JOURNAL OF ECONOMICS AND ADMINISTRATIVE SCIENCES

E-ISSN: 2227-703X P-ISSN: 2518-5764

Volume 32 | Number 1 26 - 37

4-30-2026

Comparison of 1D-Convolution Neural Network and Hilbert Huang Transform-Based Features for ECG Signal Classification

Baqir Ahmed Hussein
Department of Statistics, College of Administration and Economics, University of Baghdad, Iraq,
baqer.ahmed2201m@coadec.uobaghdad.edu.iq

Rabab Abdulrida Saleh
Department of Statistics, College of Administration and Economics, University of Baghdad, Iraq,
rabab.saleh@coadec.uobaghdad.edu.iq

JOURNAL OF ECONOMICS AND ADMINISTRATIVE SCIENCES 2026;32(1):26–37

RESEARCH ARTICLE

Comparison of 1D-Convolution Neural Network and Hilbert Huang Transform-Based Features for ECG Signal Classification

Baqir Ahmed Hussein *, Rabab Abdulrida Saleh

Department of Statistics, College of Administration and Economics, University of Baghdad, Iraq

Abstract

The electrocardiogram (ECG) signal is one of the primary biological signals used in the diagnosis of cardiac arrhythmias. However, conventional analysis and classification methods are often limited by the nonlinear and non-stationary nature of ECG signals. This research aims to evaluate the performance of a one-dimensional convolutional neural network (1D-CNN) and hybrid models that integrate Hilbert–Huang Transform (HHT)-based features with 1D-CNN for ECG signal classification.

Five standard diagnostic categories (N, S, V, Q, and F) were considered. A comparative framework was developed, including a standalone 1D-CNN model and three hybrid models based on different HHT-derived features. Model performance was assessed using confusion matrices, ROC-AUC curves, and key statistical metrics, including accuracy, precision, recall, and F1-score.

The results indicate that the standalone 1D-CNN model achieved the highest classification performance when applied to raw ECG signals. Among the hybrid models, the Mean_E + 1D-CNN model demonstrated superior performance compared to the other feature-based models. These findings suggest that while HHT features provide valuable time–frequency interpretability, direct learning from raw signals using deep learning models remains more effective for ECG classification.

Keywords: Electrocardiogram (ECG) signal, Deep learning, One-dimensional convolutional neural network (1D-CNN), Hilbert–Huang Transform (HHT), Nonlinear signal analysis, Instantaneous energy, Instantaneous frequency, Performance evaluation metrics

1. Introduction

The electrocardiogram (ECG) signal is considered one of the most important diagnostic tools in modern medicine, as it provides precise data reflecting the electrical activity of the heart and plays a central role in diagnosing cardiac arrhythmias and monitoring various pathological conditions. The widespread use of this signal in clinical and research applications has contributed to increasing interest in developing advanced techniques for its analysis and classification of its diverse cardiac patterns with high levels of accuracy and reliability (Ahmed et al., 2023).

However, the structural characteristics of the ECG signal, as a nonlinear and non-stationary signal, impose complex analytical challenges on conventional methods. This signal exhibits instantaneous variations in both amplitude and frequency, in addition to significant time–frequency interactions, as well as being affected by multiple sources of noise and interference resulting from biological interactions within the cardiac system or from digital recording devices. Recent literature indicates that classical statistical models and algorithms based on limited features often fail to represent the full dynamic behavior of these signals or to extract their underlying subtle discriminative characteristics (Tang, 2024).

Received 24 February 2026; revised 31 March 2026; accepted 7 April 2026.
Available online 30 April 2026

* Corresponding author.
E-mail addresses: baqer.ahmed2201m@coadec.uobaghdad.edu.iq (B. A. Hussein), rabab.saleh@coadec.uobaghdad.edu.iq (R. A. Saleh).

https://doi.org/10.33095/2227-703X.4342

In contrast, deep learning models, particularly convolutional neural networks, have emerged as effective tools in medical signal processing due to their ability to automatically learn features and detect complex nonlinear relationships within data. Nevertheless, exclusive reliance on the raw temporal signal may limit the physical interpretability of the results, highlighting the need to integrate these models with advanced signal processing techniques capable of providing a clearer representation of the signal's time–frequency characteristics (Zhao et al., 1051).

Based on these considerations, the present research proposes a hybrid analytical framework that integrates the Hilbert–Huang Transform (HHT) with a one-dimensional convolutional neural network (1D-CNN) to improve the classification accuracy of ECG signals. The Hilbert–Huang Transform is distinguished by its unique ability to analyze nonlinear and non-stationary signals through decomposing them into intrinsic mode functions (IMFs) and extracting time–frequency and spectral features with physical significance that reflect the instantaneous behavior of energy and frequency.

So, the research focuses on classifying ECG signals into five standard diagnostic categories: normal beat (N), supraventricular beat (S), ventricular beat (V), unclassified beat (Q), and fusion beat (F), which are categories adopted in global cardiac databases. To achieve this objective, a set of classification models was constructed, including a baseline model based on 1D-CNN using the raw signal, in addition to a hybrid model that employs features extracted from the Hilbert–Huang Transform as inputs to the convolutional neural network. A comprehensive set of statistical performance evaluation metrics was also adopted, including overall accuracy, precision, sensitivity, and F1-score, along with confusion matrices and ROC-AUC curves, in order to provide a thorough and accurate assessment of the reliability and efficiency of the proposed models in classifying electrocardiogram signals.

2. Literature review and hypothesis development

The analysis and classification of electrocardiogram (ECG) signals have received attention in recent studies due to their clinical importance and their fundamental role in diagnosing cardiac arrhythmias.

In this context, Kiranyaz et al. (2016) proposed a patient-specific 1D-CNN model for real-time ECG signal classification using the MIT-BIH database. Their results demonstrated significant superiority in terms of accuracy and computational efficiency compared with traditional classification algorithms such as SVM and KNN, thereby enhancing the feasibility of deploying the model in portable medical systems. Similarly, Chourasia et al. (2020) developed a 1D-CNN model that directly relied on the raw signal without conventional feature extraction, using a multilayer architecture to automatically learn patterns. The model achieved an accuracy of approximately 97.4%, particularly in distinguishing ventricular beats such as PVC and VEB.

