

Squeeze-ICNN architecture for cardiovascular disease detection using ECG image: training of shape, deep, and LGXP feature extractors

Thoutireddy Shilpa* and T. Priyanka

Department of Computer
Science and Engineering, Koneru
Lakshmaiah Education Foundation,
Hyderabad 500075, Telangana,
India

*E-mail: shilpathoutireddy@gmail.
com

Received for publication
June 19, 2025.

Abstract

Nowadays, the leading causes of early mortality, especially in countries with low or middle incomes are cardiovascular diseases (CVDs), which include heart failure, stroke, hypertension, and coronary artery disease. Premature death rates may be reduced if certain illnesses are identified early. Numerous methods, including data mining and machine learning (ML) have been proposed by researchers for early identification as well as tracking of cardiac patients in the context of CVD prediction. Yet there remains significant concern about the effectiveness of these strategies in scenarios where there is a high error rate and uncertain precision. Thereby, choosing a prediction method that may produce high accuracy and less errors is essential. This article presents the Squeeze-improved convolutional neural network (ICNN) model for detecting CVD via electrocardiogram (ECG) images. During preprocessing, the ECG image noise is reduced using an improved median filter (IMF). The model extracts shape, deep, and local Gabor XOR pattern (LGXP) features. For classification features, SqueezeNet and improved CNN (IMP-CNN) are combined within a hybrid model architecture (Squeeze-ICNN), with a Comb SigHyper-Hsine activation function in the classifier. The model provided experimental results proving accuracy (95.6%), precision (97.9%), sensitivity (94.9%), specificity (92.2%), and lower false negative rate (FNR) (5.1%) than conventional models (e.g., convolutional neural network [CNN], deep neural network [DNN], LSTM, DenseNet). The study clearly proves the effectiveness and performance of the model to detect CVD from ECG images.

Keywords

improved median filter, LGXP, improved hierarchy of skeleton features, ICNN, SqueezeNet

Nomenclature

Abbreviation	Description		
AC	Accuracy	CHS	Comb Hyper-Sine
AUC	Area Under the Curve	CNN	Convolutional Neural Network
BGRU	Bidirectional Gated Recurrent Unit	CVD	Cardiovascular Disease
Bi-GRU	Bidirectional Gated Recurrent Unit	DCNN	Deep Convolutional Neural Network
CDC	Centers for Disease Control and Prevention	DCT	Discrete Cosine Transform
		DL	Deep Learning

DNN	Deep Neural Network
DSR	Dynamic Source Routing
ECG	Electrocardiogram
EMD	Empirical Mode Decomposition
EMG	Electromyography
EPS	Episodes per Second
ETT	Exercise Tolerance Test
FFT	Fast Fourier Transform
FN	False Negative
FPRs	False Positive Rates
ICNN	Improved Convolutional Neural Network
IoT	Internet of Things
LBP	Local Binary Pattern
LGXOR	Local Gabor XOR Pattern
LSTM	Long Short-Term Memory
LXP	Local XOR Pattern
MCC	Matthews Correlation Coefficient
NPV	Negative Predictive Value
MI	Myocardial Infarction
ML	Machine Learning
NN	Neural Network
PSNR	Peak Signal-to-Noise Ratio
QRS	QRS Complex
R-CNN	Region-Based Convolutional Neural Network
REL	Routing by Energy and Link Quality
ReLU	Rectified Linear Unit
SSD	Single Shoot Detection
SSIM	Structural Similarity Index Measure
ST	ST Segment
SVM-RFECV	Support Vector Machine with Recursive Feature Elimination and Cross-Validation
TL	Transfer Learning
VGG	Visual Geometry Group
WHO	World Health Organization
XOR	Exclusive OR

I. Introduction

The World Health Organization (WHO) reports that cardiovascular disease (CVD) is the most commonly occurring long-term illness. In a similar vein, the Centers for Disease Control and Prevention (CDC) [1] pointed out that CVDs caused over 690,882 deaths

in the United States in 2020. On the other hand, early detection and accurate diagnoses can greatly reduce the risk of CVD [13, 22]. Innovative medical research, aided by information technology, has yielded practical approaches to address CVD concerns in contemporary society [7]. Electrocardiography (ECG), ETT, Echo, blood tests, and angiography screening are the instruments that are recommended for cardiac diagnostics [14]. The most economical and widely used diagnostic method for checking cardiac problems is the ECG [8]. Moreover, it is regarded as a typical instrument for assessing CVD in individuals residing in isolated regions [5, 16].

ECG is a common and inexpensive diagnostic tool that is used to find cardiovascular abnormalities. Historically, these recordings have been interpreted by clinicians, but more recently the advent of artificial intelligence (AI) and deep learning (DL) techniques have allowed for the automation of ECG interpretation, strictly for arrhythmia and myocardial infarction (MI). Many ECGs are still in the format of an image, which makes converting images into forms that the AI facility can use to enable faster, consistent, and scalable diagnosis [2]. In many other cases, healthcare providers and medical researchers employ paper-based ECG images of heart diseases, because a large amount of ECG recordings are still stored or shared in paper format, due predominantly to the limitations of clinical function in low resource or remote settings or countries [12, 20]. The work of manually reviewing paper-based ECG images appears to be inefficient, inconsistent across practitioners, and influenced by human error. Digitizing these ECG images to leverage computer-aided diagnostic (CAD) capabilities is a more reliable, fast, and accurate method of determining whether cardiovascular conditions exist [4, 15]. Not only does this reduce variability in diagnosis, but it increases the speed of clinical workflows. Image focusing AI onto CAD can be important for preventing early deaths from CVD, particularly in rural or limited healthcare environments. It provides a sense of urgency to address the issue of implementing an image-based DL framework to review the paper-based ECG images and potentially harness the framework as an automated CVD classification system where raw ECG signal data may not be available. Many methods for automatically identifying CVD were studied by researchers by utilizing machine learning (ML) [9, 17, 21] as well as DL methods [10, 19]. Generally, this involved utilizing ECG in a one- or two-dimensional voltage amplitude information expressed as time-series data signals [27]. Various techniques are investigated and yield significant

outcomes for classifying ECG signals over time series [25, 26]. It is essential to use a network of deep neural network [DNNs] for automated detection of heart disorders utilizing 12-lead ECG image processing. Based on mathematical equations and models, DNNs function similarly to the human brain [3, 18]. The fundamental component of a DNN is the neuron, which develops through tasks that are repeated and gains experience through learned information during training, exactly as a human brain learns [6, 23]. Making a link among input and output is the primary objective of training and learning [24]. Following training, the system can identify the things that it has been taught to recognize [11]. With this advent, this paper introduces a new model for CVD detection using DL strategy. The main key contribution of the proposed model is summarized as follows:

- First, this study creates a new hybrid DL model called Squeeze-improved convolutional neural network (ICNN), which incorporates the light-weight efficiency of SqueezeNet and the better learning functionality of an ICNN, which has a new Comb SigHyper-Hsine, activation function.
- Second, the model combines many features shape features through improved skeleton hierarchy, deep features (ResNet, VGG16, and InceptionV3), and texture features (Local Gabor XOR Pattern [LGXP]), which overall improves performance and classification.
- Third, the proposed method added a new adaptive median filtering preprocessing step that greatly improved ECG image quality, which is evidenced with its better PSNR (37.804) and SSIM (0.915) quality scores.
- Fourth, large experimental results show that the proposed method directly outperforms previous works in terms of higher precision (97.9%), higher accuracy (95.6%), and lower false negative rate (FNR) (5.1%), while still showing robust generalization and robustness with very limited training data.

The proposed model is original because it applies the image-based ECG analysis to show an original process of multilevel feature extraction (texture, shape, and deep representations) for accurate manifestation of CVD clinical discovery of phenomena such as ST-elevations and abnormalities represented by distortions of the QRS waves. Combined with a new activation function, the discovery process is clinically relevant and stabilized.

This paper is organized as follows. Section II reviews the literature on the identification of CVD.

Section III proposes a hybrid identification method to detect CVD using images of ECG signals. Section IV presents the outcomes and reviews the Squeeze-ICNN model, while Section V presents the conclusion of the paper.

II. Literature Review

Many previous studies have examined the automated detection of CVD based on ECG images using DL and ML methods.

Ali Haider Khan et al. [1] used a DNN, based on SSD MobileNet v2 to classify four types of heart conditions (MI, arrhythmia, previous MI, and normal). Their DNN achieved a 98% accuracy on a dataset of 11,148 annotated 12-lead ECG images, but did not include generalization for larger, heterogeneous datasets.

Kaniz Fatema et al. [2] used a hybrid DL method by combining InceptionV3 and ResNet50 to classify ECG images into five different types. Their classification performance improved, however, they did not take into account a highly imbalanced dataset and did not perform complete image preprocessing, only using some simple image enhancements.

Changling Li et al. [3] introduced DeepECG, an architecture based on deep convolutional neural network (DCNN) employing Inception-V3 and transfer learning (TL). Their study provided 51,579 ECG for the model to classify seven arrhythmias with a balanced accuracy precision of 98.56%, specificity of 96.75%, and recall of 95.43%. Although the performance was acceptable, the need for an extension to incorporate raw and full 12-lead ECG images for further generalization was required.

Weibo Song [4] employed a Faster R-CNN model to analyze 2D representations of ECG images to improve visual diagnosis. They achieved a diagnostic accuracy of 98.94%, better visualizing the electrical conduction waveforms from the heart, although there were gaps in analysis at detailed lead-wise levels for multiple beats.

Zeynep Hilal Kilimci et al. [5] explored the use of transformer models (i.e., Swin-Tiny, BEiT and ViT) for ECG image classification. Their use of vision transformers was novel, but subsequently provided promising classifications with high accuracy, precision, and recall; models with robustness in terms of characteristics were also required as well as wrested by hardware limitations which precluded any useful deployment in practice.

Tariq Sadad et al. [6] proposed a lightweight CNN with attention modules in their study on ECG image

classification using an architecture based on Internet of Things (IoT) based routing. Their model was able to classify cardiac anomalies, but further improvement was needed with the robustness of segmentation and extraction of features across heterogeneous datasets.

Irin Sherly and Mathivanan [7] created hybrid DCNN which contained the arithmetic optimization algorithm (AOA). They merged clinical data with ECG images, and used Fourier and DCT transforms in the amalgamated data. Although the system gave improved classification accuracy, it lacked the ability to scale in real-world applications, because of limited generality in optimizations.

El-Habibi [8] used a standard CNN that was implemented in TensorFlow and Keras, for automated ECG image classification. The model produced poor training accuracy, but also severe overfitting, and little to no generalization with additional unseen data samples.

Muthukumar et al. [40] outlined a multimodal diagnostic framework in their 2025 publication that combines ECG signals and retinal fundus images using FFT and EMD feature extraction and a neural network (NN) classifier. Their preliminary study indicated an overall accuracy of 84% in identifying CVD.

Zhu et al. [41] present a deep residual network with dual-scale (DDR-Net), combining electrocardiogram (ECG) and phonocardiogram (PCG) signals into a single signal using feature aggregation with SVM-RFECV for the classification of heartbeats. On the training, a set of the 2016 PhysioNet/CinC Challenge competition, the model obtained 91.6% accuracy with a 0.962 AUC, exceeding the single-modality ECG or PCG models (Table 1).

