Eduvest – Journal of Universal Studies
Volume 4 Number 12, December, 2024
p- ISSN 2775-3735- e-ISSN 2775-3727

ARRHYTHMIA CLASSIFICATION USING CNN-SVM FROM ECG SPECTROGRAM REPRESENTATION

Abdul Daffa Fakhrudin[1], Putu Harry Gunawan[2]
[1,2] Fakultas Informatika, Universitas Telkom, Bandung, Indonesia
Email: abduldaffa@student.telkomuniversity.ac.id, phgunawan@telkomuniversity.ac.id

ABSTRACT

Arrhythmia, a critical subset of cardiovascular diseases and a leading cause of morbidity and mortality, is caused by irregular heartbeats that disrupt the normal rhythm of the heart. Detecting arrhythmias accurately is essential for timely diagnosis and treatment, which can be achieved through electrocardiogram (ECG) signals. This study presents a hybrid Convolutional Neural Network (CNN) and Support Vector Machine (SVM) model for arrhythmia classification, leveraging spectrogram representations of ECG signals. The CNN extracts spatial and temporal features from the spectrograms, while the SVM classifies five arrhythmia classes: Normal (N), Supra-ventricular premature (S), Ventricular escape (V), Fusion of ventricular and normal (F), and Unclassified (Q). Preprocessing techniques such as wavelet denoising and Short-Time Fourier Transform (STFT) were applied to improve signal quality and facilitate robust feature extraction. The proposed model was trained and evaluated on the MIT-BIH Arrhythmia Database, achieving a weighted F1-score of 0.985, demonstrating its ability to handle the imbalanced dataset effectively. Class-wise metrics highlighted high precision, recall, and F1-scores for majority classes and commendable performance for underrepresented classes, despite the inherent imbalance. These findings underscore the hybrid model's potential for arrhythmia classification by integrating the feature extraction strengths of CNNs with the precise classification capabilities of SVMs. Future research could address dataset imbalance through augmentation techniques and explore the model's generalizability by testing on larger and more diverse datasets, paving the way for its application in real-world clinical scenarios.

KEYWORDS *CNN-SVM, Spectrogram, STFT, Arrhythmia*

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International

How to cite:	Abdul Daffa Fakhrudin, Putu Harry Gunawan. (2024). Arrhythmia Classification Using CNN-SVM from ECG Spectrogram Representation. *Journal Eduvest. 4*(12): 11245-11254
E-ISSN:	2775-3727

INTRODUCTION

Cardiovascular Diseases (CVDs) are a group of conditions that affect the heart and vascular system. CVDs have become the leading cause of death worldwide, presenting a major challenge in the field of healthcare (Kementerian Kesehatan Republik Indonesia, n.d.). In 2019, it was estimated that around 17.9 million people died from CVDs, accounting for 32% of the total deaths globally (World Health Organization, n.d.).

Arrhythmia is a group of diseases that are an important and significant part of CVDs. Arrhythmia is caused by irregular heartbeats and occurs when the electrical impulses from the nervous system, which coordinate heartbeats, do not function properly. This results in the heart beating too fast, too slow, or even skipping beats (Huang et al., 2019; Neha et al., 2021). Irregular heartbeats can be detected by analyzing electrocardiogram (ECG) signals, an ECG being a test that generates medical signals by measuring and recording the electrical impulses of the heart (Huang et al., 2019). Therefore, research into classifying arrhythmia from ECG signals can contribute to reducing cases of CVDs.

ECG signals exhibit shape variations due to individual differences, physical conditions, and the equipment used to record the electrical impulses of the heart. Therefore, a data representation of ECG is needed that can capture the essential features of the ECG signal without being affected by these factors. One method to represent ECG data is by using spectrogram images. A spectrogram is a 2-dimensional representation of a signal that is represented in the time-frequency domain (Arpitha et al., 2022; Huang et al., 2019). The advantage of spectrograms for representing ECG signals lies in their ability to depict how frequency changes over time, which helps capture essential features of the ECG signal (Arpitha et al., 2022).

