INDIVIDUAL-LEVEL PUBLIC-PRIVATE DECOUPLING FOR CONTINUOUS BLOOD PRESSURE WAVEFORM PREDICTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Existing methods for continuous arterial blood pressure (ABP) waveform prediction from photoplethysmography (PPG) do not adequately distinguish inter-individual differences in blood pressure. To address this limitation, we propose Individual-level Public-Private Decoupling (IPPD), a network that decomposes PPG representations into statistically decorrelated Public and Private features and expresses the ABP prediction as the sum of a public waveform and an individual residual. To better distinguish within-individual consistency from between-individual variation, we introduce dual-view contrastive learning into the Private branch. On VitalDB-360, IPPD achieves the best individual-level prediction performance. Without target-case calibration or fine-tuning, it attains a waveform RMSE of 11.65 mmHg and a waveform MAE of 9.91 mmHg across 30 held-out test cases, outperforming existing blood-pressure prediction baselines. The code will be released.

1 INTRODUCTION

Blood pressure is a key indicator of human health. As demand for continuous health monitoring grows, future systems are expected to become increasingly noninvasive and personalized. Cuffless photoplethysmography (PPG) has therefore emerged as an important modality for noninvasive blood-pressure estimation (Ballinger et al., 2018). PPG provides a potential route to portable, continuous hemodynamic monitoring and makes continuous arterial blood pressure (ABP) estimation possible. Unlike scalar prediction of systolic blood pressure (SBP) or diastolic blood pressure (DBP), continuous ABP waveform prediction requires a model to recover a pointwise pressure sequence from PPG (Pillai et al., 2025). The model must estimate the overall pressure level while preserving temporal structure and amplitude variation. Accurate window-level SBP or DBP estimates therefore do not necessarily imply accurate reconstruction of the complete ABP waveform.

When deployed to unseen cases, a model must contend with inter-individual variability in the PPG-to-ABP mapping. Training data may contain both shared patterns that generalize across cases and individual information specific to the current case (Jia et al., 2024). If a model learns only a single prediction pathway, these two sources of information may become entangled in the same representation, making it difficult to preserve a stable shared mapping while capturing individual variation. Foundation models such as UniCardio (Chen et al., 2025) predict ABP waveforms from four-second windows, but their training and test protocols do not explicitly distinguish individuals. Consequently, they do not explicitly model inter-individual variation or report performance for each individual case. When shared and individual information are entangled, the model may struggle to determine which prediction components should remain stable across cases and which should vary with the PPG characteristics of the current case.

An alternative is to calibrate or fine-tune the model using paired PPG-ABP samples from each new case. Although this may improve adaptation to the target case, it requires additional labeled ABP data and increases the data-acquisition and adaptation costs of deployment (Jia et al., 2024). If ABP labels are obtained through invasive arterial-pressure monitoring, the procedure also increases the burden on the patient. Few-shot calibration may fail to cover within-case blood-pressure variation, whereas test-time fine-tuning requires additional computation and quality control and may become

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

unstable or overfit when data are scarce or the distribution changes. Moreover, data collection and adaptation must be repeated for every new case.

We propose IPPD, a dual-branch network for Individual-level Public-Private Decoupling in continuous blood-pressure waveform prediction. IPPD separates individual-sensitive information from shared information to support case-specific PPG-to-ABP waveform estimation without calibration or fine-tuning. Our main contributions are as follows:

(1) We propose IPPD, a continuous ABP waveform prediction architecture with Public and Private branches. The Public branch predicts a shared baseline waveform, whereas the Private branch predicts an individual-sensitive correction. Their fusion produces the individual-level ABP waveform prediction.

(2) We develop a dual-view contrastive learning strategy that explicitly distinguishes within-individual consistency from between-individual variation. Two independently augmented views of the current individual's PPG window form a positive pair, encouraging the model to focus on morphological properties that remain stable within the individual and improving generalization to unseen individuals. Features from different individuals serve as negative samples, increasing the separation between individual representations and preventing representations from different individuals from becoming excessively similar.

