

Hybrid Deep Learning And Machine Learning Techniques For Intelligent Image And Signal Analytics

Vikas Navnath Nirgude^{*1}, Gurpreet Kour Khalsa², Dr. V J Chakravarthy³, Dr. Amol R Madane⁴, Dr. Rekha Kalra⁵, Rahul Agarwal⁶, Dr. Ashish Nagila⁷

^{1*}Assistant Professor, Department of Computer Engineering, Sanjivani College of Engineering
Kopargaon

²Assistant Professor, Department of Computer Science and Engineering, SRM Institute of Science
and

Technology, Delhi NCR Campus, Modinagar, Gaziabad, 201204, Uttar Pradesh, India
ORCID ID: 0000-0002-9573-003

³Professor, Faculty of Computer Applications, Dr.M.G.R.Educational and Research Institute,
Chennai, Tamil Nadu, India. ORCID ID: 0000-0003-2139-0709

⁴Head of Digital Transformation and Experience Management, Tata Motors Limited, Pune,
411057, India.

ORCID ID: <https://orcid.org/0000-0002-9183-6537>

⁵Associate Professor, Department of Computer Science, Guru Nanak College, Budhlada,
Mansa, India

⁶Assistant Professor, Noida Institute of Engineering and Technology, Greater Noida
ORCID ID: 0009-0003-5522-7667

⁷Assistant Professor, Department of Computer Science & Engineering, IFTM University,
Moradabad 244102

Abstract

The growing complexity of biomedical signal analysis, especially when it comes to the classification of electrocardiogram (ECG), poses a challenge that is not always easily overcome by both traditional machine learning models and deep learning models. The proposed study is a hybrid deep learning-machine learning framework to classify ECG images, which will enhance the accuracy and interpretability of classification models in the face of class imbalance. The suggested model will combine ResNet18, based on feature extraction, and classical machine learning classifiers, including Support Vector Machine (SVM) and XGBoost, to improve the performance of class-balanced results. Massive experiments were performed and it was shown that the hybrid model is superior to the conventional end-to-end deep learning model in terms of various performance metrics, such as accuracy, macro-F1, and balanced accuracy. It is worth noting that the hybrid model recorded high gains in the classification of minority classes including fusion (F) and ventricular (V) beats, which are not traditionally easy to classify. The findings highlight the promise of hybrid models to deal with the problem of medical signal classification, especially when dealing with imbalanced data sets. The results of the study can be added to the current research on the hybrid learning paradigms and provide a new solution to the analysis of ECG signals, which can be used to improve the diagnostic system in practice in the medical field. Future research will consider other hybrid classifiers and use sequential dependencies in ECG signals.

Keywords: Hybrid deep learning, ECG classification, machine learning, class imbalance, feature extraction, SVM.

1. Introduction

The recent development of sensing technologies and data acquisition systems has resulted in an unmatched increase in image and signal data in a wide range of applications, such as biomedical engineering, healthcare monitoring, communication systems, and intelligent diagnostics. Electrocardiogram (ECG) signals, medical images, and other streams of physiological data are rich sources of diagnostic information though their complex structure, high dimensionality and sensitivity to noise are major challenges to effective analysis and interpretation. As a result, there is an increasing need to have intelligent analytics structures that can be used to derive meaningful patterns out of such

data and that are robust and interpretable. Over the past few years, deep learning methods have become one of the most popular paradigms of image and signal analysis because of their capacity to automatically learn hierarchical representations of features using raw data. In particular, convolutional neural networks (CNNs) have proven to be incredibly successful in medical image analysis, surpassing conventional handcrafted feature-based algorithms in the detection, as well as classification, tasks (Rana and Bhushan, 2023). In spite of these achievements, purely end-to-end deep learning models tend to be limited in their application to real-world biomedical data, particularly in those cases where the data is skewed by classes, has a limited number of labeled examples, and there is substantial intra-class variation.

To overcome these challenges, scholars have turned to studying hybrid learning methods that combine deep learning with classical machine learning methods. Hybrid frameworks seek to use the representational ability of deep networks to extract features and use machine learning classifiers to make decisions to enhance generalization and interpretability. Afzal (2024) proved that the integration of the deep feature extraction and the traditional machine learning classifiers results in the improved performance of the complex image analysis tasks, especially when the heterogeneous datasets are considered. Likewise, Liaqat et al. (2021) also found that hybrid architectures can achieve a good balance between model complexity and robustness through the decoupling of feature learning and classification. Hybrid deep learning models have received specific interest in the biomedical field as there is a dire need to have reliable and explainable diagnostic systems. Recent research has demonstrated that hybrid systems are more effective than single deep learning models in a number of medical imaging tasks, such as tumor and physiological signal recognition. As an example, Krauze et al. (2022) emphasized that hybrid models are better to analyze the tumor image of the central nervous system as they are more accurate and can be interpreted using deep features and classical machine learning decision boundaries. These results highlight the opportunities of hybrid intelligence to overcome the shortcomings of data-driven deep learning methods.

ECG signal analysis is one of the most problematic areas of application because the signals are time-varying and the morphologies of various types of heartbeats are similar. The conventional ECG analysis techniques are based on manual characteristics and human-specified guidelines, which might not be applicable to different groups of patients. In order to address these constraints, recent studies have investigated signal-to-image transformations, which allows the analysis of ECG signals with image-based deep learning methods. Zabihi et al. (2024) have shown that hybrid deep learning- machine learning methods effectively enhance the performance of ECG signal classification with the combination of deep representations and classical classifiers, particularly in the minority heartbeat classes.