Within practical applications, Cheikhrouhou et al. (2021) integrated a 1D-CNN model into a Fog-Cloud computing environment for real-time ECG signal processing using wearable devices. The results demonstrated the model's effectiveness in terms of accuracy, reduced response time, and lower energy consumption, making it suitable for continuous healthcare monitoring systems.

On the other hand, Ahmed et al. (2023) focused on arrhythmia classification using 1D-CNN with both real and noise-contaminated data, while applying techniques to reduce overfitting. Their models achieved accuracies exceeding 99% in certain settings, confirming the efficiency of one-dimensional convolutional networks in the rapid and accurate detection of cardiac disorders. Shang et al. (2025) further extended the application of 1D-CNN to home healthcare monitoring systems by developing a wearable-based framework for automatic ECG feature extraction. Their model achieved an accuracy approaching 98%, reinforcing the potential of remote healthcare services.

Similarly, Niu et al. (2023) employed HHT to extract instantaneous time–frequency components from ground signals, subsequently integrating these features with deep learning models such as CNN, LSTM, and attention networks. The findings confirmed that HHT significantly contributes to improving nonlinear pattern discrimination under various operating conditions. Tang (2024) further demonstrated that integrating HHT features with 1D-CNN for ECG signal classification achieved accuracies ranging between 97% and 99% in distinguishing different types of arrhythmias, emphasizing the importance of this integration when dealing with non-stationary signals.

Moreover, Telangore et al. (2024) reported that integrating HHT with CNN and Transformer models in early prediction of sudden cardiac death (SCD) improved sensitivity and overall accuracy to exceed 98%, confirming the importance of HHT in extracting complex instantaneous information.

Previous studies have shown that researchers have used both the 1D-CNN and HHT models, incorporating ECG data. While Tang's (2024) study employed the Hilbert-Huang transform with a 1D-

CNN network to improve ECG signal classification, it focused on only four classification categories and relied on data balancing to enhance hybrid model performance. However, a research gap remains regarding the evaluation of single and hybrid models within a comprehensive comparative framework when using the more complex medical classification system that includes five categories, as defined by the AAMI global standard. Furthermore, our research involved signal segmentation around R-peaks using fixed time windows to preserve the signal's morphological characteristics. In addition to using a single electrode channel of the MLII type only, which is the best and most medically used (Shang et al., 2025).

Unlike previous studies that focused on improving the hybrid model, this research aims to provide a systematic comparative evaluation between several models that rely on Hilbert-Huang transform features and the single model that relies on the raw signal, to examine whether extracting energy features actually adds classification value or leads to the loss of accurate temporal information, and the existence of results that differ from some research, this difference is to bridge the gap of some research not using the method of data segmentation around the peaks and not using the same database.

3. Research methodology

3.1. Hilbert–Huang Transform (HHT)

The Hilbert–Huang Transform (HHT) is considered one of the most efficient signal analysis tools when dealing with nonlinear and non-stationary signals, as it provides a high capability to represent time–frequency characteristics in an adaptive manner compared with traditional methods such as Fourier Transform and Wavelet Transform. HHT is fundamentally based on two interrelated stages: Empirical Mode Decomposition (EMD) and the Hilbert Transform (HT).

In the first stage, the original signal is decomposed into a set of Intrinsic Mode Functions (IMFs), which represent the fundamental components of the signal at different time scales. And in the second stage, Hilbert Transform is applied to each IMF individually to extract the instantaneous frequency and instantaneous energy, providing an accurate description of the dynamic behavior of the signal (Scozzese & Dall'Asta, 2024).

This analytical framework contributes to improving the understanding of complex phenomena in electrocardiogram (ECG) signals and constitutes a strong basis for extracting physically meaningful features that can later be employed in classification models.

3.1.1. Empirical mode decomposition (EMD)

The Empirical Mode Decomposition (EMD) method aims to reveal the intrinsic characteristics of a signal by decomposing it into a finite number of Intrinsic Mode Functions (IMFs), where each function represents a narrowband mono-component that varies with time. This method is fully adaptive and relies entirely on the nature of the data itself without prior assumptions regarding the signal's linearity or stationarity. For a function to be considered an Intrinsic Mode Function (IMF), two basic conditions must be satisfied (Huang & Shen, 2005):

a- The number of extrema and the number of zero crossings must either be equal or differ at most by one.

b- The local mean of the upper and lower envelopes must be equal to zero at every time point.

3.1.1.1. Sifting process

The EMD algorithm is implemented according to the following steps (Padmaja et al., 2011):

1. Identify all local maxima and minima of the original signal $x(t)$.
2. Construct the upper envelope $X_U(t)$ and lower envelope $X_L(t)$ using cubic spline interpolation (Waskito et al., 2010).
3. Compute the local mean between the two envelopes:

$$m_1(t) = \frac{X_U(t) + X_L(t)}{2} \tag{1}$$

4. Subtract the mean $m_1(t)$ from the original signal to obtain a proto-IMF:

$$h_1(t) = x(t) - m_1(t) \tag{2}$$

5. Verify whether h_1 satisfies the IMF conditions. If the conditions are not satisfied, the sifting process is repeated several times. In this stage, h_1 is treated as a new original signal and the previous steps are repeated iteratively until the IMF condition h_k is satisfied. The first IMF is written as c_1, and the first IMF is then obtained and denoted as:

$$c_1(t) = h_{1k}(t) \tag{3}$$

If $c_1(t)$ does not satisfy the IMF properties, the sifting process from step (1) is repeated. When the conditions are fulfilled, $c_1(t)$ is called the first Intrinsic Mode Function (IMF). The sifting process stops according to certain stopping criteria when the residual (The Residual) becomes either smaller than a predefined threshold or monotonic (the residual monotonic). The residual is extracted by separating the first component $c_1(t)$ from the original signal as follows:

