

Original Article

# Breast Cancer Detection on Mammographic Images using Hyper Parameter Tuning & Optimization: A Convolutional Neural Network & Transfer Learning Approach

Pratheep Kumar P<sup>1</sup>, V. Mary Amala Bai<sup>2</sup>

<sup>1</sup> Computer Science and Engineering, Noorul Islam Centre for Higher Education, Kumaracoil, Kanyakumari District, Tamilnadu, India.

<sup>2</sup> Information Technology, Noorul Islam Centre for Higher Education, Kumaracoil, Kanyakumari District, Tamilnadu, India.

<sup>1</sup>Corresponding Author : [pratheep.kp@gmail.com](mailto:pratheep.kp@gmail.com)

Received: 23 June 2022

Revised: 03 August 2022

Accepted: 15 August 2022

Published: 06 September 2022

**Abstract** - Breast cancer is probably the most well-known; it is also the leading cause of death in women worldwide. Incidentally, it can be cured if discovered early enough. The key issues for rural locations are that radiologists are significantly less likely to identify Breast Cancer (BC) utilizing mammography pictures in testing camps and that early diagnosis of these cancers is more important for all medical specialists. As a result, this study proposes an advanced deep learning-based tool for BC, with the following steps: a) Data Source: The Mammographic Image Analysis Society (MIAS) database of digital mammograms (v1.21), which contains 322 pictures and also real-world data from VPS Lakeshore Hospital Kochi which contains 4118 images, b) Filtering and histogram-based approach for preprocessing, c) feature extraction using a convolutional autoencoder for extracting features from the input, d) feature selection with Recursive Feature Elimination (RFE) for dimensionality reduction, and e) classification using a convolutional neural network with the help of transfer learning. Experiments are performed on various state-of-the-art models, and the suggested model outperforms in various measures (accuracy;0.96, precision;0.95, sensitivity;0.97, specificity:0.98).

**Keywords** - Breast Cancer, Classification, Convolutional Neural Network, Deep learning, Mammogram, Transfer Learning.

## 1. Introduction

Medical imaging professionals have begun leveraging deep learning and machine learning to improve cancer screening accuracy with the rapid growth of machine learning. Despite its benefits, mammography screening has notable risks of false positives and negatives. The United States has the second-highest rate of cancer deaths among women from breast cancer [1,2]. Screening mammography's average sensitivity and specificity in the United States are 86.9% and 88.9%, respectively [3]. Computer-Aided Detection and Diagnosis (CADD) software [4] was developed during the 1990s to increase radiologists' confidence regarding the projected accuracy of screening mammography. There was little evidence that early commercial CAD systems improved performance substantially [5,6] and that their development stagnated for more than a decade after their introduction. Creating deep learning tools to aid radiologists and improve mammography screening accuracy is gaining increasing interest as deep learning has proved too high performance in visual object

recognition and detection [7-14]. A recent study [15,16] found that CAD systems running using deep learning performed as well and even better than radiologists in independent mode and assistance mode. Figure 1 shows BC data for several countries.

By contrast, Digital Pathology (DP) provides high-quality images derived from digitizing histology slides. Through the use of image analysis tools, these digital pictures are employed for detection, segmentation, and classification. Deep Learning (DL) with CNNs requires additional procedures, such as digital staining, to comprehend patterns for picture categorization [17]. CNN's contribution to medical imaging research isn't limited to deep CNN for feature extraction. Indeed, using CNN for synthetic picture rendering is a second topic that might aid medical research. Wahab and Khan [15] employed multifaceted fused-CNN and a hybrid descriptor to show that suitable colour and textural qualities may be produced to aid mitotic count-based ROI selection at reduced resolution.



Women with higher BMIs have a higher risk of being diagnosed later than women with a normal body mass index (Women overweight were diagnosed at a later stage than women in their healthy weight range, with 42 percent diagnosed at a late stage). As a result, this article suggests hyperparameter tuning and optimization for breast cancer detection that can be useful for tuning the hyperparameters

and optimizing the risks associated with breast cancer in later stages. So, the problem is that most deep learning models try to perform breast cancer detection. Still, ultimately, the results didn't reach that efficiency as training the model, rather than increasing knowledge of the model with certain parameters and transfer learning.

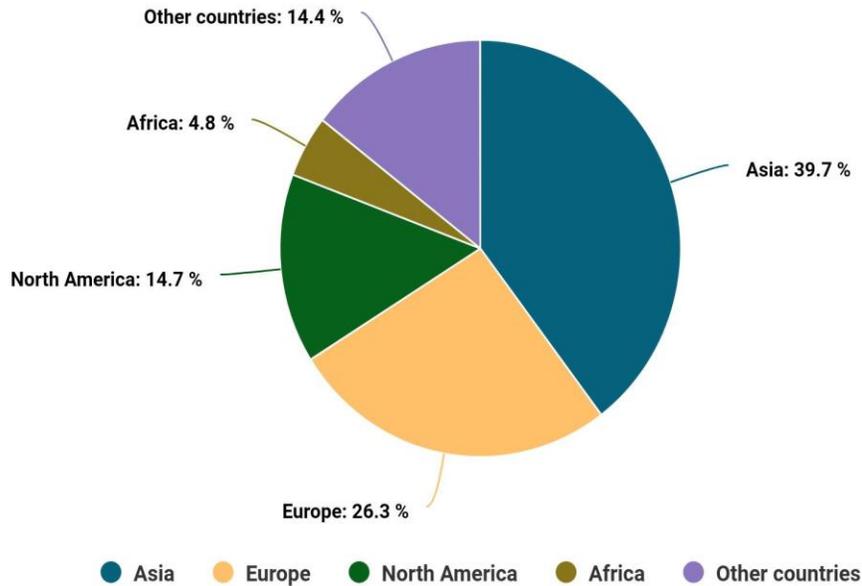


Fig. 1 Overall percentage-wise statistics of Breast Cancer

**1.1. Research Gap**

- After reviewing the existing literature, it is noticed that there is a lack of proper feature representation in the existing methods, which will reduce the overall accuracy of the classifier.
- Also, it has been identified as like, the models used are performed on their own without any additional stages with proper technique
- The models were also not trained enough to gain such knowledge to understand the data given for a model to perform the classification
- Mostly there is a great lack of improving the knowledge of models using transfer learning to improve the overall results in mammogram images
- The reliability of the existing methods needs to be improved for applications that require precise classification results.

**1.2. Key Novelty**

As part of the ongoing investigation into deep learning for breast cancer detection, this paper presents a method with the following objectives:

- To create a breast cancer detection system that is based on deep learning.
- The effective approach proposed will be worked over Mammographic images.
- A proper tool will be helpful for rural areas where radiologists' services are unavailable for BC detection.
- Autoencoder and RFE are used for feature extraction, and a better result is obtained using both.
- For early detection of BC, a convolutional neural network-based transfer learning method will be used.

**1.3. Organization of the paper**

The first section includes a BC and deep learning summary. The second section contains a literature review. Section three discusses the methodology, the fourth section describes the performance measures, and to wrap up, the conclusion is provided in section 5.

**2. Literature Review**

A novel framework for breast cancer classification and detection from cytology images was established by Khan et al. (2019) [18]. They proposed a clever, profound learning

structure for identifying and classifying breast cancer in breast cytology images using the concept of transfer learning. As a general rule, deep learning designs are displayed to be issue explicit and are acted in separation. As opposed to traditional learning ideal models, which create and yield disconnection, transfer learning is expected to use the acquired information during the arrangement of one issue into one more related issue. In the proposed structure, highlights from images are separated utilizing pre-prepared CNN designs, to be specific, GoogLeNet, Visual Geometry Group Network (VGGNet) and Residual Networks (ResNet), which are taken care of into an FC layer for characterization of malignant and benign cells using average pooling classification.

From the perspectives of transfer learning, preprocessing, and Convolutional Neural Networks (CNNs), Ayana et al. (2021) [19] provided an overview of methods for detecting and categorizing ultrasound breast images. Finally, different works were compared, and challenges and perspectives were examined.

The study by Vesal et al. (2018) [20] classifies images of breast tissue in four classes: normal, in situ cancer, benign, and invasive cancer. To accomplish this, slide preparation and color variations from the BACH 2018 grand challenge images are removed before normalizing them. With the help of recovered picture patches from the image database, ResNet50 and Inception-V3 convolutional neural networks learned domain-specific characteristics necessary to categorize images.

A transfer learning technique developed by Chang et al. (2017) [21] for detecting breast cancer using histopathology photos, based on Google's Inception v3 model originally developed for non-medical image classification, was presented [22]. Its AUC of 0.93 illustrates how successful transfer learning can identify breast cancer. Ber et al. (2021) [23] used 8020 and cross-validation techniques to develop a deep-learning method for automatically identifying and diagnosing BC suspicious regions. There are model-specific DL architectures that apply to a given set of problems. TL applies what he learns from solving one challenge to another.

Makhtar et al. (2020) [24] assess the presence of a multi-classifier put together deep learning approach for datasets. Five classifiers are involved Sequential Minimal Optimization (SMO), choice tree (J48), arbitrary timberlands (RFs), Naïve Bayes (NB) and Instance-Based for K-Nearest neighbor (IBk). These classifiers will be consolidated and dissected utilizing a profound learning approach. This technique uses models of profound neural network that is a variation of Neural Network yet with a huge estimate to the human mind utilizing a development framework contrasted with a direct neural organization.

