Journal of Machine Intelligence in Healthcare

E-ISSN: 3090-9503
P-ISSN: 3090-9511

(Research) Article

F-UTransBPNet A Transformer-Enhanced U-Net for Cuffless Blood Pressure Estimation Across Activity Scenarios

Muhammad Lutfi [1,*], and Eleni Vogiatzi [2]

[1] Department of Informatics, Universitas Muhammadiyah Semarang, Semarang, Indonesia; e-mail : m.lutfi@unimus.ac.id
[2] Data Mining and Analytics research group, School of Science and Technology, International Hellenic University, Thermi, Greece; e-mail : eleni.v@ihu.gr

* Corresponding Author : Muhammad Lutfi

Abstract: Continuous blood pressure (BP) monitoring is essential for early detection and management of hypertension, a major risk factor for cardiovascular disease and death. Traditional cuff-based devices are not suitable for continuous or ambulatory monitoring because they can cause discomfort and only provide intermittent measurements. This has led to the development of cuffless BP estimation methods that utilize physiological signals like electrocardiography (ECG) and photoplethysmography (PPG). Existing deep learning techniques, such as convolutional neural networks (CNNs) and hybrid CNN–LSTM models, have shown promising results in controlled settings; however, they often face challenges in maintaining performance under changing physiological conditions. To address this issue, we introduce F-UTransBPNet, a hybrid U-Net–Transformer architecture that combines local feature extraction with long-range temporal modeling, further optimized through selective fine-tuning. The model was trained and tested on three complementary datasets: MIMIC (ICU patients), Dataset_Drink (water intake), and Dataset_Exercise (bicycle ergometry), covering both static and activity-based scenarios. Results show that F-UTransBPNet achieves mean absolute differences of 4.4 mmHg (SBP) and 2.2 mmHg (DBP) when compared to invasive references, meeting AAMI standards, and maintains strong correlations in activity datasets (PCC up to 0.82 for SBP). Minimal fine-tuning with just 10–20% of scenario-specific data restores performance across different domains, demonstrating the feasibility of calibration-light operation. These findings indicate that F-UTransBPNet offers an effective balance of accuracy, adaptability, and computational efficiency, supporting its potential as a reliable tool for cuffless BP monitoring in both inpatient and wearable healthcare settings.

Keywords: Cuffless blood pressure estimation; Deep learning; U-Net; Transformer; Electrocardiography (ECG); Photoplethysmography (PPG); Multi-modal signals; Clinical validation

Received: June 2, 2025
Revised: July 17, 2025
Accepted: August 24, 2025
Published: August 31, 2025
Curr. Ver.: August 31, 2025

1. Introduction

Hypertension is a leading modifiable risk factor for cardiovascular disease, affecting more than one billion people worldwide and significantly contributing to global morbidity and mortality [1], [2]. Despite improvements in treatment, low awareness, poor control rates, and the asymptomatic nature of hypertension continue to increase its burden on healthcare systems [3]. Continuous and precise blood pressure (BP) monitoring is crucial for improving diagnosis, guiding treatment, and preventing adverse cardiovascular events. However, traditional cuff-based devices are unsuitable for long-term monitoring because of their intermittent measurements, discomfort, and inability to detect rapid hemodynamic changes [4].

To overcome these issues, researchers have developed cuffless BP estimation methods based on physiological signals like electrocardiography (ECG) and photoplethysmography (PPG). Early techniques relied on pulse transit time (PTT), but these were sensitive to calibration drift and lacked generalizability [5]. With advancements in deep learning, data-

driven approaches have shown promise in directly learning from ECG and PPG signals, enabling robust end-to-end BP prediction [6], [7].

Recurrent neural networks (RNNs) and LSTM-based models demonstrated the ability to model temporal dependencies in physiological signals [8], [9]. Hybrid CNN–LSTM frameworks were later developed to capture both morphological and temporal features [10]. Although effective in controlled settings, these models often need large calibration datasets and struggle to generalize in dynamic conditions such as exercise or daily activities [11].

