

PathFusion-Net: A Rough Path Theory-Based Deep Learning Model for ECG Arrhythmia Classification

Tianlong Feng, Qingchen Li, Yuanyuan Zhang, Yongzhi Liao, Di Lu, Liping Wang, Jianqin Zhao, Lei Jiang, Hao Ni, Hongying Liu and Jingjing Deng

Abstract—This study introduces a novel electrocardiogram (ECG) arrhythmia classification model, PathFusion-Net, which integrates Rough Path Theory with deep learning technologies. The model combines Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Path Signatures, and Path Development to extract spatial morphological features from ECG images and multi-order temporal representations from ECG signals. By adopting an inter-patient split paradigm, our approach more closely reflects real-world clinical diagnostic settings compared to intra-patient methods. The model demonstrates state-of-the-art overall classification performance on both the MIT-BIH Arrhythmia Database and a private clinical dataset, achieving 94.7% and 95.1% accuracy, respectively, under the AAMI four-class standard with an inter-patient split paradigm. On the MIT-BIH dataset, the proposed method attains competitive precision and recall across multiple arrhythmia types, including 95.2%/87.9% for ventricular ectopic beats (V) and 75.7%/92.3% for supraventricular ectopic beats (S), indicating balanced performance across clinically diverse categories. This research highlights the potential of Rough Path Theory in time-series analysis and offers a novel deep learning framework for automated early detection and monitoring of ECG arrhythmias.

Index Terms—Rough Path Theory, ECG Arrhythmia,

The Ethics Committee of Honghui Hospital approved this study (approval number:2025-Ethical Opinions-KY-060-01).(Corresponding author: Yuanyuan Zhang).

Hao Ni and Lei Jiang are supported by the EPSRC [grant number EP/S026347/1]. Hao Ni is also supported by The Alan Turing Institute under the EPSRC [grant number EP/N510129/1].

Tianlong Feng, Yongzhi Liao and Di Lu are with the School of Computer Science and Technology, Xidian University, Shaanxi, China (e-mail: tlfeng@stu.xidian.edu.cn; liaoyongzhi1010@stu.xidian.edu.cn; dlu@xidian.edu.cn).

Qingchen Li is with the Department of Cardiology, Honghui Hospital, Xi'an Jiaotong University, Shaanxi, China (e-mail: doc-torli0921@163.com).

Yuanyuan Zhang and Liping Wang are with the Department of Geriatrics, Honghui Hospital, Xi'an Jiaotong University, Shaanxi, China (e-mail: zhangyuanyuan.8811@126.com; docwangliping@163.com).

Jianqin Zhao is with the Department of Computer Science, Durham University, Durham, UK (e-mail: jingjing.deng@durham.ac.uk; jianqin.zhao@durham.ac.uk).

Lei Jiang and Hao Ni are with the Department of Mathematics, University College London, London, UK (e-mail: lei.j@ucl.ac.uk; h.ni@ucl.ac.uk).

Hongying Liu is with the Medical School, Tianjin University, Tianjin, China (e-mail: hylu2009@tju.edu.cn).

Jingjing Deng is with the School of Engineering Mathematics and Technology, University of Bristol, Bristol, UK (e-mail: jingjing.deng@bristol.ac.uk).

Deep Learning.

I. INTRODUCTION

Non-communicable diseases (NCDs) contribute to 41 million deaths annually, accounting for 74% of all global fatalities. Among these, cardiovascular diseases (CVDs) are the leading cause, responsible for approximately 17.9 million deaths per year [1]. This significant mortality burden underscores the critical need for accurate, timely diagnosis and management of cardiovascular conditions. One particularly challenging CVD subset is arrhythmias—a group of disorders including tachycardia, bradycardia, premature beats, and atrial fibrillation—each impacting cardiac function with varying degrees of severity and often linked to heightened risks of heart attacks and strokes. Electrocardiograms (ECGs) record the electrical activity of the heart over a period of time and are vital tools for assessing heart health [2]. With the rapid development of machine learning technologies, analyzing large amounts of ECG data can simulate the diagnostic reasoning of cardiovascular experts, enabling fast and accurate diagnoses. This not only significantly reduces the workload of medical professionals but also ensures timely and precise diagnosis of patients' conditions, thus holding great value in both practical application and research. Many previous studies adopted the intra-patient paradigm in dataset partitioning and achieved accuracies exceeding 99%. However, this approach may cause information leakage, thereby undermining the credibility of the results. As demonstrated by Cao et al. (2022) [3], the inter-patient paradigm mitigates this issue and more closely reflects real-world deployment scenarios. Therefore, this study employs the inter-patient split to avoid bias from data leakage.

In recent years, significant progress has been made in ECG-based arrhythmia classification using machine learning. Initially, traditional techniques such as Support Vector Machines (SVM) [4] and K-Nearest Neighbors (K-NN) [5] were widely used. However, with the rapid advancement of deep learning, researchers have increasingly introduced deep learning methods into automatic ECG classification. Convolutional Neural Network (CNN), as one of the prominent deep learning models, have demonstrated exceptional performance in signal analysis, image recognition, pixel data processing, and natural language processing. In recent years, CNN have also been

extensively applied to ECG classification and arrhythmia detection [6]. When applied to ECG signal analysis, CNN can automatically learn and extract relevant features from raw ECG signals, thereby enhancing the accuracy of arrhythmia detection. In addition to CNN, Long Short-Term Memory (LSTM) networks have also gained widespread attention. LSTM, a variant of Recurrent Neural Network (RNN), is equipped with unique memory cells and gating mechanisms that allow it to learn and retain sequences over extended periods. This feature is particularly crucial when processing ECG data with long-term dependencies, effectively overcoming the vanishing gradient problem common in standard RNN, making LSTM particularly well-suited for tasks involving time-series data, such as ECG signal analysis [7].

Despite the progress made with both traditional machine learning and deep learning methods, the dynamic feature extraction from time-series data remains a challenge. To address this, Rough Path Theory, proposed in the 1990s, offers a novel perspective for understanding the response of nonlinear systems to highly oscillatory input signals [8], [9]. A key component of this theory is the signature transform, which provides an efficient way to represent and extract features from high-dimensional ordered data. Path signatures have been successfully applied to feature selection in the modeling of diseases such as Alzheimer's and borderline personality disorder, demonstrating their effectiveness as features for time-series data [10], [11]. Consequently, this study utilizes Path Signatures (PS) to extract ECG features, leveraging its ability to capture temporal patterns in sequential data. However, Path Signatures may suffer from the curse of dimensionality as the path order increases. To address this, Lou et al. [12] proposed Path Development (PD) as a trainable alternative, which effectively mitigates the vanishing gradient problem when combined with LSTM. Building on these advances, this study proposes PathFusion-Net, a novel model that integrates Path Signatures, Path Development, CNN, and LSTM to extract both morphological spatial features from ECG images and temporal dynamics from ECG signals. This approach enables a more comprehensive representation of ECG data, improving classification performance. This study is the first to apply Rough Path Theory to ECG arrhythmia classification by integrating Path Signatures and Path Development within a deep learning framework. Path Signatures extract temporal features, which are fused through a fully connected layer, while Path Development enhances sequence modeling within the CNN-LSTM structure. The effectiveness of Rough Path Theory in arrhythmia classification tasks is thoroughly evaluated.

Our research not only introduces a novel ECG data classification strategy but also showcases the potential and advantages of Rough Path Theory in processing ECG signals, providing new technological means for the early detection and real-time monitoring of arrhythmias. The proposed PathFusion-Net leverages CNN for morphological feature extraction and LSTM combined with Path Development for capturing temporal dynamics. The model was evaluated on the MIT-BIH Arrhythmia Database, with arrhythmia classification conducted according to the AAMI standard under an inter-patient

paradigm. PathFusion-Net achieved an overall accuracy of 94.7%, outperforming the previous best-reported result by 2.9 percentage points, demonstrating superior classification performance and generalization capability. Specifically, the main contributions of this study are as follows:

1. This study proposes PathFusion-Net, a novel ECG arrhythmia classification model that integrates Rough Path Theory with deep learning. It combines Path Signatures and Path Development with CNN and LSTM to extract both temporal and spatial morphological features from ECG signals. The effectiveness of Rough Path Theory in arrhythmia classification tasks is evaluated.
2. The study emphasizes and employs standardized AAMI classification standards and an inter-patient paradigm for data classification training. A comparison with existing deep learning models demonstrates the superior performance of the proposed model.

II. RELATED WORK

Electrocardiogram (ECG) signals are essential for analyzing the heart's electrical activity. As a periodic signal, the ECG consists of distinct waveforms, each reflecting different phases of heart muscle contraction and relaxation, characterized by variations in frequency and amplitude. Key components, such as the P wave, QRS complex, and T wave, are crucial for accurately diagnosing heart diseases, as shown in Fig. 1. Traditional ECG diagnostic methods rely on comparing the morphological features and time intervals of these waveforms to differentiate between normal sinus rhythm and abnormal conditions [13]. However, the accuracy and efficiency of these conventional methods face significant limitations, especially when dealing with large volumes of patients. Consequently, machine learning methods for arrhythmia detection have been introduced, offering a novel solution for the automated analysis and accurate interpretation of ECG data.

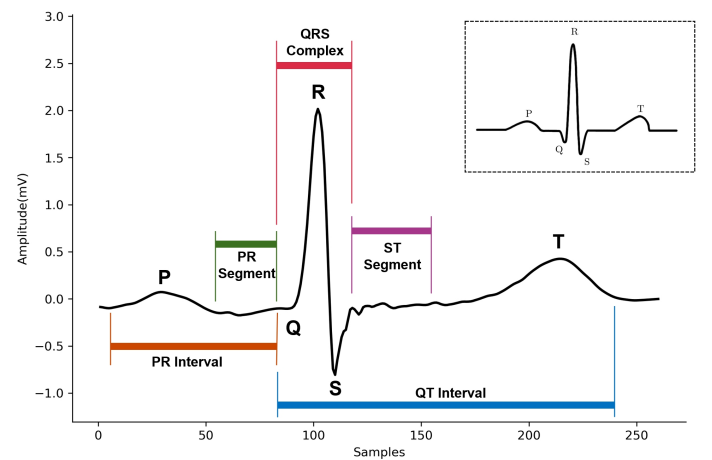


Fig. 1. A complete heartbeat cycle from the MIT-BIH dataset, illustrating key features of the ECG signal, including the PR Interval, PR Segment, QRS Complex, ST Segment, and QT Interval. The dashed box in the upper-right corner provides a schematic representation of the normal ECG waveform.