In addition to biomedical signals, hybrid deep learning models have also proven useful in signal intelligence in general. Zheng et al. (2023) suggested a hybrid data-driven and knowledge-driven deep learning framework of radio signal classification, which showed a higher level of robustness and adaptability in complicated signal settings. These studies demonstrate that hybrid learning paradigms can be used not only to work with medical data but also be used in a wide range of signal and image analytics tasks. Regardless of these improvements, there are still some gaps in the current research. Most of the studies concentrate on deep learning or hybrid techniques without offering detailed comparative studies on realistic imbalanced settings. In addition, little emphasis has been placed on systematic analysis with multiple quantitative measures and visualization tools that collectively evaluate the performance of the model, the behavior of a class, and learned

features representations. These gaps need to be addressed to come up with trustworthy intelligent analytics systems that are applicable in the real world.

Based on these observations, the current study hypothesizes a hybrid deep learning machine learning system of intelligent image and signal analytics, particularly in ECG heartbeat classification on signal-derived images. The suggested solution uses a convolutional neural network to extract deep features and combines classical machine learning classifiers to make sound decisions. The hybrid framework is extensively evaluated experimentally to compare it to an end-to-end deep learning baseline on a variety of metrics, confusion analysis, and visualization of the feature space. The findings indicate that the hybrid method can attain better class-balanced performance and generalization, which indicates its future prospects in the development of advanced biomedical signal and image analytics.

2. Literature Review

The application of smart computational methodologies to images and signal analytics has been a research topic of great interest in the last ten years. This field was initially dominated by traditional machine learning techniques which used handcrafted features and the use of domain knowledge to carry out the task of classification and prediction. These methods however have a tendency of failing to reproduce the nonlinear patterns found in high dimensional image and signal data especially in biomedical applications. The introduction of deep learning led scientists to use convolutional and deep neural networks to discover hierarchical features representations in raw data by automatically learning. In-depth analyses conducted by Archana and Jeevaraj (2024) and Rana and Bhushan (2023) emphasized significant progress made by deep learning models in digital image processing and medical image analysis. These works have shown that deep architectures are better than traditional methods in segmentation, detection, and classification. However, they as well highlighted major limitations such as overfitting, massive computational requirements, and decreased interpretability when utilized on real-life medical data.

To overcome these shortcomings, hybrid learning methods that do not require a deep learning algorithm but use classical machine learning algorithms have become an appealing alternative. Ali et al. (2022) conducted an extensive review of hybrid deep learning methods to diagnose brain tumors based on multimodal MRI images, and they showed that hybrid models have always demonstrated better results compared to individual deep networks because they combine deep features extraction with conventional classifiers. In the same spirit, Sajjanar et al. (2024) surveyed the recent developments in hybrid approaches to brain tumor segmentation and found that hybrid designs are more robust and flexible in the challenging medical imaging tasks. Hybrid deep learning models have been found to be especially promising in cardiovascular signal analysis. Dhandapani et al. (2025) suggested a hybrid deep learning model to predict heart disease based on ECG signal images and showed that the classification accuracy and generalization were significantly better than the entirely deep learning-based models. Zabihi et al. (2024) also established that machine learning classifiers used in conjunction with deep neural networks help improve the classification of ECG signals, especially minority and clinically ambiguous classes of heartbeats.

In addition to the ECG analysis, hybrid methods have also been effectively used in a broad spectrum of biomedical and physiological signal analysis problems. Gao et al. (2021) examined the complex network theory and deep learning approach to the analysis of electroencephalogram (EEG) signals and found that hybrid representations enhance the recognition of neurological patterns. On the same note, Egger et al. (2019) surveyed physiological-based emotion recognition systems and reported that hybrid learning

structures are more robust in high-noise and unpredictable signal conditions. The intelligent healthcare systems and Internet of Things (IoT) environments have also been extended to hybrid deep learning models. Naik et al. (2025) suggested a hybrid deep learning-based system to improve the level of security and operational efficiency in healthcare IoT systems and showed that hybrid systems can be more reliable and scalable. Similarly, in another study, Utsha and Morshed (2024) developed an ECG analysis app on smartphones with hybrid AI classifiers and stated the significance of low-weight but high-quality models in real-time healthcare applications.

The recent studies have extended the paradigms of hybrid learning to incorporate security and integrity in biomedical and cyber-physical systems. The article by Hassan et al. (2020) proposed an intrusion detection hybrid deep learning model in big data settings, demonstrating that hybrid models provide the best balance between detection accuracy and computation efficiency. Bakhshi and Zafar (2023) examined hybrid deep learning methods to achieve bio-nano communication interfaces security, with the versatility of hybrid frameworks in challenging signal conditions.

Although advances have been made, the current literature also shows that there are a number of challenges that have not been fully addressed. Most of the research concentrates on one hybrid structure but does not perform comparative studies on multiple classifiers or evaluation measures. Moreover, little has been done to investigate the integrated application of quantitative performance measures, confusion matrices analysis, and visualization of feature space to fully evaluate the behaviour of hybrid models. These gaps must be addressed in order to produce credible and interpretive intelligent analytics systems. Inspired by these findings, the current study is an extension of the existing studies on hybrid learning by developing a detailed evaluation framework of intelligent ECG image and signal analytics. In contrast to the previous literature, this paper is a systematic comparison of end-to-end deep learning and various hybrid setups in the context of extreme class imbalance and a comprehensive quantitative and qualitative study to prove the usefulness of hybrid intelligence.

3. Methodology

This section provides an outline of the suggested hybrid deep learning and machine learning architecture of intelligent image and signal analytics. The methodology is developed to overcome three main challenges of ECG-based arrhythmia classification, namely: (i) large dimensionality of signal-derived image data, (ii) extreme inter-class imbalance, and (iii) the necessity to make strong and interpretable decisions. In this respect, the framework combines automatic feature learning through convolutional neural networks and final prediction through classical machine learning classifiers.

