} \quad (5.19)$$

The local cost  $d(x_i, y_j)$  is the same as defined in Equation 5.16. The final sDTW score between sequences  $\mathbf{X}$  and  $\mathbf{Y}$  is  $D^\gamma(N, M)$ .

As the smoothing parameter  $\gamma \rightarrow 0$ , the  $\min^\gamma$  operator converges to the standard (hard) min operator, and consequently, sDTW converges to the standard DTW value. However, for any  $\gamma > 0$ , the sDTW score  $D^\gamma(N, M)$  is smooth and differentiable with respect to the elements  $x_i$  and  $y_j$  of the input sequences (via the local cost  $d(x_i, y_j)$ ). This differentiability allows for the computation of gradients of sDTW with respect to its inputs. A demonstration of locating the optimal warping path for two compared sequences using sDTW is shown in Figure 5.6 (b). This property is crucial for deep learning applications, as it enables sDTW to be used directly as a loss function. By optimizing a model against an sDTW loss, the network can be trained to produce output sequences that are not only close in value but also similar in shape and temporal structure to target sequences. This is particularly beneficial for tasks such as BP trend estimation, where capturing the morphological characteristics and dynamic patterns of the BP waveform is often more important than minimizing pointwise errors like MAE. The sDTW loss encourages the model to learn these shape-oriented features, potentially leading to more accurate BP trend tracking performance and enhanced model generalization.

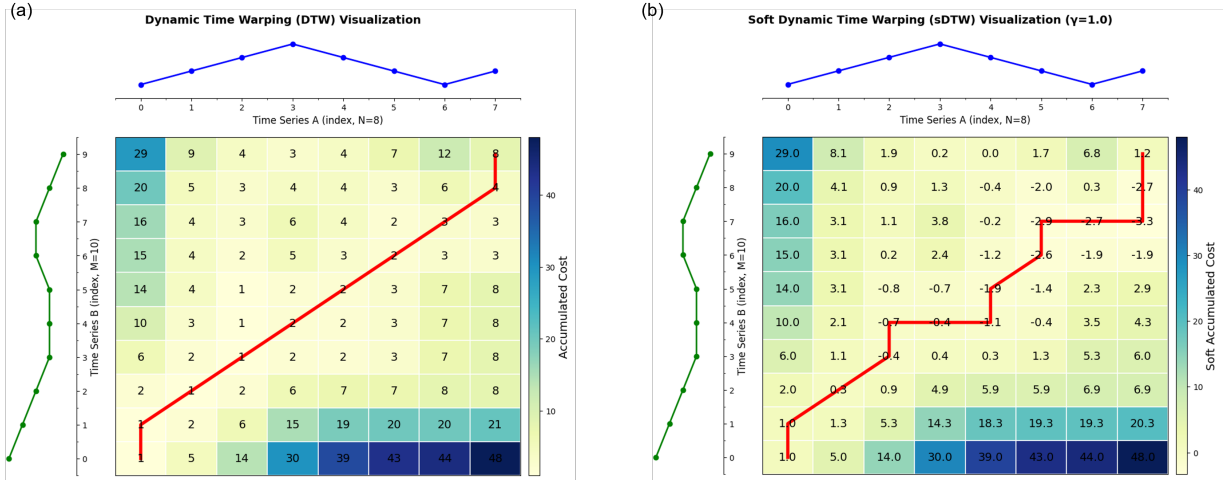


Figure 5.6: Demonstrations of (a) dynamic time warping (DTW) and (b) soft dynamic time warping (sDTW) of  $\gamma = 1.0$ .

### 5.3.2.3 Customized sDTW for Accurate BP Trend Tracking

As described in the last section that sDTW quantifies the difference of both distance and shape between two compared sequences. The previous findings indicate that the extracted features from ECG and PPG waveforms can be used as indicators for BP trend tracking (see Section 5.1), but personal calibrations are still required for accurate BP estimation. Therefore, the customized loss function is defined as the sDTW of the reference and prediction at zero mean, such that the model can exclusively focus on the shape similarity between the estimation and reference in terms of pattern and scale. Then this customized sDTW loss function is defined as:

$$\text{sDTW}_{\text{loss}} = \text{sDTW}(y_{\text{ref}} - \overline{y_{\text{ref}}}, y_{\text{pred}} - \overline{y_{\text{pred}}}) \quad (5.20)$$

where  $\overline{y_{\text{ref}}}$  and  $\overline{y_{\text{pred}}}$  are the mean of the reference and estimation respectively.

## 5.4 Experiments

### 5.4.1 Bed-Based Ballistocardiography Dataset

The Bed-based Ballistocardiography (BCG) Dataset [202] (referred as **BCG BP**) which collects three-lead ECG, finger PPG, BCG, simultaneous reference ABP and respiration waveforms. These physiological waveforms are sampled at 1 KHz which enhances the resolution of extracted features and can potentially improve BP estimation accuracy. This dataset includes 40 participants (17 males) whose ages range from 18 to 65 years and their body mass indices (BMIs) range from 18 to 48 kg/m<sup>2</sup>, and four participants indicated that they had some form of past or current cardiovascular-related conditions. Each subject was asked to recline in a customized bed with BCG sensors and an average of 6.97 minutes' data were collected. To ensure fiducial points of physiological signals and morphological features are accurately detected and extracted for RM evaluation and cuffless BP estimation, 8 subjects who has abnormal physiological waveforms are excluded from this study (e.g., a subject's PPG shows multiple ripples in diastolic phase is demonstrated in Figure 5.8 (b)). The mean and standard deviation of SBP and DBP of 32 selected subjects are 122.62±15.02 mmHg and 67.03±9.41 mmHg respectively and distributions of SBP and DBP are demonstrated in Figure 5.7.

The **BCG BP** dataset, though relatively small in size, provides high-quality waveforms sampled at 1 kHz. This high sampling rate enables more precise feature extraction and

potentially improves estimation performance (see Appendix A.5), in contrast to datasets like MIMIC datasets [203], which are sampled at only 125 Hz and also demonstrate an unknown time delay between ECG and PPG waveforms (see Appendix A.1). Additionally, the BCG BP dataset demonstrates substantial BP variability. Among the selected 32 subjects, the SBP and DBP ranges, represented by the median and interquartile range (IQR), are 18.25 (12.68) mmHg and 9.87 (6.32) mmHg, respectively. Furthermore, the dataset provides continuous recordings, which are more suitable for BP trend tracking compared to segmented datasets such as PulseDB [108], where ECG, PPG, and ABP waveforms are limited to 10-second segments.

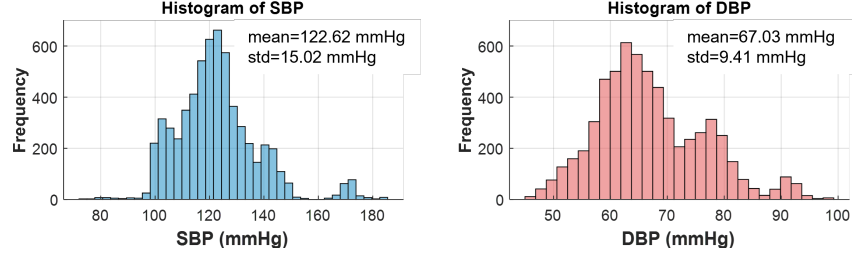


Figure 5.7: Distribution of BP readings of selected 32 subjects in BCG BP dataset.

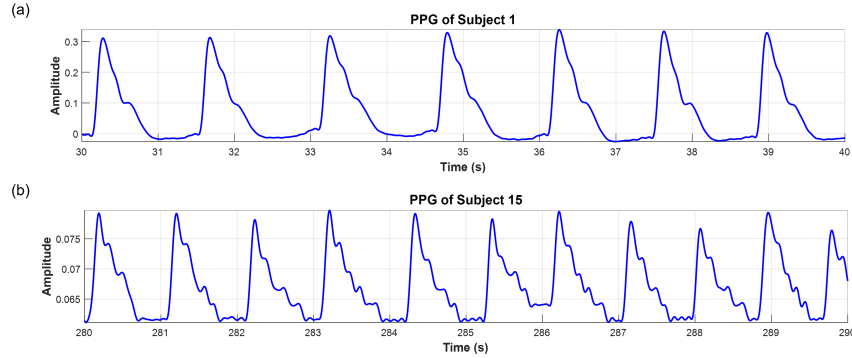


Figure 5.8: Demonstration of (a) a good quality PPG waveform segment, and (b) abnormal PPG waveform which has multiple ripples in diastolic phase is excluded from further analysis.



### 5.4.2 Pre-processing

A 4th-order Chebyshev Type II bandpass filter, with a low cut-off frequency of 0.5 Hz and a high cut-off frequency of 12 Hz, is applied to the PPG waveform to remove baseline wander and high-frequency noise. This filter design was recommended by previous studies [142,143] for more accurate detection of onsets and dicrotic notches of PPG waveforms, as it can make these PPG fiducial points more prominent. A demonstration of applying a 4-th order Butterworth bandpass filter and a 4-th order Chebyshev Type II bandpass filter to PPG waveform can be found in Figure 5.9. A low-pass filter with a cut-off frequency of 12 Hz was applied to the ABP waveform to eliminate high-frequency components. A 2-nd order Butterworth bandpass filter with low cut-off frequency of 0.5 Hz and high cut-off frequency of 30 Hz was applied for ECG waveform.

While the application of distinct filter architectures for ECG and PPG signals introduces a theoretical differential group delay, empirical validation was conducted to quantify the impact on temporal features like PAT. At the 1 kHz sampling frequency, the measured inter-signal temporal offset was minimal and confirmed to have no statistically important effect on extracted features or final estimation performance (see Appendix A.6 for details).

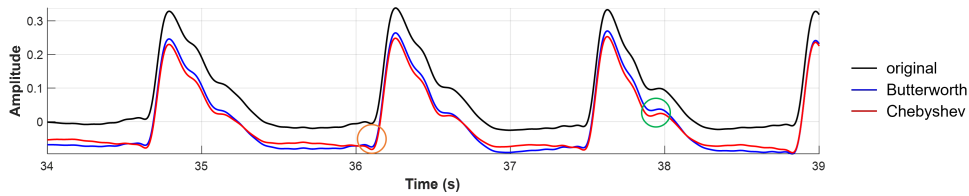


Figure 5.9: A demonstration of applying both 4-th order Butterworth bandpass filter and Chebyshev Type II bandpass filter with low cut-off frequency of 0.5 Hz and high cut-off frequency of 12 Hz to a 5-second PPG segment. Chebyshev bandpass filter can make onset and dicrotic notch of PPG more prominent.

### 5.4.3 Fiducial Points Detection

The fiducial points of ECG and PPG are detected for feature extraction, and fiducial points of ABP waveforms are detected to obtain reference SBP and DBP values.

### 5.4.3.1 ECG Fiducial Point

The ECG R peaks are detected and used for PAT calculation. The waveform of the ECG signal varies between subjects. The QRS complex is obvious and ECG R peak can be easily located for some subjects like shown in Figure 5.10(a). However, it can be challenge to locate the R peaks using a simple peak detection method which detects local maxima for some subjects as the amplitude of their ECG T wave can be higher than R wave like shown in Figure 5.10(b). In this case, an improved ECG R peak detection algorithm is developed and the procedures are explained below.

1. The high frequency components of the ECG signal are extracted by applying a Butterworth bandpass filter with the low cut-off frequency of 10Hz and high cut-off frequency of 25Hz to the original ECG signal.
2. The upper envelope of the extracted high frequency components is then detected. For example, the ECG high frequency components (blue curve) and the upper envelope (orange curve) are shown in Figure 5.11(b).
3. The peaks of the extracted upper envelope are detected by locating the local maxima, then the exact ECG R peaks are corrected by detecting the local maxima from original ECG waveform within a range of  $\pm 20$  samples around envelope peaks. A demonstration of upper envelope peak detection and ECG R peak correction is shown in Figure 5.11(c).

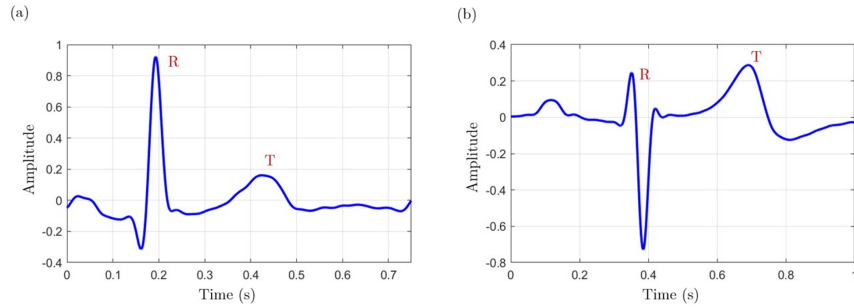


Figure 5.10: Examples of ECG pulses, (a) ECG waveform with a distinct QRS complex, (b) ECG waveform with a minor R wave.

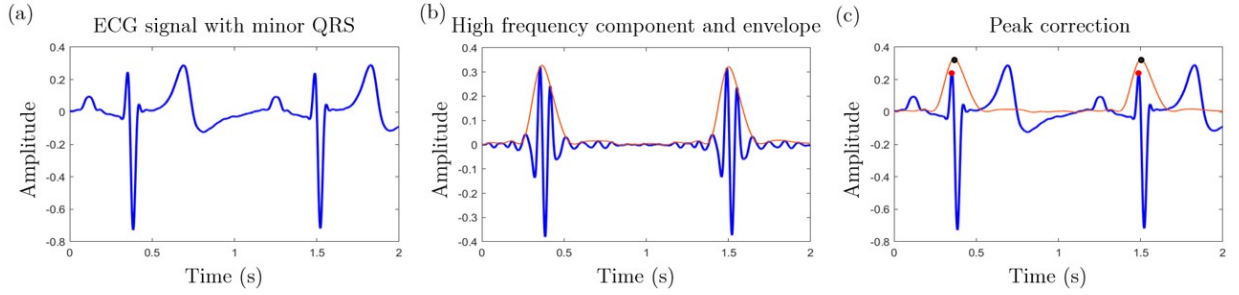


Figure 5.11: The demonstration of the proposed ECG R peak detection method. (a) an original ECG waveform with minor QRS complex, (b) the high frequency components (10-25 Hz) of the ECG signal (blue curve) and its upper envelope (orange curve), (c) the upper envelope (orange curve) with its peaks (black dots) and the corrected ECG R peaks (red dots).

#### 5.4.3.2 PPG Fiducial Points

The fiducial points (onset, systolic peak, diastolic notch and diastolic peak) of PPG, fiducial points ( $u$ ,  $v$  and  $w$  points) of the first derivative of PPG (VPG) and fiducial points ( $a$ ,  $b$ ,  $c$ ,  $d$ ,  $e$  and  $f$  points) of the second derivative of PPG (APG) are detected. These fiducial points are demonstrated in Figure 5.1.

Several different PPG fiducial points detection methods proposed, and Python libraries and MATLAB toolboxes were developed [145, 204, 205]. Even though they claimed their methods can accurately detect these fiducial points for PPG waveforms with different morphological characteristics and signal qualities. We observe their detection results are distinct, meaning the morphological features extracted from PPG and PPG derivatives may be different by using different methods. A demonstration of PPG fiducial points detection using three proposed methods—*PulseAnalyse* [205], *PPGFeat* [204] and *pyPPG* [145] is shown in Figure 5.12.

In this study, these PPG fiducial points are detected using an adjusted and improved method based on *pyPPG* algorithm [145]. The definitions of these fiducial points, as well as steps and methods to detect them from PPG and its derivatives are described in Table 5.1.

#### 5.4.3.3 ABP Fiducial Points

Similar to PPG, the systolic peak and pulse onset of ABP waveform are detected to extract SBP and DBP values.

#### 5.4.4 Feature Extraction and Processing

After detecting fiducial points, a total of 69 morphological features (see Table 5.2 for details) are extracted from ECG, PPG and derivatives of PPG (VPG and APG), as well as simultaneous SBP and DBP are detected.

These extracted beat-to-beat features and simultaneous BPs are then resampled to a fixed sampling rate of 2 Hz to form uniformly sampled time series, and the outliers caused by inaccurate fiducial point detection especially for features extracted from VPG and APG are detected using the three-sigma rule and then replaced by piecewise cubic hermite interpolating polynomial ('pchip') interpolation method. Then the variations modulated by respiration are removed from both features and BPs using the method proposed in Chapter 3 before using them for model training and evaluation.

#### 5.4.5 Implementation Details

To evaluate the model generalization and performance of BP trend tracking, two proposed models (described in Section 5.2) as well as three SOTA models, CNN-LSTM [93] (sig2bp), HGCTNet [111] (feat&sig2bp) and UNet [95] (sig2sig) are implemented. As these models predict BPs using different strategies, different data preparation methods, as demonstrated in Figure 5.13, are implemented to ensure fair comparisons. The implementation details of each model are described below:

- **TCN-SENet-KAN:** the proposed model structure is demonstrated in Figure 5.3. The kernel size, dilation factor for two stacked TCN blocks are 3 and 1 respectively. The number of filters for TCN block1 and TCN block2 is 32 and 64 respectively. ReLU activation and a dropout layer (dropout rate=0.2) are included in TCN blocks. The reduction ratio of SENet is 16. For KAN layer, SiLU and the third order spline (see Equation (5.8)) are employed for fixed based activation function and spline component respectively. Other parameters are grid size = 5, grid range = [-1,1], grid epsilon = 0.02 , noise scale = 0.1 for KAN regression layer. The last step of the SENet output is used as the input of KAN layer with output size of 1 for either SBP

or DBP prediction. 11 important morphological features  $PAT_{ms}$ ,  $AUC_{pi}$ ,  $AUC_{sys}$ ,  $A_{dp}/A_{sp}$ ,  $(A_{sp} - A_{dp})/A_{sp}$ ,  $A_{dn}/A_{sp}$ ,  $PIR$ ,  $T_{sp}/A_{sp}$ ,  $A_{sp}/\Delta T$ ,  $A_{sp}/(T_{pi} - T_{sp})$  and  $T_{dp}/A_{dp}$  (see Section 5.1 and Table 5.2 for details) which exhibit strong correlation to BP changes are selected for BP estimations, these features are windowed into segments of 4 seconds with 2 seconds overlap, resulting in 8 time steps of 11 features within each segment (as extracted features are resampled to a fixed rate of 2 Hz). The reference SBP and DBP values of the last time step of each segment are extracted for model training and evaluation. The data segmentation process is demonstrated in Figure 5.13 (a). When training this model, batch size is set to 32 and Adam optimizer with learning rate of 1e-4 is applied. Both proposed shape-oriented loss function  $sDTW_{loss}$  and MAE is used as the loss function. Additionally, to demonstrate the effect of KAN layer, it is replaced by a regular FC layer with linear activation function to predict SBP or DBP and this model is named as **TCN-SENet-Linear**.

- **A parallel model of Transformer-BiGRU with Attention Mechanisms:** the proposed model **Transformer-BiGRU-Attention** is demonstrated in Figure 5.5. For transformer branch, the expected input feature size is 32, number of heads is 2 for MHSA block and model dimension is 128. For BiGRU-Global attention branch, two BiGRU layers with filter number of 32 and 64 are used. The attention dimension for global attention mechanism is 128. The FC layer with input dimension of 32 and output dimension of 1 is added to predict SBP and DBP. Identical to TCN-SENet-KAN, the model estimate the SBP or DBP for the last time step from 4 seconds' segment of selected 11 importance features. When training this model, batch size is set to 32 and Adam optimizer with learning rate of 1e-4 is employed. MAE is used as the loss function.
- **CNN-LSTM (sig2bp):** as proposed in the original work [93], the difference of ECG and PPG waveforms (ECG-PPG) is used as the model input to predict both SBP and DBP. The data segmentation process is demonstrated in Figure 5.13 (b). The ECG and PPG waveforms of the original **BCG BP** are downsampled to 125 Hz, and a segment of 2.048 seconds' (256 samples) waveforms which is centered at BP extraction point is extracted as the model input. This is because the original model uses a global average pooling layer between the CNN-LSTM module output and FC layer for BP estimations. When training this model, batch size is set to 32 and Adam optimizer with learning rate of 1e-4 is employed. MAE is used as the loss function.
- **HGCTNet (feat&sig2bp):** similar to CNN-LSTM model, the PPG, VPG and APG waveforms of the original **BCG BP** are downsampled to 125 Hz. A 2.048 seconds' (256 samples) segment of PPG, VPG and APG waveforms which is centered

at BP extraction point is extracted, together with 26 important features (selected using Relief method) at the BP extraction point are used as the model input to predict SBP and DBP respectively. The data segmentation process is demonstrated in Figure 5.13 (c). When training this model, batch size is set to 32 and Adam optimizer with learning rate of 1e-3 is employed. MAE is used as the loss function.

- **UNet (sig2sig):** the PPG, VPG and APG waveforms of the original **BCG BP** are downsampled to 250 Hz. A segment of 4.096 seconds' (1024 samples) segment of PPG, VPG and APG waveforms which is centered at BP estimation point is extracted (as shown in Figure 5.13 (b)). The mean SBP and DBP of each predicted ABP segment are then extracted and calculated as the segment estimation results. When training this model, batch size is set to 32 and Adam optimizer with learning rate of 1e-3 is employed. MAE is used as the loss function.

For each of the subject from **BCG BP** dataset, the inputs and outputs are prepared for different models using aforementioned methods. The inputs and outputs of a same subject are then split into batches (batch size=32) according to chronological order which ensures the proposed sDTW loss function compares the reference and predicted BP trends correctly and facilitates the estimation model extracting shape-oriented features. The BP estimation models were trained for a maximum of 100 epochs. Early stopping is used as a regularization technique which monitors the validation loss with a patience setting of 10 epochs to prevent overfitting.

#### 5.4.6 Leave-One-Subject-Out Cross Validation

To evaluate the model generalization and BP trend tracking performance of the proposed methods described in this chapter, Leave-one-subject-out (LOSO) cross validation is employed. As the **BCG BP** dataset contains 32 subjects, 32 separate folds are created. In each fold, one subject is designated as the testing subject, while the remaining 31 subjects are used as training subjects. For each training subject, the first 80% of the batches are used for model training and the remaining 20% batches are used for validation. Consequently, 32 distinct models are trained and evaluated, with each model tested on the data from the testing subject which is completely unseen during model training stage.

## 5.5 Results

The mean and standard deviation of the [Absolute Error \(AE\)](#) between the reference and estimated BPs, as well as the median and interquartile range (IQR) of Pearson correlation coefficient ( $r$ ) which can implicitly quantify the similarity between the reference and estimated BP trends using different estimation models for all 32 subjects of **BCG BP** dataset using LOSO cross-validation are calculated in Table 5.3. Since the customized sDTW loss function only evaluates the shape similarity between the reference and estimated BP trends after zero-mean normalization, distance-based metrics are not meaningful for assessing the performance of TCN-SENet-Linear and TCN-SENet-KAN under the sDTW<sub>loss</sub> scenario.

Table 5.3: The mean and standard deviation of absolute error between reference and estimated BPs for all subjects, and the median (IQR) of the Pearson correlation coefficient between each subject’s reference and estimated BPs using different models and loss functions.

| Model                         | # params | Loss function        | Grand AE (mean±std) |              | $r$ (median, IQR)  |                    |
|-------------------------------|----------|----------------------|---------------------|--------------|--------------------|--------------------|
|                               |          |                      | SBP                 | DBP          | SBP                | DBP                |
| <i>SOTA models</i>            |          |                      |                     |              |                    |                    |
| CNN-LSTM                      | 181.2K   | MAE                  | 11.27 ± 10.46       | 7.72 ± 5.97  | 0.002 (0.07)       | 0.001 (0.07)       |
| HGCTNet                       | 2.4M     | MAE                  | 13.81 ± 11.70       | 10.41 ± 8.16 | 0.03 (0.12)        | 0.04 (0.10)        |
| UNet                          | 10.5M    | MAE                  | 12.68 ± 10.17       | 7.70 ± 6.67  | 0.40 (0.54)        | 0.09 (0.41)        |
| <i>Proposed by this study</i> |          |                      |                     |              |                    |                    |
| Transformer-BiGRU-Attention   | 126.7K   | MAE                  | 15.80 ± 12.82       | 8.88 ± 7.85  | 0.25 (0.76)        | 0.28 (0.69)        |
| TCN-SENet-Linear              | 25.9K    | MAE                  | 12.84 ± 11.66       | 9.07 ± 7.76  | 0.57 (0.95)        | 0.40 (0.66)        |
| TCN-SENet-Linear              | 25.9K    | sDTW <sub>loss</sub> | –                   | –            | 0.74 (0.39)        | 0.44 (0.53)        |
| TCN-SENet-KAN                 | 26.5K    | MAE                  | 19.47 ± 14.25       | 12.84 ± 8.24 | 0.48 (0.89)        | −0.04 (0.54)       |
| TCN-SENet-KAN                 | 26.5K    | sDTW <sub>loss</sub> | –                   | –            | <b>0.74</b> (0.42) | <b>0.56</b> (0.47) |

The results in Table 5.3 demonstrate that the proposed models significantly outperform existing SOTA models in tracking BP trends, particularly under a LOSO cross-validation evaluation setting. This setting reflects a realistic and challenging scenario for assessing model generalization. The SOTA models (CNN-LSTM, HGCTNet, UNet) show extremely low correlation, with correlation coefficients almost equal to zero, indicating poor trend tracking capability and suggesting that these models fail to learn meaningful temporal or shape-related information from the model inputs. In contrast, the proposed models, especially the models trained with the shape-oriented sDTW loss function, achieve substantially higher correlation values (e.g., up to 0.74 for SBP and 0.56 for DBP). This improvement highlights the benefit of optimizing for shape similarity, which encourages the model to align better with the overall BP trend. Furthermore, the application of the KAN layer, designed to enhance non-linearity modeling, consistently outperforms models with conventional fully connected (FC) layers. For example, the median correlation coefficients

for both SBP and DBP using TCN-SENet-KAN model are greater than the model with conventional FC layer.

Figures 5.14–5.19 demonstrate examples of SBP estimations from three subjects using different models and loss functions. All estimations were produced under a LOSO cross-validation setup. To allow for clearer visual comparison, each estimated SBP was adjusted to match the mean level of the reference. As shown in these examples, HGCTNet produces noisy predictions, while the CNN-LSTM model fails to capture BP fluctuations, instead generating flat, average-like estimations. Although the proposed Transformer-BiGRU-Attention and TCN-SENet-Linear models apply time series regression, they still cannot track BP trends effectively. This is mainly due to the MAE loss, which focuses on pointwise accuracy rather than shape similarity. In contrast, the TCN-SENet-KAN model trained with the shape-aware sDTW loss produces estimations that highly similar to the actual BP trends, despite having no access to the test subject’s data. These results highlight that the combination of the KAN layer and sDTW loss enables the model to capture trend-related patterns across individuals, leading to improved generalization in BP trend tracking and facilitating easier personalization for new subjects.

## 5.6 Conclusion and Discussion

In conclusion, this chapter demonstrates that the proposed TCN-SENet-KAN model effectively addresses the challenge of BP trend tracking, a critical yet underexplored aspect of cuffless BP estimation. Despite its lightweight architecture which has significantly fewer trainable model parameters compared to existing SOTA models and a parallel Transformer-BiGRU-Attention model, TCN-SENet-KAN achieves superior performance in tracking BP trends under LOSO cross-validation. This highlights its strong generalization capability across unseen subjects. A key innovation contributing to this improvement is the adoption of the shape-oriented sDTW loss function, which changes the model’s focus from pointwise accuracy to overall trend similarity. This not only enhances the model’s ability to track BP changes but also facilitates more personalized and robust long-term BP monitoring. Overall, the findings emphasize the importance of both model design and task-specific loss functions in improving the accuracy and generalization of cuffless BP estimation models.

The personalization stage which uses very few calibration measures from the new subject for subject-specific adaptation methods will be introduced in the next chapter.



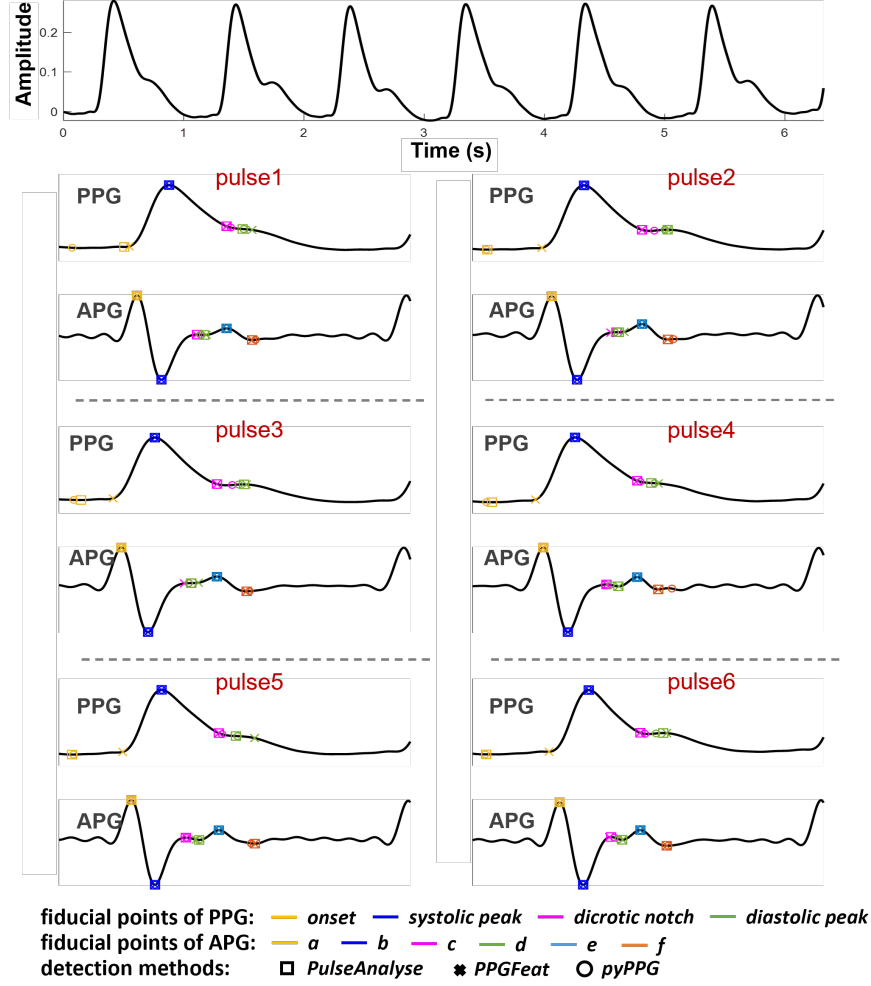


Figure 5.12: Examples of detecting fiducial points from PPG waveform and the second derivative of PPG (APG) using methods proposed by *PulseAnalyse*, *PPGFeat* and *pyPPG*, resulting in different detection results (data extracted from **BCG BP** dataset, sampling frequency is 1 KHz).

