

Article

Evaluating the Impact of Windowing Techniques on Fourier Transform-Preprocessed Signals for Deep Learning-Based ECG Classification

Niken Prasasti Martono * and Hayato Ohwada

Department of Industrial and Systems Engineering, Tokyo University of Science, Noda 278-8510, Japan; ohwada@rs.tus.ac.jp

* Correspondence: niken@rs.tus.ac.jp

Abstract: (1) Background: Arrhythmias, or irregular heart rhythms, are a prevalent cardiovascular condition and are diagnosed using electrocardiogram (ECG) signals. Advances in deep learning have enabled automated analysis of these signals. However, the effectiveness of deep learning models depends greatly on the quality of signal preprocessing. This study evaluated the impact of different windowing techniques applied to Fourier transform-preprocessed ECG signals on the classification accuracy of deep learning models. (2) Methods: We applied three windowing techniques—Hamming, Hann, and Blackman—to transform ECG signals into the frequency domain. A one-dimensional convolutional neural network was employed to classify the ECG signals into five arrhythmia categories based on features extracted from each windowed signal. (3) Results: The Blackman window yielded the highest classification accuracy, with improved signal-to-noise ratio and reduced spectral leakage compared to the Hamming and Hann windows. (4) Conclusions: The choice of windowing technique significantly influences the effectiveness of deep learning models in ECG classification. Future studies should explore additional preprocessing methods and their clinical applications.

Keywords: arrhythmia; electrocardiogram (ECG); ECG classification; signal processing; deep learning; convolutional neural network (CNN)



Citation: Martono, N.P.; Ohwada, H. Evaluating the Impact of Windowing Techniques on Fourier Transform-Preprocessed Signals for Deep Learning-Based ECG Classification. *Hearts* **2024**, *5*, 501–515. <https://doi.org/10.3390/hearts5040037>

Academic Editor: Matthias Thielmann

Received: 30 September 2024

Revised: 24 October 2024

Accepted: 25 October 2024

Published: 29 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Arrhythmias, characterized by irregular heart rhythms, are among the most prevalent heart conditions. These disorders occur when the heart beats either too fast or too slow due to abnormal electrical impulses traveling through the myocardial tissue. These impulses are responsible for controlling the contraction and relaxation of the myocardium and generating the heart's rhythm. Whereas a healthy heart maintains a consistent rhythm, arrhythmias disrupt this pattern, making the heartbeat erratic, either faster, or slower, depending on factors such as the heart's electrical activity, blood flow, and/or the strength of the myocardial tissue or damage to it [1].

The ECG signal demonstrates a unique waveform for every heartbeat cycle, with each part of the signal representing specific cyclic events. When the atria are filled with blood, the sinoatrial (SA) node fires an electrical impulse that causes atrial depolarization, leading to the formation of the P wave on the ECG [2]. After the P wave, the atria contract for approximately 100 milliseconds. The PQ interval represents the transmission of the signal from the SA node to the atrioventricular (AV) node. Ventricular depolarization follows, marked by the QRS complex, which is initiated by the AV node's firing. The Q wave reflects electrical impulses traveling through the heart's lower regions, whereas the R wave occurs as the signal passes through the ventricles' lateral sides. The S wave denotes the last phase of ventricular depolarization in the heart's lower muscles, and even though atrial repolarization happens concurrently, it is obscured by the QRS complex. The ventricle

continues contracting during the ST segment, and the T wave signals ventricular repolarization [3]. Figure 1 provides a visual representation of these components within a typical ECG waveform.

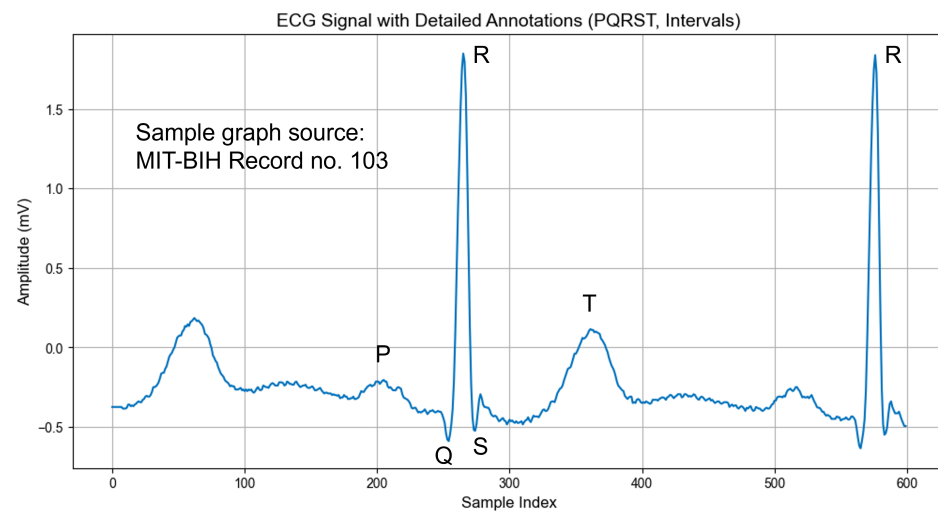


Figure 1. Illustration of PQRST wave in an ECG signal.

By carefully examining these ECG waveforms, segments, and intervals, cardiologists can diagnose various types of arrhythmias. Advances in machine learning and deep learning technologies allow automated analysis of ECG data, extracting key features and enabling the classification of different arrhythmia types. This not only helps detect arrhythmias more efficiently but also assists in making faster, more accurate diagnoses. Through these innovations, arrhythmia detection has become more accessible, improving patient outcomes by allowing for timely intervention.

Deep learning (DL) has demonstrated significant success in medical diagnoses in recent years, particularly in the automatic classification of heart abnormalities using ECG signals [4–8]. DL models learn to map ECG features to their corresponding medical categories, utilizing multiple neural layers. This mapping is optimized through a training process using datasets where neuron weights are adjusted to minimize discrepancies between the predicted and actual categories of the training data. Compared to traditional machine learning methods such as clustering [9] and support vector machines (SVMs) [10], DL-based ECG classification offers a more effective way to map ECG signal characteristics to their respective categories [1,11], due to its powerful multi-level abstraction capability for feature extraction.

Although existing research has yielded promising outcomes in the analysis of ECG signals, several challenges remain unresolved. One significant issue is the imbalance in data, where normal ECG signals far outnumber abnormal ones, leading to difficulty in effectively addressing the imbalance problem. Additionally, the generalization capabilities of many current models are limited. Due to the significant individual variations in ECG patterns, these methods are often inadequate for clinical application [12]. When applied to real-time hospital data, their performance tends to fall short compared to results from publicly available datasets.