More recently, Transformer-based and hybrid models have been created to tackle these challenges. Liu *et al.* introduced HGCTNet, combining handcrafted features with CNN and Transformer modules, achieving BHS and BHS acceptability standards [12]. Chen *et al.* [13] proposed rU-Net, which merges U-Net and ResNet using transfer learning, earning a BHS Grade A after minimal fine-tuning [13]. Wang *et al.* [14] developed cBP-Tnet, a Transformer architecture utilizing single-channel PPG, reporting mean errors of 4.3 mmHg (SBP) and 2.2 mmHg (DBP). Zhang *et al.* [15] presented MuFuBP-Net, which uses dual-feature fusion and progressive refinement, attaining state-of-the-art performance across multiple cohorts.

Furthermore, several reviews discuss broader challenges in the field, including distribution shifts across different scenarios, computational complexity, and the need for clinically validated protocols [16–18]. Other studies focus on wearable devices, domain adaptation strategies, and hybrid architectures to facilitate real-world application [19–21].

Despite these advances, significant challenges persist: (i) limited ability to generalize across different scenarios, (ii) insufficient modeling of both short- and long-term dependencies, and (iii) high computational demands that hinder implementation in wearable systems.

To address these issues, we propose F-UTransBPNet, a fine-tuned hybrid U-Net–Transformer model for cuffless BP estimation. This model combines U-Net's capability for capturing detailed local features with the Transformer's strength in modeling long-range dependencies. Additionally, it is optimized with selective fine-tuning to improve adaptability across various physiological conditions.

The main contributions of this work are:

1. A novel U-Net–Transformer hybrid architecture tailored for multi-modal signals (ECG, PPG, VPPG, APPG) to enable accurate cuffless BP estimation.
2. Comprehensive evaluation across three datasets (MIMIC, water intake, exercise), representing both static and activity-based scenarios.
3. Demonstration that limited fine-tuning significantly enhances cross-scenario performance, offering a pragmatic balance between calibration-free design and clinical reliability.

The remainder of the paper is organized as follows: Section 2 describes the methodology, Section 3 presents experimental results, Section 4 discusses findings and implications, and Section 5 concludes with future directions.

3. Proposed Method

This section introduces the proposed F-UTransBPNet, a hybrid deep learning architecture created for cuffless blood pressure estimation across different activity scenarios. The method combines the U-Net's encoder–decoder structure with Transformer encoders to capture both local and global dependencies in multi-modal physiological signals (ECG, PPG, VPPG, APPG). The workflow of the proposed system is shown in Algorithm 1.

3.1. Algorithmic Representation of F-UTransBPNe

To provide a clearer understanding of how the proposed model operates, this section presents an algorithmic representation of F-UTransBPNet. The algorithm formalizes the complete workflow of the model, beginning from raw physiological signals to the final estimation of systolic and diastolic blood pressure values. Expressing the method in pseudocode enhances reproducibility and offers a stepwise reference for researchers seeking to replicate or extend this work. Algorithm 1 presents the pseudocode of the proposed method.

Algorithm 1. F-UTransBPNet for Cuffless Blood Pressure Estimation

INPUT: $X = \{x_{ECG}(t)x_{PPG}(t)x_{VPPG}(t)x_{APPG}(t)\}, \quad X \in \mathbb{R}^{625\times4}$ where each signal is a 5-second segment sampled at 125 Hz (625 samples, 4 channels).
OUTPUT: $Y = \{S\hat{B}P, D\hat{B}P\}, \quad Y \in \mathbb{R}^2$
Initialization:

- Define U-Net encoder–decoder: 4 encoder + 4 decoder blocks, kernel size = 3, stride = 1.
- Define Transformer encoder: 12 layers, 8 attention heads, embedding dimension = 512.
- Define cross-attention modules between skip connections and Transformer outputs.
- Set training hyperparameters: learning rate η=0.0009, batch size = 32, optimizer = RMSProp, epochs = 100, early stopping patience = 20.

Steps:

1: Signal Acquisition: Load ECG, PPG, VPPG, APPG signals from datasets (MIMIC, Drink, Exercise).
2: Pre-processing: ECG:
- 0.5–30 Hz Butterworth filter (4th order).
- PPG/BP: 0.5–15 Hz filter.
- Resample to 125 Hz, normalize to [0,1].
- Segment into 5-second windows.