Over recent decades, significant progress has been made in automating arrhythmia classification from electrocardiogram

(ECG) signals, with various machine learning (ML) techniques playing a central role. Approaches such as Support Vector Machines (SVM) [14], K-Nearest Neighbors (K-NN) [15], Principal Component Analysis (PCA) [16], and adaptive back-propagation neural networks have been widely adopted. For instance, Faziludeen et al. [17] applied a multi-class SVM using a One-Against-One strategy to distinguish among left bundle branch blocks, ventricular premature beats, and normal heartbeats, while Zhu et al. [18] employed SVM after extracting morphological features from segmented P-QRS-T waves using PCA and dynamic time warping. Similarly, Lassoued et al. [19] developed a clinical decision support system based on Artificial Neural Networks (ANN), evaluating the performance impact of Levenberg-Marquardt and Bayesian Regularization algorithms. Latif et al. [20] further expanded the range of ML applications by using Random Forest and Naive Bayes for classifying ECG and electroencephalogram (EEG) signals. In addition, Zhang et al. [21] conducted a comprehensive comparison between intra-patient and inter-patient evaluation paradigms for heartbeat classification, using handcrafted statistical and morphological descriptors with a conventional classifier. This work highlighted the influence of evaluation protocol on reported performance, achieving 98.3% accuracy in the intra-patient and 93.5% in the inter-patient scenario. Similarly, Dias et al. [22] combined RR-interval, morphological, and higher-order statistical features with a lightweight classifier, explicitly testing robustness to R-wave jitter and demonstrating competitive results under inter-patient standard. While these traditional ML techniques have achieved reasonable success in arrhythmia classification, they often require manual feature extraction, adding complexity and potential bias to the process.

Unlike conventional ML models, deep learning has quickly gained widespread attention and application in ECG analysis, benefiting from increased data availability, enhanced computational power, and advancements in algorithms. Deep neural networks integrate feature extraction and classification, enabling the automatic capture of high-level features from ECG signals and achieving more accurate classifications [23]. For example, Liu et al. [24] developed a 17-layer CNN to extract deep features from ECG signals, combining them with expert-designed features for classifying nine ECG categories, and reported an accuracy of 0.81 on the training set using 5-fold cross-validation. Huang et al. [25] converted ECG signals into spectrograms using Short-Time Fourier Transform (STFT) and used a 2D CNN to classify arrhythmias, while Wang et al. [26] applied stacked denoising autoencoders to learn semantic representations of heartbeats, which were then classified with a BiLSTM network. Xiong et al. [27] designed a 1D CNN with residual connections for ECG classification, and Li et al. [28] used a six-layer 1D CNN based on the MIT-BIH Arrhythmia Database to categorize arrhythmias. Cao et al. [3] adopted a transfer learning approach by converting ECG signals into time-frequency spectrograms via STFT and fine-tuning a ResNet-18 model under inter-patient protocol, achieving 90.8% accuracy and outperforming several baseline CNNs. To address class imbalance, Alhichri et al. [29] introduced focal loss into a CNN-based heartbeat classifier, significantly

improving precision and recall for minority classes such as supraventricular ectopic beats.

Researchers have also investigated transforming ECG signals into 2D images for classification. For instance, Acharya et al. [30] implemented a CNN with three convolutional, max-pooling, and fully connected layers for heartbeat classification. Jun et al. [31] created an 11-layer 2D CNN model to analyze ECG images. Moreover, due to the temporal dependencies in ECG data, LSTM networks have become popular for capturing arrhythmic patterns. Gao et al. [32] developed an LSTM model to classify eight types of heartbeats from the MIT-BIH Arrhythmia Database, while Yildirim [33] used BiLSTM layers for detecting five heartbeat types, and Kim and Pyun [34] evaluated multiple LSTM architectures on MIT-BIH datasets. Furthermore, the work in [35] leveraged bidirectional LSTM layers to capture long-range temporal dependencies, yielding improved supraventricular arrhythmia detection compared to CNN-only approaches.

However, early single models based on CNN and LSTM for classifiers of ventricular fibrillation and atrial fibrillation (VF, AF) have not yet achieved optimal accuracy [36]. To address this problem, hybrid models of CNN and LSTM [37], [38] were proposed and showed high performance. In particular, Petmezas et al. [38] trained their model on the MIT-BIH Atrial Fibrillation Database, achieving a sensitivity of 97.87% and a specificity of 99.29% under a ten-fold cross-validation strategy, demonstrating its potential for real-time AF detection in routine ECG screening. Jin et al. [39] proposed a two-attention convolutional LSTM (TAC-LSTM) specifically for patient-independent and patient-specific AF classification and validated it in the MIT-BIH AF database. Wang et al. [40] developed a BiLSTM model with CNN and feature calibration for AF detection on a smaller dataset, where the proposed method achieved consistently higher generalization performance across geography groups than benchmark algorithms, with F1-scores ranging from 0.90 on RBDB to 0.95 on CPSC. While Chen et al. [41] combined CNN and LSTM to classify six arrhythmias including AF and ventricular fibrillation, where their four-class arrhythmia classification on the AFDB achieved an overall accuracy of 99.35% and an F1-score of 92.86%. In recent years, the hybrid CNN-LSTM design in [22] achieved an overall accuracy of 79.6% ($\delta=18$) and 80.6% ($\delta=0$), with balanced performance across all AAMI classes under the inter-patient evaluation protocol. Kachuee et al. [42] reported 81.2% accuracy for their deep transferable CNN representations in AAMI-compliant heartbeat classification, also demonstrating cross-database adaptability from MIT-BIH to PTB.

More recently, Transformer-based architectures have emerged as powerful alternatives for ECG classification due to their capability in modeling long-range dependencies and capturing global context. The ECGTransform [43] introduces a self-attention mechanism over segmented heartbeat sequences, enabling the model to focus adaptively on diagnostically relevant cardiac cycles. On the PTB dataset, it achieved an overall accuracy of 99.23%, sensitivity of 99.17%, and specificity of 99.24%, demonstrating competitive performance, particularly in improving minority

class recognition compared to conventional CNN-LSTM baselines. Beyond ECG-specific designs, generic time-series Transformers such as TimesNet [44] leverage multi-periodic temporal decomposition and convolution-enhanced attention blocks to model diverse temporal patterns. While TimesNet has shown strong performance across a variety of time-series benchmarks, its application to ECG remains limited in the literature, and the fixed multi-period decomposition it employs may be less suited for the highly variable cardiac cycles observed in arrhythmia data without substantial adaptation. These approaches highlight the potential of Transformer-based modeling for ECG analysis, motivating our inclusion of ECG-Transformer as a representative baseline in the comparative experiments.

In order to explore new and more efficient methods, we introduce path signature and path development methods from rough path theory.

III. PRELIMINARY

Rough path theory provides a mathematical framework for analyzing irregular, or ‘rough’ signals [8], [9]. Within this framework, the path signature serves as a robust tool for effectively capturing essential information about finite-length paths. Due to its strength as a feature extractor for time-series data, the path signature has been applied in various fields, including financial data modeling [45], handwritten character recognition [46], and human pose estimation [47]. This section introduces the path signatures approach, a key concept under rough path theory, highlights its limitations, and presents the improved path development approach with its advantages.

A. Preliminary of Path Signatures

Path signatures provide a principled and efficient way to extract features from time series data. Originating from rough path theory, the path signature of a path X captures the essential geometric properties and can be used to model the effects of the path on non-linear systems. A d -dimensional path X over the time interval $[0, T]$ can be represented as a continuous map $X : [0, T] \rightarrow \mathbb{R}^d$. For simplicity, in this paper, we focus on the paths of bounded variation. The path signature of X over a bounded interval $J \subset [0, T]$, denoted $S(X)_J$, is defined as:

$$S(X)_J = (1, X_J^1, X_J^2, \dots), \quad (1)$$

where X_J^k represents the k -fold iterated integral of the path:

$$X_J^k = \int_{u_1 < u_2 < \dots < u_k, u_i \in J} dX_{u_1} \otimes dX_{u_2} \otimes \dots \otimes dX_{u_k}. \quad (2)$$

for each $k \geq 1$. The truncated signature of order n , denoted as $\pi_n(S(X)_J)$, includes terms up to the n -th iterated integral:

$$\pi_n(S(X)_J) = (1, X_J^1, X_J^2, \dots, X_J^n). \quad (3)$$

Since the formal definition above may appear abstract, the reader is referred to Appendix A, which presents a worked example on a simple two-dimensional path. This example

illustrates the derivation of each signature term and shows how second-order cross terms correspond to oriented areas, following the style of prior expository work [48].

For a *linear* path $X_{[0,T]}$, its signature admits a closed-form expression:

$$S(X_{[0,T]}) = \exp(X_T - X_0), \quad (4)$$

where $\exp$ is the tensor exponential. In addition, the signature of the path has the multiplicative property known as Chen’s identity [49]: for any continuous paths X and Y of bounded variation, the signature of the concatenation of X and Y is the tensor product of the signature of X and Y . In a formula,

$$S(X * Y) = S(X) \otimes S(Y), \quad (5)$$

where $X * Y$ is the concatenation of the path.

In practice, we often observe discrete time series, which can be lifted to the piecewise linear path by linear interpolation. Thanks to Chen’s identity and Eqn. (4), the signature of the piecewise linear path is obtained by the product of its linear segments. This recursive formulation enables efficient computation of a long path.

However, the path signatures method faces several challenges, particularly when dealing with high-dimensional path features. As the network depth increases, the method tends to overfit. Specifically, the path signature method has the following three key limitations: Firstly, the Path Signature method suffers from the curse of dimensionality, leading to high computational complexity as the input dimension increases. Secondly, path signatures lack adaptability to changes in input data, limiting their flexibility and effectiveness. Lastly, truncating Path Signatures to a finite order can cause information loss, reducing the accuracy of the extracted features.

B. Preliminary of Path Development

To address these limitations, Lou et al. [12] proposed the path development method, a trainable alternative to path signatures grounded in Rough Path Theory. This method introduces trainable parameters that allow the model to better adapt to input data, mitigate overfitting when handling high-dimensional features, and more effectively preserve essential information. Building on this, the path development layer offers a solution that integrates flexibility and adaptability into feature extraction processes. To provide further clarity, the definition of path development is as follows.