Convolutional neural networks (CNNs) have established themselves as potent tools in the morphological analysis of physiological signals, particularly due to the ability of these tools to be used to discern invariant patterns and capture pivotal features across data. This paper presents an innovative approach for ECG signal recognition and classification, which utilizes the strengths of CNNs to analyze signals in both time and frequency domains. To enhance the model's performance and address the nuances of complex ECG signal patterns, different Fourier transform windowing techniques were explored.

While the Fourier transform has long been a cornerstone in the analysis of ECG signals, its application has predominantly focused on the transformation of signals from the

time domain to the frequency domain without a detailed examination of the effects of various windowing techniques. Previous studies have effectively leveraged the Fourier transform to identify fundamental frequency components and to diagnose arrhythmias, ischemic episodes, and other cardiac abnormalities [6,13–15]. Although the Hann, Hamming, and Blackman windowing techniques reduce spectral leakage and improve signal clarity, knowledge regarding how these methods impact the diagnostic capabilities of Fourier-based ECG analysis is lacking. This paper explored how these specific windowing techniques enhance the interpretability of Fourier-transformed ECG signals and, in turn, improve the performance of CNNs in classifying cardiac events. This innovative approach promises not only to refine the analysis of ECG data but also to provide more accurate and reliable diagnostic tools that are critical for clinical decision making.

The primary contributions of this work are summarized as follows: Firstly, Fourier transform techniques with different window functions—Hamming, Hann, and Blackman—were applied to convert the ECG signals into the frequency domain. This transformation facilitated the extraction of frequency-based features that provide valuable insights into the signal's characteristics. Secondly, a one-dimensional (1D)-CNN model was employed to classify the ECG signals into five categories based on the features extracted through each windowing method. This approach not only addressed data imbalances but also enhanced the analysis framework by leveraging features derived from various Fourier windowing techniques. The comparative analysis of these window functions aims to identify which method optimally enhances model performance by effectively balancing the resolution and leakage trade-offs inherent in Fourier-based frequency-domain feature extraction.

2. Materials and Methods

2.1. Dataset Description

In this study, the Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) Arrhythmia Dataset [16] is utilized. This dataset included 24-hour ECG recordings collected from 47 individuals, with subjects ranging in age from 23 to 89 years. We utilized excerpts from 30 to minute, dual-channel ECG recordings collected from 47 participants between 1975 and 1979. The dataset featured two ECG lead signals for each recording, digitized at 360 samples per second with 11-bit resolution over a 10-mV range.

For the purposes of classification, the annotations were grouped into five categories following the Association for the Advancement of Medical Instrumentation (AAMI) EC57 standard, "Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms" [17]. Class N encompasses "normal beats", as well as "left bundle branch block", "right bundle branch block", "atrial escape", and "nodal junction escape" beats. Class V includes "premature ventricular contractions" and "ventricular escape" beats. Class S contains "atrial premature" (AP) beats, "aberrated premature" beats, "nodal junction premature" beats, and "supraventricular" beats.

Class Q, sometimes referred to as "unknown beats", includes "paced beats" (P), "fusion of paced and normal beats" (fPN), and "unclassified beats." Lastly, Class F is comprised of "fusion of ventricular and normal beats" (fVN). These annotations are derived from specific symbols present in the original dataset, which were mapped into five simplified categories for classification purposes. For example, the "N" label includes symbols such as "N", "L", "R", "e", and "j", whereas the "S" label corresponds to symbols like "A", "a", and "S." This mapping was implemented to standardize the labels for the classification task. The illustration of each signal is shown in Figure 2.

The classification and understanding of these beats rely immensely on the identification of the PQRST complex, which is fundamental to interpreting ECG signals. Occasionally, a U wave may be present, following the T wave, whose origin is less clearly understood but is thought to be related to the repolarization of the Purkinje fibers. Understanding the morphology and timing of these components is critical for accurate beat classification and diagnosis, as each class of beats exhibits distinct characteristics in terms of the PQRST

sequence. For instance, ventricular contractions such as those in Class V may display an abnormal QRS complex, whereas atrial contractions in Class S alter the P wave's morphology.

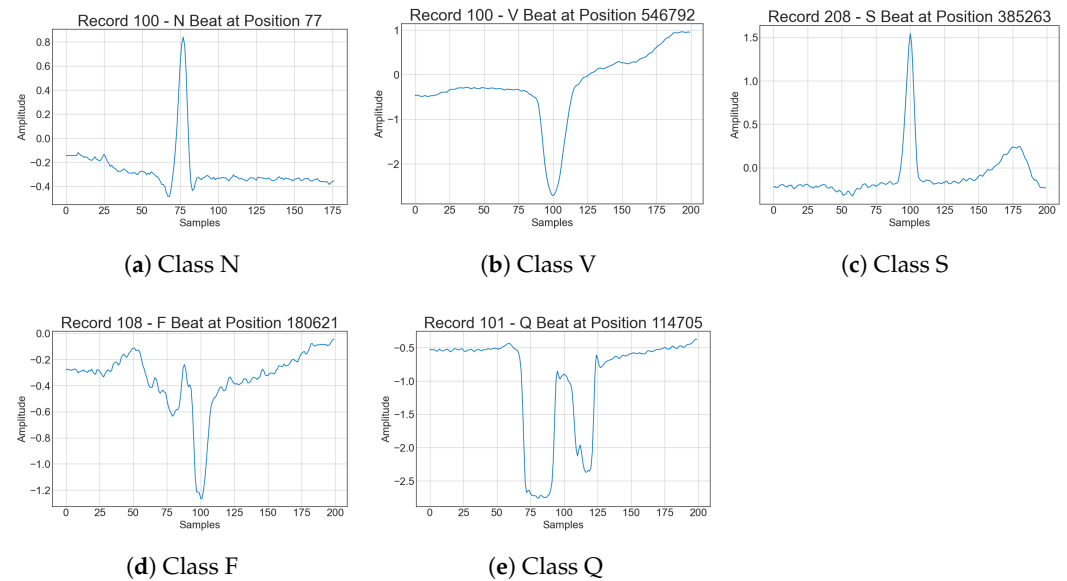


Figure 2. Illustration of time series signals for each label (N, V, S, F, Q).

2.2. Signal Preprocessing

2.2.1. Standardization for Signal Normalization

For this study, we chose standardization as the normalization technique, which is particularly suitable for data involving biological signals such as ECGs. Standardization adjusts the data to have zero mean and unit standard deviation. This is achieved by subtracting the mean and dividing by the standard deviation of each data sample. The standardization formula applied to each signal x in the dataset is expressed as:

$$x' = \frac{x - \mu}{\sigma} \quad (1)$$

where μ represents the mean of the signal, and σ denotes the standard deviation.

