.

Human Eye Cataract Detection Using Lightweight Deep Learning Models

Mohammad Shaikhul Fuad, Md. Zahidul Haque, Sayeed Miner

and Dr. Atiqur Rahman

Department of Electrical and Computer Engineering, North South University,
Bashundhara, Dhaka-1229, Bangladesh
{shaikhul.fuad, zahidul.haque10, sayeed.miner,
atiqur.rahman}@northsouth.edu

Abstract: A cataract is an eye disease that intends human vision to blurry. It forms a cloud-like over the retinas. As a result, the victim who has cataract eye face difficulties seeing, as everything around looks vague. Many times prolongation of the disease causes sheer blindness. The objective of the paper is to detect cataract images through deep learning approaches at the initial stage and reduce the severity of the patient. Therefore the study approaches some of the pre-trained models under CNN that are smaller in size, less in parameters, and faster in training speed. Like, EfficientNetB1, EfficientNetB4, MobileNet and EfficientNetV2B0. These lightweight models were deliberately used to see the outcome of the models that are low at sizes with less computational time. And the results were amazingly well. Although the training accuracy of all models reached above 99%, for the testing set, EfficientNetV2B0 with 99.54% accuracy had outdone MobileNet with 99.08% test accuracy and EfficientNetB1 and EfficientNetB4 with 98.62% and 98.17% of accuracy.

Keywords: Deep Learning, Lightweight models, CNN, EfficientNetV2B0, MobileNet, Cataract Eye Detection.

1 Introduction

An eye condition known as a cataract causes the patient's eyes to seem cloudy. A person who has cataracts in their eyes will have a vision that is cloudy or frosted over. A person afflicted with cataracts in their eyes would struggle with activities such as reading, driving, and even recognizing the face of another person [1].

Worldwide around 285 million have some vision weakening trouble out of which moderate to the severity of 246 million and total blindness of 39 million are investigated by The World Health Organization (WHO) [2]. Age-related cataracts have rendered 19.34 million persons legally blind (their best eye's eyesight is less than 3/60), as stated by the World Health Report published in 1998. This was responsible for 43 per cent of all occurrences of blindness [3].

Cataracts are an age-related condition that can also be carried on using crystalline contact lenses. The lens's transparency and optical homogeneity are maintained by many inter-reliant factors that contain the structure of the lens's microscope and its interior's chemical composition. When a yellow-brown pigment is observed, this indicates the presence of a progressive deposit in the lens. This becomes more prevalent as one gets older. This also decreases the amount of light allowed to enter the eyes. The symptoms of cataracts are determined mainly by the type of cataract a person has, his way of life and the visual demands he places on himself.

The terms "intracapsular" and "extracapsular cataract extraction" are interchangeable and refer to the same surgical procedure. Intra capsular extraction is a method that involves removing the lens in its entirety while maintaining the integrity of the capsule. This mode of therapy is utilized for very little patient care in developed countries. Because it requires a far lower number of high-priced and technically advanced devices, it is nevertheless widely practiced in developing nations. The electrical supply doesn't need to be highly stable. In addition to that, it only requires a little amount of practice to accomplish the task.

Extracapsular extraction is another procedure that can be used. It is necessary to make a relatively broad incision to extract the lens's nucleus in a single piece.

Cataract surgery has advanced to the point that it is far more effective than it was twenty years ago. Individuals who have undergone cataract surgery will have the best-corrected vision of 6/12 (20/40 or 0.5) if they don't have any ocular comorbidities like glaucoma, macular degeneration, or diabetic retinopathy [4]. This is the case for approximately 85-90% of these patients.

In the early stages of cataracts, the patient's reaction to refractive lenses for nearsightedness or farsightedness is frequently preferred. Let's say the patient receives outpatient therapy with pupillary dilatation and refractive lenses, but their vision does not improve. In that situation, the patient needs to be confronted at the hospital for intraocular lens implantation and surgical cataract removal [5].

2 Related Work

We investigated recently published articles and journals in order to have a better understanding of the issue, and we spoke about potential solutions to the challenge of enhancing the accuracy of our deep learning model. We used a preexisting dataset and examined their model so that we could compare our work to theirs.

A discrete state transition (DST) system that was built on Res-Net was presented by Zhou. They could resolve the issues with vanishing gradients and achieve an accuracy of (94.000%) in their cataract detection performance. The residual connection method is used in the recommended design for DC-NNs, which allows these problems with gradients to be solved. In addition, the image is not required to be preprocessed, and it can communicate higher-dimensional qualities [6].

In a recent study, an active shape model that leveraged SVM was used. This model was trained on more than 5000 images and achieved an accuracy of 95.00 per cent [7]. Support vector machine analysis is not the best choice when dealing with high-dimension feature maps.

Md. Sajjad Mahmud Khan recently used CNN and VGG-19 to study cataracts, and the results showed an overall accuracy of 97.47 percent, precision of 97.47 percent, and loss of 5.27 percent [8].

A project was proposed by Chandra Lekha Dondapati that would make use of CNN and a dataset containing illnesses such as glaucoma, retinal disease, common cataracts, and others. According to the criteria utilized by CNN [9], an accuracy rating of 82% or higher is regarded as satisfactory.

In addition, the results of different imaging modalities that are utilized for grading the severity of cataract disease were compared in this article. The presentation began with an overview of cataracts, covering topics such as how to distinguish between normal vision and that affected by cataracts and the various conditions that might cause cataracts [10].

Pratap and Kakoli conducted another piece of research. They gathered information from a wide variety of sources. They had a total of 800 pictures. They were able to get an accuracy of 92.91% using DL [11]. There is a possibility that the findings will change when applied to a more extensive dataset.

Mehmet Emre Sertkaya proposed conducting research on retinal diseases with convolutional neural networks and coherent optical pictures as the primary data collection methods. Their different approaches were referred to as AlexNet, LeNet, and Vgg16. They were successful in achieving their goals using the Vgg16 and AlexNet architectures. In this case, the overall accuracy was 82.9% [12].

During the study, researchers used a machine learning-based technique that detects cataracts using support vector machines. Using this method, the entire image is divided into 17 pieces, and each segment is then assigned to the SVM algorithm. The accuracy of this method is 87.52 per cent. On the other hand, partial cataracts cannot be detected [13].