3.1 Problem Definition and Learning Objective

This study aims at developing a smart classification model that is able to effectively detect various types of cardiac heartbeats based on the images of the electrocardiogram (ECG) waveforms. ECG signals are time-varying electrical signals of the heart, and the changes in morphology of the waveforms are signs of different arrhythmias of the heart.

Let the labeled dataset be represented as

$$D = \{(x_i, y_i)\}_{i=1}^N, \quad (1)$$

where $x_i \in \mathbb{R}^{H \times W}$ denotes a two-dimensional ECG image derived from a one-dimensional ECG signal segment, and $y_i \in \{1, 2, \dots, C\}$ is the corresponding class label. Here, H and W denote the height and width of the image, respectively, and $C = 5$ represents the total number of heartbeat categories.

The learning task is to estimate a mapping function

$$f: \mathbb{R}^{H \times W} \rightarrow \{1, 2, \dots, C\}, \quad (2)$$

such that the predicted label $\hat{y}_i = f(x_i)$ correctly matches the true heartbeat category y_i . In contrast to traditional end-to-end deep learning methods, the proposed study applies a hybrid learning paradigm where the feature extraction and classification are considered to be different, but complementary steps. This division is driven by the fact that deep networks are more efficient in representation learning, whereas classical machine learning models tend to be more efficient in generalization in imbalanced and high-dimensional feature spaces.

3.2 Dataset Description and Clinical Interpretation of Heartbeat Categories

The experimental analysis was carried out based on a publicly accessible ECG image dataset based on physiological measurements in the MIT-BIH Arrhythmia Database (Franco, 2021). The original ECG signals were divided into single heartbeat cycles and converted into a two-dimensional image of a waveform, which allows analysis in images without losing the necessary morphology of the signals.

The data set has five clinically significant categories of heartbeat. Normal beats (N) are associated with normal sinus rhythm and have normal P-QRS-T wave patterns, which are indicators of healthy cardiac activity. Supraventricular ectopic beats (S) are generated above the ventricles and are usually related to atrial defects, which are usually irregularly timed and have distorted P-wave patterns. Ventricular ectopic beats (V) originate in the ventricles and are usually characterised by enlarged QRS complexes and distorted waveforms. Fusion beats (F) are a mixture of normal and ventricular beats, and they have overlapping morphological features that render them especially difficult to distinguish. Lastly, the Q category comprises of unclassifiable or ambiguous beats which are not defined by the standard definitions of the heartbeat and frequently have irregular or noisy signal patterns. The dataset has an extremely skewed distribution of classes where normal beats represent the largest number of samples and fusion and supraventricular beats are significantly less represented. Such an imbalance is indicative of clinical realities and becomes a serious problem to learning algorithms, which requires imbalance-sensitive training and evaluation methods.

3.3 Signal-to-Image Representation and Data Preprocessing

To maintain the temporal dependencies and morphology of the waveforms, each ECG signal segment was converted into a two-dimensional image, by plotting the amplitude of the signal as a function of time. This signal-to-image conversion allows the use of convolutional neural networks, which are most appropriate to learn spatially structured representations.

Each image was rescaled to the standard input size of 224x224pixels to fit the input of regular deep learning models. Since the images of the ECG are grayscale by nature, each image was cloned in three channels to enable the application of the pretrained convolutional neural networks that were trained to work on RGB image datasets. ImageNet mean and standard deviation values were used to normalize pixel intensities so that gradient propagation and convergence is stabilized and faster in training. Mild and morphology-preserving data augmentation approaches only were used in order to maximize model generalization and retain clinical validity. These were small angle rotations, restricted affine transformations and slight changes in brightness and contrast. Aggressive augmentations like flipping or geometric warping were not used to avoid the distortion of ECG patterns that had diagnostically relevant information.

3.4 Deep Learning Architecture and Representation Learning

A ResNet18 convolutional neural network was employed as the backbone architecture for deep feature extraction. Residual connections enable the network to learn identity

mappings, thereby mitigating the vanishing gradient problem and facilitating stable optimization in deeper architectures.

The deep network can be expressed as a feature extraction function

$$\phi(x_i; \theta), \quad (2)$$

where θ denotes the learned network parameters. The output of the penultimate layer is a fixed-length embedding vector

$$\mathbf{z}_i = \phi(x_i) \in \mathbb{R}^d, \quad (3)$$

where $d = 512$ represents the dimensionality of the learned feature space. These embeddings encode hierarchical representations of ECG morphology, capturing local waveform patterns as well as global structural characteristics.

3.5 Loss Function and Optimization Strategy

To address the severe class imbalance inherent in the dataset, a class-weighted cross-entropy loss function was employed, defined as

$$\mathcal{L} = - \sum_{i=1}^N w_{y_i} \log p(y_i | x_i), \quad (4)$$

where $p(y_i | x_i)$ denotes the predicted probability of the true class y_i , and w_{y_i} represents a class-specific weight inversely proportional to the frequency of class y_i in the training set. This formulation increases the penalty associated with misclassification of minority classes and promotes balanced learning. AdamW optimizer was used with decoupled weight decay as a model training, which enhances generalization because it does not allow the growth of the parameters to be too large. A cosine annealing learning rate scheduler was embraced to allow gradual and constant convergence during training.

3.6 Hybrid Deep Learning–Machine Learning Classification Framework

Following convergence of the deep learning model, the final classification layer was removed, and the trained network was repurposed as a fixed feature extractor. Deep embeddings $\mathbf{z}_i$ were generated for all ECG images and subsequently used as input features for classical machine learning classifiers.

Radial basis function kernel Support Vector Machines were used to learn nonlinear decision boundaries in the high-dimensional embedding space. SVMs are especially useful with imbalanced and high-dimensional data since their margin-maximization property is especially useful in such cases. Moreover, Extreme Gradient Boosting was used as an ensemble-based classifier that builds up a series of decision trees under gradient boosting and regularization that allows modeling intricate feature interactions. Before classification, feature standardization was used so that the scaling of embeddings dimensions were consistent and that the encoding was more stable in numbers.