Standardization was selected over other normalization techniques like min-max scaling because it effectively addresses features that vary in scale and distribution. The ECG signals, which exhibit significant variations in amplitude and waveform due to factors such as heart size, electrode placement, and physiological conditions, benefit from standardization as it normalizes the range without distorting differences in values or losing information about zero entries.

2.2.2. Noise Reduction Using Moving Average Filters

Noise reduction is a critical preprocessing step in ECG signal analysis to ensure the reliability and accuracy of the data before further processing. Among various techniques available, the moving average filter is widely employed due to its simplicity and effectiveness in smoothing out short-term fluctuations and highlighting longer-term trends in the data.

The moving average filter operates by creating an average of different subsets of the total number of data points available in a signal, effectively smoothing the signal. This is particularly useful in ECG signal processing, where high-frequency noise can obscure the true heart rate signal and other important diagnostic features.

The moving average filter is applied to the ECG signal using the following formula:

$$y[i] = \frac{1}{N} \sum_{j=0}^{N-1} x[i + j] \quad (2)$$

where $x[i]$ represents the original data points in the signal, $y[i]$ represents the output of the moving average filter, and N is the number of data points in the moving average window, which determines how much the data will be smoothed. In this study, a window size of 5 samples was chosen based on the sampling rate and the expected frequency of the noise components. This window size provided a balance between smoothing the signal to remove high-frequency noise and preserving the essential characteristics of the ECG signal, such as the P wave, QRS complex, and T wave.

2.3. Feature Extraction and Windowing Techniques

Feature extraction is a pivotal aspect of signal processing, especially for tasks that involve the classification of physiological signals like ECGs. The Fourier transform is a powerful mathematical tool used for deconstructing a signal into its constituent frequencies, providing a different perspective that is particularly useful for analyzing the frequency content of signals. The Fourier transform [18] converts a time-domain signal, which is a function of time, into a frequency-domain signal, which is a function of frequency. This transformation reveals the different frequencies present in the signal and their amplitudes, which can be crucial for identifying rhythmic patterns such as those found in heartbeats.

The continuous Fourier transform of a continuous, time-dependent signal $x(t)$ is defined as:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt \quad (3)$$

where:

- $X(f)$ is the Fourier Transform of $x(t)$;
- f is the frequency in Hertz;
- t is time;
- j is the imaginary unit.

For digital signals, we use the discrete Fourier transform (DFT), especially implemented in an efficient manner through the fast Fourier transform (FFT) algorithm as shown in Equation (4).

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-j2\pi \frac{kn}{N}} \quad (4)$$

where:

- N is the total number of samples;
- $x[n]$ is the signal value at sample n ;
- $X[k]$ represents the frequency component at frequency k .

To enhance the spectral purity of the Fourier analysis, finite impulse response (FIR) filters [19] such as the Hann, Hamming, and Blackman windows are employed to precondition the ECG signal. These filters are specifically designed to address the phenomenon of spectral leakage, during which energy from strong frequency components bleeds into neighboring frequencies, potentially obscuring or altering the true spectral content [20].

The Hann, Hamming, and Blackman windows are widely adopted in ECG analysis due to the nature of physiological signals, which often contain both strong and weak components. ECG signals are characterized by cyclic waveforms (such as PQRST complexes) that demand a careful balance between frequency resolution and spectral leakage suppression. Each of these different window functions provides specific advantages. The Hann window works well when a compromise between frequency and time resolution is needed, whereas the Hamming window provides superior frequency resolution, making it ideal for distinguishing small changes between consecutive beats. In contrast, the Blackman window offers the best leakage suppression, making it a top choice in noisy environments, ensuring that even small amplitude signals are accurately captured [21,22].

Applying FIR filters in the preprocessing stages of signal analyses enhances the effectiveness of subsequent DL models. These filters are important for reducing spectral leakage and noise, thereby improving the quality of the signal. Enhanced signal quality guarantees

that DL models train on data that accurately represent the underlying physiological signals, which is particularly vital in ECG analysis. The clear and distinct representation of frequency components achieved through FIR filtering facilitates the extraction of subtle features. This is especially beneficial for CNNs, which rely strongly on high-quality inputs to detect key data patterns. By improving the signal-to-noise ratio and emphasizing important signal characteristics, FIR filters reduce the complexity of the models required to achieve high performance, thus enhancing training efficiency and predictive accuracy.

When applying the FFT, spectral leakage can occur, where energy from one frequency spreads to adjacent frequencies, distorting the frequency spectrum. This issue arises from the discontinuities introduced by the finite length of sampled signals. To mitigate this, window functions are applied to the signal before performing the Fourier transform [21,22]. Window functions taper the signal to zero at the edges, ensuring smooth transitions at the boundaries, which minimizes the leakage into adjacent frequency bins.

The window size refers to the number of consecutive samples of the signal to which the window function is applied [22]. The window defines a segment of the signal to be analyzed, and the choice of window size plays a critical role in determining the frequency resolution and time resolution of the analysis. In the context of ECG signal processing, the window size must be carefully chosen to balance the trade-off between frequency resolution and time resolution [23]. Typically, the window size matches the length of the ECG signal segment being analyzed. For example, if an ECG sample contains 360 time points, the window function will be applied to these 360 points.

2.3.1. Hann Window

The Hann window, often referred to as the Hanning window, is designed to reduce spectral leakage effectively [20,24]. It is mathematically defined as:

$$w(n) = 0.5 \left(1 - \cos \left(\frac{2\pi n}{N-1} \right) \right) \quad (5)$$

where n is the sample index ranging from 0 to $N-1$, and N is the total number of samples in the window. This window function tapers the signal to zero at both ends, thus minimizing the discontinuities at the window boundaries and reducing the resultant spectral leakage.

2.3.2. Hamming Window

The Hamming window offers a narrower main lobe compared to the Hann window [20,24], which enhances its ability to resolve close frequency components but at the cost of higher side lobes. It is defined by the expression:

$$w(n) = 0.54 - 0.46 \cos \left(\frac{2\pi n}{N-1} \right) \quad (6)$$

This characteristic provides a better frequency resolution and is beneficial when analyzing complex signals where precision in frequency component separation is crucial.

2.3.3. Blackman Window

For applications requiring significant reduction of spectral leakage, the Blackman window is a superior choice [20,24]. It incorporates a second-order cosine term to further attenuate the side lobes, as defined by:

$$w(n) = 0.42 - 0.5 \cos \left(\frac{2\pi n}{N-1} \right) + 0.08 \cos \left(\frac{4\pi n}{N-1} \right) \quad (7)$$

The inclusion of the additional cosine term results in a significantly higher attenuation of the side lobes, making the Blackman window especially effective in situations where leakage needs to be minimized to detect small-amplitude frequencies adjacent to large-amplitude components.

