Engineering Journal issn: 0125-8281

Article

Cardiovascular Classification Using Efficient Net on Electrocardiogram Images

Jothiaruna N.[1,a]**, Bandla Pavan Babu**[2,b,*]**, Nagoor Basha Shaik**[3,c,**]**, P. Arun Mozhi Devan**[4,d]**, and Kishore Bingi**[4,e]

1 School of Computer Science Engineering and Artificial Intelligence, VIT Bhopal University, Madhya Pradesh, India
2 School of Computing Science and Engineering, VIT Bhopal University, Madhya Pradesh, India
3 Department of Mining and Petroleum Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, 10330, Thailand
4 Department of Electrical and Electronics Engineering, Universiti Teknologi PETRONAS, Seri Iskandar, Malaysia
E-mail: [a]jothiaruna@vitbhopal.ac.in, [b,*]bandlapavanbabu@vitbhopal.ac.in (Corresponding author), [c,**]Nagoor.s@chula.ac.th (Corresponding author), [d]arundevaeie@gmail.com, [e]bingi.kishore@ieee.org

Abstract. Cardiovascular disease ranks among the top causes of mortality, frequently caused by sudden obstructions within blood vessels. Timely identification and intervention are essential for minimizing the impact of the disease. This research employs image augmentation techniques to correct class imbalance in an ECG image dataset divided into five categories: Normal, Abnormal Heartbeat, Myocardial Infarction, Previous History of Myocardial Infarction, and COVID-19. The balanced dataset includes 6,322 images. To improve classification accuracy for cardiovascular diseases, three pre-trained models visual Geometric Group, Residual, Dense, and Efficient Network with Version 2, were trained on the balanced ECG dataset. Critical hyper parameters were fine-tuned, yielding optimal performance with a learning rate set at 0.00001, a dropout rate of 0.3, and utilizing the Adam optimizer. EfficientNet-V2 outperformed the other models, reaching a level of accuracies of 96.22%, precision 96.34%, recall 96.31%, 95.89%, 94.75%, and an F1-Score of 96.33%, thus exceeding the performance of Densenet 161, Densenet 201, ResNet50 and VGG16.

Keywords: Cardio diseases, EfficientNet, Adam optimizer, dropout, ECG image, myocardial infraction.

ENGINEERING JOURNAL Volume 28 Issue 12
Received 9 July 2024
Accepted 6 December 2024
Published 31 December 2024
Online at https://engj.org/
DOI:10.4186/ej.2024.28.12.67

DOI:10.4186/ej.2024.28.12.67

1. Introduction

Cardiovascular disease ranks the primary cause of global mortality [1]. In middle- and low-income countries, heart disease accounts for approximately 28% of fatalities, while in high-income countries, nearly 50% of deaths are linked to lifestyle changes [2]. Heart disease typically originates from problems in the heart's blood vessels, leading to serious conditions such as heart failure, heart attacks, and blockages. Diagnosis often involves blood tests, angiograms, and electrocardiograms (ECGs). This article emphasizes the utilization of ECG images for predicting cardiovascular diseases [3].

Figure 1 depicts the conventional approach where a cardiologist diagnoses heart abnormalities. Electrodes are attached to the patient's body to capture ECG images, with the twelve-lead ECG being the most common type [4]. These leads record heart signals from different parts of the body, converting them into graphical wave patterns on a chart. The cardiologist examines the chart for irregularities and decides on the appropriate treatment to aid the patient's recovery [5].

Traditional diagnostic methods have several drawbacks, such as being time-consuming and prone to errors if the examiner lacks expertise in ECG interpretation. Automatic diagnostic methods, which offer accurate disease detection and faster prediction times, are necessary [6]. Deep learning techniques are used to automatically diagnose heart diseases from ECG images. Recent studies highlight the significant role of Medical imaging [7], signal processing, machine learning (ML) [8], and deep learning (DL) in healthcare. Signal data, often one-dimensional, is a key element in most medical studies and can be effectively analyzed using machine and deep learning methods [9]. Selecting the right algorithm is crucial, as its performance can differ across various datasets. Both one-dimensional and two-dimensional data are utilized in deep learning, significantly enhancing the diagnosis of cardiovascular diseases [10].

To improve ECG signal classification, An LSTM model that is deep and bidirectional was suggested, demonstrating strong performance in categorizing various ECG patterns, like Ventricular Premature Contraction (VPC), Normal Sinus Rhythm (NSR), Right Bundle Branch Block (RBBB), Left Bundle Branch Block (LBBB), Paced Beat (PB), and so on [17]. Additionally, one-dimensional CNN model was trained on the computerised identification of MI using 12-lead ECG signals [18]. While these deep learning approaches to time-series data offer significant advantages over traditional machine learning methods, they also present challenges. One major challenge is obtaining standardized ECG data from different healthcare providers, as variations in equipment and data formats can hinder seamless analysis [19].

In recent years, deep learning has seen a surge of interest, particularly in its medical applications [11]. Notable research has used time-series data to categorise heart conditions, introducing innovative methods like automatic signal detection with wearable devices and signal denoising using autoencoders [12]. A significant development was a Neural Network [13] model designed to categorize Myocardial Infarction and typical ECG readings. Machine learning techniques use classifiers like Random Forest, Hidden Markov Model (HMM) [15], Support Vector Machine (SVM) [14], each possessing unique benefits and drawbacks [16].

Collecting data in real time as a time series poses significant challenges while obtaining 2D data from printouts or screenshots is considerably simpler [20]. A substantial body of research exists on utilising ECG pictures to categorise cardio illnesses. Nonetheless, when handling 2D image datasets from medical facilities, a number of preprocessing procedures are necessary [21]. Among them are normalising the images, eliminating noise, and deleting duplicate pictures. When images that have been processed are used for training, the duration of processing can be shortened yet deep learning techniques can automate many of these stages and improve system efficiency [22]. Properly resizing image data is crucial for expediting model training.

