ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

Identification of Glaucoma from Retinal Fundus Images using Deep Learning Model, MobileNet

Smita Das[1], Madhusudhan Mishra[2] and Swanirbhar Majumder[3]

ABSTRACT

Glaucoma is one of the leading causes of permanent vision impairment and blindness everywhere in the world due to high intraocular stress inside the eyes. So, early and accurate detection is crucial for preventing irreversible vision loss. Manual recognition of Glaucoma is a difficult task that requires proficiency and a highly experienced person. Computer Aided Detection (CAD) techniques assist ophthalmologists in the detection of such ophthalmologic diseases by analyzing the retinal fundus images. Deep learning(DL) algorithms have accomplished exceptionally in various computer vision applications by giving them an apparent advantage over more traditional techniques by analyzing the retinal fundus images. In this paper, a pre-trained Convolutional Neural Network, MobileNet, is used to detect Glaucoma in retinal fundus images by automatically extracting the features rather than manually classifying fundus images. Due to efficiency, speed of execution, accuracy, customizability, transfer learning, real-time applications and size of the architecture, MobileNet is proposed here. The DrishtiGS, EyePACS AIROGS-Light, BEH, REFUGE, sjchoi86-HRF, CRFO-v4, G1020, FIVES, and PAPILA datasets were utilized in this study. The suggested model is evaluated by significant factors such as accuracy, precision, sensitivity, specificity, F1 score, and Confusion Matrix to examine the model's efficacy. These studies reveal the capability of the DL approach to classify Glaucoma from retinal fundus images and recommend that the suggested approach can assist ophthalmologists in a quick, correct, and dependable diagnosis of Glaucoma.

Article information:

Keywords: Glaucoma, Fundus Image, Deep Learning, Convolutional Neural Network, MobileNet

Article history:
Received: March 24, 2024
Revised: May 30, 2024
Accepted: July 11, 2024
Published: July 20, 2024
(Online)

DOI: 10.37936/ecti-cit.2024183.256182

1. INTRODUCTION

Glaucoma is one of the neurodegenerative eye diseases; timely diagnosis may prevent vision loss, but last phase identification may lead to irreversible vision loss. It harms the optic nerve head because of extreme intra-ocular damage in the eye and results in permanent vision loss [1]. Glaucoma shows no sign in its early phase. For successful treatment, timely detection is necessary; otherwise, it leads to irreversible vision loss. The most crucial approach to fight the development of this syndrome is timely diagnosis and treatment. Glaucoma detection percentage in developed nations is very poor (less than 50%). By 2040, it is predicted that approximately 116 million people

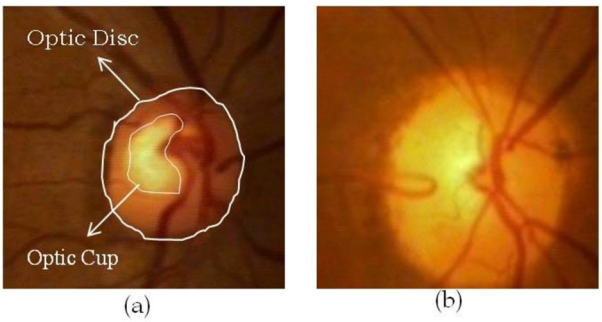

Fig. 1: *(a) Healthy image (b) Glaucoma-affected image.*

[1,3]The authors are with the Department of Information Technology, Tripura University, Suryamaninagar, Tripura, India, E-mail: ersmitadas@gmail.com and swanirbhar@ieee.org
[2]The author is with the Department of Electronics and Communication Engineering, NERIST (Deemed to be University), India, E-mail: ecmadhusudhan@gmail.com
[1]Corresponding author: ersmitadas@gmail.com

will be affected by Glaucoma globally [2]. Recognition and observation of glaucoma-effected optic nerves depend on various clinical features that are monitored and evaluated prior to making a conclusion. Presently, glaucoma detection and observation need an entire eye test and a few extra tests, which may be difficult to understand. Moreover, there may be a considerable overlap in the ocular characteristics of healthy and glaucomatous eyes [3]. Figure 1 shows the healthy image and Glaucoma-affected image.

DL model, especially convolutional neural network (CNN), helps ophthalmologists by automatically analyzing retinal fundus images. DL models overcome diagnosis challenges by maintaining a limited false positive rate [4]. DL models identify patterns in the images that are not visible to ophthalmologists. DL models incorporate methods to focus on the area of the fundus image, which has already been analyzed by ophthalmologists for glaucoma detection. The objective of the analysis is to increase categorization accurateness by automatically categorizing images as healthy or glaucomatous. By using computer-assisted techniques for retinal fundus examination, the repeatability and impartiality of fundus images should be improved. There should be a precise explanation of optic nerve representation to correct identification and to set a standard level from where transformation can be exposed through assessment by the ophthalmologist. Quantitative classification of the optic nerve using the Cup-to-Disc Ratio (CDR) is the basic process of diagnosing Glaucoma. CDR is characterized by damage to the optic disc by Glaucoma in an unpredictable way [5]. Till now, there is no proper treatment or cure for Glaucoma, so ophthalmologists should rely on timely and accurate diagnosis to keep the disease under control [6]. Numerous researchers have gone through this research on glaucoma diagnosis by utilizing various public datasets. Nearly every research is based on the feature extraction from retinal fundus images and by using various classifiers which automatically calculate the CDR [7]. Therefore, continuous evaluation and upgradation are necessary to increase the accuracy of the classification model. Pretrained Convolutional Neural Network(CNN) variant MobileNet is proposed to detect Glaucoma in retinal fundus images that classify fundus images by automatically extracting the features. To evaluate the effectiveness of the model, the DrishtiGS, EyePACS AIROGS-Light, BEH, REFUGE, sjchoi86-HRF, CRFO-v4, G1020, FIVES, and PAPILA datasets were utilized in this study. The performance of different techniques is examined by different performance metrics like Sensitivity, Specificity, F1 Score, Precision, Accuracy, etc. So, this work will provide us with a comprehensive overview of the MobileNet model for the recognition of Glaucoma and a short illustration of quantitative performance metrics like sensitivity, Specificity accuracy, etc. This paper is ordered such that in Section 2, we provide a description of associated mechanisms; Section 3 represents the proposed technique; whereas Section 4 describes results and explanation, and Section 5 provides a conclusion.