Md. Rajib Hossain and his team also worked on Cataract detection. In his research, he went for Res-Net50 model that comes under CNN. This helped him identify cataracts and non-cataract fundus. His accuracy in training was approximately 100%, while his validation accuracy was nearly as good. Overall his validation accuracy was 97.38% [14].

According to one piece of research, image processing methods can be utilized to analyze fundus images to identify eyes with cataracts. In their study of fundus photos, a team of researchers used two distinct approaches. One is the Kernel-Based, CNN and the other is NABP technique also acknowledged as Novel Angular Binary Pattern, and their submitted process had an accuracy of 0.9739 [15].

Although many deep learning approaches and models have already been used to conduct research, our study also aims to use pre-trained deep learning models under convolutional neural networks but the models that are light in weight. Lightweight models are usually very small in size because the number of trainable parameters is much much less compared to other heavy VGG and ResNet models. Also, it has been researched and determined thoroughly that many of the lightweight models not only provided better performance but also trained faster in imagenet datasets. Additionally, these models are easier to implement or deploy in production due to their low cost. As a result, we have taken four lightweight pre-trained models, EfficientNetV2B0, MobileNet, EfficientNetB1, and EfficientNetB4 to classify between normal eye and cataract eye and to evaluate their performances.

3 Methodology

In this section, specific practices and the objectives that has been taken into account for the purpose of the detection of cataract disease are described in detail. This methodology section will provide a high-level overview of each of the building components described below, as well as their significance in this study.

3.1 Data

The dataset was collected from Kaggle which includes funduscopic images of actual patients' data. It consists of images of both left and right eyes of a particular patient. The dataset 'Ocular Disease Recognition' is made up of several ocular diseases and the images were labelled by doctors [16]. As a result, the cataract and normal images were extracted from the dataset and combined together. Therefore, 1094 images have been taken into account. Among them 594 cataract images and 500 normal images. Images were varying in size, so all the images have been resized to similar pixels 224*224. Later on, the dataset has been split into 80% of training set of 876 images and 20% of testing set of 218 images. The testing set has been used for the validation aspect before going for the actual test, based on the validation performance. Data augmentation technique has been applied in the training set. Both training and validation data have been normalized by factor 1/255. Next, the data has been fed into deep learning lightweight models EfficientNetV2B0, MobileNet, EfficientNetB1 and EfficientNetB4 for training, validation and testing purposes. The figure 1, demonstrates the elaborate processes and techniques applied to conduct the study.

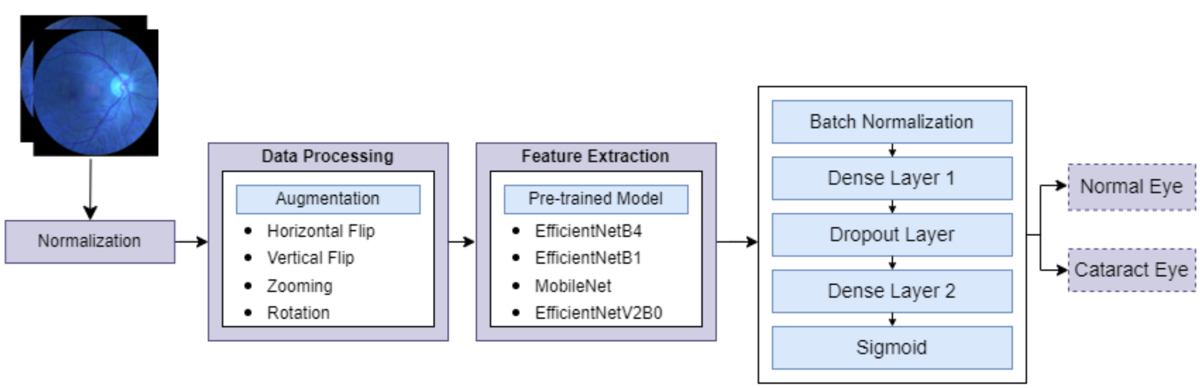

Figure 1: Workflow of the implementation of the models

3.2 Proposed Model

EfficientNetB0-B7: Just like many other models, EfficientNet was also developed by google and primarily used for image classification. Efficientnet works by scaling on depth, width, and resolution. An image with higher resolution needs more channels to extract detailed and more complex features, therefore more numbers of neurons to handle the data. Using an abundance of feature maps also degrades algorithms' training speed. As a result, in order to solve this issue there comes a scaling factor also known as compound scaling that scales width, depth, and resolution in a balanced manner. EfficientNet is a series of networks and EfficientNetB0 is a baseline model where compound scaling has been approached to create further networks of the series [17]. This baseline model is designed by neural architecture search NAS. The figure 2 shows the architectural blocks of the EfficientNetB0 model. It always scales the B0 dimensions to create other efficient models starting from EfficientNets B1 to B7. In all these networks, the layer architecture or blocks will be the same, the only change will be in resolutions, channels, and layers. In every upscaled model, there would be more channels and layers due to higher resolution. EfficientNetB1 has an input size of 240 pixels, whereas 380 resolutions for EfficientNetB4. So new efficient models will be larger than the previous model.

