



PDF Download  
3714394.3754412.pdf  
13 January 2026  
Total Citations: 0  
Total Downloads: 64

Latest updates: <https://dl.acm.org/doi/10.1145/3714394.3754412>

SHORT-PAPER

## Motion-Robust Multimodal Fusion of PPG and Accelerometer Signals for Three-Class Heart Rhythm Classification

**YANGYANG ZHAO**, University of Turku, Turku, Southwest Finland, Finland

**MATTI KAISTI**, University of Turku, Turku, Southwest Finland, Finland

**OLLI LAHDENOJA**, University of Turku, Turku, Southwest Finland, Finland

**TERO KOIVISTO**, University of Turku, Turku, Southwest Finland, Finland

Open Access Support provided by:

University of Turku

Published: 12 October 2025

[Citation in BibTeX format](#)

UbiComp '25: The 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing / ISWC ACM International Symposium on Wearable Computers  
October 12 - 16, 2025  
Espoo, Finland

Conference Sponsors:  
SIGMOBILE  
SIGCHI

# Motion-Robust Multimodal Fusion of PPG and Accelerometer Signals for Three-Class Heart Rhythm Classification

Yangyang Zhao  
yazhao@utu.fi

Department of Computing, Faculty of Technology,  
University of Turku  
Turku, Finland

Olli Lahdenoja  
olanla@utu.fi

Department of Computing, Faculty of Technology,  
University of Turku  
Turku, Finland

Matti Kaisti  
mkaist@utu.fi

Department of Computing, Faculty of Technology,  
University of Turku  
Turku, Finland

Tero Koivisto  
tejuko@utu.fi

Department of Computing, Faculty of Technology,  
University of Turku  
Turku, Finland

## Abstract

Atrial fibrillation (AF) is a leading cause of stroke and mortality, particularly in elderly patients. Wrist-worn photoplethysmography (PPG) enables non-invasive, continuous rhythm monitoring, yet suffers from significant vulnerability to motion artifacts and physiological noise. Many existing approaches rely solely on single-channel PPG and are limited to binary AF detection, often failing to capture the broader range of arrhythmias encountered in clinical settings. We introduce RhythmiNet, a residual neural network enhanced with temporal and channel attention modules that jointly leverage PPG and accelerometer (ACC) signals. The model performs three-class rhythm classification: AF, sinus rhythm (SR), and Other. To assess robustness across varying movement conditions, test data are stratified by accelerometer-based motion intensity percentiles without excluding any segments. RhythmiNet achieved a 4.3% improvement in macro-AUC over the PPG-only baseline. In addition, performance surpassed a logistic regression model based on handcrafted HRV features by 12%, highlighting the benefit of multimodal fusion and attention-based learning in noisy, real-world clinical data.

## CCS Concepts

• **Applied computing** → **Health informatics**; • **Computing methodologies** → *Neural networks*; • **Human-centered computing** → Ubiquitous computing.

## Keywords

Atrial Fibrillation, Photoplethysmography (PPG), Accelerometer (ACC), Multimodal Fusion, Heart Rhythm Classification, Deep Learning

## ACM Reference Format:

Yangyang Zhao, Matti Kaisti, Olli Lahdenoja, and Tero Koivisto. 2025. Motion-Robust Multimodal Fusion of PPG and Accelerometer Signals for Three-Class Heart Rhythm Classification. In *Companion of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Companion '25)*, October 12–16, 2025, Espoo, Finland. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3714394.3754412>

## 1 Introduction

Atrial fibrillation (AF) is the most prevalent sustained arrhythmia and a major contributor to stroke, heart failure, and mortality, particularly in elderly populations [18]. Early detection is critical for timely intervention, especially among high-risk clinical groups. Advances in wearable technology have recently enabled long-term, non-invasive cardiac monitoring through photoplethysmography (PPG), a light-based technique for tracking peripheral blood volume variations [12]. However, real-world PPG signals often suffer from motion artifacts, poor sensor contact, and physiological noise, making AF detection challenging [4, 12].