2. RELATED WORKS

Automatic diagnosis of Glaucoma from retinal fundus pictures utilizing wavelet-based denoising and Machine-Learning(ML) approaches was proposed by Khan *et al.*[1]. A 13-layer CNN is proposed by Ajitha *et al.*[2] for diagnosis of Glaucoma from fundus images. CNN-softmax and CNN-SVM(Support Vector Machine) classifiers were utilized for the classification. Akter *et al.* [3] proposed a multi-feature examination and a DL approach for Glaucoma detection by analyzing OCT (optical coherence tomography) images. Hemelings *et al.*[4] applied DL approach on fundus images for diagnosis of Glaucoma. Three pretrained CNNs, the residual network, the GoogLeNet, and the visual geometry group network, were implemented by Joshi *et al.* [5] for the categorization of Glaucoma by analyzing fundus pictures. The analysis of anatomical features of segmented optic cups and discs is utilized to diagnose Glaucoma and its different states. Kashyap et al. [6] proposed Improved U-Net DL approaches for Glaucoma detection by analyzing fundus images. A new computer-aided glaucoma recognition by utilizing CVMD (compact variational mode decomposition) from retinal fundus pictures is depicted by Kirar *et al.* [7]. A review work on automatic detection of Glaucoma by fundus image analysis using artificial intelligence was introduced by Coan *et al.* [8]. Harjoseputra *et al.*[9] proposed an Efficient CNN approach for the detection of Protected Birds by utilizing MobileNets. Howard *et al.*[10] proposed Efficient CNN, MobileNet using width multiplier and resolution multiplier for Mobile Vision Applications. Jia *et al.*[11] developed a novel rice pest and disease identification approach by utilizing an improved YOLOv7 approach. Khasoggi *et al.*[12] developed an Efficient mobilenet model for image analysis on mobile and embedded devices. Michele *et al.*[13] depicted MobileNetV2 CNN and SVM for Palmprint analysis and recognition. Sinha *et al.*[14] proposed an enhanced hybrid MobileNet model for the classification. Zhu *et al.* [15] developed an advanced MobileNet approach with Global Average Pooling and Wavelet Energy for Rotational Machineries Fault finding. Patil *et al.*[16] proposed DenseNet169, MobileNet, InceptionV3, Xception, ResNet152V2 and VGG19 for identification of glaucoma from fundus images. Vishwanatha *et al.*[17] proposed ResNet50, VGG16, Xception, ResNET-101, InceptionV3, MobileNetV2 and efficientNetB7 model for glaucoma detection. Making a case for the requirement of further research using the concepts of existing literature, we prepare new, fresh, and original research

ideas. Therefore, the main purpose of our work is to evaluate how well CNN models for Glaucoma detection might help ophthalmologists perform less stressfully and efficiently. The proposed approach was designed to help single models improve their shortcomings and minimize defects by automatically extracting features. The proposed model focuses more on the accuracy of the performance. Using the feature extraction layer of MobileNet-V3 Large, we can achieve better results without additional transfer learning or training. Random selection of fundus images ensures the efficiency of the proposed model to be independent of different variations in images.

CNN-based Deep learning model achieves significant results in binary classification. Due to the powerful characteristics of DL models, the model is also well-known in the field of medical image processing. These models achieved significant outcomes in the diagnosis of Glaucoma, but they still have some drawbacks, which are detailed below:

i) Due to a lack of interpretability, the CNN model sometimes makes wrong predictions.

ii) Prediction of the CNN model depends on training data. If training data is noisy, then the CNN model performs poorly.

iii) The CNN model faces the problem of overfitting, where the model predicts training data too well but makes wrong predictions for unseen data.

iv) In most cases, the CNN model requires powerful GPUs and large-size memory, which are expensive.

v) Sometimes, the CNN model fails to predict complicated structures.

So, the goal of this study is to overcome the issues of basic CNN models. We proposed a MobileNet-V3 model for accurate classification of fundus images to detect Glaucoma. MobileNet Model efficiently classifies fundus images with high accuracy. The MobileNet model helps minimize overfitting issues while primarily targeting efficiency. It helps single techniques remove their demerits and decrease defects by automatically extracting the features. This approach minimizes the prediction variation and generalization error. Data augmentation is employed to overcome overfitting issues. In the first step, the model adjusted the input images to the format and size the model expected. Its core convolution layers then extract features like shapes, colors, edges, and other related patterns to distinguish the substances. The inverted residual block helps the model identify more complex features. In the final step, the model analyzes and categorizes images using extracted features.

As a result, this study aims to examine how well the CNN model helps ophthalmologists in the detection of Glaucoma less stressfully and efficiently without any additional preprocessing and segmentation approaches. The novelty of our proposed approach is automatic feature extraction to detect Glaucoma in fundus images through fine-tuning model hyperparameters. It follows a streamlined architecture with efficient feature fusion techniques. The network inversion concept of MobileNet-V3 helps to extract more complex features while maintaining efficiency. Squeeze-and-excitation (SE) and Hard-Swish blocks improve learning of features. However, MobileNet's computational cost is slightly high. MobileNet gains lower accuracy as compared to other more complex models. MobileNets performance is limited when it deals with very large datasets. Despite these, MobileNetV3 gains better accuracy than its earlier versions in image classification tasks.

