SHM PUBLISHER
SURYA HIJAU MANFAAT

Journal of Soft Computing Exploration

Homepage: https://shmpublisher.com/index.php/joscex

JOSCEX

p-ISSN: 2746-7686
e-ISSN: 2746-0991

.

A hybrid VGG16 and random forest model for multi-class ischemic heart disease detection via ecg image analysis

Rizka Dian Safitri[1], Christy Atika Sari[2], Noorayisahbe Mohd Yaacob[3]
[1,2]Department of Informatics Engineering, Universitas Dian Nuswantoro, Indonesia
[3]Center for Software Technology and Management (SOFTAM), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Selangor, Malaysia

Article Info

Article history:

Received May 18, 2026
Revised July 7, 2026
Accepted July 18, 2026

Keywords:

Electrocardiogram
Myocardial ischemia
Image classification
Random forest
VGG16

ABSTRACT

A cardiac illnesses, mainly ischemic cardiac conditions, are a primary factor behind global fatality rates. Clinical diagnosis generally relies on manual interpretation of Electrocardiogram (ECG) signals, which is subjective and time-consuming, compounded by the challenge of limited access to raw ECG data on commercial devices. This research focuses on designing an automated diagnosis system based on ECG image analysis using the VGG16 Convolutional Neural Network (CNN) architecture through a knowledge transfer method. A total of 8,268 ECG images from 689 patients, divided into three categories Normal, Abnormal Heartbeat, and History of Myocardial Infarction (MI) were evaluated in this study. Test results demonstrated that the VGG16 architecture integrated with Random Forest produced the most optimal performance, with a test accuracy of 93.27%, an F1-Score of 93.25%, along with an Area Under the Curve (AUC) value of 0.991. This combined model successfully detected Normal images without any prediction errors. This computational image feature extraction approach has proven effective in reducing diagnostic subjectivity and holds strong potential for application as a fast and consistent clinical decision support system across various healthcare facilities. Nevertheless, multicenter validation on a more diverse population is still required to ensure the model's clinical generalizability.

Corresponding Author:

Christy Atika Sari,
Department of Informatics Engineering,
Universitas Dian Nuswantoro
Imam Bonjol 207, Semarang, Central Java, Indonesia
Email: christy.atika.sari@dsn.dinus.ac.id
https://doi.org/10.52465/joscex.v7i3.121

1. INTRODUCTION

Cardiac illnesses, mainly ischemic cardiac conditions, are a primary factor behind global fatality rate. According to the WHO, around 17.9 million fatalities per year or 32% of total worldwide fatalities are caused by cardiovascular diseases [1]. In Indonesia, there are 651,481 deaths per year due to this disease, with 245,343 of them caused by coronary cardiac illness [2]. Ischemic heart conditions appear due to insufficient blood delivery to the cardiac muscle and can progress to angina pectoris or life-threatening myocardial infarction (MI)[3].

The electrocardiogram (ECG) is the primary method for non-invasively diagnosing myocardial ischemia [4]. In ischemic conditions, ECG signals exhibit characteristic changes, such as abnormalities like ST-segment decline or elevation and T-wave abnormalities [5] [6]. Despite its clinical use, manual ECG evaluation needs professional skills, is relatively slow, and is susceptible to inter-physician variability. Furthermore, most commercial ECG devices provide only visual images without access to raw numerical data [7], necessitating an image-based analysis approach for the establishment of a more comprehensive automated diagnostic system.

The evolution of deep learning techniques, particularly Convolutional Neural Network (CNN)-based models, has proven effective in automatically and accurately deriving features from medical images [8] [9]. Systematic reviews also confirm the performance of deep learning in comprehensive analysis of medical images and signals [10] [11]. The VGG16 architecture, utilizing a transfer learning approach, has been implemented for various medical image analyses with promising results [12] [13]. This approach has also proven effective in transfer learning-based ECG image classification with consistent performance [14] [15]. Several related research has further shown the effectiveness of CNNs for ECG image categorization [16] [17]. Other research further confirms the potential of CNN-based methods for detecting heart disease via ECG [18] [19]. Recent studies also show similar results in multi-class ECG classification using image-based approaches [20] [21]. Based on this, this study is intended to design an ECG image-based classification model for ischemic heart disease using the VGG16 architecture with a transfer learning approach combined with machine learning classification methods, to produce a clinical decision support system that is fast, consistent, and independent of raw ECG signal data.

This study was conducted through a series of systematic stages designed to identify ischemic cardiac conditions using ECG data images. Broadly speaking, the research workflow consists of six main stages, namely: (1) dataset collection, (2) data preprocessing, (3) feature extraction using the VGG16 CNN, (4) classification model design (VGG16 Full Model and hybrid VGG16+ML), (5) model training and testing, and (6) performance assessment. Each stage is interrelated and plays a crucial role in producing an optimal classification model. Unlike previous studies that apply VGG16 solely as an end-to-end classifier, this study addresses a critical clinical and technical gap by proposing a two-stage hybrid pipeline. The novelty lies not merely in combining methods, but in specifically utilizing VGG16 exclusively as a deep feature extractor to bypass the black-box limitations and overfitting risks of pure CNNs on limited medical data, feeding directly into a Random Forest ensemble. Furthermore, this specific hybrid architecture is applied to multi-class chronic ischemic heart disease ECG images, overcoming the common limitation of commercial ECG machines that do not export raw signal data. This combination and clinical application have not been previously reported in the literature.