Let G be a finite-dimensional Lie group with Lie algebra $\mathfrak{g}$. Consider a linear map $N : \mathbb{R}^d \rightarrow \mathfrak{g}(\mathfrak{m}; \mathbb{F})$ and a path $X \in \mathcal{V}^1([0, T]; \mathbb{R}^d)$, where $\mathcal{V}^1$ denotes the space of absolutely continuous functions. The development of X on G through N is defined as the solution to the differential equation:

$$dY_t = Y_t \cdot N(dX_t), \quad \forall t \in [0, T], \quad Y_0 = e, \quad (6)$$

where Y_t is the path at time t and e is the identity element in G , with matrix multiplication implied.

For a linear path $X \in \mathcal{V}^1([0, T]; \mathbb{R}^d)$, the development on G via $N \in L(\mathbb{R}^d, \mathfrak{g})$ is given by:

$$D_N(X)_{0,t} = \exp(N(X_t - X_0)). \quad (7)$$

This approach leverages the Picard iteration technique and the contraction mapping principle. Path development shares properties with path signatures, such as time invariance and multiplicative behavior, and is effective in applications like stock price or ECG analysis, where speed of recording is predetermined.

Lou et al. (2024) introduce the *path development layer* [12], a neural network module for time series data. For discrete time series, let $x = (x_0, x_1, \dots, x_N) \in \mathbb{R}^{d \times (N+1)}$ represent a d -dimensional sequence of length $N + 1$. Given a Lie algebra $\mathfrak{g}$ and $N \in L(\mathbb{R}^d, \mathfrak{g})$, we can define the development of x by linearly interpolating x . The map N is parameterized by coefficients $\theta = (\theta_1, \theta_2, \dots, \theta_d) \in \mathfrak{g}^d$, so $N_\theta : \mathbb{R}^d \ni x = (x^{(1)}, x^{(2)}, \dots, x^{(d)}) \mapsto \sum_{j=1}^d \theta_j x^{(j)} \in \mathfrak{g}$.

The path development layer is a mapping $D_\theta : \mathbb{R}^{d \times (N+1)} \rightarrow G^{N+1}$, yielding $z = (z_0, z_1, \dots, z_N)$ such that:

$$z_{n+1} = z_n \exp(N_\theta(x_{n+1} - x_n)), \quad z_0 = Id_m, \quad (8)$$

where $\exp$ is the matrix exponential, Id_m is the identity matrix, G^{N+1} denotes the Cartesian product of G , and $\theta \in \mathfrak{g}^d$ are trainable parameters.

The recursive structure in this layer resembles RNNs and signature layers. However, unlike RNNs, it does not require a fully connected network for hidden states. Compared to signature layers, it includes trainable weights adaptable to the data. By choosing a matrix Lie algebra $\mathfrak{g} \subset GL(m, \mathbb{F})$, the development can be represented within $GL(m, \mathbb{F})$, where the dimension $\dim_{\mathbb{F}} GL(m, \mathbb{F}) = m^2$ is fixed, independent of the path dimension d . This differs from the signature approach, in which the dimensionality increases geometrically with d .

Building on this definition, the path development layer offers several advantages that address the limitations of path signature methods, particularly in time series data processing.

- **Mathematical basis:** Like RNN and signature, the path development layer has the universality, as it is rich enough to approximate any continuous function on the path space.
- **Handling high-dimensional data:** Traditional path signature methods face challenges such as the curse of dimensionality, where the number of required features increases exponentially with the dimensionality of the data. The path development layer solves this problem by using Lie group matrices to represent the data, which significantly reduces the dimensionality without losing critical information.
- **Trainable and Data Adaptive:** Unlike path signatures, the path development layer is trainable. It can learn the optimal transformation of the input data, such as classification or regression tasks in high-dimensional spaces.
- **Using in Neural Networks:** The path development layer can be plugged into neural networks, especially in architectures that process sequential data such as RNN or LSTM. It helps mitigate problems such as vanishing gradients and explosions by providing a more stable structure for backpropagation.

The path development layer radically enhances a model's ability to capture and exploit the complex properties of sequential data, making it a powerful tool in areas involving time-series

prediction, speech recognition, and any field where it is critical to understand the underlying dynamics of data over time.

While the above subsections present the mathematical definitions of Path Signatures and Path Development, their suitability for ECG analysis can also be understood from a signal-properties perspective. ECG waveforms are inherently non-stationary their morphology and baseline shift over time due to physiological and measurement variations. They are also oscillatory, with distinct repetitive structures such as the P wave, QRS complex, and T wave, whose fine-scale variations carry diagnostic significance. Moreover, ECG patterns are multi-scale, as both short-term features (e.g., QRS width) and long-term features (e.g., RR interval trends) are clinically relevant. Path Signatures encode the ordered geometry of the signal trajectory through iterated integrals, preserving temporal ordering and multi-scale interactions in a fixed, mathematically principled representation. Path Development extends this by introducing trainable transformations that adapt to data characteristics, reducing dimensionality while retaining critical structure. Compared with standard LSTM or Transformer models which learn internal state representations via optimization the path-based approach provides a more explicit and stable encoding of temporal structure, which can be especially advantageous for biomedical signals with rich morphology and irregular dynamics like ECGs.

IV. METHOD

This section presents PathFusion-Net, a novel ECG arrhythmia classification model that integrates Rough Path Theory with deep learning. The model combines Path Signatures and Path Development with a CNN-LSTM framework. The following subsections detail its architecture and implementation.

Due to the complex and variable structures reflected in ECG signals, features from a single domain are often insufficient to capture all the subtle changes. In addition, the biometric nature of ECG data poses a significant challenge for data augmentation. Therefore, in the classification of arrhythmias, multi-feature fusion typically yields better classification performance. This paper proposes a multi-feature fusion network, PathFusion-Net. After denoising, the ECG signals undergo R-wave detection and heartbeat segmentation. The segmented one-dimensional ECG signals are then transformed into 256×256 grayscale images, which are processed by CNN to extract morphological features. Concurrently, the original one-dimensional ECG signals are processed through both the DevLSTM network and the Path Signature method to capture temporal features from different perspectives. The extracted features are concatenated and passed through Batch Normalization and a Dense layer before final classification. Then a dropout layer is applied prior to the final arrhythmia classification, achieved using a Softmax layer. In summary, this study combines the capabilities of CNN, LSTM networks, and Rough Path method to effectively capture the spatio-temporal features of ECG data. This approach provides a novel perspective and method for ECG analysis. Detailed descriptions of the DevLSTM CNN and Path Signatures modules are provided below.

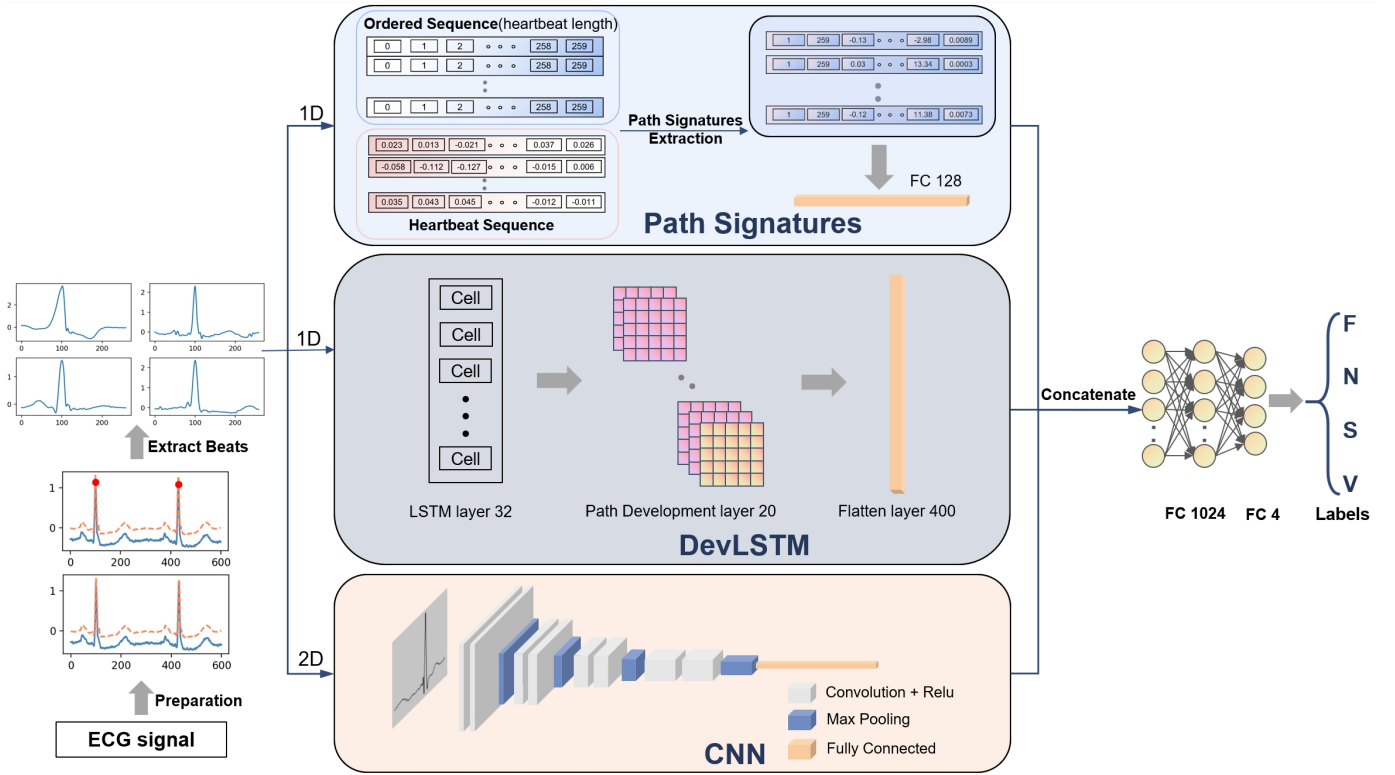


Fig. 2. PathFusion-Net Methodology Diagram.

As shown in Fig. 2, the network integrates two-dimensional morphological features from a CNN and one-dimensional temporal features from Path Signatures and the combined DevLSTM module.