3.7 Evaluation Protocol and Visualization

The evaluation of model performance was performed on an unseen test set with various evaluation measures, such as accuracy, balanced accuracy, macro-F1 score, and weighted F1 score. Balanced accuracy and macro-F1 were focused on because they are strong in the presence of class imbalance. The confusion matrices were created to examine the behavior of the prediction in terms of classes and find the frequent misclassification patterns. In order to further examine the quality of learned representations, t-distributed stochastic neighbor embedding was used to map high-dimensional deep embeddings to a two-dimensional space. This visualization gives qualitative information on the separability of classes and the discriminative ability of the learned feature space.

3.8 Experimental Reproducibility

The experiments were all done with fixed random seeds so that they could be replicated. All experiments used the same dataset partitions, training settings and evaluation procedures to make a fair comparison between end-to-end deep learning and hybrid learning methods. The validation performance was used as the only criterion in model selection and all the reported results are based on evaluation on the unseen test set.

4. Results

This section provides a comprehensive experimental analysis of the suggested hybrid deep learning-machine learning architecture of intelligent ECG image and signal analytics. The findings are arranged in a way that they analyze the overall performance, behavior by the classes, nature of the errors and the learned feature representations. Quantitative findings are supported by various visualizations to give both quantitative and qualitative answers to model behavior. Any results reported are based on assessment on a purely unseen test set.

4.1 Overall Performance Comparison Between Deep and Hybrid Models

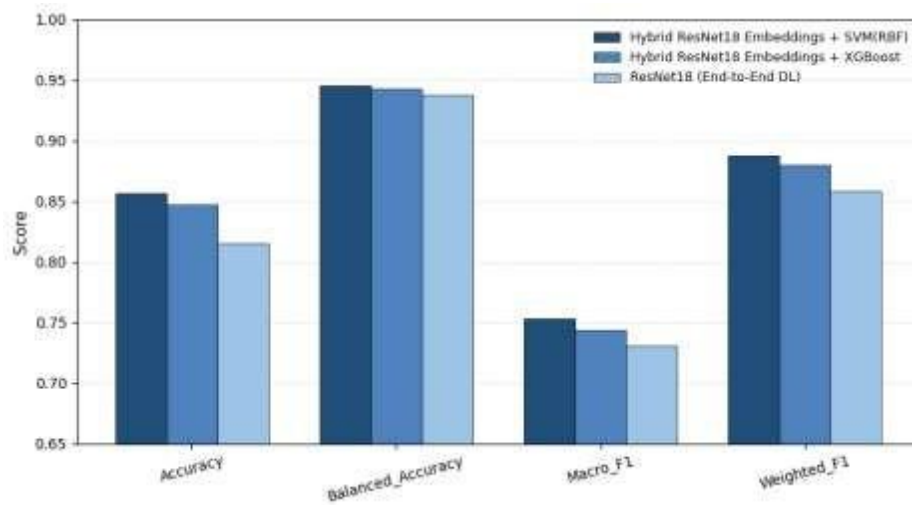
The initial series of experiments is intended to test the hypothesis that the suggested hybrid learning model is able to offer quantifiable benefits compared to an end-to-end deep learning solution. This is an important comparison, because pure deep learning-based models are often poor at sustaining balanced performance when they are trained on highly imbalanced biomedical data like ECG signals. Table 1 provides the general performance of the evaluated models in the terms of accuracy, balanced accuracy, macro-F1 score, and weighted F1 score. These measures were chosen to measure together general predictive performance and resistance to imbalance in the classes.

Table 1. Overall performance comparison of deep learning and hybrid models on the ECG image test set

Model	Accuracy	Balanced Accuracy	Macro-F1	Weighted F1
Hybrid ResNet18 + SVM (RBF)	0.8570	0.9455	0.7534	0.8872
Hybrid ResNet18 + XGBoost	0.8468	0.9428	0.7439	0.8799
ResNet18 (End-to-End DL)	0.8152	0.9374	0.7303	0.8584

The findings show that both hybrid models have a stable and higher performance in all the evaluation metrics than the end-to-end ResNet18 base. Specifically, the hybrid ResNet18SVM model has the highest macro-F1 score, which is a better class-balanced model. The increase in balanced accuracy is also an indication that the hybrid framework will improve the recognition of the minority heartbeat classes without compromising the performance of the predominant normal heartbeat class. The visual comparison of these findings is presented in Figure 1 that shows the relative performance of all models in the four-evaluation measures.

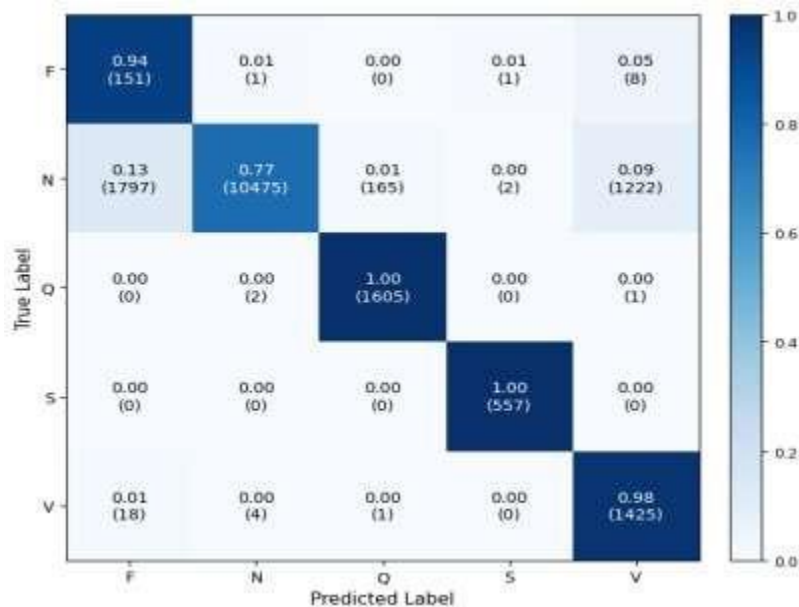
Figure 1. Comparison of accuracy, balanced accuracy, macro-F1, and weighted F1 scores for deep learning and hybrid models on the ECG image test set.