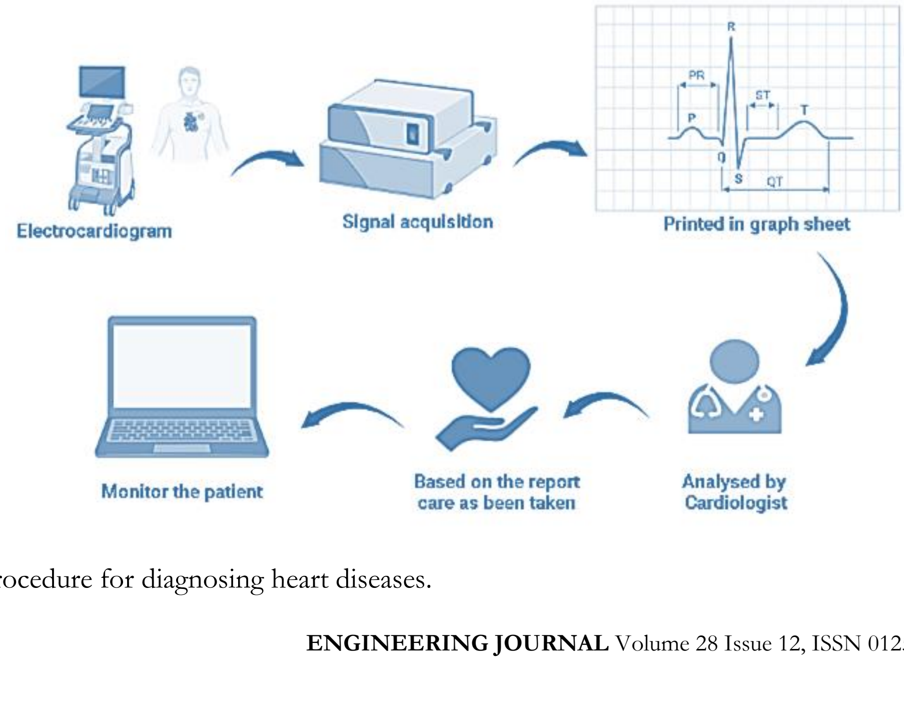

Fig. 1. Standard procedure for diagnosing heart diseases.

Segmentation, which involves dividing an image into specific sections, is essential in this process [23]. For ECG images, the 12 different leads or waveforms need to be individually segmented and processed to ensure accurate information extraction [24]. In previous work, all 12 leads were segmented separately. The objective is to utilize deep learning techniques for the automatic detection of diseases from ECG images [25], using detection and classification methods to predict diseases [26]. Surprisingly, there has been limited exploration into using DL to classify CVD based on ECG images [27].

Researchers have adopted various approaches for precise disease categorization [28]. For instance, a 2D CNN model was trained using utilising Shallow Neural Network (SNN) in conjunction with DenseNet as a feature extractor and compared against one-dimensional CNN models, Dense Net-201, and Res Network-34. To efficiently classify diseases, a hybrid deep learning designs was implemented [29], involving the selection of specific features from various pre-trained models. Notable models included Convolutional Neural Network (R-CNN) regions, Single Shot Detector (SSD) [33], You Look Only Once (YOLO) [34], fast R-CNN [31], faster R-CNN [32], and Feature Pyramid Network (FPN) [35], along with Pre-trained models such as VGG-16, Residual Network [36], Mobile Network, and others [37].

This study's novelty is outlined as follows:

- The study focuses on predicting utilising learned algorithms to classify data on unusual heart rate, COVID, myocardial infarction (MI), history of MI, and healthy.
- The study's dataset is openly accessible [30], characterized by imbalanced class distribution. To address this, a variety of image augmentation techniques have been applied to balance the classes.

Diverse hyperparameters, including the optimizer, learning rate, weight decay, and batch sizewas consistently utilised to improve the algorithm's functionality.

The paper's discourse primarily centers around the application of applying deep learning methods to information from ECG images. It thoroughly addresses various aspects, including an overview of related work, detailed descriptions of data repositories and dataset acquisition methods, preprocessing procedures, segmentation, model training and testing, performance metrics, and a dedicated section focused on the challenges and solutions related to the dataset.

2. Dataset Acquisition

The ECG database comprises a total of 3,034 images, distributed across various categories as follows: 779 images in the category of abnormal heartbeats, 250 images representing Covid cases, 375 images indicating Myocardial Infarction (MI), 387 images associated with a history of MI, and 1,243 images in the normal class. You can find specific dataset details in Table 1.

The ECG images, originated from the Chaudhry Pervaiz Elahi Institute of Cardiology in Multan, Pakistan and were stored in.jpg form [30]. A subset of the dataset—70% over training, 10% over testing, and 20% over validation—has been separated out for efficient training and assessment. A few examples of pictures from all the categories are illustrated in Fig. 2, revealing that the unaided eye, it can be difficult to discern among those categories.

Table 1. Details of Cardiovascular classes.

Disease type	Number of ECG image
Abnormal Heart Beat (AHB)	779
Covid	251
History of Myocardial Infarction	375
Myocardial Infarction	387
Healthy	1,244

3. Proposed Methodology

In this research, it's essential to provide a theoretical foundation for comprehending the deep learning technique. Subsequently, there's a concise overview of Convolutional Neural Networks and pre-trained model architectures. Figure 3 outlines the workflow for this study. A subsequent section delves into the principles of CNN along with Transfer Learning (TL) methods.

3.1. Convolutional Neural Network

In the realm of modern technology, extensive research is underway to address various medical challenges using Convolutional Neural Networks (CNN) [38]. These applications encompass cardio disease classification [22], breast cancer classification [39], and brain disease detection, among others.

CNN, when presented with an input, predominantly expects image data, through which it can discern attributes like color, texture, and edges, as well as spatial and temporal features via the utilization of filters [40]. A significant benefit of CNNs is their capability to enhance computational efficiency through the use of weight-sharing techniques, which require fewer parameters. Each layer in the network utilizes particular filters and the kernel dimensions.

A crucial part of the procedure is pooling layers, which significantly reduce the dimensions of feature maps and enhancing the independent recognition of features obtained through previous convolution layers. Two common pooling techniques include segmenting the inputs into defined squares by means of the max-pooling layer, subsequently providing the maximum value within each square. Conversely, the average pooling layer computes the average values within the squares, aiding in computational reduction and dimensionality reduction.