In conclusion, although previous work has demonstrated promising results, most of the work suffers

Table 1: Features and challenges on existing works

Author & year	Preprocessing method	Feature extraction type	Model type	Accuracy/precision/F1	Limitations
Khan et al. [1]	Not specified	Raw ECG images	DNN (SSD MobileNet)	Accuracy: 98%	Low scalability to noisy datasets
Fatema et al. [2]	Basic denoising	Deep features (InceptionV3 + ResNet50)	Hybrid DL	Not specified	Dataset imbalance, limited filtering
Li et al. [3]	Standard augmentation	Deep features + TL	DeepECG (InceptionV3)	Precision: 98.56%	Limited raw ECG format handling
Song et al. [4]	2D signal representation	Region-based waveform features	Faster R-CNN	Accuracy: 98.94%	Incomplete beat/lead recognition
Kilimci et al. [5]	Image resizing	Vision transformer embeddings	Transformer models	High (exact not listed)	Complex, not edge-friendly
Sadad et al. [6]	Basic filtering	CNN + attention features	Lightweight CNN + IoT	High (not specified)	Weak segmentation, dataset-dependent
Sherly & Mathivanan [7]	DCT + Fourier	Clinical + image features	Hybrid CNN + AOA	Higher accuracy (est.)	Optimization lacks generalization
El-Habibi [8]	Not specified	Deep features	CNN	High (train/val accuracy)	Overfitting risk
Proposed (This Study)	IMF	LGXP + improved skeleton + ResNet/VGG/inception	Hybrid (SqueezeNet + ICNN)	Accuracy: 95.6% Precision: 97.9% FNR: 5.1%	Superior in preprocessing, hybrid design, and feature fusion

AOA, arithmetic optimization algorithm; CNN, convolutional neural network; DL, deep learning; DNN, deep neural network; FNR, false negative rate; ICNN, improved convolutional neural network; IMF, improved median filter; IoT, Internet of Things; LGXP, local Gabor XOR pattern; SSD, single shoot detection; TL, transfer learning.

from challenges associated with improper preprocessing, inadequate features, imbalanced datasets and model overfitting (i.e., memorization). Few models describe the use of DL with manually applied pre-selected features or optimizing preprocessing. This study proposes a framework to overcome these challenges. We propose a Squeeze-ICNN model that incorporates adaptive preprocessing (improved median filter [IMF]), a hybrid method of feature extraction (LGXP + improved skeleton + ResNet/VGG/Inception), and a custom activation function. Our comparative results show that our proposed model had an accuracy of 95.6%, precision of 97.9%, sensitivity of 94.9%, and an FNR of 5.1%, indicating improved performance against previously reported approaches.

III. Proposed a Squeeze-ICNN Model of CVD Detection using ECG Signal Images

ECG signals are frequently utilized for detecting and diagnosis of CVDs. The automatic detection of disease using ECG images are in the trending progress. This work proposes a new hybrid approach, the Squeeze-ICNN method, for the detection of CVD from ECG signal images. The three steps for identifying CVDs are as follows: the preprocessing step, the feature extraction step, and the disease detection step. The suggested model's process is explained in the following manner.

- In the beginning, the input ECG signal image is preprocessed using an improved median filtering method to enhance the image quality, eliminate noise, and replace each pixel value with the median value.
- The next step involves extracting features in the preprocessed image. Here, shape features, deep features, and the LGXOR patterns are extracted. Together with the extraction of ResNet, VGG 16, and Inception V3-based deep features, an improved hierarchy of skeleton features regarding the shape features has been extracted.
- In the end, the SqueezeNet and improved CNN (IMP-CNN) model combined to create the Squeeze-ICNN classification model to categorize CVDs. Additionally, the IMP-CNN model uses the comb sig Hyper-H-Sine activation function to reduce the disappearing and exploding gradient issues. The general design of the suggested strategy for detecting CVD is depicted in Figure 1.

a. Preprocessing by IMF

Preprocessing the incoming ECG images is a crucial step in improving its quality and efficiency for further analysis or classification work. An ECG signal image is typically exposed to many kinds of noise. Applying filters, like a median or bandpass filter, is helpful in lowering noise and improving the quality of the signal. Let us assume that the input ECG image is I^{ECG} . An improved median filtering technique is used as part of the preprocessing step to decrease the undesired or irregular value of pixels variance in an image. The IMF's step-by-step procedure is described below.

a.i. Improved median filtering process

A non-linear technique for image processing that eliminates noise from an ECG signal image is termed as median filtering [28]. It may replace the median value that is displayed as a mask, for each position's pixel value. Let $f(x,y)$ be the pixel value of the input ECG image signal. If the average filtering algorithm and the median filtering algorithm are capable of adaptively resizing the mask based on the noise density, then the median filtering performance needs to be enhanced. An improved median filtering algorithm is developed in the study based on this. The following steps illustrate the enhanced median filtering procedure.

Step 1: Search center element of the given image, I^{ECG} (pixel value is denoted as $f(x,y)$).

Step 2: Adaptively resizing the mask. Here, it extracts the neighborhood of the image from $f(x,y)$.

Step 3: Compute noise variance of filter (NVF). Eq (1) is used to compute the noise variation of the median filtering of an image having zero average noise in normal distribution. where, σ_i^2 represents the image variance and n denotes the improved median filtering mask size, which is taken as 3.

$$NVF = \left[\frac{\sigma_i^2}{n + \frac{\pi}{2} - 1} \right] * \frac{\pi}{2} \quad (1)$$

Step 4: Compute mask pixel 1 (MP_1) and mask pixel 2 (MP_2) to adaptively resize the mask in accordance with the noise level of the mask, and these mask pixels P_1 and P_2 are calculated in Eqs (2) and (3), respectively. Here, $\min(x)$ indicates the minimum value of the ECG image,

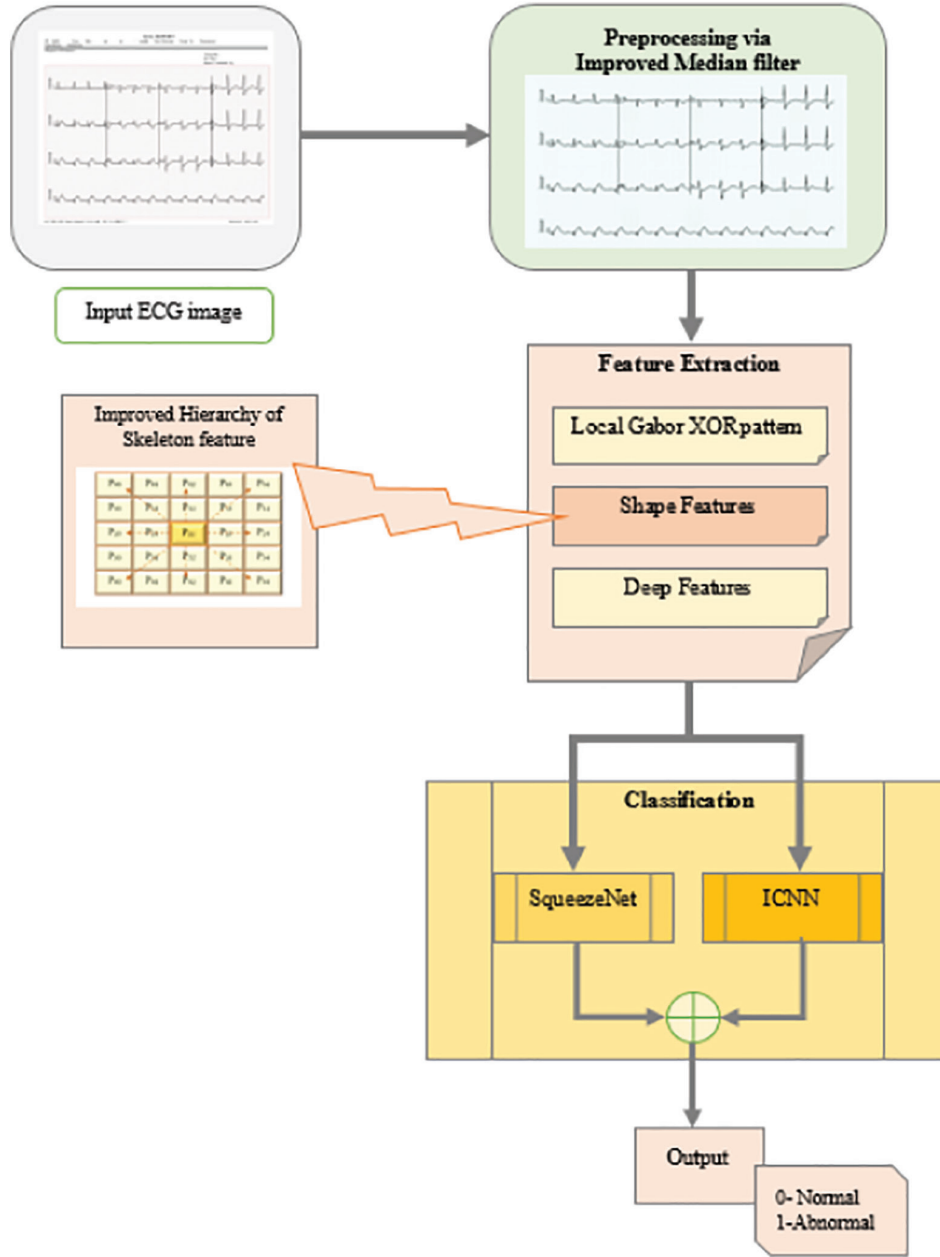


Figure 1: Basic structure of the suggested mechanism for CVD. CVD, cardiovascular disease; ICNN, improved convolutional neural network.

while $\max(x_i)$ indicates the maximum value of the ECG image:

$$MP_1 = \text{med}(x_i) - \min(x_i) \quad (2)$$

$$MP_2 = \text{med}(x_i) - \max(x_i) \quad (3)$$

Step 5: Compare each pixel with mask pixels, which is based on Algorithm 1

Algorithm 1: Improved Median Filtering process
Algorithm

if $\text{Mask Pixel}_1 > 0$ and $\text{Mask Pixel}_2 < 0$
 $f(x,y) = \text{med}(\text{Input})$
 else if
 $\text{Adaptive Pixel}_1 > 0$ and $\text{Adaptive Pixel}_2 < 0$
 $f(x,y) = f(x,y)$
 else if
 $f(x,y) < NVF$

```

f(x,y) = med(Input)
else
return the pixel value of f(x,y)

```

According to the algorithm, the calculation of adaptive pixel 1 is given in Eq. (4) as well as calculation of adaptive pixel 2 is expressed in Eq. (5), wherein, x_i indicates the input data, $f(x,y)$ indicates the input ECG image's pixel value, and $med(input)$ denotes the ECG image's median value:

$$Adaptive\ Pixel_1 = f(x,y) - \min(x_i) \quad (4)$$

$$Adaptive\ Pixel_2 = f(x,y) - \max(x_i) \quad (5)$$

Finally, the outcome of the improved median filtering is represented as I_p^{ECG} , which is the preprocessed data of the given ECG signal image.

b. Feature extraction process

The method of extracting features from preprocessed ECG images, I_p^{ECG} entails choosing specific information or features from images that can be utilized to illustrate and examine the fundamental trends. In this work, LGXP features, Shape, as well as Deep features are extracted from I_p^{ECG} .

b.i LGXP F^{LGXP}

The Local Gabor pattern and the XOR operation are combined to create the texture description known as LGXP [29]. In image processing, Gabor filters are frequently employed for texture analysis, and bit-wise operations like XOR are utilized to improve specific features.

In this proposed model, Gabor filter is initially applied in the preprocessed ECG image, I_p^{ECG} . They perform especially well at capturing data on texture at various scales and orientation. Then, several Gabor filters are combined, each with varying sensitivity to frequencies and orientations, to create a filtration bank. As a result, a series of filtered images are produced, each of which highlights a different aspect of the ECG signal.

For every filtering image, apply a local binary pattern. A texture descriptor called LBP compares the intensity of the central pixel to that of its neighbors to encode local patterns. For every pixel, this technique provides a binary code. The XOR operator is applied to the binary codes that are acquired from the LBP. The discriminating data in the encoded

pattern is improved by applying this technique pixel-by-pixel. Combining the XOR patterns for every single local region in the image is done to create a feature vector. The local texture data that the XOR operation and Gabor filters are able to extract is represented by this feature vector. Eq. (6) shows the LGXP calculation between the center (Z_c) and the neighbor (Z_i) pixel location in the Gabor phase map with orientation (η) and scale (ν), wherein $i = 1, 2, \dots, P$, LXP operator is represented as $\otimes$, Gabor phase is represented as $\phi_{\nu,\eta}(\bullet)$, and quantization operator is represented as $Q(\bullet)$:

$$LGXP_{\nu,\eta} = Q(\phi_{\nu,\eta}(Z_c) \otimes \phi_{\nu,\eta}(Z_i)) \quad (6)$$

The histograms are calculated based on the feature vectors for the preprocessed image. A condensed representation of the texture information is given by these combined features. To make sure the feature vectors are not affected by changes in intensity, they are normalized. Thus, the extracted features of the LGXP are represented as F^{LGXP} .

b.ii. Shape feature F^{SHAPE}

Extraction of shape features is an essential part of image processing as it collects the geometrical properties of components in an image. In this paper, an improved hierarchy of skeleton features are extracted from the I_p^{ECG} . Features are extracted from the skeleton representation of the preprocessed image and these features are organized or analyzed hierarchically [30]. To enhance the histogram features of the image, an improved hierarchy of skeleton features is proposed in the shape features.

b.ii.i. Improved hierarchy of skeleton features

According to the proposed improved hierarchy of skeleton features, we considered 5×5 sub blocks of images, which is shown in Figure 2. The following steps are followed in the improved hierarchy of skeleton features.