The selection of specific FIR filters can be strategically aligned with the requirements of the deep learning model, taking into account the characteristics of the ECG signal and the diagnostic objectives. For instance, a Blackman window, known for its superior attenuation of side lobes, can be particularly useful in noisy environments where detailed feature analysis is critical [25]. This strategic alignment ensures that the preprocessing not only augments the capabilities of deep learning models but also optimizes them for more effective learning and improved diagnostic outcomes. Therefore, integrating FIR filters into the preprocessing stage is instrumental in maximizing the potential of advanced machine learning techniques for medical signal analysis, thereby enhancing the overall diagnostic capabilities in clinical settings.

2.3.4. Data Balancing Using SMOTE

In many real-world datasets, especially in medical applications like ECG signal analysis, class imbalance poses a significant challenge. This occurs when some classes have substantially fewer samples than others, which can lead to biased models that perform poorly on minority classes. To address this, we used the Synthetic Minority Oversampling Technique (SMOTE) [26] to balance the dataset.

SMOTE is an oversampling technique that creates synthetic samples for the minority class by interpolating between existing minority samples [26]. Unlike simple random oversampling, which duplicates samples and can lead to overfitting, SMOTE generates new synthetic samples, making the model more generalizable and robust.

The SMOTE algorithm operates as follows:

- For each minority class sample, a set of its k -nearest neighbors is identified.
- A new synthetic sample is generated by randomly selecting one of these neighbors and creating a sample along the line segment connecting the original sample and the neighbor.

Mathematically, if x_i is a sample from the minority class, and x_j is one of its k -nearest neighbors, the synthetic sample x_{new} is given by Equation (8).

$$x_{new} = x_i + \lambda \cdot (x_j - x_i) \quad (8)$$

where λ is a random number between 0 and 1.

By implementing SMOTE, we ensured that the machine learning model does not favor the majority class, allowing for more balanced predictions across all classes. In the context of ECG signal analysis, where it is crucial to accurately detect rare events such as arrhythmias, the use of SMOTE ensures that minority events are adequately represented during training. This reduces the risk of misclassification and improves the overall diagnostic accuracy of the model. However, it is important to apply SMOTE cautiously, especially for time-series data, such as ECG signals, as generating synthetic samples can sometimes introduce noise. To mitigate this, SMOTE was applied only after signal preprocessing steps, including noise reduction and normalization, to ensure high-quality synthetic samples. Additionally, model performance on minority classes was closely monitored during training to validate the effectiveness of the balancing strategy.

2.4. 1D-CNN Model Architecture

A CNN is a type of deep learning architecture that excels at extracting high-level features from input data. Conceptually similar to a multilayer perceptron (MLP), each neuron in a CNN has an activation function that processes weighted inputs to produce outputs. CNNs are composed primarily of three types of layers: convolutional layers, pooling layers, and fully connected layers [27]. These networks, when adequately trained, find application in various fields such as speech recognition, structural engineering, and image processing [28].

A 1D-CNN is a variant of the traditional CNN, tailored specifically for handling one-dimensional data, making it particularly effective for sparse datasets. Unlike 2D-CNNs,

which use two-dimensional convolutional filters, 1D-CNNs utilize one-dimensional filters to extract the data features. This allows 1D-CNNs to be more suitable for processing audio or text data and more computationally efficient due to having fewer parameters. Their design grants the capture of local features within the signal, making them robust to slight shifts in time, particularly for analyzing time-series data with temporal components [29].

The first 1D-CNN classification model used in this study is designed to process the ECG signals and classify them into one of the five heartbeat categories. The architecture, as shown in Table 1, consists of:

- Three convolutional layers with filter sizes of 32, 64, and 64, respectively, and a kernel size of 3. Each convolutional layer is followed by a max pooling layer with a pool size of 2 to reduce the spatial dimensions of the data.
- Flattening layer, which transforms the 1D convoluted data into a flat vector for the fully connected layers.
- Two dense layers: the first dense layer has 64 units, followed by another dense layer with 32 units, both using the ReLU activation function.
- Dropout layer (with a dropout rate of 0.5) is added after the first dense layer to prevent overfitting by randomly deactivating neurons during training.
- The final output layer uses a softmax activation function to predict the probability of each class, with five output units corresponding to the five heartbeat categories.

Table 1. Parameter Tuning of the Proposed 1D CNN Model.

Parameter	Value
Pooling Type	Max Pooling
Pooling Size	2
Units in First Dense Layer	64
Units in Second Dense Layer	32
Activation Function	ReLU
Dropout Rate	0.5
Output Layer Activation	Softmax
Number of Output Units	5
Optimizer	Adam

This architecture was chosen for its ability to capture local patterns in the ECG signals while maintaining a manageable number of parameters. The model was trained using the Adam optimizer, with categorical cross-entropy as the loss function since this is a multi-class classification problem. The model was trained for 15 epochs with a batch size of 32, in line with prior research settings. The training dataset was split into 60% training and 40% testing to ensure the model was evaluated on a substantial portion of the data. Validation data were taken from the test set to monitor the model's performance during training. Additionally, the class distribution in the training and test sets was adjusted to reflect the balanced dataset achieved through SMOTE.

3. Results

3.1. Effectiveness of FIR Window Functions in Signal Preprocessing Quantitative Analysis

The signal-to-noise ratio (SNR) is a metric in signal processing used to quantify the clarity and quality of a signal relative to the background noise. It is particularly important in the context of ECG signal analysis where distinguishing true signal from noise can influence diagnostic decisions. The SNR is expressed in decibels (dB) and calculated using the following Equation [30]:

$$\text{SNR (dB)} = 10 \cdot \log_{10} \left(\frac{P_{\text{signal}}}{P_{\text{noise}}} \right) \quad (9)$$

where P_{signal} represents the power of the desired signal, calculated as the sum of the squares of the signal amplitudes, and P_{noise} denotes the power of the noise, which is determined by the difference between the original and the FIR-filtered signal:

$$P_{\text{signal}} = \sum x(t)^2, \quad P_{\text{noise}} = \sum (x(t) - x_{\text{filtered}}(t))^2 \quad (10)$$

Here, $x(t)$ is the amplitude of the original ECG signal at time t , and $x_{\text{filtered}}(t)$ is the amplitude after applying an FIR filter, such as Hann, Hamming, or Blackman. The application of these filters typically aims to enhance the SNR by reducing the noise components without distorting the essential features of the ECG signal. By improving the SNR, the filtered signal can offer clearer and more discernible cardiac events, facilitating more accurate analysis and interpretation.