(3) We design a joint objective for individual-level decoupling. A waveform reconstruction loss constrains the primary prediction, an identity-classification loss helps the model distinguish training cases, and a difference loss decorrelates the Public and Private representations. IPPD enables accurate prediction for unseen individuals without fine-tuning.

2 Related Work

Blood Pressure Scalar Regression. Early studies typically extracted handcrafted features from PPG, including pulse intervals, rise time, waveform amplitude, area, slope, and frequency-domain statistics, and then regressed SBP, DBP, mean arterial pressure (MAP), or pulse pressure (PP). With the development of deep learning, convolutional neural networks, recurrent neural networks, and temporal Transformers have increasingly been used to learn blood-pressure representations directly from raw PPG. Ballinger et al. (2018) used a semi-supervised multitask LSTM to predict cardiovascular risks, including hypertension, from large-scale wearable heart-rate sequences, showing that pretraining on unlabeled physiological sequences can outperform handcrafted biomarkers. Zhang et al. (2026) transferred language-modeling principles to PPG and jointly modeled temporal dynamics and local and global morphology, extending the design space for PPG foundation models. To address temporal misalignment in supervisory signals, Hong et al. (2026) introduced phase correction and bi-level meta-learning, improving the reliability of blood-pressure estimation across datasets.

Continuous Waveform Prediction. Continuous waveform prediction takes a segment of PPG as input and outputs an ABP sequence aligned with the same time axis. Its value lies in recovering the complete time-varying blood-pressure pulse morphology rather than only scalar values such as SBP and DBP within a window, thereby retaining richer details of blood-pressure variation. To address latent-space shifts caused by individual and hardware differences, Bian et al. (2024) introduced a quantized codebook and multispectral knowledge enhancement to improve reconstruction stability for PPG-to-ABP transformation. Kundu et al. (2025) used an invertible network to learn a bidirectional mapping between PPG and ABP gradients; recovering high-frequency details such as the systolic upstroke improved blood-pressure accuracy. Pillai et al. (2025) developed an encoder based on a one-dimensional ResNet and a mixture-of-experts architecture, together with morphology-aware self-supervised objectives for improved representation learning. Kim & Kim (2026) jointly used PPG and acceleration-derived PPG and applied Mamba to model long-range dependencies, producing ABP estimates that more closely resemble clinically interpretable waveforms. Chen et al. (2025) treated multiple cardiovascular signals as observations of a common physiological process and used a unified diffusion model (Peebles & Xie, 2023) for cross-modal generation, imputation, and translation. However, diffusion models (Sohl-Dickstein et al., 2015) often require many sampling steps (Song & Ermon, 2019), resulting in substantial computational and deployment costs (Song et al., 2021a). Flow-matching models (Lipman et al., 2023; Geng et al., 2025), which can operate with fewer sampling steps, may therefore receive increasing attention (Calvo-Ordóñez et al.,

2026). Compared with multimodal models, IPPD requires only the current PPG window, does not depend on other physiological modalities, focuses specifically on PPG-to-ABP prediction, and explicitly models shared structure and individual variation.