Step 1: Initially, iterate through each pixel in the image.

Step 2: If the current pixel is the target pixel, proceed to the next step to evaluate the next point.

Step 3: Count the number of target 8-neighbor pixels. Here, the modified neighbor pixel is calculated as per Eqs (7)–(14). P_{22} is represented as the center pixel, P_{ij} indicates the Pixel of the j^{th} row in

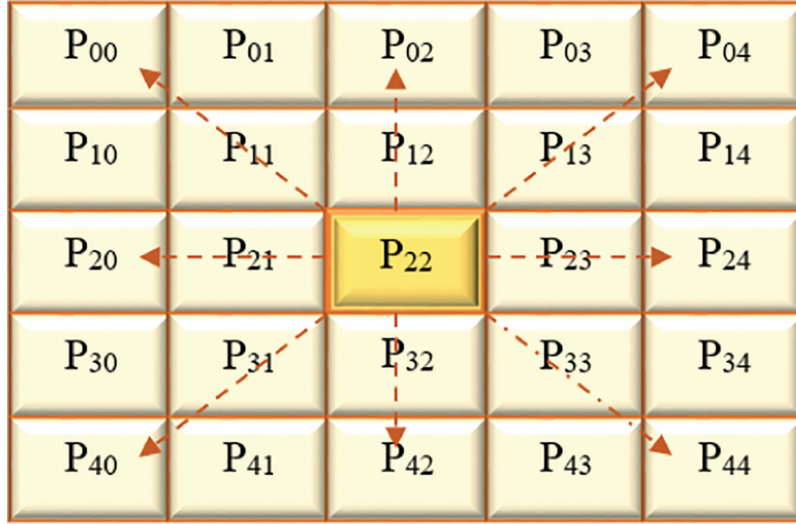


Figure 2: Neighborhood diagram of improved hierarchy of skeleton feature.

the j^{th} column, and np_l indicates the neighboring pixel ($l = 1, 2, 3, 4, 5, 6, 7, 8$):

$$np_1 = \text{median}(P_{11}, P_{00}) \quad (7)$$

$$np_2 = \text{median}(P_{12}, P_{02}) \quad (8)$$

$$np_3 = \text{median}(P_{13}, P_{04}) \quad (9)$$

$$np_4 = \text{median}(P_{23}, P_{24}) \quad (10)$$

$$np_5 = \text{median}(P_{33}, P_{44}) \quad (11)$$

$$np_6 = \text{median}(P_{32}, P_{42}) \quad (12)$$

$$np_7 = \text{median}(P_{31}, P_{40}) \quad (13)$$

$$np_8 = \text{median}(P_{20}, P_{21}) \quad (14)$$

Step 4: Independently compute the length of lines formed by the target pixels and its median of the two neighbors.

Therefore, the histogram shape features from the preprocessed image is represented as F^{SHAPE} .

b.iii. Deep features, F^{DEEP}

Deep feature extraction is the process of automatically recognizing and extracting hierarchical feature representations from images using pretrained DL models. In this proposed model, pretrained NNs such as ResNet [31], VGG 16 [32], and Inception V3 [33] model's features are extracted from the preprocessed image. To acquire deep features, feed the preprocessed image through the selected layers

of the pretrained CNN. An image representation in a high-dimensional feature space is produced by these layers.

b.iii.i. ResNet

Residual learning is introduced, in which residue functionalities are learned through the application of shortcut linkages. This facilitates the training of extreme DNNs and aids to mitigate the problem of vanishing gradients. Here, the information can move directly via the network due to identity shortcut links. The basic building blocks, known as residual blocks, have connections that skip. ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152 are some of the variations of the network; the number denotes its depth.

b.iii.ii. VGG 16

VGG16 is simple and uniform while highlighting the network's depth. It consists of 3 fully connected layers and 13 convolutional layers, making 16 weighted layers. It uses 3×3 convolution filters, which are tiny reception fields that are all throughout the architecture. It significantly down-samples using max-pooling layers.

b.iii.iii. Inception V3

The inception uses parallel filters of various sizes to catch features at various scales. There is a parallel

concatenation of 1×1 , 3×3 , and 5×5 convolutions in inception modules. Prior to fully connected layers, spatial dimensions are reduced using global average pooling. Normalization in batches is done for quicker convergence.

Thus, these features capture a high-level representation of the image, which is valuable for feature extraction. The extraction of deep features is represented as F^{DEEP} . Lastly, the final feature is represented as F , $F = [F^{LGXP} \ F^{SHAPE} \ F^{DEEP}]$, which is extracted from the preprocessed image:



When preprocessing has been carried out, ECG images are passed through ResNet, VGG16, and InceptionV3 pretrained models with the classification layers at the end of the models removed. Pretrained models extract deep features representing various levels of abstraction. VGG16 captures spatial information, ResNet extracts deep residual characteristics, and InceptionV3 provides multiscale representations of the ECG images. The extracted features beyond what comes through a Squeeze-ICNN classifier are then concatenated with handcrafted features to build a hybrid input.

c. Classification of CVD using hybrid model

Classifying CVDs using a hybrid model typically involves combining models to improve the overall performance. In this proposed work, the Squeeze-ICNN model is proposed for the classification of CVDs. The extracted LGXP features, Shape features, and Deep features are combined (F) and given as the input to Squeeze-ICNN model, which combines the SqueezeNet and IMP-CNN models. Thus, the average of the classification models gives the output of classifying the CVD as 0 and 1. Here “0” indicates the normal patient and “1” indicates the abnormal patient, who is affected with CVD. The process architecture of the classification models is shown in Figure 3, which is explained in this section.

c.i. SqueezeNet

The SqueezeNet model [34] is used for classifying CVD. Similar to numerous other DL models, it is used to train various data types, encompassing medical images or other pertinent features linked to CVD.

Utilizing 1×1 convolutions, often referred to as point-by-point convolutions, is done to lower the network’s parameter number while keeping expressive power is Squeeze Net’s primary feature by using Fire Modules. SqueezeNet provides a computationally efficient solution that can be deployed on low-resource devices because of these tiny filters. The final retrieved features F are fed to the SqueezeNet model. Convolution layer (conv1) is the first in SqueezeNet’s design, while convolution layer

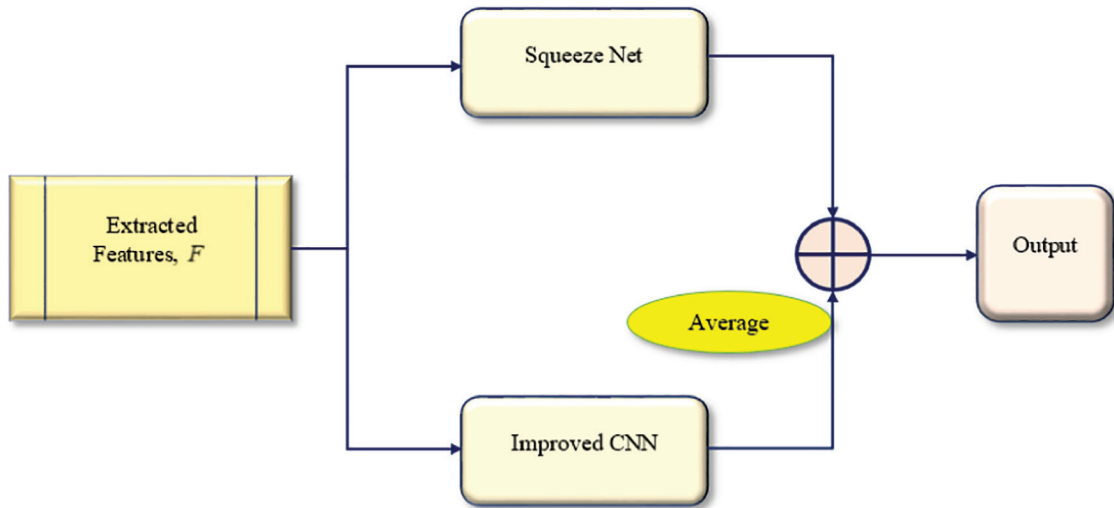


Figure 3: Architecture of classification process. CNN, convolutional neural network.

(conv10) is the last. It contains eight fire modules in it. Additionally, it carries out max-pooling with a stride of two following conv1, fire4, fire8, and conv10. Following the fire module, a 50% dropout ratio is utilized. Figure 4 shows the SqueezeNet design.

c.ii. ICNN model

CNNs are a class of DNNs that are particularly effective in tasks involving images and spatial data. In the conventional CNN model [35] is included three main layers namely, Convolutional Layers, Pooling Layers, and Fully connected Layers. Figure 5 shows the conventional CNN model. Here, the conv. block of CNNs is the convolutional layer. ReLU is the most common activation function in CNN. The pooling layer is used to down-sample the spatial dimensions of the input volume. Lastly, the fully connected layer follows the convolutional and pooling layers, which is used for making the final classification. As per this proposed work, an IMP-CNN model is proposed to improve its performance and efficiency. An ICNN model consists of five convolution layers that include the input layer with batch normalization (BN) and Leaky ReLU

activation function. The architecture of the proposed ICNN model consists of five convolution layers, a flat-ten layer with 0.5 dropout layer, three dense layers with ReLU activation, and an output layer with the proposed activation layer, which is shown in Figure 6.

c.ii.i. Convolutional layer

A convolution layer [36] performs a convolution on the given final feature, F . There are five convolutional layer blocks with strides 2 and the kernel is (1, 1) involved in the convolution part. Each convolution block contains the Convolution Layer, Leaky-ReLU Activation function, and BN. The convolution operation in every stage contains element-wise multiplication of the filter with the input feature, F . Let the kernel size be $n \times m$, which is applied on the input feature F and is denoted as f_k . There are $n \times m$ input connections involved in the CNN layer and their output is calculated as per Eq. (15), wherein (i,j) indicates the position of the feature map:

$$C(F_{u,v}) = \sum_{i=-n/2}^{n/2} \sum_{j=-m/2}^{m/2} f_k(i,j) * F_{u-i,v-j} \quad (15)$$

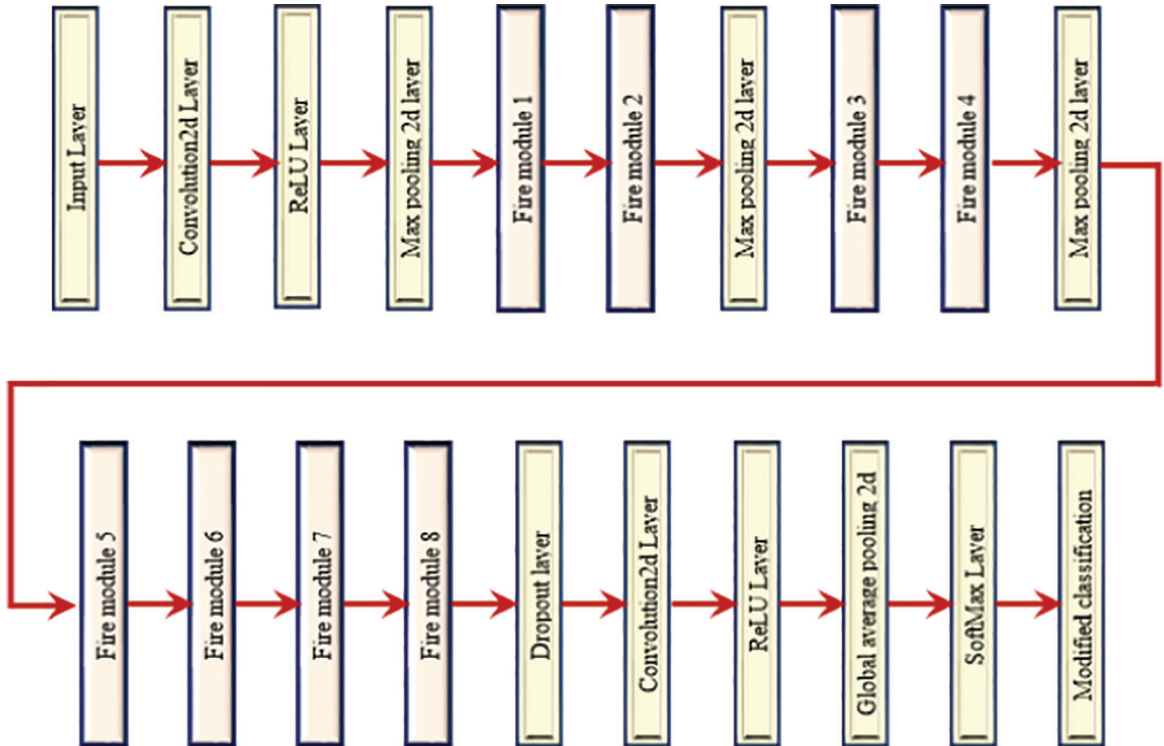


Figure 4: SqueezeNet model.