Existing AF detection approaches generally fall into two main categories. The first relies on handcrafted features derived from heart rate variability (HRV), such as the root mean square of successive differences (RMSSD) [7, 19], the standard deviation of successive differences (SDSD)[11], and the percentage of NN intervals differing by more than 40 milliseconds (pNN40) [6, 19]. While physiologically interpretable, these features are highly sensitive to signal quality and often perform poorly in the presence of noise or irregular rhythms. The second category involves deep learning methods that learn discriminative features directly from raw PPG signals [3, 5, 16]. Although such methods typically outperform traditional techniques in terms of accuracy, they suffer from several limitations: (1) they treat AF detection as a binary classification problem, thus failing to capture the full spectrum of clinically relevant arrhythmias; and (2) they rely exclusively on single-channel PPG data, which reduces robustness under motion-induced distortions [3, 15, 16]. Additionally, (3) these models often discard low-quality signal segments, leading to inefficient utilization of continuous monitoring data [2, 5, 10, 14].



This work is licensed under a Creative Commons Attribution 4.0 International License.  
*UbiComp Companion '25, Espoo, Finland*  
© 2025 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-1477-1/2025/10  
<https://doi.org/10.1145/3714394.3754412>

To address this, we propose a multimodal framework that combines PPG and ACC signals via a residual neural network augmented with temporal and channel attention mechanisms. We evaluate it on a clinical-grade dataset of 1,000 hours of synchronized PPG, ECG, and ACC recordings from 49 elderly cardiac inpatients. To reflect real-world conditions, test samples are stratified by ACC-derived motion intensity without segment exclusion. We compare RhythmiNet against a PPG-only deep learning model and a logistic regression baseline using HRV features, with performance evaluated across varying levels of motion intensity.

## 2 Methods

### 2.1 Dataset and Preprocessing

**Data Collection:** This study uses a clinical-grade dataset collected between September 2022 and August 2023 from 49 elderly inpatients at the Heart Center of Turku University Hospital, Finland. Patients were monitored using two synchronized devices (Figure 1): a wrist-worn Philips Datalogger (PPG and tri-axial ACC at 32 Hz) and a Bittium Faros™ 360 Holter (ECG at 125 Hz). Rhythm annotations (AF, SR, Other) were derived from the ECG signals. The "Other" category includes arrhythmias such as VT, IVR, arrest, COUP, GEM, and SALV, which do not fit into standard AF or SR classes. The dataset was split into patient-independent training ( $n = 25$ ; 17 SR, 8 AF) and testing sets ( $n = 24$ ; 18 SR, 6 AF), totaling 473.73 hours of PPG recordings for training and 461.73 hours for testing.

**Signal Preprocessing and Motion Stratification:** PPG signals were segmented into 30-second windows (960 samples at 32 Hz), followed by third-order Butterworth bandpass filtering (0.5–8 Hz) and normalization to zero mean and unit variance on each segment. ACC signals were segmented using the same temporal windows. Within each segment, raw tri-axial ACC data were bandpass-filtered (0.5–5.0 Hz, second-order Butterworth), and motion magnitude was computed as  $\sqrt{acc_x(t)^2 + acc_y(t)^2 + acc_z(t)^2}$ . Segment-level motion scores were defined as the variance of the magnitude signal [1, 2]. All test segments were ranked according to their motion scores and stratified into percentile bins (e.g., 0–10%, 10–20%, ..., 90–100%) to evaluate model performance under varying levels of motion interference without excluding any data.

### 2.2 Proposed Model Architecture

We introduce RhythmiNet, a lightweight residual neural network specifically designed to jointly process photoplethysmography (PPG) and tri-axial ACC signals for heart rhythm classification. As illustrated in Figure 2, the architecture is composed of four main components: (1) a convolutional stem, (2) a residual backbone [8] enhanced with Squeeze-and-Excitation (SE) blocks [9], (3) a temporal attention module [17], and (4) a final classification head. The model accepts four synchronized input channels—PPG and three-axis ACC—sampled over 30-second segments (960 samples at 32 Hz). The convolutional stem, equipped with stride-2 projection, reduces the temporal resolution and extracts low-level features. The residual backbone comprises two stages, each containing two SE-enhanced BasicBlocks. Each block consists of two 1D convolutional layers, batch normalization, dropout, and skip (residual) connections. The SE modules perform channel-wise reweighting by applying global average pooling followed by two fully connected

layers with a reduction ratio of 2. To model long-range temporal dependencies, a temporal attention module is appended after the second residual stage. This module implements scaled dot-product attention across the temporal dimension and integrates contextual information via residual connections. Finally, the extracted features are aggregated using adaptive average pooling and passed through a fully connected layer.