3: Encoding – Local Features:
- Feed X into U-Net encoder.
- Extract feature maps $F_e \in \mathbb{R}^{125\times64}$.

4: Global Context – Transformer:
- Input F_e into Transformer encoder.
- Compute attention:

$$A = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

- Generate contextual representation F_t.

5: Feature Fusion:
- Fuse U-Net skip connections with Transformer outputs using cross-attention:

$$F_f = \text{concat}(F_e, F_t)$$

6: Decoding:
- Pass F_f into U-Net decoder.
- Reconstruct predicted BP waveform $\hat{y}(t)$.

7: BP Value Extraction:
- Detect peaks of $\hat{y}(t)$.
- Derive $S\hat{B}P, D\hat{B}P$.

8: Training Loop:
- For epoch = 1 to 100:
  - Compute combined loss:

$$L = \alpha \cdot MAE + \beta \cdot (1 - PCC), \quad \alpha = 0.6, \beta = 0.4$$

  - Backpropagate gradients, update weights with RMSProp.
  - Validate on 20% of dataset.
  - Apply early stopping if no improvement in 20 epochs.

9: Fine-Tuning (Optional):
- Use 10–20% scenario-specific samples.
- Re-train last two Transformer blocks + decoder layers.

10: Evaluation:
- Report MAD, PCC, RMSE.
- Verify performance with AAMI standard (≤5 mmHg mean error, ≤8 mmHg SD).
- Grade performance with BHS protocol.
- Generate Bland–Altman plots for clinical interpretability.

11: Final Output:
- Return $Y = \{S\hat{B}P, D\hat{B}P\}$ with performance metrics.

3.2. Mathematical Components

To formally describe the proposed F-UTransBPNet, the key mathematical components are presented in this section. These equations complement the algorithmic workflow in

Section 3.1 and ensure that both local feature extraction (U-Net) and global context modeling (Transformer) are mathematically defined.

3.2.1. Multi-Head Self-Attention (MHSA)

The Transformer encoder models long-range temporal dependencies using self-attention. Given queries Q, keys K, and values V derived from input features, the attention function is defined as:

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

where d_k is the dimension of the key vector. Eq. (1) computes the weighted representation of input features, allowing the network to highlight important temporal dependencies across the physiological signals.

For multiple heads, the outputs are concatenated:

$$MHSA(Q,K,V) = \text{concat}(\text{head}_1, \dots, \text{head}_h)W^O \tag{2}$$

where each $\text{head}_i = Attention\left(QW_i^Q, KW_i^K, VW_i^V\right)$, and W^O is the projection matrix.

3.2.2. Loss Function

The model is trained to minimize a combined loss function that simultaneously accounts for prediction accuracy and correlation with the reference blood pressure values. This loss is expressed in Eq. (3):

$$L = \alpha \cdot MAE + \beta \cdot (1 - PCC) \tag{3}$$

where MAE denotes the mean absolute error, calculated as

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|\,,$$

and PCC represents the Pearson correlation coefficient, defined as:

$$PCC = \frac{\sum_i\left(\hat{y}_i - \bar{\hat{y}}_i\right)(y_i - \bar{y})}{\sqrt{\sum_i\left(\hat{y}_i - \bar{\hat{y}}_i\right)^2 \cdot \sum_i (y_i - \bar{y})^2}}.$$

The parameters α and β control the relative contribution of the two components and were empirically set to 0.6 and 0.4, respectively, in this study. By combining MAE and PCC, the loss function enforces both numerical proximity to the ground truth and strong correlation with physiological trends, thereby improving the clinical reliability of the predictions.

In summary, the proposed F-UTransBPNet integrates U-Net's hierarchical feature extraction with the Transformer's capability for capturing long-range dependencies, further optimized through a combined loss and RMSProp-based parameter updates. The algorithmic representation and mathematical formulation presented in this section ensure that the workflow is both reproducible and theoretically grounded. In the next section, we report the experimental evaluation of F-UTransBPNet on three benchmark datasets and analyze its performance under both static and activity-based scenarios, highlighting its advantages over existing approaches.