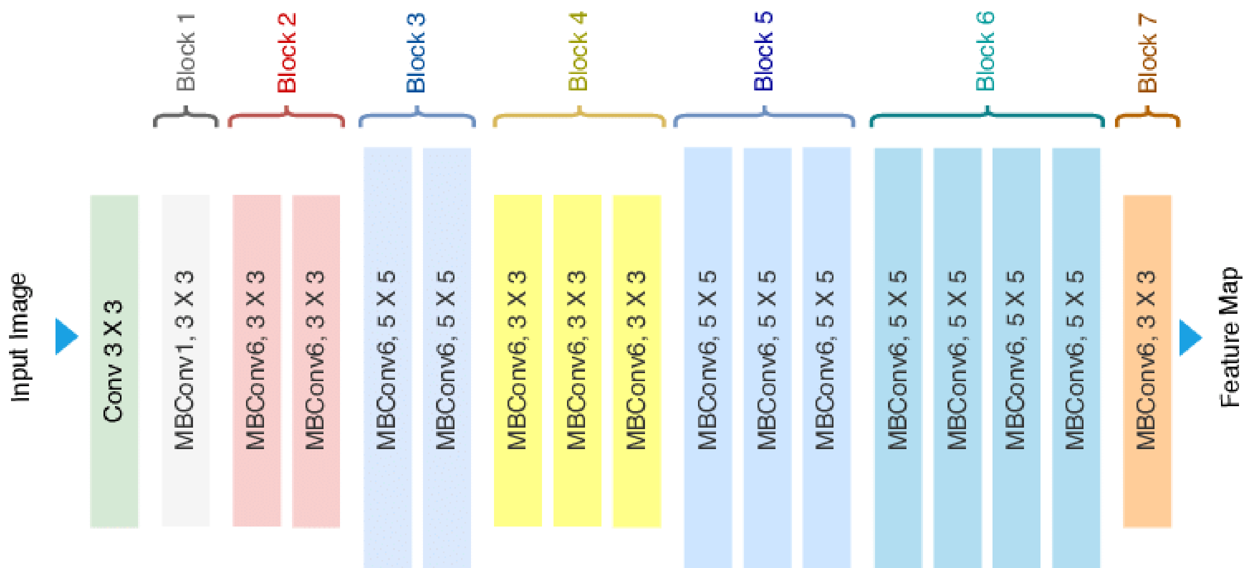

Figure 2: EfficientNetB0 architecture with MBConv blocks [18]

MobileNet: Unlike standard convolutional networks, MobileNet V1 applies depthwise separable convolution to its network. Depthwise separable convolution has two parts, first the filtering stage also known as depthwise convolution and combination stage called as pointwise convolution. The figure 3 shows the architecture of the MobileNet model. Depthwise Convolution works on an only one input channel at a time, while standard convolutional works on all channels at once. It is the progression of filtering the input deprived of accumulation of more functions. Hence, the process of developing new features is combined using point-by-point convolution. Combining the two layers produces a depth separable convolution at the end. With the help of deep convolutional, model put on a single filter to individual input channel and then using 1x1 pointwise convolutions, it generates a linear combination of deep layer outputs. Once every depthwise 3x3 convolution, Batch Normalization and ReLU activation will be used followed by each pointwise 1x1 convolutional [19], with the exception of full convolutional in first and fully connected final layer at the end.

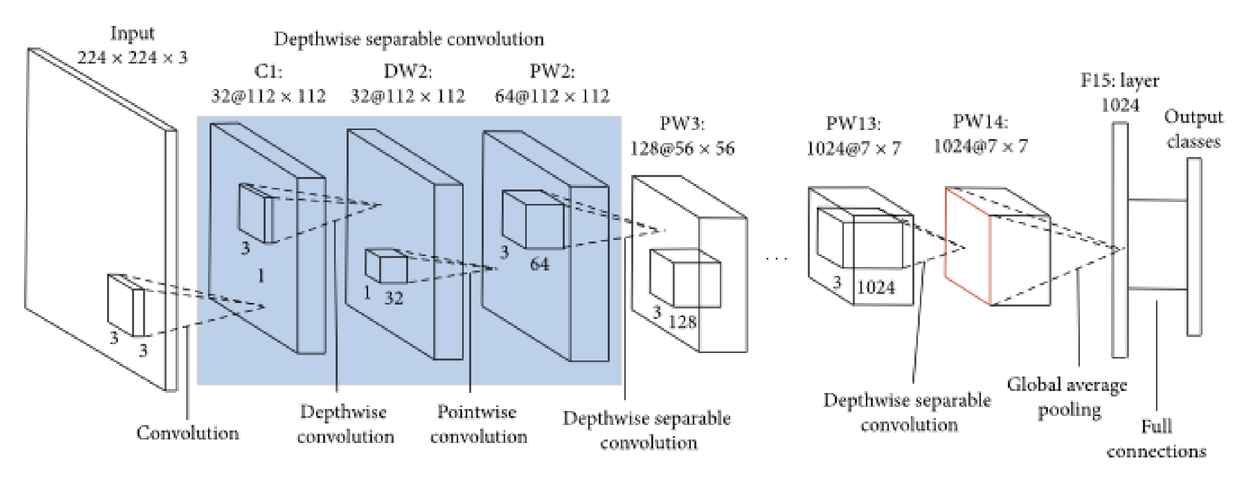

Figure 3: MobileNet architecture [20]

EfficientNetV2B0: In order to tackle the training speed EfficientNetV2 model has been created over the EfficientNet model. The batch size needs to lessen for accommodating large images on the GPU. But training becomes sub-optimal with small batches. In other ways, depthwise convolution is inefficient in GPU or TPU due to low hardware utilization. The third and major problem is the compound scaling scales depth, width, and resolution at all stages of the network with equal proportion, which leads to higher memory usage. But EfficientNetV2 addresses all these drawbacks. The figure 4 shows the architectural diagram of the EfficientNetV2 model. They introduced new fused MBConv in many blocks instead of simple MBConv block. In Fused MBConv the depthwise convolution is fused with the 1x1 convolutional layer before it. On the other hand, EfficientNetV2 memory access tends to have less overhead due to a smaller expansion ratio. Whenever it uses MBConv, kernel size seems to be 3x3 but it also increases the number of layers to make up for the lost receptive field. The final stage of the original EfficientNet, stride-1, is completely eliminated in EfficientNetV2, because of its enormous parameter size and memory access overhead [21].

Stage	Operator	Stride	#Channels	#Layers
0	Conv3x3	2	24	1
1	Fused-MBConv1, k3x3	1	24	2
2	Fused-MBConv4, k3x3	2	48	4
3	Fused-MBConv4, k3x3	2	64	4
4	MBConv4, k3x3, SE0.25	2	128	6
5	MBConv6, k3x3, SE0.25	1	160	9
6	MBConv6, k3x3, SE0.25	2	256	15
7	Conv1x1 & Pooling & FC	-	1280	1

Figure 4: EfficientNetV2 architecture [21]

4 Result and Analysis

This section explains how the proposed models has performed relying on metrics like Accuracy, F1-score, Recall, and Precision.