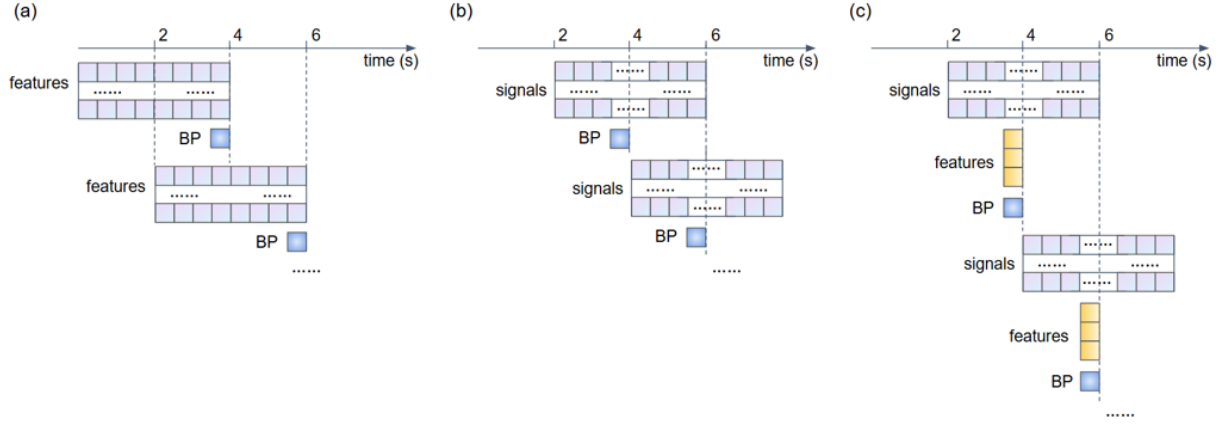


Figure 5.13: Data preparation strategies for model training, (a) strategy for time series regression models, TCN-SENet-KAN and a parallel transformer-BiGRU with attention mechanisms which takes 4 seconds' of physiological features to predict BPs, (b) strategy for sig2bp model-CNN-LSTM which uses 2.048 seconds' waveform segments to predict BPs, (c) strategy for a hybrid model-HTCTNet which uses 4.096 seconds' waveform segments as well as physiological features to predict BPs.

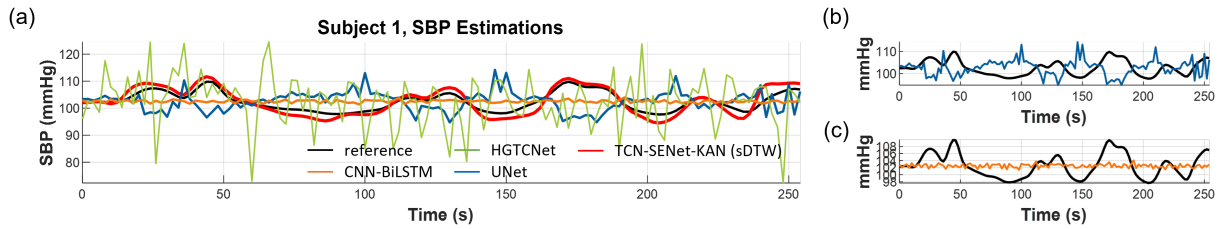


Figure 5.14: Reference and estimated SBP of Subject 1 using (a) SOTA models and TCN-SENet-KAN with sDTW loss function, (b) UNet model, (c) CNN-LSTM model.

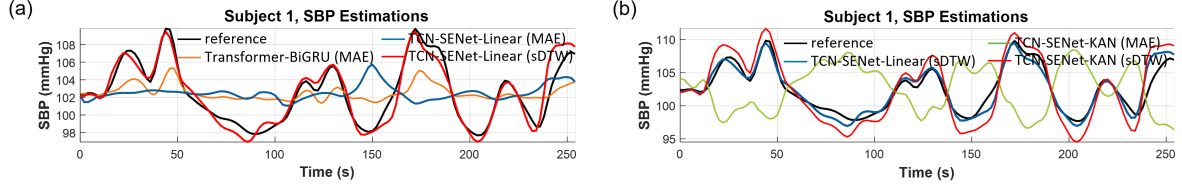


Figure 5.15: Reference and estimated SBP of Subject 1 using (a) proposed Transformer-BiGRU attention, TCN-SENet-Linear model using MAE loss function, (b) proposed TCN-SENet-Linear and TCN-SENet-KAN model with MAE and sDTW loss functions.

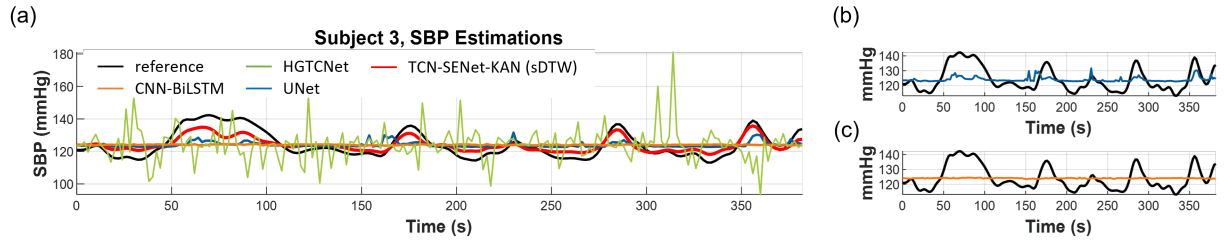


Figure 5.16: Reference and estimated SBP of Subject 3 using (a) SOTA models and TCN-SENet-KAN with sDTW loss function, (b) UNet model, (c) CNN-LSTM model.

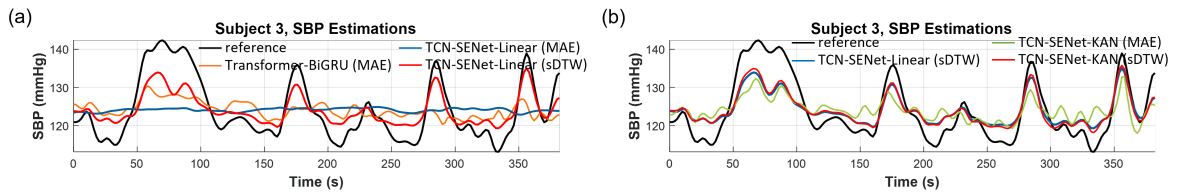


Figure 5.17: Reference and estimated SBP of Subject 3 using (a) proposed Transformer-BiGRU attention, TCN-SENet-Linear model using MAE loss function, (b) proposed TCN-SENet-Linear and TCN-SENet-KAN model with MAE and sDTW loss functions.

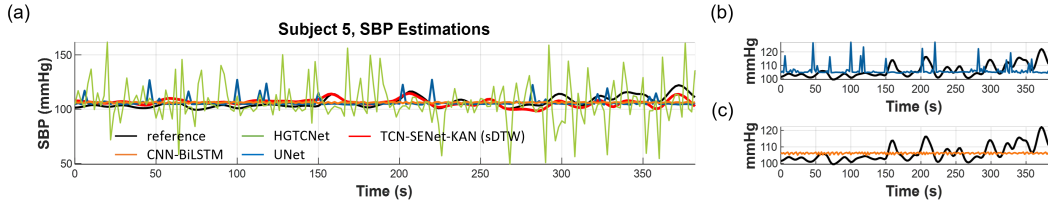


Figure 5.18: Reference and estimated SBP of Subject 5 using (a) SOTA models and TCN-SENet-KAN with sDTW loss function, (b) UNet model, (c) CNN-LSTM model.

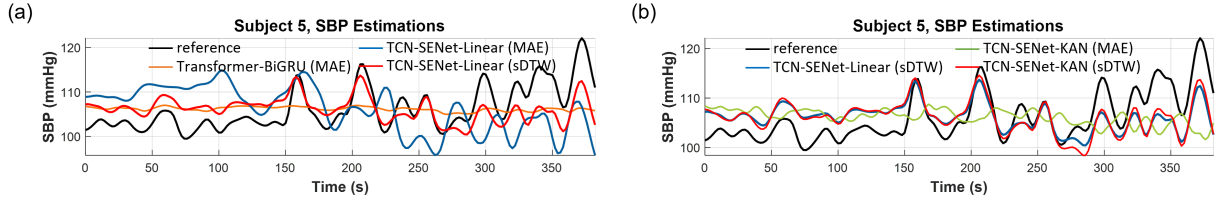


Figure 5.19: Reference and estimated SBP of Subject 5 using (a) proposed Transformer-BiGRU attention, TCN-SENet-Linear model using MAE loss function, (b) proposed TCN-SENet-Linear and TCN-SENet-KAN model with MAE and sDTW loss functions.

## Chapter 6

# Accurate BP Estimation with Few Calibrations

In this Chapter, the methods for accurately estimating BPs for a new unseen subject with only few calibrations are described. First, a Bayesian hierarchical regression based approach [206] which was proposed previously is introduced. Then, a number of new methods with different strategies, including commonly used model fine-tuning, few-shot learning methods such as meta-learning and domain adaptation, as well as simple statistic transformation methods are applied to the TCN-SENet-KAN which is proposed in the previous chapter (Chapter 5) with great model generalization and accurate BP trend tracking to quickly adapt the pre-trained models to new subjects.

This chapter presents methods for accurately estimating BP for new, unseen subjects using only a limited number of calibration samples. It begins by introducing a previously proposed approach based on Bayesian hierarchical regression. Subsequently, several new strategies are explored to enable quick adaptation of pre-trained models to individual subjects. These include conventional model fine-tuning, few-shot learning techniques such as meta-learning and domain adaptation, as well as simple statistical transformation methods. All approaches are applied to the TCN-SENet-KAN model which demonstrated strong generalization ability and accurate BP trend tracking in the previous chapter, making it a promising method for subject-specific adaptation.

## 6.1 A Bayesian Hierarchical Regression Model for Few-Shot BP Estimation

**Publication:** a version of the method described in this section was previously published by IEEE EMBC 2021 [206].

A Bayesian hierarchical regression based few-shot BP estimation model was proposed so that the BP can be accurately estimated for a new patient with just few measurements and avoid overfitting.

### 6.1.1 Theoretical Introduction

In Bayesian hierarchical regression modelling, the regression coefficients are themselves given a model, whose parameters are also estimated from data. Compared with classical regression, hierarchical modelling is always an improvement, as shown for example in the prediction of home radon levels in U.S counties [207]. In this section, the BP will be estimated using a one-level hierarchical model by optimizing the parameters of the non-linear relationship (equation (2.3),  $BP = \alpha + \beta \cdot \ln(PAT)$ ) between PAT and BP at both individual level (each subject) and group level (between subjects), and the comparison between this approach and classical regression methods will be demonstrated.

To estimate BP using a Bayesian hierarchical regression model, the equation (2.3) can be algebraically expressed as:

$$BP \sim \mathcal{N}(\mu = \alpha + \beta \cdot \ln(PAT), \epsilon) \quad (6.1)$$

That is, the data vector BP is assumed to be distributed as a normal distribution with a mean of  $\alpha + \beta \cdot \ln(PAT)$ , and a standard deviation of the measurement noise  $\epsilon$ .

The probabilistic model and the [Ordinary Least Squares \(OLS\)](#) method should have similar performance in optimizing model parameters on a large dataset. However, the parameters  $\alpha$  and  $\beta$  are unbounded when the dataset is extremely small, for example, only one data point but OLS method requires at least two points to fit a line. This extreme problem can be solved by the a hierarchical model and setting strong priors for the unknown parameters  $\alpha$  and  $\beta$ . Another way to convey information is by defining hierarchical models, since hierarchical models allow information to be shared between subjects, shrinking the plausible values of the estimated parameters [208].

In a hierarchical model, the parameters can be estimated directly from the data by placing shared priors over them, instead of fixing the parameters of our priors to some constant numbers. These higher-level priors are known as hyperpriors and their parameters are hyperparameters, and informative priors will be assigned based on the fact that the hyperpriors will shrink to the group mean. The mean and standard deviation of  $\alpha$  and  $\beta$  derived from each subject's PAT and SBP values are  $-1 \pm 30.85$  and  $-116.35 \pm 32.24$  using OLS method. Thus, the framework of the one-level hierarchical model for this study with the definition of the parameters and priors for each level is shown in Figure 6.1(a). The parameters  $\alpha$  and  $\beta$  that mainly determine the distribution of BP have their priors in normal distributions. The hyperparameters  $\mu_\alpha$  and  $\mu_\beta$  have their hyperpriors in normal distributions and the hyperparameters  $\sigma_\alpha$  and  $\sigma_\beta$  have their hyperpriors in half-normal distributions.

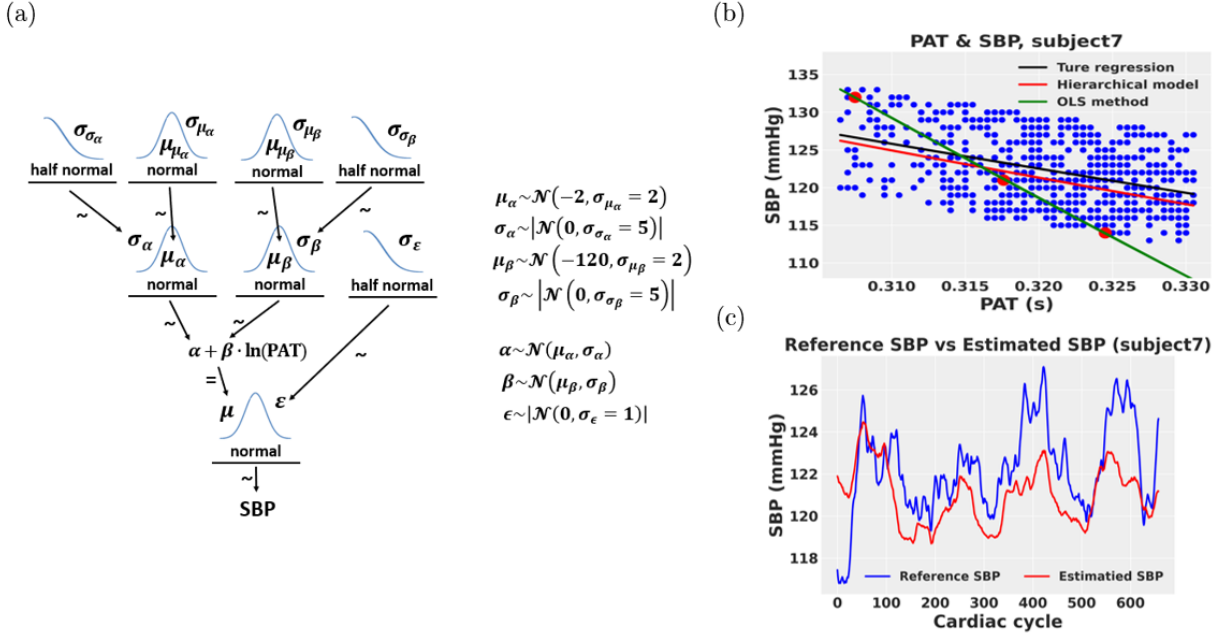


Figure 6.1: (a) The Kruschke diagram of the proposed hierarchical model and the selection of the hyperpriors. (b) A demonstration of SBP estimation for one subject using the proposed hierarchical regression (red), OLS (green) with only 3 calibrations, while the true regression is the black line. (c) The estimated SBP trend (red) using the hierarchical regression and the reference SBP trend (adapted from [206]).

### 6.1.2 Implementation and Results

To demonstrate the feasibility of the proposed hierarchical regression method, a small dataset which is collected by our research team from ten healthy subjects (5 males,  $28.3 \pm 8.3$  years' old). For each subject, 10 minutes' simultaneous ECG, PPG and beat-to-beat reference BP values were measured.  $PAT_{sys}$  which is the time interval between ECG R peaks and PPG systolic peaks (as demonstrated in Figure 2.4) and corresponding SBP values were retrieved. For each subject, the PAT and SBP values of the other nine subjects and  $n$  ( $n=1,3$ ) calibration measurements of PAT and SBP were used to train the hierarchical regression model; the SBPs were then predicted using the trained model and this subject's remaining PATs. The model parameters of interest are estimated using posterior distributions which are determined by analytical integration. The [Markov Chain Monte Carlo \(MCMC\)](#) method was employed for parameter estimation and No U-turn sampler (2000 samples for each channel with 1000 tuning samples) was used to evaluate the posterior distributions. Figure 6.1 (b) illustrates a comparison of SBP estimation for a subject using hierarchical regression versus the OLS approach with only three calibrations. OLS results in overfitting, whereas the regression line derived using hierarchical regression is almost identical to the actual one. The grand mean and standard deviation of [AE](#) between the estimated SBP and reference SBP for these ten subjects using hierarchical regression and OLS method with one and three calibrations can be found in Table 6.1, it indicates that the proposed hierarchical regression can accurately just using few calibration(s) and it is feasible for quick adaption problem within a same dataset.

Table 6.1: The mean and standard deviation of the absolute error between estimated SBP and reference SBP using different approaches (Note: it is impossible to conduct OLS regression one a single measurement).

| # of calibrations | OLS regression  | Hierarchical regression |
|-------------------|-----------------|-------------------------|
| 3                 | $6.96 \pm 3.77$ | $4.35 \pm 1.43$         |
| 1                 | --              | $4.36 \pm 1.17$         |

### 6.1.3 Limitations

Although the proposed Bayesian hierarchical model has demonstrated feasibility in accurately estimating BP for a new subject using a small number of calibration measurements, it still has limitations.



- The proposed hierarchical model is feasible for intra-domain adaptation as the basic assumption is that data of one type come from one distribution. If data are acquired using different devices, the data may not have the same distribution; hence, this model is not appropriate.
- The Bayesian model is computationally intensive; for each new subject, the model must be retrained using the training dataset and calibration measurements. This problem may be more severe as the data increases and more features are added.

Therefore, several different models will be introduced in the following to address the limitation of this Bayesian hierarchical regression based method and improve the performance for estimating BPs for new subjects via very limited amount of calibrations.

## 6.2 Problem Formalization for Few-Shot Adaptation

In this section, the problem of adapting a model which is pre-trained using a group of subjects to a new subject who is completely unseen during model training with very limited number of calibrations is formalized. Since this study uses the **BCG BP** dataset which includes 32 subjects and **LOSO** cross validation is applied, the problem can be described as:

Consider a dataset  $D$  comprising data from  $N$  ( $N = 32$  for **BCG BP** dataset used in this study) distinct subjects. Let  $S_i$  denote the  $i$ -th subject, for  $i = 1, 2, \dots, 32$ . Each subject  $S_i$  has a collection of input-output pairs  $D_i = \{(\mathbf{x}_{i,m}, y_{i,m})\}_{m=1}^{M_i}$ , where  $\mathbf{x}_{i,m} \in \mathbb{R}^d$  represents a  $d$ -dimensional input vector which is pre-determined features extracted from ECG and PPG waveforms, and  $y_{i,m} \in \mathbb{R}$  corresponds to continuous scalar output which is SBP or DBP values, indicative a regression problem. The total number of samples for subject  $S_i$  is  $M_i$ . In each fold of the LOSO cross validation, one subject  $S_{\text{test}}$  is held out as the test subject, while the remaining  $N - 1$  (which is 31 subjects in this study) subjects form the training subjects. The objective is to train a robust and generalized model with great model initialization that can quickly adapt to the held-out subject  $S_{\text{test}}$  via a small calibration set  $D_{\text{test}}^S = \{(\mathbf{x}_{\text{test},l}, y_{\text{test},l})\}_{l=1}^k$  of  $k$  samples. Finally, the adapted model is used to predict for the remaining data (query set)  $D_{\text{test}}^Q = \{(\mathbf{x}_{\text{test},m}, y_{\text{test},m})\}_{m=1}^{M_Q}$  ( $M_Q$  is the number of samples for prediction) of the test subject for performance evaluation.

## 6.3 Model Fine-Tuning for Subject-Specific Adaptation

Fine tuning is a common and straightforward approach to adapt a pre-trained model to a new subject. This method leverages the knowledge encoded in the parameters of a model trained on a large dataset and then refines these parameters using a small amount of data from the target subject. This section describes the process of adapting a pre-trained model to the held-out test subject using their limited adaptation data.

**Overview:** a model  $f_{\theta^*}$  is pre-trained on the data from  $N-1$  (31 in this study) subjects. This pre-training is accomplished by minimizing the regression loss (e.g., Mean Squared Error) over all available samples from these  $N-1$  subjects. The objective of this pre-training is to learn a set of parameters  $\theta^*$  that capture general features and relationships relevant to the regression task across a diverse population. Then this pre-trained model  $f_{\theta^*}$  is personalized for the held-out test subject  $S_{\text{test}}$  via a small adaptation set  $D_{\text{test}}^S = \{(\mathbf{x}_{\text{test},l}, y_{\text{test},l})\}_{l=1}^k$  of  $k$  samples.

### 6.3.1 Adaptation via Fine-Tuning

For the test subject  $S_{\text{test}}$ , a small adaptation set (or support set)  $D_{\text{test}}^S = \{(\mathbf{x}_{\text{test},l}, y_{\text{test},l})\}_{l=1}^k$  which consists of  $k$  labeled input-output pairs specific to subject  $S_{\text{test}}$  is utilized. Fine-tuning involves continuing the training process of the pre-trained model  $f_{\theta^*}$ , but exclusively using this small, subject-specific adaptation set  $D_{\text{test}}^S$ .

The objective of fine-tuning is to adjust the model parameters  $\theta^*$  to better fit the characteristics of  $S_{\text{test}}$ , thereby updating parameters  $\theta'_{\text{test}}$ . This is achieved by minimizing the regression loss  $\mathcal{L}$  (e.g., Mean Squared Error) on the adaptation set  $D_{\text{test}}^S$  and the optimization process starts with  $\theta = \theta^*$ :

$$\theta'_{\text{test}} = \arg \min_{\theta} \sum_{l=1}^k \mathcal{L}(f_{\theta}(\mathbf{x}_{\text{test},l}), y_{\text{test},l}) \quad (6.2)$$

Several strategies are commonly employed during fine-tuning, such as full fine-tuning, partial fine-tuning, using a smaller learning rate and number of iterations (see Appendix A.7 for details). Once the fine-tuning process is complete, the adapted model  $f_{\theta'_{\text{test}}}$  is used to predict for the remaining test samples (query set) of the test subject  $S_{\text{test}}$ :  $y_{\text{test}}^{\text{pred}} = f_{\theta'_{\text{test}}}(\mathbf{x}_{\text{test},m})$ .

## 6.4 Meta-Learning for Subject-Specific Adaptation

**Overview:** within the meta-learning framework, each subject  $S_j$  of the  $N - 1$  meta-training subjects is considered as a distinct "**task**", denoted  $\mathcal{T}_j$ . The objective is to learn a model  $f_\theta$  parameterized by  $\theta$  such that it can generalize effectively on a new, unseen task which is the held-out subject  $S_{\text{test}}$  with minimal, subject-specific data. This adaptation is performed in a  $k$ -shot learning setting, where  $k > 0$  and  $k$  is a small integer. For any given task  $\mathcal{T}_j$  (either during meta-training or meta-testing on  $S_{\text{test}}$ ), two disjoint subsets of its data  $D_j$  are defined as:

1. **Support Set ( $D_j^S$ ):** a small set of  $k$  input-output pairs,  $D_j^S = \{(\mathbf{x}_{j,l}, y_{j,l})\}_{l=1}^k$ , used to adapt the general model parameters  $\theta$  to become task-specific parameter  $\theta'_j$ .
2. **Query Set ( $D_j^Q$ ):** the remaining samples from  $D_j$ ,  $D_j^Q = D_j \setminus D_j^S$ , are used to evaluate the performance of the adapted model  $f_{\theta'_j}$  on task  $\mathcal{T}_j$ .

The ultimate goal is to find initial set of parameters  $\theta^*$  that generates accurate predictions on its query set  $D_{\text{test}}^Q$ , after adaptation using only the  $k$  samples in the support set of the held-out subject  $S_{\text{test}}$ . This requires the meta-learning process to capture underlying commonalities across the 31 training subjects that are transferable to new subject.

**Model-Agnostic Meta-Learning (MAML)** [209] is a widely recognized meta-learning algorithm designed to optimize an initial set of model parameters that are well-suited for quick adaptation to new tasks with only a few samples. Its "model-agnostic" nature implies that it can be applied to any model trained with gradient descent, making it a versatile choice for various regression architectures. The **Model-Agnostic Meta-Learning (MAML)** algorithm operates via a bi-level optimization process, comprising an **inner loop** for task-specific adaptation and an **outer loop** for meta-optimization across tasks.

### 6.4.1 Inner Loop: Task-Specific Adaptation

During meta-training, for each task  $\mathcal{T}_j$  sampled from the set of  $N - 1$  training subjects, the goal of the inner loop is to quickly adapt the shared model parameters  $\theta$  to task-specific parameters  $\theta'_j$ . This adaptation is achieved by taking one or a few gradient descent steps with respect to the loss on the support set  $D_j^S$  for that specific task.

Let  $\mathcal{L}_{\mathcal{T}_j}(f_\theta, D_j^S)$  be the loss function for task  $\mathcal{T}_j$  evaluated on its support set  $D_j^S$  using the model  $f_\theta$ . For a regression problem, Euclidean metrics such as **MSE** can be used as this loss function:

$$\mathcal{L}_{\mathcal{T}_j}(f_\theta, D_j^S) = \frac{1}{k} \sum_{l=1}^k (f_\theta(\mathbf{x}_{j,l}) - y_{j,l})^2 \quad (6.3)$$

Starting from the initial parameters  $\theta$ , the adapted parameters  $\theta'_j$  for task  $\mathcal{T}_j$  after one gradient step are computed as:

$$\theta'_j = \theta - \alpha \nabla_\theta \mathcal{L}_{\mathcal{T}_j}(f_\theta, D_j^S) \quad (6.4)$$

where  $\alpha$  is the learning rate for the inner loop. If multiple adaptation steps are used in the inner loop, this process will then be repeated. It is worth noting that the  $\theta'_j$  is directly derived from  $\theta$  and is specific to task  $\mathcal{T}_j$  during inner loop optimization stage [210].

## 6.4.2 Outer Loop: Meta Optimization

The objective of the outer loop is to update the initial model parameters  $\theta$  such that the adapted parameters  $\theta'_j$  perform well on the query sets  $D_j^Q$  of their respective tasks. Essentially, the meta-learner aims to find an initialization  $\theta$  that is highly sensitive to task-specific data which will potentially improve model performance with minimal inner loop adaptation.

The meta-objective function  $J(\theta)$  is the sum (or average) of the losses on the query sets of the meta-training tasks which is evaluated using the respective adapted parameters  $\theta'_j$ :

$$J(\theta) = \sum_{\mathcal{T}_j \sim p(\mathcal{T}_{\text{meta-train}})} \mathcal{L}_{\mathcal{T}_j}(f_{\theta'_j}, D_j^Q) \quad (6.5)$$

where  $p(\mathcal{T}_{\text{meta-train}})$  represents the distribution of meta-training tasks. In the LOSO setup of this study, this involves iterating through the 31 training subjects. The loss  $\mathcal{L}_{\mathcal{T}_j}(f_{\theta'_j}, D_j^Q)$  is calculated using the query data and the adapted parameters:

$$\mathcal{L}_{\mathcal{T}_j}(f_{\theta'_j}, D_j^Q) = \frac{1}{|D_j^Q|} \sum_{(\mathbf{x}, y) \in D_j^Q} \left( f_{\theta'_j}(\mathbf{x}) - y \right)^2 \quad (6.6)$$

The meta-parameters  $\theta$  are then updated by taking a gradient step with respect to this meta-objective:

$$\theta \leftarrow \theta - \beta \nabla_{\theta} \sum_{\mathcal{T}_j \sim p(\mathcal{T}_{\text{meta-train}})} \mathcal{L}_{\mathcal{T}_j}(f_{\theta'_j}, D_j^Q) \quad (6.7)$$

where  $\beta$  is the meta-learning rate (outer loop learning rate). Notably, the computation of  $\nabla_{\theta} \mathcal{L}_{\mathcal{T}_j}(f_{\theta'_j}, D_j^Q)$  requires differentiation through the inner loop update rule, which involves second-order derivatives (as  $\theta'_j$  is a function of  $\theta$ ). This process can be computational intensive, some novel methods such as FOMAML [211] (First-Order MAML) are proposed and developed to approximate this gradient.

### 6.4.3 Meta-Testing Stage

Once the meta-training stage is complete and an optimal  $\theta^*$  is obtained, the model is then evaluated on the held-out test subject  $S_{\text{test}}$ . Given  $k$  samples in its support set  $D_{\text{test}}^S$ , the parameters  $\theta^*$  are adapted using the same inner loop update rule:

$$\theta'_{\text{test}} = \theta^* - \nabla_{\theta^*} \mathcal{L}_{\mathcal{T}_{\text{test}}}(f_{\theta^*}, D_{\text{test}}^S) \quad (6.8)$$

The performance is then evaluated on the query set  $D_{\text{test}}^Q$  of the test subject using the adapted parameters  $\theta'_{\text{test}}$ . This process is repeated for each subject when it serves as the held-out test subject, providing a robust evaluation of the meta-learned model's ability to generalize and adapt to new individuals via very limited calibration measures.

## 6.5 Domain Adaptation for Subject-Specific Adaptation

An alternative to meta-learning approaches like MAML is to explicitly train a feature representation that is invariant across the training subjects. This section describe the use of domain adaptation methods, Domain Adversarial Neural Network (DANN) specifically, to learn subject-invariant features from the  $N - 1$  training subjects, followed by a  $k$ -shot adaptation for the held-out test subject. Unlike MAML's episodic training, DANN aims to utilize all available data from the  $N - 1$  training subjects to learn the invariant features. The final adaptation on  $S_{\text{test}}$  will then use a small support set  $D_{\text{test}}^S = \{(\mathbf{x}_{\text{test},l}, y_{\text{test},l})\}_{l=1}^k$  of  $k$  samples, then the performance is assessed on a query set  $D_{\text{test}}^Q = D_{\text{test}} \setminus D_{\text{test}}^S$ .

The DANN method aims to learn features that are both discriminative for the main task (regression for this study) and indiscriminate with respect to the "domain" from which the data originates. In the LOSO cross validation context of this study, each of the  $N - 1$

training subjects is treated as a distinct domain. The DANN architecture typically consists of three main components described below and demonstrated in Figure 6.2:

1. **Feature Extractor** ( $G_f$ ): this network which is parameterized by  $\theta_f$ , maps the input data  $\mathbf{x}$  to a feature representation  $f = G_f(\mathbf{x}; \theta_f)$ . The goal is to make these features domain-invariant but task relevant.
2. **Target Predictor** ( $G_y$ ): this network which is parameterized by the  $\theta_y$ , takes the extracted features  $f$  and predicts the output for the main regression task, i.e.,  $\hat{y} = G_y(f; \theta_y)$ .
3. **Domain Classifier** ( $G_d$ ): this network which is parameterized by  $\theta_d$ , takes the extracted features  $f$  and attempts to predict the domain (i.e., via subject ID) from which the input data originated, i.e.,  $\hat{d} = G_d(f; \theta_d)$ .

### 6.5.1 Adversarial Training

- The parameters  $\theta_f$  of the feature extractor are trained to maximize the loss of the domain classifier, thereby making the features  $f$  as indistinguishable (domain-invariant) as possible.
- Simultaneously,  $\theta_f$  and  $\theta_y$  are trained to minimize the loss on the main regression task.
- The parameters  $\theta_d$  of the domain classifier are trained to minimize the domain classification loss, i.e., to be as good as possible at identifying the subject from the features.

This adversarial objective is often implemented using a Gradient Reversal Layer (GRL). The GRL is placed between the feature extractor  $G_f$  and the domain classifier  $G_d$ . During the backward pass, it multiplies the gradient flowing back from the domain classifier by  $-\lambda$  (where  $\lambda$  is a positive hyperparameter that balances the trade-off between the two objectives before passing it to the feature extractor).

Let  $D_{\text{train-all}} = \bigcup_{j=1}^{N-1} D_j$  be the combined dataset from the  $N - 1$  training subjects. For each sample  $(\mathbf{x}_{j,m}, y_{j,m})$  from subject  $S_j$  in  $D_{\text{train-all}}$ :

1. **Regression Loss** ( $\mathcal{L}_y$ ): it measures the discrepancy between the predicted output and the true output. For a regression task, MSE can be used:

$$\mathcal{L}_y(G_y(G_f(\mathbf{x}_{j,m}; \theta_f); \theta_y), y_{j,m}) = (G_y(G_f(\mathbf{x}_{j,m}; \theta_f); \theta_y) - y_{j,m})^2$$

2. **Domain Classification Loss ( $\mathcal{L}_d$ )**: it measures the accuracy of the domain classifier in predicting the subject  $S_j$  (represented by a domain label  $d_j$ , e.g., Subject ID) from the features  $G_f(\mathbf{x}_{j,m}; \theta_f)$ . This is a cross-entropy loss as subjects are treated as distinct classes:

$$\mathcal{L}_d(G_d(G_f(\mathbf{x}_{j,m}; \theta_f); \theta_d), d_j) = - \sum_{c=1}^{N-1} \mathbf{1}[d_j = c] \log P(G_d(G_f(\mathbf{x}_{j,m}; \theta_f); \theta_d) = c)$$

where  $\mathbf{1}[\cdot]$  is the indicator function and  $P(\cdot)$  is the probability determined by the domain classifier for class  $c$ .