In the analysis of FIR filters applied to ECG signals, the Hann and Hamming filters demonstrated superior performance in enhancing the signal-to-noise ratio (SNR), both achieving improvements of around 4 dB (Figure 3). The Hamming filter slightly outperformed the Hann filter, suggesting its marginally better efficacy in minimizing spectral leakage and enhancing signal clarity. Conversely, the Blackman filter, while still effective, showed a lower SNR improvement of approximately 3 dB. The observed variations in SNR improvements, denoted by error bars in the results, underscore the consistency of the filters' effects across multiple datasets. This variability is crucial for understanding the practical implications of filter selection in clinical ECG analysis, where signal integrity can significantly influence diagnostic accuracy.

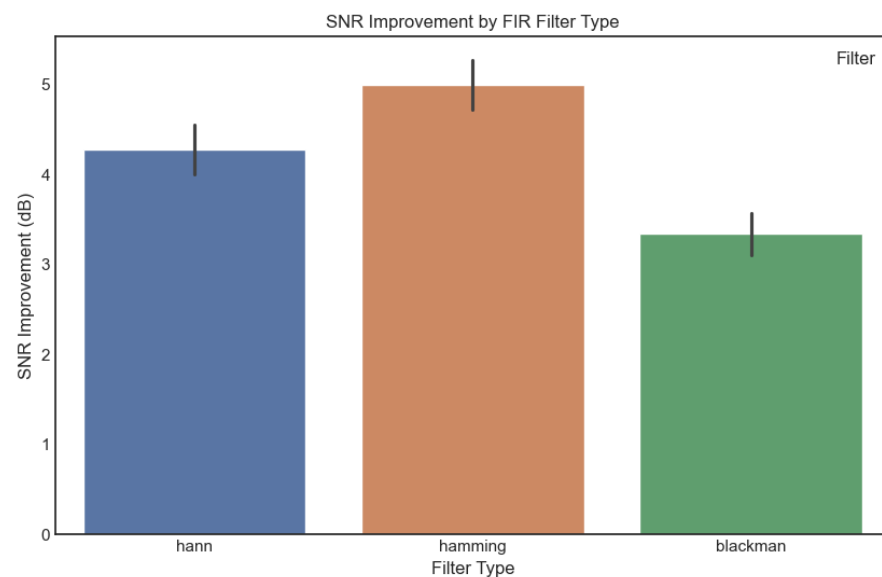


Figure 3. Comparison of spectral analysis results for ECG signals processed using Hann, Hamming, and Blackman windows.

3.2. Performance of Deep Learning Models on Preprocessed Signals

3.2.1. Confusion Matrix Analysis

In classifying ECG signals using different windowing techniques in preprocessing, the evaluation of classification performance is central, with a specific focus on confusion matrices for the Blackman, Hamming, and Hann, and “None” (no technique applied) windows. The classification matrix is an imperative metric to assess how well the classification algorithms performed under the different spectral conditions induced by the respective window types. The classification matrices of both training and test data prediction are shown in Figures 4 and 5.

The confusion matrix for the Blackman window demonstrates high classification accuracy across the board, with a notable 34,349 correct classifications for class F, indicating

strong sensitivity. However, there were 297 instances of class N being misclassified as F, which was the largest source of confusion. Similarly, class V indicated 51 instances of misclassification as F, suggesting some challenges in distinguishing between F and V. Overall, the Blackman window offered good specificity and sensitivity, with a limited spread in misclassifications across other categories.

The Hamming window’s confusion matrix also displayed robust performance, with 34,349 correct classifications for class F. There were 271 misclassifications of N as F and 228 cases where class S was also predicted as F. However, the number of mis-categorizations between other classes (such as between V and S) was relatively small. The Hamming window demonstrated effective sorting overall but performed slightly more misclassifications than the Blackman window, particularly in distinguishing between classes S and V. The Hann window demonstrated effective organization performance with 34,357 correct classifications for class F. However, it introduced some noticeable errors, such as 291 cases where class S was classified as F, and 276 cases where class N was also confused with F. In terms of distinguishing class V, the Hann window indicated 123 instances were misclassified as F, indicating similar behavior as the other windows but with a slight tendency toward more errors.

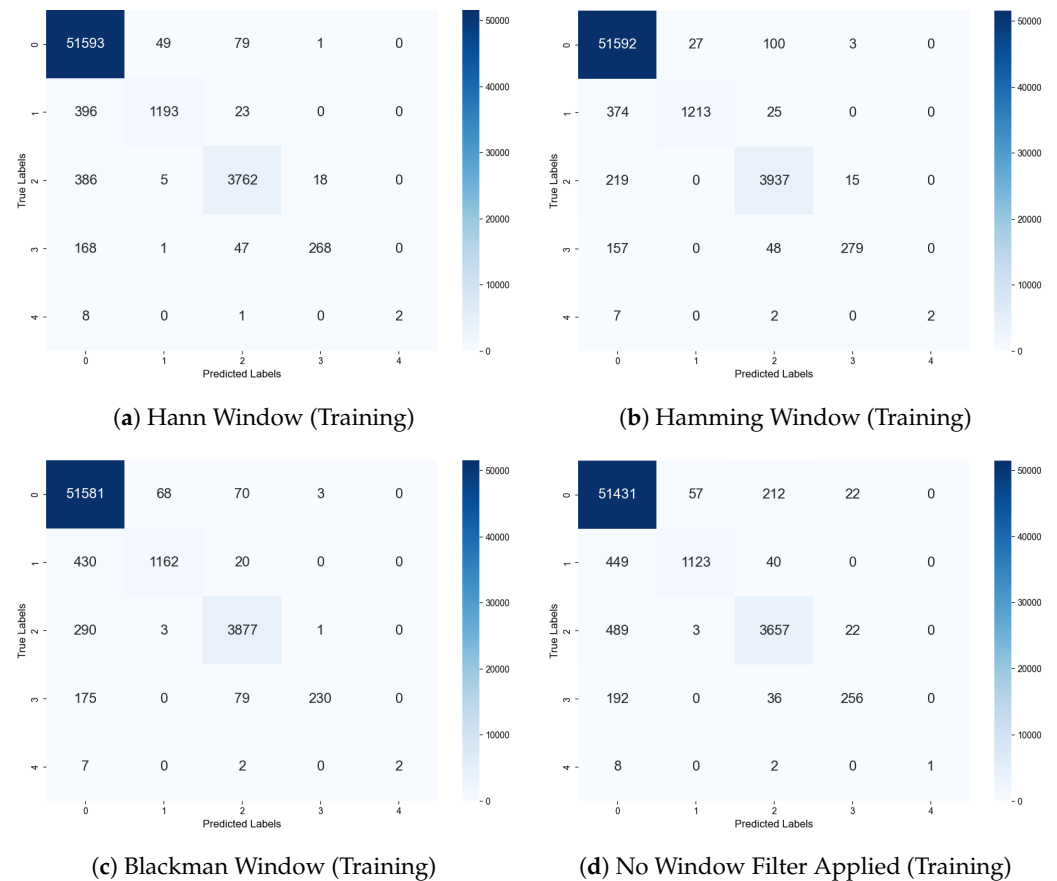


Figure 4. Comparison of confusion matrices for ECG signal classification processed using different window types: (a) Hann, (b) Hamming, (c) Blackman, and (d) No Window using training data set.