To classify arrhythmia from ECG data that has been transformed into 2-dimensional spectrogram images, an effective image data processing method is required. In this study, the authors will use a hybrid method combining Convolutional Neural Networks (CNN) and Support Vector Machines (SVM), known as the CNN-SVM method. In this hybrid method, CNN is used to extract features from images by leveraging a series of layers that contain convolutional and pooling operations, allowing the model to automatically and efficiently learn and recognize complex patterns in the images. SVM is then used to classify these features by finding the best decision boundary, or hyperplane, to separate the features extracted by CNN into their respective classes (Unlersen et al., 2022). In this study, the authors selected five arrhythmia classes based on the EC57 standard from the Association for the Advancement of Medical Instrumentation (AAMI) (Gai, 2022). This standard ensures a consistent and clinically relevant categorization of arrhythmia classes, facilitating robust evaluation and comparison across studies. The classes are Normal heartbeats (N), Supra-ventricular premature (S), Ventricular escape (V), Fusion of ventricular and normal (F), and Unclassified heartbeats (Q). Thus, the combination of CNN and SVM will allow the recognition and classification of complex patterns in the spectrogram representation of ECG data.

Previous research on arrhythmia classification has utilized various techniques, such as 2D-CNN with STFT (Huang et al., 2019) and hybrid CNN-LSTM with CWT (Madan et al., 2022), demonstrating the potential of deep learning methods for ECG signal classification. Although Ramos et al. (Ramos et al., 2021) applied the CNN-SVM hybrid model in a different domain (classifying images of plants and weeds), their results showed that this method effectively combined feature extraction and classification with high accuracy. Building on these findings, this research will apply the CNN-SVM hybrid model with STFT to the domain of arrhythmia classification, aiming to improve performance by leveraging the strengths of both techniques.

The next chapters of this paper are structured as follows. Section II outlines the approach and techniques used in this study. Section III presents the findings of the study. Finally, Section IV summarizes the conclusion of this research

RESEARCH METHOD

Classification System Design

The classification system implemented in this study follows the design illustrated in Fig. 1.

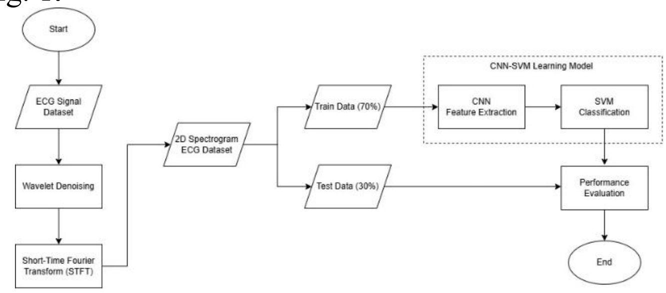

Fig. 1. Workflow of the CNN-SVM classification system.

The proposed system leverages a hybrid CNN-SVM approach to classify arrhythmias from ECG signals. The process begins with ECG signal acquisition from the MIT-BIH dataset. The raw signals undergo preprocessing steps, including wavelet denoising to remove noise and Short-Time Fourier Transform (STFT) to convert the signals into 2D spectrogram representations. These spectrograms capture the time-frequency characteristics of ECG signals, providing rich feature representations for classification.

The dataset is split into training (70%) and testing (30%) subsets. The training data is passed to a CNN-SVM learning model, where CNN extracts high-level spatial features from the spectrograms. These features are then classified by an SVM, which excels at identifying decision boundaries between classes.

Finally, the model is evaluated on the test dataset using metrics such as Accuracy, precision, recall, and F-1 Score.

MIT-BIH Dataset

The dataset used in this study is the MIT-BIH Arrhythmia Database (The Impact of the MIT-BIH Arrhythmia Database, n.d.), a widely recognized benchmark for arrhythmia classification research. It contains annotated ECG recordings from 47 individuals, each with two leads. For this study, the MLII lead was selected due to its frequent use and reliability in arrhythmia detection. The dataset was accessed through the PhysioNet platform (*PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals,* n.d.),which provides comprehensive resources and tools for analyzing physiological signals.

