Abnormality Detection in Time-Series Bio-Signals using Kolmogorov-Arnold Networks for Resource-Constrained Devices

Zhaojing Huang[†1], Jiashuo Cui[†1], Leping Yu[1], Luis Fernando Herbozo Contreras[1], Omid Kavehei[1,2]

Abstract

This study uses Kolmogorov-Arnold Networks (KANs) to analyze electrocardiogram (ECG) signals in order to detect cardiac abnormalities. These novel networks have demonstrated potential for application in biosignal analysis, particularly ECG, due to their flexibility and smaller parameter requirements, making them candidates for wearable devices. The network structure comprises a simple KAN model with a single hidden layer of 64 neurons. It was trained on the Telehealth Network of Minas Gerais (TNMG) dataset and tested for generalization on the Chinese Physiological Signal Challenge 2018 (CPSC) dataset. The KAN model delivered reasonably promising results, achieving an F1-score of 0.75 and an AUROC of 0.95 on the TNMG dataset. During the out-of-sample generalization test on the CPSC dataset, it achieved an F1-score of 0.62 and an AUROC of 0.84. It has also shown resistance to missing data channels by maintaining a reasonable performance, down to only a single lead left of ECG data instead of the initial 12 leads. Compared with traditional Multi-Layer Perceptrons (MLP) and Neural Circuit Policy (NCP, aka. Liquid Time Constant Networks), KANs exhibit superior flexibility, adaptability, interpretability, and efficiency. Their compact size and reduced computational requirements make them potential candidates for deployment on hardware, particularly in personalized medical devices.

Keywords: Abnormality Detection, Simple Network, Electrocardiogram, Generalization, Robustness, Kolmogorov-Arnold Networks.

1 Introduction

There are many techniques for monitoring cardiac activity, with Electrocardiography (ECG) being a notable option due to its non-invasive and cost-effective characteristics. The 12-lead ECG is the preferred method for evaluating the heart's electrical activity in clinical settings [1]. Detecting and treating abnormalities is a key function of cardiac care units. Annually, over 3 million ECGs are performed worldwide for patients with various conditions, ranging from heart rhythm anomalies to hormonal imbalances caused by organ failures [2, 3]. ECG signals can also detect seizures [4]. These ECGs are typically analyzed by experienced doctors who diagnose issues based on their knowledge and experience [5]. However, when morphological disturbances in the ECG signal become complex, interpretations can vary significantly among physicians [6]. This variability underscores the need for a system capable of analyzing ECG signals with high accuracy, reducing errors, and enabling timely detection and treatment of issues [7, 8].

Multi-layer Perceptrons (MLPs) are fundamental to deep learning models and have been extensively utilized in applications like ECG signal analysis. Known for their capability to approximate nonlinear functions, MLPs are effective in detecting ECG abnormalities, which is crucial for cardiac diagnosis and monitoring [9, 10, 11, 12]. However, MLPs often have limitations, such as fixed activation functions and less interpretable structures, which can hinder performance improvements and clinical applicability [13].

In response to these limitations, more compact and efficient neural network models such as Neural Circuit Policies (NCPs) have been developed. NCPs which is designed for energy-efficient computation, have demonstrated impressive performance in various tasks, including ECG signal analysis [14]. By integrating principles from neuroscience, NCPs achieve high accuracy with fewer parameters, rendering them suitable for deployment on resource-constrained devices [15]. Additionally, there is potential for merging NCPs with Spiking Neural Networks to enhance power efficiency [16].

Kolmogorov-Arnold Networks (KANs) offer a promising alternative. Inspired by the Kolmogorov-Arnold representation theorem, KANs use learnable activation functions on network edges rather than fixed node functions. This design enhances flexibility, adaptability, and interpretability, which are critical in medical applications where understanding model decisions is essential. Recent studies, such as those by [17], have shown KANs outperforming traditional MLPs in tasks like data fitting and solving partial differential equations, suggesting their effectiveness in optimizing ECG analysis and improving cardiac diagnoses.

This work examines the performance of KANs, comparing them with MLPs and NCPs in the context of ECG signal analysis, focusing on detecting cardiac abnormali-

[1] † The two authors contributed equally to this paper.

[1] Z. Huang, J. Cui, L. Yu, L.F. Herbozo Contreras, and O. Kavehei are with the School of Biomedical Engineering, The University of Sydney, Camperdown, NSW 2006, Australia. (email: {zhaojing.huang, jcui0652}@sydney.edu.au).

[2] O. Kavehei is also with The University of Sydney Nano Institute, Camperdown, NSW 2006, Australia.

Table 1: Comparsion between different networks (∗ ∗ ∗: Superior; ∗∗: Moderate; ∗: Inferior)

Network's characteristic	MLP	NCP	KAN
Learnable activation function	∗	∗	∗∗
High performance	∗	∗∗	∗ ∗ ∗
Reduced parameters	∗	∗∗	∗ ∗ ∗
Interpretability	∗	∗∗	∗ ∗ ∗
Efficiency	∗	∗∗	∗ ∗ ∗
Spline based	∗	∗	∗∗

ties. These comparisons are graphically depicted in Fig. 1 and detailed in Table 1. KANs hold promise for future hardware implementation due to their potential for reduced parameters and efficiency. Their adaptability and efficiency are important features and, hence, candidates for deployment in hardware, providing practical solutions for real-time monitoring and diagnosis in clinical environments. Assessing the architectures below numerous conditions has a look at pursuits to assess the ability of KANs to decorate ECG evaluation in cardiac healthcare. Given the increasing demand for efficient and real-time monitoring solutions, this work envisions the KANs deployment in wearable edge-devices. These devices are designed to swiftly detect anomalies and effectively identify cardiac abnormalities, leveraging KAN's compact architecture and reduced computational complexity.

1.1 Background

In the field of abnormality detection, several notable models have been developed. Petmezas et al. [18] introduced a Hybrid Convolutional Long Short-Term Memory Neural Networks (CNN-LSTM) Network with a sensitivity of 97.87% and a specificity of 99.29% for ECG heartbeat classification. A model by Gupta et al. [19] combined fractional wavelet transform, Yule-Walker autoregressive analysis, and Principal Component Analysis (PCA), achieving a mean square error of less than 0.2%, a detection accuracy of 99.89%, and an output Signal-to-Noise Ratio (SNR) of 25.25 dB. The model proposed by Chen et al. [20] utilized five Convolutional Neural Network (CNN) blocks, a bidirectional Recurrent Neural Network (RNN) layer, an attention layer, and a dense layer, achieving an F1-score of 0.84 in ECG detection. Huang et al. [21] introduced a fast compression residual CNN, achieving an average accuracy of 98.79% for ECG classification. Additionally, a shallow Diagonal State Space Sequence (S4D) model developed by Huang et al. [22] for ECG achieved a robust F1-score of 0.81 and demonstrated high robustness with input data.