Therefore, the main contributions of this study are threefold. First, we design a novel hybrid pipeline that integrates VGG16 as a deep feature extractor with Random Forest as the final classifier, specifically optimized for multi-class ischemic heart disease detection. Second, we conduct a rigorous comparative evaluation of three modeling approaches (VGG16 Full Model, VGG16+SVC, and VGG16+Random Forest) using a standardized dataset to identify the most robust architecture. Third, we demonstrate that the proposed VGG16+Random Forest hybrid achieves superior performance with 93.27% accuracy, a 93.25% F1-Score, and an Area Under the Curve (AUC) of 0.991, significantly outperforming single CNN-based approaches and offering a highly viable tool for clinical decision support.

2. METHOD

Dataset

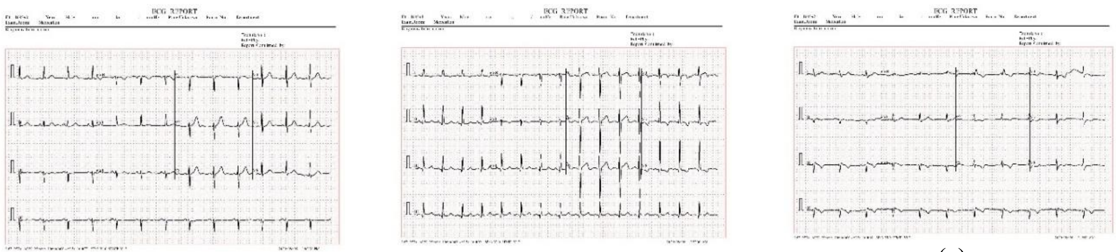
The dataset used is the Cardiac Patient ECG Image Collection (Version 2, CC BY 4.0 license) [22]. This dataset was developed at the Ch. Pervaiz Elahi medical cardiology institution in Multan, Pakistan, and initially comprised four classes of cardiac conditions: Normal, Abnormal Cardiac Rhythm, Previous Myocardial Infarction, and Acute Myocardial Infarction (AMI). Prior to data cleaning, the original dataset totaled 928 patients and 11,136 ECG images. To align the dataset strictly with the research scope, a deliberate data cleaning process was performed to exclude the AMI class, as acute conditions were irrelevant to our focus on chronic ischemic conditions. After this targeted removal, the refined dataset utilized in this study comprised 689 patients with a total of 8,268 ECG images (12 images per patient, representing the 12 standard ECG leads). The data is divided into three classes: Normal, Abnormal Heartbeat, and History of Myocardial Infarction (MI). The final distribution of the research dataset is presented in Table 1. To provide visual context for the dataset, Figure 1 presents representative ECG image samples from each of the three classes included in this study.

.

Panel (a) shows a normal ECG trace characterized by regular rhythm, normal PR interval, and physiologically normal ST-T segments. Panel (b) illustrates an abnormal heartbeat with evident conduction irregularities and morphological deviations from the normal sinus rhythm. Panel (c) displays a typical ECG recording from a patient with a history of myocardial infarction (MI), often featuring pathological Q waves, ST-segment elevation or depression, and T-wave inversions. These visual examples highlight the subtle yet distinguishable morphological differences between classes that the proposed classification model aims to learn and differentiate.

Table 1. Distribution of ECG image data across each class

Class	Number of Patients	Number of Images	Percentage (%)
Normal	148	1,776	21.48
Abnormal Heartbeat	207	2,484	30.05
History of MI	334	4,008	48.47
Total	689	8,268	100.00



(a) (b) (c)

Figure 1. Representative ECG image samples from each class: (a) normal, (b) abnormal heartbeat, and (c) history of myocardial infarction (mi).

Data Preprocessing

Data preprocessing was carried out in five main stages. First, data splitting was performed with a ratio of 70% for training, 15% for validation, and 15% for testing with stratified splitting to maintain class proportionality. Second, all ECG images were resized to 224x224 pixels using the *cv2.resize()* function, in accordance with the VGG16 architecture's input dimension requirements (224x224x3) as defined in this study's code implementation. Third, the pixel data were normalized within the interval [0, 255] to [0, 1] to accelerate training convergence. Fourth, label encoding (Normal=0, Abnormal Heartbeat=1, History of MI=2) and one-hot encoding are applied for the VGG16 Full Model. Fifth, on-the-fly data augmentation includes 10-degree rotation, 10% width/height shift, and horizontal flipping to reduce the risk of overfitting [23]. The mathematical formulations and transformation matrices of each augmentation technique applied in this study are explicitly described as follows. Image resizing transforms each ECG image (I) to a fixed dimension of 224×224 pixels, expressed in equation (1). Rotation applies a 2D transformation matrix with a rotation angle $\theta = 10°$ to each pixel coordinate (x, y), generating new coordinates (x', y') as in equation (2). Width shift translates pixel coordinates horizontally by a random offset Δx within 10% of the total image width W, as shown in equation (3). Height shift translates pixel coordinates vertically by a random offset Δy within 10% of the total image height H, formulated in equation (4). Horizontal flipping mirrors the image along the vertical axis. Assuming zero-based indexing, it maps each pixel coordinate as in equation (5). These augmentation techniques are applied strictly on-the-fly during the training phase to increase data variability without permanently modifying the original dataset.

