Heart Attack Prediction Using an Enhanced Convolutional AlexNet Model for ECG Image Classification

Omoniyi Joseph Asabia[1], Abayomi Oluwaseun Japinye[2], Olusegun J. Omotosho[3], Chigozirim U. Ajaegbu[4], *Raymond Osi Alenoghena[5], Nwamaka Grace Ajaegbu[6]

[1,3,4]Department of Computer Science, School of Computing, Babcock University, Ilishan-Remo, Ogun State, Nigeria

[2]Department of Banking Examination, Central Bank of Nigeria, Lagos

[5]Department of Economics, Caleb University Imota, Lagos

[6]Department of Social Sciences, Yaba College of Technology, Lags

DOI: https://doi.org/10.51244/IJRSI.2026.1310000056

Received: 16 October 2026; Accepted: 21 October 2026; Published: 31 October 2026

ABSTRACT

Cardiovascular disease (CVD) is still the greatest cause of death around the world, myocardial infarction (MI) making up a large number of the deaths related to CVD. Currently, diagnosis depends on manual analysis of electrocardiogram (ECG) traces, a time-consuming process prone to human error and variation between observers. The present study introduces an improved AlexNet-based convolutional neural network (CNN) which enables automated multi-class classification of ECG images into four categories: Normal Heartbeat (NH), Myocardial Infarction (MI), Previous Myocardial Infarction (PMI), and Abnormal Heartbeat (HB). The dataset consisted of 929 ECG images taken from a public Mendeley repository. We used a thorough preprocessing pipeline involving denoising, digitisation, augmentation, segmentation into thirteen ECG leads, contour-based signal extraction, Min–Max normalisation, and Principal Component Analysis (PCA), which reduced the 3,060 contour features to 400 principal components while retaining about 99% of the total variance. To modify the AlexNet architecture, we replaced the three-channel RGB input with a single-channel grayscale input; this required changing the convolutional kernel sizes and redesigning the output layer to include four Softmax neurons and a 0.5 dropout rate. The model was trained using the Adam optimiser with a learning rate of 0.001 and with a train–test split of 80:20. The results obtained were 99.44% accuracy, a macro-averaged F1-score of 99.2%, and an AUC-ROC of 96.06%, thereby exceeding the performance of a lightweight CNN, a SMOTE-enhanced CNN, and a Particle Swarm Optimisation-based neural network (PSO-NN). The fact that modified CNN architectures are able to accurately detect heart attacks from ECG data shows that they have the potential to be incorporated into clinical decision support systems in order to allow efficient and clinically relevant automated cardiac screening.

Keywords: Cardiovascular Disease Prediction; Heart Attack Detection; Deep Learning; Enhanced Alexnet; Convolutional Neural Network; ECG Image Classification; Myocardial Infarction

INTRODUCTION

Cardiovascular diseases (CVDs) are still one of the main causes of death all over the world, myocardial infarction (a heart attack) being one of the reasons for these deaths (El-Ansari & Ayoubi, 2026; Mishra & Singh, 2025). The World Health Organisation states that CVDs cause approximately 17.9 million deaths every year, which is nearly one third of all deaths worldwide. These diseases are particularly severe in low- and middle-income countries such as Nigeria (World Health Organisation, 2021). A heart attack occurs when

blood flow to the heart muscle is greatly reduced or blocked, typically because a plaque ruptures and a clot forms in a coronary artery. Predicting heart attack risk quickly and accurately remains difficult. Traditional diagnostic techniques such as electrocardiograms (ECGs), blood tests, and imaging require specialist knowledge, and manual ECG interpretation can be affected by human error and delays.

Deep learning has revolutionised medical diagnostics by being able to automatically detect subtle, complex patterns in medical images which conventional risk factor-based approaches often fail to pick up. In many instances, its performance is at least equal to that of the human expert (Mishra & Singh, 2025; Wang et al., 2025). CNNs are the go-to architecture for medical image analysis because they can extract fine spatial features directly from raw pixel data (LeCun et al., 2015; Omoniyi et al., 2026). Among CNN architectures, AlexNet (Krizhevsky et al., 2012) is one of the best-known for its early influence on deep learning in computer vision. More recent work has shown that AlexNet based models can still achieve the same level of performance as more modern and more complex networks when it comes to the classification of medical images (Hassan et al., 2025).