1) *CNN*: CNN is particularly well-suited for handling local features in images and is commonly used for extracting morphological features from data. In this study, we input the transformed 256×256 pixel grayscale images into the network. The data first passes through five layers of CNN, each with the same structure: convolution layer, ReLU activation function, batch normalization layer, and max pooling layer. To simplify hyperparameter tuning, each layer uses the same number of filters and stride values. As the network depth increases, the number of convolution kernels is gradually increased to ensure that the network can extract complex features from various perspectives and scales, thereby improving the model's performance. Finally, the data passes through a fully connected layer, where the features extracted by the CNN model are output. Table I gives the details about the CNN model.

2) *DevLSTM*: Although traditional LSTM networks are designed to address the issue of long-term dependencies, they can still encounter problems such as gradient vanishing or explosion in practical applications. The Path Development Layer, by leveraging the mathematical structure of Lie groups, helps stabilize the gradient during the training process, thereby alleviating these issues. Additionally, the Path Development Layer performs mathematical transformations based on path development theory on time series data, extracting features that effectively characterize the data. This transformation is trainable, meaning it can adaptively adjust based on specific

TABLE I
ARCHITECTURE OF PROPOSED CNN MODEL

Layer	Type	Kernel size	Stride	Kernel	Output size
Layer 0	Input				$256 \times 256 \times 1$
Layer 1	Conv	3×3	1	64	$256 \times 256 \times 64$
Layer 2	Conv	3×3	1	64	$256 \times 256 \times 64$
Layer 3	MaxPool	2×2	2		$128 \times 128 \times 64$
Layer 4	Conv	3×3	1	128	$128 \times 128 \times 128$
Layer 5	Conv	3×3	1	128	$128 \times 128 \times 128$
Layer 6	MaxPool	2×2	2		$64 \times 64 \times 128$
Layer 7	Conv	3×3	1	256	$64 \times 64 \times 256$
Layer 8	Conv	3×3	1	256	$64 \times 64 \times 256$
Layer 9	MaxPool	2×2	2		$32 \times 32 \times 256$
Layer 10	Conv	3×3	1	512	$32 \times 32 \times 512$
Layer 11	Conv	3×3	1	512	$32 \times 32 \times 512$
Layer 12	MaxPool	2×2	2		$16 \times 16 \times 512$
Layer 13	FC				2048
Layer 14	Output				2048

tasks, enhancing the model's understanding and representation of the data. Thus, by integrating LSTM with the Path Development Layer to form the new DevLSTM network, long-term dependencies in time series data can be better captured.

The proposed network receives one-dimensional ECG signals as input, which are processed through a single LSTM layer with a hidden size of 32. The resulting temporal features are then passed to the Path Development Layer, where orthogonal matrices are used as internal Lie groups, with matrix length defined by the parameter Dev Number. The last matrix in the sequence is chosen for output because it contains the complete temporal characteristics of the preceding sequence data. These features are then passed through a flatten layer for output.

3) *Path Signatures*: Using rough path theory, the Path Signature (PS) method demonstrates a strong capability to characterize temporal sequence data, capturing patterns from lower to higher orders (coarse to fine). In this study, each heartbeat sequence is treated as a time series, with features extracted via the PS method. These features are then passed through a fully connected (FC) layer and fused with spatial and temporal features within PathFusion-Net for classification.

V. DATASETS AND PREPARATION

This section provides an overview of the ECG datasets used in this study, including both public and private datasets, with a detailed description of their sources, sample sizes, and classification methods.

A. Dataset

1) *Public Dataset*: The MIT-BIH Arrhythmia Database, developed by the Massachusetts Institute of Technology and Beth Israel Hospital, is a foundational dataset for arrhythmia research and was the first globally recognized standard for evaluating arrhythmia detection methods, released in 1980 [50]. It contains 48 ECG recordings from 47 subjects (25 men aged 32-89 and 22 women aged 23-89), with recordings 201 and 202 originating from the same individual. Each 30-minute recording is sampled at 360 Hz and includes data from two leads: lead II (primarily a modified limb lead) and a second lead, typically V1, though occasionally V2, V4, or V5. For this study, only lead II data is used for consistency. Since the database provides precise R-wave annotations, these are directly utilized here, as R-wave detection is not the study's focus. Each heartbeat segment comprises 100 sample points before and 160 sample points after the R-wave peak, totaling 260 points to capture a full cardiac cycle, as shown in Fig. 1.

The MIT-BIH Arrhythmia Database contains 15 heartbeat types, which in this study are grouped into five categories: normal beat (N), supraventricular ectopic beat (SVEB), ventricular ectopic beat (VEB), fusion beat (F), and unknown beat (Q). To facilitate inter-patient data partitioning, we adopted the widely used DS1/DS2 split protocol proposed by DeChazal et al. [51], as listed in Table II. DS1 was used for training and validation, while DS2 was reserved for testing, ensuring that no ECG segments from the same patient appear in more than one subset. The validation set was obtained through an intra-patient split within DS1 to maintain balanced class proportions and enable stable early stopping. The same principle of patient exclusivity was applied to our private clinical dataset.

TABLE II
PATIENT-EXCLUSIVE DS1/DS2 SPLIT FOR THE MIT-BIH ARRHYTHMIA DATABASE.

Subset	Record IDs
DS1 (training + validation)	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 200, 201, 203, 205, 207, 208, 215, 220, 223, 230, 232
DS2 (test)	100, 103, 105, 111, 113, 117, 121, 123, 209, 210, 212, 213, 214, 219, 221, 222, 228, 231, 233, 234

TABLE III
THE DATASET OF MIT-BIH ARRHYTHMIA DATABASE

	F	N	S	V	Total
DS1	416	46366	1973	4613	53368
DS2	386	44719	808	2395	48308

The dataset consists of two parts: one-dimensional signal data and two-dimensional image data. For the two-dimensional image data, an image cropping augmentation method was applied to the training set for the three categories (F, S, and V) with smaller sample sizes. The final amount of data used is shown in Table IV.

TABLE IV
DATASET FOR PATHFUSION-NET MODEL

	F		N		S		V		Total	
	2D	1D	2D	1D	2D	1D	2D	1D	2D	1D
Training	3120	312	34024	14790	1479	34590	3459	86524	39274	
Validation	104		11342		494		1154			13094
Test	386		44719		808		2395			48308

1D refers to the original heartbeat signal data, while 2D refers to the augmented data generated by converting the signal into corresponding images.

2) *Private Dataset*: This study also utilized a private database obtained from the Department of Cardiology at Shaanxi Honghui Hospital. The database consists of 24-hour, 12-lead long-term ECG recordings sampled at 15 Hz, categorized into two groups: atrial fibrillation (AF) patients and non-AF patients. To ensure data independence, we strictly divided the training and test sets at the inter-patient split paradigm. After dividing the patient groups, multiple ECG segments from the second lead were randomly extracted, and image cropping data enhancement was also used in the training set to construct the final dataset. The detailed statistics of the dataset are presented in Table V.

TABLE V
PRIVATE ATRIAL FIBRILLATION DATASET

	AF		Non-AF		Total	
	2D	1D	2D	1D	2D	1D
Training	2800	280	9500	950	12300	1230
Validation	115		355		470	
Test	226		862		1088	

B. Preparation

1) *Denoising*: ECG signals are primarily affected by three types of noise: industrial frequency interference, electromyographic interference, and baseline drift. Traditional Fourier analysis, which uses the Fourier transform as a global transformation, faces limitations when analyzing non-smooth signals. In contrast, the wavelet transform, a local transformation in both time and frequency domains, effectively extracts information by applying multi-scale operations such as scaling and translation. This advantage enables the wavelet transform to address challenges that the Fourier transform cannot, making it particularly suitable for ECG signal denoising.

The process of denoising ECG signals using the wavelet transform involves three steps: (1) Selecting a wavelet basis function to separate noise from the signal. Based on the

quantitative analysis by Peng and Wang [52], the db6 wavelet (vanishing moment $N = 6$) achieves the minimum high-frequency energy after the first-level decomposition for ECG signals, which offers an optimal trade-off between noise suppression and preservation of waveform morphology; (2) Considering that most human ECG signals occupy the frequency range of approximately 0.05–100 Hz, and given the 360 Hz sampling rate of the MIT-BIH dataset, a 9-level discrete wavelet decomposition enables effective isolation of baseline drift (<0.5 Hz), industrial frequency interference (50/60 Hz), and high-frequency electromyographic noise, while retaining diagnostically relevant signal components; (3) Labeling the detailed coefficients as D and the approximate coefficients as A. The resulting frequency ranges of each scale after the 9-scale decomposition are presented in Table VI.

TABLE VI
SIGNAL DECOMPOSITION

signal decomposition	decomposition level	frequency range(Hz)
D1	1	180-360
D2	2	90-180
D3	3	45-90
D4	4	22.5-45
D5	5	11.25-22.5
D6	6	5.625-11.25
D7	7	2.8125-5.625
D8	8	1.40625-2.8125
D9	9	0.703125-1.40625
A9	9	0-0.703125

After scale decomposition of the ECG signal using wavelet transform, the D1 and D2 scales contain a large amount of high-frequency noise, while baseline drift noise is mainly concentrated in the A9 scale. Therefore, the wavelet coefficients on the D1, D2 and A9 scales can be set to zero, and a soft thresholding function can be applied to the wavelet coefficients at other scales to filter out other noise components, thereby achieving the goal of signal denoising. soft threshold function:

$$w_{\lambda} = \begin{cases} \text{sgn}(w)(|w| - \lambda), & \text{if } |w| \geq \lambda \\ 0, & \text{if } |w| < \lambda, \end{cases} \quad (9)$$

where w denotes the wavelet coefficients, N is the length of the signal, λ is the pre-selected threshold, and the formula for λ is

$$\lambda = \frac{\text{median}|w| \cdot \sqrt{2 \ln N}}{0.6745}. \quad (10)$$

Following noise coefficient filtering, the denoised ECG signal is reconstructed using the inverse discrete wavelet transform, enhancing signal quality and providing a clearer representation of the underlying physiological processes. As shown in Fig. 3, the raw ECG signal initially exhibited baseline drift, reflected by a -0.25mV offset. After applying a 9-level discrete wavelet transform (db6 wavelet), baseline wander was effectively removed, restoring the signal around the zero baseline. In addition, high-frequency noise was reduced without distorting key morphological features, demonstrating the suitability of wavelet-based denoising for ECG signal processing.