Purwanti & Apsari (2020) [25] characterize computerized mammograms into two classes, abnormal microcalcification and normal. The surface is one of the major mammographic qualities. The measurable textural of the Gray Level Cooccurrence Matrix (GLCM) utilized in portraying pictures are differentiation, energy and entropy. K-Nearest Neighbor (K-NN) and Fuzzy K-Nearest Neighbor (FK-NN) were proposed for arranging pictures. The aftereffect of the K-NN technique shows 77.78% exactness, 50.00% sensitivity and 100 percent specificity. The aftereffect of the FK-NN technique shows 88.89% exactness, 100 percent sensitivity and 80.00% specificity.

Kanmani et al. (2019) [26] comprises three phases, Preprocessing, Association Rule Mining and Classification. ID3 classifier is utilized for anticipating order because of responsiveness, particularity and precision. The principal objective of this strategy is to accomplish a higher exactness rate and lower blunder rate.

### 3. Methodology

In Figure 2, the general structure of the proposed system is represented as follows: a) Data collection: It is from a database of digital mammograms held by the Mammographic Image Analysis Society (v1.21) and a real-world dataset from VPS Lakeshore Hospital Kochi, which contains 4118 images. In addition to the MIAS 322 grayscale photos (161 pairs), truth data and Portable Gray Map files (161 pairs) are provided. b) Preprocessing: For removing the noises and improving the quality of the sample images taken from the dataset, models like filtering and histogram equalization are used. c) Feature extraction: Important features from the images are extracted using a convolutional autoencoder. Pretrained CNN architectures are also used for feature extraction. d) Feature selection: Relevant features are preserved, and insignificant features are removed using dimensionality reduction methods. Here, RFE is used as a selection procedure and e) Classification: Classification of normal, benign and malignant breast images is finally achieved with-a fully connected network.

#### 3.1. Data Collection

MIAS (Mammographic Institute Society Analysis database) is a collection of mammograms with a resolution of 200 microns and a size of 1024\*1024 pixels. Three hundred and twenty-two mammograms of the left and right breast from 161 women are included in this collection, which contains 54 cancerous lesions, 69 benign lesions, and 207 non-suspicious lesions. Among the information in this database is a list of mammograms stored in the MIAS database and pertinent information, such as the abnormality's class, the image's x-y coordinates, and the estimated radius (in pixels) of the circle which encloses the abnormality.

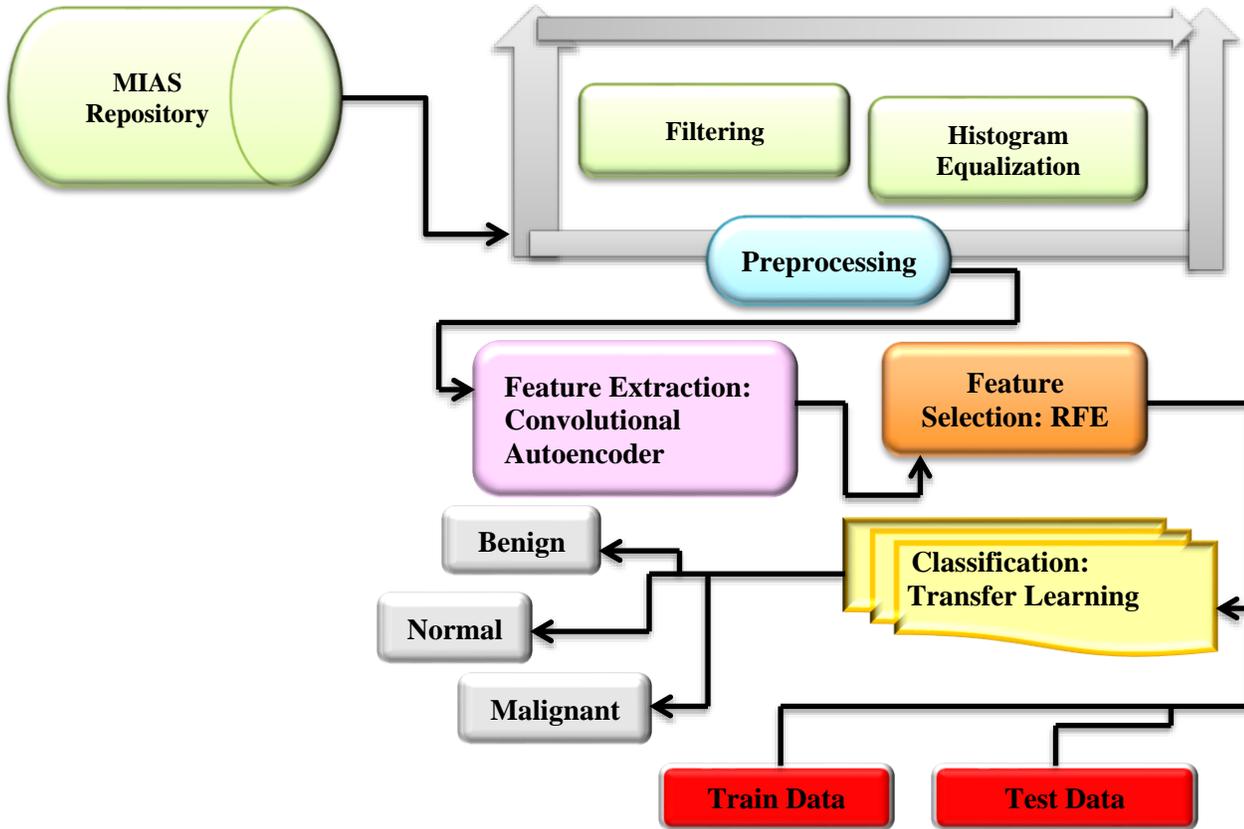


Fig. 2 The overall architecture of the proposed framework

Table 1. MIAS dataset overview

SL.No.	Description
1st	Reference number for MIAS database
2nd	Background tissues have the following characteristics: G: Fatty glandular F: Fatty D: Dense glandular
3rd	Anomaly class present: 1 CALC – Calcification 2 ARCH – Architectural distortion 3 MISC – Other, ill-defined masses 4 SPIC – Spiculated masses 5 CIRC – Well-defined/circumscribed masses 6 NORM E- Norm 7 ASYM – Asymmetry
4th	The severity of the anomaly; M E -Malignant B – Benign
5th & 6th	The centre of the abnormality of image-is X and Y coordinates.
7th	A circle enclosing the abnormality (in pixels) has a radius of approximately 9 pixels.

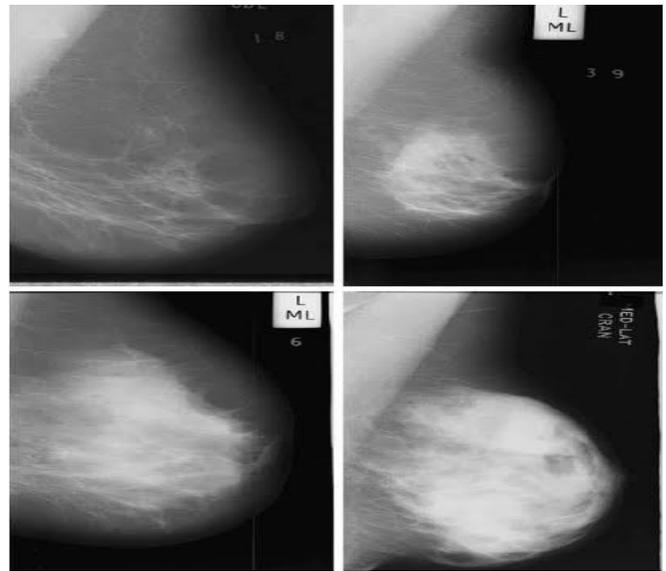


Fig. 3 Examples of MIAS Dataset instances

Different classifications are based on the type of abnormalities observed (calcifications, circumscribed masses, architectural deformities, and other ill-defined masses) [27, 28]. The real-world datasets were obtained from the VPS Lakeshore Hospital in Kochi. The dataset consists of 4118 jpg images of 1200 people (left CC and MLO and

right CC and MLO images), including 400 normal, 400 benign, and 400 malignant cases. Table 1 displays the MIAS dataset's overall features. Figure 3 shows a selection of photos from the MIAS dataset.

### 3.2. Preprocessing

Preprocessing is a crucial operation in which certain noises and abnormalities are eliminated to improve prediction.[29,30] Figure 4 illustrates the total steps and processes performed in two key phases. Initially, a filter approach is used to remove noise and abnormalities from these photos, and then the image is enhanced using histogram equalization in the second step.

#### 3.2.1. Filtering

Random fluctuations or variations in the brightness or colour information in photographs that may be created while the image is being captured are known as noise. It degrades the quality of an image by bringing variations in the original image content. [31]To avoid this noise, a Gaussian filter was applied. Gaussian filtering is a method of correcting spectral coefficients of interest and coefficients within the spectrum and the filter frame. It is based on peak detection, the assumption that peaks represent impulses. This filter has a larger relevance for pixels near the edge, which helps reduce edge blurring. It is computationally efficient, and the degree of smoothing can be adjusted.