Individual Feature Decoupling. Feature decoupling maps distinct factors of variation in an input to separate representational subspaces. This idea has been widely used in domain adaptation, personalized modeling, style-content separation, and multimodal representation learning. Wang et al. (2024b) learned subject-stable EEG representations through mutual reconstruction with multiple decoders and time-step perturbations. Wang et al. (2024a) trained a single cross-subject model using an aggregation function and cyclic fMRI reconstruction, and adapted it to new individuals through pseudo-data augmentation. Chen et al. (2024) decomposed time-series representations into domain-invariant and domain-specific components and used an intra-class hypersphere loss to improve transfer from source to target domains. Jia et al. (2024) performed both modality-level and subject-level decoupling to extract modality-shared and modality-specific information as well as subject-shared and subject-private information, providing a clear framework for decoupling in cross-individual blood-pressure estimation. Tian et al. (2025) jointly trained global and local models. Gong et al. (2025) used visual fingerprints and Low-Rank Adaptation (LoRA) to learn features for new individuals and employed images as an intermediate modality for semantic correction. Liu et al. (2025) connected shallow subject adapters to a shared deep decoder and combined visual and textual supervision to improve few-shot cross-subject decoding. Bao et al. (2025) first performed anatomical alignment and then used a mixture of brain-expert adapters to handle individual differences. Wang et al. (2025) explicitly separated subject-related and semantic factors through adversarial training, enabling zero-shot generalization without data from new individuals. Liu et al. (2026) combined multiple individual-specific encoders with a shared encoder to support rapid adaptation to new individuals from limited data. Weng et al. (2026) modeled EEG state transitions with a spatiotemporal state-space model and learned cross-subject representations through global-local attention. Lu et al. (2026) modeled multilevel information according to the hierarchy of the visual cortex and enhanced subject-independent representations through feature-level contrastive learning. IPPD does not discard individual variation; instead, it treats such variation as useful guidance for individual-specific prediction. Dual-view contrastive learning captures this information, while the joint loss constrains its contribution to the final prediction.

3 METHODOLOGY

3.1 PROBLEM FORMULATION

We partition the set of cases into three mutually exclusive subsets for training, validation, and testing, with no case IDs shared between any two subsets. The split is performed at the case level, so all windows from the same case remain within a single subset. This protocol evaluates whether a model trained on observed cases can generalize to unseen cases. During training, the model is supervised using the PPG and ABP waveforms of the training cases. For each PPG window x from a training case, the model predicts the corresponding ABP waveform $y_{pred} = f(x)$, where f denotes the IPPD waveform-prediction function. Training minimizes the mean squared error (MSE) between the predicted waveform y_{pred} and the ground-truth waveform y, while seeking to preserve generalization to new individuals at test time.

3.2 IPPD FRAMEWORK

As shown in Figure 1, IPPD is designed to model separately the shared information in PPG that can be reused across cases and the individual-sensitive information useful for prediction in the current case. IPPD takes a single PPG window as input. A ResNet1D encoder first extracts shared features while preserving temporal structure, after which the features are passed to the Public and Private branches. The Public branch generates a shared baseline waveform, and the Private branch generates an individual-sensitive waveform correction. A smoothing convolution fuses the two components to produce the final ABP prediction. The identity classifier supervises the Private representation only during training and does not contribute to the final waveform at test time. IPPD is applied directly to unseen individuals without any fine-tuning.

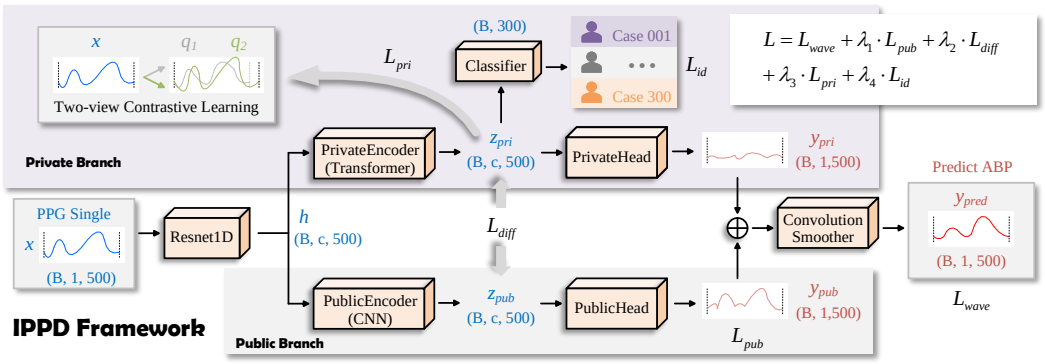


Figure 1: Framework of the proposed IPPD.

3.3 Public Branch

The Public branch learns a relatively stable and reusable PPG-to-ABP mapping from the shared temporal feature h and generates the shared baseline waveform. The shared feature h is mapped to a Public temporal representation:

$$z_{pub} = E_{pub}(h), \tag{1}$$

where E_{pub} is a temporal convolutional encoder comprising two one-dimensional convolutional layers with a kernel size of 3 and a stride of 1. The first convolution is followed by Group Normalization, GELU, and Dropout, and the second convolution is followed by another Group Normalization layer. The number of channels remains unchanged, and symmetric padding preserves the temporal length. This branch extracts information that can be reused across cases from the shared temporal features, thereby establishing a stable Public prediction pathway.