4. Results and Discussion

4.1. Experimental Setup

All experiments were implemented in PyTorch on an NVIDIA Tesla V100 GPU (32 GB memory). The proposed F-UTransBPNet was evaluated on three complementary datasets representing both static and dynamic blood pressure conditions. The MIMIC database provided invasive arterial BP recordings from 163 ICU patients, serving as the gold-standard reference for static monitoring. Two additional in-house datasets captured dynamic variations: Dataset_Drink, from 25 healthy adults undergoing a water intake protocol, and

Dataset_Exercise, from 20 participants performing graded bicycle ergometry. These datasets enable systematic evaluation across diverse hemodynamic states, as shown in Table 1.

Preprocessing was standardized across all datasets to ensure comparability. ECG signals were filtered between 0.5 and 30 Hz, and PPG signals between 0.5 and 15 Hz. All signals were resampled to 125 Hz, normalized to the range [0, 1], and segmented into 5-s windows, yielding input matrices of 625 samples × 4 channels (ECG, PPG, VPPG, APPG).

Table 1. Overview of datasets and experimental settings.

Dataset	Population	N (subjects)	Signal Types	Reference BP	Preprocessing Summary	Segments (5 s)
MIMIC	ICU patients	163	ECG, PPG, VPPG, APPG	Invasive arterial BP	0.5–30 Hz ECG, 0.5–15 Hz PPG; resample 125 Hz; normalization	15,474
Drink	Healthy adults	25	ECG, PPG, VPPG, APPG	Finapres (non-invasive)	0.5–30 Hz ECG, 0.5–15 Hz PPG; resample 125 Hz; normalization	62,678
Exercise	Healthy adults	20	ECG, PPG, VPPG, APPG	Finapres (non-invasive)	0.5–30 Hz ECG, 0.5–15 Hz PPG; resample 125 Hz; normalization	28,814

Table 2. Training configuration and evaluation metrics.

Parameter	Value/Method
Framework	PyTorch
Hardware	NVIDIA Tesla V100 GPU (32 GB)
Learning rate	0.0009
Batch size	32
Optimizer	RMSProp
Early stopping	20 epochs without improvement
Evaluation metrics	MAD, PCC, RMSE

Model training employed a learning rate of 0.0009, batch size of 32, and RMSProp optimizer, with early stopping applied after 20 epochs without improvement. The training configuration and evaluation metrics are summarized in Table 2. Model performance was assessed using mean absolute difference (MAD), Pearson correlation coefficient (PCC), and root mean square error (RMSE). Clinical acceptability was verified against the AAMI standard (≤5 mmHg mean error, ≤8 mmHg SD) and the BHS grading protocol, ensuring compliance with internationally recognized benchmarks for cuffless blood pressure estimation.

4.2. Performance Across Individual Datasets

Table 3 presents the quantitative performance of F-UTransBPNet on the three datasets. On the MIMIC cohort, which represents the gold-standard ICU monitoring environment, the model achieved mean absolute differences (MAD) of 4.4 mmHg for SBP and 2.2 mmHg for DBP, both within the AAMI acceptance thresholds. This demonstrates that the proposed architecture is clinically viable under static conditions, as confirmed by invasive reference measurements.

Table 3. Performance of F-UTransBPNet across individual datasets.

Dataset	Reference Type	MAD (SBP, mmHg)	MAD (DBP, mmHg)	PCC (SBP)	PCC (DBP)	Clinical Assessment
MIMIC	Invasive arterial BP	4.4	2.2	–	–	Meets AAMI criteria
Drink	Finapres (non-invasive)	–	–	0.61	0.62	Moderate tracking
Exercise	Finapres (non-invasive)	–	–	0.82	0.72	Strong tracking

For the activity datasets, model performance remained robust despite the physiological variability introduced by fluid intake and exercise. On Dataset_Drink, the model achieved Pearson correlation coefficients (PCC) of 0.61 (SBP) and 0.62 (DBP), reflecting moderate but consistent tracking capability. Performance further improved in Dataset_Exercise, where PCC increased to 0.82 (SBP) and 0.72 (DBP), indicating strong agreement with the reference signals even under hemodynamic stress.