Therefore, the overall objective function optimized during the DANN training stage is:

$$\begin{aligned} \min_{\theta_f, \theta_y} \max_{\theta_d} E(\theta_f, \theta_y, \theta_d) &= \sum_{j=1}^{N-1} \sum_{m=1}^{M_j} \mathcal{L}_y(G_y(G_f(\mathbf{x}_{j,m}; \theta_f); \theta_y), y_{j,m}) \\ &\quad - \lambda \sum_{j=1}^{N-1} \sum_{m=1}^{M_j} \mathcal{L}_d(G_d(G_f(\mathbf{x}_{j,m}; \theta_f); \theta_d), d_j) \end{aligned} \quad (6.9)$$

Due to GRL, this min-max problem can be solved by standard gradient descent methods where  $\theta_f, \theta_y$  are updated to minimize the combined objective and  $\theta_d$  is updated to minimize its own classification loss. The parameter  $\lambda$  controls the influence of the domain adversarial component. The parameters are updated as follows:

$$\begin{aligned} \theta_f &\leftarrow \theta_f - \mu \left( \frac{\partial \mathcal{L}_y}{\partial \theta_f} - \lambda \frac{\partial \mathcal{L}_d}{\partial \theta_f} \right) \\ \theta_y &\leftarrow \theta_y - \mu \frac{\partial \mathcal{L}_y}{\partial \theta_y} \\ \theta_d &\leftarrow \theta_d - \mu \frac{\partial \mathcal{L}_d}{\partial \theta_d} \end{aligned}$$

where  $\mu$  is the learning rate. This DANN training uses all available samples from the  $N - 1$  training subjects, contrasting with MAML's task-specific few-shot updates during its meta-training. The rationale is that more data across varied domains (subjects) allows for a more robust estimation of domain-invariant features.

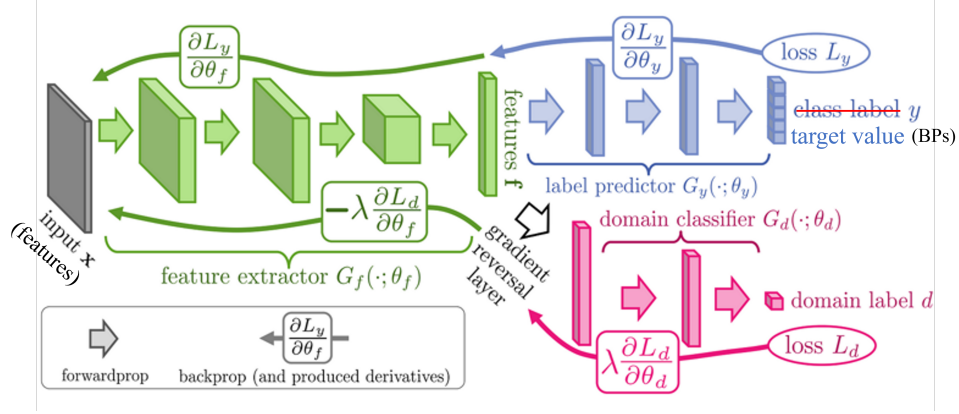


Figure 6.2: A demonstration of key process of domain adversarial neural network (DANN) (figure adapted and modified based on [212]).

### 6.5.2 K-shot Adaptation for the Test Subject

After the DANN model, specifically,  $G_f(\cdot; \theta_f^*)$  and  $G_y(\cdot; \theta_y^*)$  has been trained on the  $N - 1$  subjects, it is then adapted to the held-out test subject  $S_{\text{test}}$ . The domain classifier  $G_d$  is discarded at this stage.

Given the  $k$  support samples  $D_{\text{test}}^S = \{(\mathbf{x}_{\text{test},l}, y_{\text{test},l})\}_{l=1}^k$  from  $S_{\text{test}}$ , the pre-trained feature extractor and target predictor are fine-tuned. This adaptation aims to specialize the model to  $S_{\text{test}}$  using limited data. The fine-tuning objective is to minimize the regression loss on the support set  $D_{\text{test}}^S$ :

$$\theta_{f,\text{test}}'^*, \theta_{y,\text{test}}'^* = \arg \min_{\theta_f, \theta_y} \sum_{l=1}^k \mathcal{L}(G_y(G_f(\mathbf{x}_{\text{test},l}; \theta_f); \theta_y), y_{\text{test},l}) \quad (6.10)$$

The optimization starts from the pre-trained parameters  $\theta_f^*$  and  $\theta_y^*$ . Usually, only the target predictor  $G_y$  or the final layers of  $G_f$  are fine-tuned to prevent overfitting with small  $k$ . The performance of the adapted model  $f_{\theta_{f,\text{test}}'^*, \theta_{y,\text{test}}'^*}$  is then evaluated on the query set  $D_{\text{test}}^Q$  of the test subject.



## 6.6 Statistical Post-Hoc Methods for Subject-Specific Adaptation

Besides fine-tuning or meta-learning strategies, a pre-trained model's predictions for a new subject can be adapted using statistical post-hoc calibration techniques. This approach assumes that a model, pre-trained on data from  $N - 1$  subjects in a standard supervised manner (e.g., by minimizing a standard regression loss over all their data), captures general patterns but may exhibit systematic biases when applied to a new, held-out subject  $S_{\text{test}}$ . The adaptation then aims to correct these biases using a small number of labeled samples from  $S_{\text{test}}$ .

In the LOSO cross validation context, a model  $f_{\theta^*}$  is pre-trained using the combined data from  $N - 1$  subjects. This pre-training typically involves minimizing a standard regression loss (e.g., MAE and MSE) across all samples from these  $N - 1$  subjects, without any explicit mechanism for domain invariance or meta-learning.

For subject  $S_{\text{test}}$ , a small adaptation set  $D_{\text{test}}^S = \{(\mathbf{x}_{\text{test},l}, y_{\text{test},l})\}_{l=1}^k$  is available. For each sample in this set, the true reference output  $y_{\text{test},l}$  and the prediction from the pre-trained model,  $\hat{y}_{\text{pre},l} = f_{\theta^*}(\mathbf{x}_{\text{test},l})$  are used to learn a transformation for subject-specific adaptation. Once the transformation parameters are learned, they are applied to adjust the pre-trained model's predictions  $\hat{y}_{\text{pre},m} = f_{\theta^*}(\mathbf{x}_{\text{test},m})$  for the actual test samples (query set)  $D_{\text{test}}^Q = \{\mathbf{x}_{\text{test},m}\}_{m=1}^{M_Q}$  of subject  $S_{\text{test}}$ .

### 6.6.1 Affine Transformation

The affine transformation provides a simple linear correction involving scaling and translation. For the regression output, the adjusted prediction  $\hat{y}_{\text{adj}}$  is related to the pre-trained prediction  $\hat{y}_{\text{pre}}$  by:

$$\hat{y}_{\text{adj}} = s \cdot \hat{y}_{\text{pre}} + t \quad (6.11)$$

where  $s$  is the scaling factor and  $t$  is the translation (offset or intercept). The parameters  $s$  and  $t$  are estimated using the  $k$  samples in the adaptation set  $D_{\text{test}}^S$  by minimizing the sum of squared errors between the transformed predictions and the true reference values:

$$(s^*, t^*) = \arg \min_{s,t} \sum_{l=1}^k ((s \cdot \hat{y}_{\text{pre},l} + t) - y_{\text{test},l})^2 \quad (6.12)$$

The optimal  $s^*$  (slope) and  $t^*$  (intercept) can be found using ordinary least squares. The scaling factor  $s^*$  can be positive or negative, allowing for correction of both magnitude and potential reflection of the bias. Once  $s^*$  and  $t^*$  are determined, the predictions for the real test samples  $\hat{y}_{\text{pre},m}$  from  $Q_{\text{test}}$  are adjusted as:  $\hat{y}_{\text{adj},m} = s^* \cdot \hat{y}_{\text{pre},m} + t^*$ .

### 6.6.2 Procrustes Analysis

Procrustes analysis [213, 214] is a statistical shape analysis method to find the optimal similarity transformation via translation, rotation and scaling to map a set of points to another. To use Procrustes analysis which is typically used for multi-dimensional data, the data used for transformation need to be formed as 2D points first. A common coordinate (e.g. time stamps) is added to both pre-trained prediction and the true reference value of the adaptation samples, and this common coordinate, denoted  $u_l$  for the  $l$ -th sample in the adaptation set of the test subject  $D_{\text{tets}}^S$ . Then the adaptation set (references, referred to **source points**) is denoted as  $T_S = \{\mathbf{t}_l = (u_l, y_{\text{test},l})\}_{l=1}^k$  and the pre-trained predictions (estimations, referred to **target points**) are denoted as  $P_S = \{\mathbf{p}_l = (u_l, \hat{y}_{\text{pre},l})\}_{l=1}^k$ . Procrustes analysis aims to find an optimal non-negative scale factor  $s \geq 0$ , a 2D rotation matrix  $\mathbf{R}$ , and a 2D translation vector  $\mathbf{c} = (c_u, c_y)^T$  that minimize the sum of squared Euclidean distances between the transformed source points and the target points:

$$(s^*, \mathbf{R}^*, \mathbf{c}^*) = \arg \min_{s, \mathbf{R}, \mathbf{c}} \sum_{l=1}^k \|s\mathbf{R}\mathbf{p}_l + \mathbf{c} - \mathbf{t}_l\|^2 \quad (6.13)$$

Then the procedures of Procrustes analysis to find these parameters involves:

1. Centering the Point Sets: Calculate the centroids (means) of  $P_S$  and  $T_S$ :

$$\bar{\mathbf{p}} = \frac{1}{k} \sum_{l=1}^k \mathbf{p}_l \quad \text{and} \quad \bar{\mathbf{t}} = \frac{1}{k} \sum_{l=1}^k \mathbf{t}_l$$

Then, center the points:  $\mathbf{p}_{c,l} = \mathbf{p}_l - \bar{\mathbf{p}}$  and  $\mathbf{t}_{c,l} = \mathbf{t}_l - \bar{\mathbf{t}}$ . Let  $P_c$  and  $T_c$  be  $k \times 2$  matrices whose rows are  $\mathbf{p}_{c,l}^T$  and  $\mathbf{t}_{c,l}^T$  respectively.

2. Optimal Rotation  $\mathbf{R}^*$ : Compute the covariance-like matrix  $M = P_c^T T_c$ . Perform Singular Value Decomposition (SVD) on  $M$ :  $M = U\Sigma V^T$ . The optimal rotation matrix is  $\mathbf{R}^* = VU^T$ . If  $\det(\mathbf{R}^*) = -1$ , adjustments are required by flipping the sign

of the column of  $V$  corresponding to the smallest singular value before computing  $\mathbf{R}^*$ .

3. Optimal Scale  $s^*$ : The scale factor is computed as  $s^* = \frac{\text{trace}(\mathbf{R}^{*T} M)}{\text{trace}(P_c^T P_c)} = \frac{\text{trace}(\Sigma)}{\sum_{l=1}^k \|\mathbf{p}_{c,l}\|^2}$ , which ensures  $s^* \geq 0$ .
4. Optimal Translation  $\mathbf{c}^*$ : The translation vector is:  $\mathbf{c}^* = \bar{\mathbf{t}} - s^* \mathbf{R}^* \bar{\mathbf{p}}$ .

Once the optimal parameters  $(s^*, \mathbf{R}^*, \mathbf{c}^*)$  are determined from the adaptation set  $D_{\text{test}}^S$ , they are applied to adjust the pre-trained predictions  $\hat{y}_{\text{pre},m}$  for the real test samples in  $D_{\text{test}}^Q$ . For each query sample  $m$ , a common coordinate  $u_m$  is added to form 2D points  $\mathbf{p}_m = (u_m, \hat{y}_{\text{pre},m})$ . Then the learned Procrustes transformation is applied  $\mathbf{p}'_m = (u'_m, \hat{y}_{\text{adj},m}) = s^* \mathbf{R}^* \mathbf{p}_m + \mathbf{c}^*$ . The final adjusted BP prediction is the second component of the transformed point  $\mathbf{p}'_m$ , which is  $\hat{y}_{\text{adj},m}$ .

## 6.7 Experiments

In this section, the aforementioned subject-specific adaptation methods are applied to the **TCN-SENet-KAN** model which can accurately track BP changes and has great model generalization even with LOSO cross-validation setup (as demonstrated in Chapter 5). Unlike the previous related studies which requires additional 20% data [112] or a minimal of 50 samples [113] from the test subject to fine-tune the pre-trained models for satisfactory results, this study aims to use extremely few calibrations for personal adaptation. In this study, the first 20% of samples (corresponds to the additional 20% of samples required for personal fine-tuning proposed in [112]) of the test subject are selected as the adaptation window, which also simulates a real-world application scenario where a new user conducts very few initial reference BP measures to calibrate a newly bought device. Since the test subject only has approximately 190 samples (corresponding to  $\sim 6.3$  minutes) due to the data segmentation process described in Section 5.4.5, this adaptation window ( $\sim 190 \times 0.2 \approx 35$  samples) is still considerably small. To reveal the impact of different number of calibration samples to adaptation results, from the first 20% adaptation window,  $k$  samples ( $k = 1, 2, 4, 10, 15$  and all 20%) are equally extracted from this adaptation window for subject-specific adaptation. The remaining 80% samples are used for performance evaluation. This developed data usage method is demonstrated in Figure 6.3.

To prevent the pre-trained model overfitting on these small number of calibration points, a smaller learning rate and number of iterations are set for adapting using fine-tuning, MAML and DANN methods.

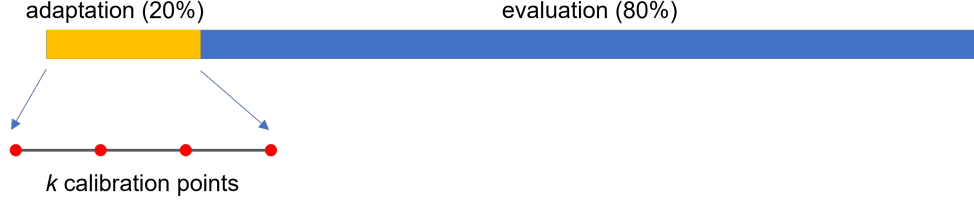


Figure 6.3: Demonstration of adaptation window, evaluation set and calibration samples selection method applied to each test subject of LOSO cross validation for this study.

## 6.8 Results

In this section, the results of the experiments by applying different adaptation methods with different number of adaptation samples are demonstrated. The mean and standard deviation of AE between the reference and estimated BPs for all 32 subjects using different subject-specific adaptation methods with different number of calibration samples are demonstrated in Table 6.2.

The results presented in Table 6.2 and 6.3 highlight the effectiveness of various subject-specific adaptation methods applied to both TCN-SENet-KAN and TCN-SENet-Linear models under different numbers of calibration samples. Among all methods, the DANN pre-training with subject-specific adaptation via fine-tuning with sDTW loss function and Procrustes analysis consistently delivers the best performance in terms of both estimation accuracy and trend tracking, across all adaptation settings. This approach combines three key components: domain adversarial training to encourage subject-invariant feature learning, shape-aware fine-tuning using sDTW loss to preserve BP trends, and a simple Procrustes transformation to align predictions with the reference BP levels.

From Table 6.2, this method achieves lowest MAEs and standard deviations across different calibration sizes. Notably, even with just 2 calibration samples, the model achieves  $4.34 \pm 4.34$  mmHg (SBP) and  $2.70 \pm 2.67$  mmHg (DBP), which meets the IEEE cuffless BP estimation standard ( $MAE \leq 7$  mmHg) and even can be considered as "Grade A" ( $MAE \leq 5$  mmHg). This performance is better than other methods, including the DANN and mean adjustment method (4th row), which yields higher errors (e.g.,  $4.39 \pm 4.38$  mmHg for SBP and  $2.63 \pm 2.45$  mmHg for DBP at 2 samples). Moreover, in settings with 4 or 10 calibration samples, the combined method maintains high accuracy, showing its robustness in low-data personalization.

Table 6.3 further confirms these findings in terms of trend tracking performance. The DANN + Procrustes method consistently achieves the highest median Pearson correlation

coefficients across different calibration settings. For example, with 4 calibration samples, it achieves 0.790 (IQR: 0.410) for SBP and 0.603 (IQR: 0.423) for DBP, indicating strong alignment with true BP fluctuations. In comparison, the MAML method fails to provide meaningful estimates which constantly producing negative correlation values (e.g.,  $-0.383$  for SBP at 2 samples), suggesting the model cannot even capture the correct direction of BP trends. This poor performance stems from the nature of MAML: during both meta-training and adaptation, only a few samples (k-shots) are used, which is insufficient for a model to learn complex physiological relationships needed for BP estimation. Even with the maximum 20% sample setting ( $\sim 30$  samples), MAML still yields poor accuracy (e.g.,  $24.00 \pm 21.03$  mmHg for SBP,  $13.41 \pm 14.65$  mmHg for DBP), indicating it practically unusable for this task.

The DANN method alone without Procrustes transformation also outperforms regular pre-training strategies, even with limited calibration data. For example, with 2 samples, DANN achieves  $4.39 \pm 4.38$  mmHg for SBP compared to  $9.50 \pm 20.04$  mmHg from fine-tuning the TCN-SENet-Linear baseline. This confirms that DANN’s subject-invariant representation learning is more suitable for generalizing to new subjects. However, by incorporating Procrustes analysis, the estimation can be further aligned to the subject’s BP range, leading to additional improvements in both MAE and correlation coefficient, particularly in small-sample scenarios.

As demonstrated in Figure 6.5, which illustrates box plot of subject-wise AEs using only two calibration measures, the MAML adaptation method performed notably poorly. MAML consistently showed the highest median MAEs for both SBP and DBP, approximately 13.22 mmHg and 9.17 mmHg respectively, coupled with the widest error distributions. This indicates significant inaccuracies and a lack of reliability across subjects with minimal calibration settings. In contrast, several other adaptation strategies, including DANN, affine transformation, Procrustes analysis, and particularly the combined DANN+Procrustes method, consistently achieved the lowest median MAEs with two calibration samples. The DANN+Procrustes approach, for instance, yielded a median SBP MAE of approximately 3.58 mmHg and a DBP MAE of about 2.28 mmHg. Standard fine-tuning techniques generally yielded moderate performance.

These findings with two calibration samples are almost identical with the observations from Figure 6.4, which utilized four calibration measures. While using more calibration data generally led to slightly lower MAEs for all adaptation methods, the relative performance ranking of the adaptation methods remained similar, with MAML performing worst and approaches like DANN+Procrustes performing best.

As demonstrated in Figure 6.6 and 6.7. The proposed 3-step DANN+Procrustes adap-

tation method and DANN with fine-tuning framework generates similar BP estimation results, and both estimations are almost identical to the reference BP trend. In contrast, the BP estimations using MAML and fine-tuning methods are inaccurate in terms of both trend and scale.

Finally, it is important to emphasize that the TCN-SENet-KAN model is lightweight, with only  $\sim 26.5K$  trainable parameters, allowing full-model fine-tuning without significant computational burden. The statistical Procrustes step is simple and fast to compute. Together, these properties make the entire adaptation process efficient and practical for real-time implementation on edge devices such as smartwatches.

Table 6.2: Mean and standard deviation of AE between the reference and estimated BPs for all 32 subjects using different subject-specific adaptation methods with different number of calibration samples

| Model & adapt method                | Loss function                                                                   | 20% samples   |               | 15 samples    |               | 10 samples    |               | 4 samples     |               | 2 samples     |              | 1 sample      |               |
|-------------------------------------|---------------------------------------------------------------------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|--------------|---------------|---------------|
|                                     |                                                                                 | SBP           | DBP           | SBP           | DBP           | SBP           | DBP           | SBP           | DBP           | SBP           | DBP          | SBP           | DBP           |
| TCN-SENet-Linear fine-tuning (last) | pre-train: MAE<br>fine-tune: MAE                                                | 9.55 ± 20.08  | 4.58 ± 9.58   | 9.52 ± 20.00  | 4.57 ± 9.58   | 9.50 ± 19.94  | 4.56 ± 9.55   | 9.52 ± 20.05  | 4.57 ± 9.54   | 9.50 ± 20.04  | 4.58 ± 9.54  | 10.23 ± 20.17 | 5.20 ± 9.56   |
| TCN-SENet-Linear fine-tuning (all)  | pre-train: MAE<br>adapt: MAE                                                    | 9.02 ± 17.28  | 4.37 ± 8.09   | 9.01 ± 17.28  | 4.36 ± 8.16   | 8.97 ± 17.14  | 4.36 ± 8.17   | 8.98 ± 17.13  | 4.36 ± 8.10   | 8.99 ± 17.12  | 4.38 ± 8.25  | 10.09 ± 19.37 | 5.14 ± 9.18   |
| TCN-SENet-Linear MAML               | pre-train: MAE<br>adapt: MAE                                                    | 24.40 ± 21.23 | 14.01 ± 13.49 | 24.11 ± 34.63 | 13.14 ± 14.05 | 22.06 ± 25.03 | 14.58 ± 15.32 | 19.06 ± 15.55 | 13.08 ± 13.75 | 24.43 ± 23.62 | 11.52 ± 9.93 | 27.26 ± 39.69 | 13.02 ± 12.31 |
| TCN-SENet-KAN DANN                  | pre-train: sDTW <sub>loss</sub><br>adapt: sDTW <sub>loss</sub>                  | 3.99 ± 4.62   | 2.40 ± 2.67   | 3.97 ± 4.57   | 2.42 ± 2.66   | 4.03 ± 4.57   | 2.45 ± 2.66   | 4.15 ± 4.51   | 2.52 ± 2.64   | 4.39 ± 4.38   | 2.65 ± 2.63  | 4.71 ± 5.05   | 2.91 ± 3.20   |
| TCN-SENet-KAN Affine transform      | pre-train: sDTW <sub>loss</sub><br>adapt: squared errors                        | 4.62 ± 5.11   | 2.50 ± 3.02   | 4.61 ± 5.08   | 2.47 ± 3.04   | 4.60 ± 5.08   | 2.46 ± 3.08   | 4.86 ± 5.90   | 2.81 ± 3.70   | 9.06 ± 14.59  | 4.14 ± 6.07  | –             | –             |
| TCN-SENet-KAN Procrustes analysis   | pre-train: sDTW <sub>loss</sub><br>adapt: squared errors                        | 4.18 ± 4.67   | 2.54 ± 2.73   | 4.18 ± 4.59   | 2.55 ± 2.73   | 4.22 ± 4.65   | 2.58 ± 2.74   | 4.32 ± 4.62   | 2.69 ± 2.79   | 4.57 ± 4.48   | 2.92 ± 2.87  | –             | –             |
| TCN-SENet-KAN DANN+Procrustes       | pre-train: sDTW <sub>loss</sub><br>adapt: sDTW <sub>loss</sub> & squared errors | 3.98 ± 4.60   | 2.40 ± 2.67   | 3.99 ± 4.59   | 2.40 ± 2.64   | 4.03 ± 4.58   | 2.45 ± 2.67   | 4.18 ± 4.54   | 2.53 ± 2.66   | 4.34 ± 4.34   | 2.70 ± 2.67  | –             | –             |

Table 6.3: Median and interquartile range (IQR) of Pearson correlation coefficient ( $r$ ) between the reference and estimated BPs for each of the 32 subjects using different subject-specific adaptation methods with different number of calibration samples

| Model & adapt method                | Loss function                                                                   | 20% samples    |                | 15 samples     |                | 10 samples     |                | 4 samples      |                | 2 samples      |                | 1 sample       |                |
|-------------------------------------|---------------------------------------------------------------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
|                                     |                                                                                 | SBP            | DBP            | SBP            | DBP            | SBP            | DBP            | SBP            | DBP            | SBP            | DBP            | SBP            | DBP            |
| TCN-SENet-Linear fine-tuning (last) | pre-train: MAE<br>fine-tune: MAE                                                | -0.038 (0.732) | 0.088 (0.653)  | -0.038 (0.732) | 0.088 (0.653)  | -0.038 (0.732) | 0.088 (0.653)  | -0.038 (0.732) | 0.088 (0.653)  | -0.038 (0.732) | 0.088 (0.653)  | -0.038 (0.732) | 0.088 (0.653)  |
| TCN-SENet-Linear fine-tuning (all)  | pre-train: MAE<br>adapt: MAE                                                    | 0.016 (0.722)  | 0.111 (0.635)  | 0.017 (0.722)  | 0.088 (0.577)  | 0.006 (0.720)  | 0.100 (0.610)  | 0.010 (0.702)  | 0.094 (0.587)  | 0.020 (0.696)  | 0.077 (0.609)  | -0.008 (0.714) | 0.092 (0.695)  |
| TCN-SENet-Linear MAML               | pre-train: MAE<br>adapt: MAE                                                    | -0.479 (0.625) | -0.476 (0.500) | -0.635 (0.394) | -0.504 (0.525) | -0.638 (0.289) | -0.322 (0.584) | -0.695 (0.574) | -0.463 (0.495) | -0.721 (0.520) | -0.383 (0.496) | -0.683 (0.677) | -0.381 (0.590) |
| TCN-SENet-KAN DANN                  | pre-train: sDTW <sub>loss</sub><br>adapt: sDTW <sub>loss</sub>                  | 0.789 (0.393)  | 0.602 (0.428)  | 0.788 (0.398)  | 0.600 (0.434)  | 0.786 (0.401)  | 0.590 (0.429)  | 0.789 (0.403)  | 0.593 (0.466)  | 0.782 (0.413)  | 0.601 (0.405)  | 0.783 (0.435)  | 0.549 (0.537)  |
| TCN-SENet-KAN Affine transform      | pre-train: sDTW <sub>loss</sub><br>adapt: squared errors                        | 0.783 (0.496)  | 0.524 (0.495)  | 0.783 (0.496)  | 0.524 (0.495)  | 0.783 (0.623)  | 0.545 (0.427)  | 0.653 (1.317)  | 0.470 (0.802)  | -0.187 (1.560) | 0.285 (1.233)  | –              | –              |
| TCN-SENet-KAN Procrustes analysis   | pre-train: sDTW <sub>loss</sub><br>adapt: squared errors                        | 0.783 (0.435)  | 0.549 (0.537)  | 0.783 (0.435)  | 0.549 (0.537)  | 0.783 (0.435)  | 0.549 (0.537)  | 0.783 (0.435)  | 0.549 (0.537)  | 0.783 (0.435)  | 0.549 (0.537)  | –              | –              |
| TCN-SENet-KAN DANN+Procrustes       | pre-train: sDTW <sub>loss</sub><br>adapt: sDTW <sub>loss</sub> & squared errors | 0.788 (0.393)  | 0.602 (0.432)  | 0.788 (0.396)  | 0.598 (0.437)  | 0.786 (0.398)  | 0.590 (0.410)  | 0.790 (0.410)  | 0.598 (0.418)  | 0.781 (0.411)  | 0.603 (0.403)  | –              | –              |

Table 6.4: Median and interquartile range (IQR) of the proposed distance, trend and composite similarities between the reference and estimated BPs for each of the 32 subjects using different adaptation method with different number of adaptation samples ( $k = 1, 2, 4$ ).

| Model & adapt method                | Loss function                                                                   | 4 samples    |             |              |              |             |              | 2 samples    |              |              |              |             |              | 1 sample      |             |              |              |             |              |
|-------------------------------------|---------------------------------------------------------------------------------|--------------|-------------|--------------|--------------|-------------|--------------|--------------|--------------|--------------|--------------|-------------|--------------|---------------|-------------|--------------|--------------|-------------|--------------|
|                                     |                                                                                 | SBP          |             |              | DBP          |             |              | SBP          |              |              | DBP          |             |              | SBP           |             |              | DBP          |             |              |
|                                     |                                                                                 | DS           | TS          | CS           | DS           | TS          | CS           | DS           | TS           | CS           | DS           | TS          | CS           | DS            | TS          | CS           | DS           | TS          | CS           |
| TCN-SENet-Linear fine-tuning (last) | pre-train: MAE<br>fine-tune: MAE                                                | -0.08 (0.66) | 0.22 (0.31) | 0.06 (0.34)  | 0.83 (0.30)  | 0.23 (0.33) | 0.56 (0.26)  | 0.02 (0.81)  | 0.22 (0.31)  | 0.09 (0.43)  | 0.82 (0.35)  | 0.23 (0.32) | 0.56         | -0.003 (0.75) | 0.22 (0.31) | 0.07 (0.42)  | 0.72 (0.60)  | 0.22 (0.32) | 0.50 (0.32)  |
| TCN-SENet-Linear fine-tuning (all)  | pre-train: MAE<br>adapt: MAE                                                    | -0.05 (0.62) | 0.22 (0.26) | 0.09 (0.34)  | 0.82 (0.32)  | 0.22 (0.35) | 0.55 (0.22)  | 0.01 (0.84)  | 0.19 (0.33)  | 0.08 (0.44)  | 0.82 (0.36)  | 0.19 (0.37) | 0.54 (0.25)  | 0.004 (0.80)  | 0.18 (0.21) | 0.07 (0.45)  | 0.71 (0.44)  | 0.22 (0.29) | 0.50 (0.28)  |
| TCN-SENet-Linear MAML               | pre-train: MAE<br>adapt: MAE                                                    | -0.89 (0.58) | 0.07 (0.39) | -0.46 (0.42) | -0.69 (1.40) | 0.16 (0.31) | -0.29 (0.77) | -0.93 (0.38) | 0.006 (0.30) | -0.48 (0.28) | -0.74 (1.37) | 0.26 (0.34) | -0.32 (0.74) | -0.96 (0.58)  | 0.08 (0.28) | -0.46 (0.25) | -0.78 (1.29) | 0.10 (0.32) | -0.42 (0.86) |
| TCN-SENet-KAN DANN                  | pre-train: sDTW <sub>loss</sub><br>adapt: sDTW <sub>loss</sub>                  | 0.73 (0.77)  | 0.49 (0.26) | 0.56 (0.51)  | 0.96 (0.29)  | 0.37 (0.21) | 0.70 (0.21)  | 0.59 (0.72)  | 0.47 (0.26)  | 0.47 (0.51)  | 0.93 (0.31)  | 0.36 (0.22) | 0.67 (0.17)  | 0.71 (0.71)   | 0.48 (0.22) | 0.58 (0.38)  | 0.94 (0.42)  | 0.32 (0.28) | 0.64 (0.27)  |
| TCN-SENet-KAN Affine transform      | pre-train: sDTW <sub>loss</sub><br>adapt: squared errors                        | 0.70 (0.90)  | 0.40 (0.29) | 0.56 (0.60)  | 0.93 (0.29)  | 0.29 (0.26) | 0.66 (0.23)  | 0.14 (1.08)  | 0.28 (0.64)  | 0.12 (0.81)  | 0.86 (0.65)  | 0.28 (0.29) | 0.62 (0.35)  | –             | –           | –            | –            | –           | –            |
| TCN-SENet-KAN Procrustes analysis   | pre-train: sDTW <sub>loss</sub><br>adapt: squared errors                        | 0.68 (0.90)  | 0.48 (0.24) | 0.57 (0.50)  | 0.96 (0.29)  | 0.32 (0.33) | 0.64 (0.22)  | 0.50 (0.80)  | 0.47 (0.22)  | 0.45 (0.48)  | 0.94 (0.36)  | 0.32 (0.30) | 0.65 (0.20)  | –             | –           | –            | –            | –           | –            |
| TCN-SENet-KAN DANN+Procrustes       | pre-train: sDTW <sub>loss</sub><br>adapt: sDTW <sub>loss</sub> & squared errors | 0.73 (0.76)  | 0.50 (0.28) | 0.57 (0.52)  | 0.96 (0.28)  | 0.39 (0.22) | 0.70 (0.24)  | 0.58 (0.71)  | 0.47 (0.31)  | 0.49 (0.48)  | 0.94 (0.32)  | 0.37 (0.20) | 0.68 (0.20)  | –             | –           | –            | –            | –           | –            |

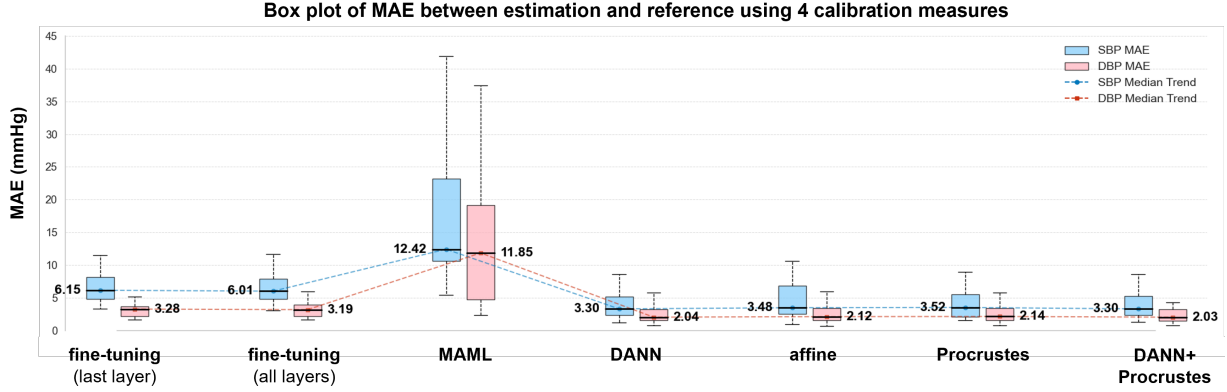


Figure 6.4: Box plot of MAE between reference and estimated BPs using different adaptation methods with 4 calibration measures.