DOI:10.4186/ej.2024.28.12.67

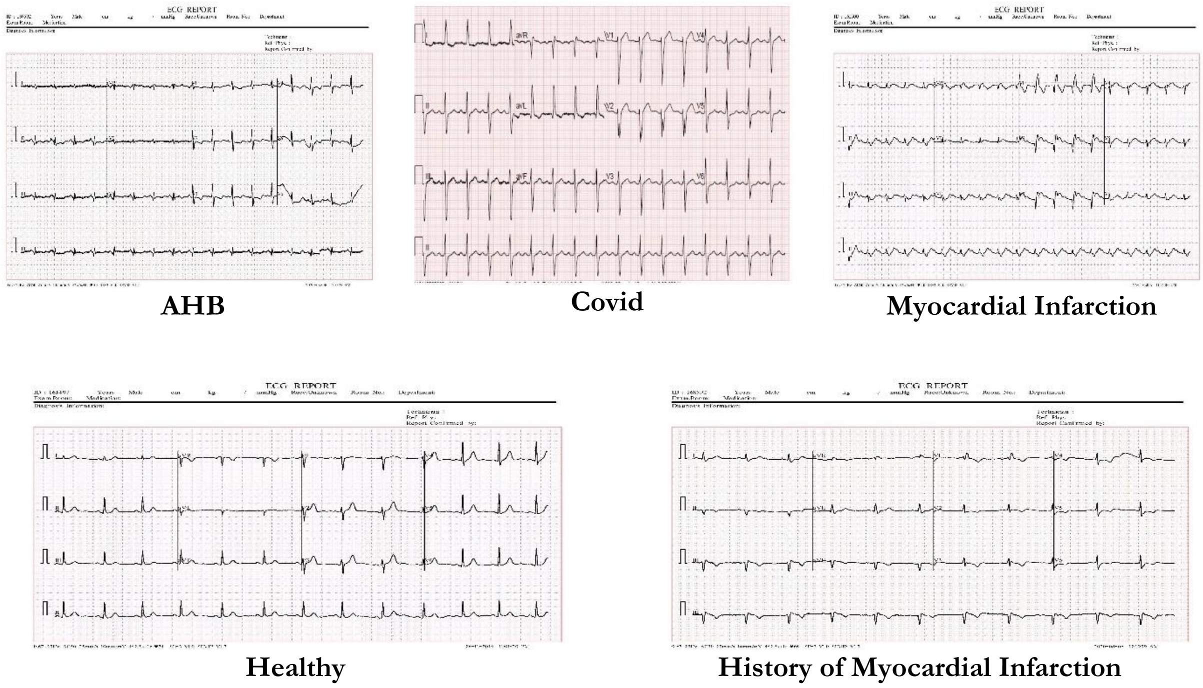

Fig. 2. ECG samples on various cardio types.

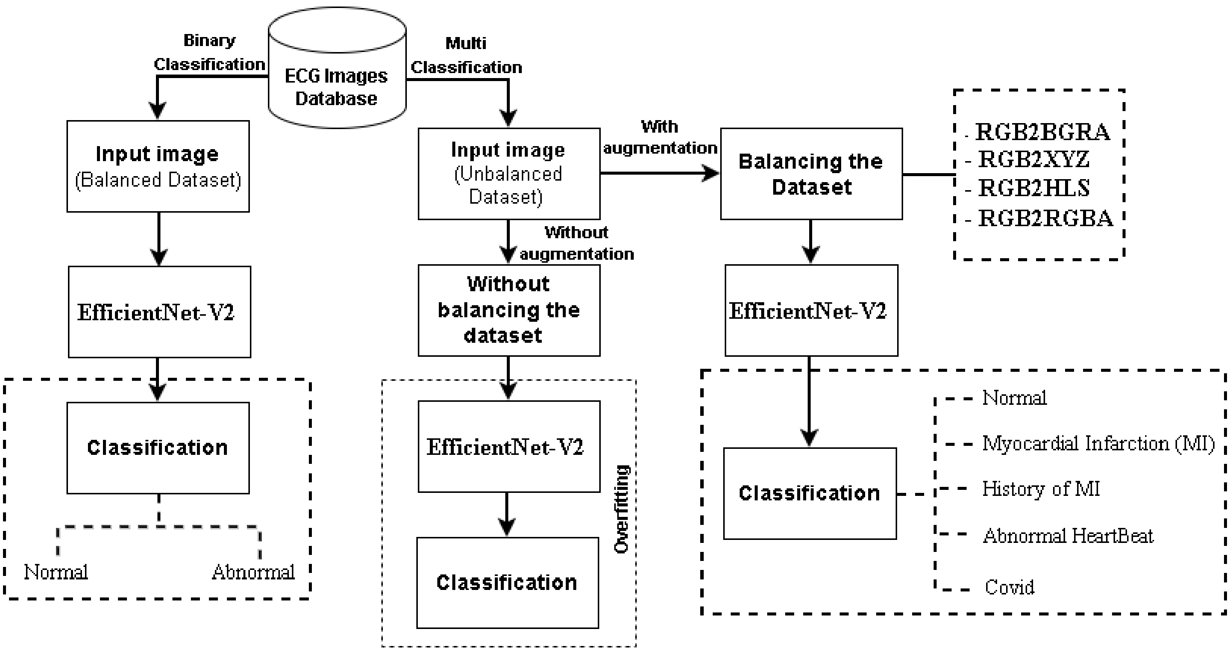

Fig. 3. Workflow of the model.

It's worth noting that the efficacy of CNN approaches in identifying heart diseases hinges on dataset size. Larger datasets tend to yield superior results, provided that appropriate precautions and considerations are taken into account [41].

3.2. Transfer Learning (TL)

To solve computer vision issues like detection and classification, researchers advised using pre-trained models because they produce superior results compared to models that must be created from scratch and trained over a longer time. Reusing a model which has undergone prior training to solve a problem is known as TL. by storing the information learned to solve a different problem. In computer vision research and many other fields, this methodology produces good results [42]. The TL process and how it is expressed mathematically are precisely defined as,

1) Recognising the origin field: $X = f_Y(Z;\ W_Y|\varphi_Y)$, Z denotes the space feature as well as the marginal distribution (f_Y) of Y. J_Y is a predictive function in task Y which is $J_Y(W_Y|Z_S, \varphi_Y)$ within domain X.
2) Transfer of Information, the researcher employed 3 methods: fixed feature extraction using a convolutional network, tuning designs and

parameterization of the convolutional neural networks network that had been trained represented as $J_Y^1(.)$.

3) The data is acquainted as it is transferred to the targeted field. The aim of TL is to improve the modelling outcome for the original field X and its learning task J_Y,, as well as for target field X' and its acquiring task J_P.

3.3. CNN's Design

There are multiple discrete layers that make up the CNN building design, forming a stack that includes the layers that are fully connected, flattened, pooled, and convolution. Layers for convolution, serves purpose in extracting features by altering the image. To achieve this, Convolution is carried out using particular filter the variables after a 2D image has been transformed into a matrix. The image being studied is reduced to a smaller size in the present research by using a filter size of 3 x 3 to create map features. Even so, this decreased dimension can lead to some loss of information, it preserves essential parts of the image as feature maps. The activation function used is Rectified Linear (ReLU), which outputs one for positive inputs and zero otherwise. In CNN models, ReLU outperforms other activation functions and effectively mitigates issue of the vanishing gradient.

Utilising subsampling processes strategies, a layer of pooling lowers the dimensionality of the image. This process maintains The details about object in the final picture, reducing the amount of math involved. In this model, sub-sampling techniques such as maximum pooling and average pooling. When maximum pooling is used, the mapping of features is given a 2 x 2 window dimensions, and the highest element inside the window's coverage area is chosen. Similarly, the layer for average pooling functions comparably to layer known as maximum pooling, computing the average of all values within the window that covers the feature map. In contrast to layers of maximum pooling and average pooling preserves a greater amount of information from the image.