Recent studies have also applied NCP models to ECG signal processing. For example, Huang et al. [15] presented two models, ConvLSTM2D-Liquid Time-Constant network (CLTC) and ConvLSTM2D-Closed-Form Continuous-time neural network (CCfC). Both models were evaluated on the ECG datasets, showing impressive generalization and robustness capabilities [15].

KANs have emerged as a novel approach in neural network architecture, inspired by the Kolmogorov-Arnold representation theorem. Unlike traditional MLPs using fixed node functions, KANs utilize learnable activation functions on network edges, enhancing flexibility and interpretability [17]. The potential of KANs in ECG analysis is underscored by their ability to handle complex signal patterns, which is essential for accurate cardiac abnormality detection.

In summary, while MLPs and their variants have been widely used for ECG signal analysis, introducing KANs represents a significant advancement. By leveraging the unique properties of the Kolmogorov-Arnold representation theorem, KANs offer a promising alternative that could enhance the accuracy, interpretability, and efficiency of ECG analysis models, ultimately improving cardiac healthcare outcomes.

2 Prerequisite

2.1 Kolmogorov-Arnold Networks

KANs represent a significant advancement over traditional MLPs by utilizing learnable activation functions on edges rather than fixed activation functions on nodes [17]. This approach allows KANs to achieve higher accuracy and interpretability, making them suitable for various applications, including time series forecasting and ECG signal classification.

2.2 Characteristics of KANs

KANs are based on the Kolmogorov-Arnold representation theorem, which states that any multivariate continuous function can be represented as a finite composition of continuous univariate functions and addition operations [17]. This theorem forms the foundation of KANs, enabling them to replace traditional linear weights with spline-parametrized univariate functions. The key characteristics of KANs include:

- **Learnable Activation Functions:** KANs place learnable univariate activation functions on the network's edges, enhancing the model's flexibility and accuracy.
- **Spline-based Representation:** Using spline functions allows KANs to adapt to the data dynamically, providing more accurate representations.
- **Improved Efficiency and Interpretability:** KANs achieve superior performance with fewer parameters than MLPs, and the learnable functions can be visualized for better interpretability.

The following equation represents the fundamental operation of a KAN layer:

$$\hat{x}_i = \sum_{j=1}^{n} \varphi_{ij}(x_j)$$

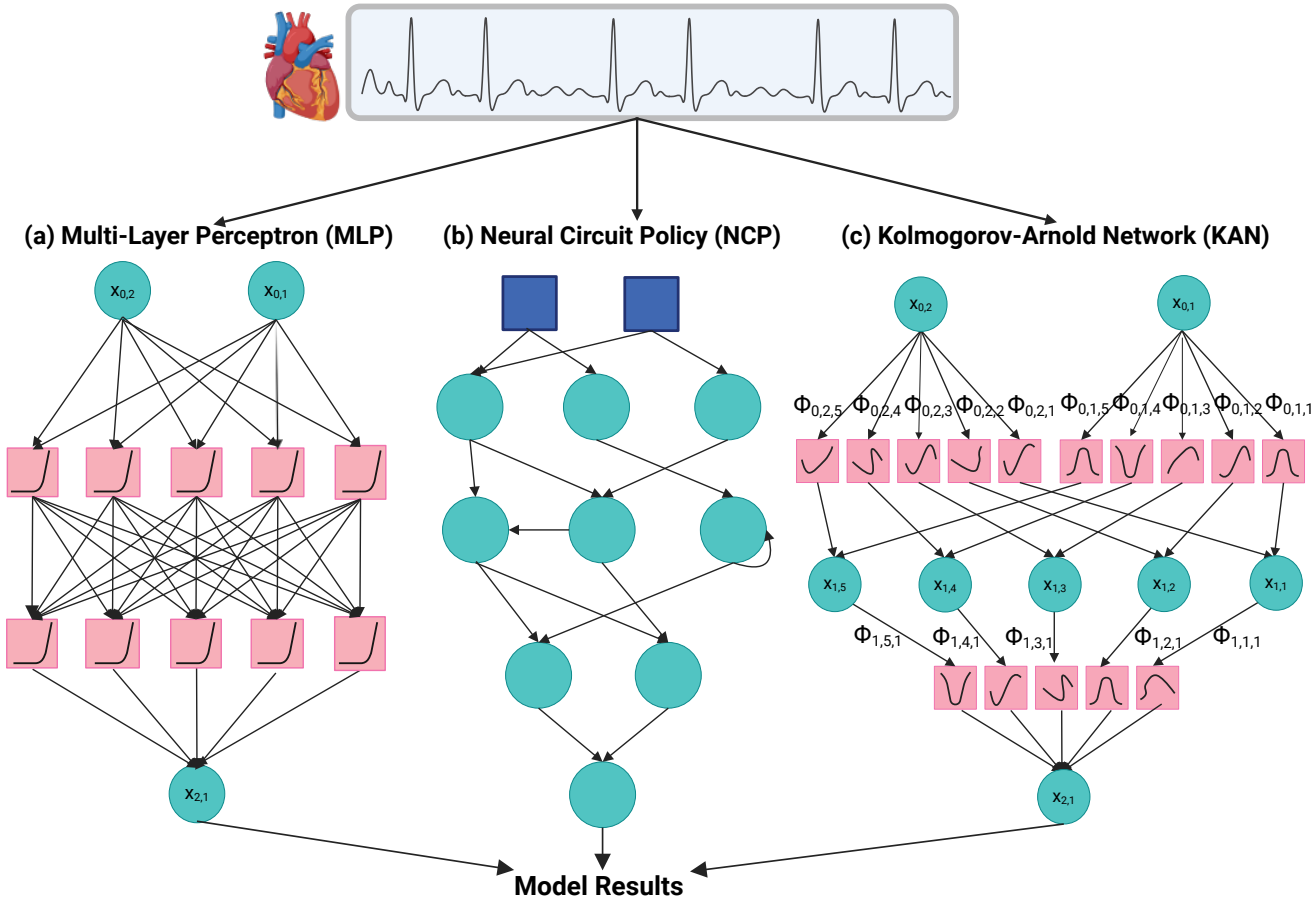

Figure 1: The figure illustrates the distinctions between three types of networks: a traditional Multilayer Perceptron (MLP) with a pre-set activation function (a), a Neural Circuit Policy (NCP) designed to emulate biological neural circuits (b), and a KAN network capable of learning activation functions (c). (a) Represents a standard MLP architecture where the activation functions are predetermined and fixed. (b) Depicts an NCP model that mimics biological neural circuits' modularity and information flow, emphasizing adaptability and task-specific decision-making. (c) Shows a KAN network with the unique ability to learn and adjust its activation functions based on the input data and learning objectives, enhancing its flexibility and performance across diverse tasks.

where $\hat{x}_i$ is the activation value at node i and φ_{ij} is the learnable activation function on the edge from node j to node i.