## 6.9 Conclusion and Discussion

This chapter presented several subject-specific adaptation methods to quickly personalize a pre-trained BP estimation model using very few calibration samples. The main goal was to reduce the amount of calibration required while maintaining accurate BP prediction and trend tracking for new, unseen subjects.

Among all methods evaluated, the 3-step DANN pre-training and fine-tuning using sDTW loss function and Procrustes analysis framework demonstrated the best overall performance. It consistently achieved the lowest MAE and the highest correlation coefficients across various calibration settings. This method improves model generalization by using domain adversarial training, which encourages the model to learn subject-invariant features, in contrast to regular pre-training that treats all subject data equally. This makes the model more robust when applied to new subjects with limited data.

To further personalize the model, shape-aware fine-tuning using sDTW loss was applied. This helped the model capture subject-specific BP trends, although improvements in MAE were limited when only a few calibration samples were available. Finally, Procrustes analysis was used to align the estimated BP values with the subject’s reference range through a simple transformation. This step significantly improved both estimation accuracy and trend similarity, especially in few-calibration settings.



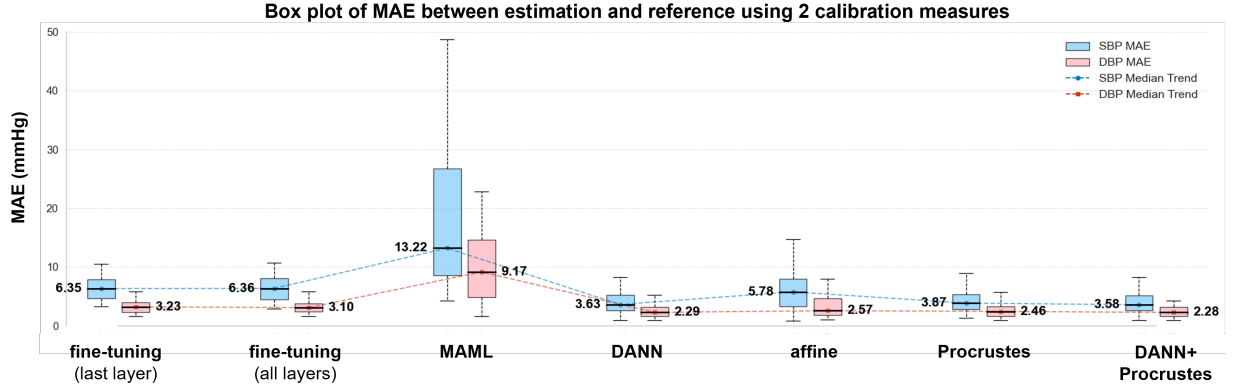


Figure 6.5: Box plot of MAE between reference and estimated BPs using different adaptation methods with 2 calibration measures.

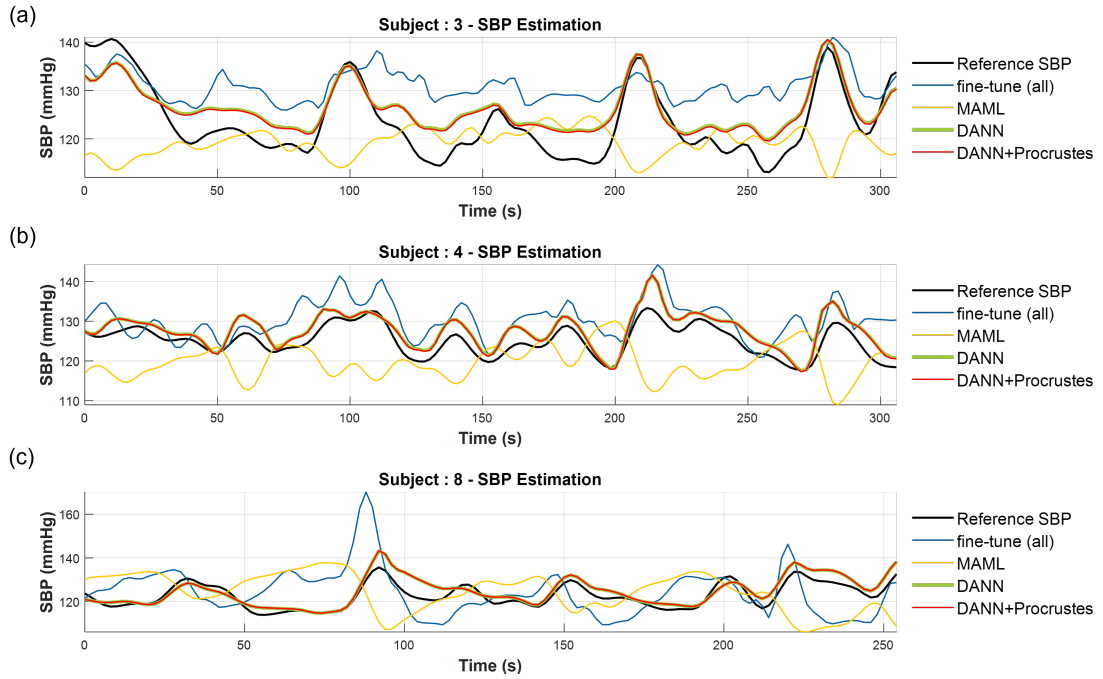


Figure 6.6: SBP estimation for three selected subjects using different adaptation methods with only 2 calibration measures.

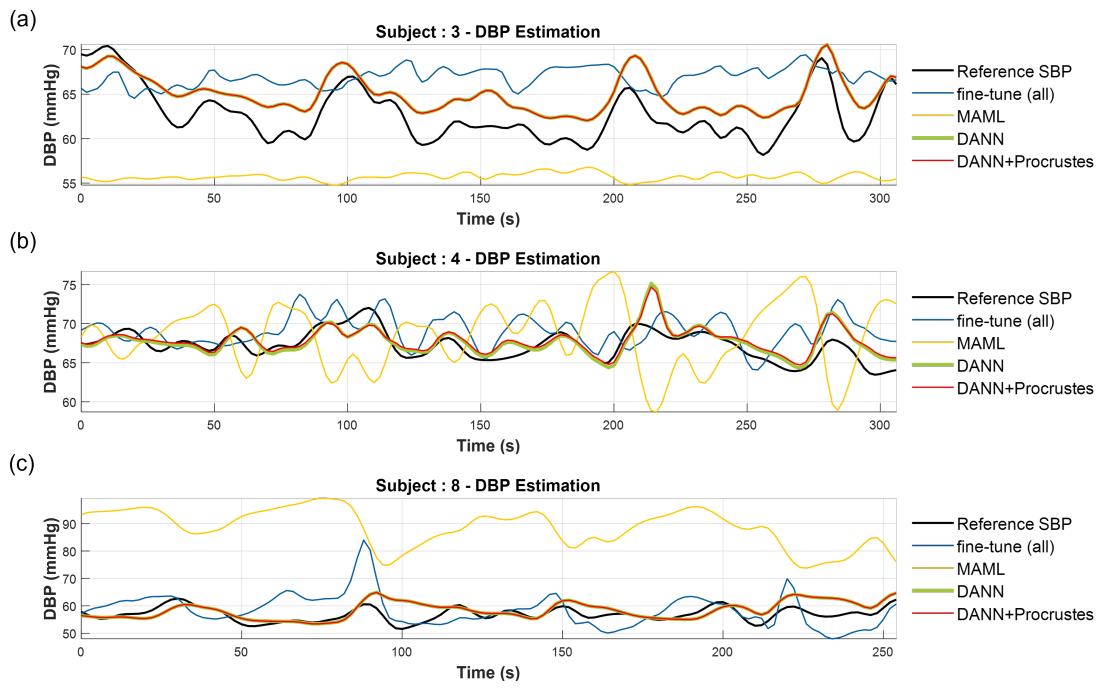


Figure 6.7: DBP estimation for three selected subjects using different adaptation methods with only 2 calibration measures.

Compared to traditional fine-tuning or meta-learning methods like MAML, this combined approach exhibits significant improvements. MAML failed to provide useful estimates, especially with 1–4 calibration samples, and constantly produced predictions with negative correlation values, indicating an inability to track correct BP trends.

The proposed method meets the IEEE cuffless BP estimation standards and even can be considered as "Grade A" for both SBP and DBP estimation using only 2 calibration samples ( $\text{MAE} \leq 5$  mmHg, [25]), making it practical for real-world use. The overall framework remains lightweight and efficient. The TCN-SENet-KAN model has a small number of parameters and allows full-model fine-tuning without high computational cost. The Procrustes transformation is a simple statistical adjustment that can be easily implemented on edge devices such as smartwatches.

In summary, the DANN-based adaptation method with shape-aware fine-tuning and Procrustes transformation provides a simple yet effective solution to reduce calibration burden while maintaining high accuracy and trend tracking performance in subject-specific BP estimation.

# Chapter 7

## Conclusions and Future Work

### 7.1 Conclusions

This thesis addresses four critical challenges in cuffless BP estimation: respiration-induced variations in morphological features and BPs, unsatisfactory trend tracking performance, poor generalization and inefficient personalization and inadequate evaluation metrics. A novel deep learning model TCN-SENet-KAN is introduced, incorporating temporal, feature-specific, and non-linear modeling components, and is optimized using a shape-oriented sDTW loss to enhance trend learning. To improve model generalization and adaptation, several personalization strategies are explored. A 3-step framework is proposed, including DANN-based pre-training, sDTW-based fine-tuning, and Procrustes analysis, which is highly effective in adapting a pre-trained model for completely new subjects with extremely limited calibrations (e.g. 2 and 4 initial calibrations) while maintaining accurate BP value estimation and trend tracking. Additionally, a VMD method is proposed to remove respiratory modulation effects from physiological features and BPs before model training and consequently enhances model performance. A new composite evaluation metric is also developed to assess both trend and distance similarity, addressing limitations of conventional performance metrics. Overall, these contributions form a robust and practical framework for real-world cuffless BP monitoring.

The key findings corresponding to these objectives are summarized below.

In Chapter 3, the often-overlooked impact of respiration on cuffless BP estimation accuracy is investigated. The experimental results indicate that respiration can modulate a wide variety of physiological features and BPs to varying degrees and at different phases,

which makes accurate BP estimation more challenge. To address this problem, a VMD based method with frequency analysis was proposed to separate and remove respiratory-induced variations from these features. Experiments comparing BP estimations derived from both original respiration-modulated features and the RM-suppressed features showed notable performance improvements in terms of both estimation accuracy and trend tracking, as measured by MAE and Pearson correlation coefficient. These findings underscore the critical importance of accounting for and mitigating respiratory effects on physiological features prior to model training to enhance the reliability of cuffless BP estimation in future research and practical applications.

Chapter 4 addresses the inadequacy of current evaluation protocols for continuous cuffless BP estimation, which primarily rely on distance-based metrics and fail to properly assess a model’s ability to track BP changes over time. A comprehensive evaluation framework is proposed, incorporating novel normalized metrics to quantify both distance and trend similarity, along with a composite metric. By applying these metrics to SOTA algorithms, the results demonstrate that even models with low distance errors can poorly track short-term BP variations. The proposed metrics range from -1 to 1, offering an intuitive interpretation similar to correlation coefficients. These new evaluation metrics are recommended for regular use in cuffless BP estimation.

Chapter 5 aims to enhance model generalization particularly for BP trend tracking. In Chapter 5, two time series regression models, TCN-SENet-KAN and Transformer-BiGRU-Attention models are proposed to improve BP trend tracking. The TCN-SENet-KAN which utilizes TCN to capture temporal dependencies from input physiological features, SENet to amplify informative features and KAN to enhance non-linearity modeling, incorporating a shape-oriented loss function [sDTW](#) to encourage the estimation model to focus on BP trend instead of pointwise error like most existing studies. The experiment results indicate the proposed TCN-SENet-KAN model with sDTW loss function significantly enhanced model generalization for BP trend tracking even under LOSO cross-validation settings, with the median Pearson correlation coefficient between the reference and estimated SBP and DBP trends of 0.74 and 0.56 respectively. The results also indicate that the shape-oriented loss function can also notably improve trend tracking performance for a same model architecture.

In Chapter 6, several few-shot adaptation methods were introduced, including model fine-tuning, meta learning using MAML, domain adaptation using [DANN](#) methods and statistical transformation methods, affine transformation and Procrustes analysis. The main goal is to reduce the amount of calibration required while maintaining accurate BP prediction and trend tracking for new, unseen subjects. Among all methods evaluated, the DANN pre-training and fine-tuning with sDTW loss and Procrustes analysis approach

demonstrated the best overall performance. It consistently achieved the lowest MAE and the highest correlation coefficients across various calibration settings. This method improves model generalization by using domain adversarial training, which encourages the model to learn subject-invariant features. To further personalize the model, shape-oriented fine-tuning using sDTW loss was applied. This helped the model capture subject-specific BP trends, although improvements were limited when only a few calibration samples were available. Finally, Procrustes analysis was used to align the estimated BP values with the subject's reference range through a simple transformation. The proposed method meets the IEEE cuffless BP estimation standards and even can be considered as "Grade A" for both SBP (MAE=4.34 mmHg) and DBP (MAE=2.70) estimation using only 2 calibration samples (MAE $\leq$ 5 mmHg [25]), making it practical for real-world use.

## 7.2 Future Work

Though this thesis successfully addresses the two primary objectives outlined at the outset, several limitations remain that require further investigation in future work.

This study utilizes the BCG BP dataset comprising 32 subjects, where the proposed model integrated with a shape-oriented loss function and efficient adaptation strategies, demonstrates excellent performance. However, its real-world applicability remains to be fully validated. Future work should focus on evaluating the model using larger, publicly available datasets such as MIMIC [203] and VitalDB [215], which encompass a broader study population with more diverse cardiovascular conditions and greater inter-subject variability. These datasets can better simulate real-world complexities and provide a more rigorous assessment of the model's generalizability. Moreover, the concept of pre-training the model on large-scale datasets and transferring it to target domains with different distributions, such as the BCG BP dataset, provides a promising direction for improving model robustness and adaptability in practical deployment scenarios.

As demonstrated in Section 5.1, some features such as PPG pulse amplitude, PAT and PIR are highly related to BP changes. However, the accuracy of BP estimation models that rely heavily on features extracted from PPG waveforms can be significantly affected by external factors, particularly contact force between the sensor and the skin. Studies, such as [216], have demonstrated that increasing contact force reduces PPG pulse amplitude due to restricted blood flow, and may even distort or block the signal entirely. Additionally, factors like skin temperature, sensor placement, subject motion, and light wavelength further alter waveform morphology. For example, lower temperatures decrease pulse amplitude, and differences in sensor configurations (transmission vs. reflection) can

shift diastolic peak positions. These morphological changes can lead to inconsistencies in extracted features even for the same subject, ultimately degrading BP estimation accuracy. Future work should systematically investigate the effects of contact force and other external influences on PPG waveform integrity. This will help in developing signal quality assessment or correction techniques to enhance model robustness and ensure reliable cuffless BP monitoring in real-world applications.

# References

- [1] R. N. Pittman, “Regulation of tissue oxygenation,” 2016.
- [2] J. S. Shahoud, T. Sanvictores, and N. R. Aeddula, “Physiology, arterial pressure regulation,” 2019.
- [3] WIKIPEDIA, “Final report of a trial of intensive versus standard blood- pressure control,” [https://en.wikipedia.org/wiki/Blood\\_pressure](https://en.wikipedia.org/wiki/Blood_pressure), accessed: 2022-11-14.
- [4] R. Chaudhry, J. H. Miao, and A. Rehman, “Physiology, cardiovascular,” in StatPearls [Internet]. StatPearls Publishing, 2022.
- [5] T. D. Homan, S. J. Bordes, and E. Cichowski, “Physiology, mean arterial pressure,” StatPearls, 2024. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK538226/>
- [6] NCD Risk Factor Collaboration (NCD-RisC), “Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19.1 million participants,” The Lancet, vol. 389, no. 10064, pp. 37–55, 2017.
- [7] A. H. A. (AHA), “Health threats from high blood pressure,” American Heart Association Website, 2025. [Online]. Available: <https://www.heart.org/en/health-topics/high-blood-pressure/health-threats-from-high-blood-pressure>
- [8] M. C. Staff, “High blood pressure dangers: Hypertension’s effects on your body,” Mayo Clinic Website, 2024. [Online]. Available: <https://www.mayoclinic.org/diseases-conditions/high-blood-pressure/in-depth/high-blood-pressure/art-20045868>
- [9] W. H. O. (WHO), “Hypertension,” WHO Fact Sheets, 2025. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/hypertension>



- [10] J. Wang, L. Zhang, A. Wang, C. Li, J. Su, M. Zhang, Z. Wang, X. Hou, H. Lin, and Y. Sun, “Long-term cardiovascular risk associated with stage 1 hypertension defined by the 2017 acc/aha hypertension guideline,” Journal of the American College of Cardiology (JACC), vol. 72, no. 1, pp. 8–17, 2018.
- [11] P. H. Institute, “Conditions we treat: Blood pressure disorders,” <https://www.pulseheartinstitute.org/for-patients/conditions-we-treat/blood-pressure-disorders/>, accessed: 2021-07-08.
- [12] S. R. Group, “Final report of a trial of intensive versus standard blood-pressure control,” New England Journal of Medicine, vol. 384, no. 20, pp. 1921–1930, 2021.
- [13] W. H. Organization, “Hypertension,” <https://www.who.int/news-room/fact-sheets/detail/hypertension>, accessed: 2021-07-08.
- [14] D. T. Frieden, “Taming the world’s leading killer: High blood pressure,” <https://www.cnn.com/2021/05/30/health/high-blood-pressure-frieden/index.html>, Jun 2021.
- [15] C. Ludwig, “Beitrage zur kenntniss des einflusses der respirationsbewegungen auf den blutlauf im aortensysteme,” Arch Anat Physiol Leipzig, vol. 13, pp. 242–302, 1847.
- [16] S. Hales, Statical essays, containing haemastaticks, or, An account of some hydraulick and hydrostatical experiments made on the blood and blood vessels of animals: also an account of some experiments on stones in the kidneys and bladder: with an enquiry into the nature of those anomalous concretions: to which is added, an appendix, containing observations and experiments relating to several subjects in the first volume, the greater part of which were read at several meetings before the Royal Society: with an index to both volumes. Vol. II. W. Innys and R. Manby, and T. Woodward, 1733.
- [17] A. Dornhorst, P. Howard, and G. Leathart, “Respiratory variations in blood pressure,” Circulation, vol. 6, no. 4, pp. 553–558, 1952.
- [18] X.-R. Ding, Y.-T. Zhang, J. Liu, W.-X. Dai, and H. K. Tsang, “Continuous cuffless blood pressure estimation using pulse transit time and photoplethysmogram intensity ratio,” IEEE Transactions on Biomedical Engineering, vol. 63, no. 5, pp. 964–972, 2015.
- [19] C. Liu, C. Liu, L. Li, Q. Zhang, and B. Li, “Systolic and diastolic time interval variability analysis and their relations with heart rate variability,” in 2009 3rd

International Conference on Bioinformatics and Biomedical Engineering. IEEE, 2009, pp. 1–4.

- [20] M. Gavriilaki, P. Anyfanti, B. Nikolaidou, A. Lazaridis, E. Gavriilaki, S. Douma, and E. Gkaliagkousi, “Nighttime dipping status and risk of cardiovascular events in patients with untreated hypertension: a systematic review and meta-analysis,” The Journal of Clinical Hypertension, vol. 22, no. 11, pp. 1951–1959, 2020.
- [21] H. Ø. Hjortkjær, F. Persson, S. Theilade, S. A. Winther, N. Tofte, T. S. Ahluwalia, and P. Rossing, “Non-dipping and higher nocturnal blood pressure are associated with risk of mortality and development of kidney disease in type 1 diabetes,” Journal of Diabetes and its Complications, vol. 36, no. 9, p. 108270, 2022.
- [22] P. A. Lempiäinen, A. Ylitalo, H. Huikuri, Y. A. Kesäniemi, and O. H. Ukkola, “Non-dipping blood pressure pattern is associated with cardiovascular events in a 21-year follow-up study,” Journal of Human Hypertension, vol. 38, no. 5, pp. 444–451, 2024.
- [23] Association for the Advancement of Medical Instrumentation and others, “American national standard. manual, electronic or automated sphygmomanometers,” ANSI/AAMI SP10-2002/A1, 2003.
- [24] G. S. Stergiou, B. Alpert, S. Mieke, R. Asmar, N. Atkins, S. Eckert, G. Frick, B. Friedman, T. Graßl, T. Ichikawa et al., “A universal standard for the validation of blood pressure measuring devices: Association for the advancement of medical instrumentation/european society of hypertension/international organization for standardization (aami/esh/iso) collaboration statement,” Hypertension, vol. 71, no. 3, pp. 368–374, 2018.
- [25] “Ieee standard for wearable, cuffless blood pressure measuring devices - amendment 1,” IEEE Std 1708a-2019 (Amendment to IEEE Std 1708-2014), pp. 1–35, 2019.
- [26] “Non-invasive sphygmomanometers - part 3: Clinical investigation of continuous automated measurement type,” ISO 81060-3:2022, pp. 1–36, 2022.
- [27] B. H. McGhee and E. J. Bridges, “Monitoring arterial blood pressure: what you may not know,” Critical care nurse, vol. 22, no. 2, pp. 60–79, 2002.
- [28] D. Perloff, C. Grim, J. Flack, E. D. Frohlich, M. Hill, M. McDonald, and B. Z. Morgenstern, “Human blood pressure determination by sphygmomanometry.” Circulation, vol. 88, no. 5, pp. 2460–2470, 1993.

- [29] Y. Ostchega, J. P. Hughes, G. Zhang, T. Nwankwo, and M. M. Chiappa, “Mean mid-arm circumference and blood pressure cuff sizes for us adults: National health and nutrition examination survey, 1999–2010,” Blood pressure monitoring, vol. 18, no. 3, pp. 138–143, 2013.
- [30] NewPak Surgical, “Blood pressure monitor ad mercury-free um-101,” 2025, accessed: 2025-10-19. [Online]. Available: <https://newpaksurgical.com/product/blood-pressure-monitor-ad-mercury-free-um-101/>
- [31] Pinterest User, “Blood pressure remedies, low blood pressure, blood pressure pregnancy,” 2025, accessed: 2025-10-19. [Online]. Available: <https://www.pinterest.com/pin/blood-pressure-remedies-low-blood-pressure-blood-pressure-pregnancy--825636544156588237/>
- [32] B. S. Alpert, D. Quinn, and D. Gallick, “Oscillometric blood pressure: a review for clinicians,” Journal of the American Society of Hypertension, vol. 8, no. 12, pp. 930–938, 2014.
- [33] G. Drzewiecki, R. Hood, and H. Apple, “Theory of the oscillometric maximum and the systolic and diastolic detection ratios,” Annals of biomedical engineering, vol. 22, no. 1, pp. 88–96, 1994.
- [34] E.-J. Marey, “Pression et vitesse du sang,” Physiologie experimentale, pp. 307–343, 1876.
- [35] M. Forouzanfar, H. R. Dajani, V. Z. Groza, M. Bolic, S. Rajan, and I. Batkin, “Oscillometric blood pressure estimation: past, present, and future,” IEEE reviews in biomedical engineering, vol. 8, pp. 44–63, 2015.
- [36] C. F. Babbs, “Oscillometric measurement of systolic and diastolic blood pressures validated in a physiologic mathematical model,” Biomedical engineering online, vol. 11, no. 1, p. 56, 2012.
- [37] A. Chandrasekhar, M. Yavarimanesh, J.-O. Hahn, S.-H. Sung, C.-H. Chen, H.-M. Cheng, and R. Mukkamala, “Formulas to explain popular oscillometric blood pressure estimation algorithms,” Frontiers in physiology, vol. 10, p. 1415, 2019.
- [38] M. Forouzanfar, S. Ahmad, I. Batkin, H. R. Dajani, V. Z. Groza, and M. Bolic, “Coefficient-free blood pressure estimation based on pulse transit time–cuff pressure dependence,” IEEE Transactions on Biomedical Engineering, vol. 60, no. 7, pp. 1814–1824, 2013.

- [39] K. Lakhal, S. Ehrmann, and T. Boulain, “Noninvasive bp monitoring in the critically ill: time to abandon the arterial catheter?” Chest, vol. 153, no. 4, pp. 1023–1039, 2018.
- [40] B. P. Imholz, W. Wieling, G. A. van Montfrans, and K. H. Wesseling, “Fifteen years experience with finger arterial pressure monitoring: assessment of the technology,” Cardiovascular research, vol. 38, no. 3, pp. 605–616, 1998.
- [41] K. Wesseling, “Physiocal, calibrating finger vascular physiology for finapres,” Homeostasis, vol. 36, pp. 67–82, 1995.
- [42] B. Silke and D. McAuley, “Accuracy and precision of blood pressure determination with the finapres: an overview using re-sampling statistics,” Journal of human hypertension, vol. 12, no. 6, pp. 403–409, 1998.
- [43] Finapres Medical Systems, “Finapres® nova,” <https://www.finapres.com/products/hardware/finapres-nova>, 2025, accessed on June 2, 2025.
- [44] BIOPAC Systems, Inc., “Noninvasive blood pressure - nibp monitoring systems,” <https://www.biopac.com/product-category/research/amplifiers/noninvasive-blood-pressure/>, 2025, accessed on June 2, 2025.
- [45] Edwards Lifesciences Corporation, “Clearsight system,” <https://www.edwards.com/ca-en/healthcare-professionals/products-services/hemodynamic-monitoring/clearsight-system>, 2025, accessed on June 2, 2025.
- [46] G. M. Drzewiecki, J. Melbin, and A. Noordergraaf, “Deformational forces in arterial tonometry,” in Proceedings of a Biomedical Engineering Conference, 1984, pp. 642–645.
- [47] H. J. Griffiths, “Encyclopedia of medical devices and instrumentation,” Radiology, vol. 170, no. 3, pp. 1016–1016, 1989.
- [48] B. Saugel, A. Meidert, A. Hapfelmeier, F. Eyer, R. Schmid, and W. Huber, “Non-invasive continuous arterial pressure measurement based on radial artery tonometry in the intensive care unit: a method comparison study using the t-line tl-200pro device,” British journal of anaesthesia, vol. 111, no. 2, pp. 185–190, 2013.
- [49] R. Dueck, O. Goedje, and P. Clopton, “Noninvasive continuous beat-to-beat radial artery pressure via tl-200 applanation tonometry,” Journal of clinical monitoring and computing, vol. 26, no. 2, pp. 75–83, 2012.

- [50] S. Hansen and M. Staber, "Oscillometric blood pressure measurement used for calibration of the arterial tonometry method contributes significantly to error," European journal of anaesthesiology, vol. 23, no. 9, pp. 781–787, 2006.
- [51] O. Healthcare, "Blood pressure watch: Omron heartguide wrist bp monitor," <https://omronhealthcare.com/products/heartguide-wearable-blood-pressure-monitor-bp8000m/>.
- [52] R. Warren, [https://www.accessdata.fda.gov/cdrh\\_docs/pdf18/K182166.pdf](https://www.accessdata.fda.gov/cdrh_docs/pdf18/K182166.pdf).
- [53] HUAWEI, "Huawei watch d," <https://consumer.huawei.com/cn/wearables/watch-d/>.
- [54] T. Arakawa, "Recent research and developing trends of wearable sensors for detecting blood pressure," Sensors, vol. 18, no. 9, p. 2772, 2018.
- [55] Y. Sun, Y. Quan, J. Xing, Z. Tan, X. Sun, L. Lou, C. Fei, J. Zhu, and Y. Yang, "Lead-free potassium sodium niobate-based wearable ultrasonic patches for blood pressure detection," Micromachines, vol. 16, no. 4, p. 392, 2025.
- [56] S. Zhou, G. Park, K. Longardner, M. Lin, B. Qi, X. Yang, X. Gao, H. Huang, X. Chen, Y. Bian et al., "Clinical validation of a wearable ultrasound sensor of blood pressure," Nature Biomedical Engineering, vol. 9, pp. 865–881, 2025.
- [57] T. H. Huynh, R. Jafari, and W.-Y. Chung, "An accurate bioimpedance measurement system for blood pressure monitoring," Sensors, vol. 18, no. 7, p. 2095, 2018.
- [58] K. Sel, D. Osman, N. Huerta, A. Edgar, R. I. Pettigrew, and R. Jafari, "Continuous cuffless blood pressure monitoring with a wearable ring bioimpedance device," npj Digital Medicine, vol. 6, no. 1, p. 59, 2023.
- [59] C.-T. Hsiao, S. Hong, K. L. Branan, J. McMurray, and G. L. Coté, "Predicting blood pressure without a cuff using a unique multi-modal wearable device and machine learning algorithm," Computers in Biology and Medicine, vol. 192, p. 110357, 2025.
- [60] K. Wang, J. Tang, Y. Fan, J. Ji, Y. Shi, and Y. Wang, "Memory-efficient low-latency remote photoplethysmography through temporal-spatial state space duality," arXiv preprint arXiv:2504.01774, 2025.
- [61] A. Dasari, L. A. Jeni, and C. S. Tucker, "Video-based estimation of blood pressure," PloS one, vol. 20, no. 1, p. e0311654, 2025.