3. THE PROPOSED METHODOLOGY

This section presents the suggested technique MobileNet based on CNN to detect Glaucoma. Due to efficiency in speed and size, versatility in deployment, transfer learning, and pre-trained models (trained on large-scale ImageNet data set) [9], scalability and architectural innovations of MobileNet were chosen here for the analysis. MobileNet provides us with high-performance output with the efficient use of computer resources along with minimum computational load. Figure 2 depicts the basic architecture of the suggested model.

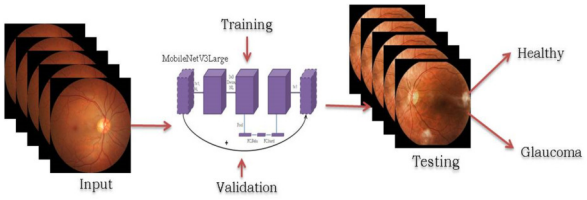

Fig.2: *Architecture of the suggested Model.*

3.1 Dataset

To examine the efficacy of the model, the DrishtiGS, EyePACS AIROGS-Light, BEH, REFUGE, sjchoi86-HRF, CRFO, G1020, FIVES, and PAPILA datasets were utilized in this study. The Datasets were split into three subcategories: training, validation, and testing. DrishtiGS [18] dataset contains 101 images with 2047 X 1760 resolutions in PNG format and digitized with 30° field-of-view (FOV). EyePACS AIROGS-Light[19] dataset contains 6540 images in JPG Format, and the resolution is 256 X 256. BEH [20] dataset captured by TopCon Retinal Camera TRC-50DX and comprises 634 images in JPG format. The retinal Fundus Glaucoma (REFUGE) [21] dataset was captured by Zeiss Visucam 500 (2124×2056 pixels) and Canon CR-2 (1634×1634 pixels) and comprises 1200 images. The dataset sjchoi86-HRF [22] consists of 601 images in PNG format. CRFO [23] dataset contains 79 images. G1020 [24] dataset consists of 1020 fundus images in JPG format with a resolution between 1944 X 2108 and 2426

X 3007 pixels. FIVES [25] dataset pictures were acquired by a TRC-NW8 fundus camera with an FOV of 50°. It contains 800 images in PNG format with 2048 X 2048 pixels. PAPILA [26] dataset images were captured by Topcon TRC-NW400, and all images were in JPEG with 2576 X 1934 pixels.

3.2 MobileNet Architecture

The MobileNet is a lightweight CNN model designed by Google in 2017. MobileNet is a type of CNN mainly based on depthwise separable convolutions instead of standard convolution in traditional neural networks [15]. MobileNet consists of several layers: the input layer, convolution layer (depthwise convolution and pointwise convolution), downsampling (stride and pooling), fully connected layers, and softmax layers. MobileNet has 3 variants. Each variant has its own individual characteristics. MobileNetV3 is an advanced version of the basic MobileNet architecture over MobileNet-V1 and MobileNet-V2, by coopting various advancements and specific optimizers to enhance performance on different databases. For example, MobileNetV3 added squeeze-and-excitation blocks and made several structural modifications to maximize the model performance. The Layer-wise architecture of the model is summarized below [10]:

1. Input Layer: Input Layer takes input of images with a size of 224x224x3.
2. Convolutional Layers: It is the initial convolutional layer with a small kernel size of 3x3 and a few filters that extract low-level features from input images.
3. Bottleneck Blocks with Inverted Residuals: MobileNetV3 contains an Inverted residual block with depthwise separable convolution, an extension layer followed by depthwise separable convolution, and a linear projection layer. It includes a Squeeze and Excitation module within a few of its bottleneck blocks.
4. Activation functions: Hard swish and Mish nonlinearity are introduced in this layer while being computationally efficient [10].
5. Efficient Last Layer: This layer is created to balance efficiency with accuracy, which includes an average pooling layer, a few convolution layers, and an output layer [10].

The basic architecture of the MobileNet model is shown in Figure 3 [11].

MobileNetV3 comes in two versions, MobileNetV3small and MobileNetV3large, which were designed for different computational requirements; the 'small' version is utilized for efficiency and speed, and the 'large' version well focuses more on the accuracy of performance. Using the feature extraction layer of MobileNet-V3 Large, we can achieve better results without additional transfer learning or training. This assures that our proposed model is exclusively designed for glaucoma identification from retinal fundus pictures [11]. Table 1 shows the MobileNet-V3 network parameter information.