#### 3.2.2. Enhancement

Histogram Equalization (HE) is a common picture-enhancing method. Because of its simplicity and significantly greater performance on practically all types of photos, this approach is often used for image improvement [32]. The HE method is commonly used in the processing of medical images and radar images. However, this approach has several drawbacks, and at the same time, it is not that computationally demanding. This approach effectively highlights the boundary and border between objects but obscures local details within the items, particularly smooth ones. As a result, background noise can be faded with this useable signal, resulting in overly enhanced images. Histogram equalization-based methods such as Contrast-Limited Adaptive Histogram Equalization can be used to

boost the contrast. In contrast to conventional histogram equalization, adaptive histogram equalization increases the contrast of the histograms. In this approach, instead of computing the full picture, multiple histograms are computed, each representing a different part of the image known as a tile. The goal is to generate a more evenly distributed pixel value distribution, so the contrast of each tile is raised. To reduce the amplification of any noise in the image, the contrast may be changed to inhomogeneous parts to eliminate any false border lines caused by adjacent tiles. Bilinear interpolation is then employed to blend the nearby tiles. As a result, this procedure is appropriate for enhancing an image's local contrast and bringing forth additional detail. Local contrast is prioritized over total contrast in this strategy. [33]

#### 3.2.3. ROI

There are many ways to filter or process a picture, but in many cases, a Region Of Interest is just a segment of the image. Since the pixels in the ROI have the same size as the pictures to be processed, the binary image has the same size as the images to be processed because all the pixels that define the ROI are set to 1, then extract one or more ROIs from the pictures. The area of interest can be identified by varying intensities. It results in an original image of 1024x1024 pixels cropped into 256x256 pixels while maintaining its location within the image.

### 3.3. Feature extraction

As part of artificial neural networks, autoencoders are artificial neural networks that acquire compressed representations of unsupervised data. The Autoencoder function consists of two parts: a function that calculates the vector of features from input data and a function that calculates the feature vector from input data. The probability models are created and then trained using a specific probability function to optimize data similarity in the decoder portion. Sparse autoencoders use large neural networks with multiple layers that are digitally linked so that the output of each successive layer can be used for dimension reduction or feature extraction.

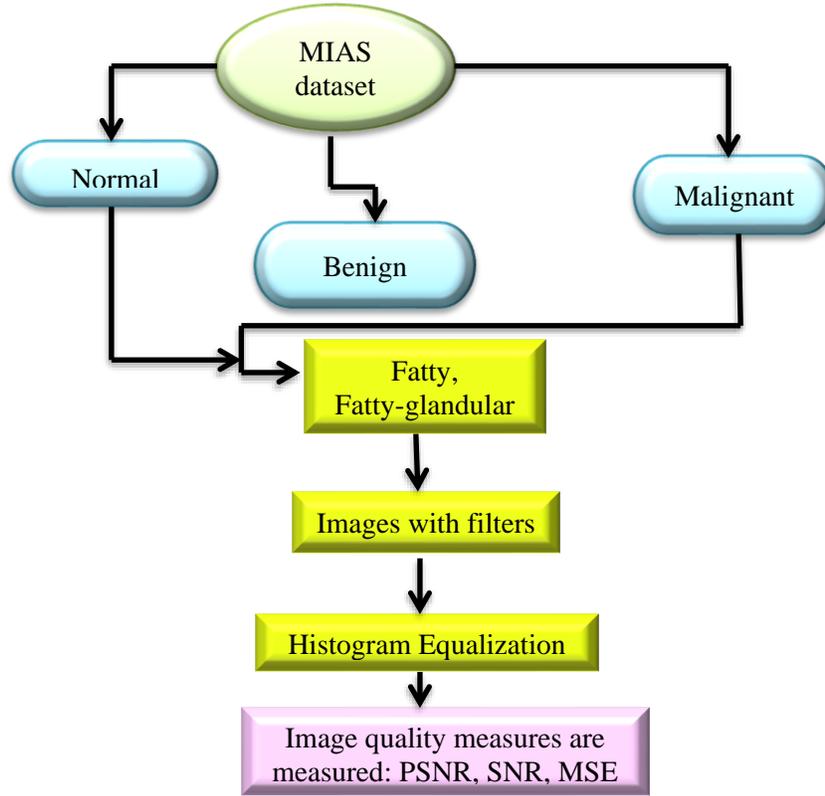


Fig. 4 Preprocessing stage

The proposed sparse autoencoder technique requires raw inputs  $X_j^l X_j^l$ , a hidden layer  $H_m^l H_m^l$ , and an output layer  $Y_n^l Y_n^l$ , where n denotes the input or output neurons, m denotes no hidden neurons, and l represents no. of sparse encoders. The input vector ln is translated into the hidden layer  $H_m^l$  through the output layer using the nonlinear function S[31, 32].

$$H_m^l = s(\sum_{m=1}^n (w_i \times X_j^l) + b_m) \quad (1)$$

Weights for the input-to-hidden layer are given by the parameter  $w_i w_i$  and hidden layer bias is given by  $b_m b_m$ . The sigmoid function is given by  $s(v)$ .

$$s(v) = \frac{1}{1+e^{-v}} \quad (2)$$

There is the same no of units in the output layer  $Y_n^l Y_n^l$  as in the input layer.

$$Y_n^l = s(\sum_{j=0}^m (w_j^l \times H_j^l) + b_n) \quad (3)$$

By categorizing breast cancer, stacked autoencoders are used to demonstrate,  $w_j^l w_j^l$  signifies the weights (or parameters) assigned to the hidden layer and the output layer,  $b_n b_n$  signifies the bias in the output layer, and S signifies the

sigmoidal function.

In figure 5a, the first sparse autoencoder determines the major feature  $H_m^l H_m^l$  by studying the information in the input layer  $X_n^l X_n^l$ . The major feature I is produced by the first sparse autoencoder. As the second sparse autoencoder is learning secondary features (Feature II), the input layer of the second encoder is sent along to the following encoder. Figures 5a, 5b, 5c, and 5d illustrate the main characteristics used as input to the following encoder. The figure shows how three layers of autoencoding are constructed using a softmax classifier once the secondary feature has been mapped to the digit labels. This autoencoder consists of two hidden layers (first and second features), followed by the activation function(softmax). Two hidden layers (the first and second features) and an output layer (the softmax classifier) make up stacked autoencoders. Features extracted after applying auto encoder are mean, variance, energy, kurtosis, contrast, mean deviation, contrast, correlation, coarseness and standard deviation.

### 3.4. Feature Selection

As the study uses many features, the computational cost and the classification accuracy will degrade. It is thus imperative to select features, which is why filtering and wrapping are the two basic feature selection algorithms. The number of characteristics exceeds the number of samples,

which increases the risk of overfitting. Filters and scores are calculated using a statistical measure that utilizes the samples' inherent features. Based on their scores, the features are sorted by highest ranking, with only the highest-ranked elements remaining to be categorized.

Moreover, the selected features may contain duplicate features due to how the scores are calculated for each feature, ignoring their dependence on other features. On the other hand, wrapping approaches select only those features with the greatest discriminating power and analyze the classifier's performance using only these selected features. Radiomic characteristics are better selected using wrapping-based feature selection methods because of their clear association [34,35].

Let  $S = [1, 2, \dots, n]$  show the subset of remaining features of  $F$  denotes the total number of features. As training samples  $X_0 = [X_1, X_2, \dots, X_k]^T$  and class labels  $Y_0 = [Y_1, Y_2, \dots, Y_k]^T$ , the RFE technique picks features by recursively examining smaller and smaller groups of features, as weighted by an external estimator. The SVM classifier is trained in each loop, and a corresponding weight

vector,  $w$ , is calculated. Through  $w$ , characteristics are rated, with the ones with the lowest rank removed. It is repeated until  $s = [ ]$ . As a result of the feature deleting process,  $r$  is obtained, with the later deleted features having a higher score. To choose how many attributes to use, cross-validation can be used. The first step is to isolate the lowest  $k$  features using cross-validation. The next step is to retrieve the highest two attributes using SVM-RFE. Until all characteristics have been removed, the process is repeated until only a few features remain, and a ranking can be assigned to determine the optimal score.

The goal of feature selection is to find an optimal subset of features that can achieve both dimension reduction and accurate prediction while balancing feature quantity and classification accuracy. In practice, feature selection aims to create subsets of features that are more discriminating and dimension-reducing than others. After collecting a list of accuracies and feature importance from RFE, there are two types of alternatives to select the ideal feature subset. The HA or a version similar to HA is the first kind, and the PreNum is the second.