The pooling layer's output is subsequently passed through flattening layers, which convert characteristics from a single layer of pooling into a 1D vectors. These flattened features are subsequently passed into a fully linked layer is essential to the categorical process. of cardiovascular diseases. Within the fully connected layer, weight calculations and classification processes are executed. The prediction is assessed through the computation of a cost function, which provides insights into how the network arrived at its predictions.

3.4. Over-Fitting Issues

Dealing with overfitting can be tackled through several strategies, such as image augmentation when dealing with imbalanced datasets and the incorporation of dropout values during model training. Additionally, an early stopping approach is implemented if there is an increase in the loss. These techniques are further detailed in the subsequent subsection.

3.4.1. Dropout

While training the model overfitting issues arises because of more number of neurons in model. To solve overfitting issues dropout technique is used. Regularization techniques encompass various approaches, with one of them being dropout. In addition to dropout, methods like L1 (Lasso) and L2 (Ridge) are frequently employed to counteract excessive fitting. Both L1 and L2 add a deduction from the loss calculation, which assists in mitigating overfitting. On the other hand, dropout dynamically affects the network during each iteration. Specifically, dropout impacts the network by deactivating some neurons during each iteration to enhance model fitting. However, it was observed that the model experienced overfitting once dropout was introduced. To counteract this, weight decay was employed to gradually reduce the model's weights toward zero, though not precisely to zero, in order to circumvent potential issues. After the weight decay and dropout were included, the model was fitted and produced positive results.

3.4.2. Image Augmentations

When our model is unbalanced, it will lead to an overfitting issue. To overcome these issues image augmentation techniques are typically used for balancing the dataset. Figure 4 depicts the conversion of RGB2BGRA, RGB2XYZ, RGB2HLS, and RGB2RGBA augmentation techniques. The number of images used for dataset balancing is shown in Table 2 for both before and after augmentation. The number of images that must be added to each class using the augmentation technique causes an imbalance in each class before augmentation. The dataset is processed using random resized crop, color jitter, and grayscale parameters from the Python library for image transformation. The number of images presenting the Covid class was 250 in order to counterbalance four augment techniques are used i.e., 250*4=1000 and 1000 + 250 (original image) = 1250). Likewise in Abnormal Heartbeat, 779 images, to balance with other types only one augmentation technique is applied So, 779*1=779 and 779+ 779 (original image) = 1556. Our samples increased from 6322 to 18966 following this augmentation technique. Instead of training random images, it will help train all the images. Both the classification outcome and training accuracy will improve as a result of this.

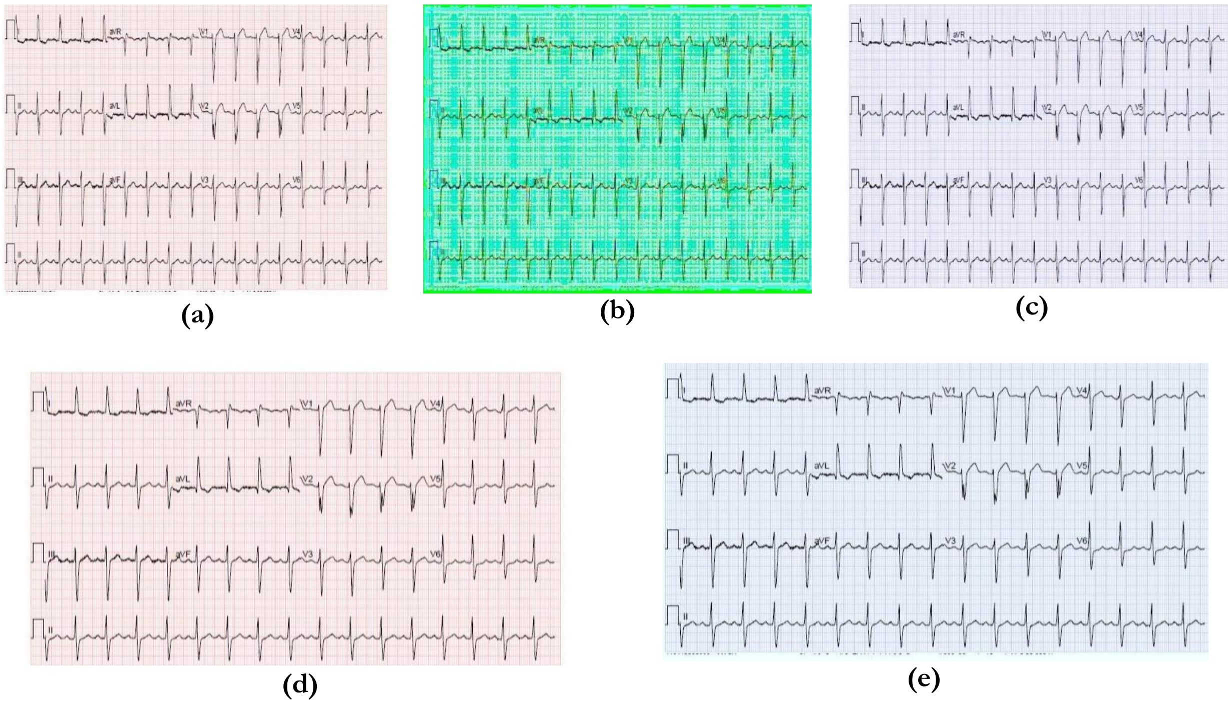

Fig. 4. A demonstration of the different image augmentation methods (a) original image (b) RGB2HLS (c) RGB2BGRA (d) RGB2RGBA (e) RGB2XYZ.

Table 2. Number of images after balancing the classes.

Disease type	Before augmentation	After augmentation
Abnormal Heart Beat (AHB)	779	1,558
Covid	251	1,255
History of Myocardial Infarction	375	1,125
Myocardial Infarction	387	1,161
Healthy	1,244	1,244

3.4.3. Model Checkpoint

To improve the performance of the model, necessary for process more epochs, which requires a significant amount of time. Periodically, the model's record is stored in checkpoint. When a model fails, this will enable us to recover the most recent saved version of the model. To reach our objective, we used checkpoints.

3.4.4. Early Stopping

Our goals are to improve accuracy while reducing training and validation loss. During training, each epoch is monitored to check if the loss is decreasing; if it starts increasing, training continues until the best accuracy is achieved and can be stopped when there are no further changes in accuracy. To tackle the issue of augmentation techniques, symptoms of overfitting can be mitigated by using decrease the model's ability, increase the training data, and increase dropout. Early stopping will halt the training process and keep the model from deteriorating.

3.4.5. Optimization

Over the model's training phase, the learning rate is lowered using the optimization approach. One of the commonly employed techniques is Adaptive Moment Estimation (Adam), which offers the advantage of reduced memory consumption. Adam essentially combines the momentum method alongside the gradient descent technique. We chose to apply the Adam optimizer [43] during the investigation..