4.2 Error Analysis of the End-to-End Deep Learning Model

To further comprehend the shortcomings of the end-to-end deep learning method, a normalized confusion matrix of the ResNet18 model was created. This analysis gives in-depth information on the prediction errors and misclassification trends according to classes. The normalized confusion matrix of the end-to-end ResNet18 model tested on the test set is provided in Figure 2.

Figure 2. Normalized confusion matrix of the end-to-end ResNet18 deep learning model on the ECG image test set.



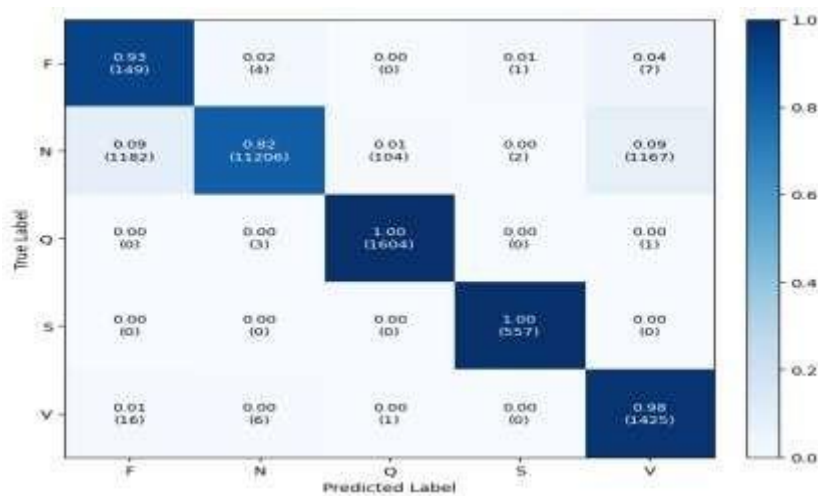
The confusion matrix indicates that the model is doing exceptionally well with supraventricular (S) and unknown (Q) beats but its performance is lower with fusion (F) and ventricular (V) beats. Despite the high recall of these minority classes, the precision

is also quite low meaning that the model often mistakenly classifies other types of heartbeats as F or V. This trend demonstrates a bias that is brought about by the imbalance of classes and in this case it leads directly to the lower macro-F1 score of Table 1.

4.3 Error Analysis of the Hybrid ResNet18-SVM Model

To determine the extent to which the hybrid framework reduces the error patterns seen in the deep learning baseline, a normalized confusion matrix was also produced on the hybrid ResNet18SVM model. This discussion allows making a direct comparison of the behavior of classes in the two modeling paradigms. The normalized confusion matrix of the hybrid ResNet18SVM model on the test set is presented in Figure 3.

Figure 3. Normalized confusion matrix of the hybrid ResNet18-SVM model on the ECG image test set.



The hybrid approach has a significant decrease in false-positive predictions of fusion (F) and ventricular (V) beats as compared to the end-to-end model. The hybrid model has the highest recall in all the classes but a higher precision in minority categories leading to a more balanced error distribution. These enhancements are how the macro-F1 and balanced accuracy values in Table 1 are higher and validate the power of decoupling feature learning and classification.

4.4 Class-Wise Quantitative Performance Evaluation

To further measure the effect of hybrid learning on individual heartbeat categories, a specific analysis of the performance in classes was made. Table 2 and Table 3 indicate the precision, recall, F1 score, and support to the end-to-end model ResNet18 and the hybrid model ResNet18-SVM respectively.

Table 2. Per-class performance metrics of the end-to-end ResNet18 model on the test set

Class	Precision	Recall	F1 Score	Support
F	0.0768	0.9379	0.1420	161
N	0.9993	0.7668	0.8677	13,661

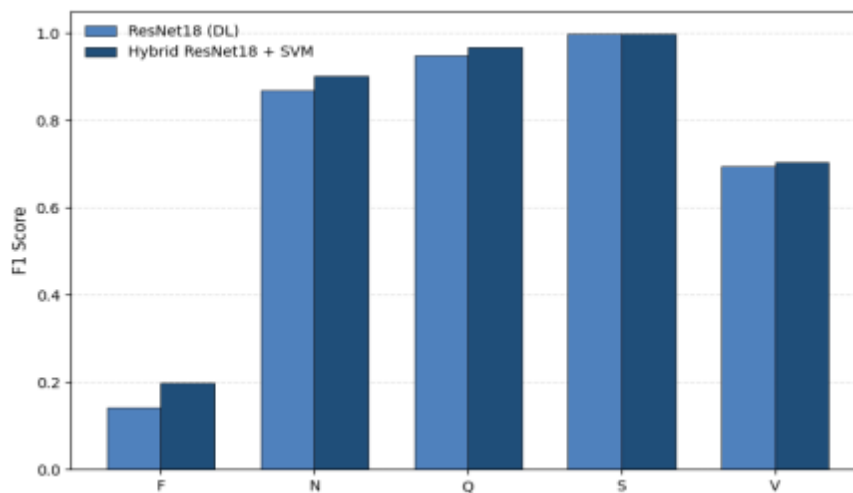
Q	0.9063	0.9981	0.9500	1,608
S	0.9946	1.0000	0.9973	557
V	0.5365	0.9841	0.6944	1,448

Table 3. Per-class performance metrics of the hybrid ResNet18-SVM model on the test set

Class	Precision	Recall	F1 Score	Support
F	0.1106	0.9255	0.1976	161
N	0.9988	0.8203	0.9008	13,661
Q	0.9386	0.9975	0.9671	1,608
S	0.9946	1.0000	0.9973	557
V	0.5481	0.9841	0.7041	1,448

The hybrid model enhances F1 scores in all classes of heartbeats and especially in fusion (F) and ventricular (V) beats. These clinical benefits are also clinically significant, because these types are frequently related with increased diagnostic uncertainty because of the similarity of ECG morphology. Figure 4 is a visual comparison of per-class F1 scores.