2.3 Mathematical Formulation of KANs

KANs leverage the Kolmogorov-Arnold representation theorem to decompose a high-dimensional function into a sum of univariate functions. This decomposition can be expressed as:

$$f(x) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{p=1}^{n} \phi_{q,p}(x_p) \right)$$

Here, Φ_q and $\phi_{q,p}$ are univariate functions parameterized as splines, and x_p represents the input features. This formulation allows KANs to effectively capture both the compositional structure and univariate functions, providing a powerful framework for function approximation.

2.4 Application of KANs to ECG Signal Detection

In this paper, to demonstrate the significance of KANs' application in bio-signal efficient and effective processing, we applied it to ECG signal abnormality classification using two separate datasets (out of distribution) of 12-lead ECGs (see Sec. 3). These two datasets, one with eight classes and another with six, have four overlapping classes on which we performed our out-of-sample tests. KANs can achieve higher accuracy in detecting cardiac abnormalities due to their ability to dynamically learn and adapt activation functions. This flexibility allows KANs to model ECG signals more effectively. Compared to traditional MLPs, this novel network architecture requires fewer parameters to perform similarly or better. This efficiency makes these networks suitable for applications where computational resources are limited. The learnable spline-based activation functions provide a more interpretable model. Clinicians can visualize these functions to understand how the model processes ECG

signals. KANs are particularly adept at capturing nonlinear relationships in ECG data, which are common in physiological signals. This capability leads to more robust and accurate classification.

3 Datasets

In this research, two independent datasets were used for out-of-distribution evaluation of the proposed models. The first was the Chinese Physiological Signal Challenge 2018 (CPSC) dataset [23], designed for automated detection of rhythm and morphology irregularities in 12-lead ECGs. The second dataset came from the Telehealth Network of Minas Gerais (TNMG) [24] and was used mainly for training. The CPSC dataset served as an independent test set to evaluate the models on unseen data, advancing our goal of prospective real-world and clinical analysis, which remains an ongoing research focus in our group.

The primary goal was to assess the models' generalization and effectiveness on unfamiliar data. The evaluation began with training on the TNMG dataset and a performance assessment on the CPSC dataset. These outcomes are crucial for determining the models' reliability and efficacy in practical applications.

3.1 TNMG Dataset

This study utilized the TNMG dataset comprising 2,322,513 annotated 12-lead ECG samples. This extensive dataset encompasses six distinct categories of cardiac irregularities, namely Atrial Fibrillation (AF), Left Bundle Branch Block (LBBB), First Degree Atrioventricular Block (1dAVb), Right Bundle Branch Block (RBBB), Sinus Tachycardia (ST), and Sinus Bradycardia (SB) [24]. The ECG recordings were obtained at a sampling rate of 400 Hz.

A systematic sampling strategy was adopted to create a balanced dataset for model training. Specifically, 3,000 samples were randomly chosen for each of the six abnormalities and an additional 3,000 normal samples, yielding a subset of 21,000 instances. Any additional samples required to complete 21,000 records were randomly drawn from the TNMG dataset for patients exhibiting multiple abnormalities. For more detailed information on these abnormalities, see Table 2. The dataset was normalized to standardize the length to 4096 readings, ensuring consistent analysis and modeling. Readings exceeding this length were excluded, simplifying data processing and enabling meaningful comparisons.

In our previous work, we provided a detailed illustration of the gender and age distribution within the subset, promoting inclusiveness and ensuring valid analyses [15]. Additionally, the subset aligns with the general population's age distribution, enhancing its representativeness for age-related studies. The sampling methodology achieved a balanced distribution of abnormalities, allowing for a comprehensive examination of their characteristics and impacts. This balance improves the model's learning process and overall performance [31].

3.2 CPSC Dataset

The CPSC dataset comprises 12-lead ECGs initially sampled at 500 Hz. To ensure compatibility between the CPSC and TNMG datasets, the TNMG data is resampled to 500 Hz. This dataset includes ECGs from patients with various cardiovascular conditions. Each ECG is labeled by experts, providing precise annotations for detected abnormalities across eight distinct categories.

This study assessed the model's generalization capabilities using four specific abnormalities from the CPSC dataset: AF, RBBB, 1dAVb, and LBBB. The additional four abnormalities—Premature Atrial Contraction (PAC), ST-segment Depression (STD), Premature Ventricular Contraction (PVC), and ST-segment Elevation (STE)—were not considered because only four abnormalities overlap with those in the TNMG dataset.

During the data curation, entries containing missing readings were filtered out, resulting in a final dataset of 6,877 distinct ECG recordings. Standardization was applied to ensure a consistent length of 4,096 readings, with any surplus readings removed during the data refinement phase. For a detailed dataset overview, please refer to our previous work [15].

Upon analysis, a noticeable gender disproportion was observed, with a higher representation of male patients than female patients. However, the age distribution closely mirrors the general population's, showcasing many patients in older age brackets. Furthermore, a slight unevenness was observed in the distribution of abnormalities, with LBBB occurrences being less frequent than other abnormalities within the dataset.

4 Methods

Our primary goal is to assess the effectiveness of KANs in processing ECG data, aiming to develop a hardware-friendly solution. To achieve this, we engineered a compact architecture that markedly reduces computational demands while upholding accuracy. Following the architecture's development, we fed data into the model for training and validation. Then, we assessed its performance in terms of generalizability to new, unseen data and its robustness. Furthermore, we explored the network's capabilities in handling single-lead ECG data, inching closer to potential wearable applications.

