DETECTION OF CATARACT USING CONVOLUTIONAL NEURAL NETWORKS

Martin Joel Rathnam[1],

Department of Electronics and Communication Engineering,

Sri Sairam College of Engineering, Bengaluru - 562106

Abstract

Among the most prevalent ocular diseases that distorts vision is cataract. The greatest strategy to reduce the danger and prevent vision is by prompt and effective identification of cataracts. Artificial intelligence-based cataract detection techniques are now the focus of study. Cataracts are one of the most prevalent visual conditions that people experience as their age. The cataract would be when the lens of the eyes develops a fog. The major clinical signs of this illness are vision changes, fading colours, and difficulties seeing in bright light. Many different tasks become challenging when these symptoms are present. Therefore, early cataract identification and prevention can aid in lowering the prevalence of impairment. Convolutional neural networks (CNN) used here to categorize cataract illness through publically available picture collection. Three distinct convolutional neural networks been used in this observation such as ResNet50, DenseNet121 and VGG 19 to assess the effectiveness of the proposed Model on the same dataset with Random forest classifier applied and TensorFlow framework. The performance of our model and the three pre-trained models (ResNet 50, DenseNet 121, and VGG-19) were evaluated and obtained through the metrics such as Precision, Recall (Sensitivity), Specificity, F1-Score and finally achieved 98% of accuracy.

Keywords: *Cataract detection, fundus images, neural network*

1 Introduction

Having a cataract causes the eyes to seem clouded. A cataract patient will have cloudy or icy vision. Having cataracts in one's eyes makes it difficult to read, drive, and even recognize faces given in [1]. The World Health Organization (WHO) estimates that there are 285 million visually impaired persons in the world, including 246 million persons who seem to be moderately to severely blind and 39 million people who really are blind [2]. According to the study conducted by Flaxman et al. in 2020, estimated that the number of people with moderate to severe vision impairment (MSVI) globally was 237.1 million, and the number of people who were blind was 38.5 million.

Additionally, the study predicted that the number of people who are blind would exceed 40 million by the year 2025. These numbers indicate the importance of continued efforts to prevent and treat vision impairment and blindness worldwide. In 2020, Flaxman *et al.* [4] predicted that the number of people suffering from Moderate to Severe Vision Impairment (MSVI) and blindness would be 237.1 and 38.5 million, respectively. The worldwide blindness will exceed 40 million by 2025 [5]. Depending upon how it manifests, cataract grouped into three categories: nuclear cataract [6], cortical cataract [7], and Posterior Sub Capsular (PSC) cataract [8]. Early cataract identification is crucial for effective therapy and thus can dramatically minimize the risk of impairment. Three elements make creating an automated service for cataract identification difficult: (i) the wide range of cataract lesions and human eye tones; (ii) the size, shape, and position of cataracts; and (iii) the dependency on age, race, and eye type.

Artificial intelligence-based systems for cataract detection have developed using a variety of feature extraction methods. Global features, such as discrete cosine transformation (DCT) [9], consider the entire image as a single entity and analyze it using mathematical transformations. Local features, on the other hand, focus on specific regions of the image and extract features based on properties such as local standard deviation [28]. Deep features, such as those extracted using deep Convolutional Neural Networks (CNNs), learned automatically from the raw data and can capture complex, high-level patterns.

Recent studies have shown that deep learning-based approaches, particularly those based on CNNs, have achieved higher accuracy in cataract detection than traditional methods based on handcrafted features. This is because deep learning models can learn features that are more representative of the underlying data distribution and can capture additional complex relationships between features. Furthermore, deep learning models can be trained end-to-end, which means that feature extraction and classification steps are learned jointly, leading to better performance overall. Despite the existence of many deep learning-based systems for automatically detecting cataracts have reported in the literature, they still have drawbacks, including low accuracy in detection, a high number of model parameters, and as a result, being computationally expensive.

