Physio-xLSTM: Unraveling Hemodynamic Dynamics via Dual-Domain KAN-xLSTM Networks for Robust PPG Biometrics

Yanchao Xiao[1,2], Chunxiao Wang[1,2,*], Yuwen Huang[3,*], Yue Zheng[4], Wenhao Li[1,2], Hao Wang[1,2], Wenzhe Zhang[1,2], Ziqiang Liu[1,2], Zhiwei Zhou[1,2]

[1]Key Laboratory of Computing Power Network and Information Security, Ministry of Education, Shandong Computer Science Center (National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan, Shandong, China
[2]Shandong Provincial Key Laboratory of Computing Power Internet and Service Computing, Shandong Fundamental Research Center for Computer Science, Jinan, Shandong, China
[3]School of Computer, Heze University, Heze, China
[4]School of Information Science and Engineering, Linyi University, Linyi, China
[*]Corresponding authors: wangchx@sdas.org, huangyuwen@hezeu.edu.cn

***Abstract*—Photoplethysmography (PPG) offers a secure and non-invasive modality for wearable biometrics. However, robust identification in the wild remains challenging because the non-linear generative dynamics of hemodynamic signals are not fully captured by existing efficient architectures. While State Space Models (SSMs) and Transformers have advanced long-range dependency modeling, they still exhibit notable limitations: standard SSMs (e.g., Mamba) compress history into fixed-size vector states, which can create an information bottleneck and obscure fine-grained morphological details. Conversely, standard Multi-Layer Perceptrons (MLPs) rely on piecewise linear activations, which are not ideal for approximating the continuous and smooth manifolds induced by vascular elasticity. To alleviate these limitations, we propose *Physio-xLSTM*, a physics-informed framework that combines the Matrix Memory of Extended LSTM (xLSTM) with the functional approximation capability of Kolmogorov-Arnold Networks (KANs). Our approach is built on three components: (1) *Biomechanical Phase Space Reconstruction:* instead of relying solely on raw waveform matching, we embed the signal's first- and second-order derivatives to capture kinematic structures that are more stable under temporal warping; (2) *High-Capacity Matrix Memory:* we leverage xLSTM's matrix-based state to preserve high-order pairwise correlations over long temporal contexts, helping retain fine-grained biometric characteristics; and (3) *Vascular Compliance Modeling:* we replace fixed linear layers with learnable B-splines (KANs), which provide a flexible way to model the non-linear stress-strain behavior of arterial walls. Extensive experiments on four public benchmarks (CAPNOBASE, BIDMC, PPG-DaLiA, and PTTP-PPG) demonstrate that Physio-xLSTM achieves state-of-the-art performance and consistently outperforms recent Mamba and Transformer baselines in motion-corrupted scenarios.**

***Index Terms*—Biometrics, Photoplethysmography (PPG), xLSTM, Matrix Memory, Kolmogorov-Arnold Networks (KAN), Hemodynamics.**

I. Introduction

PHOTOPLETHYSMOGRAPHY (PPG) has become an attractive biometric modality for wearable devices because it is non-invasive, easy to acquire, and naturally tied to physiological activity [1]. In contrast to static traits such as fingerprints, PPG signals are generated by dynamic interactions among cardiac output, vascular elasticity, and peripheral blood flow. These properties make PPG promising for continuous authentication, but they also make the signal considerably harder to model.

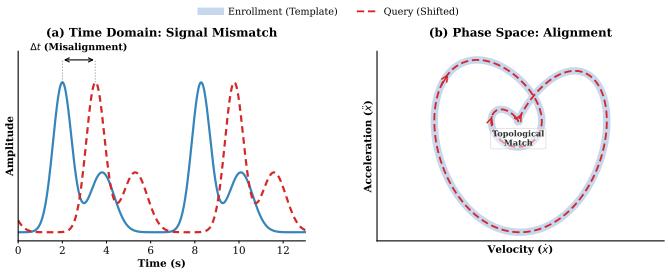

Fig. 1. **Conceptual Illustration of Kinematic Invariance.** (a) In the time domain, physiological instability causes noticeable phase mismatch between enrollment and query signals. (b) In the biomechanical phase space (Velocity vs. Acceleration), the overall topological structure remains more stable under temporal warping. This observation motivates Physio-xLSTM to model derivative-domain dynamics rather than relying only on direct waveform alignment.

A central difficulty is that PPG identity patterns are not expressed as perfectly repeatable waveforms. Instead, they are embedded in pulse dynamics that vary with motion, hemodynamic state, and acquisition conditions. As illustrated in Fig. 1(a), even signals from the same subject may exhibit substantial temporal displacement and local morphology changes. Consequently, models that rely mainly on direct shape matching in the time domain may struggle to preserve identity cues under realistic perturbations.

From this perspective, two limitations of current sequence backbones are especially relevant.

1. Information loss under vector-state summarization:

Many efficient temporal architectures, including Recurrent Neural Networks (RNNs) and recent State Space Models (SSMs) such as Mamba [2], summarize history through a fixed-dimensional vector state ($h_t \in \mathbb{R}^d$). Although effective for compact sequence modeling, this representation may dilute subtle but discriminative beat-level morphology, such as fine variations around the dicrotic notch. Transformer models [3] alleviate this issue through global attention, but their quadratic complexity ($O(L^2)$) remains undesirable for long or continuous physiological monitoring.