Taken together, these results confirm that the hybrid U-Net–Transformer design is effective in capturing both short-range dependencies (local morphological features extracted by U-Net) and long-range dependencies (temporal dynamics captured by Transformer encoders). Importantly, the model demonstrated adaptability across static ICU monitoring and dynamic activity-induced fluctuations, highlighting its potential as a generalizable solution for cuffless blood pressure estimation.

4.3. Cross-Scenario Generalization

To better understand the robustness of F-UTransBPNet, we conducted cross-scenario experiments, where we trained the model on one dataset and tested it on another without any adjustments. The results, presented in Table 4, confirmed our expectations: the performance dropped significantly with direct transfer, with PCC values approaching zero. This illustrates the challenges of generalizing across diverse physiological states. It highlights the challenge of distributional shifts in cuffless blood pressure estimation, especially since signal patterns can vary significantly between static ICU monitoring and activity-related changes.

Table 4. Cross-scenario generalization performance of F-UTransBPNet.

Training Dataset	Testing Dataset	Fine-tuning (%)	PCC (SBP)	PCC (DBP)	MAD (SBP, mmHg)	MAD (DBP, mmHg)
Drink	Exercise	0%	0.12	0.08	–	–
Drink	Exercise	10%	0.76	0.68	–	–
Exercise	Drink	0%	0.15	0.10	–	–
Exercise	Drink	20%	0.71	0.65	–	–
Drink+Exercise	MIMIC	0%	0.05	0.03	7.2	4.9
Drink+Exercise	MIMIC	10%	0.62	0.55	4.9	2.8

To address this limitation, we performed fine-tuning with 10–20% of scenario-specific data, retraining only the last Transformer blocks and decoder layers. With this minimal adaptation, performance improved substantially. For example, when trained on the Drink dataset and fine-tuned with Exercise samples, PCC increased from 0.12 to 0.76 for SBP and from 0.08 to 0.68 for DBP. Similarly, fine-tuning with a small subset of MIMIC data yielded improvements in both MAD and PCC, restoring the model's clinical viability across domains.

These results demonstrate that while true calibration-free performance remains elusive, limited fine-tuning provides a practical pathway toward generalizable cuffless BP estimation. The Transformer modules likely facilitate transfer by capturing global dependencies that are preserved across scenarios, while U-Net layers adapt to scenario-specific local features.

4.4. Comparison with State-of-the-Art

To contextualize the performance of F-UTransBPNet, we compared it with state-of-the-art (SOTA) models, as shown in Table 5.

HGCTNet, proposed by Liu *et al.* [12], combines handcrafted feature guidance with CNN and Transformer modules. It achieved errors within both AAMI and BHS thresholds, although its reliance on engineered features increases design complexity. rU-Net, introduced by Chen *et al.* [13], integrates U-Net and ResNet with transfer learning, demonstrating superior cross-scenario adaptability and achieving BHS Grade A classification. However, its dependence on transfer learning introduces additional computational cost.

In contrast, Transformer-based single-modal approaches such as cBP-Tnet [14] reported MAEs of 4.3 mmHg (SBP) and 2.2 mmHg (DBP), fully satisfying AAMI criteria, but their reliance on single-channel PPG renders them vulnerable to noise and motion artifacts. MuFuBP-Net [15], employing dual-feature fusion and progressive enhancement, achieved

state-of-the-art performance across multiple cohorts with MAEs of 2.99 ± 4.37 mmHg (SBP) and 2.63 ± 4.19 mmHg (DBP), but at the expense of higher model complexity.

Table 5. Comparison of state-of-the-art cuffless BP estimation models.

Model	Architecture Type	Input Signals	Dataset(s)	MAE / MAD (SBP, mmHg)	MAE / MAD (DBP, mmHg)	Clinical Assessment
HGCTNet [12]	CNN + Transformer + handcrafted	PPG	Private cohort	0.7 ± 8.3	0.9 ± 6.5	Meets AAMI & BHS
rU-Net [13]	U-Net + ResNet + Transfer	PPG	Multi-cohort	4.49 ± 4.86	2.69 ± 3.10	Meets AAMI; BHS A
cBP-Tnet [14]	Transformer	PPG (single)	Public datasets	4.3	2.2	Meets AAMI
MuFuBP-Net [15]	Dual-feature fusion + PFE	PPG (dual features)	Multi-cohort	2.99 ± 4.37	2.63 ± 4.19	Meets AAMI & BHS
F-UTransBPNet (Ours)	U-Net + Transformer Hybrid	ECG, PPG, VPPG, APPG	MIMIC, Drink, Exercise	4.4	2.2	Meets AAMI; adaptable across scenarios