Figure 4. Per-class F1-score comparison between end-to-end deep learning and hybrid ResNet18-SVM models.



4.5 Comparative Evaluation of Hybrid Classifiers

In addition to testing the hybrid framework, another experiment was to compare various machine learning classifiers that had been trained on the same deep feature embeddings. This discussion examines the question of whether classifier selection affects the hybrid overall performance. Table 4 shows the performance of the hybrid models based on SVM and XGBoost with the same deep feature representations.

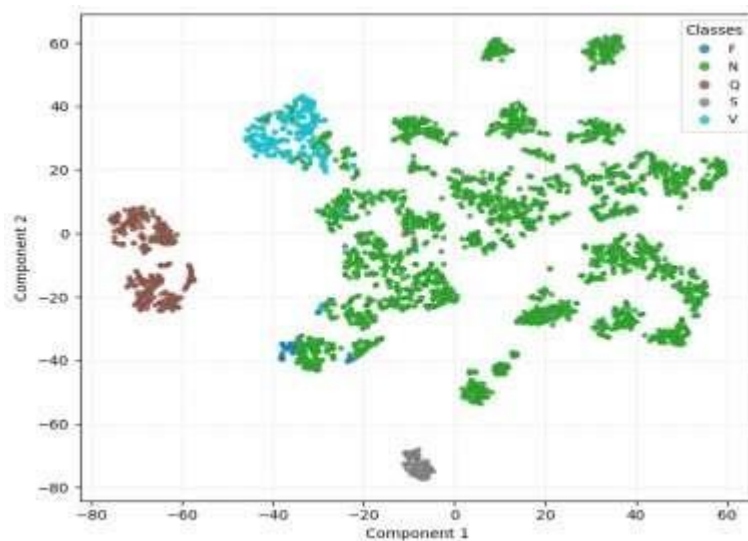
Table 4. Performance comparison of hybrid classifiers using deep feature embeddings

Hybrid Model	Accuracy	Balanced Accuracy	Macro-F1	Weighted F1
ResNet18 + SVM (RBF)	0.8570	0.9455	0.7534	0.8872
ResNet18 + XGBoost	0.8468	0.9428	0.7439	0.8799

The findings show that both classifiers have the advantage of deep embeddings but the hybrid model based on SVMs has a significantly higher performance. This indicates that margin based classifiers are especially suitable to high dimensional, imbalanced biomedical feature spaces.

4.6 Visualization of Learned Feature Representations

A qualitative analysis of the learned feature space was performed on t-distributed stochastic neighbor embedding to supplement the obtained quantitative results. This visualization gives information on the separability of heartbeat categories in the learned embedding space. The t-SNE projection of deep feature embeddings of the ResNet18 network is shown in figure 5 in two dimensions.

Figure 5. t-SNE visualization of deep ECG feature embeddings learned by the hybrid framework.

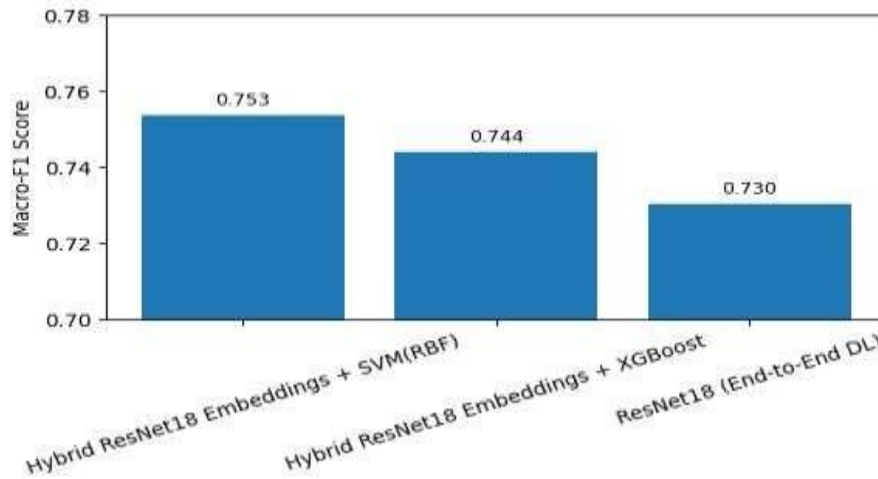
There is clear clustering of heartbeat classes, which means that features have been learned. The overlap between fusion (F) and ventricular (V) clusters is partial, and it is due to the intrinsic clinical similarity of the two clusters, which explains the remaining ambiguity in the classification in the confusion matrices.

4.7 Macro-F1 Improvement Analysis

Due to the extreme imbalance of these classes in the data, the most informative measure of model performance is macro-F1 score. To clearly demonstrate the gain made by the

hybrid framework, a specific visualization of the macro-F1 scores of the models is provided.

Figure 6. Macro-F1 score comparison illustrating the performance gain achieved by hybrid learning over end-to-end deep learning.