- [62] G. Hwang and S. J. Lee, "Phase-shifted remote photoplethysmography for estimating heart rate and blood pressure from facial video," arXiv preprint arXiv:2401.04560, 2024.
- [63] S.-H. Liu, Y. Sun, B.-Y. Wu, W. Chen, and X. Zhu, "Using machine learning models for cuffless blood pressure estimation with ballistocardiogram and impedance plethysmogram," Frontiers in Digital Health, vol. 7, p. 1511667, 2025.
- [64] C.-S. Kim, A. M. Carek, O. T. Inan, R. Mukkamala, and J.-O. Hahn, "Ballistocardiogram-based approach to cuffless blood pressure monitoring: Proof of concept and potential challenges," IEEE Transactions on Biomedical Engineering, vol. 65, no. 11, pp. 2384–2391, 2018.
- [65] C. K. Bradley, D. Shimbo, D. A. Colburn, D. N. Pugliese, R. Padwal, S. K. Sia, and D. E. Anstey, "Cuffless blood pressure devices," American journal of hypertension, vol. 35, no. 5, pp. 380–387, 2022.
- [66] E. Finnegan, S. Davidson, M. Harford, P. Watkinson, L. Tarassenko, and M. Villarroel, "Features from the photoplethysmogram and the electrocardiogram for estimating changes in blood pressure," Scientific Reports, vol. 13, no. 1, p. 986, 2023.
- [67] H. Chen, L. Lyu, Z. Zeng, Y. Jin, and Y. Zhang, "Beat-to-beat continuous blood pressure estimation with optimal feature set of ppg and ecg signals using deep recurrent neural networks," Vessel Plus, vol. 7, pp. N–A, 2023.
- [68] Y. Zheng, H. Huang, J. Gao, J. Hong, S. Wu, Y. Zhang, and Q. Liu, "Utransbpnet for cuffless and calibration-free blood pressure estimation under dynamic conditions," Scientific Reports, vol. 15, no. 1, pp. 1–12, 2025.
- [69] E. Finnegan, S. Davidson, M. Harford, J. Jorge, P. Watkinson, D. Young, L. Tarassenko, and M. Villarroel, "Pulse arrival time as a surrogate of blood pressure," Scientific reports, vol. 11, no. 1, pp. 1–21, 2021.
- [70] Wikipedia, "Moens–korteweg equation," [https://en.wikipedia.org/wiki/Moens–Korteweg\\_equation](https://en.wikipedia.org/wiki/Moens%E2%80%93Korteweg_equation), Jun 2020.
- [71] J. Shao, P. Shi, S. Hu, Y. Liu, and H. Yu, "An optimization study of estimating blood pressure models based on pulse arrival time for continuous monitoring," Journal of Healthcare Engineering, vol. 2020, 2020.

- [72] W. Chen, T. Kobayashi, S. Ichikawa, Y. Takeuchi, and T. Togawa, "Continuous estimation of systolic blood pressure using the pulse arrival time and intermittent calibration," Medical and Biological Engineering and Computing, vol. 38, no. 5, pp. 569–574, 2000.
- [73] Z. Tang, T. Tamura, M. Sekine, M. Huang, W. Chen, M. Yoshida, K. Sakatani, H. Kobayashi, and S. Kanaya, "A chair-based unobtrusive cuffless blood pressure monitoring system based on pulse arrival time," IEEE Journal of biomedical and health informatics, vol. 21, no. 5, pp. 1194–1205, 2016.
- [74] Y. Zheng, C. C. Poon, B. P. Yan, and J. Y. Lau, "Pulse arrival time based cuff-less and 24-h wearable blood pressure monitoring and its diagnostic value in hypertension," Journal of medical systems, vol. 40, no. 9, pp. 1–11, 2016.
- [75] A. Esmaili, M. Kachuee, and M. Shabany, "Nonlinear cuffless blood pressure estimation of healthy subjects using pulse transit time and arrival time," IEEE Transactions on Instrumentation and Measurement, vol. 66, no. 12, pp. 3299–3308, 2017.
- [76] Y.-C. Hsu, Y.-H. Li, C.-C. Chang, and L. N. Harfiya, "Generalized deep neural network model for cuffless blood pressure estimation with photoplethysmogram signal only," Sensors, vol. 20, no. 19, p. 5668, 2020.
- [77] Y.-H. Li, L. N. Harfiya, K. Purwandari, and Y.-D. Lin, "Real-time cuffless continuous blood pressure estimation using deep learning model," Sensors, vol. 20, no. 19, p. 5606, 2020.
- [78] G. Thambiraj, U. Gandhi, U. Mangalanathan, V. J. M. Jose, and M. Anand, "Investigation on the effect of womersley number, ecg and ppg features for cuff less blood pressure estimation using machine learning," Biomedical Signal Processing and Control, vol. 60, p. 101942, 2020.
- [79] M. H. Chowdhury, M. N. I. Shuzan, M. E. Chowdhury, Z. B. Mahbub, M. M. Uddin, A. Khandakar, and M. B. I. Reaz, "Estimating blood pressure from the photoplethysmogram signal and demographic features using machine learning techniques," Sensors, vol. 20, no. 11, p. 3127, 2020.
- [80] S. Suzuki and K. Oguri, "Cuffless blood pressure estimation by error-correcting output coding method based on an aggregation of adaboost with a photoplethysmograph sensor," in 2009 Annual international conference of the IEEE engineering in medicine and biology society. IEEE, 2009, pp. 6765–6768.

- [81] Y. Kurylyak, F. Lamonaca, and D. Grimaldi, "A neural network-based method for continuous blood pressure estimation from a ppg signal," in 2013 IEEE International instrumentation and measurement technology conference (I2MTC). IEEE, 2013, pp. 280–283.
- [82] F. Miao, N. Fu, Y.-T. Zhang, X.-R. Ding, X. Hong, Q. He, and Y. Li, "A novel continuous blood pressure estimation approach based on data mining techniques," IEEE journal of biomedical and health informatics, vol. 21, no. 6, pp. 1730–1740, 2017.
- [83] S. G. Khalid, J. Zhang, F. Chen, and D. Zheng, "Blood pressure estimation using photoplethysmography only: comparison between different machine learning approaches," Journal of healthcare engineering, vol. 2018, 2018.
- [84] P. Su, X.-R. Ding, Y.-T. Zhang, J. Liu, F. Miao, and N. Zhao, "Long-term blood pressure prediction with deep recurrent neural networks," in 2018 IEEE EMBS International conference on biomedical & health informatics (BHI). IEEE, 2018, pp. 323–328.
- [85] S. G. Khalid, H. Liu, T. Zia, J. Zhang, F. Chen, and D. Zheng, "Cuffless blood pressure estimation using single channel photoplethysmography: a two-step method," IEEE Access, vol. 8, pp. 58 146–58 154, 2020.
- [86] K. Song, K.-y. Chung, and J.-H. Chang, "Cuffless deep learning-based blood pressure estimation for smart wristwatches," IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 7, pp. 4292–4302, 2019.
- [87] C. El-Hajj and P. A. Kyriacou, "Deep learning models for cuffless blood pressure monitoring from ppg signals using attention mechanism," Biomedical Signal Processing and Control, vol. 65, p. 102301, 2021.
- [88] Y. Liang, Z. Chen, R. Ward, and M. Elgendi, "Photoplethysmography and deep learning: enhancing hypertension risk stratification," Biosensors, vol. 8, no. 4, p. 101, 2018.
- [89] C. Landry, S. D. Peterson, and A. Arami, "Nonlinear dynamic modeling of blood pressure waveform: Towards an accurate cuffless monitoring system," IEEE Sensors Journal, vol. 20, no. 10, pp. 5368–5378, 2020.



- [90] M. Panwar, A. Gautam, D. Biswas, and A. Acharyya, “Pp-net: A deep learning framework for ppg-based blood pressure and heart rate estimation,” IEEE Sensors Journal, vol. 20, no. 17, pp. 10 000–10 011, 2020.
- [91] M. S. Tanveer and M. K. Hasan, “Cuffless blood pressure estimation from electrocardiogram and photoplethysmogram using waveform based ann-lstm network,” Biomedical Signal Processing and Control, vol. 51, pp. 382–392, 2019.
- [92] N. Aguirre, E. Grall-Maës, L. J. Cymberknop, and R. L. Armentano, “Blood pressure morphology assessment from photoplethysmogram and demographic information using deep learning with attention mechanism,” Sensors, vol. 21, no. 6, p. 2167, 2021.
- [93] D. U. Jeong and K. M. Lim, “Combined deep cnn-lstm network-based multitasking learning architecture for noninvasive continuous blood pressure estimation using difference in ecg-ppg features,” Scientific Reports, vol. 11, no. 1, p. 13539, 2021.
- [94] G. Martínez, N. Howard, D. Abbott, K. Lim, R. Ward, and M. Elgendi, “Can photoplethysmography replace arterial blood pressure in the assessment of blood pressure?” Journal of clinical medicine, vol. 7, no. 10, p. 316, 2018.
- [95] N. Ibtehaz and M. S. Rahman, “Ppg2abp: Translating photoplethysmogram (ppg) signals to arterial blood pressure (abp) waveforms using fully convolutional neural networks,” arXiv preprint arXiv:2005.01669, 2020.
- [96] D. Nachman, Y. Gepner, N. Goldstein, E. Kabakov, A. B. Ishay, R. Littman, Y. Azmon, E. Jaffe, and A. Eisenkraft, “Comparing blood pressure measurements between a photoplethysmography-based and a standard cuff-based manometry device,” Scientific reports, vol. 10, no. 1, pp. 1–9, 2020.
- [97] J. H. Moon, M.-K. Kang, C.-E. Choi, J. Min, H.-Y. Lee, and S. Lim, “Validation of a wearable cuff-less wristwatch-type blood pressure monitoring device,” Scientific reports, vol. 10, no. 1, pp. 1–9, 2020.
- [98] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, “Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time,” in 2015 IEEE international symposium on circuits and systems (ISCAS). IEEE, 2015, pp. 1006–1009.
- [99] D. Lee, H. Kwon, D. Son, H. Eom, C. Park, Y. Lim, C. Seo, and K. Park, “Beat-to-beat continuous blood pressure estimation using bidirectional long short-term memory network,” Sensors, vol. 21, no. 1, p. 96, 2020.

- [100] H. Eom, D. Lee, S. Han, Y. S. Hariyani, Y. Lim, I. Sohn, K. Park, and C. Park, “End-to-end deep learning architecture for continuous blood pressure estimation using attention mechanism,” Sensors, vol. 20, no. 8, p. 2338, 2020.
- [101] C. El-Hajj and P. A. Kyriacou, “Cuffless blood pressure estimation from ppg signals and its derivatives using deep learning models,” Biomedical Signal Processing and Control, vol. 70, p. 102984, 2021.
- [102] M. Rong and K. Li, “A multi-type features fusion neural network for blood pressure prediction based on photoplethysmography,” Biomedical Signal Processing and Control, vol. 68, p. 102772, 2021.
- [103] T. Athaya and S. Choi, “An estimation method of continuous non-invasive arterial blood pressure waveform using photoplethysmography: A u-net architecture-based approach,” Sensors, vol. 21, no. 5, p. 1867, 2021.
- [104] L. Wang, W. Zhou, Y. Xing, and X. Zhou, “A novel neural network model for blood pressure estimation using photoplethysmography without electrocardiogram,” Journal of healthcare engineering, vol. 2018, 2018.
- [105] S. S. Mousavi, M. Firouzmand, M. Charmi, M. Hemmati, M. Moghadam, and Y. Ghorbani, “Blood pressure estimation from appropriate and inappropriate ppg signals using a whole-based method,” Biomedical Signal Processing and Control, vol. 47, pp. 196–206, 2019.
- [106] S. González, W.-T. Hsieh, and T. P.-C. Chen, “A benchmark for machine-learning based non-invasive blood pressure estimation using photoplethysmogram,” Scientific Data, vol. 10, no. 1, p. 149, 2023.
- [107] F. LaBerge, “Blood pressure estimation through photoplethysmography using deep learning in clinical setting: Critical survey and solutions,” Ph.D. dissertation, Concordia University, 2024.
- [108] W. Wang, P. Mohseni, K. L. Kilgore, and L. Najafizadeh, “Pulsedb: A large, cleaned dataset based on mimic-iii and vitaldb for benchmarking cuff-less blood pressure estimation methods,” Frontiers in Digital Health, vol. 4, p. 1090854, 2023.
- [109] M. Moulaeifard, P. H. Charlton, and N. Strodthoff, “Generalizable deep learning for photoplethysmography-based blood pressure estimation—a benchmarking study,” arXiv preprint arXiv:2502.19167, 2025.

- [110] T. B. D. S. Costa, F. M. Dias, D. A. C. Cardenas, M. A. F. De Toledo, D. M. De Lima, J. E. Krieger, and M. A. Gutierrez, “Blood pressure estimation from photoplethysmography by considering intra-and inter-subject variabilities: guidelines for a fair assessment,” Ieee access, vol. 11, pp. 57 934–57 950, 2023.
- [111] Z.-D. Liu, Y. Li, Y.-T. Zhang, J. Zeng, Z.-X. Chen, J.-K. Liu, and F. Miao, “Hgctnet: Handcrafted feature-guided cnn and transformer network for wearable cuffless blood pressure measurement,” IEEE Journal of Biomedical and Health Informatics, 2024.
- [112] G. Slapničar, N. Mlakar, and M. Luštrek, “Blood pressure estimation from photoplethysmogram using a spectro-temporal deep neural network,” Sensors, vol. 19, no. 15, p. 3420, 2019.
- [113] J. Leitner, P.-H. Chiang, and S. Dey, “Personalized blood pressure estimation using photoplethysmography: A transfer learning approach,” IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 1, pp. 218–228, 2021.
- [114] F. Fan, Y. Gu, J. Shen, F. Dong, and Y. Chen, “Fewshotbp: Towards personalized ubiquitous continuous blood pressure measurement,” Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 7, no. 3, pp. 1–39, 2023.
- [115] E. O’Brien, J. Petrie, W. Littler, M. De Swiet, P. L. Padfield, D. Altman, M. Bland, A. Coats, N. Atkins et al., “The british hypertension society protocol for the evaluation of blood pressure measuring devices,” J hypertens, vol. 11, no. Suppl 2, pp. S43–S62, 1993.
- [116] E. O’Brien, N. Atkins, G. Stergiou, N. Karpettas, G. Parati, R. Asmar, Y. Imai, J. Wang, T. Mengden, A. Shennan et al., “European society of hypertension international protocol revision 2010 for the validation of blood pressure measuring devices in adults,” Blood pressure monitoring, vol. 15, no. 1, pp. 23–38, 2010.
- [117] N. Ibtehaz, S. Mahmud, M. E. Chowdhury, A. Khandakar, M. Salman Khan, M. A. Ayari, A. M. Tahir, and M. S. Rahman, “Ppg2abp: Translating photoplethysmogram (ppg) signals to arterial blood pressure (abp) waveforms,” Bioengineering, vol. 9, no. 11, p. 692, 2022.
- [118] G. S. Stergiou, R. Mukkamala, A. Avolio, K. G. Kyriakoulis, S. Mieke, A. Murray, G. Parati, A. E. Schutte, J. E. Sharman, R. Asmar et al., “Cuffless blood pressure measuring devices: review and statement by the european society of hypertension

- working group on blood pressure monitoring and cardiovascular variability,” Journal of hypertension, vol. 40, no. 8, pp. 1449–1460, 2022.
- [119] “Ieee standard for wearable cuffless blood pressure measuring devices,” IEEE Std 1708-2014, pp. 1–38, 2014.
  - [120] G. S. Stergiou, A. P. Avolio, P. Palatini, K. G. Kyriakoulis, A. E. Schutte, S. Mieke, A. Kollias, G. Parati, R. Asmar, N. Pantazis et al., “European society of hypertension recommendations for the validation of cuffless blood pressure measuring devices: European society of hypertension working group on blood pressure monitoring and cardiovascular variability,” Journal of Hypertension, pp. 10–1097, 2023.
  - [121] T. Weber, A. D. Protogerou, M. Agharazii, A. Argyris, S. Aoun Bahous, J. R. Bane-gas, R. K. Binder, J. Blacher, A. Araujo Brandao, J. J. Cruz et al., “Twenty-four-hour central (aortic) systolic blood pressure: reference values and dipping patterns in untreated individuals,” Hypertension, vol. 79, no. 1, pp. 251–260, 2022.
  - [122] A. E. Schutte, A. Kollias, and G. S. Stergiou, “Blood pressure and its variability: classic and novel measurement techniques,” Nature Reviews Cardiology, vol. 19, no. 10, pp. 643–654, 2022.
  - [123] D. Di Raimondo, G. Musiari, and A. Pinto, “Nocturnal blood pressure patterns and cardiac damage: there is still much to learn,” Hypertension Research, vol. 43, no. 3, pp. 246–248, 2020.
  - [124] G. Parati, J. E. Ochoa, C. Lombardi, and G. Bilo, “Assessment and management of blood-pressure variability,” Nature Reviews Cardiology, vol. 10, no. 3, pp. 143–155, 2013.
  - [125] G. Parati, C. Torlasco, M. Pengo, G. Bilo, and J. E. Ochoa, “Blood pressure variability: its relevance for cardiovascular homeostasis and cardiovascular diseases,” Hypertension Research, vol. 43, no. 7, pp. 609–620, 2020.
  - [126] E. O’Brien, K. Kario, J. A. Staessen, A. de la Sierra, and T. Ohkubo, “Patterns of ambulatory blood pressure: clinical relevance and application,” The Journal of Clinical Hypertension, vol. 20, no. 7, pp. 1112–1115, 2018.
  - [127] A. Pal, F. Martinez, A. P. Aguila, M. A. Akey, R. Chatterjee, M. G. E. Conserman, R. S. Aysola, L. A. Henderson, and P. M. Macey, “Beat-to-beat blood pressure variability in patients with obstructive sleep apnea,” Journal of Clinical Sleep Medicine, vol. 17, no. 3, pp. 381–392, 2021.

- [128] C. L. Phillips and D. M. O'Driscoll, "Hypertension and obstructive sleep apnea," Nature and science of sleep, pp. 43–52, 2013.
- [129] M. Dani, A. Dirksen, P. Taraborrelli, D. Panagopolous, M. Torocastro, R. Sutton, and P. B. Lim, "Orthostatic hypotension in older people: considerations, diagnosis and management," Clinical Medicine, vol. 21, no. 3, p. e275, 2021.
- [130] R. Freeman, W. Wieling, F. B. Axelrod, D. G. Benditt, E. Benarroch, I. Biaggioni, W. P. Cheshire, T. Chelimsky, P. Cortelli, C. H. Gibbons et al., "Consensus statement on the definition of orthostatic hypotension, neurally mediated syncope and the postural tachycardia syndrome," Autonomic Neuroscience, vol. 161, no. 1-2, pp. 46–48, 2011.
- [131] K. Wielemborek-Musial, K. Szmigielska, J. Leszczynska, A. Jegier et al., "Blood pressure response to submaximal exercise test in adults," BioMed research international, vol. 2016, 2016.
- [132] T. Ohkubo, A. Hozawa, K. Nagaie, M. Kikuya, I. Tsujia, S. Itoc, H. Satoh, S. Hisamichi, and Y. Imaid, "Prediction of stroke by ambulatory blood pressure monitoring versus screening blood pressure measurements in a general population: the ohasama study," Journal of hypertension, vol. 18, no. 7, pp. 847–854, 2000.
- [133] E. Dolan, A. Stanton, L. Thijs, K. Hinedi, N. Atkins, S. McClory, E. D. Hond, P. McCormack, J. A. Staessen, and E. O'Brien, "Superiority of ambulatory over clinic blood pressure measurement in predicting mortality: the dublin outcome study," Hypertension, vol. 46, no. 1, pp. 156–161, 2005.
- [134] J. Trinder, J. Kleiman, M. Carrington, S. Smith, S. Breen, N. Tan, and Y. Kim, "Autonomic activity during human sleep as a function of time and sleep stage," Journal of sleep research, vol. 10, no. 4, pp. 253–264, 2001.
- [135] W.-H. Lin, H. Wang, O. W. Samuel, G. Liu, Z. Huang, and G. Li, "New photoplethysmogram indicators for improving cuffless and continuous blood pressure estimation accuracy," Physiological measurement, vol. 39, no. 2, p. 025005, 2018.
- [136] G. Slapničar, M. Luštrek, and M. Marinko, "Continuous blood pressure estimation from ppg signal," Informatica, vol. 42, no. 1, 2018.
- [137] Ö. F. Ertuğrul and N. Sezgin, "A noninvasive time-frequency-based approach to estimate cuffless arterial blood pressure," Turkish Journal of Electrical Engineering and Computer Sciences, vol. 26, no. 5, pp. 2260–2274, 2018.

- [138] C. Yan, Z. Li, W. Zhao, J. Hu, D. Jia, H. Wang, and T. You, “Novel deep convolutional neural network for cuff-less blood pressure measurement using ecg and ppg signals,” in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2019, pp. 1917–1920.
- [139] S. Baek, J. Jang, and S. Yoon, “End-to-end blood pressure prediction via fully convolutional networks,” IEEE Access, vol. 7, pp. 185 458–185 468, 2019.
- [140] A. Fasano and V. Villani, “Baseline wander removal in ecg and aha recommendations,” in Computing in Cardiology 2013. IEEE, 2013, pp. 1171–1174.
- [141] C. Chourpiliadis and A. Bhardwaj, “Physiology, respiratory rate,” <https://www.ncbi.nlm.nih.gov/books/NBK537306/>, 2022, accessed: 2025-03-10.
- [142] Y. Liang, M. Elgendi, Z. Chen, and R. Ward, “An optimal filter for short photoplethysmogram signals,” Scientific data, vol. 5, no. 1, pp. 1–12, 2018.
- [143] M. Elgendi, R. Fletcher, Y. Liang, N. Howard, N. H. Lovell, D. Abbott, K. Lim, and R. Ward, “The use of photoplethysmography for assessing hypertension,” NPJ digital medicine, vol. 2, no. 1, p. 60, 2019.
- [144] E. Mejia-Mejia, J. Allen, K. Budidha, C. El-Hajj, P. A. Kyriacou, and P. H. Charlton, “Photoplethysmography signal processing and synthesis,” in Photoplethysmography. Elsevier, 2022, pp. 69–146.
- [145] M. Á. Goda, P. H. Charlton, and J. A. Behar, “pyppg: a python toolbox for comprehensive photoplethysmography signal analysis,” Physiological Measurement, vol. 45, no. 4, p. 045001, 2024.
- [146] P. H. Charlton, D. A. Birrenkott, T. Bonnici, M. A. Pimentel, A. E. Johnson, J. Alastruey, L. Tarassenko, P. J. Watkinson, R. Beale, and D. A. Clifton, “Breathing rate estimation from the electrocardiogram and photoplethysmogram: A review,” IEEE reviews in biomedical engineering, vol. 11, pp. 2–20, 2017.
- [147] C. Carlson, V.-R. Turpin, A. Suliman, C. Ade, S. Warren, and D. E. Thompson, “Bed-based ballistocardiography dataset,” 2020. [Online]. Available: <https://dx.doi.org/10.21227/77hc-py84>
- [148] H. Sedghamiz, “Matlab implementation of pan tompkins ecg qrs detector,” Code Available at the File Exchange Site of MathWorks, 2014.

- [149] D. L. Eckberg, C. Nerhed, and B. Wallin, “Respiratory modulation of muscle sympathetic and vagal cardiac outflow in man.” The Journal of physiology, vol. 365, no. 1, pp. 181–196, 1985.
- [150] F. Shaffer and J. P. Ginsberg, “An overview of heart rate variability metrics and norms,” Frontiers in public health, vol. 5, p. 258, 2017.
- [151] J.-M. Grégoire, C. Gilon, S. Carlier, and H. Bersini, “Autonomic nervous system assessment using heart rate variability,” Acta cardiologica, vol. 78, no. 6, pp. 648–662, 2023.
- [152] X. Ding and Y.-T. Zhang, “Pulse transit time technique for cuffless unobtrusive blood pressure measurement: from theory to algorithm,” Biomedical engineering letters, vol. 9, pp. 37–52, 2019.
- [153] C. Varon, J. Lázaro, J. Bolea, A. Hernando, J. Aguiló, E. Gil, S. Van Huffel, and R. Bailón, “Unconstrained estimation of hrv indices after removing respiratory influences from heart rate,” IEEE journal of biomedical and health informatics, vol. 23, no. 6, pp. 2386–2397, 2018.
- [154] M. Ramsey, “Blood pressure monitoring: automated oscillometric devices,” Journal of clinical monitoring, vol. 7, pp. 56–67, 1991.
- [155] F. Yasuma and J.-i. Hayano, “Respiratory sinus arrhythmia: why does the heartbeat synchronize with respiratory rhythm?” Chest, vol. 125, no. 2, pp. 683–690, 2004.
- [156] R. Cassani, P. Mejia, J. A. Tavares, J. C. Sanchez, and R. Martinez, “Adaptive filtering for respiration influence reduction on heart rate variability,” in 2011 8th International Conference on Electrical Engineering, Computing Science and Automatic Control. IEEE, 2011, pp. 1–5.
- [157] M. Kircher, G. Lenis, and O. Dössel, “Separating the effect of respiration from the heart rate variability for cases of constant harmonic breathing,” Current Directions in Biomedical Engineering, vol. 1, no. 1, pp. 46–49, 2015.
- [158] J. Lázaro, R. Bailón, P. Laguna, V. Marozas, A. Rapalis, and E. Gil, “Difference in pulse arrival time at forehead and at finger as a surrogate of pulse transit time,” in 2016 Computing in Cardiology Conference (CinC). IEEE, 2016, pp. 269–272.
- [159] M. Saeed, M. Villarroel, A. T. Reisner, G. Clifford, L.-W. Lehman, G. Moody, T. Heldt, T. H. Kyaw, B. Moody, and R. G. Mark, “Multiparameter intelligent

- monitoring in intensive care ii: a public-access intensive care unit database,” Critical care medicine, vol. 39, no. 5, pp. 952–960, 2011.
- [160] A. E. Johnson, T. J. Pollard, L. Shen, L.-w. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. Anthony Celi, and R. G. Mark, “Mimic-iii, a freely accessible critical care database,” Scientific data, vol. 3, no. 1, pp. 1–9, 2016.
  - [161] K. M. M. H. Kachuee, Mohamad and M. Shabany, “Cuff-Less Blood Pressure Estimation,” UCI Machine Learning Repository, 2015, DOI: <https://doi.org/10.24432/C5B602>.
  - [162] T. Kuusela et al., “Methodological aspects of heart rate variability analysis,” Heart rate variability (HRV) signal analysis: Clinical applications, pp. 10–42, 2013.
  - [163] G. Hotho, D. von Bonin, D. Krüerke, U. Wolf, and D. Cysarz, “Unexpected cardiovascular oscillations at 0.1 hz during slow speech guided breathing (om chanting) at 0.05 hz,” Frontiers in physiology, vol. 13, p. 875583, 2022.
  - [164] H. R. K. Putra, N. Yudistira, and T. N. Fatyanosa, “Variational mode decomposition and linear embeddings are what you need for time-series forecasting,” arXiv preprint arXiv:2408.16122, 2024.
  - [165] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” IEEE transactions on signal processing, vol. 62, no. 3, pp. 531–544, 2013.
  - [166] S. He and M. Bolić, “Novel metrics for tracking blood pressure changes incontinuous cuffless blood pressure estimations,” Scientific Reports, vol. 14, no. 1, p. 27478, 2024.
  - [167] H. Sakoe and S. Chiba, “Dynamic programming algorithm optimization for spoken word recognition,” IEEE transactions on acoustics, speech, and signal processing, vol. 26, no. 1, pp. 43–49, 1978.
  - [168] H. Alt and M. Godau, “Computing the fréchet distance between two polygonal curves,” International Journal of Computational Geometry & Applications, vol. 5, no. 01n02, pp. 75–91, 1995.
  - [169] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, “Comparing images using the hausdorff distance,” IEEE Transactions on pattern analysis and machine intelligence, vol. 15, no. 9, pp. 850–863, 1993.
  - [170] I. Lawrence and K. Lin, “A concordance correlation coefficient to evaluate reproducibility,” Biometrics, pp. 255–268, 1989.



- [171] M. Mukaka, “Statistics corner: a guide to appropriate use of correlation in medical research,” Malawi Med J, vol. 24, no. 3, pp. 69–71, 2012.
- [172] B. R. Overholser and K. M. Sowinski, “Biostatistics primer: part 2,” Nutrition in clinical practice, vol. 23, no. 1, pp. 76–84, 2008.
- [173] G. McBride et al., “A proposal for strength-of-agreement criteria for lin’s concordance correlation coefficient,” NIWA client report: HAM2005-062, vol. 45, pp. 307–310, 2005.
- [174] Y. Liu, K. S. Kim, R. M. Beresford, and D. H. Fleisher, “A generic composite measure of similarity between geospatial variables,” Ecological Informatics, vol. 60, p. 101169, 2020.
- [175] C. Cassisi, P. Montalto, M. Aliotta, A. Cannata, A. Pulvirenti et al., “Similarity measures and dimensionality reduction techniques for time series data mining,” Advances in data mining knowledge discovery and applications, pp. 71–96, 2012.
- [176] K. B. Pratt and E. Fink, “Search for patterns in compressed time series,” International Journal of Image and Graphics, vol. 2, no. 01, pp. 89–106, 2002.
- [177] E. Keogh, S. Chu, D. Hart, and M. Pazzani, “Segmenting time series: A survey and novel approach. data mining in time series databases,” World Scientific, vol. 57, pp. 1–22, 2004.
- [178] U. Ramer, “An iterative procedure for the polygonal approximation of plane curves,” Computer graphics and image processing, vol. 1, no. 3, pp. 244–256, 1972.
- [179] E. Keogh, “Fast similarity search in the presence of longitudinal scaling in time series databases,” in Proceedings Ninth IEEE International Conference on Tools with Artificial Intelligence. IEEE, 1997, pp. 578–584.
- [180] M. Khatami and F. Akbarzadeh, “Algorithms for segmenting time series,” Global Analysis and Discrete Mathematics, vol. 3, no. 1, pp. 65–73, 2018.
- [181] D. WangDa et al., “Pattern distance of time series,” WIT Transactions on Information and Communication Technologies, vol. 29, 2003.
- [182] K. Zhao, M. A. Wulder, T. Hu, R. Bright, Q. Wu, H. Qin, Y. Li, E. Toman, B. Mallick, X. Zhang et al., “Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A bayesian ensemble algorithm,” Remote sensing of Environment, vol. 232, p. 111181, 2019.