4.1 Confusion Matrix

It is a 2-dimensional n x n matrix, n is the class numbers where it classifies between actual and predicted classification with the help of numeric values of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) [22].

$$TP_i = a_{ii}$$
$$FP_i = \sum_{j=1,\, j\neq i}^{n} a_{ij} \tag{1}$$
$$FN_i = \sum_{j=1,\, j\neq i}^{n} a_{ij}$$
$$TN_i = \sum_{j=1,\, j\neq i}^{n} \sum_{k=1,\, k\neq i}^{n} a_{jk}$$

The figure 5 gives the confusion matrix for EfficientNetB4 model. It classifies 124 cataract images as cataract and 90 normal images as normal. On the other hand, 3% of normal images are classified as cataract, and 1% of cataract images as normal.

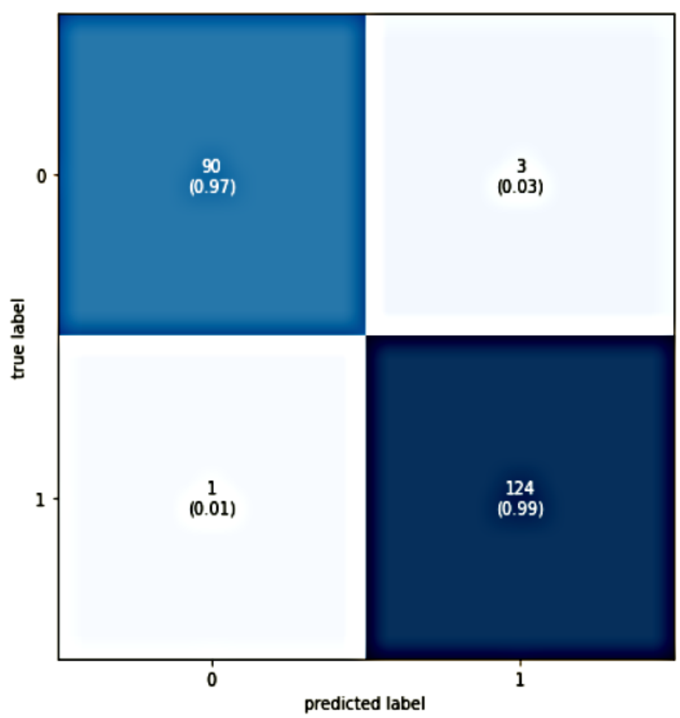

Figure 5: The confusion matrix for EfficientNetB4

The figure 6 illustrates the EfficientNetB1 model confusion matrix. It wrongly conveys 3 normal images as cataract and 0 cataract images as normal. However, 118 cataract images and 97 normal images are categorized truly.

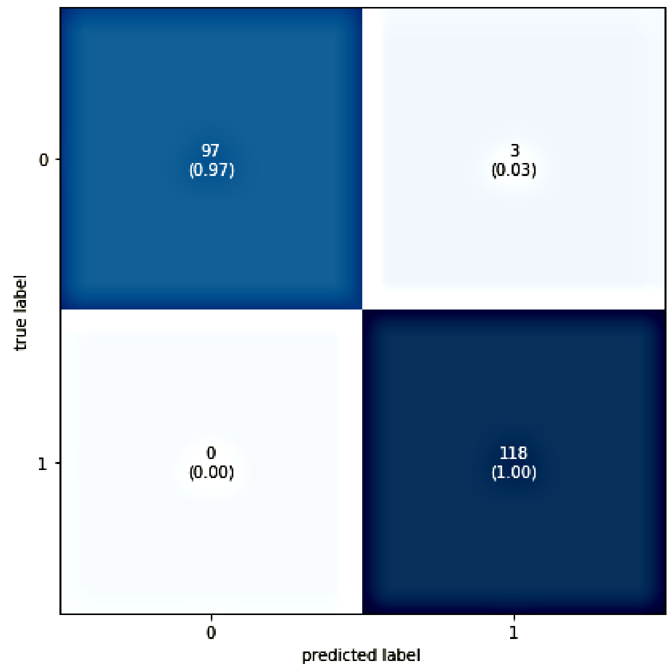

Figure 6: The confusion matrix for EfficientNetB1

The figure 7 represents the confusion matrix for MobileNet that delivered, 95 normal images and 121 cataract images were categorized truly by the network. Conversely, a normal image has been classified as cataract and a cataract image has been classified as normal.

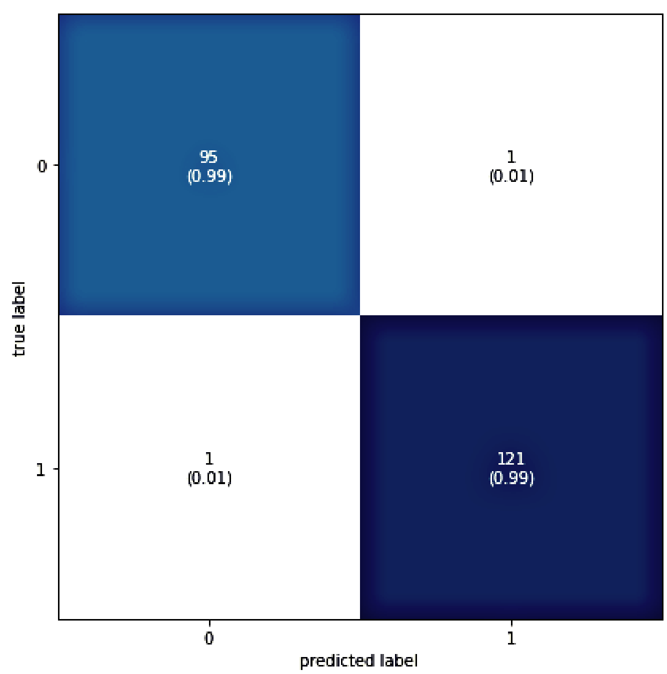

Figure 7: The confusion matrix for MobileNet

EfficientNetV2B0 confusion matrix has been shown in the figure 8. The model truly classified a total of 217 images, 103 normal and 114 cataract images classified correctly. The misclassified ratio of the model is substantially less which is 0.01 as 1 normal image has been wrongly classified into cataract.