4.1 Data Preprocessing

Several preprocessing steps are necessary to ensure signal quality and consistency before applying the Short-Time Fourier Transform (STFT) to ECG data. These steps include bandpass filtering, wavelet denoising, and normalization.

Table 2: Abnormalities in the TNMG dataset

Type	Description
LBBB +	Left ventricle is activated after the right, affecting perfusion, mechanical function, and workload within the left ventricle [25].
AF ⋄	The most common arrhythmia, characterized by an irregular and rapid heartbeat [26].
SB †	Slower-than-normal heart rate due to decreased electrical signals from the sinus node [27].
ST ‡	Elevated resting heart rate and an increased heart rate response to minor physical activities or changes in posture, indicative of tachyarrhythmia [28].
1dAVb *	ECG's PR interval exceeding 200 ms, indicating a first-degree atrioventricular block [29].
RBBB §	The condition disrupts the heart's normal electrical activity, delaying right ventricle depolarization due to interrupted signal transmission in the His-Purkinje system [30].

+ Left Bundle Branch Block (LBBB); ⋄ Atrial Fibrillation (AF); † Sinus Bradycardia (SB); ‡ Sinus Tachycardia (ST); ∗ First Degree Atrioventricular Block (1dAVb); § Right Bundle Branch Block (RBBB)

By using bandpass filtering between 0.5 Hz and 40 Hz, removes noise outside the desired frequency range, enhancing signal quality by eliminating low-frequency drift and high-frequency noise. Wavelet denoising further cleans the ECG signals by decomposing them into different frequency components and reconstructing them while removing noise. Finally, min-max normalization scales the ECG data to a specified range, ensuring equal contribution of all features and preventing any single feature from dominating the analysis. Min-max normalization is believed to enhance the model's generalizability [32].

The STFT is a powerful tool for analyzing non-stationary signals like ECG data, providing time-frequency representations that enable the detection of transient features that might be missed by traditional Fourier Transform methods [33, 34, 35]. By dividing the signal into overlapping segments and applying the Fourier Transform to each segment, STFT creates a two-dimensional representation with time and frequency dimensions. The window size and overlap choice are crucial for resolution and accuracy [36].

STFT has been used effectively in various studies to enhance ECG abnormality detection and classification. For example, Acharya et al. [37] used STFT to detect arrhythmias with improved performance compared to traditional methods. Similarly, De Chazal and Reilly [38] used STFT for automated ECG beat classification, achieving high accuracy in distinguishing between normal and abnormal heartbeats.

4.2 Simple Model Architecture

With our focus on understanding the effectiveness of KANs, we aimed to keep the network size small. In this study, we designed and evaluated four distinct model architectures to assess their efficacy in detecting abnormalities in ECG signals. These architectures were designed with varying levels of complexity to investigate the performance trade-offs linked to different network depths and sizes. The models included:

- A single hidden layer with 64 neurons.
- A two hidden layers with 32 neurons per layer.
- A four hidden layer with 16 neurons per layer.
- A single hidden layer with 128 neurons.

Our primary aim was to utilize a network featuring an intermediate architecture of 64 neurons. We will then identify the best-performing networks within this architecture.

As illustrated in Fig. 2, all the proposed architectures are built upon the foundational principles of KANs, employing learnable activation functions on network edges rather than fixed node functions. This design choice amplifies flexibility, adaptability, and interpretability, qualities crucial for medical applications where comprehending model decisions is paramount. The training process for these models entailed using ECG data for cardiac abnormality detection. These models underwent training to maximize performance and minimize losses. Their performance will be assessed using various metrics to understand the model's overall performance comprehensively.

4.3 Evaluation Metrics

In evaluating model performance, key metrics such as precision, recall, and their combined measure, the F1-score, play pivotal roles. Precision gauges the accuracy of positive predictions, while recall indicates the proportion of actual positives correctly identified, offering insights into the model's performance trade-offs.

When dealing with binary classification, metrics like the AUROC (Area Under the Receiver Operating Characteristic curve) measure a model's efficacy in determining negative and positive cases through various threshold levels. Conversely, the AUPRC (Area Under the Precision-Recall Curve) emphasizes the model's accuracy in pinpointing positive cases, which is particularly useful in handling imbalanced datasets.

Metrics such as F1-score, precision, recall, AUROC, and AUPRC are crucial for evaluating machine learning model performance, especially in tasks like abnormality detection.

5 Experiment

5.1 The Training Process

In the experimental phase of our study, we utilized the TNMG subset data as the primary training dataset for

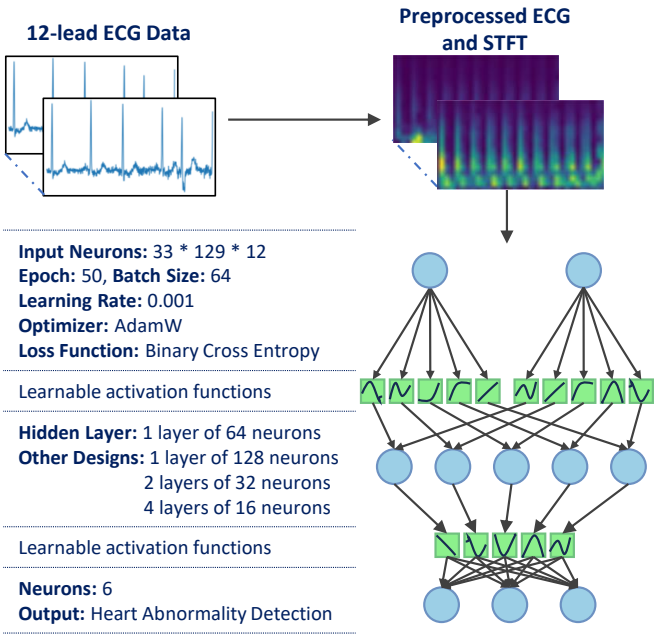

Figure 2: The model architecture consists of KAN neurons tailored for STFT-processed ECG data. The input layer comprises $33 \times 129 \times 12$ neurons, matching the STFT pre-processed ECG data points. It then passes through intermediate layers with 64 neurons, with additional intermediate architectures also tested. Finally, six output neurons are corresponding to the six abnormalities being detected.

our KAN models. This dataset provided diverse ECG recordings encompassing various cardiac abnormalities. The models underwent training for 50 epochs with a batch size of 64. We utilized the AdamW optimizer with a weight decay of 10^{-4} and employed binary cross-entropy loss. A scheduler (ExponentialLR) was also implemented with a gamma of 0.9.