2 Related Works

A computer-aided cataract identification method for routine examination either as a pre-processing phases in cataract rating was developed by Gao et al. [6] and introduced an improved texture feature and trained the Linear Discriminant Analysis with it (LDA). A medical databases study's study found an accuracy of 84.8 percent. Yang et al. [29] presented a three-step automated cataract detection approach. To increase the foreground/background contrast a top-bottom hat transformation had used. Features included the brightness and texture. The classifier build with Back Propagation Neural Network (BBNN), which was able to divide cataract severity into mild, medium, and severe phases. Guo et al. [9] published a computer-aided cataract categorization using on fundus pictures. With the help of wavelet transform and sketch-based techniques, the feature extraction process was completed. Following that, a multiclass discriminant analysis approach had used to identify and grade cataracts, with correct classification rates (CCRs) for feature extraction based on wavelet transforms being 90.9 percent and 77.1 percent, respectively, and for feature extraction based on sketches being 86.1 percent and 74.0 percent.

It is not an automated system since users or professionals manually clipped and extracted the pupil area during the processing stage. Yet, this approach is without the use of a slit light or fundus camera utilizing ocular pictures captured by a regular camera. By combining machine learning and ultrasound methods, Caixinha et al. [2] developed automated method to detect Nuclear Cataract and classification. In the classification stage, support vector machine (SVM), Bayes with multilayer perceptron for classification random forest classifiers used for examined, this has retrieved 27 characteristics with in time and frequency domains. While ultrasound-based cataract detection approaches are very accurate, the imaging technologies are pricy due to the challenging procedure. Using an SVM classifier, the fundus pictures in [10], had classified as cataract images, and RBF Network subsequently evaluated the severity with a precision of 93.33 percent. Sigit et al. [22] described an Android smart phone-based technique for cataract diagnosis. An algorithm called single-layer perceptron had used to classify the data.

A deep learning-based technique for assessing the severity of Nuclear Cataracts using slit-lamp pictures had examined by Gao et al. [7]. Through grouping the picture-patched input into a convolutional neural network (CNN), local filters were produced. To retrieve higher - level features, a group of Recursive Neural Networks (RNNs) had then deployed. Support vector regression had used for the cataract grading. An approach for six-level cataract grading based on a mixture of DCNN and Ran et al. [21] proposed Random Forests (RF). The suggested DCNN for extracting features at various levels from fundus pictures are made of three modules. From the other side, RF implemented a more complex six-level cataract grading using a feature dataset that DCNN developed and utilized. This approach has a 90.69 percent accuracy rate on average. The professionals may be better able to comprehend the patients' conditions with the use of this six-level grading system.

Pratap and Kokil [19] presented computer-aided approach based on fundus pictures, for classifying cataract severity stages. This technique used transfer learning which automatically classify cataracts using a pre-trained CNN. Utilizing SVM classifier with four stages of CCR to obtain the classification as 92.91 percent. A binary CNN model and tournament structure had used by Jun et al. [13] as a method for assessing cataracts. Khan et al. [14] used a dataset recently released by KAGGLE and applied for VGG-19 model with a transfer learning technique. Additionally latest study looked at cataract detection in a loud setting by Pratap and Kokil [20]. For feature extraction, a collection of independently trained, (locally and globally) support vector networks was combined into a pre-trained CNN.

The data attained demonstrated its noise resistance. The resilience of the cataract detection devices was the subject of the first study. It was noted that while numerous works had been conducted using traditional machine learning techniques, only a small number of deep learning-based works on cataract identification and grading had been published. As a result, there are still a number of issues to resolve, including increasing the models' accuracy while lowering their complexity by lowering the quantity of training parameters, layers, depth, running time, and total model size. Discrete Meyer filters have broad application, which detects the connecting vessels in the image and helps in extraction of Region Of Interest. The Discrete Meyer transform has the advantage of rapid decay along with infinite differentiability and it gives compact support in frequency domain, which is the main reason to select this wavelet transform [22].

3 Proposed Method

3.1 Pre-Processing

The suggested system dataset comprises images of individuals with normal conditions as well as those with diabetes, glaucoma, cataracts, pathological myopia, hypertension, age-related macular degeneration, and other illnesses/abnormalities. Therefore, we segregated every fundus image taken in the first phase, with the exception of cataract and regular fundus photographs. Labels filtered the data. Experimental fundus photos have diverse image sizes since they take with various cameras. As a result, we reduced the image's size to 224×224 pixels using OpenCV. Upon import, the dataset is changed into an array format for training using the NumPy library.