- [183] S. Chen, B. Ma, and K. Zhang, “On the similarity metric and the distance metric,” Theoretical Computer Science, vol. 410, no. 24-25, pp. 2365–2376, 2009.
- [184] C. Landry, E. T. Hedge, R. L. Hughson, S. D. Peterson, and A. Arami, “Wearable physiological and blood pressure measurements during activities of daily living,” 2021. [Online]. Available: <https://dx.doi.org/10.21227/wysp-gt69>
- [185] M. Kachuee, M. Kiani, H. Mohammadzade, and M. Shabany, “Cuff-less blood pressure estimation,” UCI Machine Learning Repository, 2015, DOI: <https://doi.org/10.24432/C5B602>.
- [186] RudderStack, “What is generalization in machine learning?” <http://www.rudderstack.com/learn/machine-learning/generalization-in-machine-learning/>, 2024, accessed on 2025-05-07.
- [187] X.-R. Ding and Y.-T. Zhang, “Photoplethysmogram intensity ratio: A potential indicator for improving the accuracy of ptt-based cuffless blood pressure estimation,” in 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, 2015, pp. 398–401.
- [188] H. H. Asada, P. Shaltis, A. Reisner, S. Rhee, and R. C. Hutchinson, “Mobile monitoring with wearable photoplethysmographic biosensors,” IEEE engineering in medicine and biology magazine, vol. 22, no. 3, pp. 28–40, 2003.
- [189] C. Chua and C. Heneghan, “Continuous blood pressure monitoring using ecg and finger photoplethysmogram,” in 2006 International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2006, pp. 5117–5120.
- [190] E. C.-P. Chua, S. J. Redmond, G. McDarby, and C. Heneghan, “Towards using photo-plethysmogram amplitude to measure blood pressure during sleep,” Annals of biomedical engineering, vol. 38, no. 3, pp. 945–954, 2010.
- [191] S. Bai, J. Z. Kolter, and V. Koltun, “An empirical evaluation of generic convolutional and recurrent networks for sequence modeling,” arXiv preprint arXiv:1803.01271, 2018.
- [192] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

- [193] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 7132–7141.
- [194] Z. Liu, Y. Wang, S. Vaidya, F. Du, Y. Halpern, X. Liu, and M. Tegmark, "Kan: Kolmogorov-arnold networks," arXiv preprint arXiv:2404.19756, 2024.
- [195] A. D. M. Ibrahim, Z. Shang, and J.-E. Hong, "How resilient are kolmogorov-arnold networks in classification tasks? a robustness investigation," Applied Sciences, vol. 14, no. 22, p. 10173, 2024.
- [196] Blealtan, "efficient-kan: A pytorch implementation of "kolmogorov-arnold networks"," <https://github.com/Blealtan/efficient-kan>, Apr. 2024. [Online]. Available: <https://github.com/Blealtan/efficient-kan>
- [197] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in Advances in Neural Information Processing Systems 30 (NIPS 2017), 2017, pp. 5998–6008. [Online]. Available: <http://papers.nips.cc/paper/7181-attention-is-all-you-need.pdf>
- [198] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," arXiv preprint arXiv:1406.1078, 2014. [Online]. Available: <http://arxiv.org/abs/1406.1078>
- [199] Y. Liu, Z. Shao, and N. Hoffmann, "Global attention mechanism: Retain information to enhance channel-spatial interactions," arXiv preprint arXiv:2112.05561, 2021.
- [200] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 26, no. 1, pp. 43–49, 1978.
- [201] M. Cuturi and M. Blondel, "Soft-dtw: a differentiable loss function for time-series," in Proceedings of the 34th International Conference on Machine Learning - Volume 70, ser. ICML'17. JMLR.org, 2017, pp. 894–903. [Online]. Available: <http://proceedings.mlr.press/v70/cuturi17a.html>
- [202] C. Carlson, V.-R. Turpin, A. Suliman, C. Ade, S. Warren, and D. E. Thompson, "Bed-based ballistocardiography: Dataset and ability to track cardiovascular parameters," Sensors, vol. 21, no. 1, p. 156, 2020.

- [203] B. Moody, M. Craig, A. Johnson, T. Kyaw, G. Moody, M. Saeed, and M. Villarroel, “The mimic-iii waveform database matched subset, physionet. org, 2017. doi: 10.13026.”
- [204] S. Abdullah, A. Hafid, M. Folke, M. Lindén, and A. Kristoffersson, “Ppgfeat: a novel matlab toolbox for extracting ppg fiducial points,” Frontiers in Bioengineering and Biotechnology, vol. 11, p. 1199604, 2023.
- [205] P. H. Charlton, J. Mariscal Harana, S. Vennin, Y. Li, P. Chowienczyk, and J. Alastruey, “Modeling arterial pulse waves in healthy aging: a database for in silico evaluation of hemodynamics and pulse wave indexes,” American Journal of Physiology-Heart and Circulatory Physiology, vol. 317, no. 5, pp. H1062–H1085, 2019.
- [206] S. He, H. R. Dajani, and M. Bolic, “Novel cuffless blood pressure estimation method using a bayesian hierarchical model,” in 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2021, pp. 898–901.
- [207] A. Gelman, “Multilevel (hierarchical) modeling: what it can and cannot do,” Technometrics, vol. 48, no. 3, pp. 432–435, 2006.
- [208] O. Martin, Bayesian Analysis with Python: Introduction to statistical modeling and probabilistic programming using PyMC3 and ArviZ. Packt Publishing Ltd, 2018.
- [209] C. Finn, P. Abbeel, and S. Levine, “Model-agnostic meta-learning for fast adaptation of deep networks,” in International conference on machine learning. PMLR, 2017, pp. 1126–1135.
- [210] Y. Zou, F. Liu, and Q. Li, “Unraveling model-agnostic meta-learning via the adaptation learning rate,” in International Conference on Learning Representations, 2022. [Online]. Available: <https://openreview.net/forum?id=3rULBvOJ8D2>
- [211] A. Nichol, J. Achiam, and J. Schulman, “On first-order meta-learning algorithms,” arXiv preprint arXiv:1803.02999, 2018.
- [212] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. March, and V. Lempitsky, “Domain-adversarial training of neural networks,” Journal of machine learning research, vol. 17, no. 59, pp. 1–35, 2016.

- [213] W. J. Krzanowski, Principles of multivariate analysis. Oxford University Press, 2000.
- [214] J. C. Gower, “Generalized procrustes analysis,” Psychometrika, vol. 40, no. 1, pp. 33–51, 1975.
- [215] VitalDB, “Vitaldb database,” <https://vitaldb.net/dataset/?query=overview&documentId=13qqajnNZzkN7NZ9aXnaQ-47NWy7kx-a6gbrcEsi-gak&gionId=h.foizq7qqcyzk>.
- [216] A. Chandrasekhar, M. Yavarimanesh, K. Natarajan, J.-O. Hahn, and R. Mukkamala, “Ppg sensor contact pressure should be taken into account for cuff-less blood pressure measurement,” IEEE Transactions on Biomedical Engineering, vol. 67, no. 11, pp. 3134–3140, 2020.
- [217] D. Liu, M. Görges, and S. A. Jenkins, “University of queensland vital signs dataset: development of an accessible repository of anesthesia patient monitoring data for research,” Anesthesia & Analgesia, vol. 114, no. 3, pp. 584–589, 2012.
- [218] B. Moody, G. Moody, M. Villarroel, G. D. Clifford, and I. Silva, “Mimic-iii waveform database,” Apr 2020. [Online]. Available: <https://physionet.org/content/mimic3wdb/1.0/>
- [219] E. Seitsonen, I. Korhonen, M. Van Gils, M. Huiku, J. Lötjönen, K. Korttila, and A. Yli-Hankala, “Eeg spectral entropy, heart rate, photoplethysmography and motor responses to skin incision during sevoflurane anaesthesia,” Acta Anaesthesiologica Scandinavica, vol. 49, no. 3, pp. 284–292, 2005.
- [220] L. Wang, E. Pickwell-MacPherson, Y. Liang, and Y. T. Zhang, “Noninvasive cardiac output estimation using a novel photoplethysmogram index,” in 2009 annual international conference of the IEEE engineering in medicine and biology society. IEEE, 2009, pp. 1746–1749.
- [221] W. B. Murray and P. A. Foster, “The peripheral pulse wave: information overlooked,” Journal of clinical monitoring, vol. 12, no. 5, pp. 365–377, 1996.
- [222] J. Dorlas and J. Nijboer, “Photo-electric plethysmography as a monitoring device in anaesthesia: application and interpretation,” British journal of anaesthesia, vol. 57, no. 5, pp. 524–530, 1985.

- [223] I. Korhonen and A. Yli-Hankala, “Photoplethysmography and nociception,” Acta Anaesthesiologica Scandinavica, vol. 53, no. 8, pp. 975–985, 2009.
- [224] U. Rubins, A. Grabovskis, J. Grube, and I. Kukulis, “Photoplethysmography analysis of artery properties in patients with cardiovascular diseases,” in 14th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics. Springer, 2008, pp. 319–322.
- [225] M. Elgendi, “On the analysis of fingertip photoplethysmogram signals,” Current cardiology reviews, vol. 8, no. 1, pp. 14–25, 2012.

# APPENDICES

# Appendix A

## Database, Signal Preprocessing and Feature Extraction

### A.1 Research Database

Currently, there are several open access database that includes BP waveforms together with several different physiological waveforms, such as ECG, PPG, respiration. A summary of some popular databases can be found in Table [A.1](#).

However, the drawbacks of these databases limit the use for accurate cuffless BP estimation. As described by MIMIC-III waveform database, the waveform data contain unspecified/unknown filtering delays and/or inter channels delays and there may be a changing time delay of up to 500 ms between different waveforms in the data. For example, Figure [A.1](#) shows the ECG and PPG waveforms from MIMIC-II database and collected by our devices respectively. As there are no physical filters employed in our data collection system and both ECG and PPG were pre-processed by a  $3^{rd}$  order Butterworth bandpass filter with different cutoff frequencies, it can be considered there is no extra time delay for ECG and PPG at data collection stage. Comparing Figure [A.1](#) (a) and Figure [A.1](#) (b), it can be observed that the PATs (time differences between ECG R peaks and PPG systolic peaks) are smaller than regular PATs collected using a system without additional delays which indicates the PAT collected from these databases may be unreliable for PAT-based cuffless BP estimation [\[218\]](#).



Table A.1: Four widely used open access databases for cuffless BP estimation

| Database                                              | Population                            | Sampling Rate     | Contents                                                                                                                                                                                                                      | Length                |
|-------------------------------------------------------|---------------------------------------|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| MIMIC-III waveform database [203]                     | 30,000 ICU patients                   | 125 Hz (waveform) | Waveform (up to 8) <ul style="list-style-type: none"> <li>• PPG</li> <li>• ECG</li> <li>• ABP</li> </ul> Numeric records <ul style="list-style-type: none"> <li>• HR</li> <li>• <math>SpO_2</math> ...</li> </ul>             | several hours         |
| VitalDB databank [215]                                | 6,388 surgical patients (non-cardiac) | 500 Hz (waveform) | Waveform <ul style="list-style-type: none"> <li>• ECG</li> <li>• PPG</li> <li>• ABP</li> </ul> Numeric record <ul style="list-style-type: none"> <li>• HR</li> <li>• <math>SpO_2</math></li> <li>• temperature ...</li> </ul> | several hours         |
| UCI Cuff-Less Blood Pressure Estimation Data Set [98] | 12,000 patients                       | 125 Hz            | <ul style="list-style-type: none"> <li>• ECG</li> <li>• PPG</li> <li>• ABP</li> </ul>                                                                                                                                         | Less than 10 minutes  |
| University of Queensland Vital Signs Dataset [217]    | 32 patients                           | 100 Hz (waveform) | Waveform <ul style="list-style-type: none"> <li>• ECG</li> <li>• PPG</li> <li>• ABP ...</li> </ul> Numeric record <ul style="list-style-type: none"> <li>• HR</li> <li>• temperature</li> <li>• RR ...</li> </ul>             | 13 minutes to 5 hours |

## A.2 Features Employed for Cuffless BP Estimation

Although there is a great number of features extracted and employed for machine-learning or deep-learning based cuffless BP prediction [78, 79], only parts of them are related to BP physiologically. In this section, several important features with physiological support are introduced.

### A.2.1 PPG Features

#### 1. Pulse Area (Area\_Sys, Area\_Dia and IPA)

The pulse area is measured as the total area under the PPG curve. The research [219] found the PPG area response to skin incision to differ between movers and non-

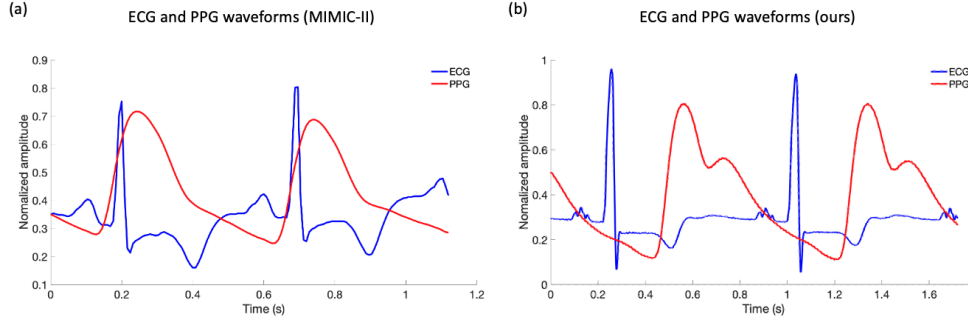


Figure A.1: (a) a demonstration of unknown time delay during data collection which causes unreliable PAT values (MIMIC database); (b) regular PAT extracted from ECG and PPG signals without extra time delay.

movers. In [220], the PPG pulse was divided into two areas at the dicrotic notch and the ratio of the two areas can be used as an indicator of total peripheral resistance, a demonstration of systole area (Area\_Sys, shown as A1 in Figure A.2(b)) and diastole area (Area\_Dia, shown as A2 in Figure A.2(b)) can be found in Figure A.2(b). This ratio is called the inflection point area ratio (IPA) and is defined as:

$$\text{IPA} = \frac{\text{Area\_Dia}}{\text{Area\_Sys}} \quad (\text{A.1})$$

## 2. Systolic Amplitude (Amp\_Sys)

As shown in Figure A.2(a), Systolic amplitude is the PPG value at systolic peak. It has been explored as an indicator of the pulsatile changes in blood volume caused by arterial blood flow around the measurement site in [188, 189]. In [221], systolic amplitude has been proved to be related to stroke volume which is proportional to blood pressure. Dorlas and Nijboer found that systolic amplitude is directly proportional to local vascular distensibility over a remarkably wide range of cardiac output [222]. It has also been observed that systolic amplitude is potentially a more suitable measure than pulse arrival time in continuous BP estimation [190]. However, the systolic amplitude of PPG can be affected by several different factors, such as relative elevation of measurement site and sympathetic activation (e.g. stress, cold) [223].

## 3. Diastolic Amplitude (Amp\_Dia)

Diastolic amplitude is the PPG value at diastolic peak, as shown in Figure A.2(a).

## 4. Augmentation Index (AI )

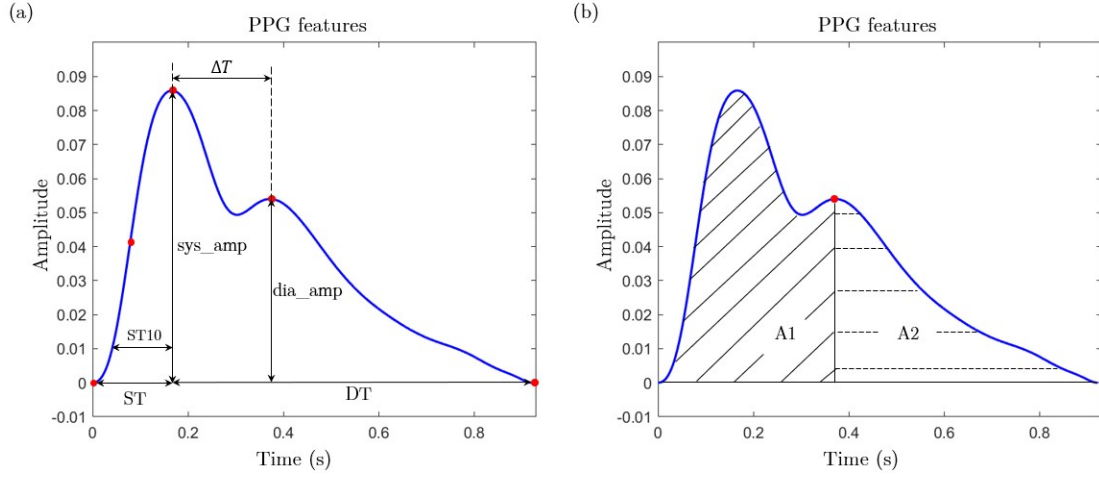


Figure A.2: The demonstration of PPG feature extraction

The augmentation pressure is the measure of the contribution that the wave reflection makes to the systolic arterial pressure, and it is obtained by measuring the reflected wave propagating from the periphery to the center. Reduced compliance of the elastic arteries in an earlier return of the 'reflected wave', which arrives in systole rather than in diastole, causing a disproportionate rise in systolic pressure and an increase in pulse pressure, with a consequent increase in ventricular after load and a decrease in diastolic blood pressure and impaired coronary perfusion. In this study, the augmentation index is defined as [224]:

$$AI = \frac{Amp\_Dia}{Amp\_Sys} \quad (A.2)$$

##### 5. $\Delta T$

As shown in Figure Figure A.2(a), the time delay between systolic and diastolic peaks (or, in the absence of a second peak, the point of inflection) is related to the transit time of pressure waves from the root of the subclavian artery to the apparent site of reflection and back to subclavian artery [225].

##### 6. Systolic Time (ST)

The systolic time is defined as the time interval from foot point of PPG pulse to the systolic peak, as shown in Figure A.2(a).

##### 7. Diastolic Time (DT)

The diastolic time is defined as the time interval from systolic peak to the end point of PPG pulse, as shown in Figure A.2(a).

### A.3 Examples of Respiratory Effects on Extracted Features

Examples of respiratory effects on extracted features and BPs for two subjects (subject1 and subject8 of BCG BP dataset) are shown in Figure A.3 to Figure A.14.

### A.4 Supplementary Material for Novel Evaluation Metrics

#### A.4.1 MLP Model for Example Shown in Figure 2.9(a)

A simple multilayer perceptron (MLP) model (hidden layer 1-4: 32-128-64-32 nodes, ReLU activation; output layer: 2 nodes, linear activation) was implemented to estimate SBP and DBP using 20 features extracted from ECG and PPG waveforms. A total of 147 subsets were selected from UCI cuffless BP estimation dataset [161], 117 subsets ( $\sim 80\%$ ) were used for training and 30 subsets ( $\sim 20\%$ ) were used for testing (no overlap between training and testing set). The mean and standard deviation of absolute difference ( $\text{MAD} \pm \text{SDAD}$ ) of the testing set is  $9.00 \pm 7.31$  and  $7.05 \pm 6.60$  mmHg for SBP and DBP, respectively. The  $\text{MAD} \pm \text{SDAD}$  of a 543 seconds' testing subset shown in Figure 1(a) is  $3.77 \pm 2.51$  mmHg for SBP. However, the model is unable to track one acute BP change occurred between 325 and 375 seconds.

#### A.4.2 Pearson Correlation Coefficient (PCC) and Concordance Correlation Coefficient (CCC)

The relation between CCC ( $\rho_c$ ) and PCC ( $\rho$ ) can be represented as:

$$\rho_c = C_b \cdot \rho \quad (\text{A.3})$$

where  $C_b = \frac{2}{v + \frac{1}{v} + u^2}$ ,  $v = \frac{\sigma_1}{\sigma_2}$  is a scale shift factor and  $u = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1 \sigma_2}}$  is a location shift factor relative to the scale.  $\sigma_{12}$  is the covariance of two variables,  $\text{BP}_{\text{est}}$  and  $\text{BP}_{\text{ref}}$ .  $\sigma_1$  and  $\sigma_2$  are

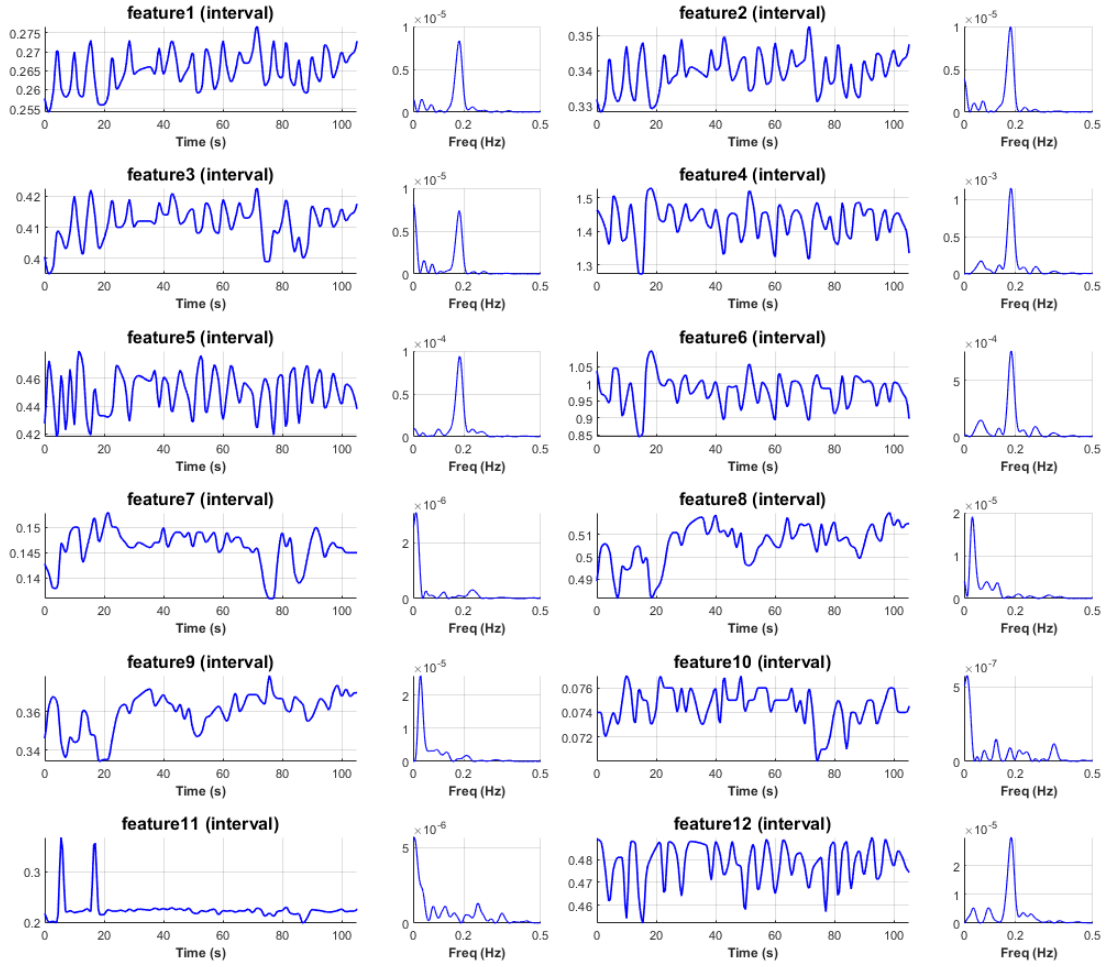


Figure A.3: The demonstration of extracted features without respiratory effect removal (feature1-feature12, exemplary on subject1 from BCG dataset).

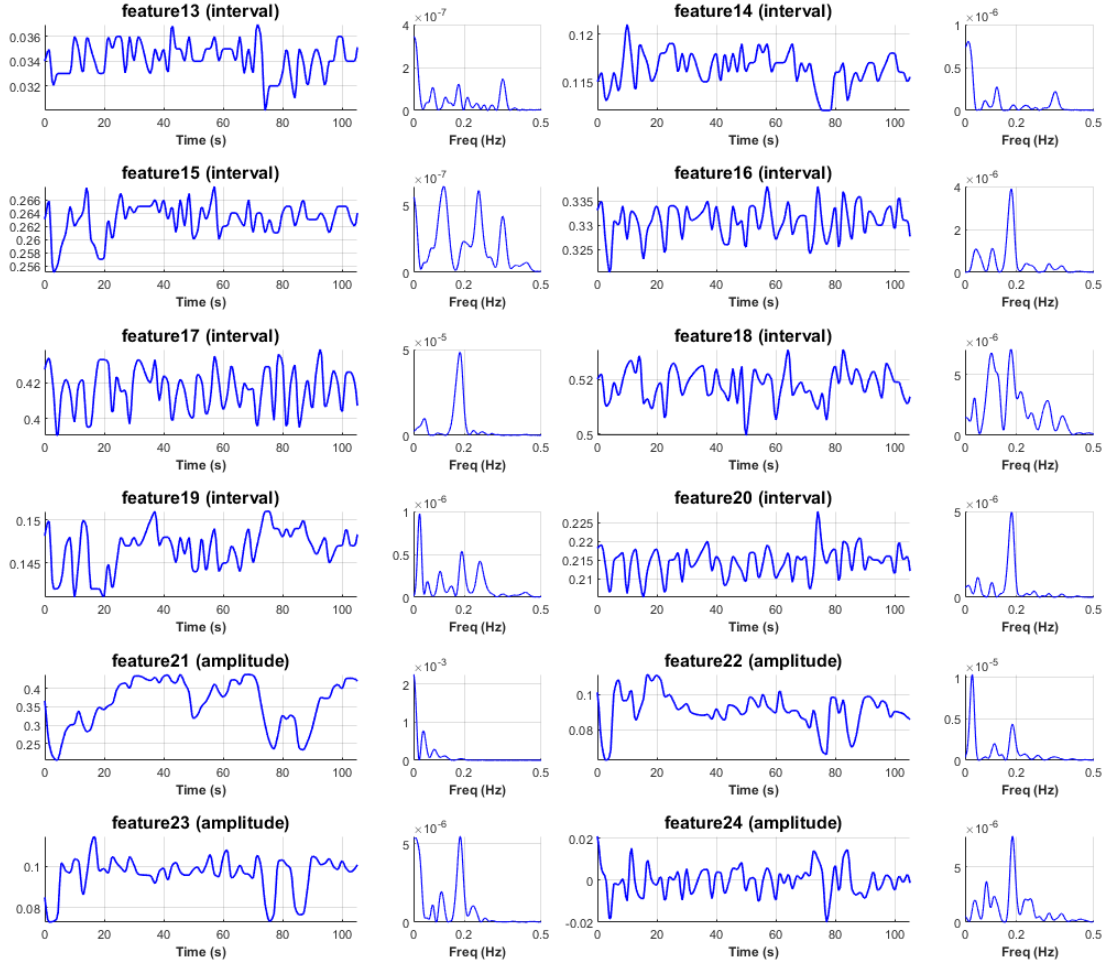


Figure A.4: The demonstration of extracted features without respiratory effect removal (feature13-feature24, exemplary on subject1 from BCG dataset).

the standard deviations of  $BP_{est}$  and  $BP_{ref}$ , respectively.  $\mu_1$  and  $\mu_2$  are the mean values of  $BP_{est}$  and  $BP_{ref}$ , respectively.

An experiment was implemented to demonstrate that CCC is more robust in detecting scale and location shifts than PCC.  $BP_{ref}$  is a reference measure of systolic blood pressure. The  $BP_{est}$  in Figure (a) to (d) to represent (1) location shift ( $BP_{est} = BP_{ref} - 2.5$ ), (b) scale shift ( $BP_{est} = 1.05 * BP_{ref}$ ), (c) location and scale shift ( $BP_{est} = 1.05 * BP_{ref} - 5$ ), and (d) no concordance ( $BP_{est} = \text{mean}(BP_{ref})$ ), as shown in Figure A.15.

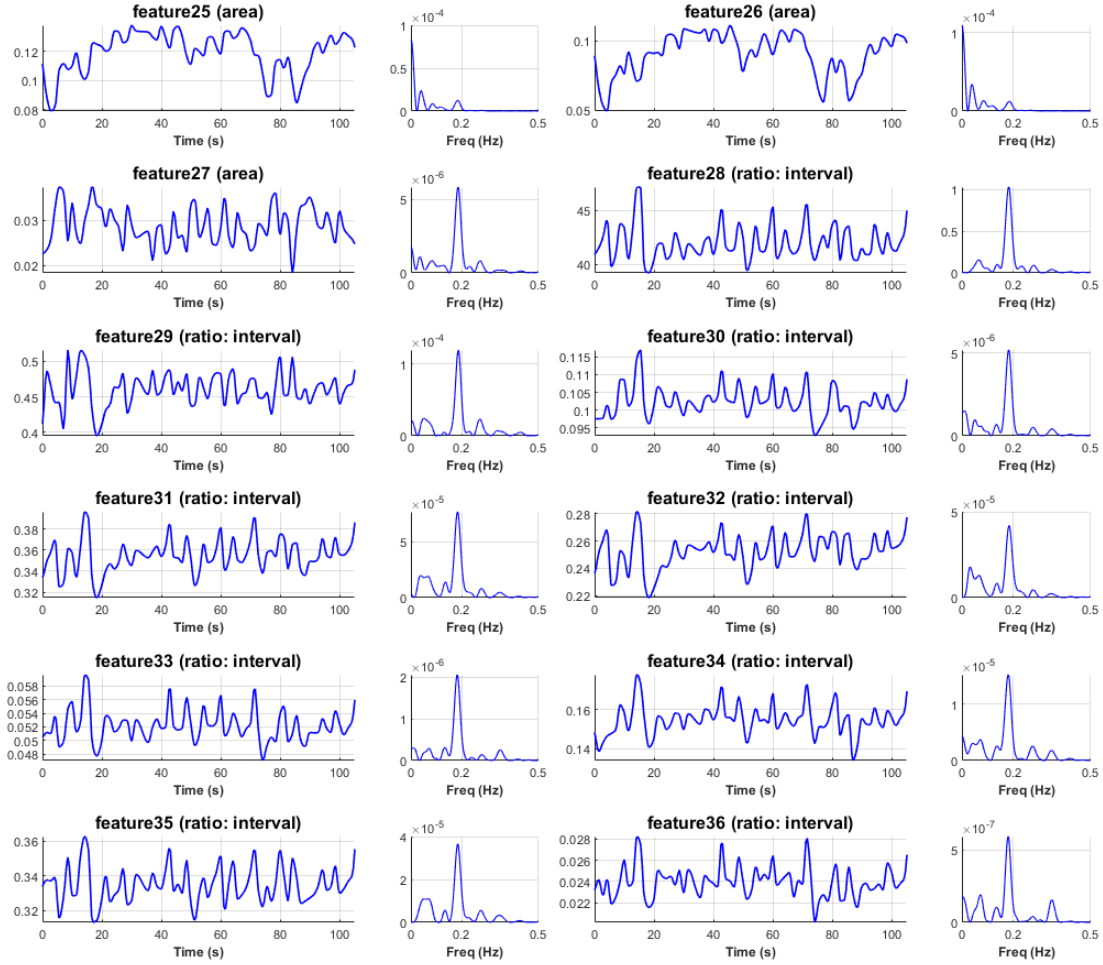


Figure A.5: The demonstration of extracted features without respiratory effect removal (feature25-feature36, exemplary on subject1 from BCG dataset).

### A.4.3 Re-implementation of SOTA Models

Two SOTA cuffless BP estimation models were selected and re-implemented using **Daila BP** [184]. Their results were used to demonstrate and compare the performance of conventional metrics and proposed metrics for BP pattern tracking.

Different training strategies were applied for ApproximateNet and DeepRNN-4L models with different training strategies and due to the fact that The model diagrams and parameters are demonstrated in Figure A.16 and experiment details are described below:

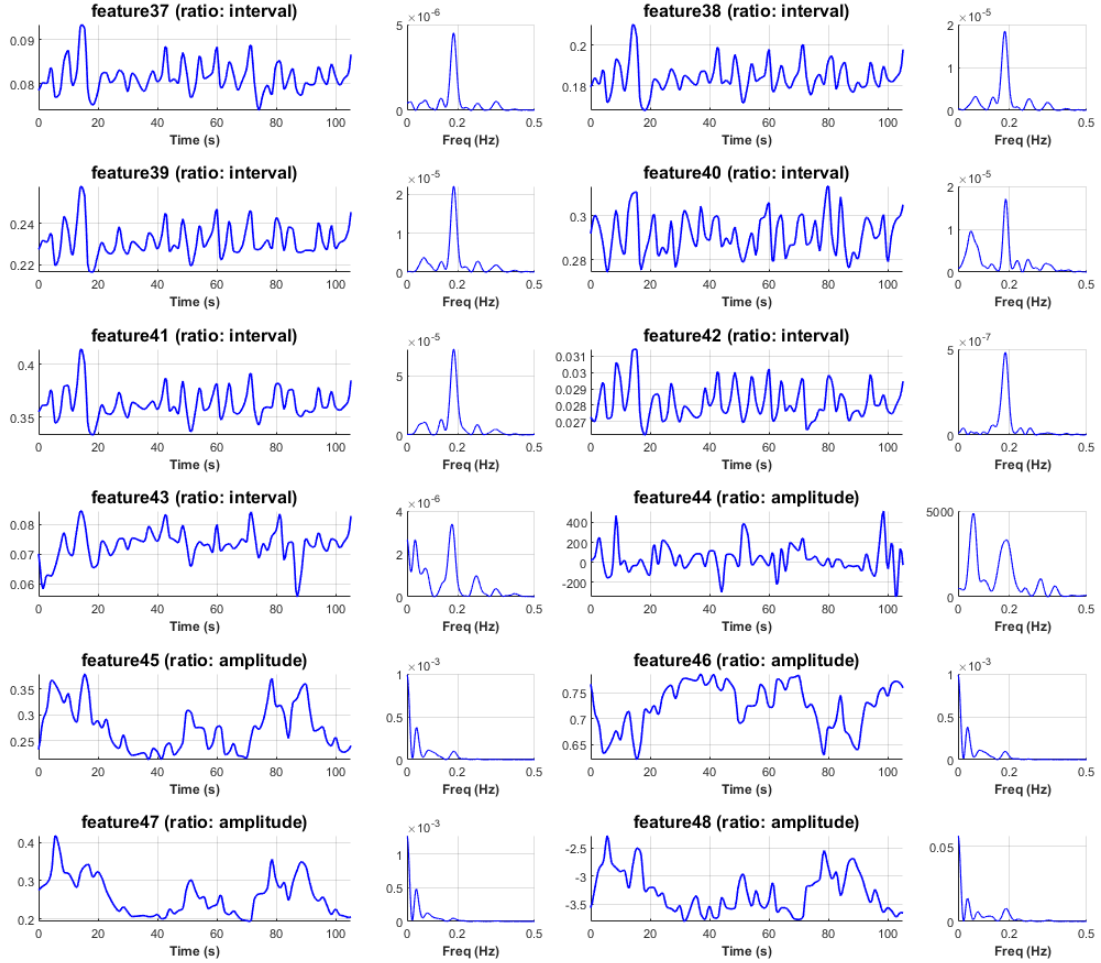


Figure A.6: The demonstration of extracted features without respiratory effect removal (feature37-feature48, exemplary on subject1 from BCG dataset).