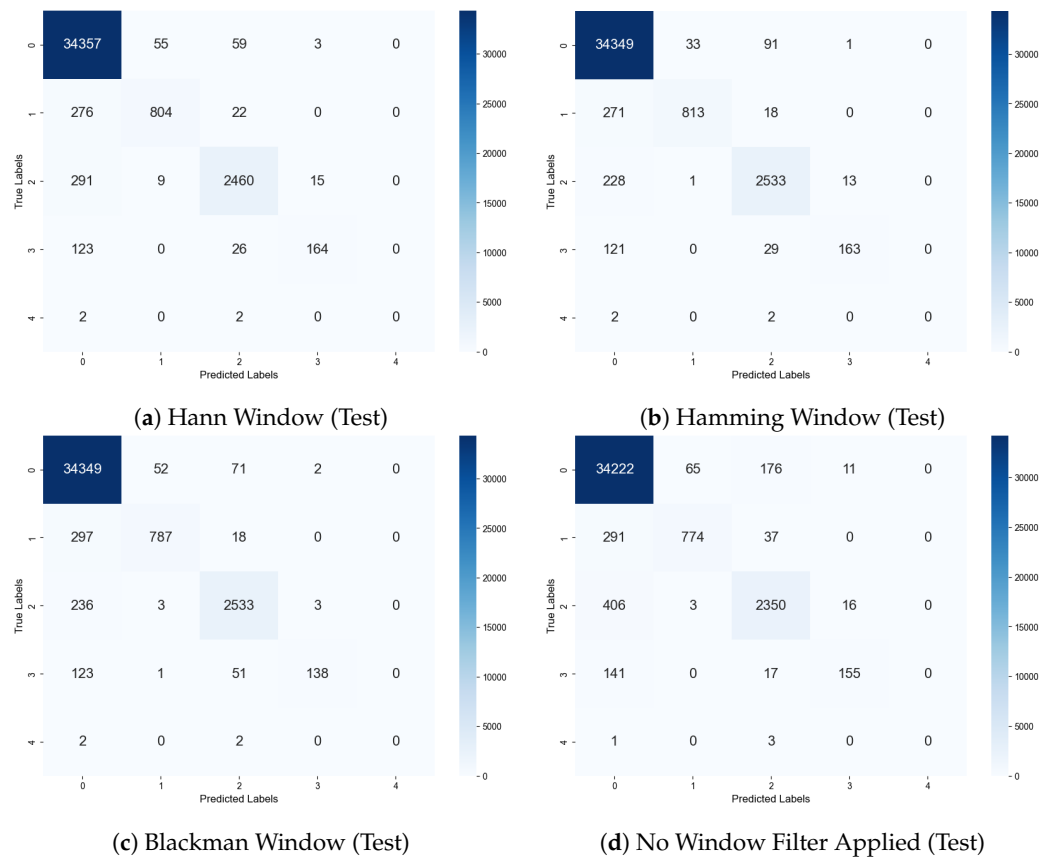


Figure 5. Comparison of confusion matrices for ECG signal classification processed using different window types: (a) Hann, (b) Hamming, (c) Blackman, and (d) No Window using test data set.

The absence of a windowing function (None) altogether resulted in the highest number of misclassifications, highlighting the importance of windowing for signal preprocessing. Without the benefits of tapering, 34,222 samples of class F were correctly classified, which is slightly lower than the other window techniques. Additionally, 406 instances of class S were miscategorized as F, and 141 instances of V were also misclassified as F, highlighting the limitations of the use of the raw signals in achieving high classification accuracy. The lack of windowing exacerbated spectral leakage and high-frequency noise, leading to reduced precision and specificity. This result emphasizes the need for windowing techniques to reduce noise and enhance feature preservation, particularly for models that rely substantially on subtle signal patterns for classification.

All three windowing techniques—Blackman, Hamming, and Hann—exhibited robust classification capabilities, but subtle differences in their performances make certain windows more suitable for specific applications. The Blackman window emerged as the most reliable option, with the least number of misclassifications and the highest specificity, making it ideal for high-stakes, high-precision clinical diagnostics. Similarly, the Hamming window offered strong performance but introduced slightly more misclassifications, particularly between related classes, suggesting that careful tuning may be required when precision is critical. The Hann exhibited slightly more spectral leakage, leading to more frequent class confusion, but it may be suitable for applications where the focus is on capturing long-term trends rather than precise transitions.

The “None” configuration, while still functional, highlighted the necessity of utilization of windowing techniques for achieving optimal performance. The increased number of misclassifications underscored the role of windowing in reducing noise and spectral leakage, improving the model’s ability to learn from the data. Therefore, the choice of window function should be guided by the specific requirements of the application—whether it

is a precise class distinction, resolution of closely spaced frequency components, or overall trend capture.

3.2.2. Overall Performance

The performance analysis of the three window functions—Hamming, Hann, and Blackman—on ECG signal classification provides a detailed insight into their effectiveness across five classes (F, N, S, V, Q), evaluated through metrics such as precision, recall, and F1-Score (Table 2). These metrics are defined as follows:

- **Precision:** the ratio of correctly predicted positive observations to the total predicted positive observations.

$$\text{Precision} = \frac{TP}{TP + FP}$$

where *TP* is the number of true positives, and *FP* is the number of false positives.

- **Recall:** the ratio of correctly predicted positive observations to all observations in the actual class.

$$\text{Recall} = \frac{TP}{TP + FN}$$

where *FN* is the number of false negatives.

- **F1-Score:** the harmonic mean of precision and recall, providing a balance between the two.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Table 2. Performance Metrics for Each Window Type and Dataset (Training vs. Test).