3.2 CNN Model

A Convolutional Neural Network (CNN) is a type of deep neural network that are commonly used for image and video recognition, natural language processing, and other applications that involve structured data. The key feature of a CNN is the use of convolutional layers, which apply filters to the input data to extract relevant features. These filters learned through the training process, allowing the network automatically identify patterns in the data. In addition to convolutional layers, CNNs also typically include pooling layers, which reduce the size of the feature maps and help to make the network more computationally efficient. Non-linear activation functions, such as ReLU (Rectified Linear Unit), used to introduce non-linearity into the model, allowing it to learn more complex representations of the input data.

By stacking multiple layers of convolutional, pooling, and activation functions, CNNs are able to build a hierarchical representation of the input data, with each layer learning increasingly complex features. This makes CNNs highly effective at tasks such as image classification, where they are able to learn to identify objects and patterns in images with remarkable accuracy [23], [24]. Deep learning methods integrate feature extraction and classification, while manual feature extraction techniques divide these processes into separate phases. The classification of the cataract photos from the database is done using an unique deep learning model that makes use of the Random Forest classifier. Lowering the prevalence of impairment, it aided by early detection and prevention of cataracts. The TensorFlow object identification framework had utilized to apply CNNs in this observation through ResNet50, DenseNet128, and VGG 19 with Random Forest Classifier. The maximum accuracy rate is achieved using the random forest classifier are combined in deep learning-based algorithms. Figure 1 depicts the process for selecting the best model and procedure for cataract or Normal image detection.

After pre-processing, the image fitted to the best-trained model for diagnosis, which then indicates whether the image exhibits cataract disease. CNN is a deep neural network that uses convolutional and pooling layers as well as non-linear activation functions to build a complicated hierarchy of features. The phase of feature extraction and the process of categorization. By using different hyperparameters throughout the training and validation phases, the model may predict the illness from the raw data. Because we are using the sigmoid activation function, it is necessary to utilize the binary cross-entropy loss. To address overfitting problems encountered during training, the optimizer chosen was Adam with a low learning rate of 0.0001. Fig. 3 displays the whole architecture of my CNN.

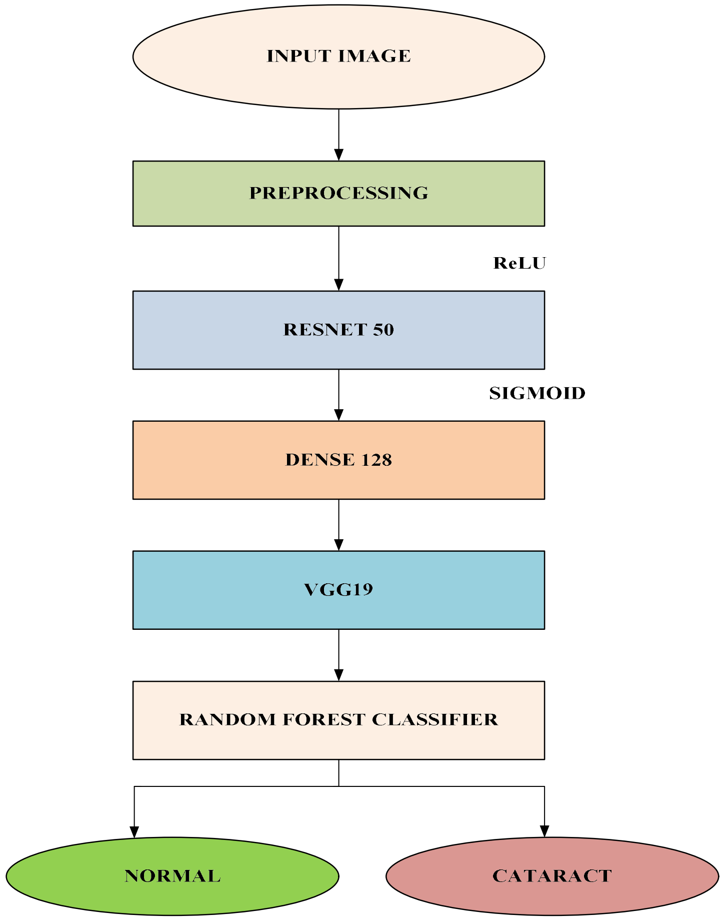

Figure 1.Proposed CNN Model

3.3 ReLU

The rectified linear activation unit (ReLU) [24] is a significant milestone in the advancement of deep learning and considered a crucial element of this field. It is straightforward yet far better to earlier activation mechanisms like the sigmoid and provided as

$$f(input) = \max(0, input) \tag{1}$$

The result of ReLU is determined by the formula that returns the largest value among 0 and the input value. The output of the function will be 0 for negative input values and identical to the input value for positive input values.