While the dataset seems balanced overall, there remains an imbalance between total positive and negative cases due to the multi-labeled nature of abnormalities. This means a person can have more than one abnormality. To rectify this skewness and enhance the model's ability to detect positive cases, a technique called repetition-of-4 was implemented. This approach involves replicating all data instances with positive cases four times within the training data. This technique has improved the model's performance in detecting positive cases [15].

5.2 In-sample Training and Validation Procedures

For assessing model performance, we set aside a dedicated dataset comprising 20% of the TNMG subset for validation.

These data points were intentionally withheld from the model's training phase and reserved solely for performance evaluation. This separate validation approach directly compared the models' individual performances.

5.3 Evaluation on Unseen Data

We evaluated our model's performance using the CPSC dataset, which closely represents real-world scenarios. This assessment involved comparing the models' predictions against data labels using a variety of performance metrics. This in-depth analysis will help us grasp the models' strengths and limitations, providing valuable insights into their effectiveness. Additionally, these findings will form the basis for identifying potential areas to enhance ECG analysis. We specifically selected the four abnormalities in both CPSC and TNMG for comparison.

5.4 Model Robustness

In assessing our model's resilience, we'll conduct a series of deliberate tests involving the removal of channels from the 12-lead ECG data. This evaluation targets the model's performance when dealing with missing inputs, especially within the KAN networks' framework. The tests will be methodically carried out, gradually removing varying numbers of channels, spanning from 1 to 6 leads. These experiments aim to understand the model's capacity to sustain accuracy across a spectrum of scenarios. We will utilize a range of performance metrics to assess its performance objectively. These metrics will facilitate a thorough evaluation and comparison of the model's efficacy under different conditions. The findings from these assessments will help identify areas for improvement and offer direction for future enhancements.

5.5 Model for Single-Lead ECG

In addition to the evaluations conducted with the 12-lead ECG data, we designed an experiment to assess the model's performance using data with only a single lead. Specifically, we utilized Lead II data to train the KAN model. This experiment aimed to investigate whether the model could maintain its performance when trained with data of just one lead, thereby evaluating its potential for use in scenarios where only single-lead ECG data is available. The training procedure, hyperparameters, and validation methods remained consistent with those used in the 12-lead ECG experiments to ensure a fair comparison. The only change we made was reducing the input layer from $33 \times 129 \times 12$ to $33 \times 129 \times 1$, reflecting the reduction in input data.

6 Results

In this section, we will comprehensively analyze the models' performance, specifically focusing on utilizing KAN models for training and evaluation. Our discussion will explore the models' generalizability capabilities and robustness. Additionally, we will showcase the model's potential for single-lead ECG data analysis.

6.1 Training with Different Architecture

As we seek to understand the effectiveness of KAN neurons and explore their potential for use in hardware applications that demand compact network sizes and fewer parameters, we are conducting a study focused on small networks. Accordingly, we have designed a network comprising 64 neurons evenly distributed across 1, 2, and 4 hidden layers. This design encompasses a model with a single intermediate layer of 64 neurons, two intermediate layers of 32 neurons each, and four intermediate layers of 16. We trained these models using 50 epochs, and Fig. 3 illustrates the training process of these models.

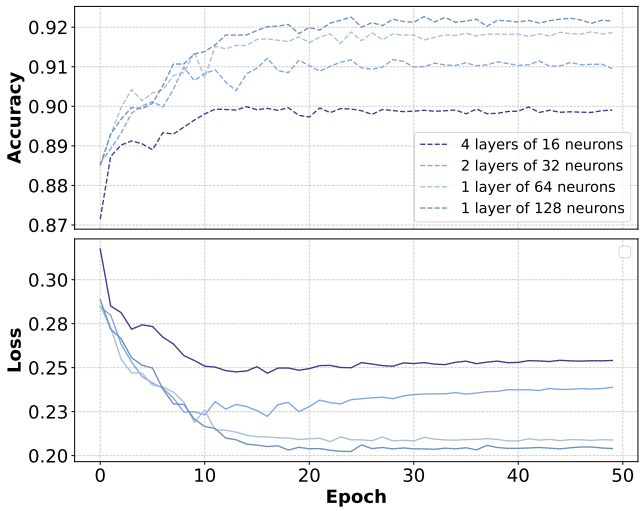

Figure 3: The training processes of various KAN architectures show increasing accuracy and decreasing loss throughout. However, discrepancies can be observed among different network architectures.

We can observe that with the same number of neurons, the model trains more effectively with a single layer of 64 neurons, as evidenced by the lower stabilization of loss and higher accuracy. For completeness, we also tested a model with an intermediate layer of 128 neurons, and it's noticeable that as the number of neurons increases, the model's performance improves.

While a more complex network yields better performance, we opted for the model with one intermediate layer of 64 neurons for further analysis. Our focus is on understanding the effectiveness of these neurons in ECG applications, with an eye toward future hardware implementations.

6.2 In-Sample Performance Evaluation

This research study focuses on training the proposed KAN model using the TNMG subset dataset, as detailed earlier in this paper. The TNMG subset dataset serves as the primary resource for developing and evaluating the models in this investigation. The model's validation results on in-sample pre-processed data are summarized in Table 3. The model exhibits robust performance, achieving an F1-score of 0.75. Moreover, the AUROC value stands at 0.95, with an AUPRC value of 0.84.

These results highlight that even a single hidden layer KAN network model performs robustly, indicating significant potential for ECG analysis tasks. The observed performance metrics underscore the importance of data pre-processing methods tailored to the specific application.

Table 3: In-Sample Validation Results on TNMG dataset (R: Recall, P: Precision)

Class	R	P	F1	AUROC	AUPRC
1dAVb	41.3%	52.6%	46.3%	80.6%	46.6%
RBBB	77.8%	83.9%	80.7%	95.1%	88.8%
LBBB	87.3%	91.2%	89.2%	98.8%	95.7%
SB	87.7%	88.6%	88.2%	98.7%	94.4%
AF	41.6%	54.1%	47.1%	81.8%	50.8%
ST	86.1%	88.3%	87.2%	98.3%	93.4%
Average	71.2%	78.8%	74.8%	94.6%	84.1%

6.3 Model Generalization

Our evaluation strategy involved applying our trained models from the TNMG subset to make predictions on the CPSC dataset, as outlined in the data methodology. This assessment was designed to ascertain how effectively our models could handle new data instances from the CPSC dataset, which may possess inherent differences compared to the training dataset. Notably, the CPSC dataset encompasses eight distinct types of abnormalities, with only four overlapping with those in the TNMG dataset.