In comparison, F-UTransBPNet achieved competitive MAD values of 4.4 mmHg (SBP) and 2.2 mmHg (DBP) on the MIMIC dataset and maintained high PCC (>0.7) under activity-based scenarios. Unlike prior works, it leverages multi-modal signals (ECG, PPG, VPPG, APPG) and incorporates selective fine-tuning, enabling consistent adaptability across static and dynamic conditions. This balance between accuracy and computational feasibility represents a pragmatic advancement toward clinically deployable cuffless BP monitoring.

4.5. Clinical Implications

The results of this study show that F-UTransBPNet achieves mean absolute difference values within the thresholds set by the AAMI standard and demonstrates strong correlation with invasive references across both static and activity-based scenarios. These findings indicate that the proposed architecture has significant potential for clinical translation.

From an inpatient perspective, accurate and continuous non-invasive monitoring is crucial in intensive care settings where early detection of hemodynamic instability is essential. The ability of F-UTransBPNet to monitor blood pressure without the need for cuffs or invasive arterial lines could reduce patient discomfort, decrease the risk of infection, and provide clinicians with a reliable alternative in resource-limited or high-risk environments.

In outpatient and ambulatory care, integrating such models into wearable devices offers the possibility of long-term, unobtrusive monitoring. This could transform hypertension management by enabling continuous blood pressure variability, enhancing treatment adherence, and allowing for timely adjustments to therapy. Additionally, real-time data collection may support personalized medicine approaches, where therapy is customized based on individual hemodynamic responses during daily activities.

A significant implication is the model's demonstrated ability to adapt with minimal fine-tuning. Unlike traditional models that need extensive calibration for each subject or scenario, F-UTransBPNet delivers clinically acceptable performance with limited adaptation. This feature greatly improves its practicality for real-world use, where extensive per-patient calibration is not feasible.

Overall, the proposed approach could play a crucial role in addressing unmet needs in cardiovascular care by facilitating the early detection of hemodynamic instability, guiding fluid and medication management, and promoting long-term hypertension control. Its potential spans from ICU monitoring to home-based disease management, emphasizing its clinical relevance and translational potential.

4.6. Limitations and Future Directions

Although the proposed F-UTransBPNet shows promising performance across multiple datasets and activity scenarios, several limitations should be recognized. First, the datasets used in this study, though diverse, are limited in size and demographic scope. The MIMIC

cohort mainly represents critically ill ICU patients, while the Drink and Exercise datasets were collected from young, healthy adults. This results in underrepresented populations, including elderly patients, individuals with chronic cardiovascular conditions, and people from diverse ethnic backgrounds. Broader validation with heterogeneous cohorts is necessary to confirm true generalizability.

Second, although the hybrid U-Net–Transformer architecture effectively captures both local and global dependencies, the computational cost of the Transformer modules remains significant. While acceptable for research and high-performance GPU use, real-world deployment on wearable devices or mobile platforms may require model compression, pruning, or edge optimization techniques.

Third, while the fine-tuning strategy effectively restores cross-scenario performance, it still depends on limited scenario-specific data. Although this is a practical compromise compared to exhaustive calibration, future research should explore self-supervised learning, continual learning, or domain adaptation methods to minimize further or eliminate the need for recalibration.

Ultimately, clinical validation remains a crucial next step. This study benchmarked performance against established standards (AAMI, BHS), but prospective trials in actual clinical settings are necessary to evaluate usability, reliability, and clinical impact. This includes integrating with wearable hardware, assessing the stability of long-term monitoring, and examining patient adherence and comfort. In the future, research will focus on four key areas: (i) large-scale, multi-center validation across diverse populations; (ii) improving computational efficiency for wearable and edge deployment; (iii) developing adaptive learning strategies to enable calibration-free operation; and (iv) integrating into prospective clinical studies to evaluate translational feasibility and impact on patient outcomes.