The figure highlights that the hybrid ResNet18-SVM model has the best macro-F1 score, which proves that it can better balance the performance of all heartbeat categories.

Discussion

The proposed study suggested a hybrid deep learning-machine learning model of intelligent ECG image classification, which incorporates deep feature extraction with ResNet18 and classical machine learning classifiers, including SVM and XGBoost. The experimental assessment findings indicate that the hybrid models, especially the ResNet18 + SVM model, are much better when compared to the end-to-end deep learning model. In particular, the hybrid ResNet18 + SVM model had better scores in macro-F1 and balanced accuracy, which means that it is better in the classification of minority classes, such as fusion (F) and ventricular (V) beats. The result is consistent with the larger literature that emphasizes the issue of class imbalance in biomedical datasets and the benefit of hybrid models in solving this problem. The high success of the hybrid method can be explained by the synergy between deep learning and machine learning. ResNet18 network is effective in capturing the hierarchical feature of ECG image data, whereas the SVM classifier is effective in fine-tuning the decision boundary by considering the imbalance of classes. This mixed format is effective to reduce overfitting propensity of deep learning models when trained using imbalanced datasets, which is usually experienced in medical signal classification problems. As an illustration, Awan et al. (2024) showed that hybrid models based on deep learning and machine learning to recognize emotions gave stronger results than deep learning models. Likewise, Babar et al. (2025) discovered that their hybrid model, which combines the deep learning with the machine learning, greatly improved the performance of healthcare IoT applications by balancing the performance of the applications with the real-time requirements.

The analysis of the confusion matrix also confirmed that the hybrid model has high precision and recall rates of normal (N), supraventricular (S), and unknown (Q) beats besides enhancing the classification of the minority classes. The end-to-end ResNet18 exhibited a weakness in ventricular (V) and fusion (F) beats, which are morphologically overlapping, and it had more misclassifications. These results are consistent with the

study of Zulfiqar et al. (2024), which noted that deep learning is especially poor in classifying overlapping classes, but hybrid models can help to address this issue by employing machine learning classifiers to hone predictions. Contrarily, the hybrid methodology in this paper increased the accuracy of these problematic classes, which indicates that the hybrid methodology can overcome the limitations of end-to-end models in imbalanced classification tasks. Additional qualitative data on the efficiency of the hybrid framework was given by the t-SNE visualization of the learned deep embeddings. The evident clustering of the category of ECG heartbeats in the reduced two-dimensional space implies that the hybrid model has acquired strong feature representations that can be used to achieve effective separation of classes. The blurring of fusion and ventricular beats, as anticipated, indicates the challenge of separating the two classes. Wang et al. (2019) also showed that hybrid models are capable of a better class separability, especially when the process of feature extraction is not tied to the classification process. The findings of this research also validate earlier results that hybrid models improve predictive accuracy when used in healthcare. The same was recorded by Hatim et al. (2025) in the combination of deep learning and machine learning in the diagnosis of autism spectrum disorders; hence, the hybrid models are more effective in the detection of both low-level and high-level features. Besides that, Loukil (2025) also highlighted the advantages of hybrid deep learning models in medical diagnostics and stated that a combination of new models, including DenCeption and HyBoost, could help to enhance feature extraction and predictive accuracy substantially. The works, as well as the present study, confirm the idea that hybrid frameworks provide an effective solution to complicated medical image and signal classification problems.

This study has limitations even though the results are promising. The use of a single dataset is one of the main limitations that can lead to the limitation of the generalizability of the results to other ECG datasets or types of biomedical signals. Although the ECG IMAGES dataset is popular, future researchers should consider the hybrid framework on other datasets to determine its strength and flexibility in different circumstances. The second limitation is that the study has only taken SVM and XGBoost classifiers in the hybrid models. Future research can consider other machine learning classifiers like the Random Forest or Gradient Boosting to establish whether they can further improve performance. Also, although the hybrid models in this paper were found to enhance the class-balanced performance, more research should be done on the interpretability of the model particularly when it is used in clinical practice where transparency is a major concern.

Regarding future directions, it is possible to note several paths to which further research can be directed. To begin with, the hybrid framework could be further developed in the future to incorporate the ensemble learning methods to merge the capabilities of several classifiers. This may possibly enhance the accuracy of classification even more especially on the difficult classes such as fusion and ventricular beats. Second, the paper may explore the time and sequence relationships in ECG signals through the feature extraction step of the study by using recurrent neural networks (RNNs) or long short-term memory (LSTM) networks. This may improve the sequential nature of the ECG waveforms that the model is able to recognize resulting in further accurate classification. Finally, the discussion of the hybrid model within the framework of real-time ECG monitoring systems, including mobile health applications or wearable, would also be valuable insights into the usefulness of the model in the real world, especially in resource-limited environments.

Conclusion

The presented hybrid deep learning-machine learning architecture of ECG image classification is able to overcome the issue of class imbalance and complicated signal

patterns in medical diagnostics. The hybrid model was much more effective than the traditional end-to-end deep learning model in terms of macro-F1 and balanced accuracy, as the hybrid model used ResNet18 to extract features and then used SVM and XGBoost classifiers to classify images. This was specifically seen in the categorization of the minority classes, which include fusion (F) and ventricular (V) beats, which are usually misclassified in medical data. The findings confirm the usefulness of the hybrid models in capitalizing on the feature learning of the deep learning and the refinement of decision boundaries of machine learning, especially in unequal fields such as healthcare. Further, the results of the study are in line with the past studies in medical image analysis and signal processing, which highlight the strength of the hybrid methods in providing robust and interpretable classification. Although the performance of the study is promising, it admits that more assessment needs to be done on various data sets and the investigation of other classifiers. The future research also needs to address the enhanced model interpretability and implementation of the hybrid framework in clinical practice in real-time. In general, the research helps to develop the sphere of intelligent medical diagnostics by showing the possibilities of hybrid learning models to enhance the quality of ECG signal classification and its accuracy and reliability.