$$I_{resized} = \text{resize}(I, 224 \times 224) \quad (1)$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2)$$

$$x' = x + \Delta x, \quad \Delta x \in [-0.1W, 0.1W] \quad (3)$$

$$y' = y + \Delta y, \quad \Delta y \in [-0.1H, 0.1H] \quad (4)$$

$$x' = W - x - 1, \quad y' = y \quad (5)$$

Feature Extraction Using the VGG16 CNN

VGG16 is a CNN model from Oxford University's Visual Geometry Group, consisting of 16 trained layers (13 convolutional + 3 fully connected) utilizing consistent 3×3 convolutional kernels [14]. The architectural configuration applied in this study is detailed in Table 2. This research utilizes VGG16 weights previously trained on the ImageNet dataset via a transfer learning approach. The VGG16 base layers are loaded without the fully connected head (include_top=False), and a Global Average Pooling layer is appended to decrease the spatial feature dimensions into a flat 512-dimensional vector. This extracted vector serves as the direct input for both the VGG16 Full Model's custom head and the hybrid models (SVC and Random Forest) [24].

Table 2. VGG16 architectural structure

Block	Layer	Filter/Unit	Output Shape	Activation
Block 1	Conv2D x2 + MaxPool	64	112x112x64	ReLU
Block 2	Conv2D x2 + MaxPool	128	56x56x128	ReLU
Block 3	Conv2D x3 + MaxPool	256	28x28x256	ReLU
Blocks 4 & 5	Conv2D x3 + MaxPool	512	7x7x512	ReLU
Output*	GlobalAvgPool2D	512	1x512	-

Model Training

The experimental design involves three classification setups, as summarized in Table 3. The VGG16 Full Model utilizes a fine-tuning strategy by freezing the initial layers of VGG16 and retraining only the last 10 layers, alongside a custom classification head (Dense 256, Dropout 0.3, and a 3-class Softmax activation). The VGG16+SVC hybrid model employs a Support Vector Classifier with a Radial Basis Function (RBF) kernel (C=10, gamma='scale') following feature standardization. Finally, the proposed VGG16+Random Forest model relies on an ensemble of 100 decision trees (n_estimators=100) with a maximum depth of 15 (max_depth=15), intended to produce more stable predictions by reducing variance [25].

Table 3. Model configuration

Model	Description	Key Parameters
VGG16 Full Model	Fine-tuned VGG16 with custom classification head	Last 10 layers trainable, Dense (256), Dropout (0.3)
VGG16 + SVC	VGG16 feature extraction + SVM classifier	Kernel = RBF, C = 10, gamma = 'scale'
VGG16 + Random Forest	VGG16 feature extraction + Random Forest	n_estimators = 100, max_depth = 15

Training Implementation

The models were trained and tested using the finalized dataset of 8,268 ECG images, allocated into 5,779 training images, 1,241 validation images, and 1,248 testing images. The VGG16 Full Model was constructed using TensorFlow 2.x, utilizing the Adam optimizer with a starting learning rate of 1×10^{-4}, Categorical Cross-entropy as the loss function, and a batch size of 32. Training was configured for a maximum of 50 epochs alongside the EarlyStopping callback (patience=10) to mitigate overfitting, and the ReduceLROnPlateau callback (factor=0.2, patience=5, min_lr=1×10^{-7}) for adaptive optimization [16]. This network comprises 14,846,787 parameters, out of which 13,701,379 parameters in the last 10 layers are trainable. The hybrid models (VGG16+SVC and VGG16+Random Forest) were trained via scikit-learn utilizing the 512-dimensional feature vectors extracted by the VGG16 base, with random_state=42 set to guarantee reproducibility. The core Python implementation snippet is depicted in Figure 2, and a visual summary of the entire systematic workflow is presented in Figure 3.

```
def load_images_from_folder(base_path, class_folders, img_size=(224, 224)):
    for class_idx, folder_name in enumerate(class_folders):
        img = cv2.imread(img_path)
        img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
        img = cv2.resize(img, img_size)
        img = img / 255.0
base_model = VGG16(weights='imagenet', include_top=False, input_shape=(224, 224, 3))
feature_extractor = Model(inputs=base_model.input,
    outputs=GlobalAveragePooling2D()(base_model.output))
X_train_features = feature_extractor.predict(X_train, batch_size=32)
for layer in vgg_base.layers[:-10]:
    layer.trainable = False
model_vgg = Sequential([vgg_base, GlobalAveragePooling2D(),
    Dense(256, activation='relu'), Dropout(0.3),
    Dense(3, activation='softmax')])
model_vgg.compile(optimizer=Adam(lr=1e-4),
    loss='categorical_crossentropy', metrics=['accuracy'])
model_vgg.fit(train_datagen.flow(X_train, y_train, batch_size=32),
    epochs=50, callbacks=[early_stop, reduce_lr])
rf_model = RandomForestClassifier(n_estimators=100, max_depth=15,
    min_samples_split=10, random_state=42)
rf_model.fit(X_train_features, y_train_int)
y_pred_rf = rf_model.predict(X_test_features)
```

Figure 2. Code implementation snippet of the proposed model

.