To align with the AAMI EC57 standard, the dataset was filtered to include ECG beats corresponding to five arrhythmia classes: Normal (N), Supra-ventricular premature (S), Ventricular escape (V), Fusion of ventricular and normal (F), and Unclassified (Q). Each beat in the dataset was labeled and categorized into one of these classes.

The dataset was split into training and testing subsets, with 70% of the data allocated for training and 30% for testing. The distribution of arrhythmia classes in the dataset is shown in Fig. 2, highlighting a significant imbalance across classes. While this study does not apply techniques to address the imbalance during training, weighted F1-scores were used during evaluation to provide a fair assessment of model performance. Future work will explore methods such as class weighting, oversampling, and focal loss to improve the handling of minority classes.

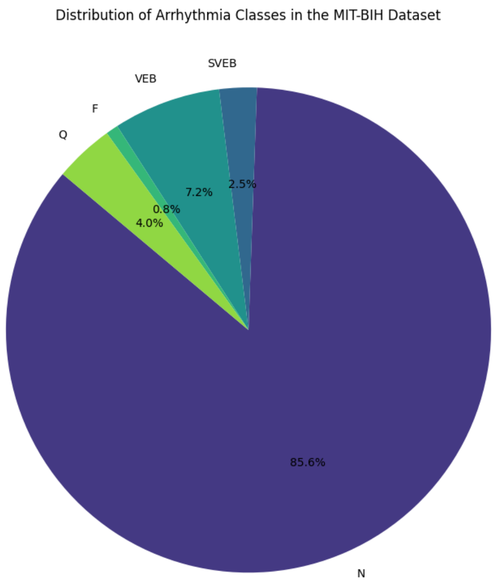

Fig. 2. Distribution of arrhythmia classes in the MIT-BIH dataset

Preprocessing Data

Preprocessing is a crucial step in ensuring that the raw ECG signals are clean, standardized, and suitable for feature extraction. In this study, the following preprocessing techniques were applied:

Wavelet Denoising: The raw ECG signals were denoised using Discrete Wavelet Transform (DWT) with the 'db5' wavelet. This method effectively removed high-frequency noise while preserving essential signal features, such as the QRS complex. A soft thresholding technique was applied to wavelet coefficients, suppressing noise components without distorting the signal. Fig. 3 shows a comparison of a raw ECG signal and its denoised counterpart, demonstrating the effectiveness of wavelet denoising.

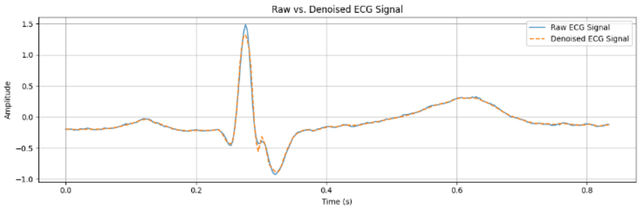

Fig. 3. Raw vs. denoised ECG signal

Normalization: The denoised signals were normalized to zero mean and unit variance to standardize the data. This step ensures that the signals are scaled consistently, improving the stability and convergence of the CNN model during training.

Spectrogram Generation: The normalized 1D ECG signals were transformed into 2D spectrograms using Short-Time Fourier Transform (STFT). A window size of 64 samples with an overlap of 32 samples was used to balance time and frequency resolution. Spectrograms provide a time-frequency representation of the ECG signals, highlighting the variation of frequency components over time. An example spectrogram generated from an ECG signal is shown in Fig. 4.

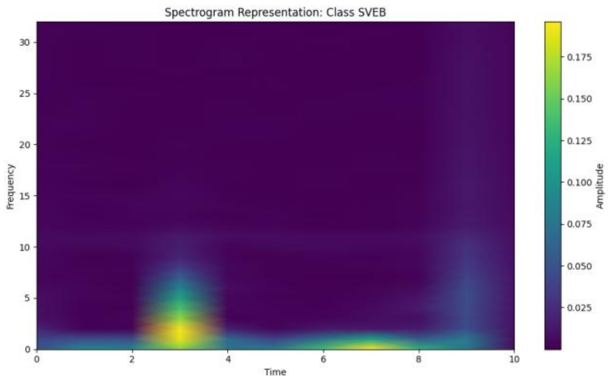

Fig. 4. Spectrogram representation of an ECG signal

CNN-SVM Hybrid Model

The classification system employs a hybrid Convolutional Neural Network (CNN) and Support Vector Machine (SVM) approach to leverage the strengths of both models. Fig. 5 illustrates the architecture of the CNN-SVM hybrid model, highlighting its two main stages: feature extraction and classification.