### 2.3 Baseline Model

We implemented a conventional HRV-based method as the baseline model. Specifically, three widely used time-domain features were extracted from inter-beat intervals (IBIs): the RMSSD [7, 19], pNN40 [11, 19], and the SDDSD [6]. Pulse peaks were then detected using the Automatic Multiscale Peak Detection (AMPD) algorithm [13]. IBIs were computed as the time differences between adjacent peaks, and the HRV features were subsequently derived for each segment using standard analytical formulas. Each of the three features was used independently to train a logistic regression classifier. This classical approach serves as a practical baseline.

### 2.4 Experimental Setup

The model was trained on the full train dataset using three random seeds for robustness. Each run used a batch size of 64 for 60 epochs. To reduce temporal overfitting, signal segments were shuffled at each epoch. Optimization used Adam (learning rate  $1 \times 10^{-4}$ , weight decay  $1 \times 10^{-5}$ ) with a StepLR scheduler (factor 0.1 every 20 epochs). Cross-entropy loss was used for training. Experiments were conducted on a Linux workstation with Python 3.11.9, PyTorch 2.3.0, and an NVIDIA TITAN RTX GPU (24 GB, CUDA 12.1). We report macro-AUC, micro-AUC, and accuracy. Macro-AUC treats all classes equally, while micro-AUC reflects true label distribution. Accuracy denotes the proportion of correct predictions.

## 3 Experiment Results

Figure 3 presents a comparison of model performance across motion percentiles using three evaluation metrics: macro-AUC, micro-AUC, and accuracy. The proposed model using combined PPG and ACC input (orange) consistently outperforms all other methods across all metrics and motion levels. Specifically, it achieves a macro-AUC exceeding 0.83 in low-motion conditions (0–10%) and maintains robust performance (around 0.76) even in high-motion segments (90–100%).

In terms of micro-AUC, the PPG+ACC model remains consistently above 0.91, highlighting its strong generalization ability despite motion interference. The PPG-only variant (blue) also shows competitive performance, though with a more noticeable decline under increased motion. Accuracy follows a similar trend: the PPG+ACC model maintains above 0.81 across all percentiles, while the PPG-only model dips slightly but stays above 0.80.

In contrast, all HRV-based baselines, including pNN40, SDDSD, and RMSSD, show steep performance degradation as motion intensity increases. Both SDDSD and RMSSD fall below 0.60 in macro-AUC and accuracy in the top motion percentile, indicating their vulnerability to motion artifacts and limited robustness in realistic wearable settings.