2. Limited flexibility of standard pointwise nonlinearities: Most modern backbones still rely on MLP blocks with fixed nonlinear activations, e.g., ReLU or GELU. However, vascular and hemodynamic responses evolve along smooth and highly nonlinear trajectories. Such behavior is not always captured efficiently by piecewise linear transformations, especially when the goal is to preserve small physiological variations that carry identity information.

To address these issues, we propose *Physio-xLSTM*, a physics-aware framework that models PPG biometrics from the viewpoint of *biomechanical dynamics*. Rather than treating the waveform as a static shape to be matched, our method reconstructs derivative-domain descriptors and combines them with a high-capacity recurrent memory and spline-based nonlinear approximation. The resulting framework is built around three key components:

- **Kinematic derivative embedding:** We augment the raw signal with its first- and second-order derivatives, enabling the model to encode identity information in a phase-space-inspired representation that is less sensitive to baseline drift and local temporal misalignment.
- **xLSTM with matrix memory:** We employ the Extended LSTM (xLSTM) [4], whose matrix-valued memory state provides richer pairwise storage than conventional vector-state sequence models. This helps retain fine-grained pulse morphology over longer temporal contexts.
- **KAN-based physiological approximation:** We integrate Kolmogorov-Arnold Networks (KANs) [5] into the feature extractor so that spline-based nonlinearities can better fit the smooth transformations associated with vascular compliance and pulse propagation.

Experiments on four public datasets, namely CAPNOBASE, BIDMC, PPG-DaLiA, and PTTP-PPG, show that Physio-xLSTM consistently achieves strong recognition performance, with the clearest advantages appearing under motion-contaminated conditions. These results indicate that combining derivative-domain priors with matrix-memory temporal modeling offers an effective route toward more reliable PPG biometrics.

II. Related Work

A. PPG Biometrics with Deep Models

Early PPG biometric systems were largely based on handcrafted descriptors and fiducial-point analysis [1], which were often sensitive to noise, imperfect peak detection, and acquisition variability. Deep learning methods later improved representation learning by replacing manual feature engineering with data-driven encoders. In particular, CNN-based solutions [6] became popular for extracting local pulse morphology. However, their finite receptive fields can limit the modeling of longer-range rhythm dependencies.

Transformer-based methods [3], [7] improve global context modeling through self-attention, but they introduce substantial computational overhead and may be sensitive to temporal jitter in ambulatory recordings. Hybrid models that combine multiple feature domains or backbone types [8] have also shown promise, yet many of them still operate primarily on raw waveform representations. In contrast, our work emphasizes derivative-domain modeling, where identity cues are described through the velocity- and acceleration-like evolution of the pulse rather than through direct waveform alignment alone.

B. Temporal Memory for Physiological Sequences

Sequential physiological analysis has long relied on RNNs and LSTMs [9], which remain useful for ordered temporal modeling but are often constrained by sequential computation and limited memory capacity. More recently, State Space Models (SSMs), particularly Mamba [2], have attracted attention because they offer linear-time sequence processing and favorable hardware efficiency. Their applicability has also been demonstrated in biomedical settings [10].

Nevertheless, many efficient temporal models still depend on compressed vector-state summaries. For physiological biometrics, this can be problematic because identity-relevant structure may reside in subtle local morphology that should not be overly smoothed during long-range aggregation. xLSTM [4] provides an appealing alternative by replacing conventional vector-style memory with a matrix-based mechanism. This richer state representation is well suited to preserving pairwise interactions among features, which is particularly relevant when small beat-shape variations carry discriminative information.

C. KANs for Smooth Physiological Modeling

Standard MLPs approximate nonlinear functions through fixed activation rules such as ReLU or GELU. While broadly effective, these blocks may require greater depth or width to fit smooth biological transformations with high fidelity. Kolmogorov-Arnold Networks (KANs) [5] offer a different design by placing learnable spline-based functions on edges instead of relying only on fixed node activations.

For hemodynamic signals, this idea is especially relevant because vascular compliance and pulse propagation are governed by smooth nonlinear processes. The B-spline parameterization in KANs therefore provides a flexible tool for representing physiological response curves. Motivated by this property, our framework incorporates KAN components into the front-end feature extractor so that nonlinear pulse transformations can be modeled more adaptively than in standard MLP-based pipelines.

III. METHODOLOGY

A. Problem Formulation and Overview

Let $\mathcal{X} = \{x_1, x_2, \ldots, x_T\}$ denote a raw PPG signal segment of length T, where $x_t \in \mathbb{R}$ represents the normalized amplitude at time step t. The objective of biometric identification is to learn a mapping function $\mathcal{F}_\theta : \mathbb{R}^T \to \mathbb{R}^{d_{id}}$ parameterized by θ. This function projects a noisy and non-stationary input signal into a compact identity-discriminative embedding space that should remain robust under motion artifacts and physiological state variations.