- **Data preparation:** For each subject, an average of 2.3 hours of continuous ECG, PPG and BP waveforms were extracted from DailA BP. The sampling rate is 64 Hz and 2<sup>nd</sup> order Butterworth bandpass filters with pass band of 0.5-30 Hz and 0.5-10 Hz were applied to ECG and PPG signals for noise removal respectively.
- **ApproximateNet:** each subject's continuous PPG and BP waveforms were divided into a number of 16 seconds' ( $f_s=64$  Hz,  $16 \times 64 = 1024$  samples, same to the segment length of the original work [117]) segments. Five-fold cross validation was applied for model training (5 folds as DailA BP includes 5 subjects). For each iteration,



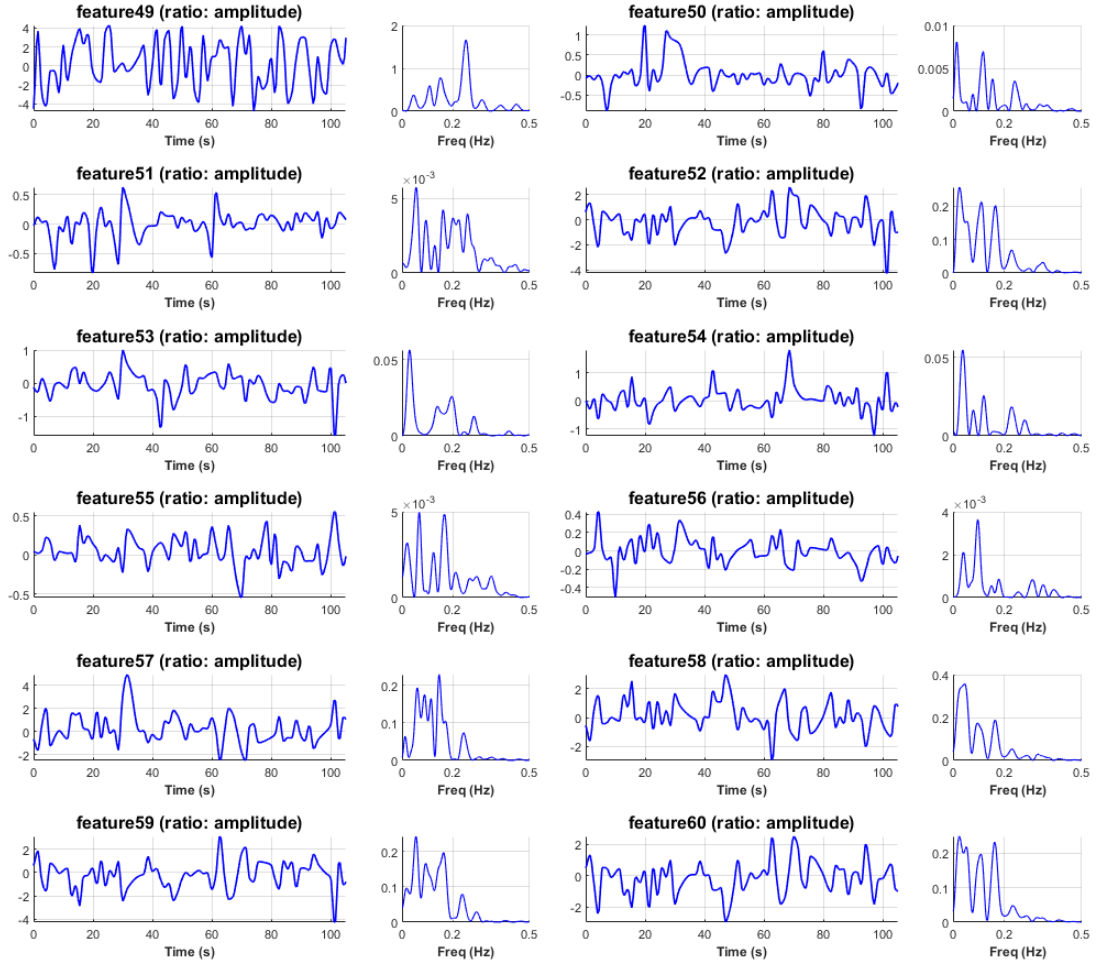


Figure A.7: The demonstration of extracted features without respiratory effect removal (feature49-feature60, exemplary on subject1 from BCG dataset).

the model was trained on four subjects and estimate BP for the left out subject. Then, the first three segments of reference BP waveform were used to calibrate the estimated BP waveforms,  $BP_{est} = BP_{est} - (\text{mean}(BP_{est_{1-3}}) - \text{mean}(BP_{ref_{1-3}}))$ . The SBP and DBP values were then detected from the calibrated BP waveforms and then resampled to uniformly sampled time series sequences ( $fs=1$  Hz). The  $MAD \pm SDAD$  is  $9.88 \pm 8.65$  and  $6.50 \pm 5.02$  mmHg for SBP and DBP respectively.

- **DeepRNN:** as the data of different subjects are independent, a personalized model was trained for each subject. Seven features (demonstrated in Appendix A.4.4)

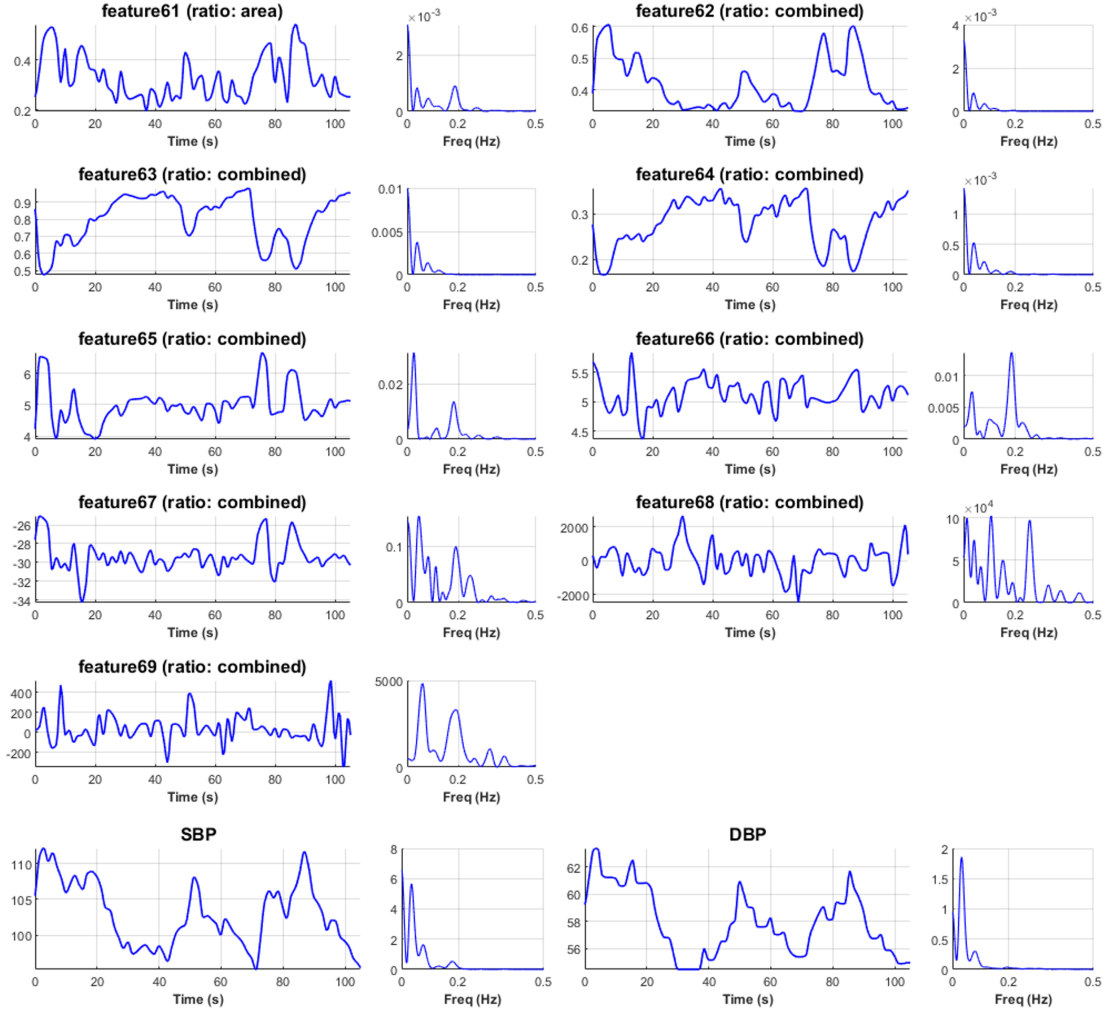


Figure A.8: The demonstration of extracted features, SBP and DBP without respiratory effect removal (feature61-feature69, SBP and DBP, exemplary on subject1 from BCG dataset).

were extracted from the first 25 minutes of ECG and PPG waveforms, as well as corresponding SBP and DBP values were used for model training, then predict SBP and DBP with the rest data. Then the delay between estimation and reference which is caused by the lagged training set was compensated, and the estimated SBP and DBP values were resampled to uniformly sampled time series sequences ( $f_s=1$  Hz) for performance evaluation. The  $MAD \pm SDAD$  is  $4.14 \pm 3.29$  and  $4.15 \pm 3.30$  mmHg

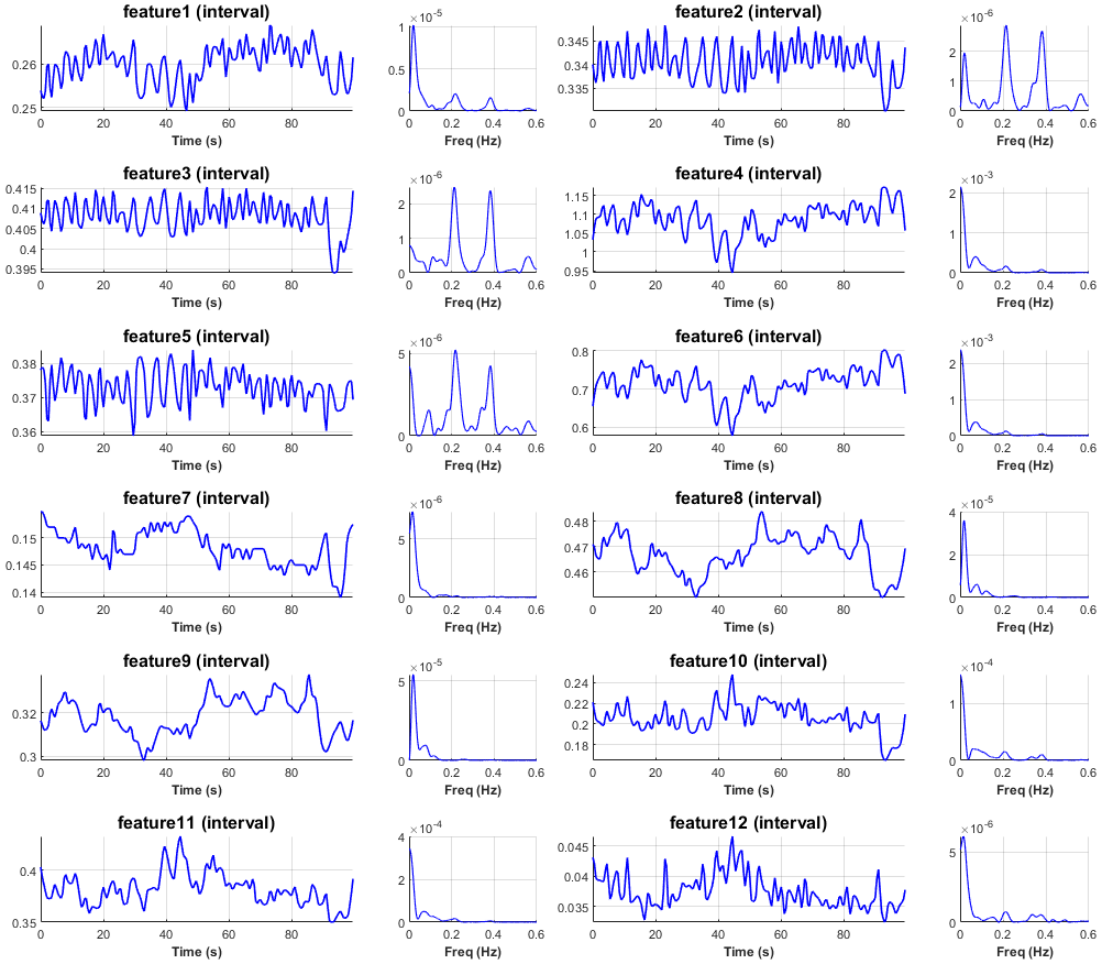


Figure A.9: The demonstration of extracted features without respiratory effect removal (feature1-feature12, exemplary on subject8 from BCG dataset).

for SBP and DBP respectively.

**\*Please note:** the purpose of this study is not to compare which model is better for cuffless BP estimation. Since two models were trained in different ways, their performance does not indicate DeepRNN-4L model outperforms ApproximateNet for BP estimation. In this study, we just intend to use different results of two SOTA models to demonstrate the workflow and performance of the proposed methods.

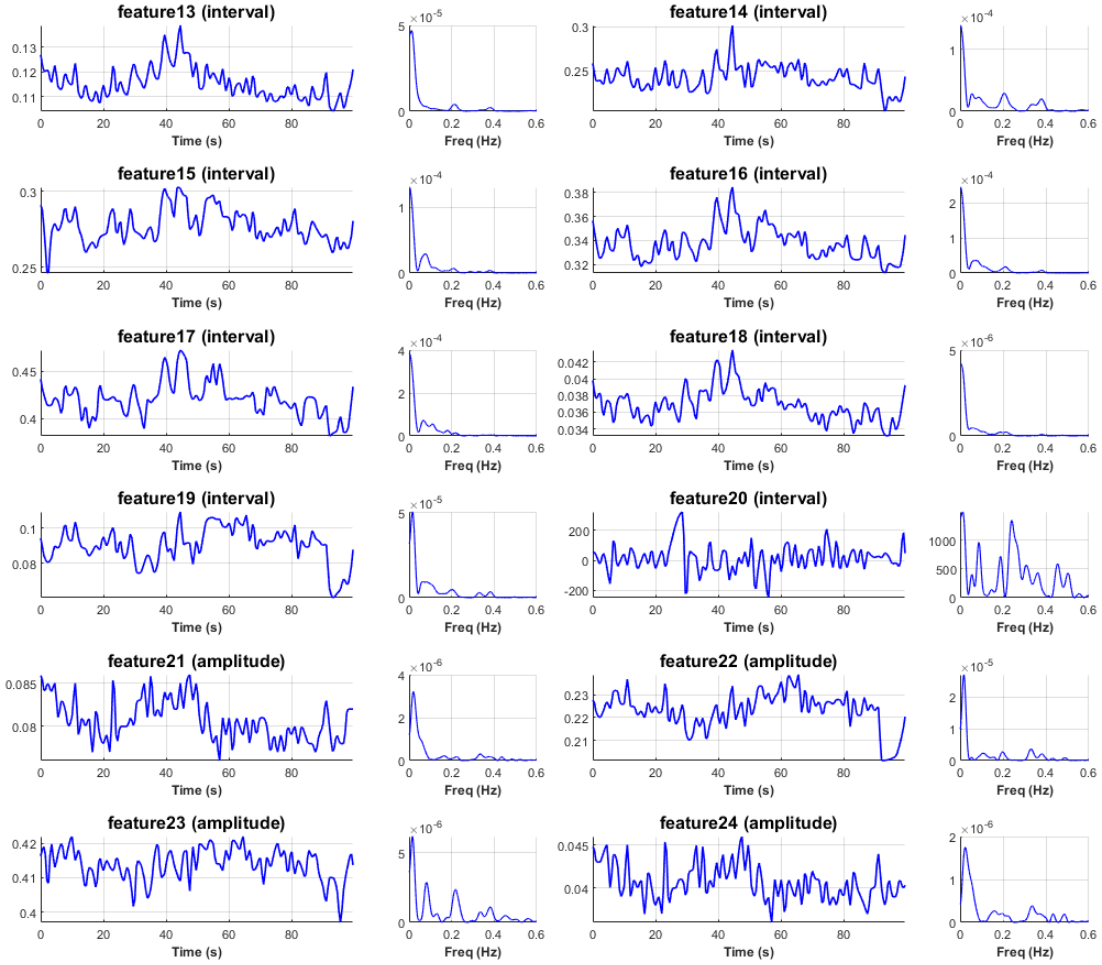


Figure A.10: The demonstration of extracted features without respiratory effect removal (feature13-feature24, exemplary on subject8 from BCG dataset).

#### A.4.4 Features Used for DeepRNN-4L

Seven features extracted from simultaneous ECG and PPG waveforms for BP estimation using DeepRNN-4L model is demonstrated in Figure A.17

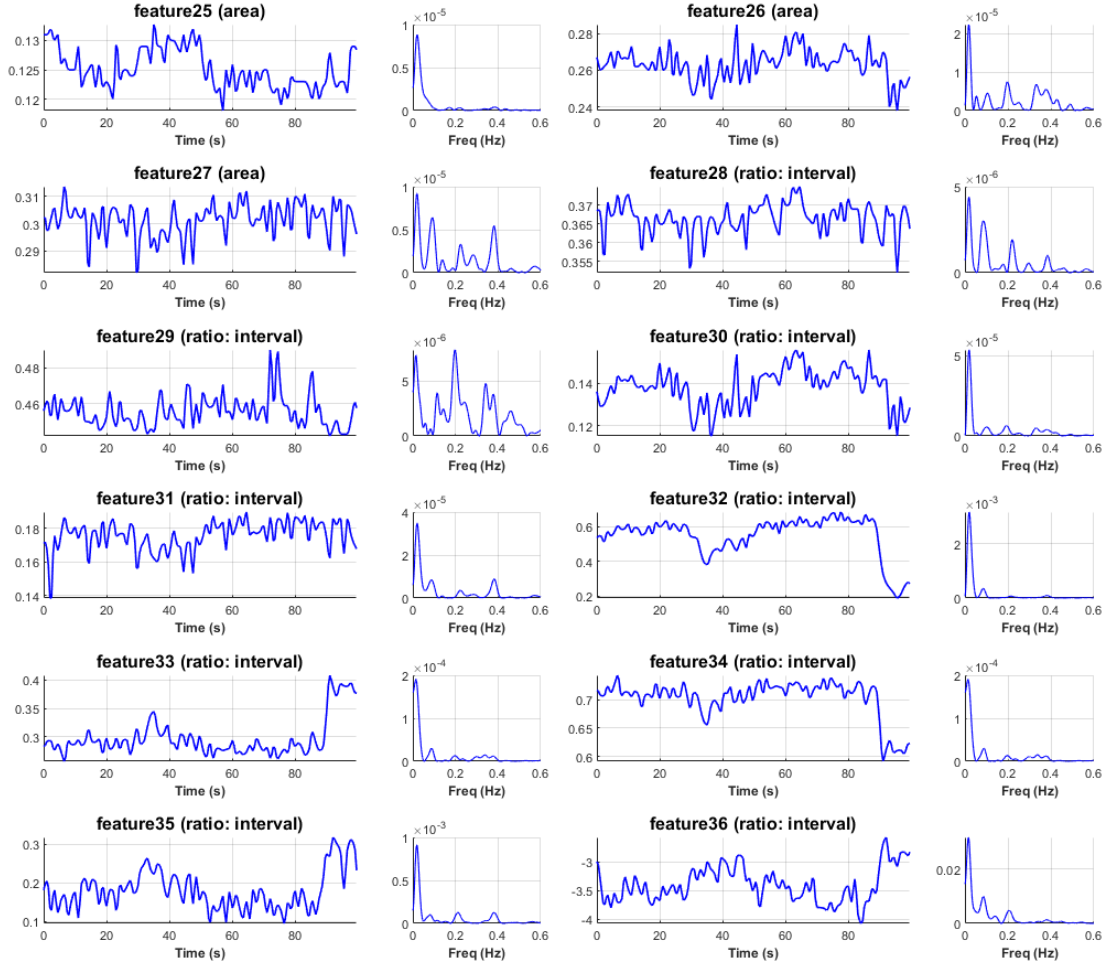


Figure A.11: The demonstration of extracted features without respiratory effect removal (feature25-feature36, exemplary on subject8 from BCG dataset).

#### A.4.5 Benchmark and Diagnose Estimation Algorithms

The proposed methods and metrics can be used to benchmark and diagnose the BP estimation algorithms. For example, two SOTA models, ApproximateNet [117] and DeepRNN-4L [84], were re-implemented with a same dataset **DailA BP** [184]. Therefore, the performance of these two estimation algorithms can be compared in terms of the similarities of distance, trend and the composite of both. For example, the percentages (the percentage is obtained by calculating the ratio between the duration of all segments meet the

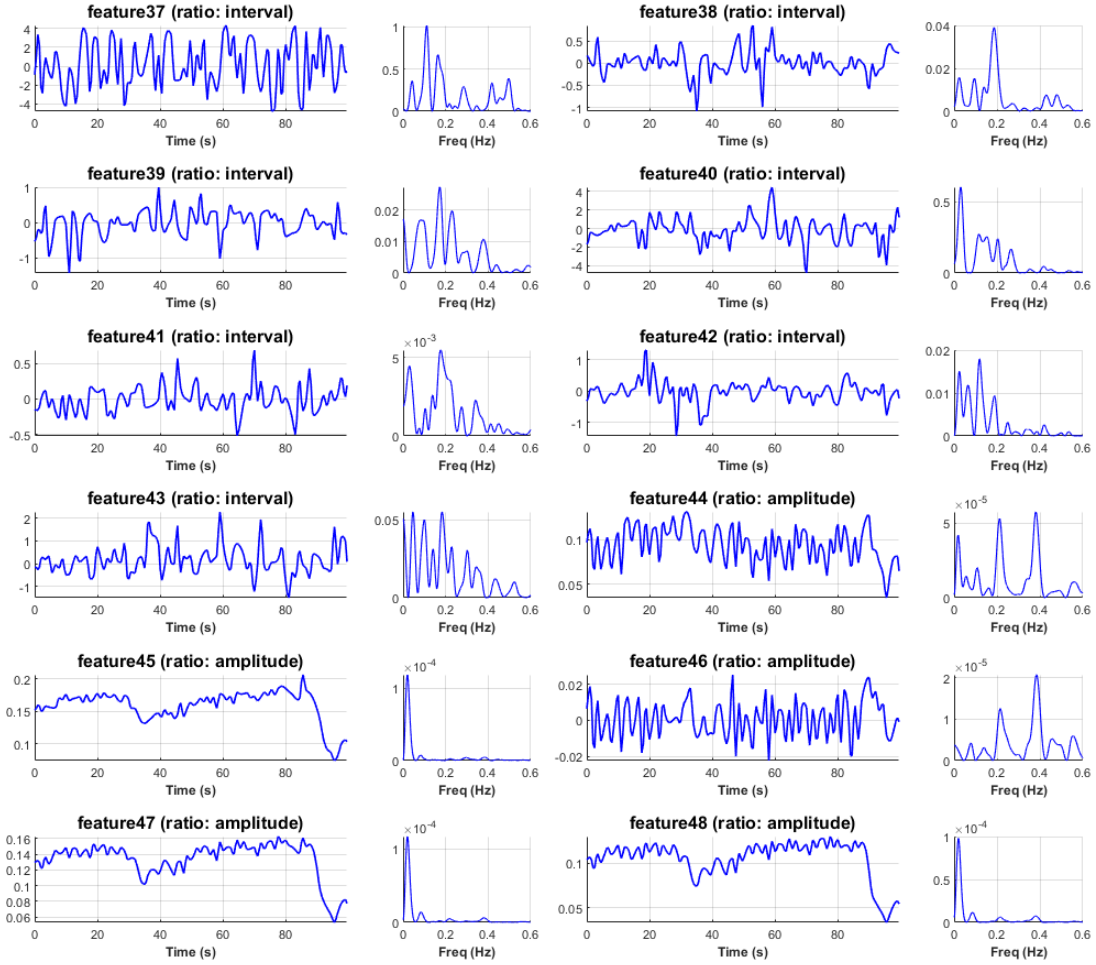


Figure A.12: The demonstration of extracted features without respiratory effect removal (feature37-feature48, exemplary on subject8 from BCG dataset).

requirement and the duration of the sequence) of segments with different MAD levels of two estimation models are shown in Figure A.18.

It can be observed that the majority of both SBP and DBP estimations obtained using DeepRNN-4L can be considered as "Grade A" of IEEE Standard (SBP:  $> 80\%$ , DBP:  $> 70\%$ ). However, only a small proportion of the estimations using ApproximateNet can be considered as "Grade A" of IEEE Standard (SBP:  $\approx 25\%$ , DBP:  $\approx 25\%$ ) and even  $\approx 50\%$  of SBP estimations using ApproximateNet dissatisfy IEEE Standard ( $MAD \geq 7$  mmHg).

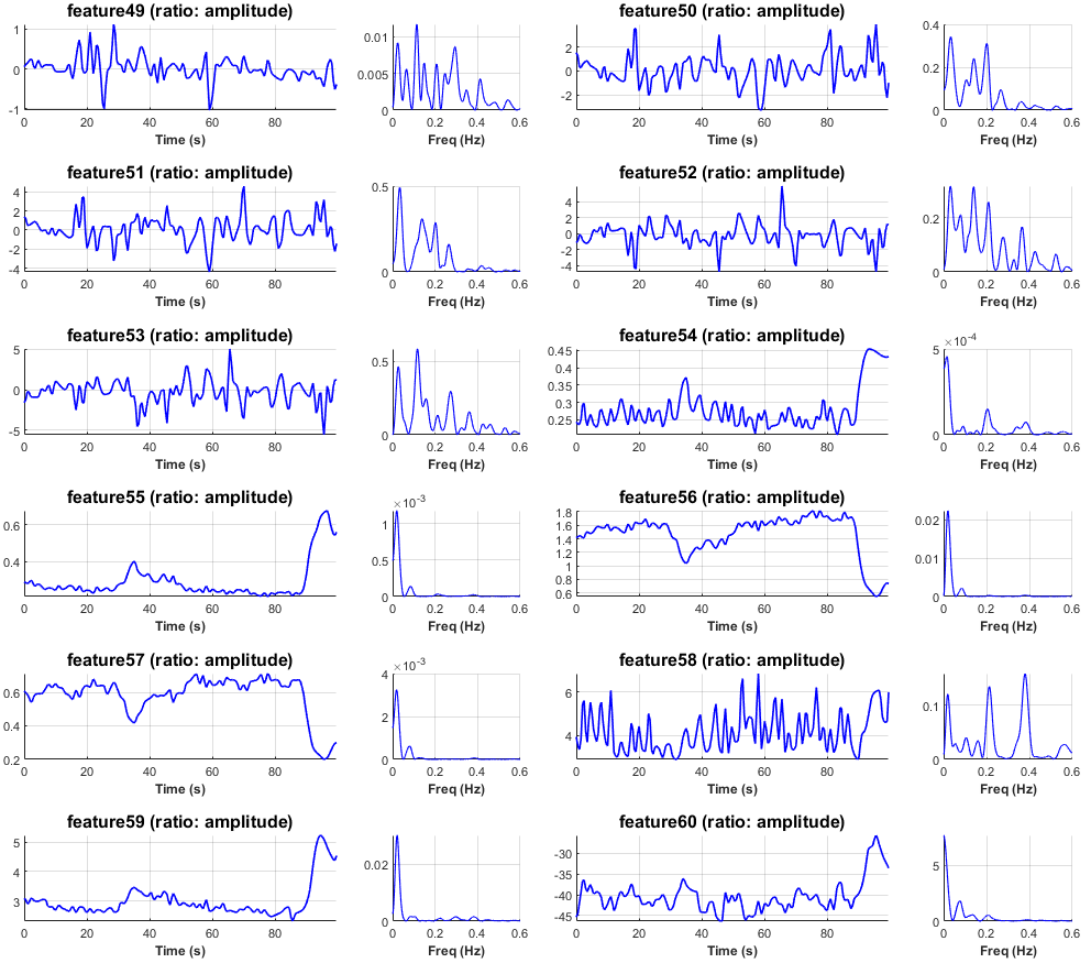


Figure A.13: The demonstration of extracted features without respiratory effect removal (feature49-feature60, exemplary on subject8 from BCG dataset).

The percentages of segments with correct trend tracking ( $TS(x_i, y_i) \geq 0$ ) for SBP and DBP estimations using ApproximateNet and DeepRNN-4L are shown in Figure A.19. It can be observed that DeepRNN-4L performs better than ApproximateNet in terms of trend tracking for both SBP and DBP.