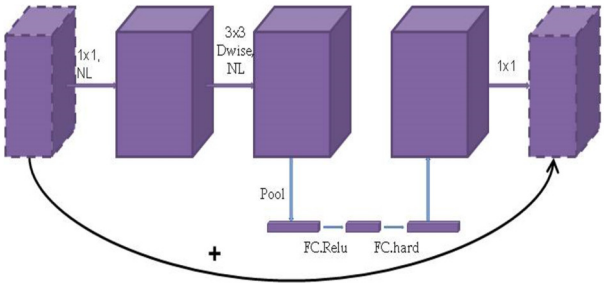

Fig.3: *Basic Architecture of MobileNetV3.*

Table 1: *Parameter information of MobileNet-V3 model [10,11].*

Input	Operator	exp size	#out	SE	NL	s
224^2 x 3	conv2d	-	16	-	HS	2
112^2 x 16	bneck, 3x3	16	16	-	RE	1
112^2 x 16	bneck, 3x3	64	24	-	RE	2
56^2 x 24	bneck, 3x3	72	24	-	RE	1
56^2 x 24	bneck, 3x3	72	40	✓	RE	2
28^2 x 40	bneck, 3x3	120	40	✓	RE	1
28^2 x 40	bneck, 3x3	120	40	✓	RE	1
28^2 x 40	bneck, 3x3	240	80	-	HS	2
14^2 x 80	bneck, 3x3	200	80	-	HS	1
14^2 x 80	bneck, 3x3	184	80	-	HS	1
14^2 x 80	bneck, 3x3	184	80	-	HS	1
14^2 x 80	bneck, 3x3	480	112	✓	HS	1
14^2 x 112	bneck, 3x3	672	112	✓	HS	1
14^2 x 112	bneck, 3x3	672	160	✓	HS	2
7^2 x 160	bneck, 3x3	960	160	✓	HS	1
7^2 x 160	bneck, 3x3	960	160	✓	HS	1
7^2 x 160	conv2d, 1x1	-	960	-	HS	1
7^2 x 960	pool, 7x7	-	-	-	-	1
1^2 x 960	conv2d 1x1, NBN	-	1280	-	HS	1
1^2 x 1280	conv2d 1x1, NBN	-	k	-	-	1

Here, SE = Squeeze-And-Excite presence in the block, NL = nonlinearity, HS = h-ewish, and RE = ReLU. NBN = no batch normalization, s = stride.

3.3 ALGORITHM OF THE PROPOSED METHODOLOGY

Following is a summary of the algorithm followed by the proposed methodology

- **Step 1:** Importing of all necessary libraries and modules. Set the target size of images.
- **Step 2:** Define data generators for training, validation, and testing of images. In this analysis, neural networks utilized Keras with TensorFlow backend.
- **Step 3:** Preprocessing steps include noise injection and data augmentation for training images.

- **Step 4:** Defining the proposed model architecture and loading the pre-trained MobileNet-V3 model for the binary classification task. It includes dropout regularization, performs global max pooling, applies convolutional layers, pooling layers, and output layers, and incorporates callbacks for training.
- **Step 5:** After compilation of the model, training of the model is performed, followed by validation and testing.
- **Step 6:** Training, validation metrics, and Confusion Matrix are generated for further analysis.

3.4 HYPER-PARAMETER USED

The following hyper-parameters were set for fine-tuning of the proposed model:

Loss Function: Binary cross-entropy loss function is utilized here for the Binary-class classification. It is also known as Logarithmic Loss or Log Loss. The loss function is calculated as [27]:

$$log\ loss = \frac{1}{N}\sum_{i=1}^{N} -(y_i * \log(p_i) + (1-y_i) * \log(1-p_i)) \quad (1)$$

Y_i is the actual label/class, P_i is the probability of Class 1, and (1- P_i) is the probability of Class 0. N is the number of predictions in the output list.

Optimizer: Optimizer minimizes the training loss value for better model performance. Due to the robustness, Adam Optimizer is utilized here. [28]

Learning Rate: The learning rate of 1e-4 is set for this analysis.

Implementation Framework: The hardware configuration used for the experiment is AMD Ryzen 9 5900Hx with Radeon Graphics card, NVIDIA GeForce RTX 3060 Laptop GPU GDDR6 @ 6GB(192 bits), and the experiment was conducted in Jupyter Notebook 6.5.2 platform.

4. RESULTS AND DISCUSSION

The performance of the suggested model is examined using performance metrics like Accuracy, Precision, Sensitivity, Specificity, F1Score, and Confusion Matrix. The fundus images are classified into four categories: true positive(TP), true negative(TN), false positive(FP), and false negative(FN). The TP marks the picture as glaucomatous, which is an accurate hypothesis. The TN marks the image as a normal fundus, and it is also a correct classification. The FP incorrectly marks a normal fundus as glaucomatous, and the FN incorrectly marks a glaucomatous fundus as a normal fundus.

Mobilenet model is pre-trained on large image datasets like ImageNet and can recognize a large number of visual patterns and features. The performance of the models was found to fluctuate across repeated experiments, which can be due to several factors, such as the quality of the input data, the quantity of the input data, the choice of hyperparameters, and the random initialization of the model weights. Multiple experiments are often conducted to account for this variability, and the results are averaged to obtain a more reliable estimate of the model's performance. In this experiment, the model is trained for 25 epochs with a batch size of 32. Overall, the use of pre-trained CNNs for retinal fundus image classification is a promising approach, and the choice of model and dataset split can have a vital effect on the performance of the suggested model. Additional experimentation and fine-tuning of the models may be necessary to optimize their performance for specific applications [29].

- **STUDY 1:** The experiment was conducted using the DRISHTI-GS dataset with 108 pictures for training, 24 images for validation, and 23 images for testing.
- **STUDY 2:** The experiment was conducted by combining EyePACS AIROGS-Light, BEH, REFUGE, sjchoi86-HRF, CRFO-v4, G1020, FIVES, and PAPILA dataset by randomly selecting fundus images from each dataset, where 2458 images were utilized for training, 397 pictures for validation and 647 pictures for testing.
- **STUDY 3:** The experiment was conducted using the EyePACS AIROGS-Light dataset by utilizing 5000 pictures for training, 540 pictures for validation, and 1000 pictures for testing.
- **STUDY 4:** The experiment was conducted by the EyePACS AIROGS-Light dataset and DRISHTI-GS dataset, where the EyePACS AIROGS-Light dataset was utilized for training and testing, and the DRISHTI-GS dataset of 101 images was used only for validation of the training data.