Despite recent advances, choosing the most suitable CNN architecture for predicting heart attacks from ECG data remains difficult. This difficulty arises because of the large number of competing models (for example, VGGNet, ResNet, DenseNet, and their variants), the absence of large annotated medical datasets, the limited transparency of deep learning models, and the high level of computational power required by the more advanced architectures. To address these issues, this study introduces an Enhanced AlexNet model designed to predict heart attacks from ECG images. The model is evaluated using a standard experimental setup and is compared with existing lightweight and metaheuristic-optimised options.

This work is concerned with the development of a heart attack prediction model using ECG data and an improved Convolutional AlexNet. The key aims are: (i) to develop an improved AlexNet for the prediction of heart attacks; (ii) to apply this model to a dataset of ECG images which has been carefully selected and pre-processed; and (iii) to compare the performance of the model against a number of existing benchmark methods.

Related Work

Electrocardiography is still one of the most commonly used non-invasive methods for detecting cardiac abnormalities, but since manual interpretation leads to diagnostic variability, automated approaches have been developed (Goldberger et al., 2000). Using convolutional filters, pooling layers, and fully connected layers, Convolutional Neural Networks can automatically extract hierarchical spatial features. Thus, manual feature engineering is no longer needed (LeCun et al., 2015). They have shown remarkable ability in recognising the subtle morphological changes in ECG waveforms linked to myocardial infarction (Litjens et al., 2017; Alenoghena et al., 2026).

AlexNet (Krizhevsky et al., 2012) attained a Top-5 error rate of 15.3% in the ImageNet Large Scale Visual Recognition Challenge, thus marking the beginning of the modern deep learning era in the field of computer vision by means of various innovations such as the use of ReLU activations, local response normalisation, dropout regularisation, and overlapping max-pooling. Later architectures adopted different compromises between depth, accuracy, and computational cost: VGGNet-16 (Simonyan & Zisserman, 2015) achieved greater depth by employing stacked 3×3 filters but still needed approximately 138 million parameters; GoogLeNet (Szegedy et al., 2015) incorporated Inception modules which brought the number of parameters down to about 6.8 million while maintaining accuracy; ResNet-50 (He et al., 2016) introduced residual connections in order to mitigate the vanishing-gradient problem in very deep networks; DenseNet-121 (Huang et al., 2017) encouraged the reuse of features by means of dense connections between layers; and EfficientNet-B4 (Tan & Le, 2019) used compound scaling of depth, width, and resolution. Recently, the Vision Transformer (Dosovitskiy et al., 2021) substituted convolution with self-attention but did so at the cost of requiring considerably larger training datasets.

In ECG-based cardiovascular disease prediction, three recently published studies are directly comparable to the one described here. The study by Hasan et al. (2025) presented an ensemble of a lightweight deep learning

approach, which increased computational efficiency but failed to include a thorough ECG preprocessing pipeline or to integrate a clinical decision-support system. Trigka and Dritsas (2025) combined a CNN with Correlation-Aware SMOTE in order to deal with the problem of class imbalance, achieving an accuracy of about 91.0% and an AUC of 0.90, yet they did not adopt a lead-specific segmentation strategy nor develop a deployable CDSS architecture. Reddy and Murthy (2025) advanced a hybrid Particle Swarm Optimisation–Neural Network (PSO-NN) framework and successfully obtained an accuracy of approximately 92.7% when deploying structured clinical data. The authors' approach excluded ECG-specific preprocessing steps such as noise removal and dimensionality reduction. The more extensive benchmarking exercise, the PhysioNet/Computing in Cardiology Challenge (Perez Alday et al., 2020), showed that ensemble machine learning could be used for 12-lead ECG classification, although it had concerns regarding generalizability on the hidden test data and did not incorporate the resulting models into a clinical decision-support setting.

Together, these studies show that CNN-based ECG classification is feasible, but they leave a gap which the current study fills by simultaneously incorporating (i) an improved AlexNet architecture together with adaptation to grayscale ECG input, (ii) a full thirteen-lead contour-extraction preprocessing pipeline, (iii) near-perfect four-class ECG classification accuracy, and (iv) an integrated design of a clinical decision-support system.

MATERIALS AND METHODS

Dataset

The ECG image dataset employed in this study was acquired from a publicly available repository on Mendeley Data (Khan & Hussain, 2021). It includes 929 colour ECG images in JPEG format and covers four diagnostic categories: 240 images of patients with Myocardial Infarction (MI), 233 images of patients having an Abnormal Heartbeat (HB), 172 images of patients with a recorded history of MI (Previous Myocardial Infarction, PMI), and 284 images of patients with Normal ECG results (NH).