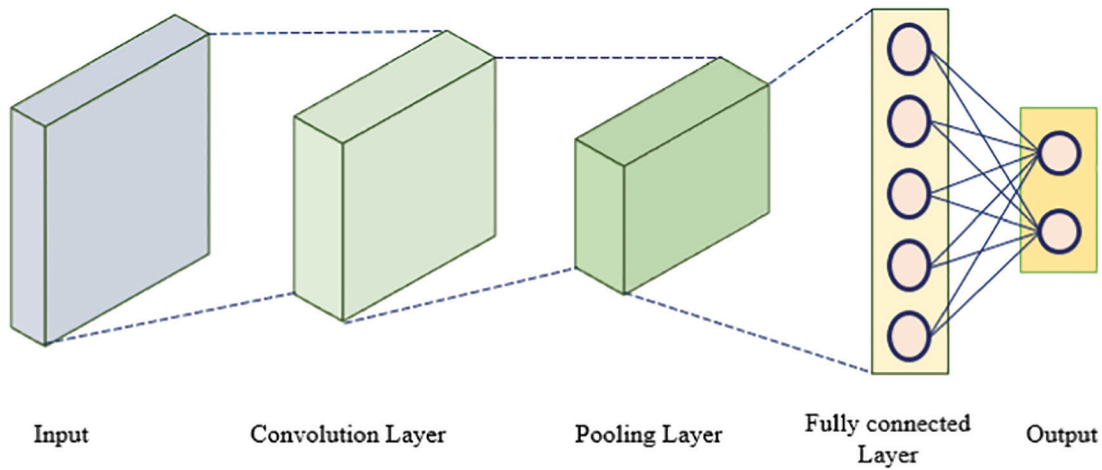


Figure 5: Conventional CNN model. CNN, convolutional neural network.

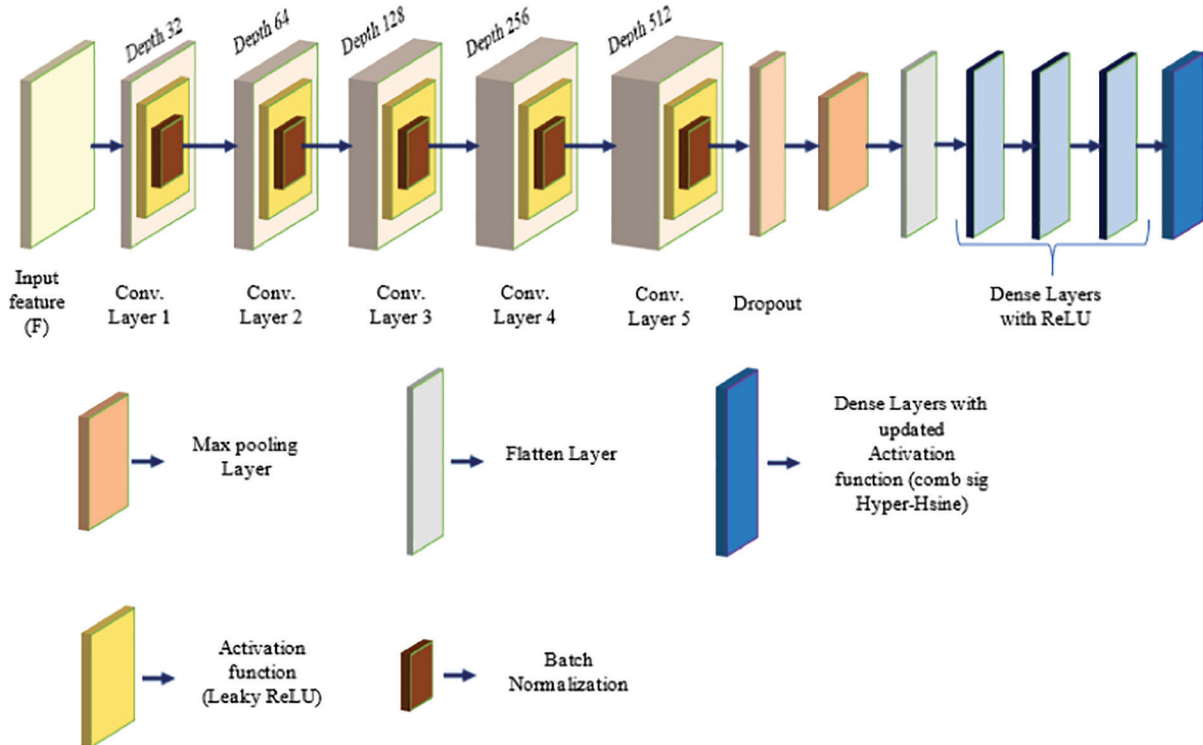


Figure 6: Architecture of ICNN model. ICNN, improved convolutional neural network.

Here, Leaky ReLU is the non-linear activation function, which is used as an extension of the traditional ReLU activation function. It is applied to the resultant features obtained by the convolution layer.

c.ii.ii. Pooling layer

The convolution layer's output that passed into the pooling layer generates the feature map. The pooling layer [36] applies the subsampling on the feature

map. A pooling layer offers a standard downsampling procedure that lowers the feature maps' in-plane dimensions to create a translation invariance to slight shifts and distortions and decrease the number of subsequently trainable parameters. The filter size, stride, and padding in pooling operations are the hyperparameters, similar to those in convolution operations, but it is significant that none of the pooling layers are equipped with learnable parameters. Averaging pooling produces the average of the values obtained from overlapped kernel areas and the input image, whereas maximal pooling produces the maximum value. Translational invariance of this max pooling is dependent on the filter size. Let us consider F to be the input and m to be the kernel size of the i^{th} value. Eq. (16) provides the computation of the maximum pooling output:

$$M(F_i) = \max \left\{ F_{i+k,j+1} \mid |k| \leq m/2, |l| \leq m/2, k, l \in \mathbb{N} \right\} \quad (16)$$

c.ii.iii. Fully connected layer

Rich information from each neuron in the preceding layer (the Flatten Layer) is gathered by a fully connected layer. In the proposed work, three Dense Layers with ReLU activation Fully connected layers are included. And the Last layer is the Dense Layer with updated Activation function, which is the output layer.

A Comb SigHyper-Hsine activation function is proposed in the Last layer of ICNN. It is the combination of SigHyper [37] and comb-H-sine [38] (CHS) activation function to enhance the classification process. Eq. (17) shows the comb-H-sine activation function derivatives, where the value of β is taken as 0.5. Eq. (18) shows the proposed *comb sigHyper-Hsine* activation function.

$$CHS = \sinh(\beta x) + \sinh^{-1}(\beta x) \quad (17)$$

comb sigHyper - H sine

$$\begin{aligned} &= \frac{1}{2} * \left(\frac{e^x (2 + e^x) - 1}{e^x + e^{-x} + e^{-2x} + 1} + \sinh(\beta x) + \sinh^{-1}(\beta x) \right) \\ &= \frac{1}{2} * \left(\frac{2e^x + e^{2x} - 1}{e^x + e^{-x} + e^{-2x} + 1} + \sinh(\beta x) + \sinh^{-1}(\beta x) \right) \end{aligned} \quad (18)$$

Finally, both the ICNN and SqueezeNet models are averaged together to get the output of the proposed model. The output of the classification is either in "0" or "1" as normal or abnormal patients.

d. Architectural comparison with AlexNet

The Squeeze-ICNN model proposed is structurally and functionally different from AlexNet. It uses the Fire modules of SqueezeNet that provides the same representation with efficiency in the number of parameters and uses a Comb SigHyper-Hsine activation function instead of ReLU to improve the stability of the gradient at a minimum. Adopting a more programmable approach, Squeeze-ICNN's model is designed to have multiple branches with fusion of deep, shape, and texture features as well as the use of dropout and BN; resulting in a lighter, more robust model suitable for classification of ECG images with limited datasets.

The proposed architecture is designed to tackle the challenges of ECG image noise and limited data. A motion, EMG, and powerline noise-based adaptive median filter was employed to suppress low-to-moderate amplitude noise while preserving possible waveform details. We used a hybrid architecture blending the lightweight Fire modules of SqueezeNet with an IMP-CNN that allowed for efficient feature extraction with robust learning; the IMP-CNN uses a customized activation function that enhanced gradient flow to the classifier block with the small training ECG datasets, all while keeping the model slim and deployable.

IV. Results and Discussion

a. Simulation procedure

The proposed CVD detection using ECG signal images were implemented in PYTHON, especially version "3.7." The computational tasks were carried out on an "AMD Ryzen 5 3450U processor with Radeon Vega Mobile Gfx, operating at 2.10 GHz," and with a total installed RAM of "16.0 GB." In addition, the assessment of the performance of CVD detection using ECG signal images was executed by utilizing the ECG Images dataset of Cardiac Patients [39].

a.i. Handling dataset class imbalance

To address the moderate class imbalance noted in the ECG dataset (i.e., 284 normal vs 172 history of MI), a few design choices were made to lessen the performance bias:

The proposed Squeeze-ICNN architecture contains multiple distinctions from traditional architectures (LGXP, shape, and deep features), which enhances the representation of the minority class in the feature space.

The combination SigHyper-Hsine activation function provides stability for learning and allows for better gradient flow, which extends generalization over under-represented classes. While model training, the class-weighted loss function was used to penalize the misclassification of minority classes more heavily.

Lastly, stratified sampling was applied to the training/testing splits so that class proportion is preserved and allows for balanced learning.

These methods together contribute to the high sensitivity of 94.9% (and 5.1% FNR), suggesting performance was learned effectively across classes, despite the class imbalance.

a.ii. Enhancing generalizability with limited ECG data

While the proposed model is built on a small sample of 928 images of ECGs [39], various means were put in place to improve generalizability and limit overfitting. TL with pretrained models (ResNet, VGG16, InceptionV3), hybrid feature fusion (deep, shape, and texture features), and adaptive preprocessing by means of an IMF that enhances signal quality (PSNR = 37.8) were utilized. Dropout, BN, and a custom activation function (Comb SigHyper-Hsine) aided the model during training to counter and stabilize learning based on low data. Low variance was noticed in learning across folds of $\sigma = 0.028$ in learning and the model's predictions scored well in performance metrics including 95.6% accuracy, 97.9% precision, and 5.1% FNR, indicating that the model is making somewhat stable, and reproducible predictions with limited sample size.