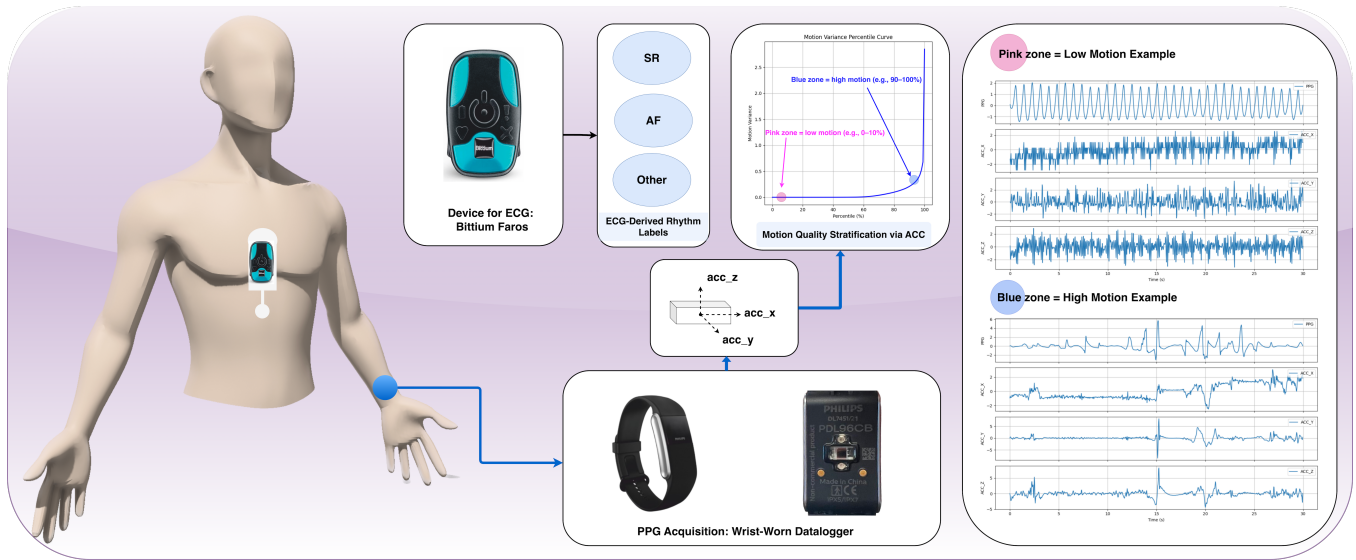


Figure 1: Overview of the data collection and motion stratification process. A wrist-worn Philips Datalogger recorded both PPG and tri-axial ACC signals, while a chest-worn Bittium Faros device recorded reference ECG. ECG provided rhythm labels (AF, SR, Other) used for training and testing. Segment-level motion scores were computed from the ACC magnitude and used to stratify test data into percentile-based motion levels. Example PPG and ACC traces are shown for two motion conditions: low-motion (top) and high-motion (bottom).

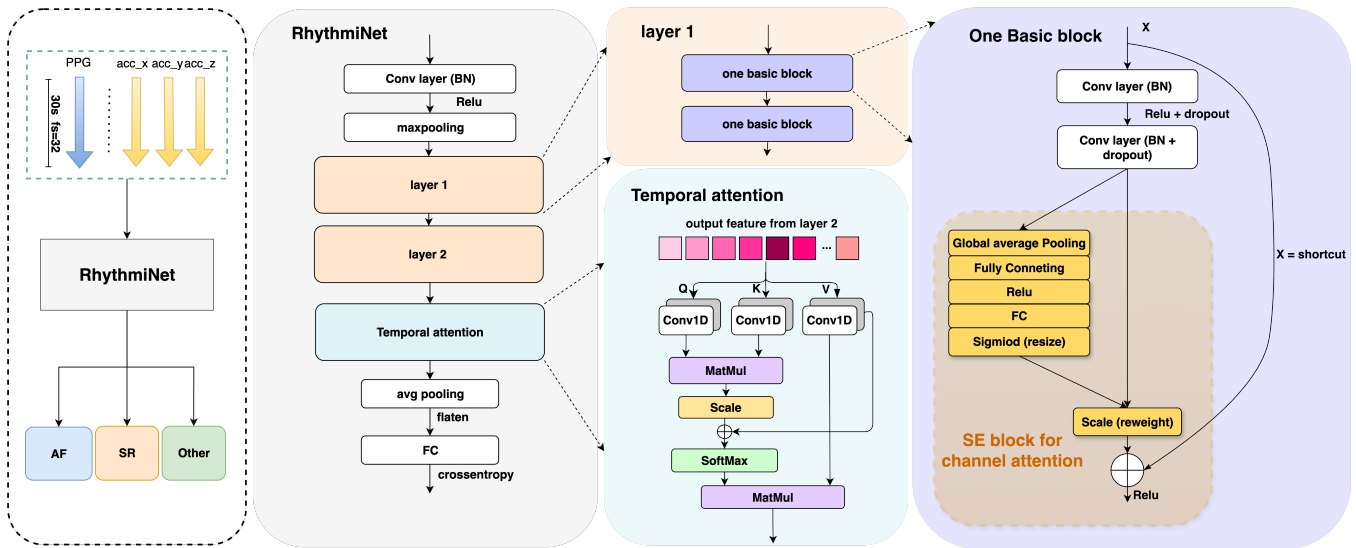


Figure 2: Overview of the RhythmiNet architecture for three-class heart rhythm classification (SR, AF, Other). The model processes 30-second segments of synchronized PPG and tri-axial ACC signals (960 samples at 32 Hz). Inputs are concatenated and passed through a convolutional stem, followed by two residual blocks enhanced with Squeeze-and-Excitation (SE) modules to emphasize informative channels. A temporal attention module captures long-range dependencies across time. Finally, global average pooling and a fully connected layer form the classification head. Dashed boxes illustrate the internal structure of selected modules (e.g., residual blocks and attention).