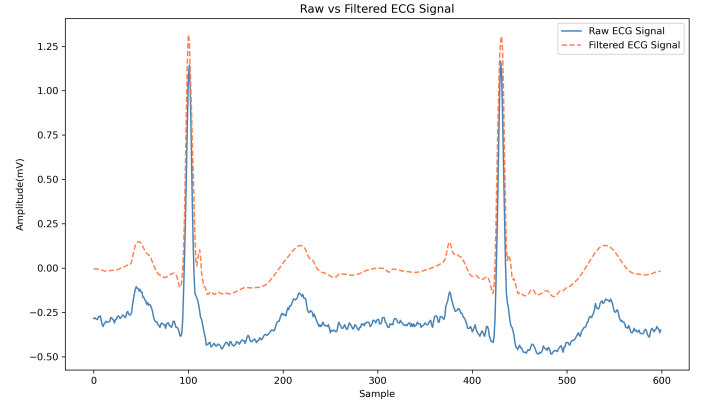


Fig. 3. Comparison between the raw and filtered ECG signals in category F.

2) Augmentation: ECG data, as a type of biosignal, is highly sensitive to small fluctuations in sampling points, which could potentially indicate the presence of a pathological condition. In one-dimensional form, it is difficult to apply data augmentation without risking the loss or distortion of critical physiological information. For the proposed model, both one-dimensional and two-dimensional representations are required. The raw ECG signals were first segmented into single beats using the annotation files provided by the MIT-BIH Arrhythmia Database and our private clinical dataset. Each beat was centered on the R-peak and contained the complete P wave, QRS complex, and T wave. For the CNN branch, each 1D beat was transformed into a 2D grayscale image of size 256×256 pixels, with the horizontal axis representing time and the vertical axis representing amplitude.

To balance the number of samples across classes and improve generalization, we applied a nine-region cropping strategy with overlapping windows (Fig. 4). Once the 1D ECG segments are transformed into 2D grayscale images, spatially shifting the crop region can simulate variations in signal alignment and scaling. This type of augmentation has been successfully applied in previous ECG image-based classification studies and has been shown to improve robustness to such variations [53]. While certain cropped regions may omit portions of specific waveform components (e.g., partial R-wave truncation), other crops in the nine-region set preserve the complete morphological structure of the heartbeat.

Importantly, both the original (uncropped) images and the cropped versions were included in the training set, ensuring that the model was exposed to complete waveform morphologies as well as their spatially shifted variants. This approach did not lead to a loss of diagnostic information; instead, it increased data diversity and improved the model's robustness to spatial variability.

VI. EXPERIMENT RESULT

A. Model Size and Computational Complexity

The proposed PathFusion-Net contains approximately 275 million parameters, corresponding to a storage size of 1.1 GB in FP32 precision. The computational complexity is estimated

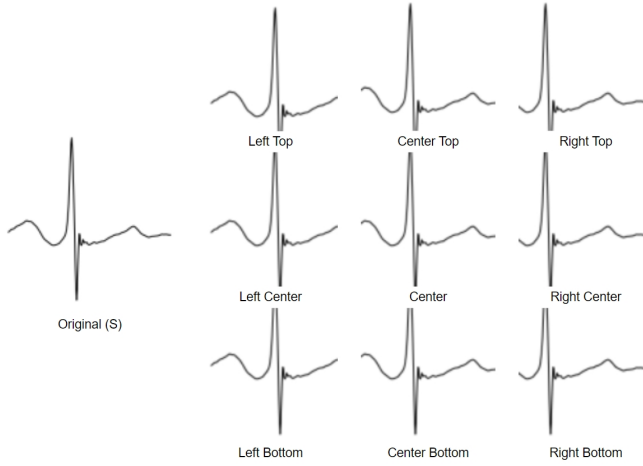


Fig. 4. Original ECG image and nine cropped images.

at 23.4 GFLOPs per heartbeat sample in the inter-patient paradigm. This relatively large size arises from the hybrid multi-branch design, which integrates a CNN branch for 2D morphological feature extraction, a DevLSTM branch for temporal dependencies, and a Path Signature branch for structured time-series features. To mitigate potential overfitting, we employed extensive data augmentation, dropout regularization (0.2 in the DevLSTM branch and 0.5 in the fusion MLP), batch normalization, and early stopping based on validation performance. Despite the parameter count, no signs of overfitting were observed, and the model achieved strong generalization to both the MIT-BIH dataset (over 100,000 labeled beats across four classes) and an independent private clinical dataset. Inference speed was measured at 62 ± 6 ms per heartbeat on GPU and 189 ± 10 ms on CPU, demonstrating practical feasibility for both real-time monitoring and offline Holter ECG analysis.

To further examine deployment feasibility, we additionally tested the model on two low-power embedded devices. Direct deployment without optimization yielded latencies of 12.8 s/beat on a Raspberry Pi 4B (4 GB) and 4.2 s/beat on a Jetson Nano (4 GB). Lightweight optimizations improved performance to 6.3 s/beat on the Raspberry Pi (via dynamic quantization) and 1.9 s/beat on the Jetson Nano (via TensorRT). Power consumption during inference was approximately 9W and 7W, respectively, much lower than desktop CPU (96W) and GPU (111W). While these results confirm the energy efficiency of embedded platforms, their inference latency remains above the strict real-time threshold, suggesting that further techniques such as pruning, aggressive quantization, or compact backbone redesign will be required for wearable or edge deployment.

B. Experimental Setup

This research utilized Python 3.10.13 for implementation. The deep learning models were developed and trained using PyTorch 2.0.1, supported by the iisignature library (version 0.24). The system featured an Intel i9-13900H CPU with 14 cores, 20 threads, and a base frequency of 2.60GHz. Additionally, an NVIDIA RTX 4070 Laptop GPU with 8GB of

video memory, driver version 31.0.15.4680, and CUDA 11.7 compatibility provided computational support for the deep learning algorithms.

Training configuration and convergence stability: The model was trained using the Adam optimizer (initial learning rate of 1×10^{-3} , $\beta_1 = 0.9$, $\beta_2 = 0.999$) and the cross-entropy loss function, with a mini-batch size of 32. To stabilize training and prevent overfitting, Batch Normalization was applied after each convolutional or fully connected layer, and Dropout was employed with a rate of 0.5 in the fusion MLP and 0.2 in the DevLSTM branch. A ReduceLROnPlateau scheduler monitored the validation loss with *mode* set to min, reducing the learning rate by a factor of 0.5 after 10 consecutive epochs without improvement. An early stopping strategy was also adopted, terminating training if the validation loss did not decrease for 15 consecutive epochs ($\text{min_delta} = 10^{-4}$), with the model reverted to the checkpoint achieving the highest validation accuracy. These strategies ensured smooth and stable convergence without divergence or gradient explosion.

C. Evaluation Metrics

To assess classification performance, this study employed four metrics: Precision, Recall, Accuracy, and F-score. These metrics are based on the concepts of true positive (TP), true negative (TN), false positive (FP), and false negative (FN). While Precision, Recall, and Accuracy are widely used, the F-score provides an additional measure by evaluating the balance between precision and recall.

- **Macro-averaged F-score**

macro-averaged F1-score (*F1-score*) specifically reflects the relationship between the actual positive labels and those predicted by the classifier, averaged across each class.

$$F1\text{-score} = \frac{Pre \times Rec}{Pre + Rec} \times 2$$

D. Proposed Model on Public Dataset

While Path Signatures (PS) are effective in extracting temporal features from ECG signals, they suffer from the curse of dimensionality, leading to high computational costs and potential information loss as the signature order increases. Additionally, their fixed nature limits adaptability, making it challenging to capture complex temporal patterns in ECG data. To address these limitations, this study proposes PathFusionNet, which integrates Path Signatures, Path Development, CNN, and LSTM for a more comprehensive feature extraction approach. CNN extracts morphological features from 2D ECG images, while DevLSTM, combining Path Development and LSTM, dynamically learns temporal features from 1D ECG signals. This design overcomes the rigidity and high-dimensional constraints of standalone Path Signatures, enhancing classification performance. The proposed model employs an early stopping strategy on the validation set, halting training automatically when the minimum loss in the validation set reaches a specified threshold. The model's final evaluation is then conducted on the test set. Results indicate that, despite the

TABLE VII
EVALUATION OF PATHFUSION-NET WITH DIFFERENT ORDERS ON PUBLIC DATASET

	PS-only method			PathFusion-Net without PS			Order 3			Order 4		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
F	6.81%	5.70%	6.21%	3.91%	18.12%	6.41%	1.23%	1.04%	1.13%	2.33%	1.55%	1.87%
N	98.11%	91.09%	94.47%	97.34%	95.38%	96.35%	97.82%	90.80%	94.18%	97.66%	92.85%	95.19%
S	69.90%	83.08%	75.92%	75.73%	92.34%	83.38%	11.24%	59.65%	18.92%	17.29%	74.01%	28.04%
V	31.59%	69.19%	43.38%	93.2%	87.9%	90.44%	65.50%	59.96%	62.65%	80.70%	69.85%	74.89%
accuracy		89.02%			91.52%			88.03%			90.67%	
	Order 5			Order 6			Order 7			Order 8		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
F	1.15%	0.52%	0.71%	6.81%	5.70%	6.21%	3.90%	18.13%	6.41%	5.75%	3.37%	6.41%
N	97.77%	94.55%	96.13%	98.69%	95.76%	97.20%	98.87%	95.76%	97.29%	97.94%	95.13%	96.51%
S	18.35%	64.23%	28.54%	24.64%	86.88%	38.39%	75.74%	92.33%	83.21%	20.86%	70.17%	32.16%
V	85.37%	73.36%	78.91%	94.96%	69.19%	80.05%	95.16%	87.93%	91.41%	91.81%	73.95%	81.91%
accuracy		92.25%			93.57%			94.69%			92.93%	

We compared the precision, recall, F1-score, and overall accuracy of PathFusion-Net with different Path Signature orders, as well as the performance of PS-only method and PathFusion-Net without PS, on the MIT-BIH arrhythmia dataset across different arrhythmia categories. The **best**, **second** and **third** scores are highlighted.

significant imbalance between different categories, the model maintains high classification accuracy across major categories.