The proposed model is not a standard case of overfitting with multiple mitigating strategies in place. First, all dropout layers, BN, and early stopping were regularly used in this training and thus served as regularizers. Second, TL from pretrained models (ResNet, VGG, Inception) limited our liability of overfitting to small data sets. Third, the model was evaluated using stratified data splits and multiple evaluation metrics. On the seven folds of validation conducted here, the amount of variance from fold to fold in accuracy was low $\sigma = 0.028$. The model consistently produced high precision (97.9%) and low FNR (5.1%), providing strong evidence for the maintenance of generality outside of the training set.

a.iii. Handling ECG artifacts and ensuring model robustness

There are a number of interference sources of ECG such as powerline interference, muscular (EMG)

activity, as well as motion artifacts within the ECG signal. To combat this interference, *the model employed in this study used an IMF* to selectively remove noise by adaptively responding to the noise in the local region. It is effective against not only high-frequency noise (e.g., EMG) and low-frequency fit artifacts (e.g. motion, AC), which result in deterring noise removal from ECG signals. Subsequently the cleaned ECG images achieved PSNR of 37.804 and SSIM of 0.915. Based on these metrics, the cleaned ECG is of significant higher quality and cleaner feature input from which to classify. Due to the improvements, the model exhibits solid results and the robustness of the model allows to make predictions on other noisy ECG data-sets with some level of generality.

a.iv. Testing methodology and dataset disjointness

The performance of the model was evaluated using a disjoint training and testing setup to ensure fair and unbiased evaluation. The convenience dataset of 928 ECG images was split using a stratified 80:20 train-test split while maintaining the proportions of each class. There was no overlap between the training and testing samples, due to complete disjointness. Moreover, to evaluate consistency, the model was evaluated in repeated runs using different random-seed, with accuracy, precision, F1-score, and FNR performance metrics reported based on the average performance. This approach will demonstrate that the results of the model is due to actual generalization and not overfitting to the training data.

b. Dataset description

A set of ECG images from individuals with cardiac conditions has been compiled at the "Ch. Pervaiz Elahi Institute of Cardiology in Multan, Pakistan." The primary aim of this initiative is to support the scientific community in guiding study related to CVDs. The dataset comprises a total of 928 ECG images, which include Normal Person ECG Images (284), ECG Images of MI Patients (239), ECG Images of Patient that have abnormal heartbeat (233), and ECG Images of Patient that have History of MI (172). Furthermore, sample ECG images are depicted in Figure 7.

c. Performance analysis

The proposed model's performance is assessed on several performance metrics including accuracy, precision, recall (sensitivity), F1-score, specificity, and FNR. While

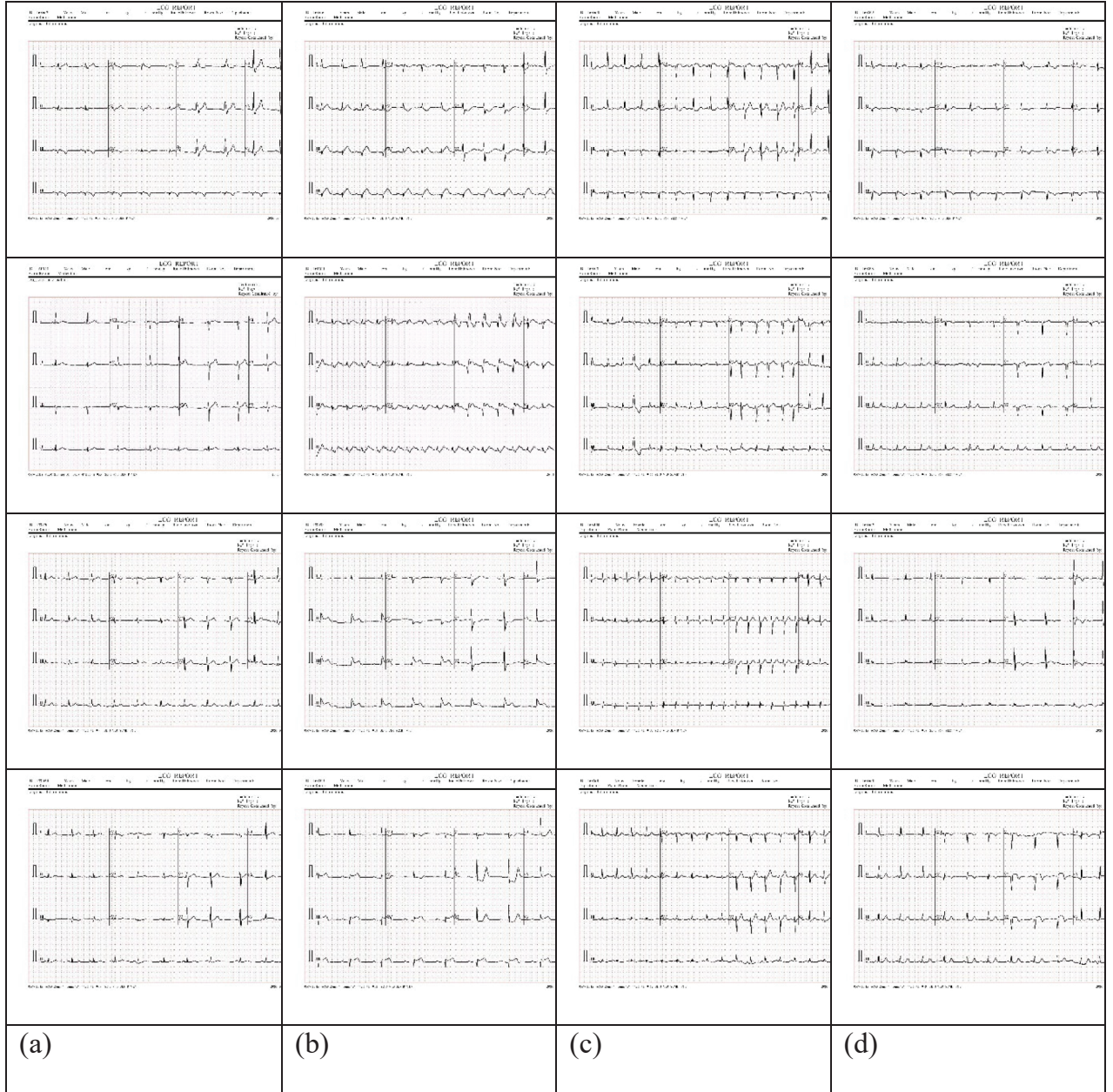


Figure 7: Sample images (A) Normal person ECG images (B) ECG images of MI patients (C) ECG images of patient that have abnormal heartbeat (D) ECG images of patient that have history of MI. MI, myocardial infarction.

all of these performance metrics are not only standard, they are highly relevant for medical diagnosis application, where not just a correctly detected case (sensitivity), but especially the ability to minimize missed cases (low FNR), is important. For example, the model attained an accuracy of 95.6%, which indicates that the overall performance is reliable. However, the model's precision of 97.9% is of particular importance because it demonstrates that it avoided false positives – a particularly

useful metric for avoiding undesirable follow-ups. The low FN of 5.1%, and sensitivity of 94.9% indicates that the model is good at detecting actual CVD cases, while minimizing a miss rate of potentially affected people, which specifically addresses a significant limitation presented by many other systems. The proposed system also demonstrated improved detection reliability to the extent of both the CNN (FNR = 19.0%) and DNN (FNR = 21.5%) failing to detect numerous likely cases of

CVD, particularly among the minority classes. The values that the proposed model calculated contribute in technical detail to the manuscript by demonstrating that the model is not only accurate, but also clinically reliable and applicable in a real-world sense to ECG-based CVD screening.

d. Preprocessing analysis

Figure 8 depicts the identification of CVDs using ECG signal images, utilizing diverse filtering methods, including Improved Median, Gaussian, Mean, Wiener, and Conventional Median Filtering.

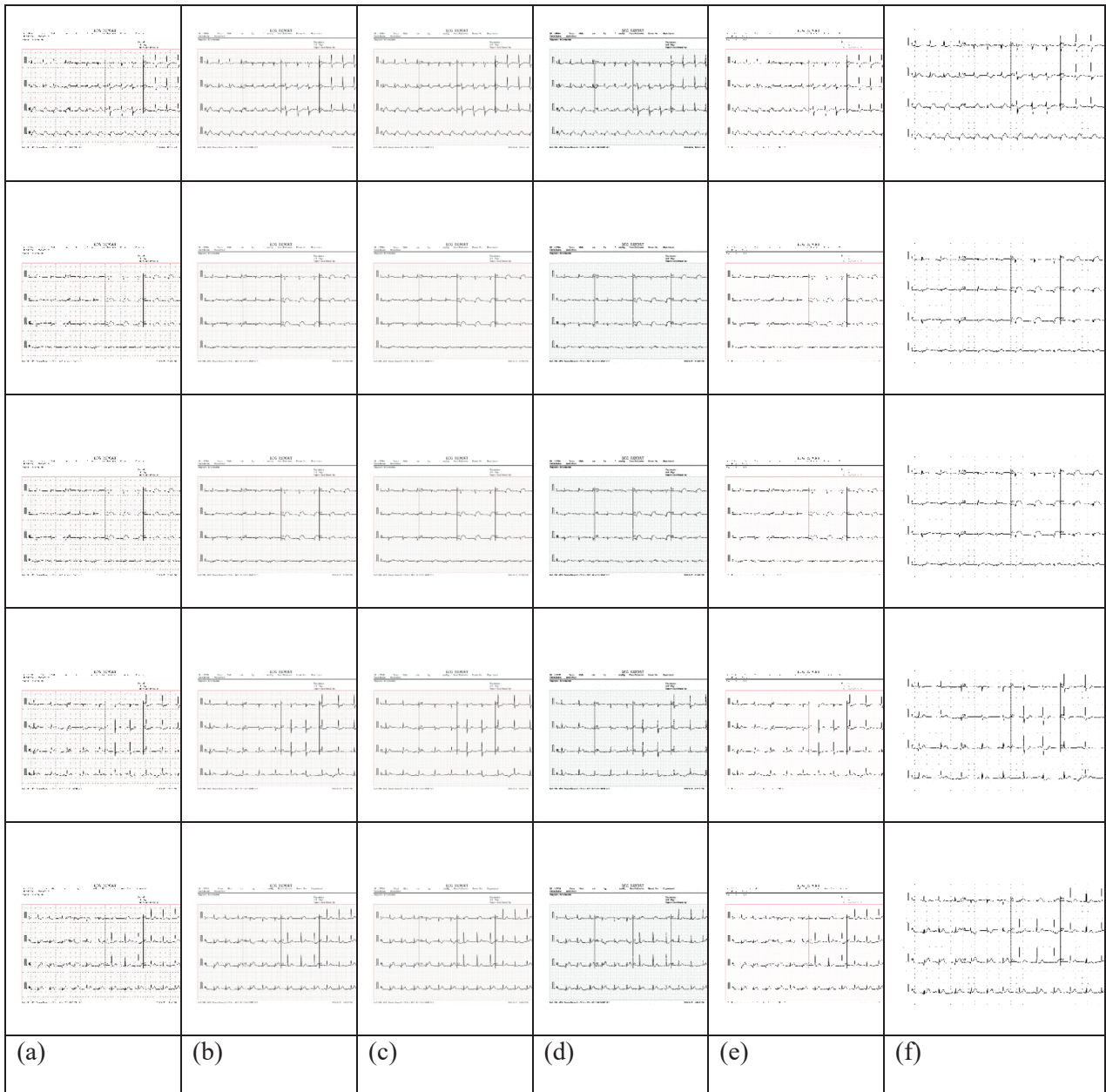


Figure 8: Images for CVD detection using ECG signal (A) Original images, (B) Gaussian filtering, (C) Mean filtering, (D) Wiener filtering, (E) Conventional median filtering, and (F) Improved median filtering. CVD, cardiovascular disease.

d.i. PSNR and SSIM analysis

PSNR is a metric measuring the quality of a reconstructed signal by comparing it to the original signal, considering the ratio of the extreme possible signal rate to the mean squared error. SSIM, a metric assessing the structural similarity between two images, considers luminance, contrast, and structure, providing an index within the range of -1 to 1 . In Table 2, the analysis of PSNR and SSIM for Improved Median Filtering is presented, comparing its performance with Mean filter, Gaussian filter, Wiener filter, and Conventional Median Filter. Furthermore, achieving higher PSNR and SSIM value is desirable for efficacious performance of CVD detection. Generally, the PSNR of the IMF is notably high at 37.804, while the traditional methods scored least PSNR values, notably, Mean Filter = 23.586, Gaussian Filter = 22.136, Weiner Filer = 19.997, and Conventional Median Filter = 30.656, respectively. Likewise, the Improved Median Filtering demonstrates the highest SSIM value at 0.915, while conventional filtering models exhibit comparatively lower SSIM ratings.

e. Comparative assessment on positive metric

The Squeeze-ICNN model represents a noteworthy progression in the domain of medical image investigation, precisely for the detection of CVDs through ECG signal images. In order to assess its effectiveness, a comprehensive analysis comparing various metrics with established models such as LSTM, DCNN, Bi-GRU, SqueezeNet, DenseNet, DNN [1], and CNN [8] is imperative. One crucial aspect of this evaluation is the consideration of positive metric values, which play a pivotal role in determining the model's ability to accurately detect instances of CVD. As we delve

into the analysis, the focus will be on elucidating how the Squeeze-ICNN model outperforms its counterparts, showcasing higher positive metric values as a testament to its efficacy in accurately identifying and classifying cardiovascular abnormalities in ECG signal images. Figure 9 visually represents the results of our comparison, showing the detailed differences in performance among the various models. Here, for the training data 90%, the SqueezeNet and DenseNet achieved detection accuracies exceeding 90%, yet the Squeeze-ICNN scheme outperformed them with a remarkable 95% detection accuracy. Likewise, the Squeeze-ICNN approach consistently demonstrated superior performance, surpassing 92% detection accuracy at the 70%, 80%, and 90% training data.