#### 4 Discussion and Future Work

RhythmiNet improves AF detection by leveraging multimodal inputs—PPG and ACC—to enhance robustness in motion-intensive

and noise-prone conditions. Instead of relying on complex fusion strategies, the model simply concatenates the input channels and processes them through a compact residual backbone. Temporal

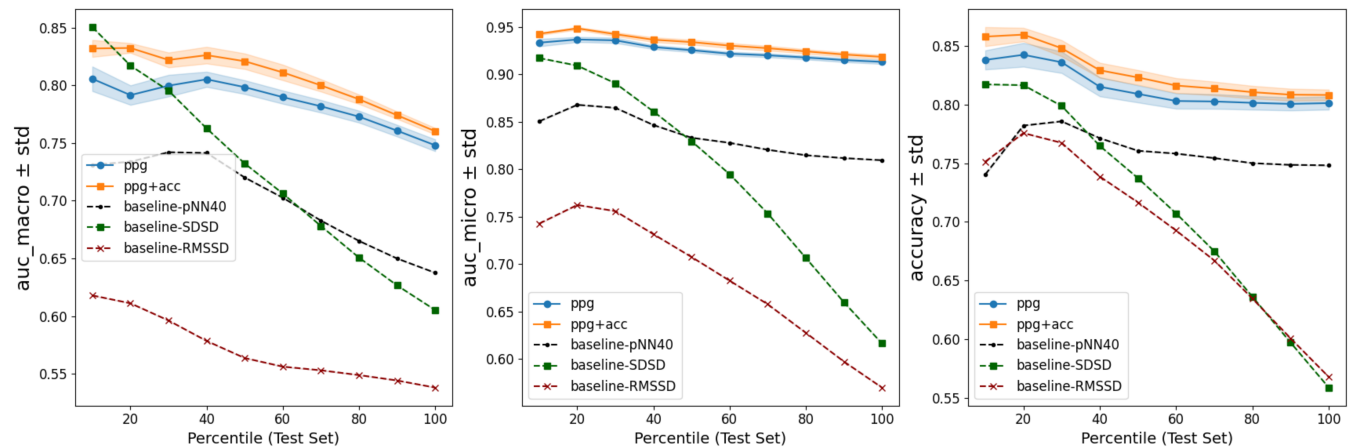


Figure 3: Performance across motion intensity percentiles: (Left) Macro-AUC, (Middle) Micro-AUC, and (Right) Accuracy.

attention and SE modules assist in extracting relevant features without adding complexity. We hypothesize that the model’s robustness to motion interference arises from three factors: (1) the inclusion of ACC signals to provide motion context, (2) attention-based reweighting of noisy temporal segments, and (3) inclusive training using motion-contaminated data without exclusion. This effective design outperforms traditional PPG-only methods and demonstrates strong potential for deployment in wearable cardiac monitoring.

To further enhance the clinical applicability of RhythmiNet, future work will focus on three main directions. First, we aim to improve model interpretability by developing intuitive visualization tools that highlight which input segments most influence decision-making. Second, we will investigate model compression and acceleration techniques, such as pruning and quantization, to support real-time inference on low-power, resource-constrained wearable devices. Finally, we intend to benchmark RhythmiNet against a wider range of state-of-the-art models across diverse, publicly available datasets to comprehensively evaluate its generalizability and clinical utility.

## 5 Acknowledgments

This study was part of the clinical trial CARE-DETECT (ClinicalTrials.gov ID: NCT05351775). It was approved by the Ethics Committee of the Hospital District of Southwest Finland and conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants.

This study was funded by the Moore4Medical project, supported by the ECSEL JU and Business Finland (Grant Agreements H2020-ECSEL-2019-IA-876190 and 7215/31/2019), and by the ITEA project RM4HEALTH, supported by Business Finland (Grant 8139/31/2022).