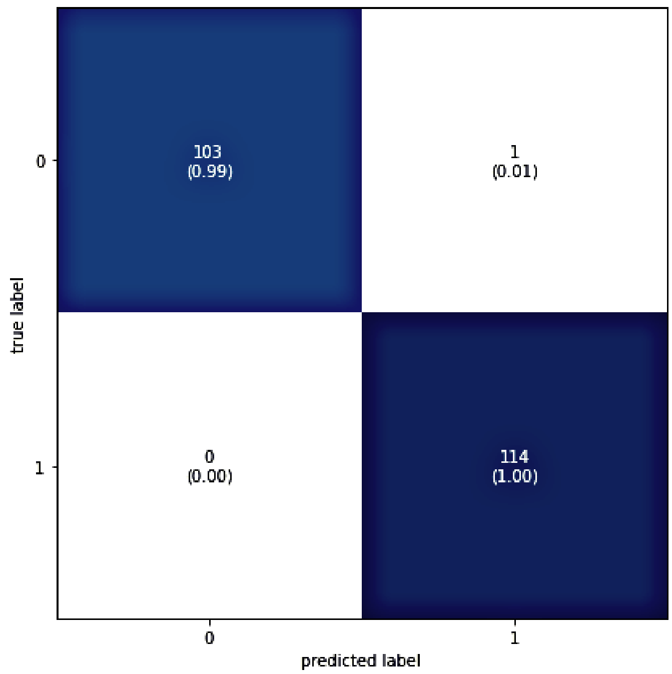

Figure 8: The confusion matrix for EfficientNetV2B0

4.2 Accuracy

Accuracy is the calculation of a model that correctly predicts classes out of the entire dataset or total predicted data.

$$Accuracy = (TP + TN)/ (TP + TN + FP + FN) \tag{2}$$

4.3 Precision

It is how many positive values have rightly been predicted from all positive values, also called positive predictive values.

$$Precision = (True\ Positive)/ (True\ Positive + False\ Positive) \tag{3}$$

4.4 Recall

The recall is another metric to determine how successfully the model finds true positives. A recall value close to one signifies that the model successfully detected true positives. The model gets several false negatives due to the low recall value.

$$Recall = (True\ Positive)/ (True\ Positive + False\ Negative) \tag{4}$$

4.5 F1-score

The F1-score is essentially needed to have a balance between precision and recall. As we previously observed, True Negatives significantly improve accuracy. The F1-score becomes a superior metric when there is an unequal class dispersion or the number of actual negatives is greater.

$$F1\text{-}score = (2 \times Precision \times Recall)/ (Precision + Recall) \tag{5}$$

4.6 Model Accuracy and Loss

EfficientNetB4: Figure 9 shows loss and accuracy graph of validation set and training set. The graph on the left showcases training and validation accuracy over the period of 15 epochs. Training accuracy gradually improves, whereas validation accuracy increased largely at the beginning then followed a fixed pattern due to constant learning rate and climbed up to a mark where there was an insignificant difference between validation and training accuracy. As well as, the loss in training dropped steadily as the epochs moved on while validation loss dropped initially and then slightly varied at the end.

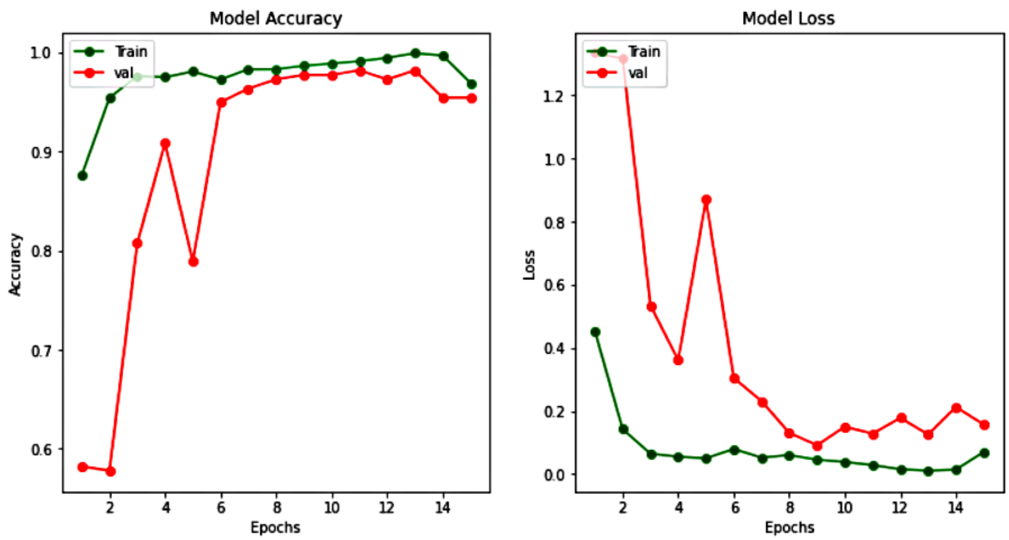

Figure 9: EfficientNetB4 validation and training accuracy and loss graphs.

EfficientNetB1: Figure 10 shows graphs of validation and training loss and accuracy. Due to the increase in learning rate training accuracy increases following a pattern, but validation accuracy does not follow the trend at the initial phase. After the 3rd epoch, it increases, but soon after begins fluctuating. Similarly in the loss graph, the training loss decreases at an approximate rate but validation loss increases in the beginning, significantly drop in between, and then starts increasing due to overfitting issues.