Window Type	Class	Training				Test			
		Precision	Recall	F1-Score	Samples	Precision	Recall	F1-Score	Samples
None	F	0.9784	0.9944	0.9863	51,722	0.9761	0.9927	0.9843	34,474
	N	0.9493	0.6967	0.8036	1612	0.9192	0.7024	0.7963	1102
	S	0.9265	0.8768	0.9010	4171	0.9098	0.8468	0.8772	2775
	V	0.8533	0.5289	0.6531	484	0.8516	0.4952	0.6263	313
	Q	1.0000	0.0909	0.1667	11	0.0000	0.0000	0.0000	4
Hamming	F	0.9855	0.9975	0.9915	51,722	0.9822	0.9964	0.9892	34,474
	N	0.9782	0.7525	0.8506	1612	0.9599	0.7377	0.8343	1102
	S	0.9574	0.9439	0.9506	4171	0.9476	0.9128	0.9299	2775
	V	0.9394	0.5764	0.7145	484	0.9209	0.5208	0.6653	313
	Q	1.0000	0.1818	0.3077	11	0.0000	0.0000	0.0000	4
Hann	F	0.9818	0.9975	0.9896	51,722	0.9803	0.9966	0.9884	34,474
	N	0.9559	0.7401	0.8343	1612	0.9263	0.7296	0.8162	1102
	S	0.9617	0.9019	0.9308	4171	0.9576	0.8865	0.9207	2775
	V	0.9338	0.5537	0.6952	484	0.9011	0.5240	0.6626	313
	Q	1.0000	0.1818	0.3077	11	0.0000	0.0000	0.0000	4
Blackman	F	0.9828	0.9973	0.9900	51,722	0.9812	0.9964	0.9887	34,474
	N	0.9424	0.7208	0.8169	1612	0.9336	0.7142	0.8093	1102
	S	0.9578	0.9295	0.9434	4171	0.9469	0.9128	0.9295	2775
	V	0.9829	0.4752	0.6407	484	0.9650	0.4409	0.6053	313
	Q	1.0000	0.1818	0.3077	11	0.0000	0.0000	0.0000	4

All three models performed excellently in classifying class F, with notably high precision and recall, suggesting their strong capability in accurately identifying this category with minimal false positives and negatives. Classes N and S were also observed with relatively good performance, although the recall for class N was slightly reduced, indicating fewer true positives detected relative to the actual number of positives present. This outcome could impact the clinical utility of the model where missing a positive case can have significant consequences. The performance on class V was moderate across all models, with precision slightly lower than other classes, hinting at a higher rate of false positives for this class. This finding might suggest a common challenge in the classification of this class using window-based spectral analysis, possibly due to the overlap in spectral characteristics between class V and other classes.

A notable observation across all three models was the complete lack of detection of class Q, which exhibited zero values in precision, recall, and F1-score. This uniform

failure to detect class Q likely indicates that class Q represents a type of signal that is undetectable under the current feature extraction and classification methodologies employed. This consequence could be due to the absence of distinct or adequate features within the spectral domain that these window functions analyze, or possibly the absence of sufficient representative samples of class Q in the training dataset.

Overall, while all three window functions provide robust classification capabilities for most ECG signal classes, the choice among them should consider the specific classification needs and the unique challenges presented by each class. The slight differences in performance metrics between the models can guide the selection of a window function based on whether higher precision or a better-balanced F1-Score is more critical for the intended application.

The results clearly indicate the utility of FIR window functions in enhancing the signal-to-noise ratio and reducing artifacts in ECG signals, which in turn significantly improved the performance of the deep learning models. The particular use of Blackman window filtering facilitated more refined feature extraction, thereby aiding in more accurate heart condition diagnostics. The findings suggested that integrating advanced FIR filtering techniques with deep learning frameworks can significantly advance the field of biomedical signal analysis.

4. Conclusions

This study investigated the impact of applying different FIR window functions—Hann, Hamming, and Blackman—on ECG signal classification using convolutional neural networks. The results established that the choice of window function plays a critical role in balancing frequency resolution, leakage suppression, and classification performance across different ECG classes.

From the performance metrics and confusion matrices, all FIR filters utilized enhanced the signal-to-noise ratio and improved the clarity of ECG signals, which in turn facilitated more accurate feature extraction by the 1D-CNN models. However, differences in precision, recall, and F1-scores among window functions suggested that specific windows perform better under certain conditions, as highlighted below.

4.1. Window Function Performance

- **Hamming Window: Best Overall Balance**
The Hamming window achieved high F1-scores and recall, especially for smaller classes (S and V) in both training and test datasets. This suggests that the Hamming window's narrow main lobe helped the CNN distinguish subtle ECG features, such as T-wave variations and atrial fibrillation events. It also showed the best generalization performance, with consistent results across training and test sets, indicating that it effectively balances frequency resolution and leakage suppression.
- **Hann Window: Consistent General Purpose Window**
The Hann window provided a balanced trade-off between precision and recall for most classes, particularly in identifying normal beats (F-class), while not excelling in any one metric, the Hann window demonstrated stable performance across all classes, making it a practical choice for general ECG analysis tasks where both time and frequency information are critical.
- **Blackman Window: Superior Noise Suppression but Limited Generalization**
The Blackman window, known for its excellent side-lobe suppression, performed well in training but showed lower recall and F1-scores on the test data, particularly for rarer classes like V and Q.
- **No Window: Moderate Performance Across Metrics**
The absence of a window function resulted in lower F1-scores and recall, particularly for smaller classes (Q and V). This outcome highlights the importance of preconditioning ECG signals with window functions to reduce spectral leakage and enhance classification performance.

4.2. Class-Specific Observations

Normal beats (F-class) consistently achieved the highest precision, recall, and F1-scores across all window types due to their abundance in the dataset. Rare Classes (Q-class) exhibited low recall and F1-scores across all window types, indicating that class imbalance remains a challenge despite the use of FIR filters. The performance for Q-class was especially poor on the test set, with zero recall and precision in some cases. PVC beats (V-class) showed improved detection when using the Hamming and Hann windows, but struggled with the Blackman window on unseen test data.

4.3. Implications

The results suggest that Hamming and Hann windows are the most suitable for ECG signal analysis, with Hamming performing better for smaller classes and Hann providing consistent results across all classes. The Blackman window, while effective for noise suppression, may require further regularization techniques to reduce overfitting. The use of no window function is not recommended, as it led to poorer model performance due to the lack of leakage control. The findings highlight the importance of selecting appropriate window functions during preprocessing to ensure that the deep learning models receive high-quality input signals. This is especially relevant in clinical settings, where accurate detection of rare cardiac events can significantly impact patient outcomes.

4.4. Limitations and Future Work

Although this study provides valuable insights, further research is needed to address the class imbalance problem observed in the detection of rare ECG events. Additionally, exploring adaptive windowing techniques that dynamically adjust window parameters based on the signal characteristics could further enhance model performance. Future work could also investigate the impact of other window functions or hybrid filtering methods on different types of ECG signals to optimize classification accuracy across all conditions.

Author Contributions: N.P.M.: Conceptualization, Data analysis, Writing—review and editing; H.O.: Review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available at <https://www.physionet.org/content/mitdb/1.0.0/> [16] accessed on February 2023.