## References

- [1] Bahaa Al-Sheikh. 2023. Adaptive algorithm for motion artifacts removal in wearable biomedical sensors during physical exercise. *IEEE Sensors Journal* 23, 9 (2023), 9491–9499.
- [2] Alireza Aliamiri and Yichen Shen. 2018. Deep learning based atrial fibrillation detection using wearable photoplethysmography sensor. In *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*. IEEE, 442–445.
- [3] Pavel Antiperovitch, David Mortara, Joshua Barrios, Robert Avram, Kimberly Yee, Armeen Namjou Khaless, Ashley Cristal, Geoffrey Tison, and Jeffrey Olgin. 2024. Continuous Atrial Fibrillation Monitoring From Photoplethysmography: Comparison Between Supervised Deep Learning and Heuristic Signal Processing. *Clinical Electrophysiology* 10, 2 (2024), 334–345.
- [4] KR Arunkumar and M Bhaskar. 2020. Heart rate estimation from wrist-type photoplethysmography signals during physical exercise. *Biomedical Signal Processing and Control* 57 (2020), 101790.
- [5] Syed Khairul Bashar, Dong Han, Shirin Hajeb-Mohammadalipour, Eric Ding, Cody Whitcomb, David D McManus, and Ki H Chon. 2019. Atrial fibrillation detection from wrist photoplethysmography signals using smartwatches. *Scientific reports* 9, 1 (2019), 15054.
- [6] Linda M Erikäinen, Lukas Dekker, Alberto G Bonomi, Rik Vullings, Fons Schipper, Jenny Margarito, Helma M de Morree, and Ronald M Aarts. 2017. Validating features for atrial fibrillation detection from photoplethysmogram under hospital and free-living conditions. In *2017 Computing in Cardiology (CinC)*. IEEE, 1–4.
- [7] Sven Geurts, Martijn J Tilly, Banafsheh Arshi, Bruno HC Stricker, Jan A Kors, Jaap W Deckers, Natasja MS de Groot, M Arfan Ikram, and Maryam Kavousi. 2023. Heart rate variability and atrial fibrillation in the general population: a longitudinal and Mendelian randomization study. *Clinical research in cardiology* 112, 6 (2023), 747–758.
- [8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.
- [9] Jie Hu, Li Shen, and Gang Sun. 2018. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 7132–7141.
- [10] Duc Huy Nguyen, Paul C-P Chao, Chih-Chieh Chung, Ray-Hua Horng, and Bhaskar Choubey. 2022. Detecting atrial fibrillation in real time based on PPG via two CNNs for quality assessment and detection. *IEEE Sensors Journal* 22, 24 (2022), 24102–24111.
- [11] Michele Orini, Stefan van Duijvenboden, William J Young, Julia Ramirez, Aled R Jones, Alun D Hughes, Andrew Tinker, Patricia B Munroe, and Pier D Lambiase. 2023. Long-term association of ultra-short heart rate variability with cardiovascular events. *Scientific Reports* 13, 1 (2023), 18966.
- [12] Tania Pereira, Nate Tran, Kais Gadhomi, Michele M Pelter, Duc H Do, Randall J Lee, Rene Colorado, Karl Meisel, and Xiao Hu. 2020. Photoplethysmography based atrial fibrillation detection: a review. *NPJ digital medicine* 3, 1 (2020), 3.
- [13] Felix Scholkmann, Jens Boss, and Martin Wolf. 2012. An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals. *Algorithms* 5, 4 (2012), 588–603.
- [14] JL Selder, T Proesmans, L Breukel, O Dur, W Gielen, Anne C van Rossum, and CP Allaart. 2020. Assessment of a standalone photoplethysmography (PPG) algorithm for detection of atrial fibrillation on wristband-derived data. *Computer Methods and Programs in Biomedicine* 197 (2020), 105753.
- [15] Supreeth Prajwal Shashikumar, Amit J Shah, Qiao Li, Gari D Clifford, and Shamim Nemat. 2017. A deep learning approach to monitoring and detecting atrial fibrillation using wearable technology. In *2017 IEEE EMBS international conference on biomedical & health informatics (BHI)*. Ieee, 141–144.

- [16] Yichen Shen, Maxime Voisin, Alireza Aliamiri, Anand Avati, Awni Hannun, and Andrew Ng. 2019. Ambulatory atrial fibrillation monitoring using wearable photoplethysmography with deep learning. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*. 1909–1916.
- [17] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems* 30 (2017).
- [18] Fons J Wesselius, Mathijs S van Schie, Natasja MS De Groot, and Richard C Hendriks. 2021. Digital biomarkers and algorithms for detection of atrial fibrillation using surface electrocardiograms: A systematic review. *Computers in Biology and Medicine* 133 (2021), 104404.
- [19] Yangyang Zhao, Olli Lahdenoja, Ismail Elnaggar, Tuija Vasankari, Samuli Jaakkola, Tuomas Kiviniemi, Juhani Airaksinen, Matti Kaisti, and Tero Koivisto. 2025. Validation of a MEMS-Based Pressure Sensor System for Atrial Fibrillation Detection from Wrist and Finger. *IEEE Sensors Journal* (2025).