The Public waveform head uses a 1×1 Conv1d layer to project the channel-wise feature vector at each time point to a single blood-pressure value:

$$y_{pub} = G_{pub}(z_{pub}), \tag{2}$$

where G_{pub} denotes a temporal convolution that mixes channels at each time point without mixing adjacent time points. The Public branch therefore outputs a 500-sample shared baseline waveform, y_{pub}, which anchors the final waveform prediction.

To ensure that the Public branch learns a relatively stable baseline prediction, we impose auxiliary supervision on the shared baseline waveform:

$$L_{pub} = MSE(y_{pub}, y). \tag{3}$$

L_{pub} stabilizes and constrains the Public pathway, whereas the final prediction is still determined jointly by the shared baseline waveform and the individual residual correction.

3.4 Private Branch

The Private branch extracts temporal information related to individual physiology and waveform morphology from the current PPG window and transforms it into an individual-sensitive correction to the shared baseline waveform. Unlike methods that construct a fixed embedding for each target case in advance, IPPD dynamically generates the Private features from the current window:

$$z_{pri} = E_{pri}(h), \tag{4}$$

where E_{pri} is a temporal convolutional encoder with the same architecture as E_{pub}. This branch separates case-specific information from the shared temporal features that helps distinguish different individuals, providing a basis for subsequent identity classification and waveform correction.

The Private waveform head uses a 1×1 Conv1d layer to project the channel-wise feature vector at each time point to a single blood-pressure value:

$$y_{pri} = G_{pri}(z_{pri}), \tag{5}$$

where G_{pri} denotes a temporal convolution with the same architecture as G_{pub}. The weights and bias of the Private waveform head are initialized to zero, so the model initially predicts from the Public baseline waveform and gradually learns the required individual-sensitive correction. This initialization prevents a randomly generated Private correction from disrupting the Public prediction pathway at the beginning of training.

To prevent the two branches from learning highly redundant features, we regularize them by computing the channel-wise cross-covariance matrix between z_{pub} and z_{pri} and penalizing the mean squared value of its elements:

$$L_{diff} = Mean\left[\left(\frac{z_{pub}^{\top} z_{pri}}{B\,(N-1)}\right)^{2}\right], \tag{6}$$

where $N = BT$ is the total number of batch-time observations, B is the batch size, T is the sequence length, and $N - 1$ is the degrees-of-freedom correction used for an unbiased sample-covariance estimate. This loss reduces second-order statistical correlation between the Public and Private features and discourages the two branches from redundantly encoding the same information.

3.5 Dual-View Contrastive Learning

In addition to identity-classification supervision, IPPD applies dual-view morphological contrastive learning to the Private branch. This strategy improves the stability of the Private representation under perturbations in amplitude, baseline, and temporal scale while retaining its ability to distinguish inputs from different cases.

For each standardized PPG window, two independently augmented views are generated during training. The augmentations include amplitude scaling, baseline shifting, and temporal-speed variation. The amplitude-scaling range is $[0.5, 1.5]$, the baseline-shift range is $[-0.5, 0.5]$, and the speed range is $[0.5, 1.5]$. Temporal-speed variation is implemented by linear interpolation around the center of the window. The two views are passed through ResNet1D and the Private encoder to obtain Private features. A shape projector then maps each Private feature to a high-dimensional descriptor q_i. The descriptors from the two views of the same original window, $q_i^{(1)}$ and $q_i^{(2)}$, form a positive pair, whereas descriptors q_j from windows of other cases in the batch serve as negative samples. IPPD uses a symmetric InfoNCE loss to constrain this contrastive learning process:

$$L_{pri} = -\frac{1}{2B}\sum_{i=1}^{B}\left[\log\frac{exp\left(sim(q_{i_1}, q_{i_2})/\tau\right)}{\sum_{j=1}^{B} exp\left(sim(q_{i_1}, q_j)/\tau\right)} + \log\frac{exp\left(sim(q_{i_2}, q_{i_1})/\tau\right)}{\sum_{j=1}^{B} exp\left(sim(q_{i_2}, q_j)/\tau\right)}\right], \tag{7}$$

where τ is the temperature coefficient. The symmetric bidirectional InfoNCE loss brings the two augmented views of the same case closer in the descriptor space while separating them from the descriptors of other cases. Dual-view contrastive learning is used only during training; no augmented views are generated at inference time.

3.6 Identity Classifier

To encourage the Private branch to encode case-related information, IPPD introduces a case-identity classifier during training. Given the Private temporal feature z_{pri}, adaptive average pooling first divides the temporal dimension into k segments and aggregates the features within each segment. The pooled representation is flattened and passed to a multilayer perceptron (MLP) to produce the identity logits l_{id}:

$$l_{id} = MLP(z_{pri}). \tag{8}$$

The identity classifier uses auxiliary supervision to make the Private representation sensitive to differences among training cases. Identity-classification accuracy is used only to diagnose whether the Private representation contains information related to training cases and should not be interpreted as identity-recognition performance on unseen cases. The identity-classification loss is computed as cross-entropy over the 300 training-case classes:

$$L_{id} = CE(l_{id}, s), \tag{9}$$

where s denotes the training-case label. L_{id} is used only during training to encourage z_{pri} to retain case-related information.

The identity classifier is optimized using the identities of the training cases and does not contribute directly to the final ABP waveform. At test time, neither the target-case identity label nor the predicted identity is used to select a prediction model or adjust the waveform.

3.7 Loss Function

The IPPD training objective comprises five terms: final waveform reconstruction, Public baseline-waveform supervision, case-identity classification, Public-Private difference regularization, and Private dual-view morphological contrastive learning:

$$L = L_{wave} + \lambda_1 \cdot L_{pub} + \lambda_2 \cdot L_{diff} + \lambda_3 \cdot L_{pri} + \lambda_4 \cdot L_{id}, \tag{10}$$

where L_{wave} is the pointwise MSE between the final prediction and the normalized ground-truth ABP and serves as the primary objective. Its weight is 1.0, whereas each of the four auxiliary losses is weighted by 0.01, keeping waveform reconstruction dominant while guiding the separation of the Public and Private branches. Identity classification and dual-view contrastive learning are used only during training; at test time, the model predicts ABP from the current PPG window without fine-tuning.

4 Experiments

4.1 Dataset and Evaluation Metrics

We curated the VitalDB-360 dataset from VitalDB (Lee et al., 2022). It contains 360 cases, with an average of 2.7 hours of paired PPG–ABP data per case. Each case was segmented into paired PPG–ABP windows of 500 samples (4 s at 125 Hz), yielding approximately 2,430 window pairs per case on average. Cases were split by case ID into mutually exclusive training, validation, and test sets containing 300, 30, and 30 cases, respectively. All windows from the 30 test cases were used for evaluation, and no case IDs overlapped across the three sets. This protocol assesses prediction performance on unseen cases. Validation and test PPG windows were passed directly to the model to generate ABP estimates, without target-case calibration or fine-tuning.

We evaluated performance at two levels: continuous-waveform reconstruction and blood-pressure statistics. The waveform metrics were Wave RMSE, Wave MAE, and KS, which measure the overall pointwise error, the mean absolute pointwise error, and the difference between the amplitude distributions of the predicted and reference ABP waveforms, respectively. We also extracted SBP, DBP, MAP, and PP from each predicted waveform and computed MAE and RMSE for each statistic. For every metric, we first aggregated results over all test windows within each case and then computed an equally weighted macro-average across the 30 test cases, avoiding bias from differences in the number of windows per case.