Preprocessing Pipeline

Before model training, the multi-stage preprocessing pipeline was used. By means of the Python Imaging Library (PIL). It was found that all the images were RGB JPEG files with a uniform resolution; mean image noise, measured by the mean and standard deviation of the pixel-noise array, was reduced by an average of 67% as a result of the wavelet-based denoising. The images were then digitised and augmented using the Albumentations library, horizontal flipping, random adjustments of brightness and contrast, rotation, and shift-scale-rotate transformations being applied so as to increase the effective size of the training set by roughly four times. Each ECG image was divided into thirteen leads (that is, the standard twelve leads I, II, III, aVR, aVL, aVF and V1 to V6, together with one derived lead), and the electrical pulse waveform for each lead was isolated through contour-based signal extraction (using the find_contours function from the scikit-image library). The values obtained from the extracted contours were normalised using Min–Max scaling and assembled into a comprehensive feature dataframe containing 3,060 features per sample.

Feature Selection and Dimensionality Reduction

Since the 3,060 raw contour features lead to a high dimensionality and a risk of overfitting, four different feature-selection methods were compared in an empirical study using a total-variance-retained criterion: Recursive Feature Elimination (94.71%), Lasso (88.48%), SelectKBest in conjunction with the f_classif score function (91.08%), and Principal Component Analysis (99.70%). Principal Component Analysis was chosen as the dimensionality-reduction technique for this study and the feature space was reduced to 400 principal components (Jolliffe & Cadima, 2016).

Enhanced AlexNet Architecture

The model suggested makes alterations to the standard AlexNet architecture (Krizhevsky et al., 2012) so as to handle grayscale ECG data and produce a four-class diagnostic result. Table 1 shows the architecture on a

layer-by-layer basis. Compared with the standard setup, the improved model (i) changes the input layer from 3-channel RGB to a single grayscale channel which is then replicated over three channels in order to maintain compatibility with the pretrainened weights; (ii) keeps the convolutional kernels that have been tuned (11×11, 5×5, and 3×3) since they are optimised for the morphological features of ECG signals; (iii) reduces the final fully connected layer (FC8) from 1,000 ImageNet classes to four output neurons representing the NH, MI, PMI, and HB classes; and (iv) uses dropout regularisation (at a rate of 0.5) in both FC6 and FC7 to reduce overfitting, together with a Softmax activation at the output layer.

Table 1. Enhanced AlexNet layer-by-layer architecture

Layer	Type	Filters	Kernel	Activation / Detail	Output Size
Conv1	Convolutional	96	11×11, s=4	ReLU + LRN + MaxPool 3×3	55×55×96
Conv2	Convolutional	256	5×5, p=2	ReLU + LRN + MaxPool 3×3	27×27×256
Conv3	Convolutional	384	3×3, p=1	ReLU	13×13×384
Conv4	Convolutional	384	3×3, p=1	ReLU	13×13×384
Conv5	Convolutional	256	3×3, p=1	ReLU + MaxPool 3×3	6×6×256
FC6	Fully Connected	4096	—	ReLU + Dropout (0.5)	4096
FC7	Fully Connected	4096	—	ReLU + Dropout (0.5)	4096
FC8 (Modified)	Fully Connected	4	—	Softmax (4-class output)	4

Authors' Computation

Model Training and Implementation

The PCA feature vectors, which consisted of 400 components, were reshaped into 20 by 20 single-channel matrices, replicated three times to produce pseudo-RGB tensors, and then rescaled to 224 by 224 pixels by means of bilinear interpolation in order to meet AlexNet's input requirements. The dataset was divided into 80 percent for training and 20 percent for testing. The model was implemented in Python 3.9 using PyTorch and trained for 50 epochs with a batch size of 32, employing the Adam optimizer (with a learning rate of 0.001; Kingma & Ba, 2015) and a categorical cross-entropy loss function. The training was carried out on a CUDA-enabled GPU through Google Colaboratory.

Evaluation Metrics

The model's performance was evaluated in terms of classification accuracy, precision, recall (that is, sensitivity), specificity, the F1-score, the confusion matrix, and the area under the Receiver Operating Characteristic Curve (AUC-ROC), with the latter being calculated using one-vs-rest multi-class averaging (Sokolova & Lapalme, 2009; Fawcett, 2006). Together, these measures reflect both the model's ability to discriminate and its clinical reliability for all four diagnostic categories.