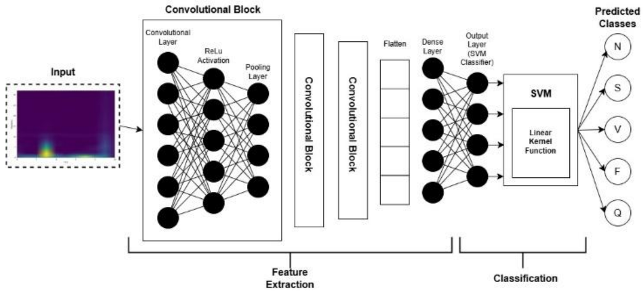

Fig. 5. CNN-SVM hybrid architecture

In the feature extraction stage, spectrogram representations of ECG signals serve as input to the CNN. The CNN architecture comprises three convolutional blocks, each consisting of a convolutional layer, a ReLU activation function, and a max-pooling layer. These blocks progressively extract hierarchical spatial and temporal features from the spectrograms, capturing essential patterns and reducing dimensionality. After the final convolutional block, the output is flattened into a one-dimensional vector, which is passed to a dense layer. This dense layer produces a compact feature representation that encapsulates the most relevant information required for classification.

In the classification stage, the feature vectors generated by the dense layer are passed to the SVM. The SVM uses a linear kernel to construct decision boundaries that separate the data into five arrhythmia classes: Normal (N), Supra-ventricular premature (S), Ventricular escape (V), Fusion of ventricular and normal (F), and Unclassified (Q). The hybrid approach capitalizes on the CNN's ability to learn complex spatial patterns and the SVM's precision in defining class boundaries, resulting in a robust model for arrhythmia classification.

The CNN was trained using 70% of the dataset. Within this, 80% of the training data (56% of the entire dataset) was used for learning, while the remaining 20% (14% of the entire dataset) was reserved for validation. The Adam optimizer and sparse categorical cross-entropy loss function were employed for training, with validation metrics monitored to prevent overfitting. After training, the CNN acted as a feature extractor, and the extracted features were used to train the SVM. The remaining 30% of the dataset was used as testing data to evaluate the model's performance on unseen data.

Performance Evaluation

The CNN-SVM hybrid model was evaluated on the testing dataset, which constituted 30% of the MIT-BIH dataset and was completely unseen during training and validation. The evaluation aimed to assess the model's generalization ability and classification performance across all five arrhythmia classes.

Several metrics were used for performance evaluation, including accuracy, precision, recall, and F1-score, with a particular emphasis on the weighted F1-score. This metric accounts for the imbalanced class distribution and provides a balanced assessment of the model's performance across both majority and minority classes. The precision, recall, and F1-score for each class were calculated and summarized in a table to highlight the model's performance for individual arrhythmia categories.

To monitor the training process, the CNN's performance was assessed on both training and validation datasets after each epoch. Metrics such as training loss, validation loss, training accuracy, and validation accuracy were tracked and plotted to ensure convergence and to detect any signs of overfitting. These curves provide insights into the model's learning behavior and its ability to generalize to unseen data.

This evaluation framework ensures a comprehensive assessment of the CNN-SVM hybrid model's effectiveness. Detailed results, including the metric values and training-validation performance curves, are discussed in Section III.

RESULT AND DISCUSSION

Overall Performance

The CNN-SVM hybrid model demonstrated strong performance, achieving a weighted F1-score of 0.985. This metric, which accounts for the dataset's imbalanced class distribution, highlights the model's ability to classify both majority and minority classes effectively. The high precision and recall values further confirmed the robustness of the hybrid model, indicating that it successfully extracted meaningful features from spectrograms and performed accurate classification.