$$r_1(t) = x(t) - c_1(t) \tag{4}$$

6. Steps (1) to (5) are repeated for the residual $r_1(t)$, since it still contains information about longer-period components. It is treated as new data and subjected to the same sifting process. This procedure is repeated for all subsequent residuals $r_n's$:

$$r_2(t) = r_1(t) - c_2(t)\ldots,\ r_n(t) = r_{n-1}(t) - c_{j-1}(t) \tag{5}$$

The final residuals are not always zero; therefore, if the data exhibits a trend, the final residual represents that trend. According to Eqs. (4) and (5), the original signal can be reconstructed precisely through linear superposition, where the input signal is decomposed into several IMFs $c_j(t)$ plus the final residual $r_n(t)$:

$$x(t) = \sum_{j=1}^{n} c_j(t) + r_n(t) \tag{6}$$

Thus, the EMD method decomposes complex signal data into a series of intrinsic oscillatory modes (IMFs) that reflect the signal characteristics, enabling complex signals such as ECG to be analyzed and filtered from unimportant noise data.

3.1.2. Hilbert transform (HT)

After obtaining the Intrinsic Mode Functions using EMD, the Hilbert Transform (HT) is applied to each function $c_j(t)$ independently to extract instantaneous signal characteristics (Bowman & Lees, 2013).

Hilbert Transform is expressed as follows (Houda, 2018):

$$\mathcal{H}_j(\omega, t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{c_j(t)}{t-\tau}\, d\tau \tag{7}$$

This process produces a complex analytical signal denoted as $sig(t)$, where the original IMF $c_j(t)$ represents the real part and its Hilbert Transform $\mathcal{H}_j(\omega, t)$ represents the imaginary part:

$$\begin{aligned} sig(t) &= \sum_{j=1}^{n} IMF_j(t) + i\mathcal{H}_j(\omega, t) \\ &= \sum_{j=1}^{n} a_j(t)\, e^{-i\theta_j(t)} + r_j(t) \end{aligned} \tag{8}$$

The instantaneous amplitude $a(t)$ and instantaneous phase $\theta(t)$ are computed as follows (Özdemir, 2014):

$$a(t) = \sqrt{c_j^2(t) + \mathcal{H}_j^2(\omega, t)} \tag{9}$$

$$\theta(t) = \arctan\frac{\mathcal{H}_j(\omega, t)}{c_j(t)} \tag{10}$$

The instantaneous frequency is obtained as the time derivative of the instantaneous phase:

$$\omega(t) = \frac{d\theta}{dt} \tag{11}$$

The instantaneous energy can be defined as:

$$\mathrm{IE}(t) = |a(t)|^2 = x^2(t) + y^2(t) \tag{12}$$

Since instantaneous energy is a function of time, it can be used to identify energy fluctuations in the time domain and detect real anomalies associated with the recorded signal (Kumar et al., 2025).

After extracting the features from the Hilbert–Huang Transform, these features are introduced into the 1D-CNN model to enhance its ability to distinguish between different ECG signal patterns. This integration benefits from the physical interpretability provided by HHT and the powerful automatic learning capability of convolutional neural networks in detecting complex patterns.

3.2. Deep learning (DL)

Deep Learning (DL) is considered one of the most advanced branches of artificial intelligence, as it relies on multi-layer neural networks capable of learning complex nonlinear representations from raw data. Deep learning has demonstrated remarkable superiority in digital signal processing applications due to its ability to detect fine patterns that are difficult for traditional or statistical methods to represent (Mozaffari & Tay, 2020).

3.2.1. One-dimensional convolutional neural network (1D-CNN)

Since ECG signals represent one-dimensional time series, one-dimensional convolutional neural networks are considered an appropriate choice for processing them directly without the need to convert them into two-dimensional images (2D). 1D-CNN is characterized by lower computational complexity compared to 2D-CNN networks, in addition to faster training and applicability in medical systems with

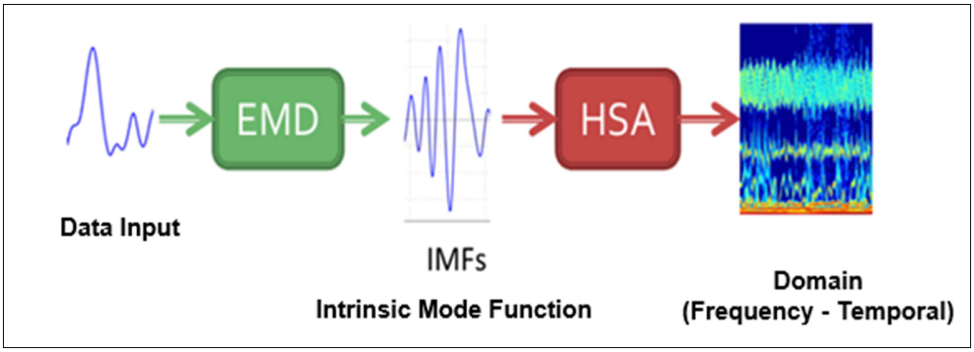

Fig. 1. Hilbert Huang transform (HHT) (Waskito et al., 2010).

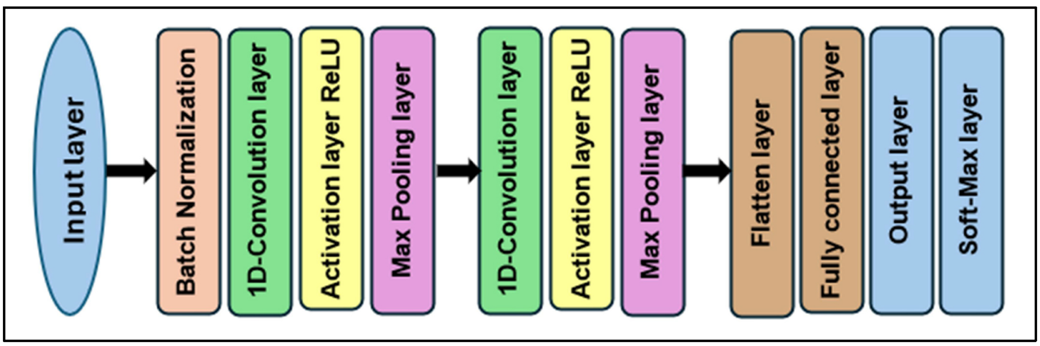

Fig. 2. Diagram of the one-dimensional convolutional neural network.
***Source:** Prepared by the researcher.*

limited resources. Fig. 2 illustrates the architecture of the 1D-CNN model (Kiranyaz et al., 2021).