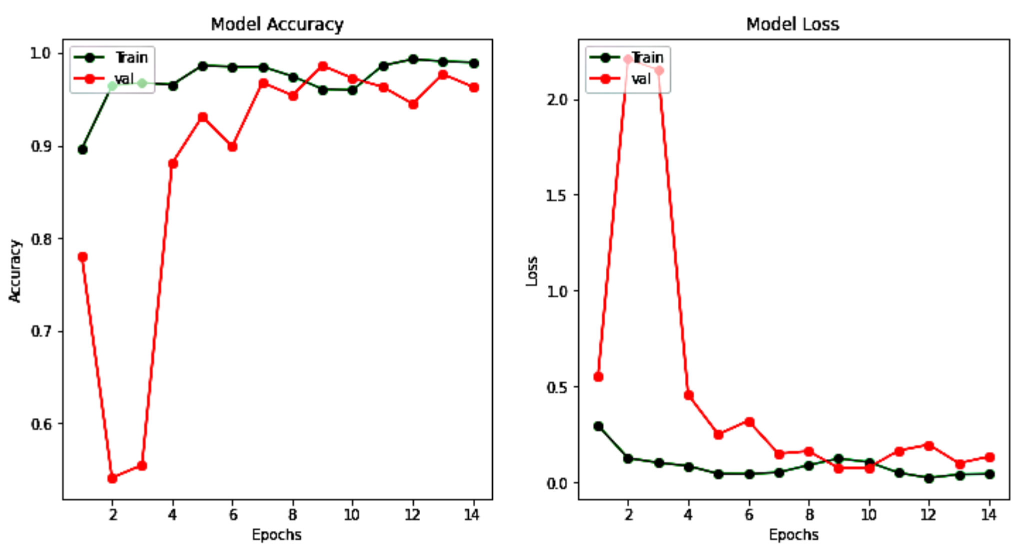

Figure 10: EfficientNetB1 validation and training accuracy and loss graphs.

MobileNet: Loss and accuracy graph of validation and training in figure 11. Through the sequence of 12 epochs, the accuracy progressively improves of the training set, simultaneously training loss on the left decreases constantly after each passing epoch. However, large jumps in validation accuracy in the beginning and afterward improves at an estimated rate. At the 8th epoch, it reaches its highest validation accuracy. Additionally, the validation loss drastically falls and keeps declining across the epochs.

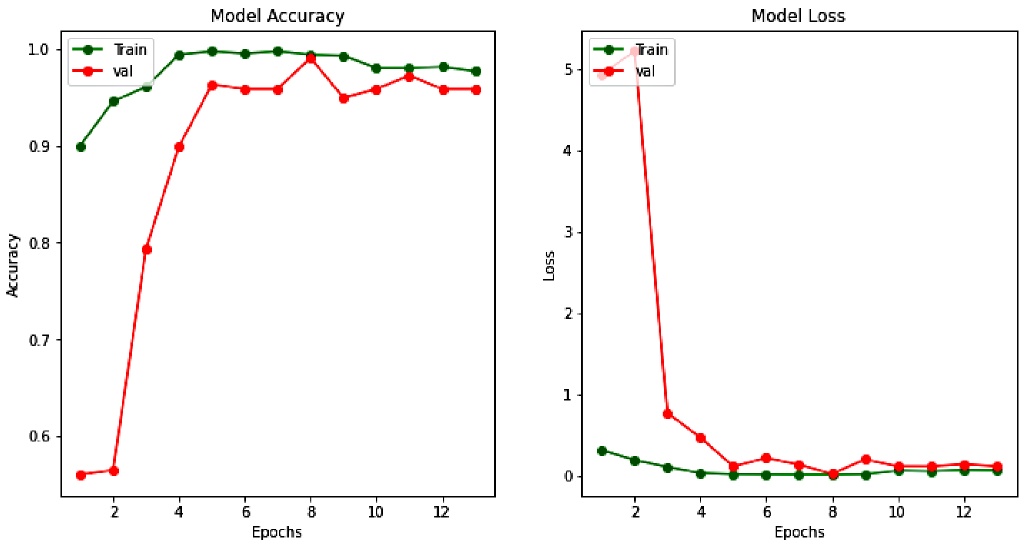

Figure 11: MobileNet validation and training accuracy and loss graphs.

EfficientNetV2B0: Validation accuracy, training accuracy, validation loss and training loss in figure 12. The accuracy in training of EfficientNetV2B0 increases and keeps increasing until the 8th epoch. It drops, then again keeps on increasing, reaches the highest accuracy which is just above 99% and after 12 epochs it fluctuates due to the overfitting problem. On the validation set, we can see there is a series of increases and fluctuations in accuracy. Both training and validation accuracy gets their highest accuracy at the 11th epoch. In addition, training loss drops at an increasing rate, then loss increases at the 8th epoch again fall and minimizes to less than 0.05. On the contrary, loss in validation drops at starting phase, and later inconsistently falls due to overfitting issues.

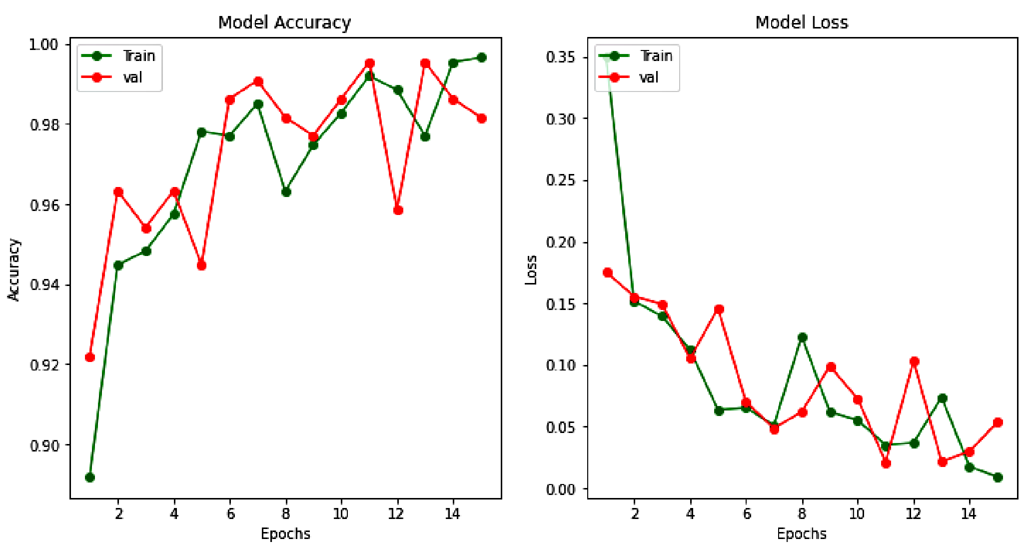

Figure 12: EfficientNetV2B0 validation and training accuracy and loss graph.