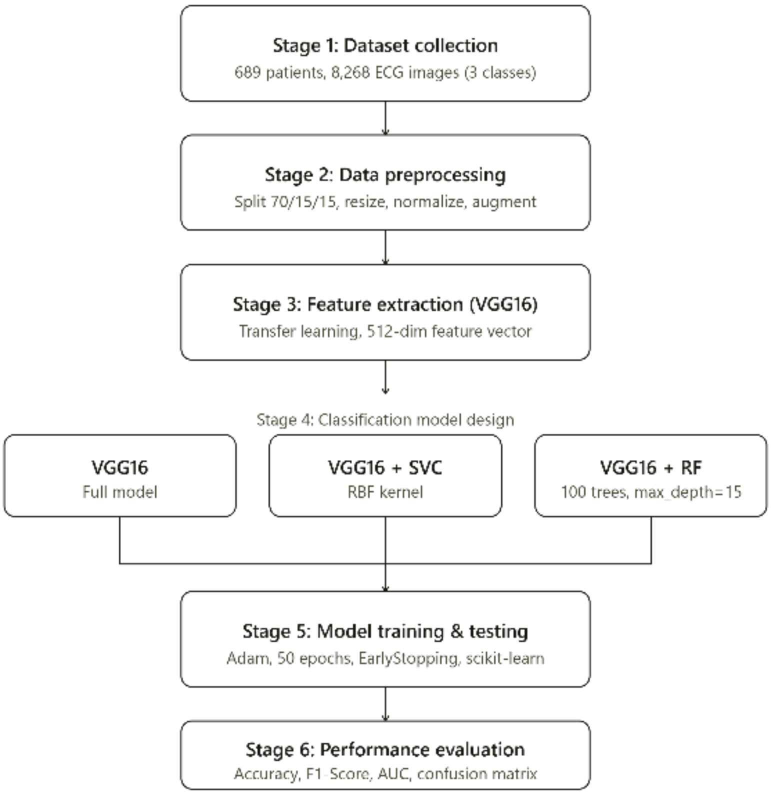


Figure 3. Block diagram of the proposed methodology

Model Performance Evaluation

The performance of the evaluated models is quantified using standard classification metrics derived from the confusion matrix, including Accuracy, Precision, Recall (Sensitivity), and the F1-Score, as mathematically defined in Table 4. In addition to these scalar metrics, the confusion matrix is utilized to identify specific misclassification patterns among the classes. Furthermore, Receiver Operating Characteristic (ROC) curves and the AUC are calculated using a One-vs-Rest (OvR) approach to evaluate the multi-class discriminative capacity of each model [26] [27]. A notable limitation of this current experimental design is the reliance on a single train-validation-test split without cross-validation, combined with the image-wise splitting approach. These factors limit the ability to fully guarantee model stability and pose a potential risk of data leakage. To definitively validate the clinical generalizability and stability of the reported performance metrics, future work must incorporate rigorous patient-wise k-fold cross-validation.

Table 4. Model performance evaluation metrics

Metric	Formula	Description
Accuracy	(TP + TN) / (TP + FP + FN + TN)	Proportion of correct predictions out of all data
Precision	TP / (TP + FP)	The model's positive prediction accuracy
Recall	TP / (TP + FN)	Ability to identify all positive cases
F1-Score	2 × (Precision × Recall) / (Precision + Recall)	Harmonic mean of precision and recall

3. RESULTS AND DISCUSSIONS

Performance evaluation was carried out on 104 test images that the model had never seen before. The metrics involved include sensitivity, precision, F1 Score, test accuracy, and generalization capability analysis via the Confusion Matrix along with the ROC curve for three modeling scenarios: VGG16 Full Model, VGG16 integrated with Support Vector Classifier (SVC), and VGG16 integrated with Random Forest (RF). Training the VGG16 Full Model demonstrated highly stable convergence. Training data accuracy consistently increased to 92.93% by the end of the epoch, in line with validation accuracy reaching 87.50%. The loss value was significantly reduced from 1.10 to 0.23 on the training data and to 0.41 on the validation data. Overfitting was effectively prevented through the implementation of the EarlyStopping and ReduceLROnPlateau callbacks, which adaptively reduced the learning rate from 1×10^{-4} to 4×10^{-6} [16]. This parameter adjustment strategy proved to optimize the convergence process without triggering a spike in error during the validation stage. The comparison results confirm that the VGG16 hybrid architecture integrated with Random Forest outperforms the others, achieving a test accuracy of 93.27% and an average F1 score of 93.25%. The reliability of this ensemble approach is demonstrated by achieving perfect sensitivity (100%) for identifying Normal images and perfect precision (100%) for the Abnormal Heartbeat class. In terms of computational efficiency, the hybrid models (SVC and Random Forest) demonstrated a substantial advantage, requiring only 0.2 and 0.5 seconds for training respectively, compared to 255.5 seconds for the VGG16 Full Model, highlighting that feature extraction prior to classification significantly reduces training overhead without compromising accuracy.