3.2.2. Architecture of the 1-D convolutional neural network (Cheikhrouhou et al., 2021)

The architecture of the proposed network consists of three main stages: the feature extraction stage (Cheikhrouhou et al., 2021), the integration and classification stage, and the output stage. The details of 1D-CNN are as follows (Chourasia et al., 2020):

1. Input Layer:
The input layer receives signals, where each signal represents a one-dimensional vector that is directly fed into the convolutional layers.

2. Convolutional Layers:
Multiple convolutional layers are used to extract local patterns using filtering kernels that slide over the time signal. The convolution operation in the layer can be expressed as:

$$C_i^{lj} = \sigma\left(b_j + \sum_{m=1}^{M} W_m^j x_{i+m-1}^j\right) \tag{13}$$

where W_m^j represents the kernel weights, b_j is the bias term, and σ is the activation function.

3. Activation Function (ReLU):
The Rectified Linear Unit (ReLU) activation function is used to introduce nonlinearity and improve convergence speed during training. It sets negative outputs to zero while positive values remain unchanged. The formulation is:

$$\sigma = \max(0, x) \tag{14}$$

4. Pooling Layers:
Pooling layers aim to reduce the dimensions of feature maps while preserving the most important values. Max-Pooling was adopted according to the following formulation:

$$y_k = \max\ (x_k) \tag{15}$$

5. Fully Connected Layers:
After flattening the feature maps (Flatten), the data are passed to fully connected layers that integrate the extracted features in preparation for the final classification stage.

6. Output Layer and Softmax Function:
The Softmax function is used in the output layer to convert outputs into probability distributions over the five diagnostic classes, according to the equation

(Mattioli et al., 2021):

$$P_i = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \tag{16}$$

where P_i represents the probability that the signal belongs to class.

The resulting values fall within the range [0,1], i.e., $0 \le P_i \le 1$,and their sum equals one ($\sum_{i=1}^{n} P_i = 1$).

The exponentiation step (e^{x_i}) amplifies differences between input values, emphasizing larger values and reducing the importance of smaller ones. Consequently, the class with the highest probability is considered the predicted class for the classification process (Cacciari & Ranfagni, 2024).

3.2.3. Training and optimization algorithm

During the training phase of artificial intelligence and deep learning models, the backpropagation algorithm is commonly used along with an optimization method known as Adam (Adaptive Moment Estimation). Adam is a highly efficient stochastic optimization algorithm (Reyad et al., 2023). It gained popularity due to its empirical acceleration compared with simple stochastic gradient descent (SGD) (Tian et al., 2023), and it computes adaptive learning rates for each parameter, making it computationally efficient and memory-efficient.

Adam maintains two moving averages: the first moment estimate (mean of gradients) and the second moment estimate (variance of gradients). These estimates are initialized as zero vectors.

At each iteration, Adam computes the gradient of the loss function with respect to the parameters, updates the moving averages of gradients and squared gradients, and updates the weights according to the following equations:

$$g_t = \nabla_{\mathrm{w}_t - 1} - \mathcal{L}_t\left(\mathrm{w}_t - 1\right) \tag{17}$$

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\, g_t \tag{18}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)\, {g_t}^2 \tag{19}$$

$$\widehat{m_t} = \frac{m_t}{1 - {\beta_1}^t} \tag{20}$$

$$\widehat{v_t} = \frac{v_t}{1 - {\beta_2}^t} \tag{21}$$

$$\Delta W_t = -\alpha \frac{\widehat{m_t}}{\sqrt{\widehat{v_t}} + \varepsilon} \tag{22}$$

$$W_t = W_{t-1} + \Delta W_t \tag{23}$$

where:

g_t: the gradient of the loss with respect to the weight W_t

m_t: the first moment estimate (mean of gradients)

v_t: the second moment estimate (variance of gradients)

$\widehat{m_t}$: bias-corrected first moment estimate m_t

$\widehat{v_t}$: bias-corrected second moment estimate v_t

β_1: exponential decay rate for the first moment (commonly 0.9)

β_2: exponential decay rate for the second moment (commonly 0.999)

ε: small constant to avoid division by zero (usually 10^{-8})

α: learning rate (commonly 0.001)

3.2.4. Performance evaluation

In any mathematical model or artificial intelligence-based system, the efficiency of the 1D-CNN model is determined by using key evaluation metrics for ECG signal classification. These indicators reflect the model's performance in handling multiclass classification problems (Ahmed et al., 2023).

After obtaining the results, the Confusion Matrix and the ROC-AUC curve are used. The confusion matrix compares the true class labels with predicted labels. From it, the following values are obtained:

True Positive (TP): correctly predicted positive cases

True Negative (TN): correctly predicted negative cases

False Positive (FP): incorrectly predicted positive cases

False Negative (FN): incorrectly predicted negative cases

Based on these, the following evaluation metrics are calculated:

1. Accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{24}$$

2. Precision:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{25}$$

3. Recall (Sensitivity):

$$\text{Recall} = \frac{TP}{TP + FN} \tag{26}$$

4. F1-score:

$$\text{F1} - \text{score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{27}$$

The harmonic mean of precision and recall.

5. Specificity:

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{28}$$

6. AUC-ROC:

AUC represents the area under the ROC curve, which illustrates the trade-off between true positive rate and false positive rate. The mathematical expression is:

$$AUC = \frac{1}{2}\left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right) \tag{29}$$

4. Results

4.1. Analytical results in the practical application

In this research, data analysis and programming were performed using the Python programming language, specifically on the Colab platform on the Internet. The programmer used this platform because it provides sufficient space and cloud memory equipped with a GPU that efficiently processes code and executes complex operations at high speed. The classification accuracy results of the standalone 1D-CNN model, including its architecture and parameters, will be presented. In addition, the results of the hybrid HHT + 1D-CNN model features will also be presented, along with graphical representations for each model.