As illustrated in Fig. 2, the proposed Physio-xLSTM framework processes the signal through a parallel dual-stream architecture. The *Time Stream* captures fine-grained biomechanical dynamics using a kinematic-enhanced input and a KAN-xLSTM backbone. In parallel, the *Frequency Stream* extracts shift-invariant spectral patterns to dynamically gate temporal features. Finally, a *Disentanglement Module* separates the latent space into orthogonal identity ($\mathbf{z}_{id}$) and noise ($\mathbf{z}_{noise}$) components to improve robustness during verification.

B. Kinematic Decomposition and Augmentation (Module A)

Standard deep learning models typically treat PPG signals as static position sequences. However, the identity-specific signatures of the cardiovascular system are intrinsically linked to fluid dynamics. To better capture these dynamics, we propose a *Kinematic Decomposition* strategy that reconstructs a phase-space view of the signal.

As shown in the "Biomechanical Derivative" block of Fig. 2, we augment the raw input sequence x_t with its first-order discrete derivative Δx_t and second-order discrete derivative $\Delta^2 x_t$. In the figure, these channels are visualized as $f_{Original}$, $\Delta f_{Original}$, and $\Delta f'_{Original}$, respectively. In our discrete formulation,

$$\Delta x_t = x_t - x_{t-1}, \tag{1}$$

$$\Delta^2 x_t = \Delta x_t - \Delta x_{t-1}. \tag{2}$$

We refer to Δx_t and $\Delta^2 x_t$ as velocity-like and acceleration-like descriptors, respectively. These three components are concatenated to form the augmented kinematic tensor $\mathbf{X}_{kin} \in \mathbb{R}^{3\times T}$.

Physiologically, these derivatives capture complementary information: Δx_t correlates with instantaneous pulse-wave variation, while $\Delta^2 x_t$ reflects finer changes related to cardiac contractility and vascular compliance. By explicitly modeling these higher-order dynamics, the network is encouraged to learn from the generative behavior of the pulse wave instead of relying only on waveform shape matching, which is more sensitive to baseline drift.

C. Physiologically-Informed KAN Backbone

1) KAN-based Multi-Scale Stem (Module B): To extract low-level morphological features from $\mathbf{X}_{kin}$, we design a KAN-MultiScale Stem. Unlike standard Convolutional Neural Networks (CNNs) that use linear kernels followed by piecewise linear activations (e.g., ReLU), we employ *KAN-Convolutions*. For a kernel size k, the operation is defined as a linear combination of B-spline basis functions. The output y for an input x is given by:

$$y = \sum_{j=1}^{k} \Phi_j(x_j), \quad \Phi_j(u) = \sum_{i=1}^{n} w_i \cdot B_{i,p}(u), \tag{3}$$

where w_i are learnable coefficients. Here, $B_{i,p}(u)$ denotes the i-th B-spline basis function of degree p, defined recursively via the Cox-de Boor formula:

$$B_{i,p}(u) = \frac{u - t_i}{t_{i+p} - t_i} B_{i,p-1}(u) + \frac{t_{i+p+1} - u}{t_{i+p+1} - t_{i+1}} B_{i+1,p-1}(u), \tag{4}$$

with base case $B_{i,0}(u) = 1$ if $t_i \le u < t_{i+1}$ and 0 otherwise. This spline formulation allows KANs to approximate smooth physiological non-linearities more flexibly than fixed activations. We utilize four parallel branches with kernel sizes $k \in \{1, 3, 7, 11\}$ (denoted as Conv_1 to Conv_{11} in Fig. 2) to capture diverse temporal granularities, from sharp systolic upstrokes to gradual diastolic decays.

2) Bidirectional Matrix-Memory Encoding (Module C): The core temporal modeling is handled by the **BixLSTM** module. To better preserve transient details than standard vector-state sequence models, we leverage the *Matrix Memory* mechanism of xLSTM.

Let $\mathbf{h}_t \in \mathbb{R}^d$ be the input at step t. The xLSTM updates its memory matrix $\mathbf{C}_t \in \mathbb{R}^{d\times d}$ through gated accumulation. Specifically, the input gate $\mathbf{i}_t$ and forget gate $\mathbf{f}_t$ are computed using KAN layers:

$$\mathbf{i}_t = \sigma(\text{KANLinear}(\mathbf{h}_t)), \quad \mathbf{f}_t = \sigma(\text{KANLinear}(\mathbf{h}_t)). \tag{5}$$

The memory update is then written as

$$\mathbf{C}_t = \mathbf{f}_t \odot \mathbf{C}_{t-1} + \mathbf{i}_t \otimes (\mathbf{u}_t \mathbf{k}_t^\top), \tag{6}$$

where $\odot$ denotes element-wise multiplication, $\otimes$ denotes the gated outer-product update, and $\mathbf{u}_t$ and $\mathbf{k}_t$ are the learned value and key vectors, respectively. In this way, $\mathbf{C}_t$ stores higher-order pairwise interactions between features, providing richer memory than a compressed vector state.

Given the non-causal nature of offline biometric identification, we use a bidirectional encoding strategy:

$$\mathbf{H}_{out} = [\text{xLSTM}_{\rightarrow}(\mathbf{H}_{in}) \oplus \text{xLSTM}_{\leftarrow}(\mathbf{H}_{in})]. \tag{7}$$

The sequence is then aggregated into a global vector $\mathbf{z}_{time}$ via Global Average Pooling (GAP).