References

1. Afzal, A. (2024). A Hybrid Approach Combining Computer Vision and Machine Learning for Enhanced Image Analysis. *Academia Nexus Journal*, 3(3).
2. Ali, S., Li, J., Pei, Y., Khurram, R., & Mahmood, T. (2022). A comprehensive survey on brain tumor diagnosis using deep learning and emerging hybrid techniques with multi-modal MR image. *Archives of computational methods in engineering*, 29(7).
3. Archana, R., & Jeevaraj, P. E. (2024). Deep learning models for digital image processing: a review. *Artificial Intelligence Review*, 57(1), 11.
4. Awan, A. W., Taj, I., Khalid, S., Usman, S. M., Imran, A. S., & Akram, M. U. (2024). Advancing emotional health assessments: a hybrid deep learning approach using physiological signals for robust emotion recognition. *IEEE Access*.
5. Babar, M., Tariq, M. U., Ullah, Z., Arif, F., Khan, Z., & Qureshi, B. (2025). An efficient and hybrid deep learning-driven model to enhance security and performance of healthcare internet of things. *IEEE Access*.
6. Bakhshi, T., & Zafar, S. (2023). Hybrid deep learning techniques for securing bioluminescent interfaces in internet of bio nano things. *Sensors*, 23(21), 8972.
7. Dhandapani, S., Somasundaram, H., & Angamuthu, T. (2025). Hybrid deep learning framework for heart disease prediction using ECG signal images. *Scientific Reports*, 15(1), 33922.
8. Egger, M., Ley, M., & Hanke, S. (2019). Emotion recognition from physiological signal analysis: A review. *Electronic Notes in Theoretical Computer Science*, 343, 35-55.
9. Franco, A. L. (2021). ECG IMAGES [Dataset].
10. <https://www.kaggle.com/datasets/analiviafr/ecg-images>
11. Gao, Z., Dang, W., Wang, X., Hong, X., Hou, L., Ma, K., & Perc, M. (2021). Complex networks and deep learning for EEG signal analysis. *Cognitive Neurodynamics*, 15(3), 369-388.
12. Hassan, M. M., Gumaei, A., Alsanad, A., Alrubaian, M., & Fortino, G. (2020). A hybrid deep learning model for efficient intrusion detection in big data environment. *Information Sciences*, 513, 386-396.
13. Hatim, H. A., Alyasseri, Z. A. A., & Jamil, N. (2025). A recent advances on autism spectrum disorders in diagnosing based on machine learning and deep learning. *Artificial Intelligence Review*, 58(10), 313.

14. Krauze, A. V., Zhuge, Y., Zhao, R., Tasci, E., & Camphausen, K. (2022). AI-driven image analysis in central nervous system tumors-traditional machine learning, deep learning and hybrid models. *Journal of biotechnology and biomedicine*, 5(1), 1.
15. Liaqat, S., Dashtipour, K., Arshad, K., Assaleh, K., & Ramzan, N. (2021). A hybrid posture detection framework: Integrating machine learning and deep neural networks. *IEEE Sensors Journal*, 21(7), 9515-9522.
16. Loukil, Z. (2025). A Hybrid Approach to Intelligent Prediction of Medical Conditions A Framework for Advancing Medical Diagnostics through Novel Hybrid Deep Learning Models DenCeption and HyBoost for Enhanced Feature Extraction and Predictive Accuracy in Medical Image Analysis (Doctoral dissertation, University of Gloucestershire).
17. Naik, N., Surendranath, N., Raju, S. A. B., Madduri, C., Dasari, N., Shukla, V. K., & Patil, V. (2025). Hybrid deep learning-enabled framework for enhancing security, data integrity, and operational performance in Healthcare Internet of Things (H-IoT) environments. *Scientific Reports*, 15(1), 31039.
18. Rana, M., & Bhushan, M. (2023). Machine learning and deep learning approach for medical image analysis: diagnosis to detection. *Multimedia Tools and Applications*, 82(17), 26731-26769.
19. Sajjanar, R., Dixit, U. D., & Vagga, V. K. (2024). Advancements in hybrid approaches for brain tumor segmentation in MRI: a comprehensive review of machine learning and deep learning techniques. *Multimedia Tools and Applications*, 83(10), 30505-30539.
20. Utsha, U. T., & Morshed, B. I. (2024). CardioHelp: A smartphone application for beat-by-beat ECG signal analysis for real-time cardiac disease detection using edge-computing AI classifiers. *Smart Health*, 31, 100446.
21. Wang, H., Li, S., Song, L., Cui, L., & Wang, P. (2019). An enhanced intelligent diagnosis method based on multi-sensor image fusion via improved deep learning network. *IEEE Transactions on Instrumentation and measurement*, 69(6), 2648-2657.
22. Zabihi, F., Safara, F., & Ahadzadeh, B. (2024). An electrocardiogram signal classification using a hybrid machine learning and deep learning approach. *Healthcare Analytics*, 6, 100366.
23. Zheng, S., Zhou, X., Zhang, L., Qi, P., Qiu, K., Zhu, J., & Yang, X. (2023). Toward next-generation signal intelligence: A hybrid knowledge and data-driven deep learning framework for radio signal classification. *IEEE Transactions on Cognitive Communications and Networking*, 9(3), 564-579.
24. Zulfiqar, Z., Zulfiqar, Z., Khan, N. A., Khan, M. A., Din, I. U., Almogren, A., ... & Martínez, M. G. (2024). Enhancing Hyperspectral Image Analysis With Deep Learning: An Innovative Hybrid Approach for Improved Detection and Classification Using Deep-Detect. *IEEE Access*.