4.2. Data description

ECG signal data recorded from the MLII lead for 46 patients were used. The data were obtained from the PhysioNet website, specifically from the PhysioBank ATM section of the MIT-BIH Arrhythmia Database (Moody & Mark, 2001). The ECG signals recorded from the MLII lead were segmented into fixed-length time windows of 216 samples centered around the R peak (R−108, R+108), The length of 216 samples was selected based on previous studies that rely on window splitting around the R-top to fully contain the QRS compound, but they used different split lengths by a small margin.

The total number of windows reached 107,707 (each window representing one row). These windows were also analyzed using the Hilbert–Huang Transform to decompose the complex signal into several intrinsic mode functions (We kept the data of the categories without using the process of balancing them to obtain realistic and natural medical results without changing the nature of their composition).The independent variable (X) represents the ECG signal windows (columns), where each window consists of 216 features. The dependent variable (Y) represents the diagnostic class, which belongs to five categories encoded as: (Q=0, N=1, S=2, V=3, F=4); (N = Normal beat, S = Supraventricular beat, V = Ventricular beat, Q = Unclassified beat, F = Fusion beat). The percentage of training and testing data in the single model and hybrid models was (70% training, 30% testing)

4.3. Practical application in the standalone model (1D-CNN)

A one-dimensional convolutional neural network (1D-CNN) was used to automatically extract features from the signals, while the Softmax function was used in the output layer to perform multiclass classification. Hyperparameters were adjusted for the training and testing processes.

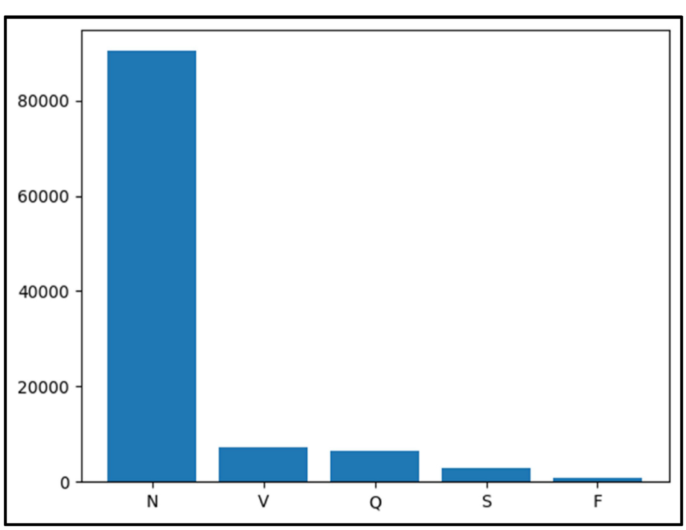

Fig. 3. Graphical columns describing the five diagnostic heartbeat categories.
Source: *Prepared by the researcher.*

Table 1. Details of the hyperparameters used in the 1D-CNN model.

Hyperparameters	Values
No. of classes	5
Batch Size	32
No. of Epochs	100
ConvlD	
Filters	16
Kernels	3
Regularization!	0.01
Activation	ReLU
Max_pooling	2
ConvlD	
Filters	32
Kernels	3
Regularization	0.01
Activation	ReLU
Max_pooling	2
Optimizer used	Adam
FC Layer	512 neurons

Source: Prepared by the researcher.

Predicted Classes

True Classes	Q	N	S	V	F
Q	1884	38	5	29	1
N	30	26930	112	39	1
S	1	282	547	4	0
V	14	140	12	1983	20
F	0	73	0	18	150

Fig. 4. Confusion Matrix of the 1D-CNN Model for Arrhythmia Classification.
Source: *Prepared by the researcher.*

The predicted outputs were compared with the true values using the Categorical Cross-Entropy loss function, and the model performance was evaluated using Accuracy, Recall, F1-score, and the Confusion Matrix. Fig. 4 shows the confusion matrix based on the test data, illustrating the distribution of correct and incorrect predictions for each class. The values on the main diagonal represent correctly classified samples, whereas the off-diagonal values represent misclassified samples.

The classes N, Q, and V maintained high performance with large values on the main diagonal, indicating model stability. In contrast, confusion remained noticeable in classes S and F, where most misclassifications were directed toward class N.

Table 2. Confusion matrix report of the 1D-CNN model.

Five Classes	Performance Evaluation			
	Accuracy	Precision	Recall	Fl-score
Q	99.63%	97.7%	96.3%	97.0%
N	97.79%	98.1%	99.3%	98.7%
S	98.71%	80.9%	65.6%	72.5%
V	99.15%	95.7%	91.4%	93.5%
F	99.65%	87.2%	62.2%	72.6%

Source: Prepared by the researcher.

This behavior confirms that the main issue does not lie in the overall weakness of the model, but rather in the difficulty of distinguishing between certain pathological patterns that are close in temporal characteristics and morphological features of ECG signals.

From the results in Table 2, which present the main evaluation metrics used to assess classification performance on the test data:

For class Q, the model achieved a very high overall accuracy of 96.3% in testing, with high values for the remaining metrics: Precision = 97.7%, Recall = 99.63%, F1-score = 97.0%.

This reflects the model's ability to accurately recognize this class and maintain stability when transitioning from training data to previously unseen data.

Class N (normal beats) recorded the lowest performance among all classes. Precision = 98.7% was slightly lower than Recall. This indicates the presence of some cases classified as normal despite belonging to other classes. This result is expected since this class is the most represented in the dataset, allowing the model to effectively learn its characteristics, with possible class overlap favoring the majority class.

For class S, good performance was observed, with an overall test accuracy of 98.71%. However, Recall reached 65.6%, while Precision remained relatively high at 80.9%. This behavior indicates that the model is relatively accurate when predicting this class but fails to detect a significant portion of actual cases, leading to a decrease in sensitivity. For class V, strong results were achieved: Recall = 91.4% and Precision = 95.7%.