The evolution of the loss function and accuracy function during the training of a model can be seen in a graph of training loss per epoch and training accuracy per epoch, respectively. Figure 4-7 shows the graph of training loss per epoch, validation loss per epoch, training accuracy per epoch, and validation accuracy per epoch. The x-axis of the graph denotes the epoch, which indicates how many times the model has iterated through the whole dataset during training, whereas the y-axis represents the training loss, training accuracy, validation loss, and validation accuracy, respectively. Figure 8 (a)-(d) depicts the confusion matrix achieved by the proposed model in all four experiments.

For effectiveness analysis, the Sensitivity, Specificity, Precision, Accuracy, and F1score of the suggested Model have been computed. The accuracy is

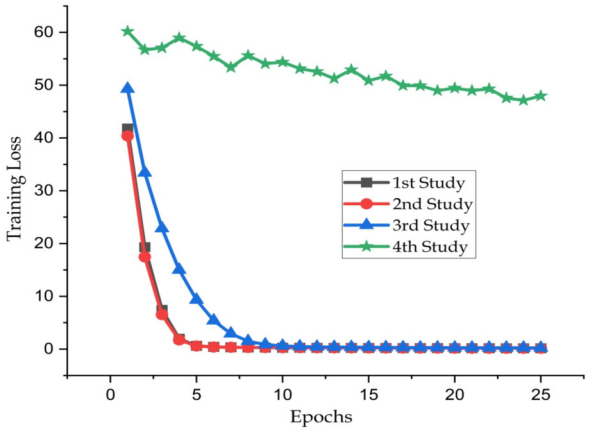

Fig.4: *Training Loss per epoch graph.*

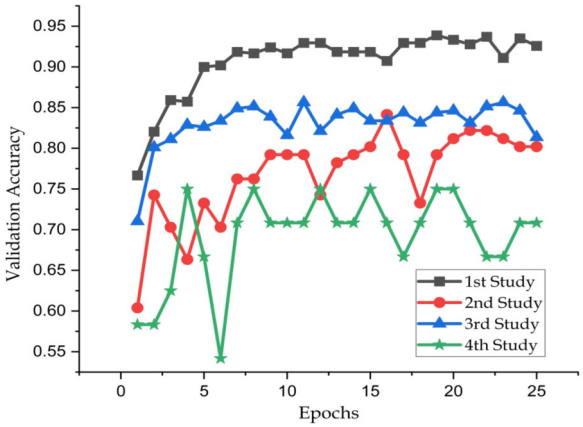

Fig.7: *Validation Accuracy per epoch graph.*

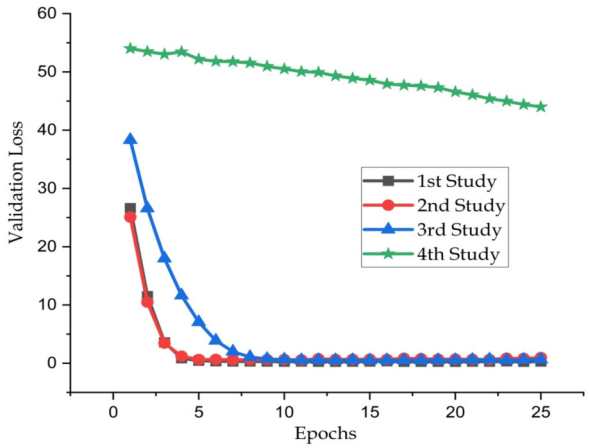

Fig.5: *Validation Loss per epoch graph.*

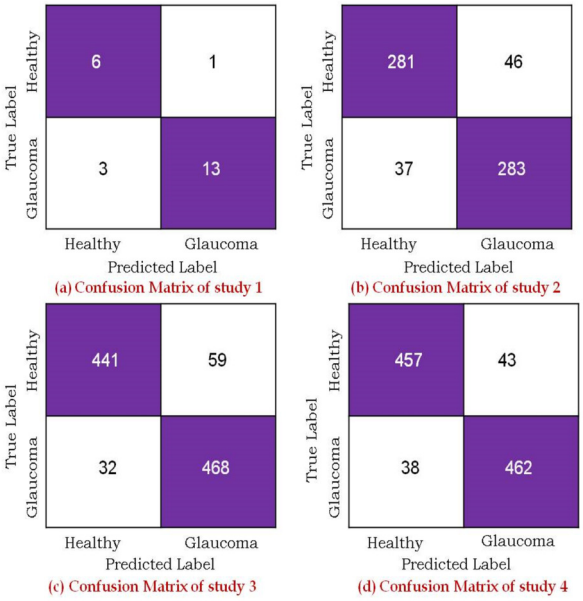

Fig.8: *(a)-(d) Confusion Matrix for data validation on studies 1, 2, 3, and 4.*

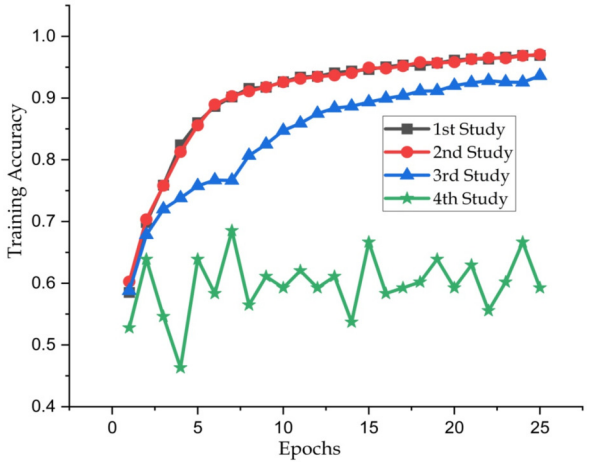

Fig.6: *Training Accuracy per epoch graph.*

calculated as the number of accurately identified fundus pictures among a total number of pictures. The sensitivity measures the accurately identified glaucomatous images. The specificity measures the accurately identified healthy images [30]. Precision is the measurement of instances predicted as positive by the model that are actually true positives. F1Score measures the model's accuracy by combining the sensitivity and precision values. The Precision, Sensitivity, Specificity, Accuracy, and F1Score of the confusion matrix can be calculated using the formula (2), (3), (4), (5) and (6) respectively. Table 2 shows the performance of the proposed model in each experiment.