Table 5. Training and Validation Results of Each Model

Model	Training Accuracy (%)	Validation Accuracy (%)	Test Acc (%)	AUC	Training Time (s)
VGG16 Full Model	92.93	87.5	95.00	0.991	255.5
VGG16 + SVC	91.35	91.35	91.35	0.990	0.2
VGG16 + Random Forest	93.27	93.27	93.27	0.991	0.5

VGG16 full model

The VGG16 Full Model reached an accuracy rate of 92.93% on the training data and 87.50% on the validation data. The loss value was successfully reduced from 1.10 to 0.23 on the training data and to 0.41 on the validation data. Figure 4 shows the accuracy and loss graphs during training, which consistently converged. The classification report shows macro-average precision, recall, and F1-Score of 0.95, 0.94, and 0.94, successively. The best performance was achieved in the Abnormal Heartbeat class (F1-Score 0.96), followed by Normal (0.95), while History of MI recorded the lowest F1-Score of 0.88. The confusion matrix demonstrates that prediction outputs predominantly appear on the main diagonal with an overall accuracy of 95%. Classification errors primarily occur in the History of MI class, which is sometimes classified as Normal or Abnormal Heartbeat due to overlapping feature patterns. Figure 5 presents the ROC curves for the VGG16 Full Model (left) and VGG16 + SVC (right). The VGG16 Full Model achieved a Macro AUC of 0.991 with class-specific AUCs: Normal 0.993, Abnormal Heartbeat 1.000, and History of MI 0.975. VGG16 + SVC achieved a Macro AUC of 0.990 with AUCs of: Normal 0.996, Abnormal Heartbeat 0.991, and History of MI 0.978.

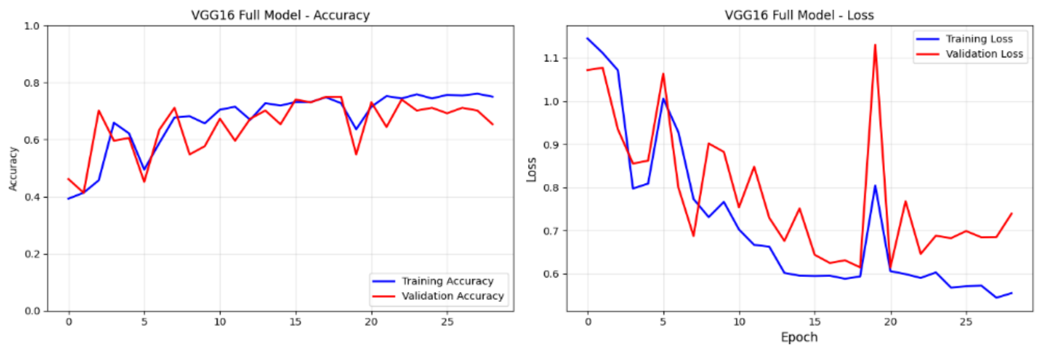


Figure 4. VGG16 full model accuracy and loss graph

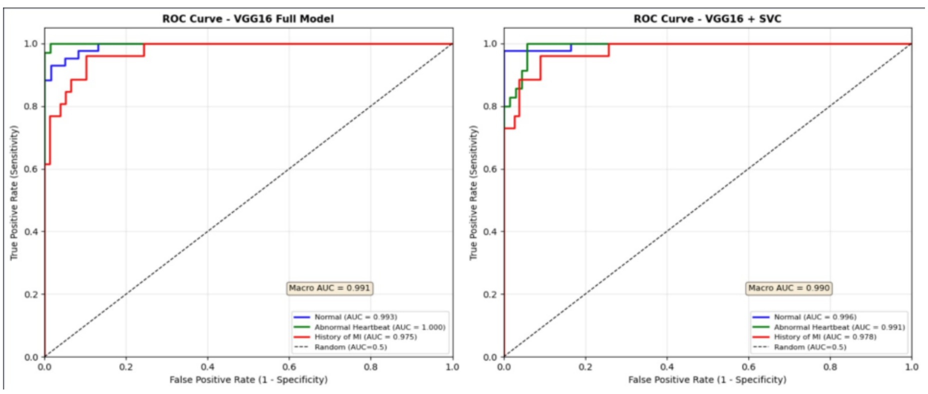


Figure 5. ROC curves for the VGG16 full model (left) and VGG16 + SVC (right)

VGG16 + SVC

The VGG16 + SVC model uses 512-dimensional features extracted from VGG16 as input for the Support Vector Classifier with an RBF kernel (C=10, gamma='scale'). This model attained an accuracy of 91.35% on the test data. The VGG16 + SVC classification report shows macro-average precision, recall, and F1-Score of 0.91, 0.91, and 0.91, successively. The ROC curve in Figure 5 (right panel) demonstrates excellent discriminatory performance across all classes, with a Macro AUC of 0.990.

VGG16 + Random Forest

The VGG16 + Random Forest model is a hybrid method that integrates deep visual features of VGG16 with the Random Forest ensemble algorithm (100 trees, max_depth=15). This model obtained the top test performance with an accuracy of 93.27% with an average F1-Score of 93.25%. The classification report shows superior precision, recall, and F1-Score values, including perfect sensitivity (100%) for the Normal class and perfect precision (100%) for the Abnormal Heartbeat class, demonstrating the reliability of the ensemble approach in minimizing prediction errors. Figure 6 presents the confusion matrix for VGG16 + Random Forest. The model successfully classified all 43 Normal images correctly (0 errors), 32 out of 35 Abnormal Heartbeat images, and 22 out of 26 History of MI images. Figure 7 shows the ROC curve for VGG16 + RF, which achieved a maximum Macro AUC of 0.991. AUC per class: Normal 0.993, Abnormal Heartbeat 0.995, and History of MI 0.978. This achievement surpasses previous research results, which generally fall within the range of 0.94–0.98 for cardiac ischemia detection [6][28].