3.5. Model Architectures

The three different models underwent testing and training using ECG images, with a comprehensive description of these models provided in the following subsection.

3.5.1. ResNet-50

ResNets outperform AlexNet, and generally, better performance is achieved in networks with more layers. ResNet, despite its relatively fewer layers, can effectively scale up the layer count while addressing the vanishing gradient problem. In our case, we employed ResNet50 [44]. To mitigate the vanishing gradient problem and reduce both errors and computational time, ResNet incorporates skip connections within its architecture. These skip connections essentially bypass some layers positioned between the activation layer and subsequent layers. This feature enables the network to effectively adapt to residual mapping, represented by H(b). Non-linear layer's further adjusts an additional mapping, denoted as F(b)=H(b)-b, which is subsequently incorporated as H(b)=F(b) + a.

3.5.2. EfficientNet

Efficient Network is neural network model, which is higher is speed. In EfficientNet MBconv is used and to increase the speed of the model MBconv is replaced by Fused MBconv. But Fused MBconv will increase the number of paramters, if parameter increases training training time will also increases. To solve this issues EfficinetNet-V2 is used, it uses combination of MBconv and Fused MBconv (inverted residual block) [45]. Comparison to earlier models, uses a nonuniform scaling approach, may expand the number of deeper layers, and has a faster training speed and better parameter efficiency [46].

3.5.3. VGG 16

Integrating the VGG16 model network causes the dimension of the network to rise., thereby significantly enhancing the model's performance. The VGG16 model comprises a combination of convolutional kernels, pooling kernels, and Rectified Linear Units (ReLU). This architecture encompasses five convolutional layers, three completely interconnected layers, a Softmax layer. Notably, the layers are distinguished using max pooling, and ReLU serves as the activation function for all hidden layers [47]. One prominent advantage of VGG16 is its streamlined neural network structure. The resultant feature maps are subsequently fed into the five distinct forms of CVD and their layers that are fully linked, depends on the Softmax activation function [39].

3.5.4. Dense Network 161 and 201

DenseNet-161 has 161 layers. It is designed to balance model complexity and performance, making it suitable for various tasks where computational resources might be a constraint. Due to its depth and dense connectivity, DenseNet-161 can capture intricate patterns in the data, making it effective for tasks such as image classification.

DenseNet-201 has 201 layers. With additional layers compared to DenseNet-161, it can learn more complex representations but at the cost of increased computational requirements. The extra depth allows DenseNet-201 to achieve higher accuracy on complex tasks, making it suitable for more challenging datasets.

DenseNet architectures, including DenseNet-161 and DenseNet-201, have been widely adopted for different tasks related to visual processing, like segmentation, identifying objects, and classification of images. Their ability to promote feature reuse and efficient gradient flow makes them particularly effective for these applications. DenseNet-161 and DenseNet-201 are powerful deep-learning models that leverage dense connectivity to achieve high performance while maintaining parameter efficiency. The choice between them typically depends on the specific requirements of the task and the computational resources available.

4. Performance Evaluation

In this section, we present the experimental results related to the classification of CVD from ECG images. For these experiments, a Python environment was employed, utilizing a system with an i9-9900K, 3.60 GHz CPU, a Quadro RTX 5000 with 32 GB of RAM, and a 1 TB HDD, running on the GNOME 3.28.2 interface within the Ubuntu 20.04 Linux operating system. The model was implemented using the PyTorch framework, specifically, version 1.1.0. Each of the three different pre-trained models underwent training for a total of 100 epochs.

Table 3. Various hyperparameters on different Models.

Models	Optimizations	Weight Decay	Batch Size	Learning rate
Dense Network 161	Adaptive Momentum	0.00001	16	0.00001
Dense Network 201	Adaptive Momentum	0.00001	16	0.00001
VGG-16	Adaptive Momentum	0.00001	16	0.00001
Residual Network	Adaptive Momentum	0.00001	16	0.00001
Efficient Network V2	Adaptive Momentum	0.00001	16	0.00001

Table 4. Comparison of different model.

Model	Learning rate	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	MCC (%)
EfficientNetV2	0.00001	96.22	96.34	96.31	96.33	95.24
	0.0001	95.97	96.01	95.99	96.01	95.11
Dense Network 161	0.00001	95.89	95.97	96.02	96.10	95.02
	0.0001	95.09	96.78	96.33	96.18	95.12
Dense Network 201	0.00001	94.75	96.41	96.98	94.61	93.54
	0.0001	94.23	96.04	96.55	94.12	92.04
ResNet-50	0.00001	93.33	94.23	93.33	93.33	89.77
	0.0001	92.30	93.27	92.28	92.43	85.06
VGG-16	0.00001	92.00	93.12	92.03	92.03	84.98
	0.0001	91.32	92.46	91.06	91.05	82.89

12 Lead ECG images are used to categorize images of cardiovascular diseases in which there are three categories for ECG images: training (70%), validation (20%), and testing (10%). Employing popular models that have been trained and transfer learning methodologies, recent advancements in the processing of images and computer vision. Table 3 various hyperparameters on different pre-trained models had been fitted and trained to get a better result. The early stopping technique is employed to reduce overfitting during training, and learning rate of 0.00001 is applied while using the optimizer developed by Adam. Figure 5 illustrates the accuracy as well as loss in multiple classification respectively. These figures demonstrate that the model was effectively trained and completed successfully.

Performance indicators such as F1-score, recall, accuracy, and precision are measured into consideration when assessing our model. Equations (1) to (4) define the equations for each of those metrics. The primary metric accuracies, will describe the proportion of accurate predictions and overall prediction. Precision measures the accuracy of positive predictions, indicating both normal and abnormal changes in the ECG. Recall identifies the positive predictions among all positive instances. The F1-score combines precision and recall to assess the model's accuracy.

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (2)$$

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \times 100\% \quad (3)$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \times 100\% \quad (4)$$

In the context of our evaluation, the abbreviations False Positive, False Negative, True Positive, and True Negative are represented by the letters TP, TN, FP, and FN, respectively. Figure 6 illustrates the confusion matrices, providing a clear and intuitive understanding of the classification performance of EfficientNetV2 on the test data. Among our three proposed models, EfficientNetV2 achieved the highest accuracy. Figure 6 serves as evidence of EfficientNet's effectiveness in accurately classifying CVD from ECG images, with fewer instances of misclassification.

In our experiment, we assessed five classes of ECG images: Normal, MI (Myocardial Infarction), History of MI, Abnormal Heartbeat, and Covid. We tested ResNet, VGG16, Densenet 161, Densenet 201, and EfficientNetV2 independently using different learning rates, specifically 0.0001 and 0.00001. The corresponding accuracies for these different learning rates across the four models are summarized in Table 4.