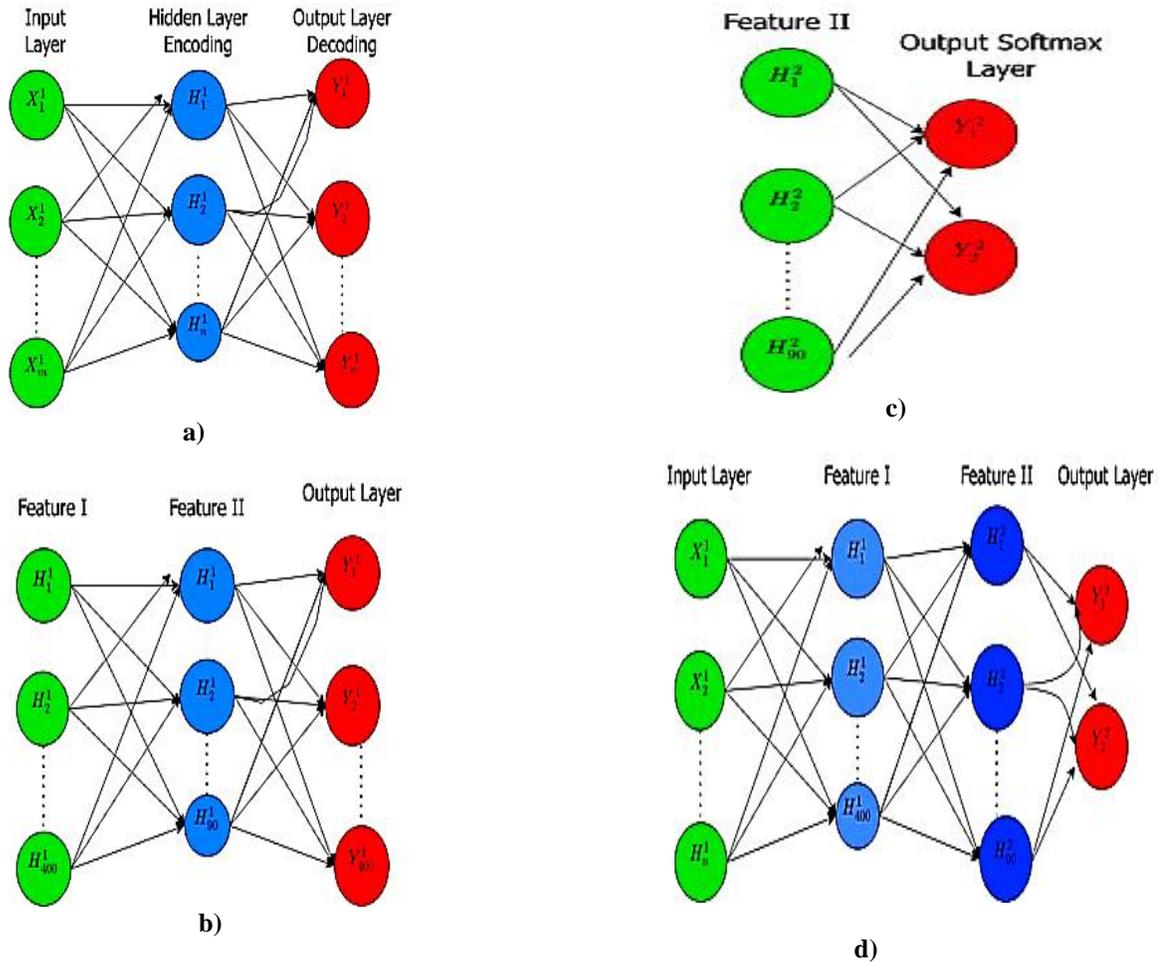


Fig. 5a The first autoencoder, Fig. 5b. The second autoencoder, Fig.5c. Softmax classifier, Figure 5d. Finalized autoencoder.

The ideal feature subset for the HA variant corresponds to the HA (or a specified proportion of the HA); for PreNum, the top-rated PreNum features sorted by relevance are picked as the optimal feature subset. To determine the number of ideal feature subsets, investigating three variants are considered: HA, 90% of HA, and PreNum. The following are the analysis results and comparisons between the three variations. Assuming there are p feature subsets in total, abbreviated as  $f_{sub}$  after RFE, the best feature subset  $F_{sf}$  for each fold is defined as follows:

$$\begin{aligned}
 F_{sf}(HA) &= f_{sub}(HA) \mid Acc(f_{sub}(HA)) = \\
 & \text{argmax } Acc(f_{sub}), \\
 F_{sf}(90\%HA) &= f_{sub}(90\%HA) \mid \\
 Acc(f_{sub}(90\%HA)) &= \text{argmax } 90\% \times \\
 & Acc(f_{sub}), \\
 F_{sf}(\text{PreNum}) &= f_{sub}(\text{PreNum}) \text{ Feature number in} \\
 f_{sub}(\text{PreNum}) &= \text{PreNum} \quad (4)
 \end{aligned}$$

**3.5. Classification**

A new approach to CNN based classification model is created in this section. Table 2 shows the details about it. It consists of three convolutional layers, one FC layer, and three max-pooling layers. Except for the last layer for output, whose activation function is the ReLU function [36], each layer has the convolutional layer as its activation function. This layer is symbolized by Conv 3-32, which indicates that it contains 32 convolutional neurons (units) with a filter size of 3x3 pixels (height x width). FC 64 refers to a layer with a 64-unit fully-connected structure. MaxPool 2 is a max-pooling layer with a 2x2 pixel window and stride 2. In the Maxpool layer, after each update during training, a softmax function is applied to transfer the output value to the [0, 1, 2] range. It should prevent overfitting. In the output layer, a fractional input rate is set to 0 for the following layer at random [37].

A CNN has been created using 13 convolutional layers based on the VGG-16, and a basic FC layer was pre-trained (Table 3).

Table 3 displays the weights from the convolutional blocks in the VGG-16 pre-trained model. The weights in the convolutional blocks used in CNN training did not change (or were "frozen"). Weights at the FC layer were varied based on the inputs throughout training. VGG-16 identifies features in input images by using FC NN-classifiers and VGG-16 to extract them.

**3.5.1. Fine Tuning**

The four convolutional blocks were imported from a VGG-16 model and frozen. Unlike the fine-tuning CNN structure in Table 3, these blocks in the pre-trained model do not have fixed weights. The FC weights were drawn from the

early training procedure for feature extraction and used to fill final convolutional blocks instead of random initialization. As a result, no weights were initially randomized during fine-tuning. Figure 6 depicts output instances obtained after the fine-tuning.

**3.5.2. Hyper Parameter Tuning**

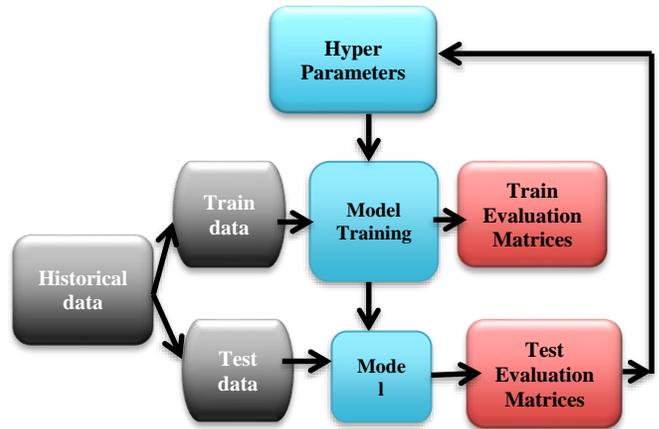
Our proposed method uses hyperparameters to manage the model's over-fitting and under-fitting. Different datasets have different optimal hyperparameters. In our method, the following stages are taken to obtain the best hyperparameters:

1. The model is assessed for each proposed hyperparameter configuration.
2. The optimum hyperparameters for the model are chosen.

**Table 2. CNN Parameters**

Input image
Conv_3-32 + RELU
Conv_3-64 + RELU
Maxpool 2
Maxpool 3
Softmax output : [0,1,2]
FC_64 + ReLU (with 0.5 dropout)

The hyperparameter is evaluated by tuning a few hyperparameter settings, evaluating the validation matrices, adjusting the hyperparameters, and evaluating the validation matrices until they reach optimum results. Figure 7 gives a diagrammatic representation of the evaluation of the hyperparameter model.



**Fig. 7 Diagrammatic representation of evolution hyperparameteric model.**

Table 3. Transfer Learning

<b>VGG16</b>	<b>Input Image</b>	
		<b>Conv_3-64 + ReLu</b>
	<b>Conv_block 1</b>	<b>Conv_3-64 + ReLu</b>
		<b>Maxpool 2</b>
		<b>Conv_3-128 + ReLu</b>
	<b>Conv_block 2</b>	<b>Conv_3-128 + ReLu</b>
		<b>Maxpool 2</b>
		<b>Conv_3-256 + ReLu</b>
	<b>Conv_block 3</b>	<b>Conv_3-256 + ReLu</b>

		<b>Maxpool 2</b>
		<b>Conv_3-512 + ReLu</b>
	<b>Conv_block 4</b>	<b>Conv_3-512 + ReLu</b>
		<b>Maxpool 2</b>
		<b>Conv_3-512 + ReLu</b>
	<b>Conv_block 5</b>	<b>Conv_3-512 + ReLu</b>
		<b>Maxpool 2</b>
	<b>FC_256 + ReLU (with Dropout = 0.5)</b>	
	<b>Softmax output: [0,1,2]</b>	



Fig. 6. Output Instances; a) Benign, b) Malignant and c) Normal

Table 4. Overall Analysis under Accuracy, Sensitivity, Specificity based on real-world images and MIAS dataset

Models	Accuracy	Sensitivity	Specificity	Image Used
VGG19	0.85	0.90	0.94	
ANN	0.81	0.86	0.90	Real World Image
Alexnet	0.86	0.91	0.95	
Inception v3	0.89	0.92	0.96	
Resnet50	0.81	0.86	0.90	
CNN - VGG16	0.96	0.97	0.98	
VGG19	0.84	0.89	0.93	
ANN	0.80	0.85	0.89	MIAS image
Alexnet	0.85	0.90	0.94	
Inception v3	0.87	0.91	0.95	
Resnet50	0.80	0.85	0.89	
CNN-VGG16	0.95	0.96	0.9	

Table 5. Overall Analysis under precision, recall and F1-score in terms of real-world images and MIAS dataset