Analyzing the KAN model's performance on the CPSC dataset provides valuable insights into its ability to transfer learned knowledge to new data domains. This examination is a critical benchmark for assessing the model's generalization capabilities beyond its original training data, demonstrating its adaptability to a broader range of clinical scenarios.

Table 4 outlines the model's generalization performance, achieving an F1-score of 0.62, an AUROC of 0.84, and an AUPRC of 0.61. While a decrease in performance is noticeable, the drop remains within an acceptable range based on our experience. Considering the simplicity of the architecture we've introduced, there is significant potential for further improvement with advancements in research in this novel network. These results highlight the potential of the KAN network model in ECG analysis, indicating its viability for real-world applications. Given that this network comprises solely KAN neurons, there is considerable room for enhancing the performance of this model with future research into the network.

6.4 Model Robustness

To evaluate the model's efficacy, we broadened our analysis to incorporate the CPSC dataset, deliberately introducing systematic variability. Furthermore, we performed supplementary evaluations by randomly omitting

Table 4: Out-of-Distribution Generalization Results on CPSC dataset (R: Recall, P: Precision)

Class	R	P	F1	AUROC	AUPRC
1dAVb	41.7%	34.0%	37.5%	78.1%	34.0%
RBBB	54.1%	85.2%	66.2%	80.3%	74.5%
LBBB	78.7%	86.0%	82.2%	97.3%	86.2%
AF	59.3%	42.4%	49.5%	78.9%	48.3%
Average	67.3%	67.4%	62.2%	83.6%	60.7%

a defined number of channels from the 12-lead ECG data. This deliberate alteration of input data enabled us to assess the model's robustness and its aptitude to manage incomplete or absent input information proficiently. These experimental maneuvers significantly enhance our comprehension of the model's adaptability and resilience in practical scenarios where data integrity might be compromised.

In Fig. 4, it can be observed that as the number of emptied leads increases, the F1 performance metric of the models declines. However, the decline is insignificant, which may be partly attributed to the min-max normalization, which is known to enhance resistance to missing channels [32].

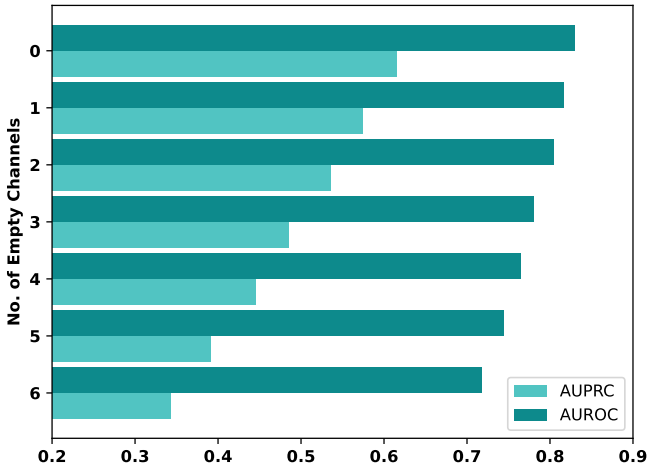

Figure 4: Assessing model robustness using AUROC and AUPRC metrics across different levels of missing channels in CPSC ECG data.

6.5 Exploring Model Performance with Single-Lead ECG Data

As part of our investigation into the model's capabilities, we conducted an experiment utilizing the TNMG dataset. Interestingly, for this specific experiment, we focused solely on Lead II data to train the KAN model. The detailed outcomes of this experiment can be found in Table 5.

Contrasting the outcomes outlined in Table 3, originating from a model trained on comprehensive 12-lead ECG data, with the results of a model exclusively trained on Lead II data reveals intriguing observations. While there's a marginal dip in the overall performance of the latter model across various abnormalities, it's remarkable that such proficiency was achieved using data from only one lead. Despite this limitation, the model's performance remains robust.

Table 5: In-sample Validation of a Model Trained on Single-lead ECG Data (R: Recall, P: Precision).

Class	R	P	F1	AUROC	AUPRC
1dAVb	41.0%	58.1%	48.0%	86.1%	55.1%
RBBB	61.4%	69.2%	65.1%	87.7%	70.8%
LBBB	67.7%	79.1%	72.9%	94.8%	83.1%
SB	88.0%	86.2%	87.1%	98.5%	93.2%
AF	36.3%	50.9%	42.4%	81.4%	46.3%
ST	85.7%	89.3%	87.5%	98.1%	93.4%
Average	63.6%	73.9%	68.3%	92.7%	77.8%

7 Discussion

This study represents the initial successful application of KAN to ECG signal analysis, particularly focusing on multi-abnormality detection. The results demonstrate promising and satisfactory outcomes. The proposed KAN models have shown strong performance and robust generalization capabilities, effectively handling situations involving incomplete or missing input data. Even a single hidden layer KAN network, with only 64 nodes in its hidden layer, exhibited excellent performance. This highlights the flexibility and adaptability of KAN networks, even in extremely simple structures.

Compared to traditional MLPs, KANs offer significant advantages. MLPs, while powerful, often struggle with fixed activation functions and less interpretable structures, which can hinder performance and clinical applicability [13]. In contrast, KANs leverage the Kolmogorov-Arnold representation theorem to implement learnable activation functions on the network's edges, enhancing flexibility and interpretability. Additionally, the study compared KANs to more recent and compact models such as NCP models. NCPs, a bio-inspired neural network, have demonstrated significant success in efficient and robust ECG abnormality detection [15]. Despite these advanced capabilities, the simple single hidden layer KAN network with 64 neurons in its hidden layer performed comparably, underscoring its potential as a highly adaptable and efficient model for ECG analysis.

A notable characteristic of KANs is their efficiency. The model's compact size and quick training make it well-suited for potential deployment on hardware, especially in personalized medical devices. This efficiency supports real-time data processing and reduces power consumption, a critical factor for continuous monitoring applications.

In our evaluation, we compared the single hidden layer KANs with NCPs [15], using metrics such as precision, recall, F1-score, and AUROC. The results are detailed in Table 6.