4.7 Model Evaluation

Table 1: Optimal results after model implementation

Models	Training Accuracy	Training Loss	Testing Accuracy	Testing Loss	Precision	Recall	F1-score
EfficientNetB4	0.9989	0.0105	0.9817	0.1254	0.98	0.99	0.98
EfficientNetB1	0.9931	0.0224	0.9862	0.0485	0.99	0.98	0.99
MobileNet	0.9977	0.0109	0.9908	0.0202	0.99	0.99	0.99
EfficientNetV2B0	0.9966	0.0093	0.9954	0.0175	0.99	1.00	1.00

The table shows the outcome of all the pre-trained models that are small in size and parameters, on the training set for accuracy and loss, and also on the testing set for accuracy, loss, precision, recall, and f1-score. From the training portion EfficientNetB4 model gives training accuracy of 99.89% along with training loss of 1.05%, whereas from the testing set, testing accuracy comes to 98.17%, loss value 12.54% with both precision and f1-score have a value of 98%, and 99% in the recall. However, EfficientNetB1 has a training accuracy of 99.31%, and training loss of 2.24%, but it performs better compared to EfficientNetB4 for testing set as testing accuracy is 98.62%, loss value minimizes to 4.85%, precision and f1-score 99% and recall reaches to 98%. MobileNet achieves the training accuracy of 99.77% with a training loss of 1.09%, on the other hand from testing data loss value drops to 2.02%, where testing accuracy reaches to 99.08% with similar precision, recall, and f1-score of 99%. Lastly, in EfficientNetV2B0 training accuracy attains 99.66% where training loss is the minimum, 0.93%. As a matter of fact, in testing data, the EfficientNetV2B0 model touches accuracy of 99.54% surpassing the previous best model MobileNet, with the very slightest loss of 1.75%. Also with a precision value of 99% and recall and f1-score of 100%.

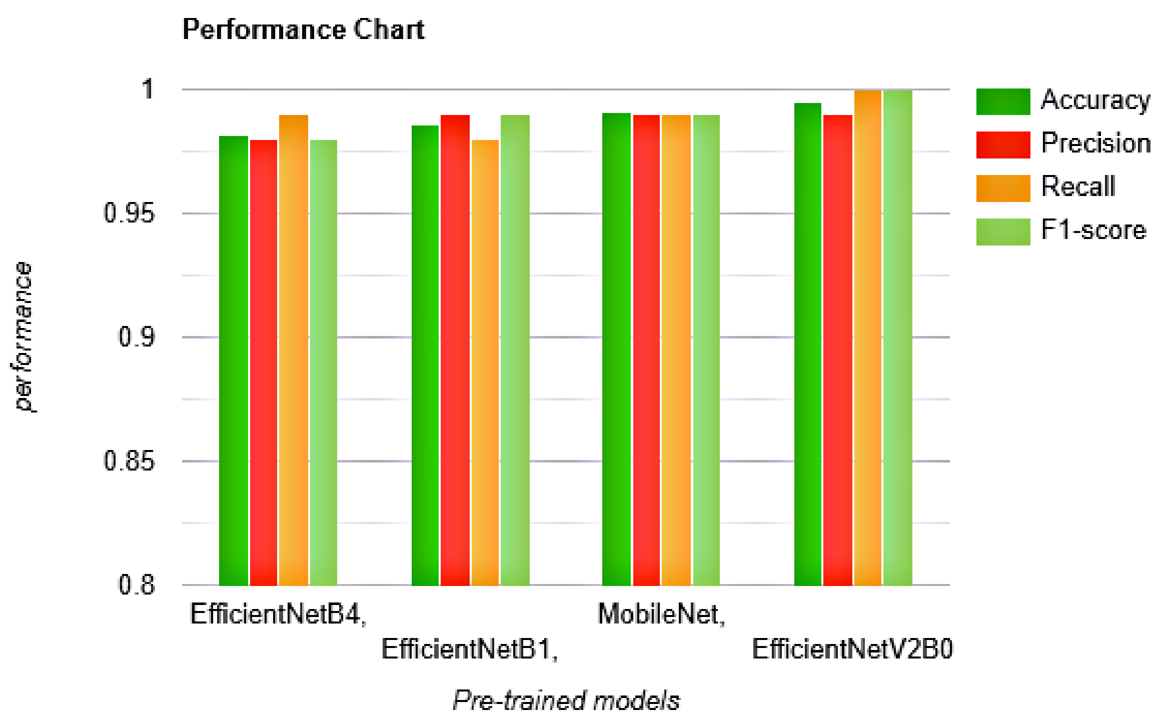

Figure 13: Comparison chart amongst the model from testing set

Figure 13 demonstrates the overview of the performance of the models in respect to the testing set. EfficientNetV2B0 performed better compare to all other models in every metrics. Here 99.54% is the highest achievable accuracy in testing set with the highest precision, recall and f1-score value.

4.8 Model Test

Afterwards, we ensured our study by creating a predictive system which has been made by saving the best-performed model in it. Next, a couple of images randomly has been taken from the testing set to clarify whether the model classifies rightly or not. Figure 14 shows, correctly categorizes slight or cloudier images as cataract and clear images as normal.