$$\text{PRECISION} = \text{TP}/(\text{TP}+\text{FP}) \quad (2)$$

$$\text{SPECIFICITY} = \text{TN}/(\text{TN}+\text{FP}) \quad (3)$$

$$\text{SENSITIVITY} = \text{TP}/(\text{TP}+\text{FN}) \quad (4)$$

$$\text{ACCURACY} = (\text{TP}+\text{TN})/(\text{TP}+\text{FP}+\text{TN}+\text{FN}) \quad (5)$$

$$\text{F1Score} = 2\text{TP}/(2\text{TP}+\text{FP}+\text{FN}) \quad (6)$$

Table 2: *Performance of the Proposed Model.*

Study	Accuracy	Sensitivity	Specificity	Precision	F1 Score
1	0.8260	0.8125	0.8571	0.9285	0.8666
2	0.8717	0.8843	0.8593	0.8601	0.8721
3	0.9090	0.9360	0.8820	0.8880	0.9110
4	0.9190	0.9240	0.9140	0.9148	0.9194

The accuracy achieved by the suggested model is 0.8260, 0.8717, 0.9090 and 0.9190. The accuracy of 0.8260 indicates that the model may have trouble generalizing due to a smaller image database. The accuracy of 0.8717 indicates a comparatively better outcome than the previous study because of the larger input database. The accuracy of 0.9090 shows that continuous improvement of ongoing training and validation processes may result in the development of accuracy. The accuracy of 0.9190 indicates a significant development, most likely due to modifications of the validation dataset, which is utilized to improve the classification by evaluating generalization, avoiding overfitting, and fine-tuning hyperparameters. The sensitivity of 0.8125 indicates that the model is executed on a smaller dataset, and the model correctly identified around 81.25% of actual positive cases. A slightly higher sensitivity of 0.8843 shows development in the accurate identification of positive cases. Higher Sensitivity values of 0.9360 and 0.9240 demonstrate the improvement in model performance to identify positive cases for advanced training and additional input data. The Specificity of 0.8571 indicates that the model is executed on a smaller dataset, and the model accurately identified around 85.71% of actual negative cases among the input dataset, and the Specificity of 0.8593 shows recognition of 85.93% of correct negative cases. A specificity of 0.8820 shows recognition of 88.20% of correct negative cases. A specificity of 0.9140 indicates a significant improvement in model performance to identify negative cases for advanced training and additional input data. A precision of 0.9285 indicates that 92.85% of the finalized positive cases by the proposed model are true positive cases. Precision values of 0.8601 and 0.8880 identify comparatively fewer positive cases, which may be due to fewer training images. A precision value 0.9148 indicates that the proposed model is more accurate in recognizing related patterns throughout the dataset. F1Score of 0.8666 and 0.8721 demonstrates model performance on a smaller dataset, whereas F1Score of 0.9110 and 0.9194 indicates that the model achieves high overall performance by proper balancing among precision and sensitivity.

Table 3: *Performance comparison of our method with other related methods.*

References	Method	Accuracy	Sensitivity	Specificity
Khan et al.[1] 2022	LS-SVM classifier	91.22%	-	-
Mojab et al[31] 2019	Deep CNN	79.04%	79.04%	88.95%
Blagus et al.[32] 2013	Multi Branch neural network	91.51%	92.33%	90.90%
Chai et al.[33] 2018	RCNN	91.51%	92.33%	90.90%
Brigatti et al.[34]	NN(Back propagation)	88	90	84
Deepika and Maheswari et al.[35] 2018	SVM	91.7%	90.0%	93.3%
Kang et al.[36]2020	SVM	85.1%	82.0%	88.3%
Parashar and Agrawal et al.[37] 2020	LS-SVM classifier	90.76%	-	-
Xu et al.[38] 2021	Simple rule on RNFLD, then SVM	-	96.1%	95.6%
Afolabi et al.[39] 2021	Extreme Gradient Boost	88.3%	-	-
Lenka and Mayaluri [40] 2023	Hybrid model(UNet, SVM)	91.83%	96.39%	95.37%
Our Method	MobileNet-V3	0.9190	0.9240	0.9140

Table 3 shows an outline of the outcomes gained from the research done by other researchers. All of these models were developed to diagnose Glaucoma from fundus images. Due to fine-tuning of the parameters and hyper-parameters of the models, the performance was well but not very successful. Our model achieves the highest accuracy of 0.9190, which is relatively high and also proves that our proposed method of binary classification is accurate like other approaches. The model achieves better generalization compared to the above mentioned previous works.

Thus, appropriate preprocessing and fine-tuning of the model for a particular dataset can improve the performance of classification.

MobileNet combines several architectural innovations to gain high performance with a minimal computational requirement. The squeeze and excitation blocks help the network to concentrate on more relevant features to improve performance. However, the suitability of a model depends on various factors, including dataset size, quality, and the specific characteristics of the problem. So, we can say that MobileNet outperforms accurately in categorizing the class of the images without using any additional preprocessing and segmentation or feature extraction techniques. Finally, we can conclude that MobileNetv3Large performs accurately on any datasets with different file formats and resolutions without any pre-processing or segmentation approaches.