6. Conclusions

This study introduced F-UTransBPNet, a hybrid U-Net–Transformer model for cuffless blood pressure estimation using multi-modal physiological signals. The model achieved clinically acceptable accuracy with invasive ICU references (MAD 4.4 mmHg for SBP, 2.2 mmHg for DBP) and maintained stable tracking during activity-induced variability (PCC up to 0.82 for SBP). These results confirm that combining the U-Net's local feature extraction with Transformer-based temporal modeling allows for reliable performance across both static and dynamic conditions.

Beyond accuracy, the strength of F-UTransBPNet lies in its ability to generalize with minimal fine-tuning, requiring only limited scenario-specific data to restore cross-domain performance. This sets it apart from previous models such as HGCTNet, rU-Net, and MuFuBP-Net, which depend heavily on handcrafted features, complex transfer learning pipelines, or increased computational demands. By balancing accuracy, adaptability, and computational efficiency, F-UTransBPNet provides a practical step toward deploying cuffless BP monitoring in clinical settings.

Clinically, these findings support the integration of hybrid AI architectures into critical care and wearable health technologies, with potential benefits including early detection of hemodynamic instability, improved hypertension management, and personalized therapy. However, broader validation on diverse populations, computational optimization for wearable devices, and prospective clinical trials are still necessary. Future research will address these aspects to move toward real-world, calibration-light, and scalable cuffless BP monitoring solutions.

Author Contributions: Conceptualization: M.L. and E.V.; Methodology: M.L.; Software: M.L.; Validation: M.L. and E.V.; Formal analysis: M.L.; Investigation: M.L.; Resources: E.V.; Data curation: M.L.; Writing: original draft preparation: M.L.; Writing: review and editing: M.L. and E.V.; Visualization: M.L.; Supervision: E.V.; Project administration: M.L.; Funding acquisition: E.V.

Funding: This research received no external funding.

Data Availability Statement: The MIMIC dataset is publicly available at PhysioNet (https://physionet.org/content/mimicdb/). The in-house Dataset_Drink and

Dataset_Exercise, containing sensitive data collected under institutional protocols, cannot be shared due to privacy restrictions. Access is possible upon reasonable request and ethical approval. No new datasets were created.

Acknowledgments: The authors thank the Department of Informatics, Universitas Muhammadiyah Semarang, for computational resources and support, and the Data Mining and Analytics Research Group at the International Hellenic University for feedback on experimental design. They also acknowledge the MIMIC database provided by PhysioNet. Grammarly was used for grammar checking and language refinement. All conceptualization, methodology, data analysis, and interpretation of results were conducted entirely by the authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] A. Chen *et al.*, "Disease burden of chronic kidney disease due to hypertension from 1990 to 2019: A global analysis," *Frontiers in Medicine*, vol. 8, p. 690487, 2021, doi: 10.3389/fmed.2021.690487.

[2] J. Liu *et al.*, "Global burden of cardiovascular diseases attributable to hypertension in young adults from 1990 to 2019," *Journal of Hypertension*, vol. 39, no. 12, pp. 2480–2490, 2021, doi: 10.1097/HJH.0000000000002958.

[3] P. Lloyd-Sherlock, J. Beard, N. Minicuci, S. Ebrahim, and S. Chatterji, "Hypertension among older adults in low- and middle-income countries: Prevalence, awareness and control," *International Journal of Epidemiology*, vol. 43, no. 1, pp. 116–128, 2014, doi: 10.1093/ije/dyt215.

[4] N. Pilz *et al.*, "Cuff-based blood pressure measurement: Challenges and solutions," *Blood Pressure*, 2024, doi: 10.1080/08037051.2024.2402368.

[5] X. Ding, B. P. Yan, Y.-T. Zhang, J. Liu, N. Zhao, and H. K. Tsang, "Pulse transit time based continuous cuffless blood pressure estimation: A new extension and a comprehensive evaluation," *Scientific Reports*, vol. 7, p. 11507, 2017, doi: 10.1038/s41598-017-11507-3.