4.2 Experimental Setup

Training used a distinct-case balanced sampler with a batch size of 256. We optimized IPPD with AdamW, a learning rate of 10^{-3}, and a weight decay of 10^{-5}, and clipped gradients at 1.0. Training ran for up to 100 epochs, with early stopping if validation waveform MSE failed to improve for 10 consecutive epochs. All experiments were run on a single NVIDIA RTX 5090 GPU with 32 GB of memory.

4.3 Main Results

We compared IPPD with DiT (Peebles & Xie, 2023), FlowMatch (Lipman et al., 2023), UniCardio (Chen et al., 2025), and PaPaGei (Pillai et al., 2025), as shown in Table 1. Methods marked with an asterisk were reproduced by us under the same case-level protocol. Under the case-disjoint test protocol, IPPD achieved the best results on all seven reported metrics: Wave RMSE, Wave MAE, KS, and the MAEs for SBP, DBP, MAP, and PP. Its Wave RMSE, Wave MAE, and KS were

Table 1: Comparison of the main results of IPPD with other methods (mmHg).

Method	Wave RMSE	Wave MAE	KS-Test	SBP	DBP	MAP	PP
DiT*	16.1336 ± 0.1835	13.6439 ± 0.1741	0.3303 ± 0.0036	17.5970	14.5611	11.0150	21.4589
FlowMatch*	16.3542 ± 0.2074	14.2837 ± 0.1995	0.3968 ± 0.0040	18.4062	10.6331	12.3340	15.0896
Unicardio*	13.3541 ± 0.6041	11.5206 ± 0.5727	0.3523 ± 0.0135	16.8356	8.4487	9.7333	15.6391
PaPaGei*	12.2069 ± 0.1462	10.4336 ± 0.1367	0.3047 ± 0.0029	13.9945	7.5914	8.7984	12.1848
IPPD (Ours)	$\mathbf{11.6505 \pm 0.1399}$	$\mathbf{9.9183 \pm 0.1303}$	$\mathbf{0.3024 \pm 0.0028}$	**13.4409**	**7.3095**	**8.2975**	**12.0572**

* denotes the method was reproduced by ourselves.

Table 2: Ablation results of each module in IPPD (mmHg).

Method	Public Branch (L_{pub})	Private Branch (L_{diff})	Two View (L_{pri})	Classifier (L_{id})	Wave RMSE	Wave MAE
IPPD	✓				12.4016	10.5836
	✓	✓			12.2273	10.4336
	✓	✓	✓		12.0526	10.2349
	✓	✓	✓	✓	11.6505	9.9183

11.6505 mmHg, 9.9183 mmHg, and 0.3024, respectively. Compared with PaPaGei, the closest competing method, IPPD reduced Wave RMSE and Wave MAE by 4.56% and 4.94%, respectively, and lowered KS from 0.3047 to 0.3024. These results indicate improved pointwise waveform reconstruction and closer agreement between amplitude distributions over a strong baseline. Relative to the updated UniCardio results, IPPD reduced Wave RMSE and Wave MAE by 12.76% and 13.91%, respectively; relative to DiT and FlowMatch, it reduced Wave RMSE by 27.79% and 28.76%, respectively. For blood-pressure statistics, IPPD achieved the lowest MAEs for SBP, DBP, MAP, and PP, at 13.4409, 7.3095, 8.2975, and 12.0572 mmHg, respectively. Compared with PaPaGei, these four MAEs improved by 3.96%, 3.71%, 5.69%, and 1.05%, respectively, with the largest relative improvement in MAP MAE. Compared with UniCardio, the corresponding reductions were 20.16%, 13.48%, 14.75%, and 22.90%. Overall, the results support the design in which the Public branch models a baseline waveform shared across cases and the Private branch learns an input-dependent, individual-sensitive residual correction, improving both continuous ABP reconstruction and the blood-pressure statistics derived from the predicted waveforms.