Models	Precision	Recall	F1-score	Images Used
VGG19	0.84	0.8	0.85	
ANN	0.84	0.78	0.81	
Alexnet	0.86	0.8	0.84	Real World Image
Inception v3	0.9	0.82	0.87	
Resnet50	0.83	0.77	0.82	
CNN - VGG16	0.95	0.83	0.92	
VGG19	0.83	0.79	0.84	
ANN	0.83	0.77	0.80	
Alexnet	0.85	0.79	0.83	MIAS Image
Inception v3	0.89	0.81	0.86	
Resnet50	0.82	0.76	0.81	
CNN - VGG16	0.94	0.82	0.91	

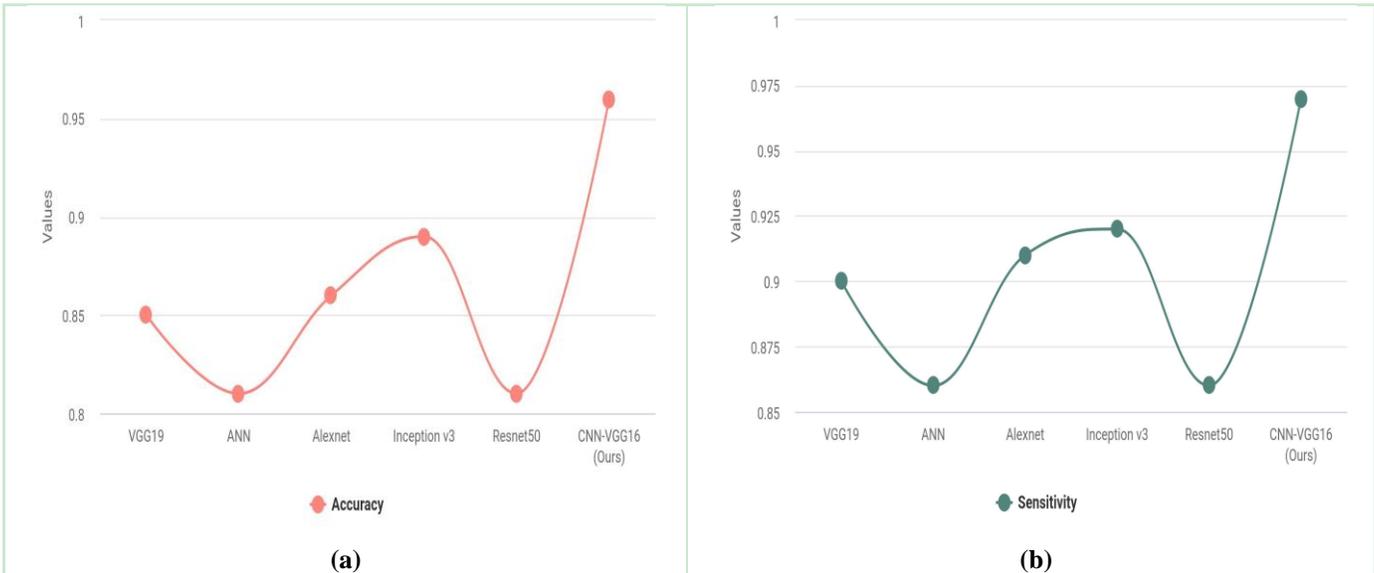
### 4. Performance Analysis

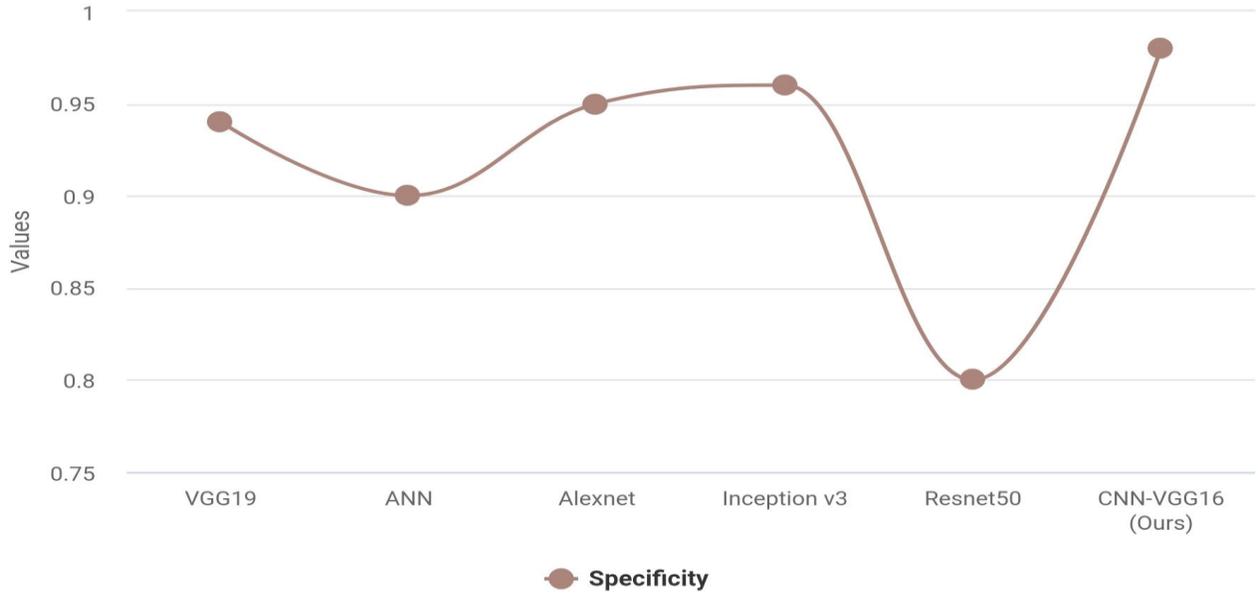
The proposed model uses hardware specifications like an Intel i5/i7 or Ryzen 5/7 series CPU, 12 GB RAM, 1TB HDD, and Windows 10 OS, as well as software such as PyTorch, an open-source python library-based programming language for deep learning frameworks, and Google Collaboratory, an open-source Google environment for developing deep learning frameworks. To train the model, a learning rate of 0.09 and a set of 10 epochs using an Adaboost hyperparameter with standard parameter value  $\mu=3$  have been considered. The proposed model combines VGG16 and CNN, where VGG16 extract certain features, while CNN works on this base model to learn and gain the knowledge to perform over Mammographic images. Since CNN is a generic model and training will take much time and transfer learning will cut short those processes, also using transfer learning will improve 20% performance than custom made model. This model is compared with ANN, VGG19, Alexnet, Inception V3, and Resnet50 in terms of accuracy, sensitivity, specificity, recall, precision, F1 score, detection rate, TPR, and FPR, AUC score, and calculate the AUC value. The value obtained after the AUC calculation is 0.823. The chart of the over-analysis models' accuracy, sensitivity and specificity is shown in Figures 8a, 8b, and 8c. The proposed model outperforms the other models (accuracy, sensitivity, and specificity: 0.96, .97and 0.98, respectively) in real-world images.

Table 5 shows the real-world database's overall study results for accuracy, recall, and F1 score. Figures 9a, 9b, and 9c show a graphical depiction of multiple models, showing that the suggested model outperforms others (precision:0.95, recall:0.83, F1-score:0.92). Other models have a generic framework for conducting classification and, as a result, require a boost in information transmission throughout the network to get even better outcomes.

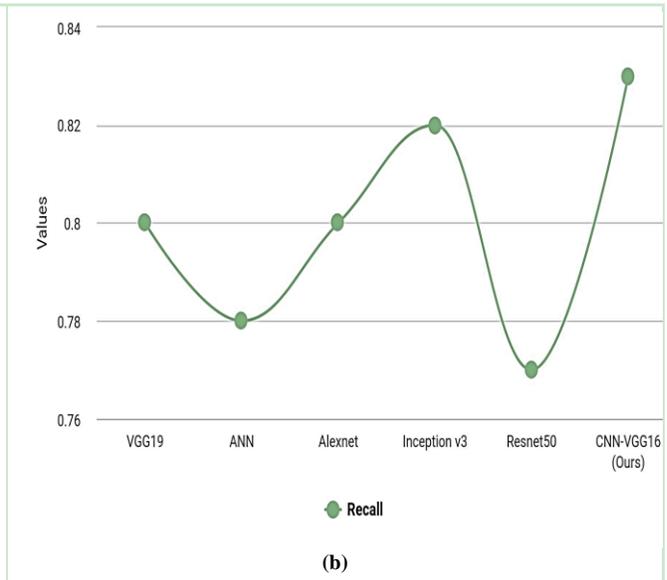
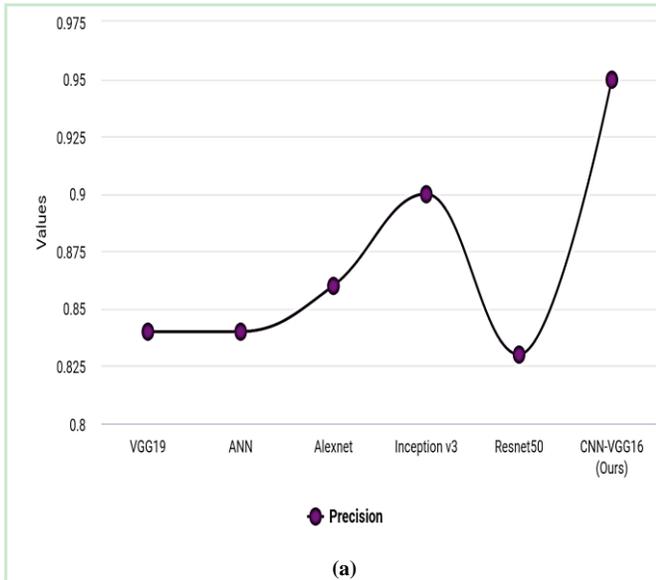
Table 6 summarises the overall Analysis of several models based on AUC and computation time in real-world images. Figure 10a and 10b show a graphical depiction of several models compared to the suggested technique, with the proposed method gaining (0.9) on the AUC score. Even while models like VGG19 and Inception V3 have a complicated structure, they achieve close ranges of 0.84 and 0.87. Regarding computing time, the suggested model has a lower CT (4.1) obtained during training.