D. Spectral-Magnitude Gating (Frequency Stream)

Motion artifacts often induce phase shifts that distort time-domain features. To mitigate this effect, we introduce a parallel Frequency Stream based on the *Magnitude Spectrum*.

We first apply Fast Fourier Transform (FFT) to the raw signal and compute the magnitude spectrum $\mathbf{S}_{freq} = |\text{FFT}(\mathbf{x})|$. As shown in the purple block of Fig. 2, this spectral input

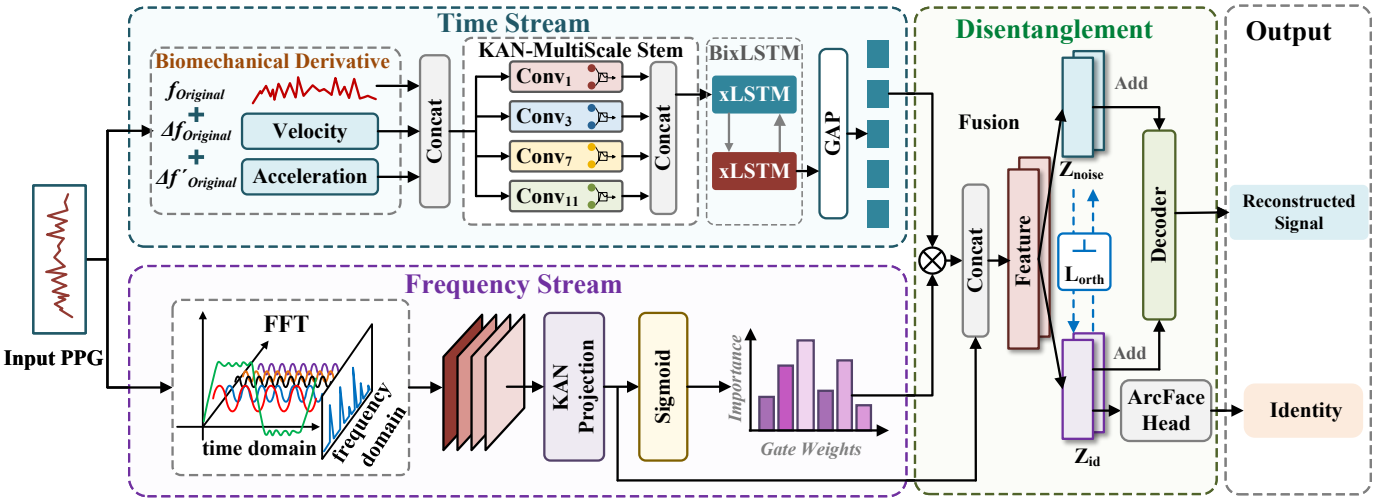

Fig. 2. **Overview of the proposed Physio-xLSTM framework.** The architecture consists of two parallel streams: (1) The **Time Stream** (top) incorporates the raw waveform and its first- and second-order discrete derivatives, visualized in the figure as $(f, \Delta f, \Delta f')$ and defined in Eqs. (1)–(2), to capture kinematic dynamics, processed by a KAN-MultiScale Stem and a Bidirectional xLSTM (BixLSTM) backbone. (2) The **Frequency Stream** (bottom) utilizes FFT and KAN Projection to generate global gate weights based on the magnitude spectrum. The features are fused and passed to a **Disentanglement Module**, where the representation is split into Identity ($\mathbf{z}_{id}$) and Noise ($\mathbf{z}_{noise}$) components, supervised by Orthogonality Loss (L_{orth}), ArcFace Loss, and Reconstruction Loss (Decoder).

is projected into a latent feature vector $\mathbf{Z}_{freq}$ via a KAN projection layer:

$$\mathbf{Z}_{freq} = \text{KANProjection}(\mathbf{S}_{freq}). \tag{8}$$

This vector captures global rhythmic patterns that are comparatively stable under time shifts. We then generate channel-wise gate weights $\mathbf{g} \in (0, 1)$ through a Sigmoid activation:

$$\mathbf{g} = \sigma(\mathbf{W}_g \mathbf{Z}_{freq} + \mathbf{b}_g). \tag{9}$$

The fused representation $\mathbf{z}_{fused}$ is obtained by modulating the time-domain features and concatenating the spectral context:

$$\mathbf{z}_{fused} = (\mathbf{z}_{time} \odot \mathbf{g}) \oplus \mathbf{Z}_{freq}. \tag{10}$$

Here, $\odot$ represents the Hadamard product. This mechanism allows stable spectral priors to re-weight temporal features dynamically and suppress segments contaminated by transient noise.

E. Disentanglement and Optimization (Module D)

To encourage a cleaner biometric representation, we employ a **Disentanglement Module**. The fused feature is projected into two latent spaces: an Identity Embedding ($\mathbf{z}_{id}$) and a Noise Embedding ($\mathbf{z}_{noise}$).