These results suggest that the combination of CNN for feature extraction and SVM for classification effectively leverages the strengths of both approaches. The model demonstrated strong performance on unseen data from the MIT-BIH dataset, highlighting the importance of preprocessing steps such as wavelet denoising and STFT in enhancing the signal quality. However, further validation on larger and more diverse datasets is required to confirm its generalizability.

Class-Wise Performance Metrics

The class-wise performance of the CNN-SVM hybrid model is summarized in Table 1, which reports the precision, recall, and F1-score for each of the five arrhythmia classes: Normal (N), Supra-ventricular premature (S), Ventricular escape (V), Fusion of ventricular and normal (F), and Unclassified (Q). These metrics provide a detailed perspective on the model's ability to correctly identify each class.

Table 1 Performance Metrics for Each Arrhythmia classes

Class	**Precision**	**Recall**	**F1-Score**
Normal	0.99	1.00	0.99
S	0.90	0.80	0.85
V	0.97	0.97	0.97
F	0.86	0.70	0.77
Q	0.99	1.00	1.00

The model demonstrates near-perfect performance for the Normal (N) and Unclassified (Q) classes, with precision and recall values close to or at 100%. These results are consistent with the higher representation of these classes in the dataset, allowing the model to learn their patterns more effectively. In contrast, challenges are evident for minority classes such as Fusion (F) and Supra-ventricular premature (S), which exhibit relatively lower recall and F1-scores. For instance, the Fusion (F) class achieves a recall of 70% and an F1-score of 77%, while the Supra-ventricular premature (S) class reaches a recall of 80% and an F1-score of 85%. These metrics indicate difficulty in correctly identifying all true positive samples for these underrepresented classes.

The model's weighted F1-score of 0.985 demonstrates strong overall performance on the dataset, particularly for majority classes such as Normal (N) and Unclassified (Q). However, challenges remain for minority classes like Fusion (F) and Supra-ventricular premature (S), which show lower recall and F1-scores. This imbalance highlights the difficulty in accurately classifying underrepresented classes, often leading to misclassifications such as Fusion (F) being confused with Normal (N) or Ventricular escape (V). Addressing these limitations requires strategies to mitigate class imbalance, such as incorporating data augmentation, oversampling techniques, or applying class weighting during training. Additionally, exploring advanced loss functions, such as focal loss, could further enhance the model's ability to prioritize learning for challenging classes and improve its overall robustness.

Training and Validation Behavior

The training and validation curves for loss and accuracy, presented in Fig. 6, show the learning behavior of the CNN during training.

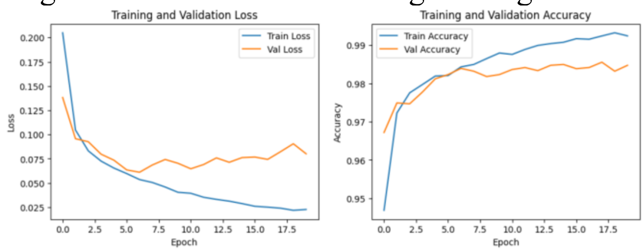

Fig. 6. Training and validation performance curves of the CNN

The steady decrease in both training and validation loss, combined with the close alignment of accuracy curves, suggests effective learning and no apparent signs of overfitting on the MIT-BIH dataset. However, further testing on external datasets is necessary to confirm the model's robustness. This behavior reflects the appropriateness of the model architecture and training process, as well as the robustness of the preprocessing pipeline. The use of validation data during training ensured the model maintained good generalization performance.

These findings suggest that the CNN-SVM hybrid model is well-suited for the specific arrhythmia classification tasks tested in this study, though additional validation on diverse datasets is required for broader applicability. The consistent convergence of the training and validation metrics reinforces the effectiveness of the chosen approach.

CONCLUSION

This study proposed a hybrid Convolutional Neural Network (CNN) and Support Vector Machine (SVM) model for arrhythmia classification using spectrogram representations of ECG signals. The CNN was employed to extract spatial and temporal features from spectrograms, while the SVM utilized these features to classify five arrhythmia classes: Normal (N), Supra-ventricular premature (S), Ventricular escape (V), Fusion of ventricular and normal (F), and Unclassified (Q).