In PathFusion-Net, the order of Path Signatures (PS) plays a crucial role in feature extraction and classification performance. Initially, as the order increases, the model benefits from richer temporal representations, leading to improved accuracy. However, beyond a certain threshold, the curse of dimensionality introduces redundancy and noise, which negatively impacts classification performance. To analyze this effect, we conduct experiments to evaluate how different PS orders influence the model's final accuracy. Table VII presents the classification performance of PathFusion-Net under different Path Signature orders, alongside the PS-only method and PathFusion-Net without PS. The results indicate that the full PathFusion-Net achieves the highest overall accuracy of 94.69% at an optimal PS order, surpassing both the PS-only method (89.02%) and the PathFusion-Net without PS (91.52%). For the Normal (N) category, PathFusion-Net maintains high precision and recall rates, exceeding 97% accuracy across different orders. In the Supraventricular Ectopic Beat (S) category, despite class imbalance, the model achieves up to 92.3% recall. However, higher PS orders do not always yield better performance. While increasing the order initially improves accuracy, excessive complexity results in diminishing returns. These findings validate the effectiveness of integrating Path Signatures within PathFusion-Net, confirming that neither the PS-only method nor removing PS entirely can achieve the same level of classification performance. The results also underscore the importance of selecting an optimal PS order, balancing feature richness with computational efficiency for robust arrhythmia classification in real-world applications.

Table VIII presents a performance comparison between the proposed PathFusion-Net and several representative state-of-the-art methods on the MIT-BIH Arrhythmia Database under the inter-patient evaluation protocol. The compared approaches span a diverse range of architectures, including traditional machine learning pipelines with handcrafted features [21], [22], CNN-based methods leveraging time-frequency representations [3], cost-sensitive learning techniques such as focal loss [29], recurrent neural networks with bidirectional LSTM

layers [35], hybrid CNN-LSTM frameworks with transferable representations [42], and recent Transformer-based models for ECG analysis such as ECG-Transformer [43].

From the results, we observe that traditional ML-based pipelines [21], [22] achieve competitive performance in detecting majority classes but generally underperform in minority class recognition, particularly for supraventricular ectopic beats (S) and fusion beats (F). CNN-only models, such as the STFT-ResNet18 approach [3], show improved overall accuracy or dominant classes but still face challenges in capturing long-range temporal dependencies, which can affect performance on morphologically similar arrhythmias. The focal loss-enhanced CNN [29] alleviates some class imbalance issues, leading to higher recall for the S class compared to standard CNN baselines, but the gain in minority classes comes at a slight cost to overall accuracy.

Recurrent architectures [35] improve sensitivity to temporally extended patterns, particularly for supraventricular arrhythmias, but their reliance on 1D sequential processing limits their ability to fully exploit multi-scale morphological variations. Hybrid CNN-LSTM designs [22], [42] demonstrate more balanced performance across classes, with [42] additionally showing robustness in cross-database adaptation scenarios. However, these methods typically do not incorporate structured time-series representations, potentially limiting their capacity to model fine-grained geometric variations in ECG signals.

Transformer-based methods, such as ECGTransform, leverage self-attention to capture global dependencies in the heart-beat sequence and have shown strong performance on long-range temporal modeling tasks. In our experiments, ECGTransform achieved competitive overall accuracy and particularly high recall for the N class, reflecting its strength in modeling contextual relationships across the full heartbeat. Nevertheless, the quadratic complexity of self-attention increases memory and computation demands, and in our inter-patient evaluation setting, ECGTransform lower pre and recall for the imbalanced class compared to PathFusion-Net, indicating potential limitations in handling extreme class imbalance without additional structural priors.

By contrast, PathFusion-Net consistently achieves the high-

est or near-highest Accuracy across all AAMI classes (absolute gain of 2.9% over the next best method). This performance gain can be attributed to the integration of CNN-based morphological feature extraction, Path Signature-based structured sequence encoding, and Path Development-enhanced LSTM modeling. The combination allows the model to capture both local morphological patterns and global temporal dependencies, while the path-based representation provides robustness to non-stationarity and waveform variability inherent in ECG signals.

Overall, the results in Table VIII demonstrate that PathFusion-Net not only matches or surpasses the best-performing existing methods, including Transformer-based approaches, in terms of average metrics, but also offers a more balanced classification performance across both majority and minority arrhythmia classes. This suggests that the proposed path-based multi-branch design provides a promising direction for ECG arrhythmia classification, particularly under strict inter-patient evaluation settings where robustness to distribution shifts and class imbalance is critical.

TABLE VIII
COMPARISON OF DEEP LEARNING MODEL PERFORMANCE USING AN INTER-PATIENT SPLIT PARADIGM

Work	Accuracy (%)	Arrhythmia types			
		F	N	S	V
		(n=386)	(n=44719)	(n=808)	(n=2395)
Romdhane [29]	62.1	0.0/0.0	95.6/64.0	0.0/0.0	12.7/79.3
Oliveira [35]	66.0	N/A	62.0/72.0	21.0/14.0	88.0/86.0
Dias($\delta=18$) [22]	79.6	4.4/81.4	99.4/77.9	36.9/90.6	92.6/87.2
Dias($\delta=0$) [22]	80.6	4.6/81.4	99.6/79.2	39.7/92.2	92.8/87.2
Kachuee [42]	81.2	1.0/1.3	94.4/84.5	0.0/0.0	30.9/ 92.4
Zhang [21]	86.6	13.7/93.8	98.9/88.9	35.9/79.1	92.7/85.5
Minh Cao [3]	90.8	1.3/0.3	95.3/95.1	13.0/9.0	68.2/88.4
El-Ghaish [43]	91.8	5.6/27.7	98.8/92.8	36.6/84.9	80.9/86.1
PathFusion-Net(Order 7)	94.7	3.9/18.1	98.9/ 95.8	75.7/92.3	95.2/87.9

We compared the performance of deep learning methods on the MIT-BIH arrhythmia dataset using an inter-patient split paradigm. The comparison includes precision (Pre) and recall (Rec) values for F, N, S, and V categories, along with overall accuracy. We highlight the **best** scores.

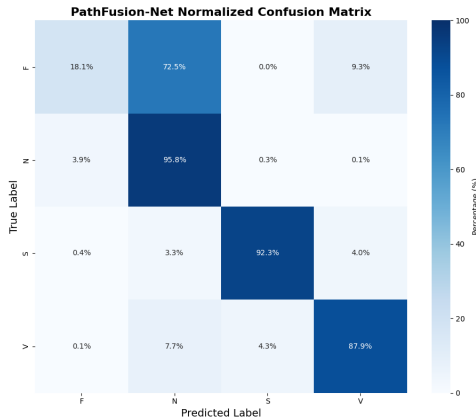


Fig. 5. PathFusion-Net Normalized Confusion Matrix.

To further evaluate classification performance under the severe class imbalance in Table III, we provide the normalized confusion matrix for the test set results (Fig. 5). The confusion matrix offers an intuitive visualization of class-wise prediction patterns. For example, 72.5% of F beats are misclassified as N

beats, explaining the relatively low F1-score for the F category. In contrast, the model achieves high discrimination between S and V beats, with only 4.3% of S beats predicted as V and 7.7% of V beats predicted as N. This analysis complements the precision/recall metrics and provides a clearer understanding of error distributions in multi-class classification.

E. Evaluation on Private Dataset

This study also utilized data from the Department of Cardiology of the Honghui Hospital in Shaanxi Province, China. Each sample consists of 24-hour 12-channel long-term ECG recordings from both patients with atrial fibrillation and healthy individuals. This dataset allows us to evaluate the effectiveness of Rough Path Theory in extracting meaningful features from real-world ECG signals.

Table IX presents the classification performance of PathFusion-Net under different Path Signature orders on the private dataset, alongside the PS-only method and PathFusion-Net without PS. Similar to the results on the public dataset, the findings confirm that introducing Path Signatures improves classification accuracy, with PathFusion-Net achieving its highest accuracy (95.13%) at order 6, outperforming both the PS-only method (91.54%) and the PathFusion-Net without PS (92.00%). However, as the PS order continues to increase beyond this point, the model's performance starts to decline, reinforcing the trade-off between capturing richer temporal features and managing high-dimensional complexity. This highlights a key limitation of the Path Signature method, which requires careful selection of the optimal order to balance feature extraction and computational efficiency.

F. Discussion

This study explores the integration of Rough Path Theory into deep learning for ECG arrhythmia classification, introducing Path Signatures (PS) and Path Development into a unified framework. To evaluate the effectiveness of this integration, we conducted experiments comparing PathFusion-Net with and without PS, as well as a PS-only model, analyzing how different PS orders influence classification performance. The results demonstrate the advantages of combining morphological features from CNN, temporal representations from DevLSTM, and structured path-based features from PS, reinforcing the importance of hybrid models for biomedical signal analysis. Integrating Path Signatures significantly enhances classification performance by improving temporal feature representation. While PS alone suffers from dimensionality issues, its structured encoding of time-series dependencies complements Path Development's adaptability. The full PathFusion-Net achieves the highest accuracy (94.69%), outperforming both PathFusion-Net without PS (91.52%) and the PS-only model (89.02%), confirming the benefits of feature fusion. Beyond overall performance improvement, the choice of Path Signature order also plays a crucial role in classification results. As the PS order increases, the model initially benefits from richer temporal representations, leading to improved accuracy (e.g., 94.69% at order 7). However, beyond a certain threshold, higher-order Path Signatures introduce excessive

TABLE IX
EVALUATION OF PATHFUSION-NET WITH DIFFERENT ORDERS ON PRIVATE DATASET

	PS-only method			PathFusion-Net without PS			Order 3			Order 4		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score	precision	Recall	F1-score
Non-AF	91.86%	98.60%	95.11%	91.13%	98.96%	94.88%	90.72%	99.77%	95.03%	92.86%	99.65%	96.14%
AF	92.81%	67.39%	78.09%	94.08%	63.27%	75.66%	98.57%	61.06%	75.41%	98.16%	70.80%	82.26%
accuracy		91.54%			92.00%			91.73%			93.66%	
	Order 5			Order 6			Order 7			Order 8		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
Non-AF	93.36%	99.54%	96.35%	94.60%	99.54%	97.00%	93.94%	98.96%	96.38%	90.79%	99.64%	94.96%
AF	97.63%	73.01%	83.54%	97.79%	78.31%	86.98%	95.00%	75.66%	84.24%	97.20%	61.50%	75.34%
accuracy		94.03%			95.13%			94.11%			91.64%	

We compared the precision, recall, F1-score, and overall accuracy of PathFusion-Net with different Path Signature orders, as well as the performance of PS-only method and PathFusion-Net without PS, on the private dataset across different arrhythmia categories. The **best**, **second** and **third** scores are highlighted.

complexity and redundancy, leading to a decline in performance (e.g., 92.93% at order 8), highlighting the trade-off between capturing fine-grained temporal details and managing feature dimensionality, and emphasizing the importance of selecting an optimal PS order in deep learning models.