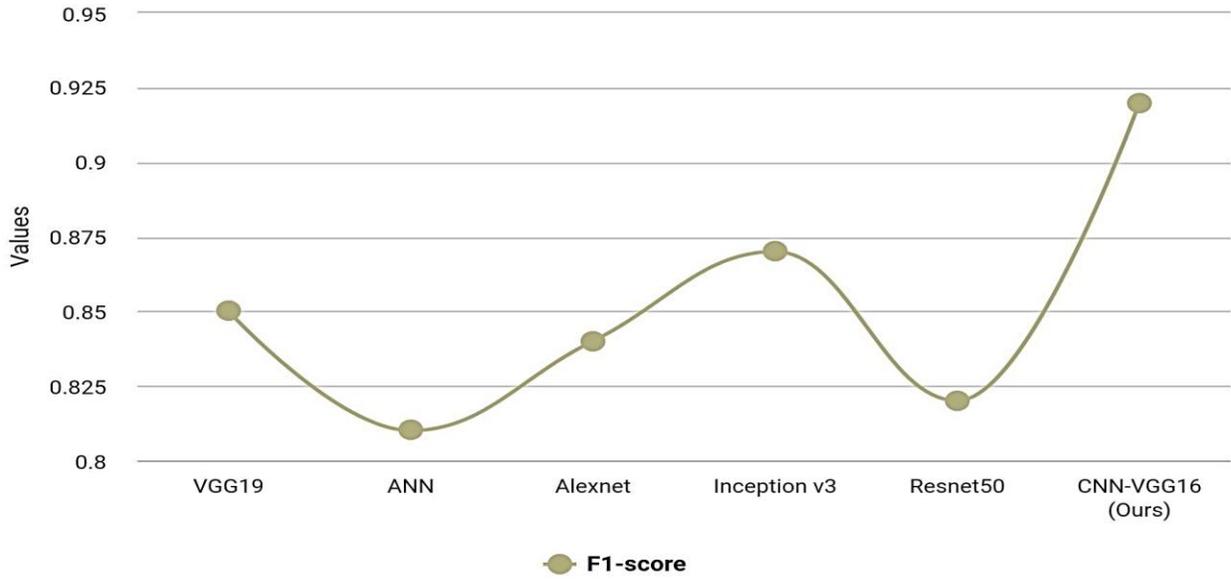
The entire Analysis of models under detection rate, TPR, and FPR in real-world images are shown in Table 7. Figures 11a, 11b, and 11c show graphical depiction of several models compared to the suggested technique, demonstrating that the proposed model outperforms others (detection rate: 0.94, TPR: 0.95, FPR: 0.5). Figures 12a, 12b, 12c, and 12d show graphical depiction of accuracy vs epoch, in which, as the epoch grows, accuracy fluctuates at particular epoch range before stabilizing. Figure 13 shows the final classifier result that has been categorized.





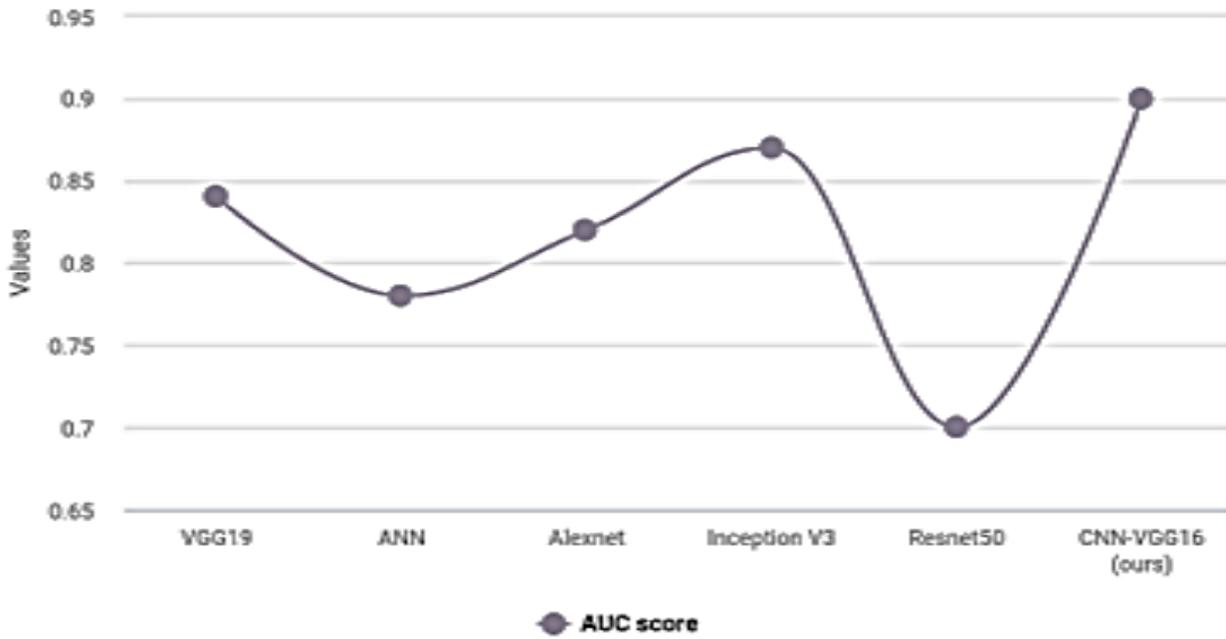
(c)  
**Fig. 8a Models vs Accuracy, 8b. Models vs sensitivity and 8c. Models vs Specificity**



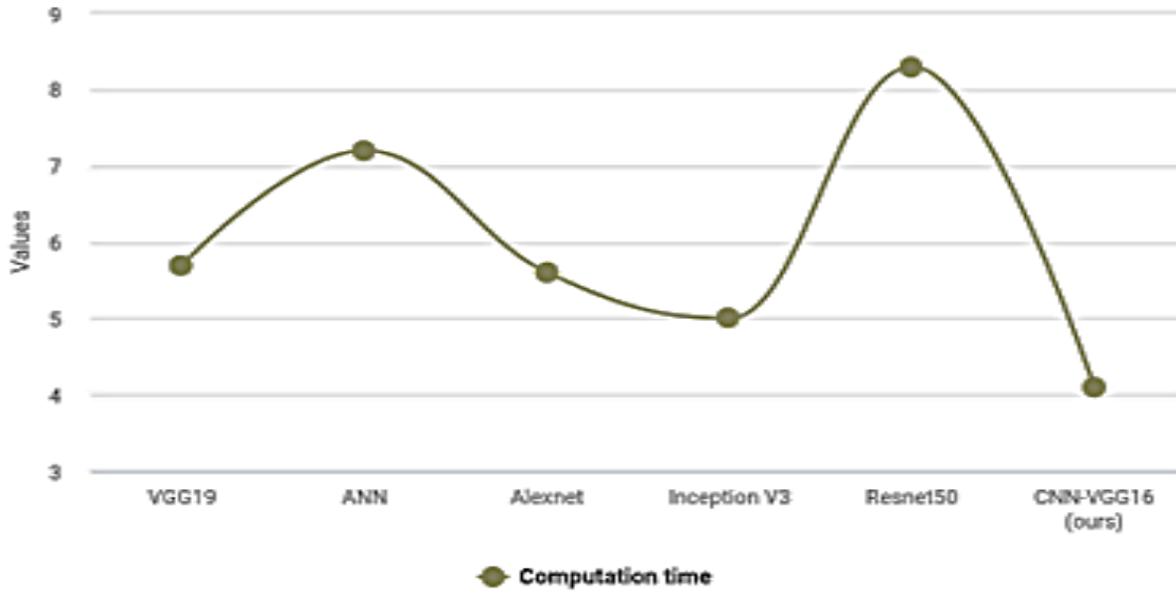


(e)