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

References

1. Antzelevitch, C.; Burashnikov, A. Overview of Basic Mechanisms of Cardiac Arrhythmia. *Card. Electrophysiol. Clin.* **2011**, *3*, 23–45. [[CrossRef](#)] [[PubMed](#)]
2. Bhattacharyya, S.; Munshi, N.V. Development of the Cardiac Conduction System. *Cold Spring Harb. Perspect. Biol.* **2020**, *12*, a037408. [[CrossRef](#)]
3. Varalakshmi, P.; Sankaran, A.P. An improved hybrid AI model for prediction of arrhythmia using ECG signals. *Biomed. Signal Process. Control* **2023**, *80*. [[CrossRef](#)]
4. Xiao, Q.; Lee, K.; Mokhtar, S.A.; Ismail, I.; bin Md Pauzi, A.L.; Zhang, Q.; Lim, P.Y. Deep Learning-Based ECG Arrhythmia Classification: A Systematic Review. *Appl. Sci.* **2023**, *13*, 4964. [[CrossRef](#)]
5. Ahmed, A.A.; Ali, W.; Abdullah, T.A.; Malebary, S.J. Classifying Cardiac Arrhythmia from ECG Signal Using 1D CNN Deep Learning Model. *Mathematics* **2023**, *11*, 562. [[CrossRef](#)]
6. Eleyan, A.; Alboghbaish, E. *Multi-Classifer Deep Learning Based System for ECG Classification Using Fourier Transform*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2023. [[CrossRef](#)]
7. Ullah, H.; Heyat, M.B.B.; Akhtar, F.; Sumbul, M.; Mueen, A.Y.; Islam, M.S.; Abbas, Z.; Pan, T.; Gao, M.; Lin, Y.; et al. An End-to-End Cardiac Arrhythmia Recognition Method with an Effective DenseNet Model on Imbalanced Datasets Using ECG Signal. *Comput. Intell. Neurosci.* **2022**, *2022*, 9475162. [[CrossRef](#)]

8. Zhang, H.; Liu, C.; Zhang, Z.; Xing, Y.; Liu, X.; Dong, R.; He, Y.; Xia, L.; Liu, F. Recurrence Plot-Based Approach for Cardiac Arrhythmia Classification Using Inception-ResNet-v2. *Front. Physiol.* **2021**, *12*, 648950. [[CrossRef](#)]
9. Yeh, Y.C.; Chiou, C.W.; Lin, H.J. Analyzing ECG for cardiac arrhythmia using cluster analysis. *Expert Syst. Appl.* **2012**, *39*, 1000–1010. [[CrossRef](#)]
10. Dhyani, S.; Kumar, A.; Choudhury, S. Analysis of ECG-based arrhythmia detection system using machine learning. *MethodsX* **2023**, *10*, 102195. [[CrossRef](#)]
11. Śmigiel, S.; Pałczyński, K.; Ledziński, D. ECG signal classification using deep learning techniques based on the PTB-XL dataset. *Entropy* **2021**, *23*, 1121. [[CrossRef](#)]
12. Ansari, Y.; Mourad, O.; Qaraq, K.; Serpedin, E. Deep learning for ECG Arrhythmia detection and classification: An overview of progress for period 2017–2023. *Front. Physiol.* **2023**, *14*, 1246746. [[CrossRef](#)] [[PubMed](#)]
13. Aziz, S.; Ahmed, S.; Alouini, M.S. ECG-based machine-learning algorithms for heartbeat classification. *Sci. Rep.* **2021**, *11*, 18738. [[CrossRef](#)]
14. Biran, A.; Jeremic, A. *ECG Based Human Identification Using Short Time Fourier Transform and Histograms of Fiducial QRS Features*; SciTePress: Setúbal, Portugal, 2020; pp. 324–329. [[CrossRef](#)]
15. Kumar, M.A.; Chakrapani, A. Classification of ECG signal using FFT based improved Alexnet classifier. *PLoS ONE* **2022**, *17*, e0274225. [[CrossRef](#)] [[PubMed](#)]
16. Moody, G.B.; Mark, R.G. MIT-BIH Arrhythmia Database. 1992. Available online: <https://physionet.org/content/mitdb/1.0.0/> (accessed on 29 September 2024).
17. Yang, M.; Liu, W.; Zhang, H. A robust multiple heartbeats classification with weight-based loss based on convolutional neural network and bidirectional long short-term memory. *Front. Physiol.* **2022**, *13*, 982537. [[CrossRef](#)] [[PubMed](#)]
18. Bracewell, R.N.; Bracewell, R.N. *The Fourier Transform and Its Applications*; McGraw-Hill New York: New York, NY, USA, 1986; Volume 31999.
19. Oppenheim, A.V. *Discrete-Time Signal Processing*; Pearson Education India: Bengaluru, India, 1999.
20. Podder, P.; Khan, T.Z.; Khan, M.H.; Rahman, M.M. Comparative Performance Analysis of Hamming, Hanning and Blackman Window. 2014. Available online: <https://www.ijcaonline.org/archives/volume96/number18/16891-6927/> (accessed on 29 September 2024).
21. Gharaibeh, K. Assessment of various window functions in spectral identification of passive intermodulation. *Electronics* **2021**, *10*, 1034. [[CrossRef](#)]
22. Prabhu, K.M. *Window Functions and Their Applications in Signal Processing*; Taylor & Francis: Abingdon, UK, 2014.
23. Zhang, M. Multi-resolution short-time Fourier transform providing deep features for 3D CNN to classify rolling bearing fault vibration signals. *Eng. Res. Express* **2024**, *6*, 035201. [[CrossRef](#)]
24. Kaur, M.; Kaur, S.P. High Frequency Noise Removal From Electrocardiogram Using Fir Low Pass Filter Based On Window Technique. *Procedia Technol.* **2018**, *8*, 27–32. [[CrossRef](#)]
25. Berryman, F.; Pynsent, P.; Cubillo, J. The effect of windowing in Fourier transform profilometry applied to noisy images. *Opt. Lasers Eng.* **2004**, *41*, 815–825. [[CrossRef](#)]
26. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [[CrossRef](#)]
27. Kiranyaz, S.; Avci, O.; Abdeljaber, O.; Ince, T.; Gabbouj, M.; Inman, D.J. 1D convolutional neural networks and applications: A survey. *Mech. Syst. Signal Process.* **2021**, *151*, 107398. [[CrossRef](#)]
28. Kiranyaz, S.; Ince, T.; Gabbouj, M. Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks. *IEEE Trans. Biomed. Eng.* **2016**, *63*, 664–675. [[CrossRef](#)] [[PubMed](#)]
29. Ige, A.O.; Sibiya, M. State-of-the-Art in 1D Convolutional Neural Networks: A Survey. *IEEE Access* **2024**, *12*, 144082–144105. [[CrossRef](#)]
30. Czanner, G.; Sarma, S.V.; Ba, D.; Eden, U.T.; Wu, W.; Eskandar, E.; Lim, H.H.; Temereanca, S.; Suzuki, W.A.; Brown, E.N. Measuring the signal-to-noise ratio of a neuron. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 7141–7146. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.