RESULTS AND DISCUSSION

Preprocessing Outcomes

It was confirmed by means of image inspection that all 929 of the source images were valid RGB JPEG files. The denoising process greatly reduced the average pixel noise without losing any of the waveform details that are relevant for diagnosis. The augmentation technique increased the effective size of the training set from 743 images (which made up the 80% training subset) to about 2,972 instances. PCA was able to retain 99.7% of the total feature variance by using 400 of the original 3,060 contour-derived features and thus performed better than Recursive Feature Elimination, Lasso, and SelectKBest as a dimensionality-reduction method for this dataset.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

Classification Performance

The improved version of AlexNet attained an overall classification accuracy of 99.44% and a macro-averaged AUC-ROC of 96.06%. Table 2 gives a summary of the performance for each class. For the Normal Heartbeat, Myocardial Infarction, and Previous Myocardial Infarction categories, both precision and recall were above 98.8%, while the Abnormal Heartbeat class had only slightly lower values but still demonstrated strong performance (F1 = 99.15%).

Table 2. Enhanced AlexNet classification performance by diagnostic class

Class	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)	AUC-ROC (%)
Normal Heartbeat (NH)	98.9	99.6	99.2	99.7	97.2
Myocardial Infarction (MI)	99.5	99.1	99.3	99.8	96.8
Previous MI (PMI)	99.4	98.8	99.1	99.6	95.8
Abnormal Heartbeat (HB)	99	99.3	99.15	99.5	96.1
Overall / Macro-Average	**99.2**	**99.2**	**99.2**	**99.7**	**96.06**

Authors' Computation

The results were confirmed by the confusion matrix: 56 NH cases, 47 MI cases, 34 PMI cases, and 47 HB cases were correctly classified, there being only two misclassifications in the whole test set (one MI case being misclassified as HB and one HB case being misclassified as NH). Since there were no classification errors for either the NH or PMI classes, this shows that the method has a strong ability to discriminate between all four categories.

Comparative Benchmarking

Table 3 exhibits the enhanced AlexNet being compared with three benchmark methods that were also evaluated under the same preprocessing and partitioning conditions. The benchmarks include a lightweight CNN (Hasan et al., 2025), a CNN enhanced using Correlation-Aware SMOTE (Trigka & Dritsas, 2025), and a neural network based on Particle Swarm Optimisation. The enhanced AlexNet achieved an accuracy that was at least 6.5 percentage points higher than all of the other methods.

Table 3. Comparative performance of the enhanced AlexNet against benchmark classification models

Model	Approach	Accuracy (%)	AUC (%)	Key Limitation
Lightweight CNN	Minimal CNN architecture	89.3–90.86	87.5	Insufficient depth for 4-class ECG
CNN + Correlation-Aware SMOTE	SMOTE class balancing	91.0–92.98	90	Balancing without architectural depth
PSO-NN	Metaheuristic optimization	92.7	91.3	PSO suited to tabular, not image data
Enhanced AlexNet (Proposed)	**Fine-tuned AlexNet + ECG pipeline**	**99.44**	**96.06**	**Best-in-class; CDSS-integrated**

Authors' Compilation

DISCUSSION OF FINDINGS

These results clearly show that an architecture-level improvement to a relatively "legacy" CNN can outperform both purely data-oriented interventions (such as SMOTE-based class balancing) and metaheuristic optimisation methods (PSO-NN) for classifying CVD images. Although the lightweight CNN and SMOTE-

CNN benchmarks achieved only slightly higher AUC-ROC values in certain evaluation runs, this suggests these models can achieve good class separability even with lower absolute accuracy.

Yet the improved AlexNet version showed better and more balanced precision, recall, and F1-score across all four diagnostic categories, which is likely more important in a clinical setting, where minimising both false negatives (missed heart attacks) and false positives (unwarranted interventions) is essential. The relatively low recall for the Previous Myocardial Infarction (minority) class across all models shows that class imbalance remains a significant problem in ECG-based CVD classification and should therefore be a priority in future data augmentation efforts.

The significance of these results is that a well-tuned, only slightly altered version of an existing architecture like AlexNet can achieve a good compromise between diagnostic accuracy and computational efficiency, which means it can be used in resource-limited clinical situations, such as at point-of-care devices and in telemedicine systems, where access to high-performance computing facilities may be restricted.