The results demonstrated that the hybrid CNN-SVM model achieved high performance, with a weighted F1-score of 0.985, showcasing its ability to handle the imbalanced dataset effectively. Class-wise metrics highlighted strong performance across all classes, with particularly high precision, recall, and F1-scores for the majority class (Normal) and commendable accuracy for underrepresented classes such as Fusion and Unclassified. Despite these results, minority classes still exhibited slightly lower recall and F1-scores, reflecting the impact of class imbalance on the model's performance.

The training and validation curves confirmed the model’s stability during training and its ability to generalize well to unseen data. These findings emphasize the potential of the CNN-SVM hybrid approach for arrhythmia classification, particularly in capturing spatial and temporal features from spectrogram representations.

While the model performed robustly, certain limitations were noted. The dataset's inherent class imbalance posed challenges, particularly for minority classes such as Fusion and Supra-ventricular premature. Future work could explore techniques such as class weighting, data augmentation, or advanced loss functions like focal loss to further improve the model's performance for underrepresented classes. Additionally, testing the model on larger and more diverse datasets could validate its generalizability and potential for real-world applications.

In conclusion, the proposed CNN-SVM hybrid model demonstrates strong potential for accurate arrhythmia classification. By combining the feature extraction capabilities of CNNs with the robust classification power of SVMs, this approach offers a promising solution for enhancing the detection and diagnosis of arrhythmias in clinical and research settings.

REFERENCES

Arpitha, Y., Madhumathi, G. L., & Balaji, N. (2022). Spectrogram analysis of ECG signal and classification efficiency using MFCC feature extraction technique. *Journal of Ambient Intelligence and Humanized Computing*, *13*(2), 757–767. https://doi.org/10.1007/s12652-021-02926-2

Gai, N. D. (2022). *ECG beat classification using machine learning and pre-trained convolutional neural networks*. https://doi.org/https://doi.org/10.48550/arXiv.2207.06408

Huang, J., Chen, B., Yao, B., & He, W. (2019). ECG Arrhythmia Classification Using STFT-Based Spectrogram and Convolutional Neural Network. *IEEE Access*, *7*, 92871–92880. https://doi.org/10.1109/ACCESS.2019.2928017

Kementerian Kesehatan Republik Indonesia. (n.d.). *Penyakit Kardiovaskular*. Retrieved November 5, 2024, from https://ayosehat.kemkes.go.id/topik/penyakit-kardiovaskular

Madan, P., Singh, V., Singh, D. P., Diwakar, M., Pant, B., & Kishor, A. (2022). A Hybrid Deep Learning Approach for ECG-Based Arrhythmia Classification. *Bioengineering*, *9*(4). https://doi.org/10.3390/bioengineering9040152

Neha, Sardana, H. K., Kanwade, R., & Tewary, S. (2021). Arrhythmia detection and classification using ECG and PPG techniques: a review. In *Physical and Engineering Sciences in Medicine* (Vol. 44, Issue 4, pp. 1027–1048). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s13246-021-01072-5

PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals,. (n.d.).

Ramos, M. Y. G., Castilla, J. S. R., & Lamont, F. G. (2021). Hybrid Algorithm of Convolutional Neural Networks and Vector Support Machines in Classification. *Communications in Computer and Information Science*, *1485 CCIS*, 263–274. https://doi.org/10.1007/978-3-030-90241-4_21

The Impact of the MIT-BIH Arrhythmia Database. (n.d.). https://doi.org/10.1109/51.932724

Unlersen, M. F., Sonmez, M. E., Aslan, M. F., Demir, B., Aydin, N., Sabanci, K., & Ropelewska, E. (2022). CNN–SVM hybrid model for varietal classification of wheat based on bulk samples. *European Food Research and Technology*, *248*(8), 2043–2052. https://doi.org/10.1007/s00217-022-04029-4

World Health Organization. (n.d.). *Cardiovascular diseases (CVDs)*. Retrieved November 5, 2024, from https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)