The accuracy improvement with increased data availability (92%–95.6% at 90% training) is consistent with the expected learning behavior in which the model can generalize better with more labeled data. The increase in accuracy through learning is arguably more important because the model escaped from potential overfitting by applying dropout, BN, and early stopping. Overall, the variance is sufficiently low across the multiple random train-test splits ($\sigma = 0.028$), suggesting that the increase in accuracy is due to learning, and not due to bias, and the model is simply learning features from more diverse training samples. These results suggest that the performance improvement are gains in true learning and not overfitting.

Additionally, the Squeeze-ICNN scheme exhibits a sensitivity of 0.923 with a training data of 70%. By contrast, traditional schemes, such as LSTM (0.730), DCNN (0.813), Bi-GRU (0.781), SqueezeNet (0.861), DenseNet (0.827), DNN [1] (0.762), and CNN [8] (0.800), yielded lower sensitivity values, underscoring the superior sensitivity of the Squeeze-ICNN approach. Moreover, the Squeeze-ICNN scheme attains the greatest precision and specificity, with rates of 0.978 and 0.922, respectively. These values markedly exceed those attained through traditional methodologies. The significant outcomes noticed in the positive measure estimation highlight the efficacy of the Squeeze-ICNN method in identifying CVD through the analysis of ECG signal images. This success is credited to the sophisticated detection strategy and enhancements applied in the preprocessing, feature extraction, and classification procedures.

In terms of clinical cardiology, it is essential to make a correct diagnosis in order to provide adequate early management and intervention in the treatment of CVD, such as MI and arrhythmias. High

Table 2: Analysis on PSNR and SSIM

Methods	PSNR	SSIM
Mean filter	23.586	0.770
Gaussian filter	22.136	0.707
Weiner filter	19.997	0.696
Conventional median filter	30.656	0.877
IMF	37.804	0.915

IMF, improved median filter.

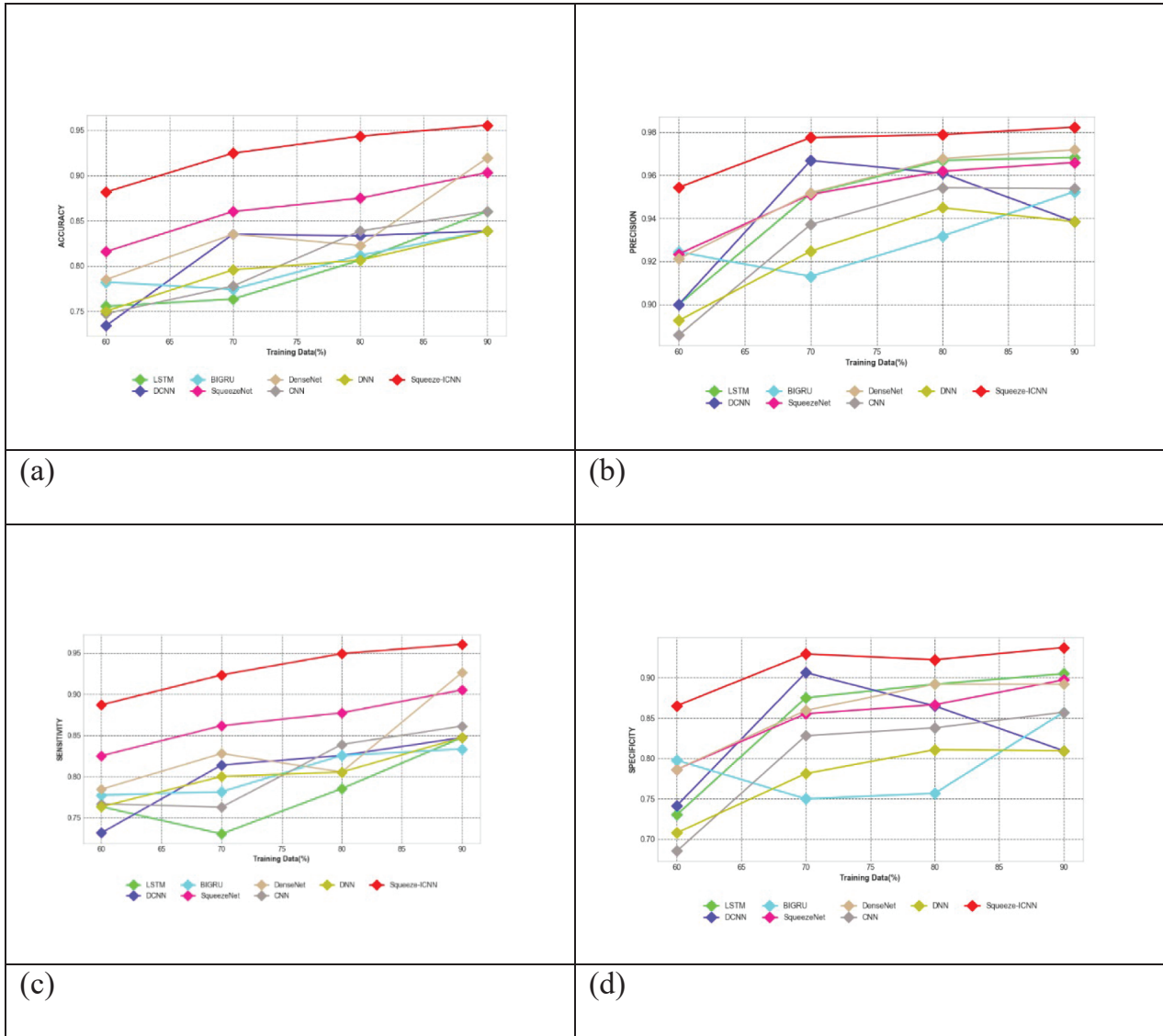


Figure 9: Comparative analysis on positive metrics for Squeeze-ICNN and traditional schemes. CNN, convolutional neural network; DCNN, deep convolutional neural network; DNN, deep neural network; ICNN, improved convolutional neural network.

sensitivity (0.923) in the Squeeze-ICNN model allows for unintended zero missed cases of those containing true abnormal cardiac condition. If a real abnormal cardiac condition was missed, there could be severe life-threatening complications when left untreated. In a similar fashion, high precision value (0.978) means the model does not have a high false alarm rate and therefore avoids wasted clinical follow-ups. More importantly, this value aligns with actual practice, since ECG chart interpretation becomes an important differentiating factor when diagnosing patients

in acute care with time constraints. The promising outcomes of DL when interpreting ECG images is not new, as described in previous studies such as DeepECG by Changling Li et al. [3], which also revealed better performance detections of arrhythmias using similar DCNN-based architectures.

They simulated its performance against the standard CNN and SqueezeNet models. The results presented in Table 3 demonstrate that ICNN was able to achieve greater accuracy and sensitivity as compared with the baseline CNN.

Table 3: Comparative assessment on positive metric

Model	Accuracy (%)	Precision (%)	Recall/sensitivity (%)	F1-score (%)
CNN (baseline)	88.0	86.5	84.3	85.4
ICNN (proposed)	91.2	92.4	90.1	91.2
SqueezeNet	90.6	91.1	89.3	90.2
Squeeze-ICNN	94.4	97.9	94.9	96.4

CNN, convolutional neural network; ICNN, improved convolutional neural network.

This demonstrates that the proposed additions to the architecture, in combination with the proposed activation function, were able to learn more effectively from the ECG patterns. Moreover, the Squeeze-ICNN model, which was an integration of SqueezeNet and ICNN, achieved the highest F1 score and precision in diagnostic performance across all tested models, even compared with the baseline model, SqueezeNet.

f. Comparative assessment on negative metric

Figure 10 illustrates the assessment of negative metrics for the Squeeze-ICNN method in comparison to other models such as LSTM, DCNN, Bi-GRU, SqueezeNet, DenseNet, DNN [1], and CNN [8], for the detection of CVD using ECG signal images. Lowering the negative metric value is crucial for improving the effectiveness of disease detection. Upon thorough examination of all the graphs, it is obvious that the Squeeze-ICNN method consistently achieved the least negative metric values across all training data, highlighting its superior performance in CVD detection. Primarily, with a training data of 90%, the Squeeze-ICNN method exhibited an FPR of 0.062. In comparison, LSTM, DCNN, SqueezeNet, DenseNet, DNN [1], and CNN [8] achieved the lowest FPRs of 0.095, 0.190, 0.102, 0.108, 0.142, and 0.190, respectively. Moreover, the FNR of the Squeeze-ICNN scheme spans from 0.113 to 0.039, demonstrating a significantly lower value compared with conventional strategies. Consequently, the Squeeze-ICNN scheme has exceeded the performance of previously utilized schemes, yielding more satisfactory negative metric ratings. This fascinating indication highlights its outstanding ability in accomplishing efficient detection of CVD through the analysis of ECG signal images.

g. Comparative assessment on other metrics

The comparative assessment of the Squeeze-ICNN method is contrasted with LSTM, DCNN, Bi-GRU, SqueezeNet, DenseNet, DNN [1], and CNN [8] concerning other metrics for the detection of CVD using ECG signal images, as depicted in Figure 11. Primarily, the goal is to maximize the values of the other metrics for the effective detection of CVD. In particular, at a training data of 60%, the Squeeze-ICNN scheme achieves an F-measure of 0.919, surpassing the F-measures of other models, namely LSTM (0.826), DCNN (0.807), Bi-GRU (0.844), SqueezeNet (0.871), DenseNet (0.847), DNN [1] (0.821), and CNN [8] (0.822), respectively. Furthermore, training data = 90, the Squeeze-ICNN scheme achieves MCC and NPV rates of 0.874 and 0.869, respectively. By contrast, LSTM, DCNN, Bi-GRU, SqueezeNet, DenseNet, DNN [1], and CNN [8] obtained the lowest ratings for both MCC and NPV. Certainly, this serves as clear evidence that the Squeeze-ICNN methodology excels in the detection of CVD. This achievement is made possible through the implementation of an enhanced preprocessing technique utilizing median filtering, coupled with a hybrid classification method and an improved hierarchy of skeleton feature-based feature extraction.

h. Ablation study on Squeeze-ICNN

A thorough investigation was systematically conducted through ablation to estimate the effect of the Squeeze-ICNN approach. This detailed process explores a complete sympathetic of the detailed improvements these elements brought to the complete success of the Squeeze-ICNN method. Table 4 displays the ablation assessment performed on the Squeeze-ICNN framework, including the model

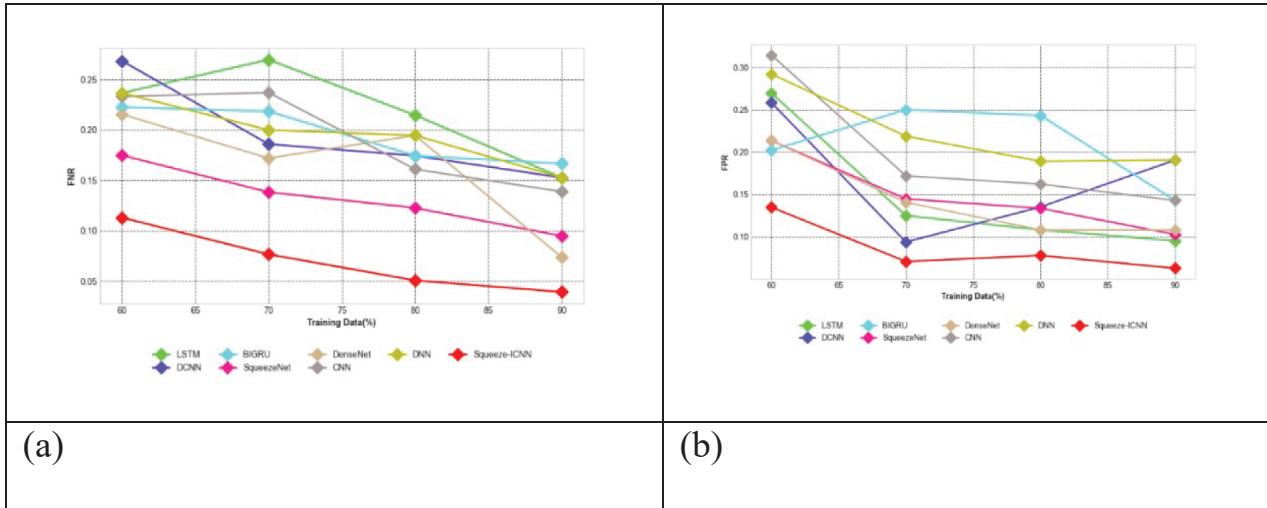


Figure 10: Comparative analysis on negative metrics for Squeeze-ICNN and traditional schemes. CNN, convolutional neural network; DCNN, deep convolutional neural network; DNN, deep neural network; ICNN, improved convolutional neural network.