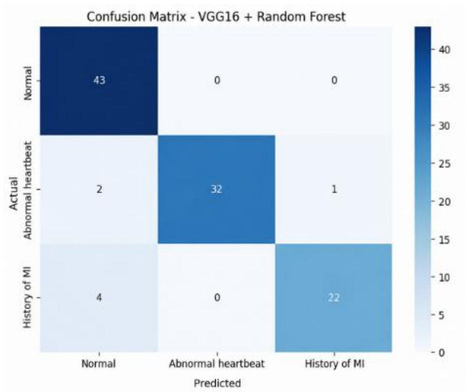


Figure 6. Confusion Matrix for VGG16 + RF

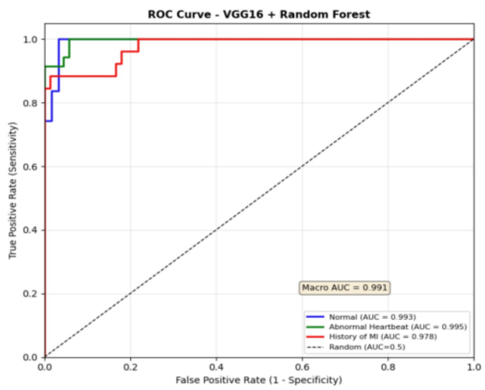


Figure 7. ROC Curve for VGG16 + RF

Comparison with Previous Studies

The outcomes of this study were contrasted with several previous studies that used a similar approach in classifying ECG images or heart disease using deep learning. Table 6 presents a comparison of the results of the best-performing model from this study with relevant prior studies. Based on the comparison in Table 6, this study obtained an accuracy of 93.27%, exceeding the results of the three compared studies that used the VGG16 architecture or other combinations of CNN transfer learning. The main difference in this study lies in the use of a hybrid approach combining VGG16 as a feature extractor and Random Forest as the final classifier for multi-class classification of cardiac ischemia ECG images. This approach is able to improve the stability and performance of classification compared to the use of a single CNN model [25], thus demonstrating that the integration of deep feature extraction and ensemble machine learning is effectively applied to the classification of ECG images of heart disease [24] [25].

Table 6. Comparison with previous studies

Title	Year	Method Differences	Dataset	Sample	Accuracy (%)
Rizaqi & Tahyudin: Comparative analysis of VGG16 and ResNet50 in cardiac ECG image classification [29]	2025	Comparing VGG16 and ResNet50 as single CNN models without separate feature extraction or ensemble classifiers	ECG Images Dataset (Mendeley)	Cardiac ECG Images	81.0
Azizah et al.: Implementation of VGG16 for Cardiomegaly Detection on Chest X-Rays	2024	Single VGG16 with Attention Mechanism for Binary Classification of Cardiomegaly Based on Chest X-rays	Kaggle Chest X-Ray Cardiomegaly Dataset	Chest X-ray images	78.75
Berliani et al.: Comparison of Lung X-ray Image Classification using ResNet-50 and VGG-16	2023	Transfer learning of VGG16 and ResNet50 for multi-class classification of lung X-ray images	COVID-19 Radiography Dataset	Lung X-ray images	91.45

Proposed Method (This Study)	2025	VGG16 + Random Forest (hybrid feature extraction + ensemble)	ECG Images (Mendeley)	8,268 ECG images	93.27

Discussion

The experimental results demonstrate that the VGG16 + Random Forest hybrid model achieved the best overall performance with an accuracy of 93.27%, an F1-Score of 93.25%, and an AUC of 0.991. These findings align with transfer learning principles, which state that pre-trained weights on large-scale datasets such as ImageNet can be effectively adapted for medical image classification tasks with limited training data [24]. The superiority of the hybrid approach over the VGG16 Full Model can be attributed to the complementary strengths of deep feature extraction and ensemble learning. While VGG16 effectively captures hierarchical visual features from ECG images through its 16-layer architecture trained on ImageNet, Random Forest provides robust decision boundaries through ensemble averaging of 100 decision trees, thereby reducing variance and improving generalization [25]. This is consistent with Napitupulu et al. [25], who reported that ensemble-based classifiers significantly improve prediction stability in ECG signal classification. The perfect sensitivity for the Normal class (100%) indicates that the model successfully identifies all healthy cases, minimizing false negatives that could lead to missed diagnoses. The perfect precision for the Abnormal Heartbeat class (100%) indicates that when the model predicts an abnormal heartbeat, it is always correct, minimizing false positives that could lead to unnecessary interventions.