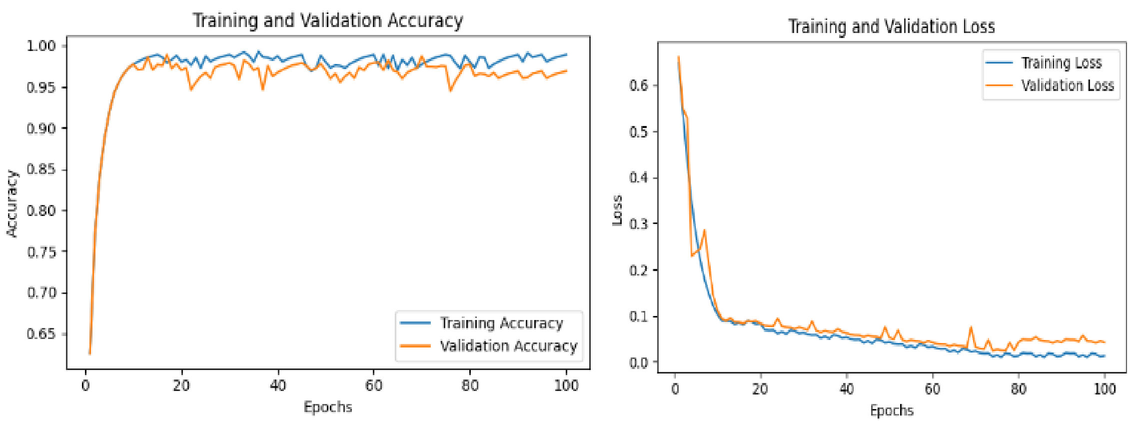

Fig. 5. Accuracy and loss graph.

Table 5. Accuracy and Loss Analysis.

Models	Test Accuracy	Validation Accuracy	Test Loss	Validation Loss
EfficientNetV2	96.22%	96.18%	0.152	0.181
Dense Network 161	95.89%	95.78%	0.165	0.178
Dense Network 201	94.75%	94.64%	0.184	0.201
ResNet-50	93.33%	93.25%	0.198	0.179
VGG-16	92.00%	91.55%	0.241	0.241

5. Discussion

The analysis was conducted on a publicly accessible dataset, focusing on multi-class classification. Table 5 provides insights into the training and validation accuracies. In multi-class classification, the presence of imbalanced classes can lead to issues of overfitting. To address this, image augmentation techniques were applied, significantly increasing the total number of available images. This augmentation approach resulted in an impressive of 95.74 for the test and 95.72 for the validation accuracy. It's worth noting that as the number of images is expanded, accuracy tends to increase, and concurrently, the loss automatically decreases as accuracy improves.

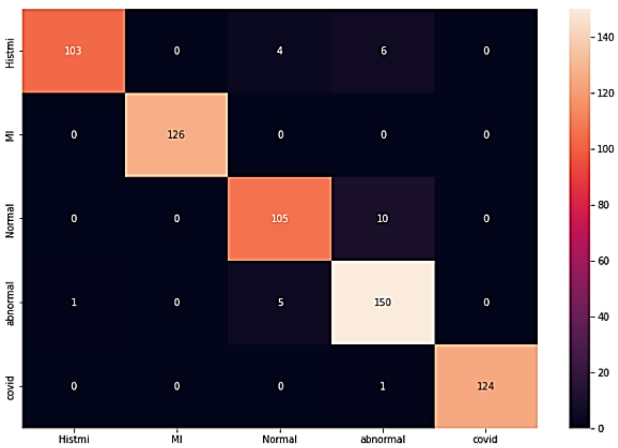

Fig. 6. Confusion metrics on EfficientNet-V2.

MI survivors are identified by their Histmi class with a classification accuracy of 79.64%; MI identifies patients related to abnormal blood flow in an artery with a classification accuracy of 96.82%; Normal identifies patients with healthy hearts with a classification accuracy of 86.95%; Abnormal identifies patients who have recently recovered from COVID and MI with breathing problems with a classification accuracy of 99.35%; and COVID-related The remaining 23 images in Histmi are incorrectly categorized as twenty images are abnormal, and three are normal., with the Abnormal class being a recovery of MI and Covid in 20 of the misclassified images. Four images in MI are incorrectly labelled as abnormal. One image in abnormal is incorrectly classified as Normal, and fifteen images in Normal are incorrectly classified as Abnormal. Images in Covid classification six are incorrectly labelled as irregular. Owing to how similar and connected all diseases are, there are frequently classification errors. ECG wave has some small variations that are challenging to distinguish in an ECG image, so adding more images is necessary to enhance the model's training for more accurate predictions of CVD.

Previous studies have investigated the classification of CVD using ECG images, and our proposed method has been compared with these efforts. As shown in Table 5, the EfficientNetV2 pre-trained model beats VGG 16 and Residual Network 50. The models have more class types than the existing models and are trained and tested on a larger number of images. The author will try to improve the hyper-parameters and parameters in the future in order to decrease misclassification of cardiovascular diseases and increase accuracy. In order to increase accuracy and aid the doctor in automatically diagnosing heart diseases, additional cardiovascular disease classes and samples have also been added.

6. Conclusion

Cardiovascular diseases are automatically classified using deep learning techniques. In this research, the classification task involves distinguishing between MI's, a previous history of MI, healthy patterns, Covid-19-related abnormalities, and irregular heartbeats using a dataset of ECG images. Numerous models with prior training, such as Dense network 161, ResNet fifty, and the VGG-16, Densenet 201, and EfficientNet-V2, leverage transfer learning methods, fine-tuning essential hyperparameters such as Dropout and Learning Rate in combination with optimization algorithms. During the training process for multi-class classification, the model exhibited signs of overfitting. To mitigate this issue and ensure robust model performance, techniques such as weight decay and image augmentation were employed to regulate weight decay and enhance training data. On ResNet-50, VGG-16, Densenet 161, Densenet 201, and EfficientNet-V2 classification accuracy for cardiovascular diseases are 94.71%, 93.33%, 95.74%, 95.89%, 94.75%, and 96.22%. Assessing how EfficientNet-V2 models perform compared to other pre-trained models in classifying five different categories, it performs better. Efficiency is capable of obtained by training the model with additional classes and images. Consequently, compared to the traditional methods, the models based on deep learning that are offered perform better in categorization and authority medication. Future enhancements could include expanding the dataset to improve generalizability, integrating models into real-time clinical workflows, and exploring novel architectures like transformers for better feature extraction. Additionally, incorporating multi-modality data and focusing on explainable AI techniques would enhance model reliability and applicability in clinical settings.

Acknowledgment

No acknowledgment to specify.