The network is trained with a multi-task objective:

$$\mathcal{L}_{total} = \mathcal{L}_{id} + \lambda_{rec}\mathcal{L}_{rec} + \lambda_{orth}\mathcal{L}_{orth}. \tag{11}$$

a) Identity Loss ($\mathcal{L}_{id}$): We apply an ArcFace head to $\mathbf{z}_{id}$ to improve inter-class separability.

b) Reconstruction Loss ($\mathcal{L}_{rec}$): To encourage $\mathbf{z}_{noise}$ to capture non-identity variations, the decoder reconstructs the original waveform from the sum of the two embeddings:

$$\hat{\mathbf{x}} = \text{Decoder}(\mathbf{z}_{id} + \mathbf{z}_{noise}), \quad \mathcal{L}_{rec} = \|\hat{\mathbf{x}} - \mathbf{x}\|_2^2. \tag{12}$$

This corresponds to the logical flow Fusion $\rightarrow$ Add $\rightarrow$ Decoder shown in Fig. 2.

c) Orthogonality Loss ($\mathcal{L}_{orth}$): To explicitly encourage disentanglement, we minimize the absolute cosine similarity between the identity and noise vectors:

$$\mathcal{L}_{orth} = \frac{1}{N}\sum_{j=1}^{N} \frac{|\mathbf{z}_{id}^{(j)} \cdot \mathbf{z}_{noise}^{(j)}|}{\|\mathbf{z}_{id}^{(j)}\|_2 \|\mathbf{z}_{noise}^{(j)}\|_2}, \tag{13}$$

where N is the batch size. This term encourages $\mathbf{z}_{id}$ and $\mathbf{z}_{noise}$ to occupy approximately orthogonal subspaces, thereby reducing the contamination of identity features by transient noise.

IV. Experiments

A. Experimental Setup

a) *Datasets*: We evaluate Physio-xLSTM on four public PPG benchmarks collected under different acquisition conditions. CAPNOBASE [11] and BIDMC [12] are treated as relatively clean clinical datasets, whereas PPG-DaLiA [13] and PTTP-PPG [14] represent more challenging ambulatory settings with stronger motion corruption and physiological variability.

b) *Physics-Preserving Preprocessing*: Because derivative cues are central to the proposed framework, preprocessing is kept conservative. Raw PPG signals are bandpass filtered within 0.5–8 Hz to suppress obvious interference while preserving subtle pulse morphology, especially around the

dicrotic-notch region. The filtered recordings are segmented into 256-point pulse representations and rescaled to $[0, 1]$, with the pipeline kept shape-preserving so that derivative-domain dynamics remain interpretable for kinematic modeling.

c) *Evaluation Setting and Metrics*: We evaluate the model in a within-dataset identity recognition setting, where each subject's earlier recordings are used to capture subject-specific pulse dynamics and later recordings are reserved for hold-out evaluation (80%/20% chronological split). The same split is applied to the proposed model and all internal baselines to keep the comparison protocol unchanged across backbones. We primarily report Top-1 identification accuracy, and we additionally summarize verification behavior with Equal Error Rate (EER). For the literature-level comparison in Table II, we report the corresponding published results under heartbeat-aggregation settings when such numbers are available, so that the comparison remains consistent with prior PPG biometric studies.

B. Comparison with Baselines

We compare Physio-xLSTM with representative backbones from several model families, including CNNs (ConvNeXt V2), recurrent models (BiLSTM), Transformer variants (PatchTST), Mamba, and xLSTM-based baselines. As reported in Table I, the proposed model yields the best overall results across all four benchmarks. On cleaner datasets such as BIDMC, performance differences among strong models become relatively small, suggesting that these settings are less informative for exposing robustness gaps.

The distinction becomes clearer on the ambulatory benchmark PPG-DaLiA, where local morphology is more easily perturbed by motion and temporal inconsistency. Under this condition, Physio-xLSTM improves over standard Mamba from 83.27% to **87.50%**. This result is consistent with our design motivation: derivative-domain descriptors and matrix-memory temporal aggregation provide a more stable representation when raw waveform alignment is unreliable.

In terms of model size, Physio-xLSTM remains moderate with 3.36M parameters and 189 MFLOPs. We note, however, that practical deployment cost is not determined by parameter count alone, since spline evaluation in KAN layers and matrix-state updates also contribute to runtime. Therefore, the present model is best viewed as a robustness-oriented architecture rather than a strictly hardware-optimized one.

C. Comparison with Recent Methods

Table II compares Physio-xLSTM with recent PPG biometric approaches under the corresponding heartbeat-aggregation setting. The proposed model achieves the strongest reported result on PTTP-PPG (**96.13%**) and also improves over prior methods on PPG-DaLiA. Rather than relying solely on increasingly complex waveform transformations, these gains support the utility of modeling subject identity through derivative-domain pulse dynamics and higher-capacity temporal retention.