5. CONCLUDING REMARKS

In this work, we have proposed MobileNetv3-large deep learning architectures for Glaucoma detection. A classifier has been used to detect the accurate class of the input image. Without using any extra image processing operations for enhancement, the MobileNetv3-large method outperforms state-of-the-art approaches in examined fundus images. This fact is significant for constructing a real-time system that helps detect Glaucoma. To evaluate the effectiveness of the model, the performances of different techniques are evaluated based on different performance parameters like Sensitivity, Specificity, Precision, Accuracy, F1Score, etc. Overall, the experiment provides valuable insights into the performance of pretrained CNNs in detecting Glaucoma, which could have important implications for the improvement of automated screening systems for this condition. Future research will concentrate on Preprocessing and Segmentation strategies and deep learning to classify fundus images. The method will be examined on available labeled online databases and datasets obtained from IGM Hospital, Agartala, Tripura. Different performance metrics, such as sensitivity, Specificity, accuracy, and ROC Curve, will be utilized to assess the suggested approach.

AUTHOR CONTRIBUTIONS

Conceptualization, SD and SM; methodology, SD and SM; programming, SD; validation, SD and SM; formal analysis, SD and SM; investigation, SD, SM and MM; data collection and curation, SD; writing—original draft-preparation, SD and SM; writing—review and editing, SD, SM and MM; visualization, SD and SM; supervision, SM and MM; funding acquisition, NOT APPLICABLE; All authors have read and agreed to the published version of the manuscript.

References

[1] S. I. Khan, S. B. Choubey, A. Choubey, A. Bhatt, P. V. Naishadhkumar and M. M. Basha, "Automated Glaucoma Detection from Fundus Images Using Wavelet-Based Denoising and Machine Learning," *Concurrent Engineering*, vol. 30, no. 1, pp. 103-115, 2021.

[2] S. Ajitha, J. D. Akkara and M. V. Judy, "Identification of Glaucoma from fundus images using deep learning techniques," *Indian Journal of Ophthalmology*, vol. 69, no. 10, pp. 2702-2709, 2021.

[3] N. Akter, J. Fletcher, S. Perry, M. P. Simunovic, N. Briggs and M. Roy, "Glaucoma Diagnosis Using Multi-Feature Analysis and a Deep Learning Technique," *Scientific Reports*, vol. 12, no. 8064, 2022.

[4] R. Hemelings, B. Elen, J. Barbosa-Breda, M. B. Blaschko, P. De Boever and I. Stalmans, "Deep Learning on Fundus Images Detects Glaucoma beyond the Optic Disc," *Scientific Reports*, vol. 11, no. 20313, 2021.

[5] S. Joshi, B. Partibane, W. A. Hatamleh, H. Tarazi, C. S. Yadav and D. Krah, "Glaucoma Detection Using Image Processing and Supervised Learning for Classification," *Journal of Healthcare Engineering*, vol. 2022, no. 2988262, 2022.

[6] R. Kashyap, R. Nair, S. M. P. Gangadharan, M. Botto-Tobar, S. Farooq and A. Rizwan, "Glaucoma Detection and Classification Using Improved U-Net Deep Learning Model," *Healthcare*, vol. 10, no. 12, p. 2497, 2022.

[7] B. S. Kirar and D. K. Agrawal, "Current Research on Glaucoma Detection Using Compact Variational Mode Decomposition from Fundus Images," *International Journal of Intelligent Engineering and Systems*, vol. 12, no. 3, pp. 1–10, 2019.

[8] L. J. Coan, B. M. Williams, V. K. Adithya, S. Upadhyaya, A. Alkafri, S. Czanner, R. Venkatesh, C. E. Willoughby, S. Kavitha and G. Czanner, "Automatic detection of glaucoma via fundus imaging and artificial intelligence: A review," *Survey of Ophthalmology*, vol. 68, no. 1, pp. 17-41, 2023.

[9] Y. Harjoseputro, I. P. Yuda and K. P. Danukusumo, "MobileNets: Efficient Convolutional Neural Network for Identification of Protected Birds," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 10, no. 6, pp. 2290-2296, 2020.

[10] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *arXiv preprin arXiv:1704.04861*, 2017.

[11] L. Jia, T. Wang, Y. Chen, Y. Zang, X. Li, H. Shi and L. Gao, "MobileNet-CA-YOLO: An Improved YOLOv7 Based on the MobileNetV3 and Attention Mechanism for Rice Pests and Diseases Detection," *Agriculture*, vol. 13, no. 7, pp. 1285–1285, 2023.

[12] B. Khasoggi, E. Ermatita and S. Samsuryadi, "Efficient Mobilenet Architecture as Image Recognition on Mobile and Embedded Devices," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 16, no. 1, pp. 389-394, 2019.

[13] A. Michele, V. Colin and D. D. Santika, "MobileNet Convolutional Neural Networks and Support Vector Machines for Palmprint Recognition," *Procedia Computer Science*, vol. 157, pp. 110–117.

[14] D. Sinha and M. El-Sharkawy, "Thin MobileNet: An Enhanced MobileNet Architecture," *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, New York, NY, USA, pp. 0280-0285, 2019.

[15] F. Zhu, C. Liu, J. Yang and S. Wang, "An Improved MobileNet Network with Wavelet Energy and Global Average Pooling for Rotating Machinery Fault Diagnosis," *Sensors*, vol. 22, no. 12, p. 4427, 2022.

[16] V. C R, V. Asha, U. M and T. M, "Glaucoma Detection using Transfer Learning," *2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, Coimbatore, India, pp. 476-481, 2023.

[17] R. Patil and S. Sharma, "Automatic glaucoma detection from fundus images using transfer learning," *Multimedia Tools and Applications*, Springer, 2024.