While the KANs model may lag in performance compared to other efficient neural networks like NCPs, their

Table 6: Validation and Generalization Results (R: Recall, P: Precision)

Model	P	R	F1	AUROC
Validation Results				
CLTC	85.7%	80.3%	82.7%	96.1%
CCfC	85.9%	80.6%	82.8%	96.3%
KAN	71.2%	78.8%	74.8%	94.6%
Generalization Results				
CLTC	69.8%	75.4%	69.9%	90.1%
CCfC	75.5%	73.9%	72.1%	91.3%
KAN	67.3%	67.4%	62.2%	83.6%

architecture has fundamental differences. The CCfC and CLTC models incorporate a layer of ConvLSTM2D to aid in the feature extraction process, whereas the proposed KANs model does not include additional layers for this purpose. We also trained an MLP model with the same architecture as the proposed KAN, but it did not yield significant results. This underscores the power and efficiency of the neurons in the KAN model, which sheds light on future applications in hardware compatibility when further optimization is done in the newly introduced KAN.

8 Limitations and Future Work

8.1 Limitations

Despite the promising results demonstrated by KANs in ECG signal analysis, several limitations need to be addressed:

- **Novelty and Integration:** KANs are a relatively new architecture and have not yet been widely tested or integrated with other advanced neural network layers such as CNNs or RNNs. This limits the current understanding of how KANs can be effectively combined with other deep-learning models to enhance performance further.
- **Code Optimization:** The existing implementations of KANs are still in the early stages and may not be fully optimized for large-scale deployment. The current codebase might lack the efficiency and robustness required for real-world applications, particularly in resource-constrained environments like wearable devices.

8.2 Future Work

Future work should focus on addressing these limitations by:

- **Integration with Other Architectures:** Exploring the integration of KANs with other neural network architectures to create more robust and versatile models.
- **Code and Algorithm Optimization:** Enhancing the performance and reducing the computational complexity of KANs by optimizing their code and algorithms.
- **Deployment on Devices:** Deploying KANs on portable and wearable health monitoring systems, which requires further minimization of power requirements and ensuring sustainable and efficient operation in real-time applications.

9 Conclusion

In this paper, we have explored the performance of MLPs, NCPs, and KANs in ECG signal analysis, focusing on detecting cardiac abnormalities. KANs, inspired by the Kolmogorov-Arnold representation theorem, offer a novel approach with learnable activation functions on network edges, enhancing flexibility and interpretability. Our findings highlight the potential of KANs to revolutionize ECG analysis in cardiac healthcare, promising significant advancements in accuracy, interpretability, and efficiency. The results show that the KAN model exhibits strong in-sample performance while maintaining good generalization in out-of-distribution scenarios. The model also shows robustness to incomplete data, which is common in real-world scenarios. When tested with single-lead ECG data instead of 12 leads, the model's performance remains robust, confirming its suitability for wearable device applications.
While KANs have shown promising results, their practical application and integration with other neural network layers need further investigation. Future research should focus on optimizing KANs for deployment on portable and wearable devices, ensuring they are efficient, robust, and capable of real-time data processing. These advancements could ultimately contribute to better patient outcomes and more effective cardiac healthcare solutions.

10 Acknowledgement

Zhaojing Huang acknowledges the Australian Government's Research Training Program (RTP) support. The authors acknowledge the support from the Australian Research Council under Project DP230100019.

11 Availability of Data

This paper uses the Chinese Physiological Signal Challenge 2018 (CPSC) dataset for analysis and experiments, which is openly available. The Telehealth Network of Minas Gerais (TNMG) dataset is not openly accessible. Access and permission to this data should be discussed with and granted by the custodian of the data.

12 Conflict of Interest Statement

The authors declare no conflicts of interest.

References

[1] Steve Meek and Francis Morris. ABC of clinical electrocardiography: Introduction. I—Leads, rate, rhythm, and cardiac axis. *BMJ: British Medical Journal*, 324(7334):415, 2002.

[2] Sidney C. Smith et al. Global variation of mortality from ischaemic heart disease and stroke: the world health organization MONICA project. *Bulletin of the World Health Organization*, 93(3):99–108, 2015.

[3] Joseph S. Alpert et al. Cardiology patient pages. the importance of cardiovascular disease in women. *Journal of the American College of Cardiology*, 36 (2):423–425, 2000.

[4] Yikai Yang, Nhan Duy Truong, Jason K Eshraghian, Christina Maher, Armin Nikpour, and Omid Kavehei. A multimodal AI system for out-of-distribution generalization of seizure identification. *IEEE Journal of Biomedical and Health Informatics*, 26(7): 3529–3538, 2022.

[5] Saeid Abolfazli et al. A survey on mobile cloud computing. *Journal of Network and Computer Applications*, 40:83–100, 2014.

[6] Sheldon Greenfield et al. Variability of interpretation of the electrocardiogram. *Journal of Chronic Diseases*, 32(9-10):747–751, 1979.

[7] Robert L. Lux et al. Accuracy of automated ECG analysis: a comparison of real-time with off-line systems. *Journal of Electrocardiology*, 21:S114–S118, 1988.

[8] George B. Moody et al. Impact of the MIT-BIH arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*, 20(3):45–50, 2001.

[9] Yann LeCun et al. Deep learning. *Nature*, 521 (7553):436–444, 2015.

[10] Jiaxin Wang et al. Deep learning for ECG data analysis: insights and applications. *IEEE Journal of Biomedical and Health Informatics*, 24(7):1920–1930, 2020.

[11] Kurt Hornik et al. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5): 359–366, 1989.

[12] Zachi I. Attia et al. An artificial intelligence-enabled ecg algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *The Lancet*, 394 (10201):861–867, 2019.

[13] Qi Zhang et al. Interpretability of deep learning models for image classification. *IEEE Signal Processing Magazine*, 35(1):44–52, 2019.

[14] Mathias Lechner, Ramin Hasani, Alexander Amini, Thomas A Henzinger, Daniela Rus, and Radu Grosu. Neural circuit policies enabling auditable autonomy. *Nature Machine Intelligence*, 2(10):642–652, 2020.

[15] Zhaojing Huang, Luis Fernando Herbozo Contreras, Wing Hang Leung, Leping Yu, Nhan Duy Truong, Armin Nikpour, and Omid Kavehei. Efficient edge-AI models for robust ECG abnormality detection on resource-constrained hardware. *Journal of Cardiovascular Translational Research*, pages 1–14, 2024.