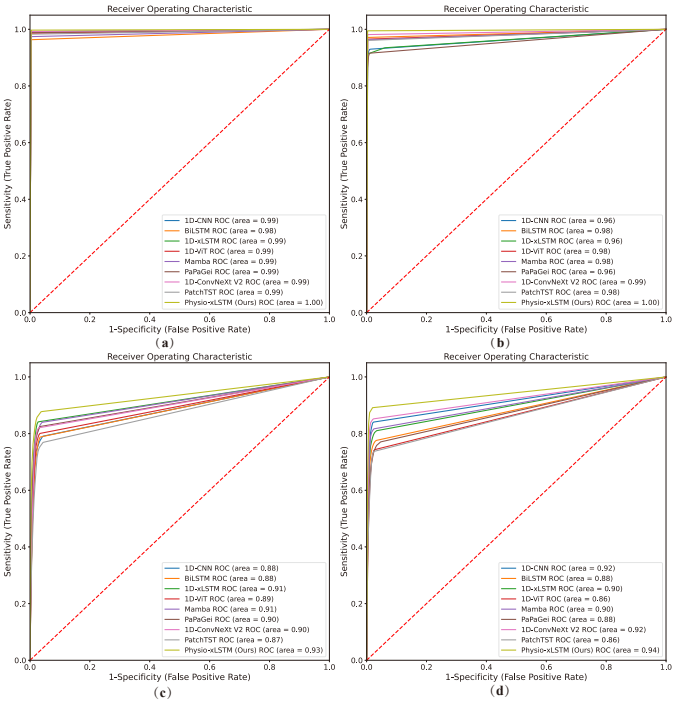

Fig. 3. ROC curves of the compared methods. Physio-xLSTM maintains the clearest separation, especially in the low-false-positive region.

D. Embedding Quality Analysis

The ROC curves in Fig. 3 provide a complementary view of the learned embedding space. On the two cleaner datasets, the AUC reaches 1.00, indicating nearly perfect separation among identities. On the noisier datasets, the AUC remains above 0.93, showing that the proposed dual-stream representation continues to preserve useful decision boundaries even when motion artifacts are present.

E. Perturbation Studies

a) *Hilbert-Based Phase Distortion*: We introduce Hilbert-based phase perturbation to emulate temporal deformation caused by physiological irregularity or unstable sensing. As shown in Fig. 4, Physio-xLSTM remains more resilient than the compared baselines. This suggests that the derivative-domain representation stays informative even when direct waveform alignment is perturbed.

b) *Cross-Domain Spectral Interference*: We further evaluate the model under a cross-dataset interference setting that disturbs the surrounding spectral statistics. The results in Fig. 5 show that Physio-xLSTM degrades more gracefully than the compared baselines under this mismatch, indicating that matrix-memory encoding and disentanglement help preserve subject-related information under distribution shifts.

F. Ablation Study

Table III quantifies the contribution of each module. Starting from a vanilla xLSTM with MLP projections, the kinematic branch provides the most notable improvement on the motion-heavy DaLiA benchmark, yielding a +2.52% gain over the

TABLE I

PERFORMANCE BENCHMARK ON FOUR DATASETS. PHYSIO-XLSTM ACHIEVES THE STRONGEST OVERALL RESULTS, WITH THE LARGEST GAINS APPEARING ON MOTION-HEAVY BENCHMARKS.

Method	CapnoBase		BIDMC		PPG-DaLiA		PTTP-PPG	
	Acc(%)	EER(%)	Acc(%)	EER(%)	Acc(%)	EER(%)	Acc(%)	EER(%)
1D-CNN	98.89	0.95	93.77	3.42	80.60	10.50	77.23	12.80
BiLSTM	98.40	1.25	97.45	1.55	81.46	10.10	77.21	13.10
1D-xLSTM	99.26	0.58	93.29	3.50	83.93	7.80	80.61	10.20
PaPaGei	98.83	0.92	93.24	3.65	82.26	9.20	76.26	13.50
1D-ViT	98.62	1.10	96.72	1.90	82.74	8.90	74.01	14.80
PatchTST	98.28	1.35	96.27	2.10	76.31	14.50	73.48	15.60
ConvNeXt V2	99.23	0.60	98.36	1.15	81.79	9.80	85.04	8.50
Mamba	99.05	0.72	96.75	1.85	83.27	8.15	81.27	10.50
Physio-xLSTM (Ours)	**99.71**	**0.20**	**99.60**	**0.35**	**87.50**	**5.10**	**89.02**	**4.85**

TABLE II

COMPARISON WITH RECENT PPG BIOMETRIC METHODS UNDER THE SAME HEARTBEAT-AGGREGATION SETTING (ACCURACY %).

Dataset	Method	Acc (%)
CAPNOBASE	PulseID [15]	97.80
	Liu et al. [16]	99.82
	Wang et al. [17]	99.50
	Zhang et al. [18]	97.62
	Xu et al. [19]	98.87
	Physio-xLSTM (Ours)	**99.94**
BIDMC	Liu et al. [8]	99.16
	Wang et al. [17]	99.30
	Ngo et al. [20]	98.80
	Liu et al. [16]	98.78
	Xu et al. [19]	98.41
	Zhao et al. [21]	96.17
	Physio-xLSTM (Ours)	**99.86**
PPG-DaLiA	Xu et al. [19]	92.67
	Ghorbani et al. [22]	87.00
	Wan et al. [23]	86.32
	Ye et al. [24]	93.74
	Physio-xLSTM (Ours)	**95.35**
PTTP-PPG	Wan et al. [23]	85.94
	Ibrahim et al. [25]	92.78
	Ye et al. [24]	88.45
	Ghorbani et al. [22]	93.00
	Physio-xLSTM (Ours)	**96.13**