Precision being higher than Recall indicates that when the model predicts this class, it is usually correct, but the number of detected cases remains slightly lower compared to the total actual cases. This reflects the model's good ability to distinguish ventricular beats despite their relative complexity.

Finally, class F recorded the highest overall performance among all classes, with test accuracy of 99.65%, but Recall was relatively low at 62.2%, while Precision reached 87.2%. This means that the model avoids false predictions for this class, but overlooks a non-negligible portion of actual cases, which is reflected

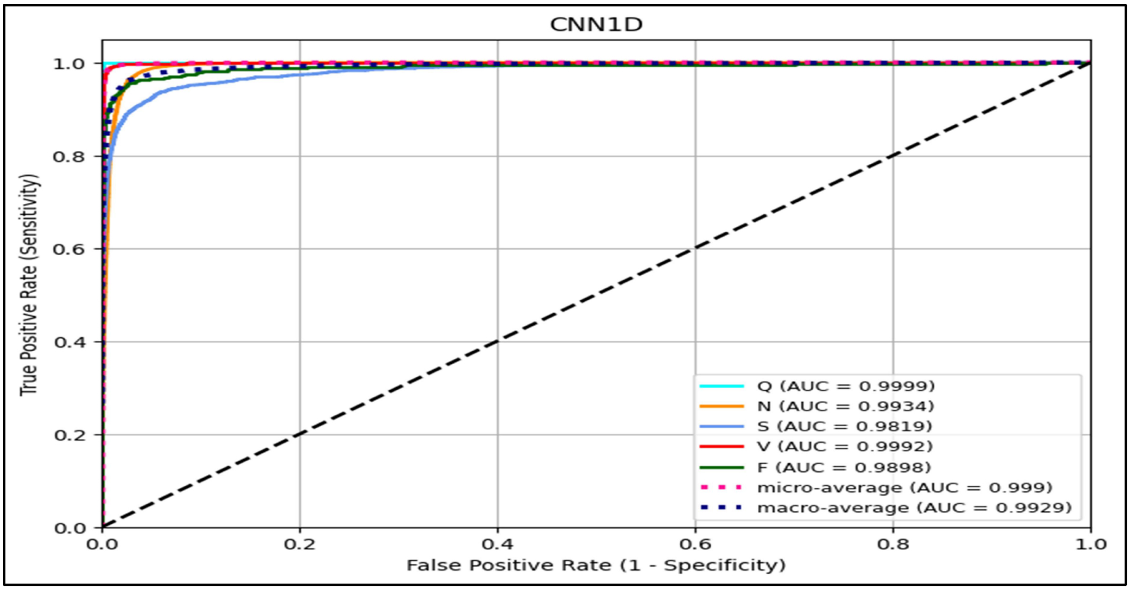

Fig. 5. ROC curve of the 1D-CNN model based on the test set.
***Source:** Prepared by the research.*

in the F1-score value of 72.6%. This clearly reveals a moderate imbalance in the dataset.

4.4. Practical application in the hybrid model (HHT + 1D-CNN)

The Hilbert–Huang Transform was used to extract features. Three features were extracted:

Mean_E/IMF (mean instantaneous energy for all IMFs)

Median_E/IMF (median instantaneous energy of all IMF)

Mean_f/IMF (mean instantaneous frequency for all IMFs)

Each feature was considered a single time-series variable and then input into the one-dimensional convolutional neural network (1D-CNN). Three hybrid models were obtained by combining HHT features with 1D-CNN to classify the five heartbeat classes.

In the following figure, one window was selected from among the window numbers (6500) as an example to illustrate the intrinsic mode functions (IMFs) in the Empirical Mode Decomposition (EMD) algorithm, as well as the original signal shape.

In Fig. 6, one of the windows is shown, where the horizontal axis represents the sample and the vertical axis represents amplitude. In Fig. 7, it is observed that IMF1 had the highest oscillations, which

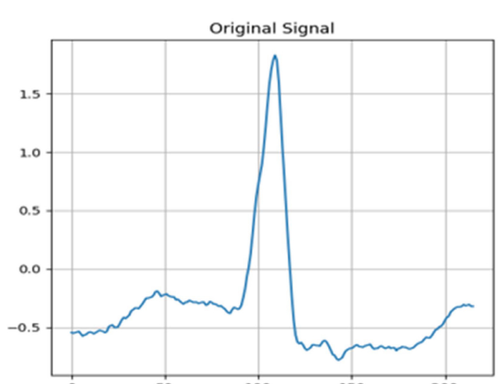

Fig. 6. Original signal in one of the windows.
***Source:** Prepared by the research.*

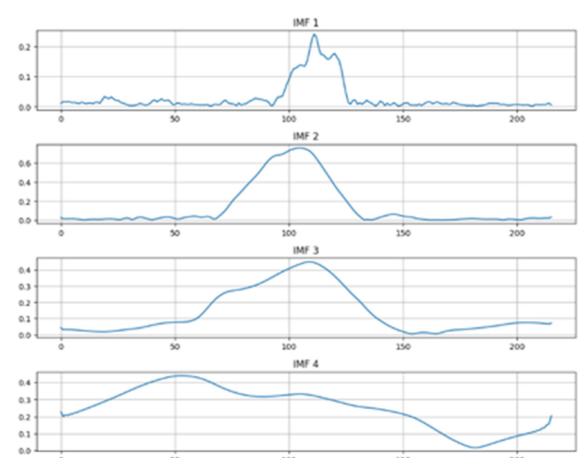

Fig. 7. Intrinsic mode functions (IMFs).
***Source:** Prepared by the research.*

decreased in IMF2, changed slightly in IMF3, and became smoother in IMF4.

Performance metrics and confusion matrices were extracted for each hybrid model. For general comparison between the standalone and hybrid models, a comparison table of performance metrics was constructed to determine the best model in terms of high classification accuracy for the five heartbeat classes.

4.5. General comparison of standalone and hybrid models

Table 3 presents a comprehensive comparison between the standalone model (1D-CNN) and the hybrid models based on HHT features. The main performance metrics (Accuracy, Precision, Recall, and F1-score) were used based on the test dataset.