4.4 Ablation Results

Table 2 reports ablations of the IPPD components. Starting from a Public-only prediction baseline, we sequentially added the Private branch, dual-view contrastive learning, and the case-identity classifier, together with their associated loss terms. Wave RMSE and Wave MAE decreased at each step. The Public-only model achieved a Wave RMSE of 12.4016 mmHg and a Wave MAE of 10.5836 mmHg. Adding the Private branch reduced these values to 12.2273 mmHg and 10.4336 mmHg, respectively, suggesting that modeling individual-sensitive features can improve prediction. Adding the dual-view constraint further reduced the errors to 12.0526 mmHg and 10.2349 mmHg, consistent with the value of multi-view information for waveform representation learning. The complete model, including the classifier, achieved the best results, with a Wave RMSE of 11.6505 mmHg and a Wave MAE of 9.9183 mmHg. These ablations provide empirical support for the proposed framework and the contributions of its components and associated losses.

5 Conclusion

We proposed IPPD, an individual-level Public–Private decoupling method for continuous blood-pressure waveform prediction that decomposes PPG representations into Public and Private features. The Public branch learns a baseline waveform shared across cases, whereas the Private branch extracts individual-sensitive information from the current PPG window and generates a waveform correction. The two components are smoothly fused to produce the final continuous ABP prediction. On 30 test cases, IPPD achieved low waveform errors and outperformed the compared methods. Ablation results support the effectiveness of the Public–Private two-branch decoupling strategy and the

contribution of the individual loss terms. Future work will investigate representation transfer, cross-case interventions, and stronger decoupling constraints to assess the stability and interpretability of the Public–Private roles.

AI USE STATEMENT

In this work, we used generative AI tools for language polishing of the manuscript and for converting the manuscript into the ICLR LaTeX template. We have not used generative AI tools for the design of the method, the experiments, or the analysis of the results, and data collection is not applicable to this work since only the publicly available VitalDB database was used. We have reviewed and verified all AI-assisted work. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This study uses only the publicly available VitalDB database, which was collected with approval from the relevant institutional ethics committee. No new human-subject experiments were conducted, and no identifiable patient information is used or reported. The proposed method is intended for research on noninvasive physiological monitoring; clinical deployment would require further validation, and the predicted waveforms should not be used for medical decision making without proper regulatory approval.

REPRODUCIBILITY STATEMENT

The VitalDB-360 dataset is curated from the publicly available VitalDB database, and the case-level splitting protocol is described in Section 4.1. The full architecture, training objective, and hyperparameters are given in Sections 3 and 4.2. The source code will be released to facilitate reproduction of all reported results.

REFERENCES

B. Ballinger, J. Hsieh, A. Singh, N. Sohoni, J. Wang, G. H. Tison, G. M. Marcus, J. M. Sanchez, C. Maguire, J. E. Olgin, and M. J. Pletcher. DeepHeart: Semi-supervised sequence learning for cardiovascular risk prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32(1), 2018. https://doi.org/10.1609/aaai.v32i1.11891.

G. Bao, Q. Zhang, Z. Gong, J. Zhou, W. Fan, K. Yi, U. Naseem, L. Hu, and D. Miao. Wills Aligner: Multi-subject collaborative brain visual decoding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2025.

C. Bian, X. Li, Q. Bi, G. Zhu, J. Lyu, W. Zhang, Y. Li, and Z. Zeng. Constraint latent space matters: An anti-anomalous waveform transformation solution from photoplethysmography to arterial blood pressure. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38(10), pp. 11087–11095, 2024. https://doi.org/10.1609/aaai.v38i10.28985.

S. Calvo-Ordóñez, M. Meunier, Á. Cartea, C. Reisinger, Y. Gal, and J. M. Hernández-Lobato. Weighted conditional flow matching. In *International Conference on Learning Representations (ICLR)*, 2026.

Y. Chen, X. Yan, Y. Yang, J. Zhang, J. Zhang, L. Pan, and J. Li. Decoupling domain and general representations for time series classification. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2024.

Z. Chen, Y. Miao, L. Wang, L. Fan, D. P. Mandic, and J. Zhu. Versatile cardiovascular signal generation with a unified diffusion transformer. *Nature Machine Intelligence*, 2025. https://doi.org/10.1038/s42256-025-01147-y.