The VGG16 + SVC model achieved a competitive accuracy of 91.35%, indicating that support vector classifiers are also capable of leveraging VGG16 features effectively. However, its performance was slightly lower than Random Forest, likely due to the sensitivity of SVM to high-dimensional feature spaces without further dimensionality reduction. The RBF kernel with C=10 and gamma='scale' provided reasonable decision boundaries, but the maximum margin principle of SVM may have been less effective than the ensemble averaging approach of Random Forest for this particular feature distribution. Compared to previous studies, the proposed model surpasses Rizaqi & Tahyudin [29] (81.00%), Azizah et al. [13] (78.75%), and Berliani et al. [14] (91.45%), confirming the advantage of combining VGG16 feature extraction with an ensemble classifier for multi-class ischemic heart disease classification. While Makhir et al. [6] reported a slightly higher accuracy of 94.20%, their study used raw ECG signal data rather than ECG images, making direct comparison inappropriate given the different input modalities. The proposed method's independence from raw signal data represents a practical advantage for deployment with existing commercial ECG devices. The confusion matrix analysis reveals that misclassification primarily occurred in the History of MI class, with 4 out of 26 images misclassified. This is likely due to overlapping visual feature patterns between History of MI and Abnormal Heartbeat classes, as both conditions may present similar ST-T wave changes on ECG images [30]. This challenge is also reported by Makhir et al. [6], who noted that chronic ischemic conditions often share visual characteristics with arrhythmia patterns in image-based analysis. Future work should investigate class-specific feature weighting or attention mechanisms to better distinguish these overlapping patterns. Despite these promising results, several limitations were identified. First, the model was trained and evaluated on a single dataset from one medical institution, which may limit its generalizability across different ECG devices and patient populations. Second, the absence of cross-validation means that model stability across different data splits has not been fully verified. Third, the image-wise splitting approach, while practical for lead-level analysis, may not fully reflect patient-level generalizability. Multicenter validation with a more diverse patient population and implementation of patient-wise cross-validation are therefore recommended to ensure the clinical applicability of the proposed system.

4. CONCLUSION

This study successfully developed an ECG image-based ischemic heart disease classification system using the VGG16 architecture with a transfer learning approach. The VGG16 + Random Forest hybrid model yielded the best performance with an accuracy of 93.27%, an F1-Score of 93.25%, and an AUC of 0.991, and was able to identify all images in the Normal class without any prediction errors. This approach has proven effective as a clinical decision support system that does not rely on raw ECG signal data. Future studies are encouraged to explore more modern architectures such as EfficientNet or Vision Transformer and to conduct multicenter validation to ensure the model's clinical generalization.

.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Rizka Dian Safitri: conceptualization, methodology, data collection, data analysis, drafting the initial manuscript, and revising and editing the manuscript. **Christy Atika Sari**: Writing – review & editing, supervision. **Noorayisahbe Mohd Yaacob**: Validation**.**

DECLARATION OF COMPETING INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY

Supporting data for this study is available via Mendeley Data at the following link: https://data.mendeley.com/datasets/gwbz3fsgp8/2.

REFERENCES

[1] G. A. Roth *et al.*, "Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019," *J. Am. Coll. Cardiol.*, vol. 76, no. 25, pp. 2982–3021, Dec. 2020, doi: 10.1016/j.jacc.2020.11.010.

[2] F. R. Muharram *et al.*, "The 30 Years of Shifting in The Indonesian Cardiovascular Burden—Analysis of The Global Burden of Disease Study," *J. Epidemiol. Glob. Health*, vol. 14, no. 1, pp. 193–212, Feb. 2024, doi: 10.1007/s44197-024-00187-8.

[3] R. Myrtha, M. A. Wijayanto, D. P. Putri, and B. D. Hermawati, "ST-Depression in ST-Elevation Myocardial Infarction–Significance of Reciprocal Changes in Subtle ST-Segment Elevation," *Jurnal Biomedika dan Kesehatan*, vol. 7, no. 3, pp. 339–344, Nov. 2024, doi: 10.18051/JBiomedKes.2024.v7.339-344.

[4] M. Martínez-Sellés and M. Marina-Breysse, "Current and Future Use of Artificial Intelligence in Electrocardiography," *J. Cardiovasc. Dev. Dis.*, vol. 10, no. 4, p. 175, Apr. 2023, doi: 10.3390/jcdd10040175.

[5] J. C. Brown, T. E. Gerhardt, and E. Kwon, *Risk factors for coronary artery disease*. StatPearls Publishing, 2023.

[6] A. Makhir, M. H. El Yousfi Alaoui, and L. Belarbi, "Comprehensive Cardiac Ischemia Classification Using Hybrid CNN-Based Models," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 20, no. 03, pp. 154–165, Feb. 2024, doi: 10.3991/ijoe.v20i03.45769.

[7] S. Mishra *et al.*, "ECG Paper Record Digitization and Diagnosis Using Deep Learning," *J. Med. Biol. Eng.*, vol. 41, no. 4, pp. 422–432, Aug. 2021, doi: 10.1007/s40846-021-00632-0.

[8] P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," *Pattern Recognit. Lett.*, vol. 141, pp. 61–67, Jan. 2021, doi: 10.1016/j.patrec.2020.07.042.

[9] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1–1, 2021, doi: 10.1109/TPAMI.2021.3059968.

[10] E. Canessa and L. Tenze, "Morphing a Stereogram into Hologram," *J. Imaging*, vol. 6, no. 1, p. 1, Jan. 2020, doi: 10.3390/jimaging6010001.

[11] K. E. C. Vidyasagar, K. R. Kumar, and G. N. K. A. Sai, "Signal to image conversion and convolutional neural networks for physiological signal processing: A review," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–18, 2024.

[12] R. Klangbunrueang, P. Pookduang, W. Chansanam, and T. Lunrasri, "AI-Powered Lung Cancer Detection: Assessing VGG16 and CNN Architectures for CT Scan Image Classification," *Informatics*, vol. 12, no. 1, p. 18, Feb. 2025, doi: 10.3390/informatics12010018.