Table 3. Performance evaluation report of the standalone and hybrid models.

Models Comparison	Performance Evaluation			
	Accuracy	Precision	Recall	Fl-score
1D-CNN	97.47%	97.35%	97.47%	97.36%
Mean_E	92.50%	91.69%	92.50%	91.17%
Median_E	90.54%	89.21%	90.54%	88.63%
Mean_f	90.34%	88.79%	90.34%	88.84%

Source: Prepared by the researcher.

The results clearly show the superiority of the standalone 1D-CNN model, which achieved the highest classification accuracy among all models. Using a single evaluation criterion (test Accuracy), it reached 97.47%, indicating high generalization capability, stable performance, and resistance to overfitting.

This superiority reflects the efficiency of the convolutional neural network in directly handling the raw temporal ECG signal and learning detailed morphological patterns without the need for manually engineered features that may lose part of the information.

In contrast, the hybrid models based on HHT features showed varying performance. The Median_E + 1D-CNN model achieved a test accuracy of 90.54%, which is relatively good but still lower than the standalone model. This may be attributed to relying on the instantaneous energy of the first IMF only, which focuses on specific frequency components and does not reflect the complete signal structure.

On the other hand, the Mean_E + 1D-CNN model achieved the best performance among hybrid models, with a test accuracy of 92.50%. This indicates that averaging instantaneous energy across all IMFs provides a more comprehensive representation of the signal compared to relying on a single IMF.

Meanwhile, the Mean_f + 1D-CNN model recorded a test accuracy of 90.34%, indicating that the mean instantaneous frequency is useful in describing the general rhythm of the signal but remains less efficient than comprehensive energy features due to lower sensitivity to sharp morphological changes in ECG signals.

Overall, although all HHT-based hybrid models achieved acceptable accuracy, none surpassed the standalone 1D-CNN model. The superiority of the standalone model can be explained by its ability to exploit the complete morphological structure of the heartbeat components (P–QRS–T), which carry precise diagnostic information associated with the temporal relationships between waves.

Furthermore, segmenting signals into windows centered around R peaks led to temporal alignment and reduced inter-sample variability, making structural features clearer and more directly learnable by the 1D-CNN model.

In contrast, reducing the signal into individual energy or frequency features derived from HHT may reduce the informational dimension and lose some fine temporal relationships between beat components, which explains the lower sensitivity of hybrid models compared to direct learning from raw signals, especially under class imbalance conditions.

5. Conclusion

The research's findings showed that, in comparing hybrid models based on Hilbert–Huang Transform (HHT) features, the one-dimensional convolutional neural network (1D-CNN) model based on the raw temporal signal achieved better classification performance and greater stability in classifying ECG heartbeats. The model's excellent capacity to learn intricate temporal ECG data patterns without the requirement for human feature extraction is demonstrated by the overall accuracy of 97.47% and the F1-score of 97.36%.

The results of this research demonstrated that the one-dimensional convolutional neural network (1D-CNN) model based on the raw temporal signal achieved superior classification performance and greater stability in classifying ECG heartbeats compared to the hybrid models based on Hilbert–Huang Transform (HHT) features. The overall accuracy reached 97.47%, and the F1-score reached 97.36%, indicating the model's high capability to learn fine temporal patterns of ECG signals without the need for manual feature extraction.

The comparison showed that the hybrid models based on Hilbert–Huang Transform (HHT) features, despite their effectiveness in describing the time–frequency structure of the signal, did not achieve additional performance improvement compared to the standalone model. The accuracy of the Mean_E

+ 1D-CNN model decreased to 92.50%. This suggests that reducing the signal to energy features derived from IMFs may lead to the loss of important temporal information that has a more direct impact on the classification process.

In contrast, the results of the Median_E + 1D-CNN model showed that relying on the IMF energy median, although achieving relatively acceptable performance, does not reflect the complete temporal structure of the signal, which resulted in a noticeable reduction in the ability to distinguish certain pathological classes. The results of the Mean_f + 1D-CNN model also indicated that the mean instantaneous frequency is capable of describing the general pattern of the signal; however, it is less sensitive to temporal variations associated with cardiac arrhythmias.

The comparison among the hybrid models based on HHT features clarified that the Mean_E + 1D-CNN model achieved the best performance among them, reaching an overall accuracy of 92.50%, outperforming the Median_E + 1D-CNN and Mean_f + 1D-CNN models. This indicates that the mean instantaneous energy across all IMFs provides a more comprehensive representation of the signal's energy content compared to relying on a single IMF or only frequency characteristics.

Future research is recommended to adopt the 1D-CNN model based on raw ECG signals due to its superior performance, while improving HHT-based hybrid models through integrating richer energy, frequency, and temporal features. It is also important to address class imbalance systematically, adopt patient-independent data splitting, and explore alternative architectures such as LSTM and GRU. In addition, further optimization of the 1D-CNN architecture and evaluation of alternative signal representations, including wavelet-based approaches, are encouraged. Finally, extending these methods to real-time clinical applications and other signal domains represents a promising direction for future work.

Acknowledgment

This work was supported by the Physionet Site for the MIT-BIH Arrhythmia Database Disorder Database. We express our appreciation for providing the necessary data.

Author contributions

The first and second authors contributed to the research, and both read and approved the published version of the manuscript.

Funding statement

This research did not receive any funding from public, private, or non-profit organizations.

Conflicts of interest

The author declares that they have no conflict of interest.

Authors' declaration

We Hereby Confirm That All the Figures and Tables in The Manuscript Are Mine and Ours. Besides, The Figures and Images, which are Not Mine, Have Been Permitted Republication and Attached to The Manuscript.

Ethical Clearance: Research Was Approved by The Local Ethical Committee in The University of Baghdad.

Data availability

Raw data is available at the source (Moody & Mark, 2001).