Z. Geng, M. Deng, X. Bai, J. Z. Kolter, and K. He. Mean flows for one-step generative modeling. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2025.

Z. Gong, Q. Zhang, G. Bao, L. Zhu, R. Xu, K. Liu, L. Hu, and D. Miao. MindTuner: Cross-subject visual decoding with visual fingerprint and semantic correction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39(13), pp. 14247–14255, 2025. https://doi.org/10.1609/aaai.v39i13.33560.

Q. Hong, C. Bian, X. Zhou, X. Li, Y. Li, and Z. Zeng. Lost in time? a meta-learning framework for time-shift-tolerant physiological signal transformation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2026.

Z. Jia, F. Zhao, Y. Guo, H. Chen, and T. Jiang. Multi-level decoupling network for cross-subject emotion recognition based on multimodal physiological signals. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2024.

S. W. Kim and J. Kim. Physio-Mamba: Physiologically-guided Mamba network for continuous calibration-free ABP reconstruction from PPG. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2026.

S. Kundu, G. Panda, S. Bhattacharya, A. Routray, and R. Guha. INN-PAR: Invertible neural network for PPG to ABP reconstruction. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2025.

H.-C. Lee, Y. Park, S. B. Yoon, S. M. Yang, D. Park, and C.-W. Jung. VitalDB, a high-fidelity multi-parameter vital signs database in surgical patients. *Scientific Data*, 9:279, 2022. https://doi.org/10.1038/s41597-022-01411-5.

Y. Lipman, R. T. Q. Chen, H. Ben-Hamu, M. Nickel, and M. Le. Flow matching for generative modeling. In *International Conference on Learning Representations (ICLR)*, 2023.

X.-H. Liu, Y.-K. Liu, T. Zhou, B.-L. Lu, and W.-L. Zheng. MindCross: Fast new subject adaptation with limited data for cross-subject video reconstruction from brain signals. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2026.

Y. Liu, Y. Ma, G. Zhu, H. Jing, and N. Zheng. See through their minds: Learning transferable brain decoding models from cross-subject fMRI. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39(6), pp. 5730–5738, 2025. https://doi.org/10.1609/aaai.v39i6.32611.

J. Lu, H. Wang, Q. Zhang, and X. Li. A cognitive process-inspired architecture for subject-agnostic brain visual decoding. In *International Conference on Learning Representations (ICLR)*, 2026.

W. Peebles and S. Xie. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023.

A. Pillai, D. Spathis, F. Kawsar, and M. Malekzadeh. PaPaGei: Open foundation models for optical physiological signals. In *International Conference on Learning Representations (ICLR)*, 2025.

J. Sohl-Dickstein, E. A. Weiss, N. Maheswaranathan, and S. Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2015.

J. Song, C. Meng, and S. Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations (ICLR)*, 2021a.

Y. Song and S. Ermon. Generative modeling by estimating gradients of the data distribution. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.

Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations (ICLR)*, 2021b.

Z. Tian, R. Quan, F. Ma, K. Zhan, and Y. Yang. BRAIN GUARD: Privacy-preserving multisubject image reconstructions from brain activities. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2025.

H. Wang, J. Lu, H. Li, and X. Li. ZEBRA: Towards zero-shot cross-subject generalization for universal brain visual decoding. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2025.

S. Wang, S. Liu, Z. Tan, and X. Wang. MindBridge: A cross-subject brain decoding framework. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11333–11342, 2024a. https://doi.org/10.1109/CVPR52733.2024.01077.

Y. Wang, B. Zhang, and Y. Tang. DMMR: Cross-subject domain generalization for EEG-based emotion recognition via denoising mixed mutual reconstruction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38(1), pp. 628–636, 2024b. https://doi.org/10.1609/aaai.v38i1.27819.

W. Weng, Y. Gu, Y. Ma, Y. Liu, Y. Zhang, and Y. Chen. State Mamba: Spatiotemporal EEG state-space model with dynamic brain alignment for cross-subject representation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2026.

Z. Zhang, H. Lu, and Q. Zhao. PPGPT: Transferring next-token modeling from language to PPG signals. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2026.