[18] https://www.kaggle.com/datasets/lokeshsaipureddi/drishtigs-retina-dataset-for-onh-segmentation

[19] https://www.kaggle.com/datasets/deathtrooper/eyepacs-airogs-light

[20] https://github.com/mirtanvirislam/Deep-Learning-Based-Glaucoma-Detection-with-Cropped-Optic-Cup-and-Disc-and-Blood-Vessel-Segmentation/tree/master/Dataset

[21] https://ieee-dataport.org/documents/refuge-retinal-fundus-glaucoma-challenge

[22] https://github.com/cvblab/retina_dataset

[23] https://data.mendeley.com/datasets/trghs22fpg/4

[24] https://www.kaggle.com/datasets/arnavjain1/glaucoma-datasets

[25] https://figshare.com/articles/figure/FIVES_A_Fundus_Image_Dataset_for_AI-based_Vessel_Segmentation/19688169/1

[26] https://figshare.com/articles/dataset/PAPILA/14798004

[27] https://medium.com/@TheDataScience-ProF/understanding-log-loss-a-comprehensive-guide-with-code-examples-c79cf5411426

[28] M. Tariq, V. Palade and Y. Ma, "Transfer learning-based classification of Diabetic Retinopathy on the Kaggle EyePACS dataset," *Medical Imaging and Computer-Aided Diagnosis*, vol. 810, pp.89-99, 2023.

[29] S. Das, M. Mishra, and S. Majumder, "A Comprehensive Analysis of Diabetic Retinopathy Detection in Retinal Fundus Images Using Different Convolutional Neural Network," *ECTI-CIT Transactions*, vol. 17, no. 4, pp. 510–521, Nov. 2023.

[30] A. Shoukat, S. Akbar, S. A. Hassan, S. Iqbal, A. Mehmood and Q. M. Ilyas, "Automatic Diagnosis of Glaucoma from Retinal Images Using Deep Learning Approach," *Diagnostics*, vol. 13, no. 10, pp. 1738–173, 2023.

[31] N. Mojab, V. Noroozi, P. S. Yu and J. A. Hallak, "Deep Multi-Task Learning for Interpretable Glaucoma Detection," *2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI)*, Los Angeles, CA, USA, pp. 167-174, 2019.

[32] R. Blagus and L. Lusa, "SMOTE for high-dimensional class-imbalanced data," *BMC Bioinformatics*, vol. 14, no. 106, 2013.

[33] Y. Chai, H Liu and J. Xu, "Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models," *Knowledge-Based Systems*, vol. 161, pp. 147-156, 2018.

[34] L. Brigatti, D Hoffman and J Caprioli, "Neural networks to identify Glaucoma with structural and functional measurements," *Am. J. Ophthalmol*, vol. 121, no. 5, pp. 511–521, 1996.

[35] E. Deepika and S. Maheswari, "Earlier glaucoma detection using blood vessel segmentation and classification," *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, Coimbatore, India, pp. 484-490, 2018.

[36] H. Kang, X. Li and X. Su, "Cup-disc and retinal nerve fiber layer features fusion for diagnosis glaucoma," *Medical Imaging 2020: Computer-Aided Diagnosis*, vol. 11314, no. 113143z, 2020.

[37] D. Parashar and D. K. Agrawal, "Automated Classification of Glaucoma Stages Using Flexible Analytic Wavelet Transform From Retinal Fundus Images," in *IEEE Sensors Journal*, vol. 20, no. 21, pp. 12885-12894, 1 Nov.1, 2020.

[38] O. J. Afolabi, G. P. Mabuza-Hocquet, F. V. Nelwamondo and B. S. Paul, "The Use of U-Net Lite and Extreme Gradient Boost (XGB) for Glaucoma Detection," in *IEEE Access*, vol. 9, pp.

47411-47424, 2021.

[39] Y. Xu, M. Hu, H. Liu, H. Yang, H. Wang, S. Lu, T. Liang, X. Li, M. Xu, L. Li, H. Li, X. Ji, Z. Wang, L. Li, R. N. Weinreb and N. Wang, "A hierarchical deep learning approach with transparency and interpretability based on small samples for glaucoma diagnosis," *NPJ digital medicine*, vol. 4, no. 48, pp. 1-11, 2021.

[40] S. Lenka and Z. L. Mayaluri, "Hybrid glaucoma detection model based on reflection components separation from retinal fundus images," *EAI Endorsed Transactions Research Article on Pervasive Health and Technology*, vol. 9, pp. 1-13, 2023.

Madhusudhan Mishra has been working as an Associate Professor in the Department of Electronics and Communication Engineering, North Eastern Regional Institute of Science and Technology (NERIST), Nirjuli, Arunachal Pradesh since 2019. He received his B.Tech. Degree in Electronics and Communication Engineering from NERIST (2004), M. Tech in Signal Processing from IIT Guwahati (2011), and holds a Ph.D. Degree from the Department of Electrical Engineering, IIT Kharagpur, India (2020). His research areas focus on Biomedical Signal Processing and Machine Learning. With more than 18 years of experience in research, teaching, designing, and leading programs, he co-edited a research book and published his work in different journals, conferences, and professional press of repute.

Smita Das received her B.Tech degree in Computer Science and Engineering from National Institute of Technology, Agartala, Tripura, India in 2006, M.Tech in Computer Science and Engineering from Tripura University (A Central University), India in 2008. Currently, she is working as an Assistant Professor in the Department of Computer Science, MBB College, Tripura, India. She is pursuing Ph.D degree from Information Technology Department of Tripura University (A Central University), Tripura, India. Her research interests include medical image processing.

Swanirbhar Majumder received his Ph.D, M.Tech, and B.Tech from Jadavpur University, in West Bengal; University of Calcutta, West Bengal; and North Eastern Regional Institute of Science and Technology (NERIST), under North Eastern Hill University (NEHU), Shillong in India respectively. Currently, since 2020 he has been working as a Professor of Information Technology at Tripura University (A Central University). His research interests include Signal Processing, Medical Image processing.