References

Ahmed, A. A., Ali, W., Abdullah, T. A. A., & Malebary, S. J. (2023). Classifying Cardiac Arrhythmia from ECG Signal Using 1D CNN Deep Learning Model. *Mathematics*, *11*(3), 562. https://doi.org/10.3390/math11030562

Bowman, D. C., & Lees, J. M. (2013). The Hilbert-huang transform: A high resolution spectral method for nonlinear and nonstationary time series. *Seismological Research Letters*, *84*(6), 1074–1080. https://doi.org/10.1785/0220130025

Cacciari, I., & Ranfagni, A. (2024). Hands-on fundamentals of 1D convolutional neural networks—A tutorial for beginner users. *Applied Sciences*, *14*(18), 8500. https://doi.org/10.3390/app14188500

Cheikhrouhou, O., Mahmud, R., Zouari, R., Ibrahim, M., Zaguia, A., & Gia, T. N. (2021). One-dimensional CNN approach for ECG arrhythmia analysis in fog-cloud environments. *IEEE.Access*, *9*, 103513–103523. https://doi.org/10.1109/ACCESS.2021.3097751

Chourasia, M., Thakur, A., Gupta, S., & Singh, A. (2020). ECG Heartbeat Classification Using CNN. *2020 IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, 1–6. https://doi.org/10.1109/UPCON50219.2020.9376451

Houda, H. (2018). Intelligent support system for ground penetrating radar data inversion [The University of Sidi Mohammed Ben Abdellah]. http://toubkal.imist.ma/handle/123456789/11898

Huang, N. E., & Shen, S. S. P. (2005). Hilbert-Huang transform and its applications. In Hilbert-Huang Transform And Its Applications. *World Scientific Publishing Co*. https://doi.org/10.1142/5862

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., & Inman, D. J. (2021). 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, *151*, 107398. https://doi.org/10.1016/j.ymssp.2020.107398

Kiranyaz, S., Ince, T., & Gabbouj, M. (2016). Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks. *IEEE Transactions on Biomedical Engineering*, *63*(3), 664–675. https://doi.org/10.1109/TBME.2015.2468589

Kumar, H. K. P., Dash, R., Reddy, K. J., Atyam, N. R., & Mohapatra, B. (2025). High resistance fault detection in DC microgrid using Hilbert-Huang transform and vector-based ensemble optimized LSTM networks. *Scientific Reports*, *15*(1), 16039. https://doi.org/10.1038/s41598-025-00977-5

Lakhdari, K., & Saeed, N. (2022). A new vision of a simple 1D Convolutional Neural Networks (1D-CNN) with Leaky-ReLU function for ECG abnormalities classification. *Intelligence-Based Medicine*, *6*, 100080. https://doi.org/10.1016/j.ibmed.2022.100080

Mattioli, F., Porcaro, C., & Baldassarre, G. (2021). A 1D CNN for high accuracy classification and transfer learning in motor imagery EEG-based brain-computer interface. *Journal of Neural Engineering*, *18*(6), 066053. https://doi.org/10.1088/1741-2552/ac4430

Moody, G. B., & Mark, R. G. (2001). The impact of the MIT-BIH arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*, *20*(3), 45–50. https://doi.org/10.1109/51.932724

Mozaffari, M. H., & Tay, L.-L. (2020). A Review of 1D Convolutional Neural Networks toward Unknown Substance Identification in Portable Raman Spectrometer. https://doi.org/10.48550/arXiv.2006.10575

Niu, J., Li, H., Liu, Z., Liu, W., & Xu, H. (2023a). Robot Ground Media Classification Based on Hilbert–Huang Transform and Attention-Based Spatiotemporal Coupled Network. *IET Signal Processing*, *2023*(1). https://doi.org/10.1049/2023/4721508

Özdemir, S. (2014). Analysis of GNSS time series obtained from Turkish national permanent GNSS stations network-active system using Hilbert-Huang transform /[The graduate school of natural and applied sciences of middle east technical university]. https://open.metu.edu.tr/handle/11511/23599

Padmaja, Varadarajan, & Swathi. (2011). Signal Processing of Radar Echoes Using Wavelets and Hilbert Huang Transform. *Signal & Image Processing: An International Journal*, *2*(3), 101–119. https://doi.org/10.5121/sipij.2011.2310

Shang, H., Yu, S., Wu, Y., Liu, X., He, J., Ma, M., Zeng, X., & Jiang, N. (2025). A noninvasive hyperkalemia monitoring system for dialysis patients based on a 1D-CNN model and single-lead ECG from wearable devices. *Scientific Reports*, *15*(1), 2950. https://doi.org/10.1038/s41598-025-85722-8

Scozzese, F., & Dall'Asta, A. (2024). Nonlinear Response Characterization of Post-Tensioned R.C. Bridges through Hilbert–Huang Transform Analysis. *Structural Control and Health Monitoring*, *2024*(1). https://doi.org/10.1155/2024/5960162

Tang, T. (2024). Classification of ECG Signals Based on Hilbert-Huang Transform and 1D Convolution Neural Network. *Advances in Engineering Technology Research*, *10*(1), 695. https://doi.org/10.56028/aetr.10.1.695.2024

Telangore, H., Azad, V., Sharma, M., Bhurane, A., Tan, R. S., & Acharya, U. R. (2024). Early prediction of sudden cardiac death using multimodal fusion of ECG Features extracted from Hilbert–Huang and wavelet transforms with explainable vision transformer and CNN models. *Computer Methods and Programs in Biomedicine*, *257*, 108455. https://doi.org/10.1016/j.cmpb.2024.108455

Waskito, P., Miwa, S., Mitsukura, Y., & Nakajo, H. (2010). Parallelizing Hilbert-Huang Transform on a GPU. 2010 First International Conference on Networking and Computing, 184–190. https://doi.org/10.1109/IC-NC.2010.44

Zhao, W., Shi, R., Tuo, X.-G., Zheng, H.-L., Yang, G., Wang, B., & Liu, M.-Z. (2023). Novel radionuclides identification method based on Hilbert–Huang Transform and Convolutional Neural Network with gamma-ray pulse signal. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, *1051*, 168232. https://doi.org/10.1016/j.nima.2023.168232