[13] D. E. Astuti and M. Hidayat, "Transfer Learning Analysis of VGG16 for Tuberculosis Detection," *TRIAC Journal of Electrical and Computer Engineering*, vol. 11, no. 1, pp. 31–36, Jun. 2024, doi: 10.21107/triac.v11i1.24371.

[14] T. Berliani, E. Rahardja, and L. Septiana, "Comparison of Lung X-ray Image Classification Capabilities Using ResNet-50 and VGG-16 Transfer Learning," *Journal of Medicine and Health*, vol. 5, no. 2, pp. 123–135, Aug. 2023, doi: 10.28932/jmh.v5i2.6116.

[15] S. N. M. Sayed Ismail, S. F. Abdul Razak, and N. A. Ab. Aziz, "Transfer learning for improved electrocardiogram diagnosis of cardiac disease: exploring the potential of pre-trained models," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 5, pp. 3288–3300, Oct. 2024, doi: 10.11591/eei.v13i5.7053.

[16] D. Siahaan, *Classification of Electrocardiogram (ECG) Signals in Cardiovascular Diseases Using a Convolutional Neural Network (CNN) Model*. 2022.

[17] S. Hadiyoso, F. Fahrozi, Y. S. Hariyani, and M. D. Sulistyo, "Image Based ECG Signal Classification Using Convolutional Neural Network," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 18, no. 04, pp. 64–78, Mar. 2022, doi: 10.3991/ijoe.v18i04.27923.

[18] M. Irhamsyah, M. Melinda, and Yunidar, "Implementation of discrete wavelet transform and Xception for ECG image classification," *JURNAL INFOTEL*, vol. 17, no. 1, 2025, doi: 10.20895/infotel.v17i1.

[19] A. Jaya Prakash, A. Nasreddine Belkacem, I. M. Elfadel, H. F. Jelinek, and M. Atef, "Advances in machine and deep learning for ECG beat classification: a systematic review," *Front. Digit. Health*, vol. 7, Nov. 2025, doi: 10.3389/fdgth.2025.1649923.

[20] O. A. Oke and N. Cavus, "Electrocardiogram image classification for six classes of heart diseases," *Iran Journal of Computer Science*, vol. 8, no. 2, pp. 419–439, Jun. 2025, doi: 10.1007/s42044-025-00227-x.

[21] R. Selvi, M. Robinson Joel, and M. Navaneethakrishnan, "SCNN-VGG16: Siamese Convolutional Neural Network-Visual Geometry Group 16 for Cardiovascular Disease Detection Using ECG," *Biomed. Eng. (Singapore).*, vol. 38, no. 05, Oct. 2026, doi: 10.4015/S1016237225500528.

[22] A. H. Khan, M. Hussain, and M. K. Malik, "ECG Images dataset of Cardiac and COVID-19 Patients," *Data Brief*, vol. 34, p. 106762, Feb. 2021, doi: 10.1016/j.dib.2021.106762.

[23] F. Purwaningtias, "Integration of transfer learning in Inceptionv3 for automatic identification of tropical fruit types," *JCOSIS*, vol. 2, no. 2, pp. 39–44, 2025.

[24] E. Puls, M. Todescato, and J. Carbonera, "An Evaluation of Pre-Trained Models for Feature Extraction in Image Classification," in *Proceedings of the 26th International Conference on Enterprise Information Systems*, SCITEPRESS - Science and Technology Publications, 2024, pp. 465–476. doi: 10.5220/0012622300003690.

[25] P. B. R. Napitupulu, E. T. D. Tambunan, R. R. Prayoga, D. Sitanggang, Emilliano, and M. Turnip, "Random Forest-Based Classification of Electrocardiogram Signals for Early Detection of Cardiac Abnormalities," in *2025 IEEE International Conference on Artificial Intelligence and Mechatronics Systems (AIMS)*, IEEE, May 2025, pp. 1–6. doi: 10.1109/AIMS66189.2025.11229733.
[26] M. M. Fahmy, "Confusion Matrix in Binary Classification Problems: A Step-by-Step Tutorial," *Journal of Engineering Research*, vol. 6, no. 5, 2022.
[27] C. Kaope and Y. Pristyanto, "The Effect of Class Imbalance Handling on Datasets Toward Classification Algorithm Performance," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 22, no. 2, pp. 227–238, Mar. 2023, doi: 10.30812/matrik.v22i2.2515.
[28] M. Khalid, C. Pluempitiwiriyawej, and S. Wangsiripitak, "The applications of deep learning in ECG classification for disease diagnosis: A systematic review," *Engineering Journal*, vol. 28, no. 4, 2024, doi: 10.4186/ej.2024.28.4.245.
[29] H. Rizaqi and I. Tahyudin, "Comparative analysis of VGG16 and ResNet50 model performance in cardiac ECG image classification," *Journal of Applied Informatics and Computing*, vol. 9, no. 1, 2025, doi: 10.30871/jaic.v9i1.9350.
[30] A. Ameen, I. E. Fattoh, T. Abd El-Hafeez, and K. Ahmed, "Advances in ECG and PCG-based cardiovascular disease classification: a review of deep learning and machine learning methods," *J. Big Data*, vol. 11, no. 1, p. 159, Nov. 2024, doi: 10.1186/s40537-024-01011-7.

.