with traditional preprocessing, the model integrating a conventional skeleton hierarchy, and the model amalgamating SqueezeNet with conventional CNN for CVD detection using ECG signal images. Furthermore, the precision of the Squeeze-ICNN approach is 0.979, whereas the model with conventional preprocessing attains a precision of 0.785, the model with a conventional hierarchy of skeleton attains 0.788, and SqueezeNet + conventional CNN demonstrates a precision of 0.836. Likewise, the FNR for Squeeze-ICNN method is 0.051, while the model with conventional preprocessing, the model with a conventional hierarchy of skeleton, and SqueezeNet + conventional CNN exhibit FNRs of 0.215, 0.212, and 0.164, correspondingly. Additionally, the proposed Squeeze-ICNN models demonstrate superior performance compared with employing models with conventional preprocessing techniques, SqueezeNet + Conventional CNN, and model with conventional hierarchy of skeleton.

i. Statistical evaluation on accuracy

To ensure the precision in results, every method undergoes a thorough statistical assessment, involving a scrupulous inspection of key statistical metrics like “Maximum, Minimum, Mean, Standard Deviation, and Median.” Table 5 describes the statistical analysis comparing the Squeeze-ICNN method with LSTM, DCNN, Bi-GRU, SqueezeNet, DenseNet,

DNN [1], and CNN [8] for the detection of CVD using ECG signal images. In terms of the median statistical metric, the Squeeze-ICNN scheme demonstrates an accuracy of 0.934. This outperforms conventional methods, with accuracy scores as follows: LSTM = 0.785, DCNN = 0.834, Bi-GRU = 0.797, SqueezeNet = 0.868, DenseNet = 0.829, DNN [1] = 0.808, and CNN [8] = 0.801, respectively. Moreover, when considering the maximum statistical measure, the Squeeze-ICNN system accomplished a higher accuracy value of 0.956. Meanwhile, LSTM, DCNN, Bi-GRU, SqueezeNet, DenseNet, DNN [1], and CNN [8] yielded lower accuracy ratings.

j. Limitations and future scope for cross-dataset validation

The present evaluation was conducted on a single dataset, but the model architecture is intended to generalize to new data because of the feature fusion, custom activation function, and regularization methods. The model showed consistent performance across multiple randomized train-test splits, and was low variance—indicating good generalization. However, a completely new and external dataset was not used in this study, and therefore, we are aware of this. Future work will validate the model on publicly available ECG image datasets, to assess its robustness and generalizability across different patient populations and acquisition settings.

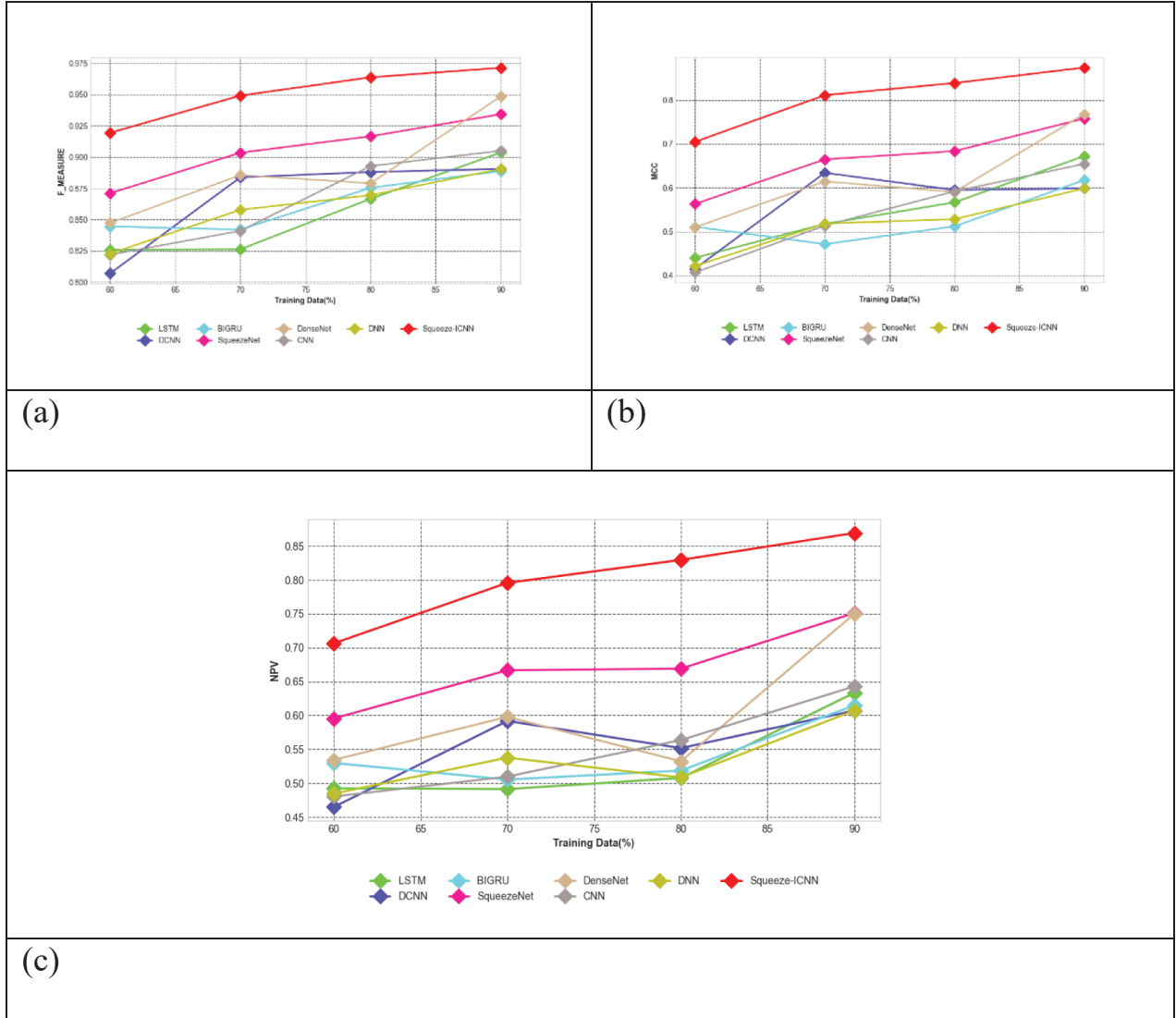


Figure 11: Comparative analysis on other metrics for Squeeze-ICNN and traditional schemes. CNN, convolutional neural network; DCNN, deep convolutional neural network; DNN, deep neural network; ICNN, improved convolutional neural network.

k. Comparative discussion: Strengths and limitations

While there are benefits to approaching the construction of model and classifier on the EPS framework, several challenges remain. To confirm, the present evaluation has relied on a singular dataset, and the need for external validation with other ECG datasets will facilitate the further understanding of utility and generalizability regarding the use of clinical ECG data. The model, while relatively lightweight when compared with deeper CNN's, will require real-time deployment and testing on edge devices in future works. The introduction of real-world variabilities and

noise types from multiple sources will further improve clinical feasibility (Table 6).

V. Conclusion

This work proposed a Squeeze-ICNN model for CVD using ECG image. In the preprocessing step, the noise in the provided input ECG image was eliminated using an IMF. The preprocessed filtered image was used to extract shape features, deep features, and LGXP. From the filtered preprocessed image, an improved Hierarchy of Skeleton form features was retrieved to improve the histogram features. The pre-trained NN models ResNet, VGG 16, and Inception

Table 4: Ablation assessment on Squeeze-ICNN, model with conventional preprocessing, model with conventional hierarchy of skeleton, and SqueezeNet + Conventional CNN

Metrics	Model with conventional preprocessing	Squeeze-ICNN	SqueezeNet + conventional CNN	Model with conventional hierarchy of skeleton
Sensitivity	0.785	0.949	0.836	0.788
F-measure	0.785	0.964	0.836	0.788
Accuracy	0.851	0.944	0.908	0.852
FPR	0.127	0.078	0.068	0.126
Specificity	0.873	0.922	0.932	0.874
FNR	0.215	0.051	0.164	0.212
NPV	0.873	0.830	0.932	0.874
Precision	0.785	0.979	0.836	0.788
MCC	0.741	0.839	0.788	0.745

CNN, convolutional neural network; FNR, false negative rate; ICNN, improved convolutional neural network.

Table 5: Statistical assessment on accuracy

Statistical metrics	LSTM	DCNN	SqueezeNet	CNN [8]	DenseNet	Bi-GRU	DNN [1]	Squeeze-ICNN
Mean	0.796	0.810	0.864	0.798	0.841	0.802	0.806	0.926
Minimum	0.755	0.734	0.816	0.750	0.785	0.774	0.747	0.882
Standard deviation	0.042	0.044	0.032	0.032	0.049	0.026	0.045	0.028
Median	0.785	0.834	0.868	0.801	0.829	0.797	0.808	0.934
Maximum	0.860	0.839	0.903	0.839	0.919	0.839	0.860	0.956

CNN, convolutional neural network; DCNN, deep convolutional neural network; DNN, deep neural network; ICNN, improved convolutional neural network.

V3 based features were extracted from the deep features. Lastly, the Squeeze-ICNN model, a hybrid model that incorporates the SqueezeNet and IMP-CNN models, was used to classify CVDs. As a result, the averaged classification model classifies the output of the suggested model as 0 or 1. Additionally, the proposed model's performance was compared with that of the conventional models with regard to metrics like sensitivity, accuracy, FPR, and MCC.

To effectively train the proposed Squeeze-ICNN model, a set of hyperparameters, including learning rate (0.001), batch size (32), dropout (0.5), and activation (Comb SigHyper-H-Sine), were finalized in advance. The model was then trained with Adam for 100 epochs. To aid in the interpretability of the model, shape analysis was performed to examine the feature

contribution on the extracted features. The analysis discovered that deep features (from ResNet, VGG16), shapes features, and LGXP features were found to have considerable importance on the classification predictions, which allowed full transparency into the model's decision-making process.

The results demonstrate that the standalone ICNN model had significant improvements compared with standard CNNs, and further benefiting from the incorporation of SqueezeNet in order to receive better classification accuracy, sensitivity, and F1-score.

Going forward, we will focus on testing the proposed model on publicly available ECG datasets to examine the robustness, generalizability, and clinical applicability across multiple acquisition settings and diverse populations.

Table 6: Comparative overview of the proposed method vs existing studies

Study/ year	Preprocessing technique	Feature type	Model used	Accuracy (%)	Key advantages	Limitations
Khan et al. [1]	Not specified	Raw ECG Images	MobileNet- based DNN	92	Lightweight architecture	Limited generalization, no hybrid features
Fatema et al. [2]	Basic filtering	Deep features (CNN)	InceptionV3 + ResNet50	~90	Combined model improved feature learning	Imbalanced dataset, minimal filtering
Li et al. [3]	Data augmentation	TL (inceptionV3)	DeepECG	91–93	Pretrained on large ECG datasets	Dataset-specific architecture
Sadad et al. [6]	Basic filtering	CNN + attention	Lightweight CNN	89–91	IoT compatible, efficient	Lacked multi-feature fusion
Proposed (This Study)	IMF	Hybrid: Deep + LGXP + shape	SqueezeNet + ICNN	95.6	Low FNR (5.1%), robust features, compact size	Limited dataset, needs cross-validation

CNN, convolutional neural network; DNN, deep neural network; FNR, false negative rate; ICNN, improved convolutional neural network; IoT, Internet of Things; IMF, improved median filter; LGXP, local Gabor XOR pattern; TL, transfer learning.