[6] C.-T. Yen, S.-N. Chang, and C.-H. Liao, "Estimation of beat-by-beat blood pressure and heart rate from ECG and PPG using a fine-tuned deep CNN model," *IEEE Access*, vol. 10, pp. 80427–80437, 2022, doi: 10.1109/ACCESS.2022.3195857.

[7] F. Delrio, V. Randazzo, G. Cirrincione, and E. Pasero, "Non-invasive arterial blood pressure estimation from electrocardiogram and photoplethysmography signals using a Conv1D-BiLSTM neural network," *Engineering Proceedings*, vol. 39, no. 1, p. 78, 2023, doi: 10.3390/engproc2023039078.

[8] U. Senturk, K. Polat, and I. Yucedag, "A non-invasive continuous cuffless blood pressure estimation using dynamic recurrent neural networks," *Applied Acoustics*, vol. 167, p. 107534, 2020, doi: 10.1016/j.apacoust.2020.107534.

[9] 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, p. 30, 2023, doi: 10.20517/2574-1209.2023.30.

[10] H. Mou and J. Yu, "CNN-LSTM prediction method for blood pressure based on pulse wave," *Electronics*, vol. 10, no. 14, p. 1664, 2021, doi: 10.3390/electronics10141664.

[11] Y. Huang, Y. He, Z. Song, K. Gao, and Y. Zheng, "Validation of deep learning models for cuffless blood pressure estimation on a large benchmarking dataset," *Connected Health and Telemedicine*, vol. 2, no. 1, p. 23, 2024, doi: 10.20517/chatmed.2023.23.

[12] Z. Liu *et al.*, "HGCTNet: Handcrafted Feature-Guided CNN and Transformer Network for Wearable Cuffless Blood Pressure Measurement," in *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2024, pp. 1–8. doi: 10.1109/BIBM.2024.10516263.

[13] 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, doi: 10.3390/s21051867.

[14] A. A. Pimentel, J.-J. Huang, and A. R. A. See, "cBP-Tnet: Continuous blood pressure estimation using multi-task transformer network with automatic photoplethysmogram feature extraction," *Applied Sciences*, vol. 15, no. 14, p. 7824, 2025, doi: 10.3390/app15147824.

[15] X. Zhang *et al.*, "MuFuBP-Net: A Multimodal Fusion Network for Cuffless Blood Pressure Estimation Using Dual-Feature Pipeline with Probabilistic Feature Encoder," in *2025 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2025, pp. 1–5. doi: 10.1109/ISCAS.2025.10975118.

[16] N. Pilz, A. Patzak, and T. L. Bothe, "Continuous cuffless and non-invasive measurement of arterial blood pressure—concepts and future perspectives," *Blood Pressure*, vol. 31, no. 5, pp. 324–334, 2022, doi: 10.1080/08037051.2022.2128716.

[17] M. M. R. K. Mamun and A. Sherif, "Advancement in the cuffless and noninvasive measurement of blood pressure: A review of the literature and open challenges," *Bioengineering*, vol. 10, no. 1, p. 27, 2022, doi: 10.3390/bioengineering10010027.

[18] B. Henry, M. Merz, H. Hoang, G. Abdulkarim, J. Wosik, and P. Schoettker, "Cuffless blood pressure in clinical practice: Challenges, opportunities and current limits," *Blood Pressure*, 2024, doi: 10.1080/08037051.2024.2304190.

[19] L. Zhao *et al.*, "Emerging sensing and modeling technologies for wearable and cuffless blood pressure monitoring," *npj Digital Medicine*, vol. 6, p. 71, 2023, doi: 10.1038/s41746-023-00835-6.

[20] J. Hua, M. Su, J. Wu, Y. Zhou, Y. Guo, Y. Shi, and L. Pan, "Wearable cuffless blood pressure monitoring: From flexible electronics to machine learning," *Wearable Electronics*, vol. 5, p. 100028, 2024, doi: 10.1016/j.wees.2024.05.004.

[21] M. Nour, K. Polat, Ü. Şentürk, and M. Arıcan, "A novel cuffless blood pressure prediction: Uncovering new features and new hybrid ML models," *Diagnostics*, vol. 13, no. 7, p. 1278, 2023, doi: 10.3390/diagnostics13071278.