[16] Zhaojing Huang, Wing Hang Leung, Leping Yu, Luis Fernando Herbozo Contreras, Ziyao Zhang, Nhan Duy Truong, Armin Nikpour, and Omid Kavehei. On-device edge-learning for cardiac abnormality detection using a bio-inspired and spiking shallow network. *APL Machine Learning*, 2(2), 2024.

[17] Ziming Liu, Yixuan Wang, Sachin Vaidya, Fabian Ruehle, James Halverson, Marin Soljačić, Thomas Y. Hou, and Max Tegmark. KAN: Kolmogorov-arnold networks, 2024.

[18] Georgios Petmezas, Kostas Haris, Leandros Stefanopoulos, Vassilis Kilintzis, Andreas Tzavelis, John A Rogers, Aggelos K Katsaggelos, and Nicos Maglaveras. Automated atrial fibrillation detection using a hybrid CNN-LSTM network on imbalanced ECG datasets. *Biomedical Signal Processing and Control*, 63:102194, 2021. ISSN 1746-8094.

[19] Varun Gupta, Monika Mittal, and Vikas Mittal. A novel FrWT based arrhythmia detection in ECG signal using YWARA and PCA. *Wireless Personal Communications*, pages 1–18, 2022.

[20] Tsai-Min Chen, Chih-Han Huang, Edward SC Shih, Yu-Feng Hu, and Ming-Jing Hwang. Detection and classification of cardiac arrhythmias by a challenge-best deep learning neural network model. *iScience*, 23(3), 2020.

[21] Jing-Shan Huang, Bin-Qiang Chen, Nian-Yin Zeng, Xin-Cheng Cao, and Yang Li. Accurate classification of ECG arrhythmia using mowpt enhanced fast compression deep learning networks. *Journal of Ambient Intelligence and Humanized Computing*, pages 1–18, 2020.

[22] Zhaojing Huang, Luis Fernando Herbozo Contreras, Leping Yu, Nhan Duy Truong, Armin Nikpour, and Omid Kavehei. S4D-ECG: A shallow state-of-the-art model for cardiac abnormality classification. *Cardiovascular Engineering and Technology*, pages 1–12, 2024.

[23] Feifei Liu, Chengyu Liu, Lina Zhao, Xiangyu Zhang, Xiaoling Wu, Xiaoyan Xu, Yulin Liu, Caiyun Ma, Shoushui Wei, Zhiqiang He, et al. An open access

database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. *Journal of Medical Imaging and Health Informatics*, 8(7):1368–1373, 2018.

[24] Antônio H Ribeiro, Manoel Horta Ribeiro, Gabriela MM Paixão, Derick M Oliveira, Paulo R Gomes, Jéssica A Canazart, Milton PS Ferreira, Carl R Andersson, Peter W Macfarlane, Wagner Meira Jr, et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nature Communications*, 11(1):1760, 2020.

[25] Andrés R Pérez-Riera, Raimundo Barbosa-Barros, Marianne PC de Rezende Barbosa, Rodrigo Daminello-Raimundo, Luiz C de Abreu, and Kjell Nikus. Left bundle branch block: Epidemiology, etiology, anatomic features, electrovectorcardiography, and classification proposal. *Annals of Noninvasive Electrocardiology*, 24(2):e12572, 2019.

[26] Zengwu Wang, Zuo Chen, Xin Wang, Linfeng Zhang, Suning Li, Ye Tian, Lan Shao, Hao Hu, Runlin Gao, et al. The disease burden of atrial fibrillation in China from a national cross-sectional survey. *The American Journal of Cardiology*, 122(5):793–798, 2018.

[27] Shuenn-Nan Chiu, Lian-Yu Lin, Jou-Kou Wang, Chun-Wei Lu, Chi-Wei Chang, Ming-Tai Lin, Yu Chuan Hua, Hung-Chi Lue, and Mei-Hwan Wu. Long-term outcomes of pediatric sinus bradycardia. *The Journal of Pediatrics*, 163(3):885–889, 2013.

[28] Aino-Maija Still, Pekka Raatikainen, Antti Ylitalo, Heikki Kauma, Markku Ikäheimo, Y Antero Kesäniemi, and Heikki V Huikuri. Prevalence, characteristics and natural course of inappropriate sinus tachycardia. *EP Europace*, 7(2):104–112, 2005.

[29] Theodora Nikolaidou, Justin M Ghosh, and Andrew L Clark. Outcomes related to first-degree atrioventricular block and therapeutic implications in patients with heart failure. *JACC: Clinical Electrophysiology*, 2(2):181–192, 2016.

[30] Yunyun Xiong, Lian Wang, Wenyan Liu, Graeme J Hankey, Biao Xu, and Shang Wang. The prognostic significance of right bundle branch block: A meta-analysis of prospective cohort studies. *Clinical Cardiology*, 38(10):604–613, 2015.

[31] Zhaojing Huang, Sarisha MacLachlan, Leping Yu, Luis Fernando Herbozo Contreras, Nhan Duy Truong, Antônio Horta Ribeiro, and Omid Kavehei. Generalization challenges in ECG deep learning: Insights from dataset characteristics and attention mechanism. *medRxiv*, pages 2023–07, 2023.

[32] Zhaojing Huang, Wing Hang Leung, Jiashuo Cui, Leping Yu, Luis Fernando Herbozo Contreras, Nhan Duy Truong, Armin Nikpour, and Omid Kavehei. Cardiac abnormality detection with a tiny diagonal state space model based on sequential liquid neural processing unit. *APL Machine Learning*, 2(2), 2024.

[33] Leon Cohen. *Time-frequency analysis*. Prentice Hall PTR, 1995.

[34] Brandon Allan et al. Electrocardiogram (ECG) signal processing and analysis techniques: A review. *Computer Methods and Programs in Biomedicine*, 120(1):1–29, 2015.

[35] Nhan Duy Truong, Anh Duy Nguyen, Levin Kuhlmann, Mohammad Reza Bonyadi, Jiawei Yang, Samuel Ippolito, and Omid Kavehei. Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Networks*, 105:104–111, 2018.

[36] Alan V. Oppenheim et al. *Frequency domain and time domain representations of signals*. Springer, 2004.

[37] U Rajendra Acharya et al. Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. *Information Sciences*, 415:190–198, 2017.

[38] Philip De Chazal and Richard B Reilly. ECG beat classification using time-frequency analysis and a combination of decision trees and support vector machines. *Computer Methods and Programs in Biomedicine*, 82(2):127–137, 2006.