For the fusion beat (F) class, our model achieves a precision of 3.9%, which is in a similar range to most compared methods in Table VIII. This consistently low performance underscores the difficulty of detecting this category, largely due to its extreme rarity in the dataset and its morphological similarity to both normal and ventricular beats. To further examine this issue, we conducted a controlled experiment by downsampling the other classes to match the F class size. Under this balanced setting, the F class accuracy improved markedly from 18.1% to 63.0%, confirming that data scarcity is the dominant factor behind its poor recognition. However, this improvement came at the cost of degraded performance for the majority classes, highlighting the trade-off between minority-class recognition and overall accuracy. Addressing this limitation will likely require more advanced strategies such as synthetic data generation, class-weighted training, or specialized sub-networks. It is worth noting that this study focuses on classifying arrhythmia types defined in the AAMI standard, all of which are present in the training data. Consequently, the model may misclassify previously unseen arrhythmia types into the most similar known category. Future work will explore open-set recognition and anomaly detection approaches to enhance the model's ability to handle arrhythmia classes not represented in the training dataset. While the model performs well on the MIT-BIH Arrhythmia Database, its effectiveness on other types of ECG data remains underexplored. To overcome these limitations and strengthen our study, we propose several directions for future research: Efficiency Enhancements: Future work should focus on optimizing model architectures to reduce computational costs while maintaining or potentially improving accuracy. Approaches such as pruning, quantization, and knowledge distillation are promising techniques to explore. Extension to Other Biomedical Signals: Investigating the applicability of Rough Path Theory and the developed methodologies to other types of biomedical signals, such as EEG, EMG, and Photoplethysmography (PPG), which is a widely used non-invasive optical technique for cardiovascular monitoring. For example, Ebrahimi and Gosselin [54] provide a comprehensive

methodological review of ultralow-power PPG sensors, highlighting their potential for continuous monitoring in wearable devices. Lastly, to address the challenge of class imbalance in ECG datasets, future studies could explore advanced data augmentation methods, such as GANs, to create synthetic ECG signals. Such techniques could enhance the model's ability to learn from underrepresented arrhythmia types, improving its classification accuracy across all classes and ensuring a more balanced performance. Alternatively, ensemble methods that incorporate multiple models with specialized strengths in different arrhythmia classes could be developed to further optimize classification outcomes. Addressing these limitations and pursuing the proposed future directions will improve our model's practical applicability and contribute meaningfully to advancements in medical signal processing.

VII. CONCLUSION

In conclusion, this study proposes a comprehensive approach to ECG arrhythmia classification by integrating Rough Path Theory into a hybrid deep learning framework. The PathFusion-Net model combines CNN for morphological feature extraction, DevLSTM for sequential modeling, and Path Signatures for structured temporal feature representation, effectively capturing both spatial and temporal patterns in ECG data. Experimental results confirm that integrating PS enhances classification performance, outperforming models without PS or those using only PS methods. Additionally, analysis of different PS orders reveals that while increasing order initially improves performance, excessive complexity can lead to diminishing returns. This study demonstrates the potential of Rough Path Theory in enhancing deep learning models for biomedical applications, offering a robust tool for automated ECG arrhythmia detection. The model's strong performance in an inter-patient setting highlights the advantages of hybrid frameworks in capturing the complex dynamics of ECG signals and reinforces the importance of balancing feature richness with computational efficiency.

Despite these promising results, further research is required to enhance computational efficiency, refine feature selection methodologies, and ensure the model's robustness across diverse datasets. Future studies will focus on extending PathFusion-Net to other biomedical time-series applications, such as EEG and EMG analysis, to evaluate its generalizability

beyond ECG classification. Additionally, improving computational efficiency through model compression techniques, such as pruning and quantization, will be a key objective to facilitate real-time deployment in clinical settings. Addressing these challenges will further solidify PathFusion-Net as a powerful tool for time-series analysis and contribute to more reliable and efficient diagnostic practices in cardiology. Furthermore, the framework could be extended to other biosignals, including PPG [54], enabling broader applications in non-invasive cardiovascular monitoring.

VIII. REFERENCES

- [1] D. J. Hunter and K. S. Reddy, "Noncommunicable diseases," *New England Journal of Medicine*, vol. 369, no. 14, p. 1336–1343, Oct. 2013.
- [2] D. Koppad, "Arrhythmia classification using deep learning: A review," *WSEAS TRANSACTIONS ON BIOLOGY AND BIOMEDICINE*, vol. 18, p. 96–105, Aug. 2021.
- [3] M. Cao, T. Zhao, Y. Li, W. Zhang, P. Benharash, and R. Ramezani, "Ecg heartbeat classification using deep transfer learning with convolutional neural network and stft technique," in *Journal of Physics: Conference Series*, vol. 2547, no. 1. IOP Publishing, 2023, p. 012031.
- [4] A. Rabee and I. Barhumi, "Ecg signal classification using support vector machine based on wavelet multiresolution analysis," in *2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)*. IEEE, 2012, pp. 1319–1323.
- [5] I. Saini, D. Singh, and A. Khosla, "Qrs detection using k-nearest neighbor algorithm (knn) and evaluation on standard ecg databases," *Journal of advanced research*, vol. 4, no. 4, pp. 331–344, 2013.
- [6] Z. Ebrahimi, M. Loni, M. Daneshlatab, and A. Gharehbaghi, "A review on deep learning methods for ecg arrhythmia classification," *Expert Systems with Applications: X*, vol. 7, p. 100033, 2020.
- [7] M. Ashfaq Khan and Y. Kim, "Cardiac arrhythmia disease classification using lstm deep learning approach," *Computers, Materials Continua*, vol. 67, no. 1, p. 427–443, 2021.
- [8] T. J. Lyons, "Differential equations driven by rough signals," *Revista Matemática Iberoamericana*, vol. 14, no. 2, p. 215–310, Aug. 1998.
- [9] T. J. Lyons, M. Caruana, and T. Lévy, *Differential Equations Driven by Rough Paths: École d'Été de Probabilités de Saint-Flour XXXIV - 2004*. Springer Berlin Heidelberg, 2007.
- [10] I. Perez Arribas, G. M. Goodwin, J. R. Geddes, T. Lyons, and K. E. A. Saunders, "A signature-based machine learning model for distinguishing bipolar disorder and borderline personality disorder," *Translational Psychiatry*, vol. 8, no. 1, Dec. 2018.
- [11] Y. Tang, Q. Wu, H. Mao, and L. Guo, "Epileptic seizure detection based on path signature and bi-lstm network with attention mechanism," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, p. 304–313, 2024.
- [12] H. Lou, S. Li, and H. Ni, "Path development network with finite-dimensional lie group representation," 2024.
- [13] P. De Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ecg morphology and heartbeat interval features," *IEEE transactions on biomedical engineering*, vol. 51, no. 7, pp. 1196–1206, 2004.
- [14] F. Melgani and Y. Bazi, "Classification of electrocardiogram signals with support vector machines and particle swarm optimization," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 5, p. 667–677, Sep. 2008.
- [15] J. Park, K. Lee, and K. Kang, "Arrhythmia detection from heartbeat using k-nearest neighbor classifier," in *2013 IEEE International Conference on Bioinformatics and Biomedicine*. IEEE, Dec. 2013, p. 15–22.
- [16] F. Castells, P. Laguna, L. Sörnmo, A. Bollmann, and J. M. Roig, "Principal component analysis in ecg signal processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 1, Feb. 2007.
- [17] S. Faziludeen and P. V. Sabiq, "Ecg beat classification using wavelets and svm," in *2013 IEEE CONFERENCE ON INFORMATION AND COMMUNICATION TECHNOLOGIES*. IEEE, Apr. 2013, p. 815–818.
- [18] W. Zhu, X. Chen, Y. Wang, and L. Wang, "Arrhythmia recognition and classification using ecg morphology and segment feature analysis," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 16, no. 1, p. 131–138, Jan. 2019.
- [19] H. Lassoued and R. Ketata, "Ecg multi-class classification using neural network as machine learning model," in *2018 International Conference on Advanced Systems and Electric Technologies (ICASET)*. IEEE, Mar. 2018, p. 473–478.
- [20] G. Latif, F. Y. Al Anezi, M. Zikria, and J. Alghazo, "Ecg-ecg signals classification for arrhythmia detection using decision trees," in *2020 Fourth International Conference on Inventive Systems and Control (ICISC)*. IEEE, Jan. 2020.
- [21] Z. Zhang, J. Dong, X. Luo, K.-S. Choi, and X. Wu, "Heartbeat classification using disease-specific feature selection," *Computers in Biology and Medicine*, vol. 46, p. 79–89, Mar. 2014.
- [22] F. M. Dias, H. L. Monteiro, T. W. Cabral, R. Naji, M. Kuehni, and E. J. d. S. Luz, "Arrhythmia classification from single-lead ecg signals using the inter-patient paradigm," *Computer Methods and Programs in Biomedicine*, vol. 202, p. 105948, Apr. 2021.
- [23] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436–444, May 2015.
- [24] Z. Liu, X. Meng, J. Cui, Z. Huang, and J. Wu, "Automatic identification of abnormalities in 12-lead ecgs using expert features and convolutional neural networks," in *2018 International Conference on Sensor Networks and Signal Processing (SNSP)*. IEEE, Oct. 2018, p. 163–167.
- [25] J. Huang, B. Chen, B. Yao, and W. He, "Ecg arrhythmia classification using stft-based spectrogram and convolutional neural network," *IEEE Access*, vol. 7, p. 92871–92880, 2019.
- [26] E. K. Wang, X. Zhang, and L. Pan, "Automatic classification of cad ecg signals with sdac and bidirectional long short-term network," *IEEE Access*, vol. 7, p. 182873–182880, 2019.
- [27] Z. Xiong, M. Stiles, and J. Zhao, "Robust ecg signal classification for the detection of atrial fibrillation using novel neural networks," in *2017 Computing in Cardiology Conference (CinC)*, ser. CinC2017. Computing in Cardiology, Sep. 2017.
- [28] D. Li, J. Zhang, Q. Zhang, and X. Wei, "Classification of ecg signals based on 1d convolution neural network," in *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*. IEEE, Oct. 2017, p. 1–6.
- [29] T. F. Romdhane, H. Alhichri, R. Ouni, and M. Atri, "Electrocardiogram heartbeat classification based on a deep convolutional neural network and focal loss," *Computers in Biology and Medicine*, vol. 123, p. 103866, Aug. 2020.
- [30] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. S. Tan, "A deep convolutional neural network model to classify heartbeats," *Computers in Biology and Medicine*, vol. 89, p. 389–396, Oct. 2017.
- [31] T. J. Jun, H. M. Nguyen, D. Kang, D. Kim, D. Kim, and Y.-H. Kim, "Ecg arrhythmia classification using a 2-d convolutional neural network," 2018.
- [32] J. Gao, H. Zhang, P. Lu, and Z. Wang, "An effective lstm recurrent network to detect arrhythmia on imbalanced ecg dataset," *Journal of Healthcare Engineering*, vol. 2019, p. 1–10, Oct. 2019.
- [33] O. Yildirim, "A novel wavelet sequence based on deep bidirectional lstm network model for ecg signal classification," *Computers in Biology and Medicine*, vol. 96, p. 189–202, May 2018.
- [34] B.-H. Kim and J.-Y. Pyun, "Ecg identification for personal authentication using lstm-based deep recurrent neural networks," *Sensors*, vol. 20, no. 11, p. 3069, May 2020.
- [35] R. F. Oliveira, G. J. P. Moreira, V. L. S. Freitas, and E. J. S. Luz, "Leveraging visibility graphs for enhanced arrhythmia classification with graph convolutional networks," 2024.
- [36] B. M. Mathunjwa, Y.-T. Lin, C.-H. Lin, M. F. Abbod, and J.-S. Shieh, "Ecg arrhythmia classification by using a recurrence plot and convolutional neural network," *Biomedical Signal Processing and Control*, vol. 64, p. 102262, Feb. 2021.
- [37] J. Wang, "Automated detection of atrial fibrillation and atrial flutter in ecg signals based on convolutional and improved elman neural network," *Knowledge-Based Systems*, vol. 193, p. 105446, Apr. 2020.
- [38] G. Petmezas, K. Haris, L. Stefanopoulos, V. Kilintzis, A. Tzavelis, J. A. Rogers, A. K. Katsaggelos, and N. Maglaveras, "Automated atrial fibrillation detection using a hybrid cnn-lstm network on imbalanced ecg datasets," *Biomedical Signal Processing and Control*, vol. 63, p. 102194, Jan. 2021.
- [39] Y. Jin, C. Qin, Y. Huang, W. Zhao, and C. Liu, "Multi-domain modeling of atrial fibrillation detection with twin attentional convolutional long short-term memory neural networks," *Knowledge-Based Systems*, vol. 193, p. 105460, Apr. 2020.
- [40] J. Wang, "An intelligent computer-aided approach for atrial fibrillation and atrial flutter signals classification using modified bidirectional lstm network," *Information Sciences*, vol. 574, p. 320–332, Oct. 2021.