The percentages of segments with  $MAD \leq 5$  mmHg and correct trend tracking for SBP and DBP estimation using ApproximateNet and DeepRNN-4L are shown in Figure A.20. It can be observed that DeepRNN-4L performs significantly better than ApproximateNet in terms of both distance and trend. Approximate 78% and 58% of SBP and DBP estimations



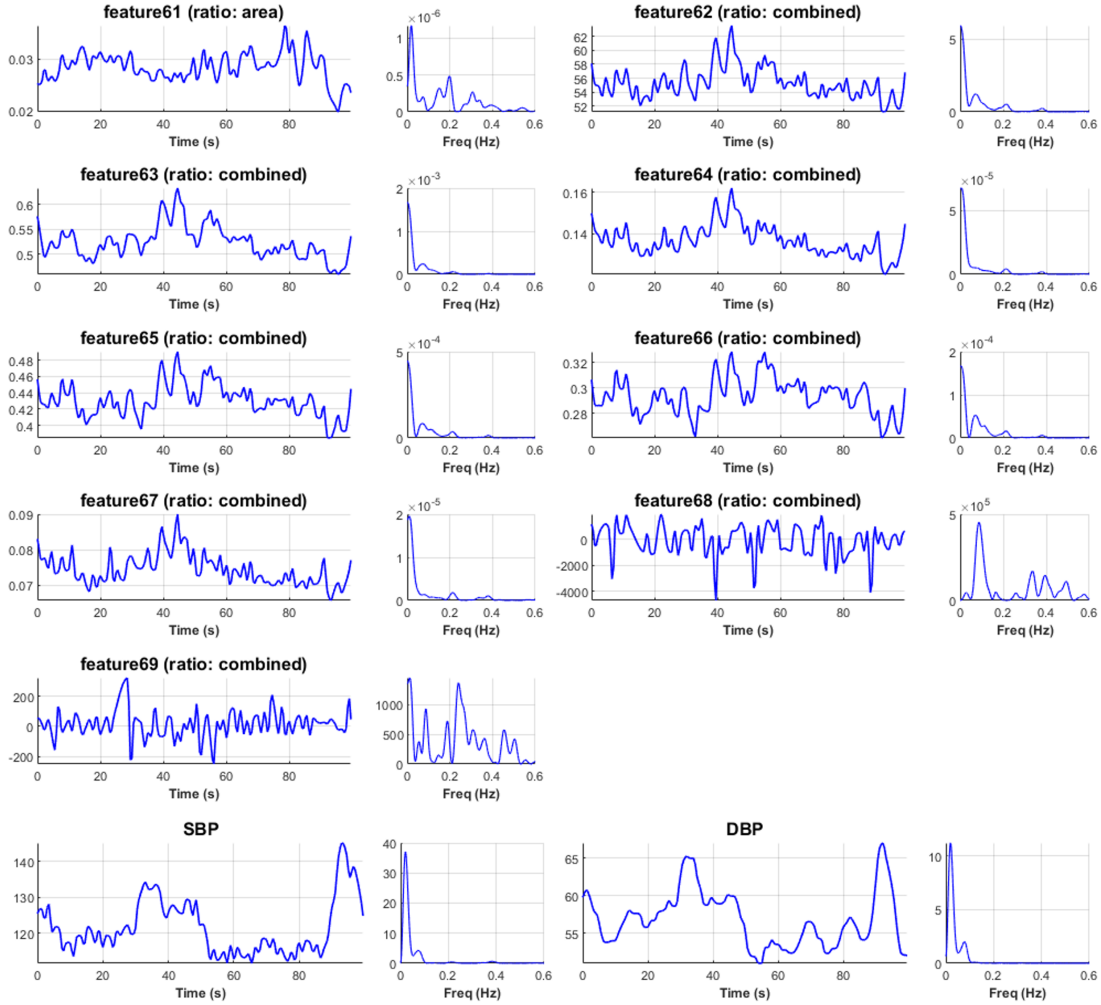


Figure A.14: The demonstration of extracted features, SBP and DBP without respiratory effect removal (feature61-feature69, SBP and DBP, exemplary on subject8 from BCG dataset).

have  $MAD \leq 5$  mmHg and correct trend tracking ( $TS(x_i, y_i) \geq 0$ ) using DeepRNN-4L. On the contrary, only around 16% and 25% of SBP and DBP estimations using ApproximateNet with  $MAD \leq 5$  mmHg and correct trend tracking.

With the examples shown in this section, the proposed methods and metrics can be used to benchmark different estimation algorithms using a same dataset. Also, an estimation algorithm can diagnose and improve its performance by analyzing the similarity metrics



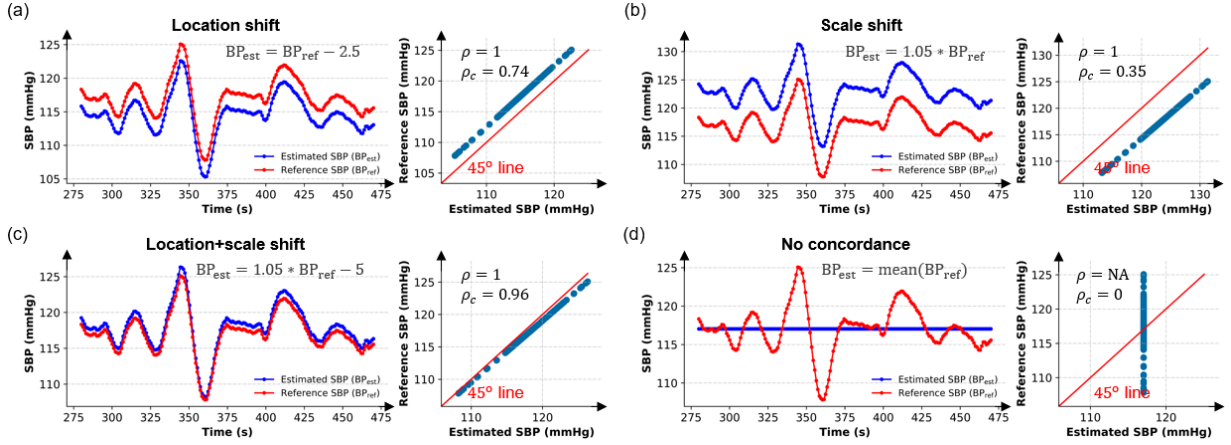


Figure A.15: Demonstration of concordance correlation coefficient ( $\rho_c$ ) and Pearson correlation coefficient ( $\rho$ ) for different scenarios, (a) location shift, (b) scale shift, (c) location and scale shift, (d) no concordance.

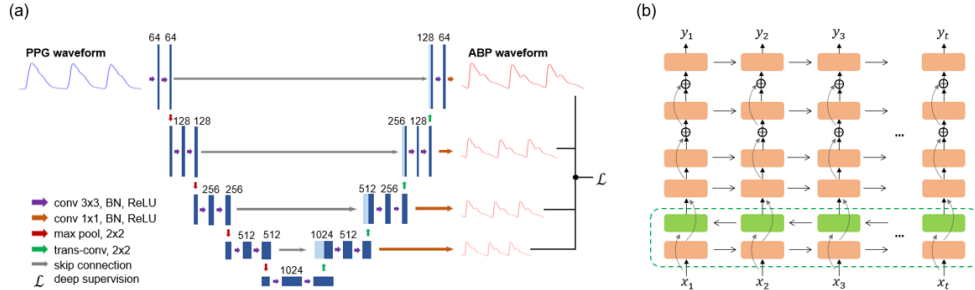


Figure A.16: (a) Diagram of ApproximateNet [117], which is a deeply supervised one-dimensional U-Net to approximate ABP waveforms from preprocessed PPG waveforms; (b) Diagram of DeepRNN-4L [84], each rectangle box is an LSTM cell (unit = 128), and the sequence length of each training sample is set to 32. The bottom green dashed block is a bidirectional LSTM layer (unit = 128); there is a dropout layer (rate = 0.2) applied after each bidirectional and forward LSTM layer; residual connections are also applied between LSTM layers. The last layer (unit = 2, activation=linear) was added for SBP and DBP prediction.

proposed in this thesis. For example, the DeepRNN-4L already performed well in BP trend tracking, and can just focus on minimizing the estimation error (e.g., MAD). However for

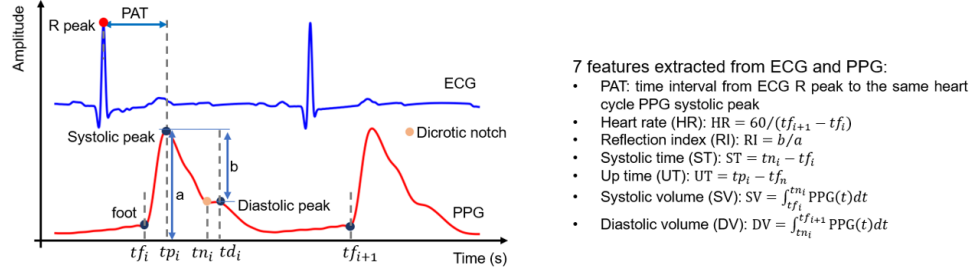


Figure A.17: Illustration of 7 ECG and PPG features used for SBP and DBP estimation proposed in [84].  $tf_i$ ,  $tp_i$ ,  $tn_i$  and  $td_i$  indicate the time index of foot point, systolic peak, dicrotic notch and diastolic peak of  $i^{th}$  PPG pulse, respectively.

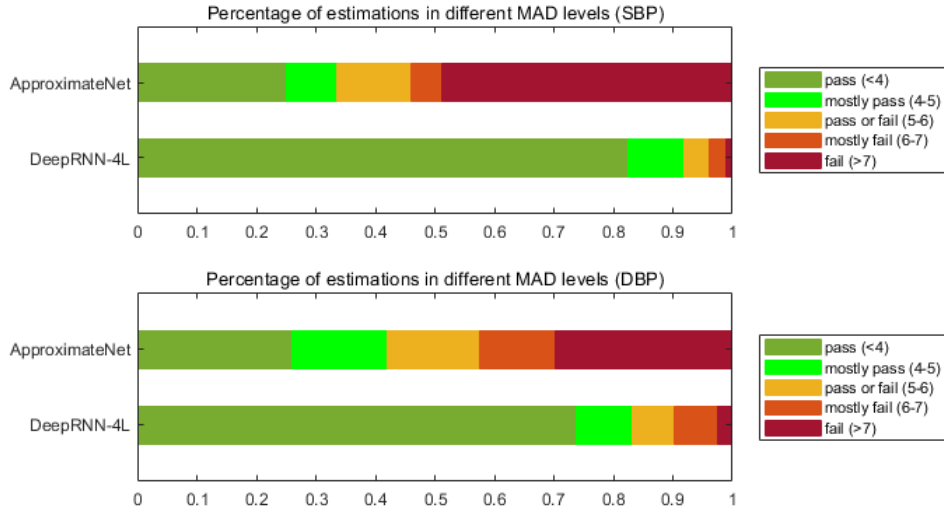


Figure A.18: Percentage of segments with different MAD levels for SBP and DBP estimation using ApproximateNet and DeepRNN-4L.

ApproximateNet, it needs to improve in terms of both distance and trend tracking accuracy.

#### A.4.6 Detailed Interpretation of the Proposed metrics

To better interpret the proposed metrics, a set of references and estimations of BP are simulated with different distance and trend similarities and different combinations of  $w_1$

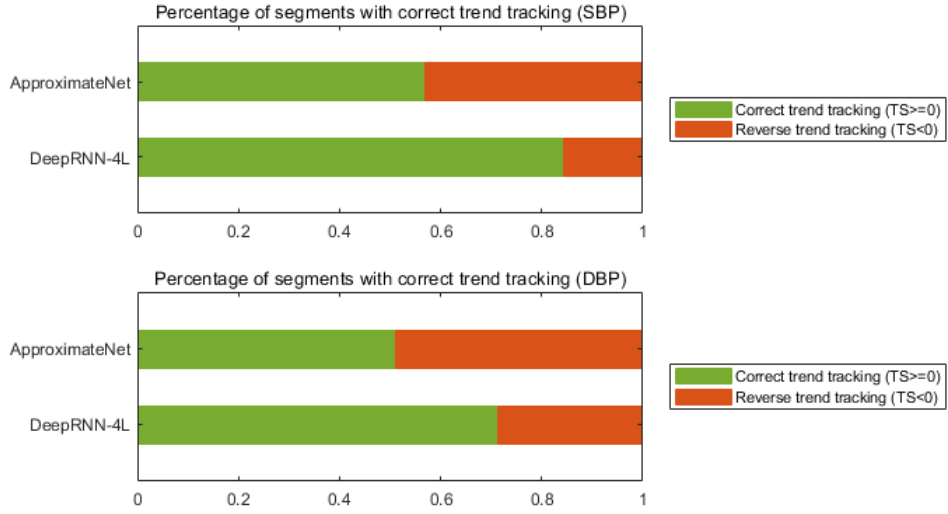


Figure A.19: Percentage of segments with correct BP trend tracking for SBP and DBP using ApproximateNet and DeepRNN-4L.

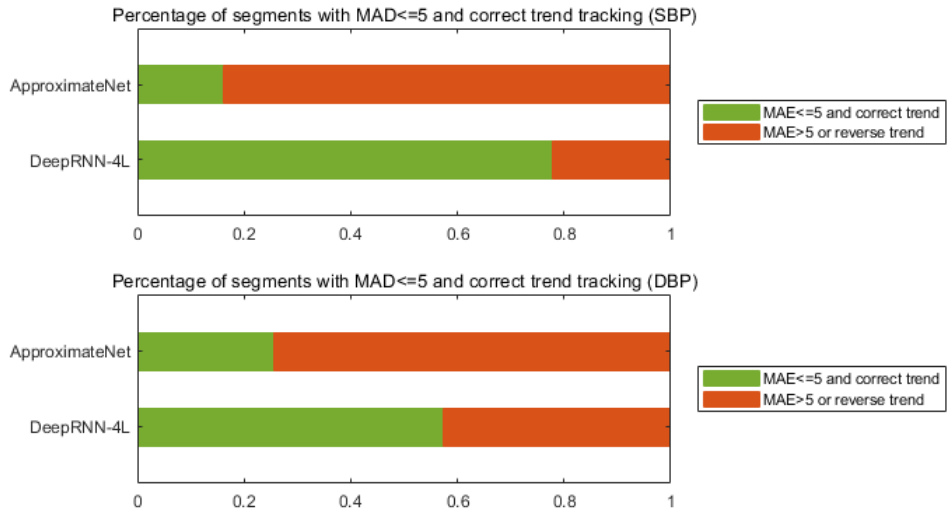


Figure A.20: Percentage of segments with  $MAD \leq 5$  mmHg and correct BP trend tracking for SBP and DBP using ApproximateNet and DeepRNN-4L.

and  $w_2$  for composite similarity calculation. The simulated reference and estimated BPs, as well as the corresponding distance (TN\_DS), trend (TN\_TS) and composite similarity (TN\_CS) with different  $w_1$  and  $w_2$  are shown in Figure A.21. Please note we assume  $MAD(x_L, y'_L) \approx MAD_i$  (in Equation 1) when calculating distance similarity, since  $x_L$  and  $y'_L$  are the approximated linear representation of the original reference and estimated BP segment  $x_i$  and  $y_i$  respectively. It can be observed in the simulation1 of Figure A.21 that both distance and trend similarities ( $DS(x_L, y'_L)$  and  $TS(x_L, y'_L)$ ) equal to 1 when the compared segments are identical. It is shown in Figure A.21 simulation11 that both distance and trend similarities ( $DS(x_L, y'_L)$  and  $TS(x_L, y'_L)$ ) equal to -1 when the MAD between two compared segments is significantly large and their trends are exactly reversed. It can also be observed from Table in Figure A.21 that the segment distance and trend similarities range from -1 to 1, and the segment composite similarity  $CS(x_i, y_i)$  always range from -1 to 1 for different combination of  $w_1$  and  $w_2$ .

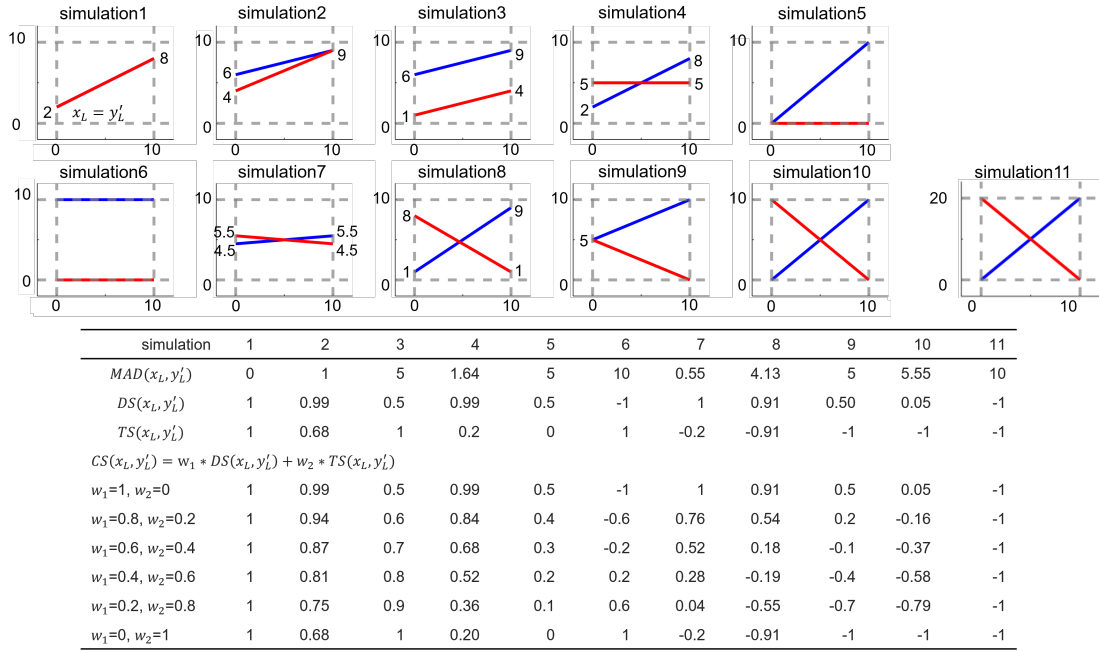


Figure A.21: Reference and estimated BPs with different distance and trend similarities, and composite measure of both distance and trend similarities with different weights combination  $w_1$  and  $w_2$  (assume  $t_{iR} - t_{iL} = 10$ ,  $\max(x_L, y'_L) - \min(x_L - y'_L) = 10$  mmHg for simulation1-10 s and  $\max(x_L, y'_L) - \min(x_L - y'_L) = 20$  mmHg for simulation11 (refer to Table4 and Figure4).

### A.4.7 Composite Similarity of Examples Shown in Figure 4.6 with Different $w_1$ and $w_2$ Combinations

In this section, we will demonstrate the composite similarity between the estimations using DeepRNN-4L, ApproximateNet and reference shown in Figure 4.6 with different  $w_1$  and  $w_2$  combinations. The temporally-normalized distance (TN\_DS) and trend (TN\_TS) similarities are 0.81 and 0.81 between estimations using DeepRNN-4L model and reference respectively. The temporally-normalized distance (TN\_DS) and trend (TN\_TS) similarities are 0.37 and 0.23 between estimations using DeepRNN-4L model and reference respectively. The temporally-normalized composite similarity (TN\_CS) combines distance and trend similarity with the weighted sum method,  $TN\_CS = w_1 \cdot TN\_DS + w_2 \cdot TN\_TS$ . The composite similarities with different  $w_1$  and  $w_2$  are shown in Figure S9 ( $w_1 = [0, 0.2, 0.4, 0.6, 0.8, 1]$ ,  $w_2 = 1 - w_1$ ). It can be observed that the distance and trend similarity between estimated BPs using DeepRNN-4L model and the reference are identical, therefore the composite similarity remains the same for different  $w_1$  and  $w_2$  combinations, as  $w_1 + w_2 = 1$ . The composite similarity between estimated BPs using ApproximateNet and the reference equals the distance similarity when  $w_1 = 1$  and equals to trend similarity when  $w_2 = 1$ . The range of the composite similarity is  $TN\_CS \in [\min(TN\_DS, TN\_TS), \max(TN\_DS, TN\_TS)]$ , according to different values assigned to  $w_1$  and  $w_2$ .

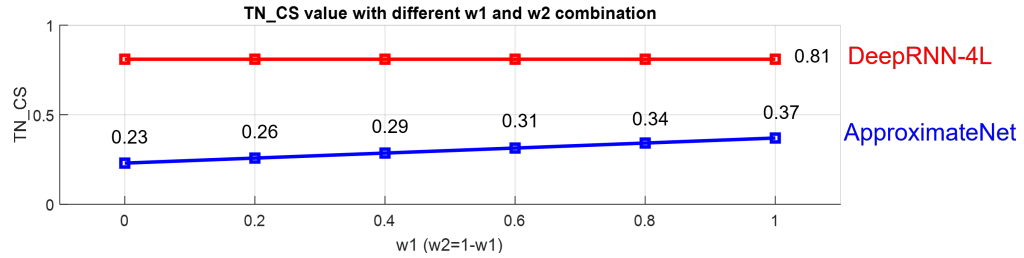


Figure A.22: the composite similarity metric (TN\_CS) between estimations of DeepRNN-4L, ApproximateNet model and reference shown in Figure7 with different combinations of  $w_1$  and  $w_2$ .

## A.5 Impact of Sampling Frequency

To investigate the impact of sampling frequency on feature precision and its relationship with BP, a downsampling experiment was conducted. ECG and PPG signals originally

sampled at 1 kHz from BCG BP dataset were downsampled to 500 Hz and 125 Hz, simulating the rates used in VitalDB dataset [215] and MIMIC dataset [203] respectively. Subsequently, key morphological features, particularly PAT (calculated as the time interval between the ECG R-peak and the PPG max-slope point,  $PAT_{ms}$ ), were extracted from the signals at three sampling frequencies (1 kHz, 500 Hz, and 125 Hz).

The analysis revealed a noticeable degradation in the precision of PAT measurements and its correlation with SBP at lower sampling frequencies. As illustrated in Figure A.23 (a), the PAT sequence extracted from signals sampled at 1 kHz exhibits finer granularity compared to those from 500 Hz and 125 Hz. Specifically, the 125 Hz PAT sequence shows distinct quantization steps, as its temporal resolution is limited to  $1/125s = 8$  milliseconds.

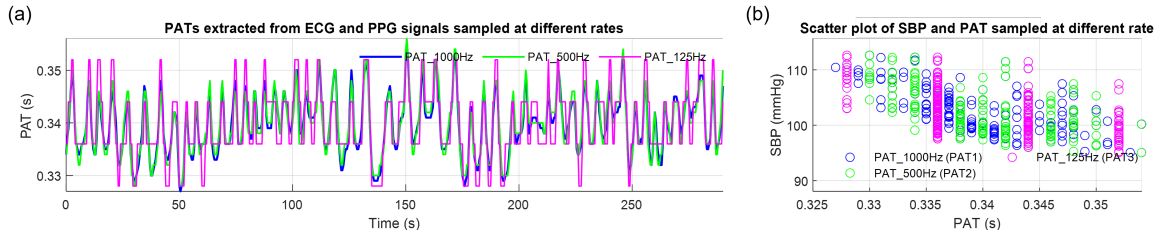


Figure A.23: Impact of sampling frequency on PAT precision and its correlation with SBP for an example subject. (a) PAT extracted from signals sampled at 1 kHz (blue), downsampled to 500 Hz (green), and downsampled to 125 Hz (magenta), showing increased quantization at lower sampling rate. (b) Scatter plot of SBP versus PAT for the three sampling frequencies, illustrating a wider range of SBP values for a given PAT measurement at lower sampling rate and decreased correlation.

The reduced resolution at lower sampling frequencies leads to increased ambiguity in the relationship between PAT and SBP. As shown in the scatter plot in Figure A.23 (b) for an example subject, a single PAT value derived from the 125 Hz signals often corresponds to a substantially wider range of SBP values compared to PAT derived from 1 kHz signals. For instance, when PAT was measured as 0.336 s from the 125 Hz downsampled signals, the corresponding SBP values spanned a range of 14.72 mmHg (from 97.54 mmHg to 112.26 mmHg). In contrast, PAT values derived from the original 1 kHz signals within the same approximate temporal vicinity exhibited a narrower SBP range of only 7.11 mmHg (from 99.11 mmHg to 106.21 mmHg). This precision difference directly impacted the correlation between PAT and SBP. For the example subject shown in Figure, the Pearson correlation coefficient ( $r$ ) between PAT and SBP decreased as the sampling rate was reduced (1 kHz:  $r = -0.62$ , 500 Hz:  $r = -0.59$ , 125 Hz:  $r = -0.525$ ).

The experimental results confirm that the 1 kHz sampling rate of the BCG BP dataset provides enhanced temporal resolution for extracting time interval based features like PAT compared to lower frequencies typical of other datasets. This higher precision leads to a stronger and less ambiguous correlation between the extracted features and BP, which can potentially improve BP estimation performance.

## A.6 Impact of Differential Filter Orders

To quantify the impact of using distinct filter configurations for ECG and PPG signals, an experiment was conducted to compare the PAT values derived using two different pre-processing pipelines applied to the same raw ECG and PPG data segments sampled at 1 kHz:

1. **Set 1 (Thesis Configuration):** ECG was filtered using a 2-nd order Butterworth bandpass filter (0.5-30 Hz), and PPG was filtered using a 4-th order Chebyshev Type II bandpass filter (0.5-12 Hz). PAT was calculated from these differentially filtered signals.
2. **Set 2 (Control Configuration):** Both ECG and PPG were filtered using 4-th order Chebyshev Type II bandpass filters with respective passbands (ECG: 0.5-30 Hz, PPG: 0.5-12 Hz). PAT was calculated from these signals filtered with the same filter type and order.

The PAT sequences derived from both filter sets were compared for the example subjects.

- **Magnitude of Difference:** The absolute difference in PAT values between the two sets was found to be minimal. For Subject 1 (Figure A.24), the absolute difference was typically around 1 ms. For Subject 2 (Figure A.25), the absolute difference ranged between 2 ms and 4 ms.
- **Correlation and Trend:** Despite the small offsets, the PAT sequences generated by both configurations exhibited almost identical temporal patterns and trends. The Pearson correlation coefficient between the PAT sequences from Set 1 and Set 2 was extremely high for both subjects (Subject 1:  $r = 0.9978$ ; Subject 2:  $r = 0.9945$ ).

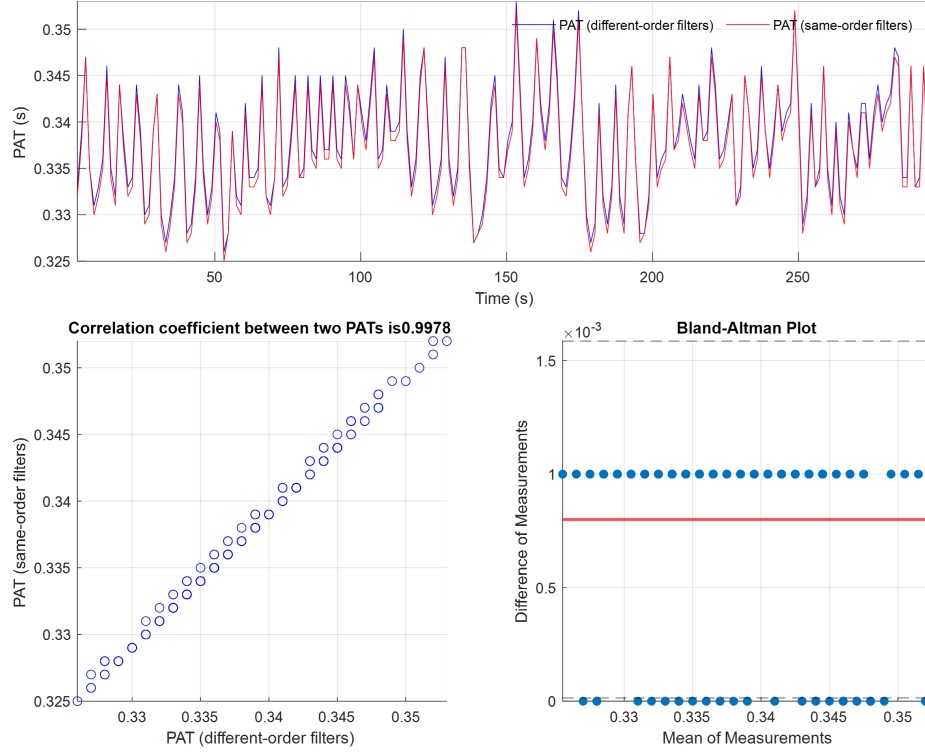


Figure A.24: Comparison of PAT calculated using different filter configurations for Subject 1. Time series plot of PATs, with blue line representing PAT extracted from the thesis configuration and red line representing PAT extracted from the control configuration, showing high temporal agreement. The scatter plot (bottom left) confirms near-perfect linear correlation ( $r > 0.99$ ). The Bland-Altman plot (bottom right) illustrates the small magnitude and near-constant nature of the difference between the two filter configurations.

- **Visual Inspection:** As shown in the time series plots and Bland-Altman plots in Figure A.24 and Figure A.25, the difference between two PAT sequences appears as a small, near-constant offset, rather than a distortion of the shape or dynamic changes.

The results demonstrated that while using different filter types and orders for ECG and PPG introduces a small difference in the absolute PAT values, this difference is highly consistent and primarily manifests as a minor offset. Crucially, the relative changes and dynamic trends within the PAT sequence, which are the primary focus for the trend-tracking models developed in this thesis, are preserved with very high fidelity (correlation  $r > 0.99$ ). Given the high sampling rate of 1 kHz and the model’s emphasis on trend



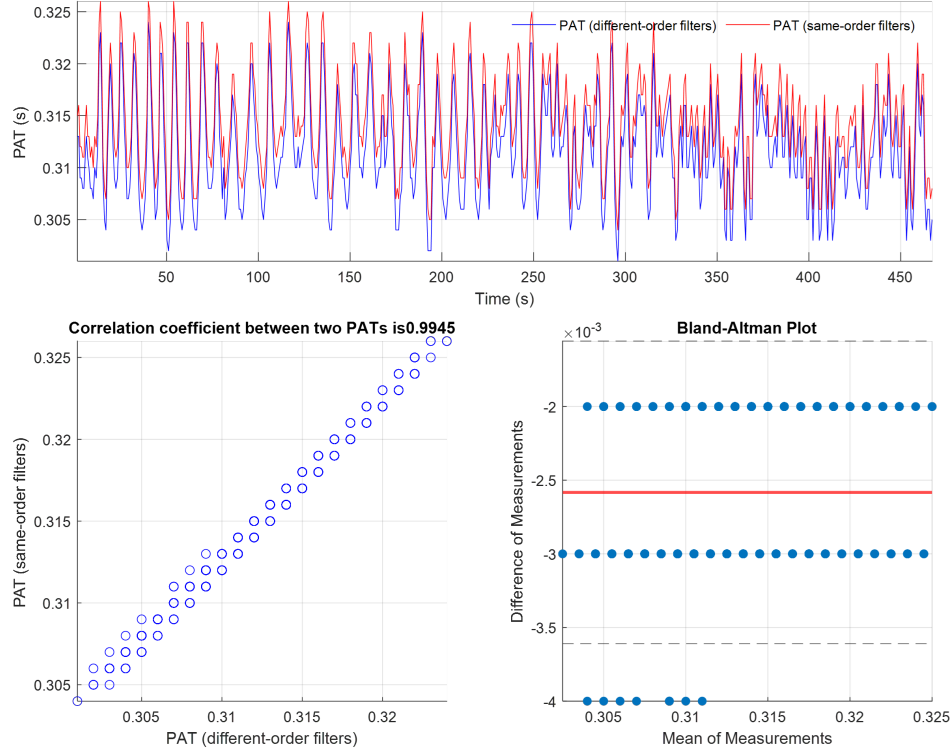


Figure A.25: Comparison of PAT calculated using different filter configurations for Subject 2. Time series plot of PATs, with blue line representing PAT extracted from the thesis configuration and red line representing PAT extracted from the control configuration, showing high temporal agreement. The scatter plot (bottom left) confirms near-perfect linear correlation ( $r > 0.99$ ). The Bland-Altman plot (bottom right) illustrates the small magnitude and near-constant nature of the difference between the two filter configurations.

similarity over absolute point-wise accuracy, the minimal and consistent temporal shift introduced by the differential filtering is considered negligible and does not significantly impact the feature extraction or the final estimation performance reported in this thesis.

## A.7 Few-Shot Fine Tuning Strategies

1. **Full Fine-tuning:** All parameters  $\theta^*$  of the pre-trained model are updated based on the gradients computed from  $D_{\text{test}}^S$ . Then the model is more flexible to adapt but

is easier to overfit with a small dataset.

2. **Partial Fine-tuning (Layer-wise Tuning):** To mitigate overfitting and preserve general features, only a subset of the model's parameters is fine-tuned. For instance, the "feature extractor" of a regression model (e.g., the TCN and SENet modules) can be frozen and only fine-tune the target predictor (e.g., the last KAN layer). The choice of which layers to fine-tune is a critical design decision and often depends on the architecture and the assumed transferability of features at different depths.
3. **Learning Rate Adjustment:** The learning rate used for fine-tuning is typically set to a value smaller than that used during the initial pre-training. A smaller learning rate ensures that the parameter updates are modest, preventing the model from drastically moving away from the well-generalized solution  $\theta^*$  and potentially forgetting the knowledge acquired from the  $N - 1$  training subjects.
4. **Number of Iterations:** Fine-tuning is generally performed for a limited number of epochs or gradient update steps. Excessive training on the small set  $A_{\text{test}}$  can quickly lead to overfitting.