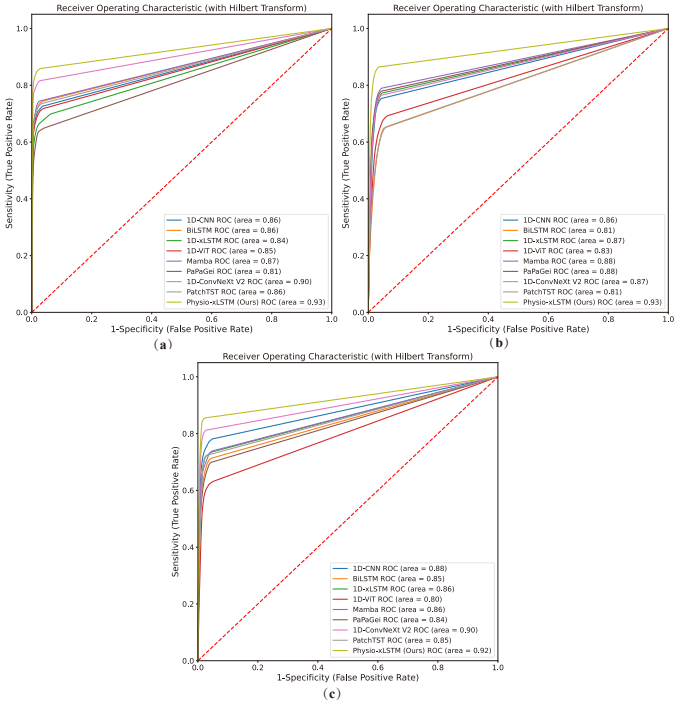

Fig. 4. Results under Hilbert-based phase distortion. The derivative-domain representation remains reliable when waveform alignment is perturbed.

baseline. This suggests that explicit first- and second-order descriptors help separate physiologically meaningful pulse evolution from motion-induced distortion. Adding KAN components further improves the clean-data results, supporting the view that spline-based nonlinearities are beneficial for smooth physiological transformations.

G. Data Efficiency

Fig. 6 shows that Physio-xLSTM remains competitive even when only 30% of the training data is used, indicating favorable data efficiency under limited-data enrollment.

V. CONCLUSION

This paper presented Physio-xLSTM, a PPG biometric framework based on derivative-domain identity modeling, matrix-memory temporal encoding, and KAN-based nonlinear approximation. By describing subjects through first- and second-order pulse dynamics rather than direct waveform similarity alone, the proposed method improves robustness to temporal deformation and motion corruption. Across four public datasets, Physio-xLSTM achieves consistently strong performance, including 95.35% on PPG-DaLiA under the heartbeat-aggregation setting. These results suggest that derivative-domain temporal modeling is a promising direction for robust wearable biometrics. Future work will further examine cross-

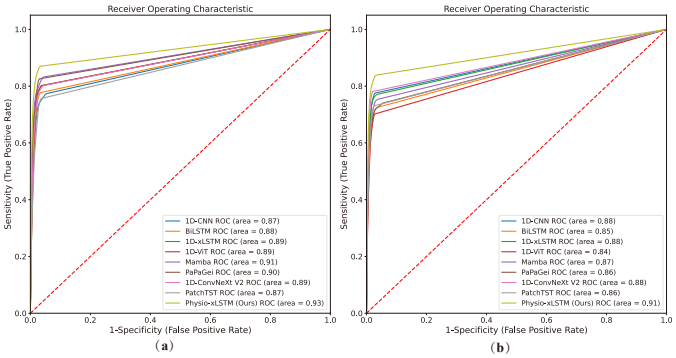

Fig. 5. Results under cross-domain spectral interference. Physio-xLSTM is less affected by cross-dataset spectral mismatch.

TABLE III
ABLATION STUDY (ACCURACY %).

Model Variant	Capno	BIDMC	DaLiA	PTTP
Baseline (1D-xLSTM)	99.26	93.29	83.93	80.61
Baseline + Kinematics	99.40	98.60	86.45	87.20
Baseline + KAN	99.35	97.80	84.80	86.10
Baseline + Disent.	99.20	97.40	84.10	84.90
Physio-xLSTM (w/o KAN)	99.58	99.10	86.85	88.35
Physio-xLSTM (w/o Kinematics)	99.45	98.90	85.20	87.50
Physio-xLSTM (Full)	**99.71**	**99.60**	**87.50**	**89.02**

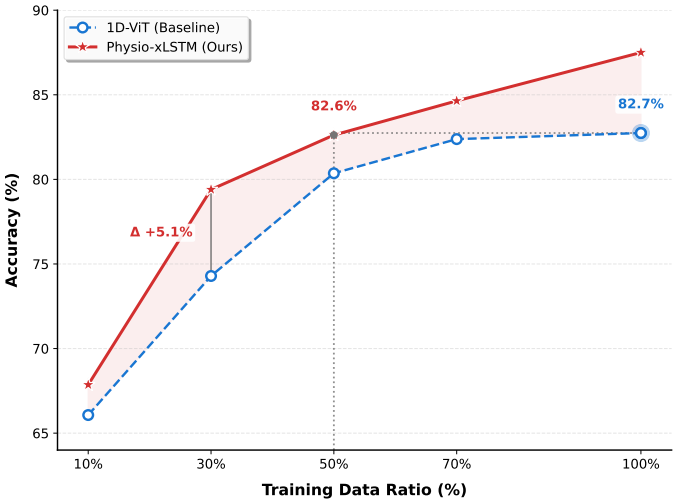

Fig. 6. Data efficiency on PPG-DaLiA. Physio-xLSTM maintains robust performance with reduced training data.

session generalization, simplified deployment, and broader sources of physiological variation.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support of project ZR2022LZH010 supported by Shandong Provincial Natural Science Foundation, project No. 2024CXGC010113 supported by Key R&D Program of Shandong Province, China, the National Natural Science Foundation of China under Grant 62276093, and the Natural Science Foundation of Shandong Province under Grant ZR2022MF286.