- [41] C. Chen, Z. Hua, R. Zhang, G. Liu, and W. Wen, "Automated arrhythmia classification based on a combination network of cnn and lstm," *Biomedical Signal Processing and Control*, vol. 57, p. 101819, Mar. 2020.
- [42] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "Ecg heartbeat classification: A deep transferable representation," in *2018 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE, Jun. 2018.
- [43] H. El-Ghaish and E. Eldele, "Ecgtransform: Empowering adaptive ecg arrhythmia classification framework with bidirectional transformer," *Biomedical Signal Processing and Control*, vol. 89, p. 105714, 2024.
- [44] H. Wu, T. Hu, Y. Liu, H. Zhou, J. Wang, and M. Long, "Timesnet: Temporal 2d-variation modeling for general time series analysis," *arXiv preprint arXiv:2210.02186*, 2022.
- [45] T. Lyons, H. Ni, and H. Oberhauser, "A feature set for streams and an application to high-frequency financial tick data," in *Proceedings of the 2014 International Conference on Big Data Science and Computing*, ser. BigDataScience '14. ACM, Aug. 2014, p. 1–8.
- [46] S. Lai and L. Jin, "Recurrent adaptation networks for online signature verification," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 6, p. 1624–1637, Jun. 2019.
- [47] W. Yang, T. Lyons, H. Ni, C. Schmid, and L. Jin, *Developing the Path Signature Methodology and Its Application to Landmark-Based Human Action Recognition*. Springer International Publishing, 2022, p. 431–464.
- [48] P. Moore, T. Lyons, J. Gallacher, and A. D. N. Initiative, "Using path signatures to predict a diagnosis of alzheimer's disease," *PloS one*, vol. 14, no. 9, p. e0222212, 2019.
- [49] K.-T. Chen, "Integration of paths—a faithful representation of paths by noncommutative formal power series," *Transactions of the American Mathematical Society*, vol. 89, no. 2, pp. 395–407, 1958.
- [50] G. Moody and R. Mark, "The impact of the mit-bih arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, p. 45–50, 2001.
- [51] P. deChazal, M. O'Dwyer, and R. Reilly, "Automatic classification of heartbeats using ecg morphology and heartbeat interval features," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, p. 1196–1206, Jul. 2004.
- [52] "Denoising of ecg signal by wavelet transform," *JOURNAL OF SIGNAL PROCESSING*, vol. 33, no. 8, pp. 1122–1131, 2017.
- [53] T. J. Jun, H. M. Nguyen, D. Kang, D. Kim, D. Kim, and Y.-H. Kim, "Ecg arrhythmia classification using a 2-d convolutional neural network," *arXiv preprint arXiv:1804.06812*, 2018.
- [54] Z. Ebrahimi and B. Gosselin, "Ultralow-power photoplethysmography (ppg) sensors: A methodological review," *IEEE Sensors Journal*, vol. 23, no. 15, pp. 16 467–16 480, 2023.

APPENDIX

A. Intuitive Demonstration of the Path Signature Concept

The goal of this section is to provide an intuitive example that illustrates how path signature features arise from a simple path, in order to make the concept less abstract. The full mathematical definition of path signatures is given in Section Preliminary of Path Signatures of the main text; here we only recall the order 2 expansion needed for demonstration. For a two-dimensional path, the truncated signature up to order 2 takes the form $[1, S^{(1)}, S^{(2)}, S^{(1,1)}, S^{(1,2)}, S^{(2,1)}, S^{(2,2)}]$, where $S^{(1)}$ and $S^{(2)}$ denote first-order increments, diagonal terms $(S^{(1,1)}, S^{(2,2)})$ are proportional to squared increments, and cross terms $(S^{(1,2)}, S^{(2,1)})$ correspond to oriented areas.

To make these quantities concrete, consider the simplified ECG segment in Fig. X, which represents the left portion of a heartbeat leading up to the R-peak. The path is described by the discrete sequence $X_1 = [1, 2, 3, 4, 5]$, $X_2 = [2, 2, 4, 3, 8]$, where X_1 indexes time and X_2 represents the signal amplitude. For this path, the order 2 signature is $[1, 4, 6, 8, 18, 6, 18]$.

First-order terms. The increments are $\Delta X_1 = 4$ and $\Delta X_2 = 6$, so the first-order terms are simply $S^{(1)} = 4$ (temporal span) and $S^{(2)} = 6$ (overall voltage rise).

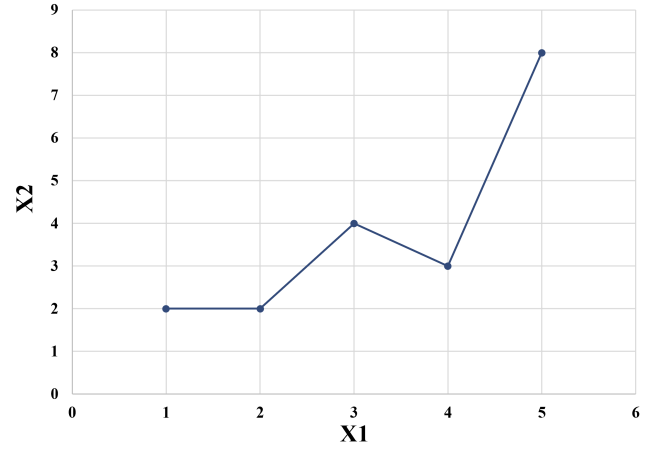


Fig. 6. Path of the trajectory.

Second-order diagonal terms. These capture squared increments: $S^{(1,1)} = \frac{1}{2}(4^2) = 8$ and $S^{(2,2)} = \frac{1}{2}(6^2) = 18$, corresponding to the magnitude of temporal duration and signal growth.

Second-order cross terms. These quantify oriented areas between the path and a piecewise-linear baseline. In this example, the steep ascent towards the peak generates a large positive cross term $S^{(1,2)} = 18$, which geometrically corresponds to the orange shaded region in Fig. 7. The opposite ordering, $S^{(2,1)} = 6$, corresponds to the blue shaded region. Intuitively, $S^{(1,2)}$ measures how strongly amplitude increases relative to time, while $S^{(2,1)}$ reflects the much smaller “vertical-then-horizontal” area.

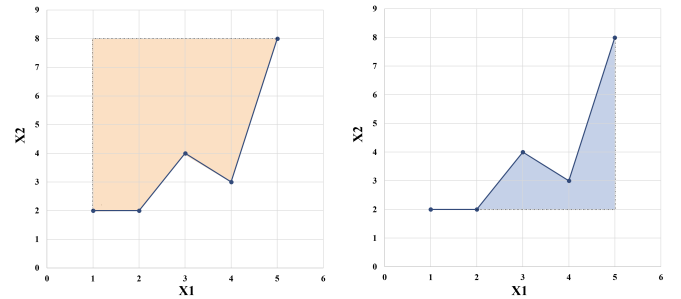


Fig. 7. Geometric interpretation of the cross terms in the signature.

Through this worked example, one can see that path signature terms have direct geometric meaning: first-order terms measure net displacements, diagonal second-order terms correspond to squared increments, and cross terms represent oriented areas that capture the interaction between axes. This demonstration clarifies how the abstract algebraic definitions translate into tangible geometric quantities.