CONCLUSION

This study designed and assessed an improved AlexNet-based convolutional neural network to classify ECG images into four classes, aiming to support the automation of heart attack prediction. By means of an extensive preprocessing procedure—which included denoising, augmentation, segmentation of the thirteen-lead images, extraction of features based on contours, and dimensionality reduction using PCA—together with specific modifications to the input, kernel, and output layers of the AlexNet architecture, the new model obtained a classification accuracy of 99.44% and an AUC-ROC of 96.06%, outperforming a lightweight CNN, a SMOTE-augmented CNN, and a PSO-optimized neural network that had been tested on the same task.

The results show that it is possible to achieve near-state-of-the-art performance when fine-tuning transfer learning for ECG images, provided careful preprocessing is carried out and a multi-class classification approach is used—a level of performance appropriate for inclusion in a clinical decision support system (CDSS). Future studies should validate the model on independent ECG datasets from hospitals, incorporate explainable AI methods such as Grad-CAM or SHAP to improve clinical interpretability, and evaluate real-time deployment on wearable ECG sensors and within electronic health record systems.

REFERENCES

1. Alenoghena, R. O., Ajaegbu, N. G., Imohi, G. D., Omachonu, F., Oguntade, F. A., & Adewale, S. A. (2026). Stock Market Development, Financial Deepening, and Industrial Growth in Nigeria. *International Journal of Research and Innovation in Social Science*, *10*(3), 4581-4596.
2. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. International Conference on Learning Representations (ICLR).
3. Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010
4. Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation, 101(23), e215–e220.
5. Hasan, M., Yasmin, F., Hassan, M. M., Yu, X., Yeasmin, S., Joshi, H., & Islam, S. M. S. (2025). Effectiveness of machine learning models in diagnosis of heart disease: A comparative study. Scientific Reports, 15, 24568.
6. Hassan, M., Ahmed, S., Ali, R., & Khan, M. (2025). Enhanced AlexNet architectures for medical image classification: A comprehensive review. Journal of Medical Artificial Intelligence, 8(2), 101–120.

7. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. https://doi.org/10.1109/CVPR.2016.90
8. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4700–4708. https://doi.org/10.1109/CVPR.2017.243
9. Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202
10. Khan, A. H., & Hussain, M. (2021). ECG images dataset of cardiac patients (Version 2) [Data set]. Mendeley Data. https://doi.org/10.17632/gwbz3fsgp8.2
11. Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. International Conference on Learning Representations (ICLR).
12. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097–1105.
13. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539
14. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60–88.
15. Oluwadamilare, J. O., Alenoghena, R. O., Ajaegbu, N. G., Oguntade, F. A., Osunlana,, K. I., & Ajisafe, H. B. (2026). Project Charter Development and Its Influence on Software Project Success. *International Journal of Humanities Social Science and Management (IJHSSM), 6*(2), 403-409.
16. Omoniyi, O. A. J., Jacob, O. O., Chigozirim, A., Alenoghena, R. O., Abayomi, J. O., & Aliu, F. O. (2026). Comparative Analysis of Custom and Pre-Trained Convolutional Neural Networks (CNNs) for Object Recognition on the Cifar-10 Dataset. *International Journal of Research and Scientific Innovation*, *13*(6), 3499-3507.
17. Perez Alday, E. A., Gu, A., Shah, A. J., Liu, C., Sharma, A., Seyedi, S., Bahrami Rad, A., Reyna, M. A., & Clifford, G. D. (2020). Classification of 12-lead ECGs: The PhysioNet/Computing in Cardiology Challenge 2020. Physiological Measurement, 41(12), 124003. https://doi.org/10.1088/1361-6579/abc960
18. Reddy, P., & Murthy, K. (2025). Particle swarm optimization enhanced neural network for cardiovascular disease prediction. Expert Systems with Applications, 269, 126095.
19. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations (ICLR).
20. Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Information Processing & Management, 45(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002
21. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1–9. https://doi.org/10.1109/CVPR.2015.7298594
22. Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning (ICML), 6105–6114.
23. Trigka, M., & Dritsas, E. (2025). Improving cardiovascular disease prediction with deep learning and correlation-aware SMOTE. IEEE Access, 13, 44590–44606. https://doi.org/10.1109/ACCESS.2025.3549417
24. World Health Organization. (2021). Cardiovascular diseases (CVDs) [Fact sheet]. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)