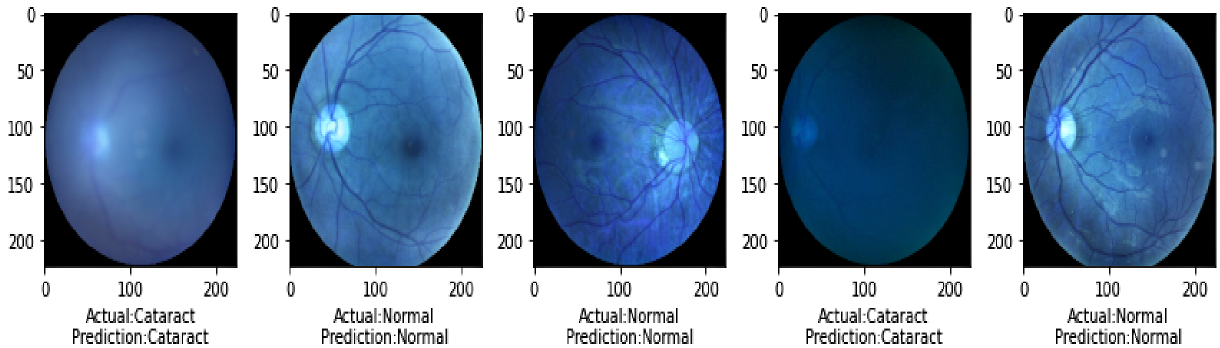

Figure 14: Screenshot of classification of Cataract and Normal eye images

5 Conclusion

This study's main objective was to discover a precise method for cataract image detection with funduscopic images. Other researchers have been fervently pursuing numerous approaches to identify cataract eye images. The importance of this study lies in the fact that it utilizes CNN's transfer learning models that are small in sizes due to a smaller number of parameters and has faster execution speed over many other deep learning models and attains outstanding results. In this study, cataract eye images are detected using four pre-trained models. The outcomes have been accurate and reliable. EfficientNetV2B0 achieved the finest in the metrics of F1-score, recall, precision and accuracy. Though, MobileNet was as good as well also. Although four of the models' outcomes were appreciable, among them EfficientNetB4 was the least performed model in the testing set compared to other models used in this study. Therefore, we can get rid of less skilled ophthalmologists for diagnosing cataract eye disease. It will also reduce time consumption, as it will give an immediate result based on the image. Furthermore, there are several practical challenges that can be solved and aid the job of clinicians easier.

For future work, there are many other models that are equally small in size and can be used for further illustration of performance. Altering hyperparameters or image processing techniques will also help to get more effective outcomes. Besides, the more complex and large datasets can also be used for further evaluation and better accuracy.

References

1. Cataracts - Symptoms and causes - Mayo Clinic, https://www.mayoclinic.org/diseases-conditions/cataracts/symptoms-causes/syc-20353790, last accessed 2022/08/15.
2. Flaxman, S.R.: Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. The Lancet Global Health. 5, e1221–e1234 (2017). https://doi.org/10.1016/S2214-109X(17)30393-5.
3. Foster, A.: Cataract and "Vision 2020—the right to sight" initiative. Br J Ophthalmol. 85, 635–637 (2001). https://doi.org/10.1136/bjo.85.6.635.
4. Allen, D.: Cataract and surgery for cataract. BMJ. 333, 128–132 (2006). https://doi.org/10.1136/bmj.333.7559.128.
5. Cataract Article, https://www.statpearls.com/ArticleLibrary/viewarticle/19005, last accessed 2022/08/15.
6. Zhou, Y.: Automatic Cataract Classification Using Deep Neural Network With Discrete State Transition. IEEE Trans. Med. Imaging. 39, 436–446 (2020). https://doi.org/10.1109/TMI.2019.2928229.
7. Huiqi Li.: A Computer-Aided Diagnosis System of Nuclear Cataract. IEEE Trans. Biomed. Eng. 57, 1690–1698 (2010). https://doi.org/10.1109/TBME.2010.2041454.
8. Mahmud Khan, M.: Cataract Detection Using Convolutional Neural Network with VGG-19 Model. IEEE World AI IoT Congress (AIIoT). 0209-0212 (2021). https://doi.org/10.1109/AIIoT52608.2021.9454244.
9. Dondapati, C.L.: Classification of Eye Disorders based on Deep Convolutional Neural Network. IJITEE. 9, 1388–1393 (2020). https://doi.org/10.35940/ijitee.F4209.049620.
10. Shaheen, I.: Survey Analysis of Automatic Detection and Grading of Cataract Using Different Imaging Modalities. Springer International Publishing. 35–45 (2019). https://doi.org/10.1007/978-3-319-96139-2_4.
11. Pratap, T., Kokil, P.: Computer-aided diagnosis of cataract using deep transfer learning. Biomedical Signal Processing and Control. 53, 101533 (2019). https://doi.org/10.1016/j.bspc.2019.04.010.
12. Sertkaya, M.E.: Diagnosis of Eye Retinal Diseases Based on Convolutional Neural Networks Using Optical Coherence Images. 2019 23rd International Conference Electronics. 1–5 (2019). https://doi.org/10.1109/ELECTRONICS.2019.8765579.
13. Qiao, Z.: Application of SVM based on genetic algorithm in classification of cataract fundus images. In: 2017 IEEE International Conference on Imaging Systems and Techniques (IST). 1–5 (2017). https://doi.org/10.1109/IST.2017.8261541.
14. Hossain, Md.R.: Automatic Detection of Eye Cataract using Deep Convolution Neural Networks (DCNNs). In: 2020 IEEE Region 10 Symposium (TENSYMP). 1333–1338 (2020). https://doi.org/10.1109/TENSYMP50017.2020.9231045.
15. Sirajudeen, A.: Detection of Cataract Through Feature Extraction by the Novel Angular Binary Pattern (NABP) and Classification by Kernel Based Convolutional Neural Networks. (2021). https://doi.org/10.21203/rs.3.rs-383419/v1.
16. Ocular Disease Recognition, https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k, last accessed 2022/08/15.
17. Tan, M..: EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. (2020). https://doi.org/10.48550/arXiv.1905.11946
18. Ahmed, T.: Classification and understanding of cloud structures via satellite images with EfficientUNet. Information and Computing Sciences. (2021).

https://doi.org/10.1002/essoar.10507423.1.
19. Dais, D.: Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning. Automation in Construction. 125, 103606 (2021). https://doi.org/10.1016/j.autcon.2021.103606.
20. Wang, W.: A New Image Classification Approach via Improved MobileNet Models with Local Receptive Field Expansion in Shallow Layers. Computational Intelligence and Neuroscience. (2020). https://doi.org/10.1155/2020/8817849
21. Tan, M., Le, Q.V.: EfficientNetV2: Smaller Models and Faster Training. (2021) https://doi.org/10.48550/arXiv.2104.00298
22. Junayed, M.S.: AcneNet - A Deep CNN Based Classification Approach for Acne Classes. In: 2019 12th International Conference on Information & Communication Technology and System (ICTS). 203–208 (2019). https://doi.org/10.1109/ICTS.2019.8850935.