Fig. 9a Models vs Precision, 9b. Models vs Recall 9c. Models vs F1-score

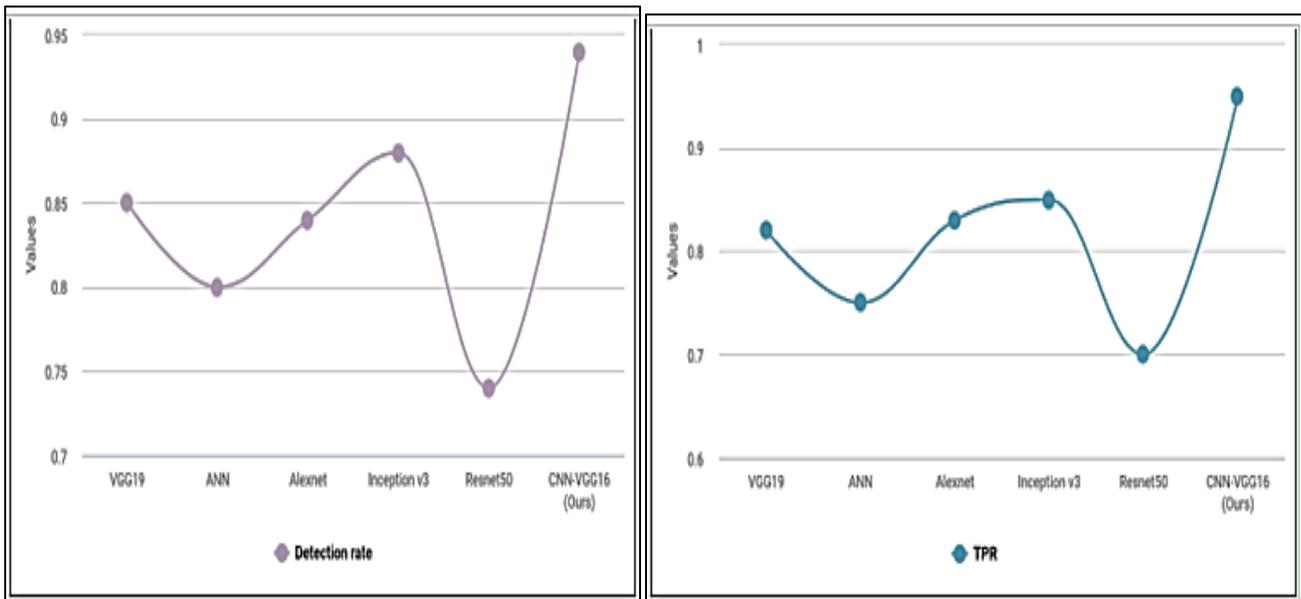


(a)



(b)

Fig. 10a Models vs AUC score, 10b. Models vs Computation time



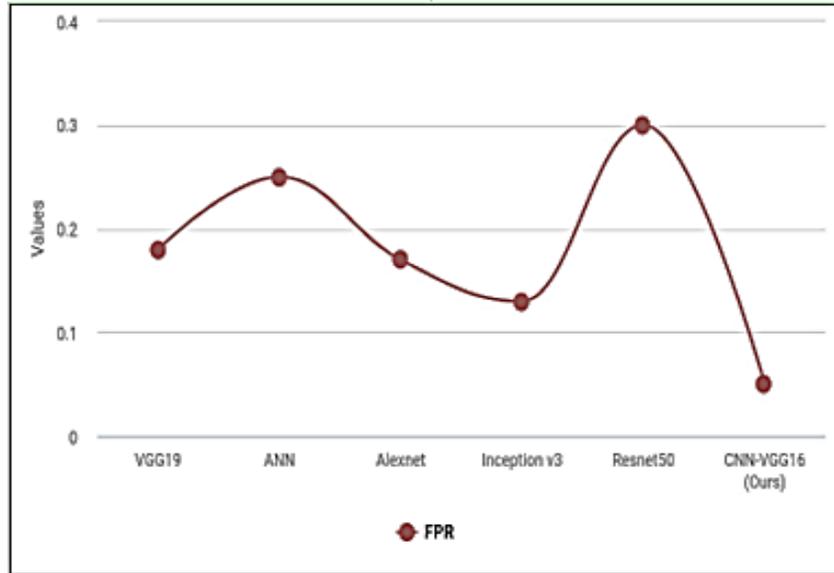


Fig. 11a Models vs Detection rate, 11b. Models vs TPR and 11c. Models vs FPR

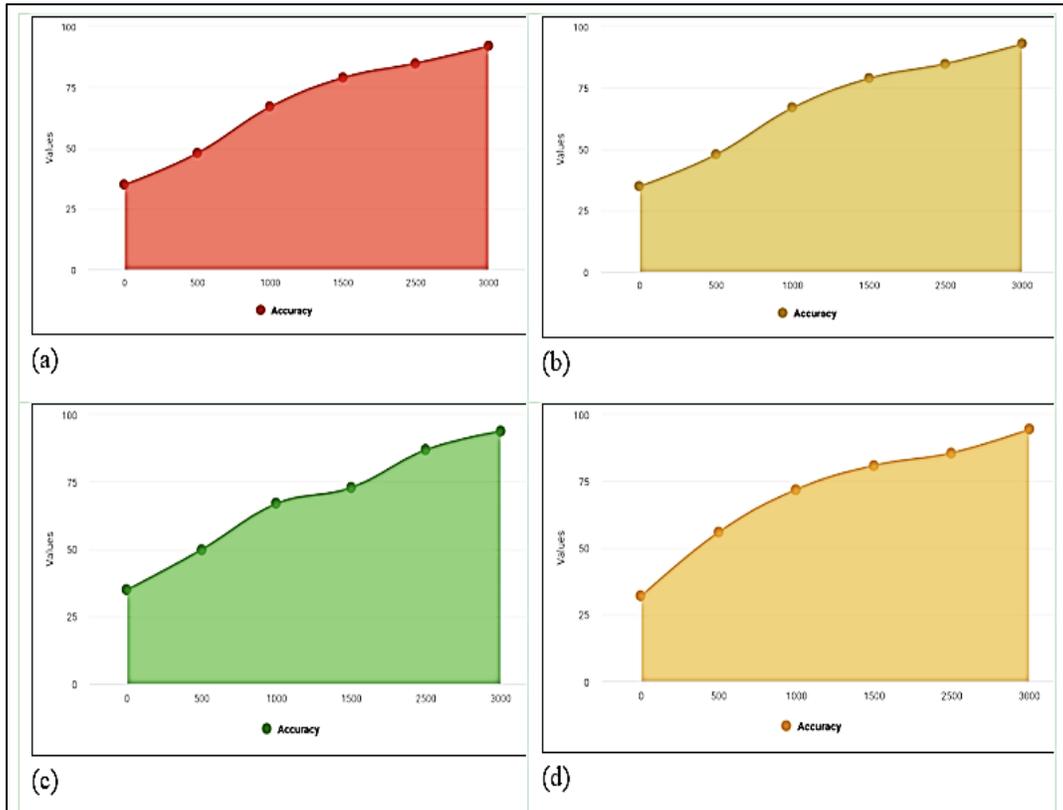


Fig. 12a, 12b, 12c, and 12d. Accuracy vs Epochs over various training periods and its respective percentage-wise variations over various image instances.

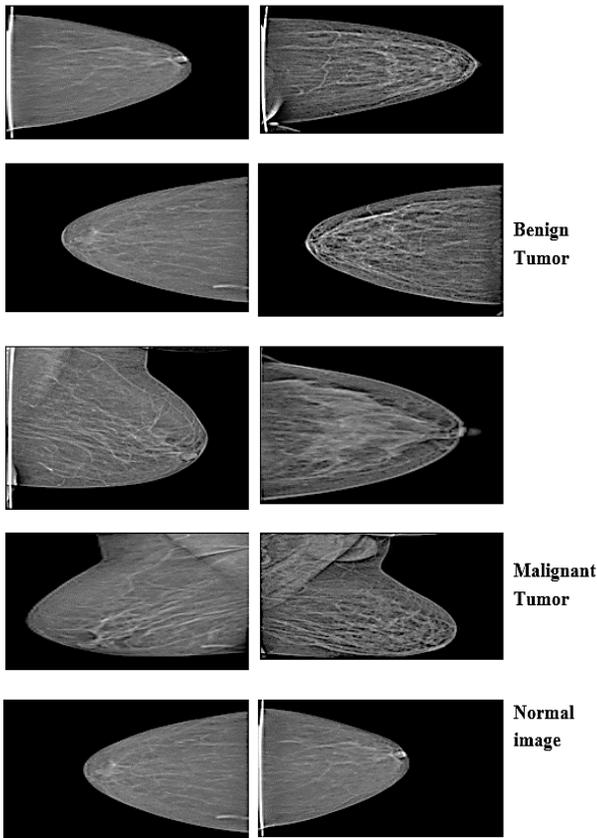


Fig. 13. Final classifier performance for benign and malignant tumors by using a Convolutional Neural Network

Table 6. Overall Analysis under AUC score and CT

Models	AUC score	Computation time	Images used
VGG19	0.84	5.7	
ANN	0.78	7.2	Real World Images
Alexnet	0.82	5.6	
Inception V3	0.87	5	
Resnet50	0.7	8.3	
CNN-VGG16	0.9	4.1	
VGG19	0.83	5.6	
ANN	0.77	7.1	MIAS Images
Alexnet	0.81	5.5	
Inception V3	0.86	4.9	
Resnet50	0.69	8.2	
CNN-VGG16	0.89	4.0	

Table 7. Overall Analysis under detection rate, TPR and FPR

Models	Detection rate	TPR	FP R	Image Used
VGG19	0.85	0.82	0.18	
ANN	0.8	0.75	0.25	Real World Images
Alexnet	0.85	0.83	0.17	
Inception v3	0.88	0.87	0.13	
Resnet50	0.78	0.7	0.3	
CNN - VGG16	0.94	0.95	0.5	
VGG19	0.84	0.81	0.17	
ANN	0.79	0.74	0.24	MIAS Images
Alexnet	0.84	0.82	0.16	
Inception v3	0.87	0.86	0.12	
Resnet50	0.77	0.69	0.29	
CNN - VGG16	0.93	0.94	0.4	

### 5. Conclusion

Researchers have studied the segmentation and classification of breast cancer using a range of imaging techniques. Breast cancer is a very common cancer among women throughout the world. This work presents an efficient hyperparameter tuning and optimization-based strategy for diagnosing breast cancer using mammographic images. And to increase network knowledge, a CNN-based transfer learning approach is also performed. Popular datasets such as MIAS and real-world datasets from VPS Lakeshore Hospital have been utilized in this proposed system. The experiments described here demonstrate that the suggested system performs well. Finally, research professionals will find this study extremely useful in analyzing the best strategies for feature extraction and selection that can be utilized to improve the network even more, as well as attempting to combine with other sophisticated techniques for even better outcomes.