References

[1] World Health Organization, *The World Health Report 2002: Reducing Risks, Promoting Healthy Life*. World Health Organization, 2002.

[2] A. D. Lopez et al., *Global Burden of Disease and Risk Factors*. Washington, DC: The International Bank for Reconstruction and Development / The World Bank, 2006.

[3] C. D. Mathers et al., *Deaths and Disease Burden by Cause: Global Burden of Disease Estimates for 2001 by World Bank Country Groups*. 2005.

[4] World Health Organization Cardiovascular Disease Programme, "Integrated management of cardiovascular risk," Report of a WHO meeting, World Health Organization, Geneva, July 9-12, 2002.

[5] N. D. Wong, "Epidemiological studies of CHD and the evolution of preventive cardiology," *Nature Reviews Cardiology*, vol. 11, no. 5, pp. 276–289, 2014.

[6] L. Wang, H. Zhang, K. C. L. Wong, H. Liu, and P. Shi, "Physiological-model-constrained noninvasive reconstruction of volumetric myocardial transmembrane potentials," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 2, pp. 296–315, 2009.

[7] R. R. Bond, D. D. Finlay, C. D. Nugent, G. Moore, and D. Guldenring, "Methods for presenting and visualising electrocardiographic data: From temporal signals to spatial imaging," *Journal of Electrocardiology*, vol. 46, no. 3, pp. 182–196, 2013.

[8] Q. Zhang, X. Zeng, W. Hu, and D. Zhou, "A machine learning-empowered system for long-term motion-tolerant wearable monitoring of blood pressure and heart rate with ear-ECG/PPG," *IEEE Access*, vol. 5, pp. 10547–10561, 2017.

[9] S. Veeragandham and H. Santhi, "Role of IoT, image processing and machine learning techniques in weed detection: A review," *International Journal of Internet Technology and Secured Transactions*, vol. 12, no. 3, pp. 185–204, 2022.

[10] N. Sobahi, A. Sengur, R. S. Tan, and U. R. Acharya, "Attention-based 3D CNN with residual connections for efficient ECG-based COVID-19 detection," *Computers in Biology and Medicine*, vol. 143, p. 105335, 2022.

[11] K. Suzuki, "Overview of deep learning in medical imaging," *Radiological Physics and Technology*, vol. 10, no. 3, pp. 257–273, 2017.

[12] Appathurai et al., "A study on ECG signal characterization and practical implementation of some ECG characterization techniques," *Measurement*, vol. 147, p. 106384, 2020.

[13] S. Veeragandham and H. Santhi, "A Review on the role of machine learning in agriculture," *Scalable Computing: Practice and Experience*, vol. 21, no. 4, pp. 583–589, 2020.

[14] S. Osowski, L. T. Hoai, and T. Markiewicz, "Support vector machine-based expert system for reliable heartbeat recognition," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 4, pp. 582–589, 2004.

[15] D. A. Coast, R. M. Stern, G. G. Cano, and S. A. Briller, "An approach to cardiac arrhythmia analysis using hidden Markov models," *IEEE Transactions on Biomedical Engineering*, vol. 37, no. 9, pp. 826–836, 1990.

[16] P. Chang, J. Lin, J. Hsieh, and J. Weng, "Myocardial infarction classification with multi-lead ECG using hidden Markov models and Gaussian mixture models," *Applied Soft Computing*, vol. 12, no. 10, pp. 3165–3175, 2012.

[17] R. Ouni, H. Alhichri, and A. Kharshid, "Light weight residual convolutional neural network for atrial fibrillation detection in single-lead ECG recordings," *Engineering Journal*, vol. 28, no. 2, pp. 67–80, 2024.

[18] P. Kanani and M. Padole, "ECG heartbeat arrhythmia classification using time-series augmented signals and deep learning approach," *Procedia Computer Science*, vol. 171, pp. 524–531, 2020.

[19] H. Khan, M. Hussain, and M. K. Malik, "Cardiac disorder classification by electrocardiogram sensing using deep neural network," *Complexity*, vol. 2021, no. 1, p. 5512243. 2021.

[20] M. A. Ozdemir, G. D. Ozdemir, and O. Guren, "Classification of COVID-19 electrocardiograms by using hexaxial feature mapping and deep learning," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, pp. 1–20, 2021.

[21] T. Rahman et al., "COV-ECGNET: COVID-19 detection using ECG trace images with deep convolutional neural network," *Health Information Science and Systems*, vol. 10, no. 1, pp. 1–16, 2022.

[22] T. Anwar and S. Zakir, "Effect of image augmentation on ECG image classification using deep learning," in *Proc. International Conference in Artificial Intelligence (ICAI)*, 2021, pp. 182–186.

[23] B. Allegranzi, E. Tartari, and D. Pittet, "Seconds save lives—clean your hands": the 5 May 2021 World Health Organization SAVE LIVES: Clean

Your Hands campaign," *Journal of Hospital Infection*, vol. 111, pp. 1-3, 2021.

[24] B. Breve, "Visual ECG analysis in real-world scenarios," in *Proceedings of the 27th International Conference on Distribution in Multimedia Systems*, 2021, pp. 46–54.

[25] A. Hassan, M. Elhoseny, and M. Kayed, "A novel and accurate deep learning-based Covid-19 diagnostic model for heart patients only," *Research Square*, pp. 1–11, 2022.

[26] E. Irmak, "COVID-19 disease diagnosis from paper-based ECG trace image data using a novel convolutional neural network model," *Physical and Engineering Sciences in Medicine*, vol. 45, no. 1, pp. 167–179, 2022.

[27] J. C. Gomes, M. A. de Santana, A. I. Masood, C. L. de Lima, and W. P. Dos Santos, "COVID-19's influence on cardiac function: a machine learning perspective on ECG analysis," *Medical & Biological Engineering & Computing*, vol. 61, no. 5, pp. 1057–1081, 2023.

[28] A. Roshan and M. Abdelrahman, "Improving aggregate abrasion resistance prediction via micro-deval test using ensemble machine learning techniques," *Engineering Journal*, vol. 28, no. 3, pp. 15–24, 2024.

[29] P. Hao, X. Gao, Z. Li, J. Zhang, F. Wu, and C. Bai, "Multi-branch fusion network for myocardial infarction screening from 12-lead ECG images," *Computational Methods and Programs in Biomedicine*, vol. 184, 2020.

[30] A. H. Khan, M. Hussain, and M. K. Malik, "ECG images dataset of cardiac and COVID-19 patients," Data Brief, vol. 34, p. 106762, 2021.

[31] R. Girshick, "Fast R-CNN," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1440–1448.