REFERENCES

[1] R. D. Labati *et al.*, "Photoplethysmographic biometrics: A comprehensive survey," *Pattern Recognition Letters*, 2022.
[2] A. Gu and T. Dao, "Mamba: Linear-time sequence modeling with selective state spaces," in *Conference on Language Modeling (COLM)*, 2024.
[3] A. Vaswani *et al.*, "Attention is all you need," in *NIPS*, 2017.
[4] M. Beck, K. Pöppel, M. Spanring, A. Auer, O. Prudnikova, M. Kopp, G. Klambauer, J. Brandstetter, and S. Hochreiter, "xlstm: Extended long short-term memory," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.
[5] Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljačić, T. Y. Hou, and M. Tegmark, "Kan: Kolmogorov-arnold networks," in *International Conference on Learning Representations (ICLR)*, 2025.
[6] A. M. Alqudah and Z. Moussavi, "A review of deep learning for biomedical signals: Current applications, advancements, future prospects, interpretation, and challenges," *Computers, Materials & Continua*, vol. 82, no. 1, 2025.
[7] Q. Wen, T. Zhou, C. Zhang, W. Chen, Z. Ma, J. Yan, and L. Sun, "Transformers in time series: A survey," in *International Joint Conference on Artificial Intelligence (IJCAI)*, 2023, pp. 6778–6786.
[8] C. Liu *et al.*, "Dual-domain and multiscale fusion deep neural network for ppg biometric recognition," *Machine Intelligence Research*, vol. 20, no. 5, pp. 707–715, 2023.
[9] F. A. Gers *et al.*, "Learning to forget: Continual prediction with lstm," *Neural computation*, 2000.
[10] J. Ma, F. Li, and B. Wang, "U-mamba: Enhancing long-range dependency for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2024.
[11] W. Karlen, S. Raman, J. M. Ansermino, and G. A. Dumont, "Multiparameter respiratory rate estimation from the photoplethysmogram," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 7, pp. 1946–1953, 2013.
[12] M. A. Pimentel *et al.*, "Toward a robust estimation of respiratory rate from pulse oximeters," *IEEE TBME*, vol. 64, no. 8, pp. 1914–1923, 2016.
[13] A. Reiss *et al.*, "Deep ppg: Large-scale heart rate estimation with convolutional neural networks," *Sensors*, vol. 19, no. 14, p. 3079, 2019.
[14] P. Mehrgardt *et al.*, "Pulse transit time ppg dataset," *PhysioNet*, 2022.
[15] R. Wei *et al.*, "Pulseid: Multi-scale photoplethysmographic identification using a deep convolutional neural network," *Biomedical Signal Processing and Control*, vol. 88, p. 105609, 2024.
[16] C. Liu *et al.*, "Time–frequency fusion learning for photoplethysmography biometric recognition," *IET Biometrics*, vol. 11, no. 3, pp. 187–198, 2022.
[17] D. Wang *et al.*, "Biometric recognition based on scalable end-to-end convolutional neural network using photoplethysmography," *Computers in Biology and Medicine*, vol. 147, p. 105654, 2022.
[18] Y. Zhang *et al.*, "A ppg-based biometric system using fuzzy min-max model," in *ICCVDM 2022*, 2023.
[19] Y. Xu, Z. Wang, and X. Liu, "Photoplethysmogram (ppg)-based biometric identification using 2d signal transformation and multi-scale feature fusion," *Sensors*, vol. 25, no. 15, p. 4849, 2025.
[20] D. Ngo *et al.*, "Precision biometrics based on the ppg measured from an iot device," *IEEE Internet of Things Journal*, 2024.
[21] F. Zhao *et al.*, "A lightweight convolutional neural network method for two-dimensional photoplethysmography signals," *Applied Sciences*, vol. 14, no. 10, 2024.
[22] R. Ghorbani *et al.*, "Personalized anomaly detection in ppg data using representation learning and biometric identification," *Biomedical Signal Processing and Control*, vol. 94, p. 106216, 2024.
[23] L. Wan *et al.*, "Deep learning-based photoplethysmography biometric authentication for continuous user verification," *Applied Soft Computing*, vol. 156, p. 111461, 2024.
[24] Y. Ye *et al.*, "Ppg-based biometric identification: Discovering and identifying a new user," in *IEEE EMBC*, 2021.
[25] M. Ibrahim *et al.*, "A novel ppg-based biometric authentication system using a hybrid cvt-convmixer architecture," *Sensors*, vol. 24, no. 1, 2023.