### 6. Acknowledgement

Expressing our gratitude to the famous oncologist Dr.V P Gangadharan, Dr.Thara Prathap and her team, and the Management and staff of the VPS Lakeshore hospital for providing the valuable real-world dataset.

## References

- [1] J Arevalo, Gonz alez fa, Ramos-poll an r, Oliveira jl, Guevara lopez ma, "Representation Learning for Mammography Mass Lesion Classification with Convolutional Neural Networks," *Comput Methods Programs Biomed*, vol. 127, pp. 248257, 2016.
- [2] Benjamin Q Huynh, Hui Li, and Maryellen L Giger, "Digital Mammographic Tumor Classification using Transfer Learning from Deep Convolutional Neural Networks," *Journal of Medical Imaging*, vol. 3, no. 3, pp. 034501, 2016.
- [3] Yuan-Pin Lin and Tzyy-Ping Jung, "Improving EEG-based Emotion Classification using Conditional Transfer Learning," *Frontiers in Human Neuroscience*, vol. 11, no. 334, 2017.
- [4] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich, "Going Deeper with Convolutions," In Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-9, 2015.
- [5] Neil J Vickers, "Animal Communication: When I'm Calling You, Will You Answer Too?," *Current Biology*, vol. 27, no. 14, pp. R713R715, 2017.
- [6] Hiroki Tanaka, Shih-Wei Chiu, Takanori Watanabe, Setsuko Kaoku, and Takuhiro Yamaguchi, "Computer-Aided Diagnosis System for Breast Ultrasound Images using Deep Learning," *Physics in Medicine & Biology*, vol. 64, no. 23, pp. 235013, 2019.
- [7] Walid Al-Dhabyani, Mohammed Gomaa, Hussien Khaled, and Fahmy Aly, "Deep Learning Approaches for Data Augmentation and Classification of Breast Masses Using Ultrasound Images," *Int. J. Adv. Comput. Sci. Appl*, vol. 10, no. 5, pp. 111, 2019.
- [8] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna, "Rethinking the Inception Architecture for Computer Vision," In Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818-2826, 2016.
- [9] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, and Quoc V Le, "Learning Transferable Architectures for Scalable Image Recognition," In Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8697-8710, 2018.
- [10] Ting Xiao, Lei Liu, Kai Li, Wenjian Qin, Shaode Yu, and Zhicheng Li, "Comparison of Transferred Deep Neural Networks in Ultrasonic Breast Mass Discrimination," *Biomed Research International*, 2018.
- [11] Heqing Zhang, Lin Han, Ke Chen, Yulan Peng, and Jiangli Lin, "Diagnostic Efficiency of the Breast Ultrasound Computer-Aided Prediction Model Based on Convolutional Neural Network in Breast Cancer," *Journal of Digital Imaging*, vol. 33, pp. 1218-1223, 2020.
- [12] Abdullah-Al Nahid and Yinan Kong. "Histopathological Breast-Image Classification using Local and Frequency Domains by a Convolutional Neural Network," *Information*, vol. 9, no. 1, pp. 19, 2018.
- [13] Abdullah-Al Nahid, Mohamad Ali Mehrabi, and Yinan Kong, "Histopathological Breast Cancer Image Classification by Deep Neural Network Techniques Guided by Local Clustering," *Biomed Research International*, 2018.
- [14] Mehedi Masud, Amr E Eldin Rashed, and M Shamim Hossain, "Convolutional Neural Network-Based Models for Diagnosis of Breast Cancer," *Neural Computing and Applications*, pp. 1-12, 2020.
- [15] Barath Narayanan Narayanan, Vignesh Krishnaraja, and Redha Ali, "Convolutional Neural Network for Classification of Histopathology Images for Breast Cancer Detection," In 2019 *IEEE National Aerospace and Electronics Conference NAECON*, pp. 291-295, 2019.
- [16] L. Tsochatzidis, P. Koutla, L. Costaridou, and I. Pratikakis, "Integrating Segmentation Information into CNN for Breast Cancer Diagnosis of Mammographic Masses," *Computer Methods and Programs in Biomedicine*, vol. 200, no. 105913, 2021.
- [17] Leung, J., Martin, J. and McLaughlin, D, "Rural-Urban Disparities in the Stage of Breast Cancer at Diagnosis in Australian women," *The Australian Journal of Rural Health*, 2016. DOI: 10.1111/ajr.12271.
- [18] Khan, S., Islam, N., Jan, Z., Din, I. U., & Rodrigues, J. J. C, "A Novel Deep Learning Based Framework for the Detection and Classification of Breast Cancer using Transfer Learning," *Pattern Recognition Letters*, vol. 125, pp. 1-6, 2019.
- [19] Ayana, G., Dese, K., & Choe, S. W, "Transfer Learning in Breast Cancer Diagnoses via Ultrasound Imaging," *Cancers*, vol. 13, no. 4, pp. 738, 2021.
- [20] Vesal, S., Ravikumar, N., Davari, A., Ellmann, S., & Maier, A, "Classification of Breast Cancer Histology Images using Transfer Learning," In *International Conference Image Analysis and Recognition*, pp. 812-819, 2018.
- [21] Chang, J., Yu, J., Han, T., Chang, H. J., & Park, E, "A Method for Classifying Medical Images using Transfer Learning: A Pilot Study on Histopathology of Breast Cancer," In 2017 *IEEE 19th International Conference on e-Health Networking, Applications and Services, Healthcom*, pp. 1-4, 2017.
- [22] Saber, A., Sakr, M., Abo-Seida, O. M., Keshk, A., & Chen, H. "A Novel Deep-Learning Model for Automatic Detection and Classification of Breast Cancer Using the Transfer-Learning Technique," *IEEE Access*, vol. 9, pp. 71194-71209, 2021.
- [23] Khuriwal, N., & Mishra N, "Breast Cancer Detection from Histopathological Images Using Deep Learning," In 2018 3rd *International Conference and Workshops on Recent Advances and Innovations in Engineering, ICRAIE*, pp. 1-4, 2018.
- [24] Makhtar, M., Rosly, R., Awang, M. K., Mohamad, M., & Zakaria, A. H., "A Multi-Classifer Method based Deep Learning Approach for Breast Cancer," *Int. J. Eng. Trends Technol*, vol. 1, pp. 102-107, 2020.
- [25] Purwanti E, & Apsari R, "Classification of Digital Mammograms Using Nearest Neighbor Techniques."

- [26] Kanmani P, Marikkannu P & Brindha M, "A Medical Image Classification using Id3 Classifier."
- [27] Charan, S., Khan, M. J., & Khurshid, K, "Breast Cancer Detection in Mammograms Using Convolutional Neural Network," In 2018 *International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, pp. 1-5, 2018.
- [28] Kwon S., Lee H., & Lee S, "Image Enhancement with Gaussian Filtering in the Time-Domain Microwave Imaging System for Breast Cancer Detection," *Electronics Letters*, vol. 52, no. 5, pp. 342-344, 2016.
- [29] Kaur M., Kaur J. and Kaur J, "Survey of Contrast Enhancement Techniques based on Histogram Equalization," *IJASA*, vol. 2, no. 7, pp. 137-141, 2011.
- [30] Ritika, Kaur S, "Contrast Enhancement Techniques for Images a Visual Analysis," *International Journal of Computer Applications*, pp. 64, no. 17, 2013.
- [31] M. Loey, A. El-Sawy, and H. El-Bakry, "Deep Learning Autoencoder Approach for Handwritten Arabic Digits Recognition," arXiv preprint arXiv:1706.06720, 2017
- [32] Feng Y, Zhang, L, & Mo J, "Deep Manifold Preserving Autoencoder for Classifying Breast Cancer Histopathological Images," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 17, no. 1, pp. 91-101, 2018.
- [33] Chen G, Xie X, & Li S, "Research on Complex Classification Algorithm of Breast Cancer Chip Based on SVM-RFE Gene Feature Screening," *Complexity*, 2020.
- [34] Abdulkareem S. A, & Abdulkareem Z. O, "An Evaluation of the Wisconsin Breast Cancer Dataset Using Ensemble Classifiers and RFE Feature Selection," *Int. J. Sci., Basic Appl. Res.*, vol. 55, no. 2, pp. 67-80, 2021.
- [35] V Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," in Proceedings of the 27th *International Conference on Machine Learning ICML-10*, pp. 807-814, 2010.
- [36] N Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929-1958, 2014.
- [37] Wahab, N., Khan, A., & Lee, Y. S, "Transfer Learning Based Deep CNN for Segmentation and Detection of Mitoses in Breast Cancer Histopathological Images," *Microscopy*, vol. 68, no. 3, pp. 216-233, 2019.