[32] R. Gavrilescu, C. Zet, C. Foşalău, M. Skoczylas, and D. Cotovanu, "Faster R-CNN: An approach to real-time object detection," in *2018 International Conference and Exposition on Electrical and Power Engineering (EPE)*, 2018 pp. 165–168.

[33] W. Liu et al., "SSD: Single shot multibox detector," in *European Conference on Computer Vision*, 2016, vol. 9905, pp. 21–37.

[34] W. Chen, H. Huang, S. Peng, C. Zhou, and C. Zhang, "YOLO-face: A real-time face detector," *Visual Computing*, vol. 37, no. 4, pp. 805–813, 2021.

[35] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2117–2125.

[36] H. Alaeddine and M. Jihene, "Deep residual network in network," *Computational Intelligence and Neuroscience*, vol. 2021, 2021.

[37] B. P. Babu, S. Khandagale, V. Shinde, S. Gargote, and K. Bingi, "Enhancing infrastructure safety: A UAV-based approach for crack detection," *Engineering Journal*, vol. 27, no. 12, pp. 11–22, 2023.

[38] K. Kallianos et al., "How far have we come? Artificial intelligence for chest radiograph interpretation," *Clinical Radiology*, vol. 74, no. 5, pp. 338–345, 2019.

[39] A. Manickam, J. Jiang, Y. Zhou, A. Sagar, R. Soundrapandiyan, and R. D. J. Samuel, "Automated pneumonia detection on chest X-ray images: A deep learning approach with different optimizers and transfer learning architectures," Measurement, vol. 184, p. 109953, 2021.

[40] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in 2017 International Conference on Engineering and Technology (ICET), 2017, pp. 1–6.

[41] C. Bailer, T. Habtegebrial, and D. Stricker, "Fast feature extraction with CNNs with pooling layers," *Computer Vision and Pattern Recognition*, 2018, *arXiv:1805.03096*.

[42] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, "Transfusion: Understanding transfer learning for medical imaging," in *NIPS'19: Proceedings of the 33rd International Conference on Neural Information Processing Systems*, 2019, vol. 3, pp. 3347–3357.

[43] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[44] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.

[45] M. Tan and Q. Le, "EfficientNetV2: Smaller models and faster training," in *International Conference on Machine Learning (ICML)*, PMLR, 2021, pp. 10096–10106.

[46] M. L. Huang and Y. C. Liao, "Stacking ensemble and ECA-EfficientNetV2 convolutional neural networks on classification of multiple chest diseases including COVID-19," Academic Radiology, vol. 30, no. 9, pp. 1915-1935, 2023.

[47] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

DOI:10.4186/ej.2024.28.12.67

Dr. Jothiaruna N. currently working in VIT University Bhopal, and received a Doctor of Philosophy (Ph.d) in School of Computer Science Engineering and Information Systems in Vellore Institute of Technology, Vellore in 2024. She has published around 10+ papers in SCI, WOS, Scopus, Book chapter, and presented various conferences. Her research field is Deep learning in various applications like medical and Agriculture, Computer Vision, Image Processing, and Machine Learning..

Dr. Bandla Pavan Babu received Doctor of Philosophy (Ph.D) in School of Electrical Engineering from Vellore Institute of Technology, Vellore in 2022. He has an academic experience of 10 years and worked in different premier institutions in India. He has published around 20 SCI, WOS, Scopus, Book chapters, Textbooks and International Conferences. Achieved a Patent Grant from Government of India in 2022 titled "E-Bicycle". He organized many conferences, workshops, FDP's and was also invited as speaker. He was elevated as a Senior Member in IEEE in 2022 and acted as IEEE Branch counselor, Faculty Advisor etc in Professional societies. With his passion towards teaching many video lectures done and published in YouTube for the benefits of students, working professionals and other learners. His current area of Research is Artificial Intelligence and Machine Learning Techniques, Electric Vehicle & charging technologies, Smart Grid and Renewable Energy Sources, Power Electronics and Electrical Machines and Power Quality

Dr. Nagoor Basha Shaik is a Post-Doctoral Researcher in the Department of Mining and Petroleum, at the Faculty of Engineering, from Chulalongkorn University. He did his Ph.D. in the Department of Mechanical Engineering, specializing in AI and ML applications - in the oil & gas sector from Universiti Teknologi PETRONAS, Malaysia (2017-2021). He finished his M. Tech in the Department of Mechanical Engineering (Specialised in Machine Design) from Jawaharlal Nehru Technological University, Kakinada, India, in 2016. He completed his Bachelor of Technology in Mechanical Engineering from Acharya Nagarjuna University, Guntur, A.P. in 2011. He worked as a mechanical engineer in India from July 2011- Nov 2014 in India. He published more than 32 articles in good impact factor journals. His areas of interest includes materials, artificial intelligence, oil and gas pipelines, energy, sustainability, piping systems, prediction techniques, manufacturing processes, engineering drawing, and material properties.

Dr. P. Arun Mozhi Devan is working as a Research Scientist in the Department of Research, Innovation & Commercialisation at Universiti Teknologi PETRONAS, Malaysia, since October 2022. He received the B.Eng. degree (Hons.) in electronics and instrumentation engineering from Muthayammal Engineering College, Rasipuram, Tamil Nadu, India, in 2012, the M.Eng. degree (Hons.) in control and instrumentation engineering from Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India, in 2016 and the Ph.D. degree in the Department of Electrical and Electronics Engineering from Universiti Teknologi PETRONAS, Perak, Malaysia. He was an Assistant Professor at the Department of Electronics and Instrumentation Engineering, Sri Ramakrishna Engineering College, from 2016 to 2018. His current research interests include control systems, fractional-order control, process modelling and control, optimization and wireless network control.

Dr. Kishore Bingi received the B.Tech. Degree in Electrical and Electronics Engineering from Acharya Nagarjuna University, India, in 2012. He received the M.Tech. Degree in Instrumentation and Control Systems from the National Institute of Technology Calicut, India, in 2014, and the Ph.D. degree in Electrical and Electronic Engineering from Universiti Teknologi PETRONAS, Malaysia, in 2019. From 2014 to 2015, he worked as an Assistant Systems Engineer at the TATA Consultancy Services Limited, India. From 2019 to 2020, he worked as Research Scientist and Post-Doctoral Researcher at the Universiti Teknologi PETRONAS, Malaysia. From 2020 to 2022, he served as an Assistant Professor at Process Control Laboratory, School of Electrical Engineering, Vellore Institute of Technology, India. Since 2022 he has been working as a faculty member at the Department of Electrical and Electronic Engineering from Universiti Teknologi PETRONAS, Malaysia. His research area is developing fractional-order neural networks, including fractional-order systems and controllers, chaos prediction and forecasting, and advanced hybrid optimization techniques.He is an IEEE and IET Member and a registered Chartered Engineer (CEng) from Engineering Council UK.