References

- [1] Ali Haider Khan, Muzammil Hussain, and Muhammad Kamran Malik., “Cardiac disorder classification by electrocardiogram sensing using deep neural network”, Complexity, pp.1-8, 2021.
- [2] Kaniz Fatema, Sidratul Montaha, Md. Awlad Hossen Rony, Sami Azam, Md. Zahid Hasan and Mirjam Jonkman, “A robust framework combining image processing and deep learning hybrid model to classify cardiovascular diseases using a limited number of paper-based complex ECG images”, Biomedicines, vol.10, p.2835, 2022.
- [3] Changling Li, Hang Zhao, Wei Lu, Xiaochang Leng, Li Wang, Xintan Lin, Yibin Pan, Wenbing Jiang, Jun Jiang, Yong Sun, Jianan Wang, Jianping Xiang, “DeepECG: Image-based electrocardiogram interpretation with deep convolutional neural networks”, Biomedical Signal Processing and Control, vol.69, 2021.
- [4] Weibo Song, “A new method for refined recognition for heart disease diagnosis based on deep learning”, Information, vol.11, p.556, 2020.
- [5] Zeynep Hilal Kilimci, Mustafa Yalcin, Ayhan Kucukmanisa and Amit Kumar Mishra, “Heart Disease

Detection using Vision-Based Transformer Models from ECG Images”, arXiv preprint arXiv:2310.12630, 2023.

- [6] Tariq Sadad, Mejdil Safran, Inayat Khan, Sultan Alfarhood, Razaullah Khan and Imran Ashraf, “Efficient Classification of ECG Images Using a Lightweight CNN with Attention Module and IoT”, Sensors, vol.23, p.7697, 2023.

- [7] S.Irin Sherly and G.Mathivanan, “ECG BASED MULTI MODAL FRAMEWORK FOR HEART DISEASE DETECTION”, Journal of Theoretical and Applied Information Technology, vol.101, 2023.

- [8] Mohammed F El-Habibi, “Automated Detection of Cardiovascular Diseases using Deep Learning and Electrocardiogram (ECG) Images: A Convolutional Neural Network Approach” International Journal of Academic Engineering Research (IJAER), vol.7, 2023.

- [9] Lotfi Mhamdi, Oussama Dammak, François Cottin and Imed Ben Dhaou, “Artificial intelligence for cardiac diseases diagnosis and prediction using ECG images on embedded systems”, Biomedicines, vol.10, p.2013, 2022.

- [10] S. Karthik, M. Santhosh, M. S. Kavitha and A. Christopher Paul, “Automated Deep Learning Based

Cardiovascular Disease Diagnosis Using ECG Signals", *Computer Systems Science & Engineering*, vol.42, 2022.

[11] V. Jahmunah, E.Y.K. Ng, Tan Ru San, U. Rajendra Acharya, "Automated detection of coronary artery disease, myocardial infarction and congestive heart failure using GaborCNN model with ECG signals", *Computers in Biology and Medicine*, vol.134, 2021.

[12] Eftal Sehirli, Muhammed Kamil Turan, "A novel method for segmentation of QRS complex on ECG signals and classify cardiovascular diseases via a hybrid model based on machine learning", *International Journal of Intelligent Systems and Applications in Engineering*, vol.9, pp.12-21, 2021.

[13] Adyasha Rath, Debahuti Mishra and Ganapati Panda., "Imbalanced ECG signal-based heart disease classification using ensemble machine learning technique", *Frontiers in Big Data*, vol.5, p.1021518, 2022.

[14] Jinyong Cheng, Qingxu Zou and Yunxiang Zhao., "ECG signal classification based on deep CNN and BiLSTM", *BMC medical informatics and decision making*, vol.21, pp.1-12, 2021.

[15] Venkataramanaiah, B., Kamala, J., "ECG signal processing and KNN classifier-based abnormality detection by VH-doctor for remote cardiac health-care monitoring", *Soft Comput*, vol.24, 17457–17466, 2020.

[16] Pławiak, P., Acharya, U.R., "Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals", *Neural Comput & Applic*, vol.32, 11137–11161, 2020.

[17] Md. Rashed-Al-Mahfuz, Mohammad Ali Moni, Pietro Lio, Sheikh Mohammed Shariful Islam, Shlomo Berkovsky, Matloob Khushi, Julian M. W. Quinn, "Deep convolutional neural networks based ECG beats classification to diagnose cardiovascular conditions", *Biomed. Eng. Lett.*, vol.11, 147–162, 2021.

[18] Amin Ullah, Sadaqat ur Rehman, Shanshan Tu, Raja Majid Mehmood, Fawad and Muhammad Ehatisham-ul-haq, "A hybrid deep CNN model for abnormal arrhythmia detection based on cardiac ECG signal", *Sensors*, vol.21, p.951, 2021.

[19] Nazrul Anuar Nayan, Hafifah Ab Hamid, Mohd Zubir Suboh, Rosmina Jaafar, Noraidatulakma Abdullah, Nurul Ain Mhd Yusof, Mariatul Akma Hamid, Nur Farawahida Zubiri, Azwa Shawani Kamalul Arifin, Syakila Mohd Abd Daud, Mohd Arman Kamaruddin, A. Rahman A. Jamal, "Cardiovascular disease prediction from electrocardiogram by using machine learning", 2020.

[20] Padmavathi Kora, Ajith Abraham, K Meenakshi, "Heart disease detection using hybrid of bacterial foraging and particle swarm optimization", *Evolving Systems* vol.11, 15–28, 2020.

[21] Abdullah Alqahtani, Shtwai Alsubai, Mohemmed Sha, Lucia Vilcekova, and Talha Javed, "Cardiovascular Disease Detection using Ensemble Learning", *Hindawi, Computational Intelligence and Neuroscience*, Volume 2022, Article ID 5267498, 9 pages, doi: 10.1155/2022/5267498

[22] Hana H. Alalawi and Manal S. Alsuwat, "Detection of Cardiovascular Disease using Machine Learning Classification Models", *International Journal of Engineering Research & Technology (IJERT)*, ISSN: 2278-0181, Volume 10 Issue 07, July-2021.

[23] SKonstantinos C. Siontis, Peter A. Noseworthy, Zach I. Attia & Paul A. Friedman, "Artificial intelligence-enhanced electrocardiography in cardiovascular disease management", *Nature Reviews Cardiology* volume 18, pages465–478 (2021), doi: 10.1038/s41569-020-00503-2

[24] SiripuriKiran,GantaRaghothamReddy,GirjaS.P., VenkatramuluS,KumarDorthiandChandraShekharRao V., "A Gradient Boosted Decision Tree with Binary Spotted Hyena Optimizer for cardiovascular disease detection and classification", *Healthcare Analytics*, Volume 3, November 2023, 100173, doi: 10.1016/j.health.2023.100173

[25] Nikki van der Velde, Cécile P.M. Janus, Daniel J. Bowen BSc a, H. Carlijnne Hassing, Isabella Kardys, Flora E. van Leeuwen, Cynthia So-Osman, Remi A. Nout, Olivier C. Manintveld, Alexander Hirsch, "Detection of Subclinical Cardiovascular Disease by Cardiovascular Magnetic Resonance in Lymphoma Survivors", *JACC: CardioOncology*, Volume 3, Issue 5, December 2021, Pages 695-706, doi: 10.1016/j.jacc.2021.09.015

[26] Kellen Sumwiza, Celestin Twizere, Gerard Rushingabigwi, Pierre Bakunzibake and Peace Bamurigire, "Enhanced cardiovascular disease prediction model using random forest algorithm", *Informatics in Medicine Unlocked*, Volume 41, 2023, 101316, doi: 10.1016/j.imu.2023.101316

[27] Stephen Mariadoss and Felix Augustin, "Enhanced sugeno fuzzy inference system with fuzzy AHP and coefficient of variation to diagnose cardiovascular disease during pregnancy", *Journal of King Saud University - Computer and Information Sciences*, Volume 35, Issue 8, September 2023, 101659, doi: 10.1016/j.jksuci.2023.101659

[28] Youlian Zhu, Cheng Huang, "An Improved Median Filtering Algorithm for Image Noise Reduction", *Physics Procedia*, Volume 25, 2012, Pages 609-616, doi: 10.1016/j.phpro.2012.03.133

- [29] Shufu Xie, Shiguang Shan, Xilin Chen and Jie Chen, "Fusing Local Patterns of Gabor Magnitude and Phase for Face Recognition", IEEE Transactions on Image Processing, Vol. 19, No. 5, May 2010, doi: 10.1109/TIP.2010.2041397
- [30] Wei Chen, Lichun Sui and Zhengchao Xu, Yu Lang, "Improved Zhang-Suen Thinning Algorithm in Binary Line Drawing Applications", 2012 International Conference on Systems and Informatics (ICSAI 2012), pp: 1947-1950, 2012.
- [31] Ma Xin and Jing Xiaojun, "Palm vein recognition method based on fusion of local Gabor histograms", The Journal of China Universities of Posts and Telecommunications, 2017, doi: 10.1016/S1005-8885(17)60242-5
- [32] D. Albashish, R. Al-Sayyed, A. Abdullah, M. H. Ryalat and N. Ahmad Almansour, "Deep CNN Model based on VGG16 for Breast Cancer Classification", 2021 International Conference on Information Technology (ICIT), Amman, Jordan, 2021, pp. 805-810, doi: 10.1109/ICIT52682.2021.9491631
- [33] Jianfang Cao, Minmin Yan, Yiming Jia, Xiaodong Tian and Zibang Zhang, "Application of a modified Inception-v3 model in the dynasty-based classification of ancient murals", EURASIP Journal on Advances in Signal Processing, (2021) 2021:49, doi: 10.1186/s13634-021-00740-8
- [34] Tushar Nayak, Krishnaraj Chadaga, Niranjana Sampathila, Hilda Mayrose, Nitila Gokulkrishnan, Muralidhar Bairy G, Srikanth Prabhu, Swathi K. S, Shashikiran Umakanth, "Deep learning based detection of monkeypox virus using skin lesion images", Medicine in Novel Technology and Devices 18 (2023) 100243, doi: 10.1016/j.medntd.2023.100243
- [35] Jing Yang and Guanci Yang, "Modified Convolutional Neural Network Based on Dropout and the Stochastic Gradient Descent Optimizer", Algorithms 2018, 11(3), 28, doi: 10.3390/a11030028
- [36] Kuppusamy P, Raga Siri P, Harshitha P, Dhanyasri M and Celestine Iwendi, "Customized CNN with Adam and Nadam Optimizers for Emotion Recognition using Facial Expressions", 2023 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), IEEE, 2023, doi: 10.1109/WISPNET57748.2023.10134002
- [37] Adamu I. Abubakar, Haruna Chiroma, Sameem Abdulkareem, Abdulsalam Ya'u Gital, Sanah Abdullahi Muaz, Jafaar Maitama, Muhammad Lamir Isah and Tutut Herawan, "Modified Neural Network Activation Function", 2014 4th International Conference on Artificial Intelligence with Applications in Engineering and Technology, 2014, doi:10.1109/ICALET.2014.12
- [38] K. Vijayaprabakaran and K. Sathiyamurthy, "Towards activation function search for long short-term model network: A differential evolution based approach", Journal of King Saud University – Computer and Information Sciences, volume 34, pp: 2637–2650, 2022, doi: 10.1016/j.jksuci.2020.04.015
- [39] <https://data.mendeley.com/datasets/gwbz3fsgp8/2>
- [40] Muthukumar, K. A., et al. "Integrating electrocardiogram and fundus images for early detection of cardiovascular diseases." Scientific Reports 15.1 (2025): 4390. doi: 10.1038/s41598-025-87634-z.
- [41] Zhu, Jiayuan, et al. "Cardiovascular disease detection based on deep learning and multi-modal data fusion." Biomedical Signal Processing and Control 99 (2025): 106882. doi: 10.1016/j.bspc.2024.106882.