3.4 Sigmoid Function

The sigmoid function [4] is a mathematical function that transforms any real number into a value between 0 and 1, forming an "S" shaped curve. This function also referred as the logistic function. The equation for the sigmoid function is

$$X = \frac{1}{1+e^{-y}} \tag{2}$$

The sigmoid function's key benefit is that it may found between the coordinates ‘0’ and ‘1’. It is hence particularly useful in models where we need to foresee probability as an outcome. The probability of anything happening is limited between 0 and 1, therefore picked this function.

3.5 ResNet 50

Multiple variations of ResNet exist that employ the same fundamental concept but with varying numbers of layers. The ResNet version that can handle 50 neural network layers, commonly referred to as ResNet50. The architecture (Resnet50), based on the concept mentioned above; however, there is one significant distinction. Concerns about the duration of layer training prompted the utilization of a bottleneck design for building blocks. This design employs a stack of three layers instead of the previous two. To create the ResNet50 model, each two-layer block in ResNet34 had replaced with a three-layer bottleneck block. The ResNet-50 model, consisting of a convolutional neural network with 50 layers, has significantly higher accuracy than the ResNet34 model. ResNet, or Residual Networks, is a widely used neural network architecture that forms the basis for numerous computer vision applications. The major advancement introduced by ResNet was its ability to train neural networks with over 150 layers.

3.6 DenseNet121

DenseNet121, a convolutional neural network that does deep analysis and produces straightforward output, is more effective than DenseNet. Each layer in Dense-Net121 has a connection to the layers before it, even without the starting layer. Each layer's output had fed as input to the subsequent layer, creating a direct connection between every layer.DenseNet121, a convolutional neural network that does deep analysis and produces straightforward output, is more effective than DenseNet. Each layer in Dense-Net121 has a connection to the layers before it, even without the starting layer. The output of one layer serves as an input for the following layer. Each layer connected directly to the following. It is designed primarily to increase the accuracy drop brought on by high-level neural networks' vanishing gradient. Dense nets are convolutional networks that are highly linked. With some essential changes, it is extremely similar to a ResNet.

DenseNet, designed to address the problem of vanishing gradient in deep neural networks, which occurs due to the significant distance between the source and destination nodes, causing the information to fade away before reaching its destination. The aim of DenseNet is to improve accuracy in such scenarios. ResNet uses an additive method, which implies those who start taking a prior outcome as that of an input for a future layer, whereas DenseNet involves all preceding outcome as an input for just a subsequent layer.

3.6 VGG-19

Several advances in picture categorization have made possible with the use of deep neural networks. These sophisticated models have also helped several other visual recognition tasks. As a result, we often enhance our understanding of problems, find solutions to them more effectively, and increase our accuracy. But when we go further into neural networks, accuracy suffers and training becomes more difficult. These problems being solve by using VGG-19 by Oda et al. [2020]. The VGG-19 model is a CNN-based model that uses 3x3 filters with a single stride and consistent padding. It also employs max pooling layers with 2x2 filters and a stride of 2. Unlike other models, VGG-19 contains a smaller number of hyperparameters.

The convolution and max pooling layers arranged similarly and the model contains two FC levels. The VGG-19 network is a sizeable model with approximately 138 million parameters that to be trained. The layout of the VGG-19 network shown in Figure 2. Following the classification layer, which included comprised a dropout layer and a densely connected classifier, many convolutional layers were used (conv1, conv2, conv3, conv4, and conv5). Every neuron in a dense layer linked to every neuron in the layer above it. A densely connected layer, as opposed to a convolutional layer, picks up information from the characteristics of the preceding layer. It must be defined how the tightly linked layer is enabled.

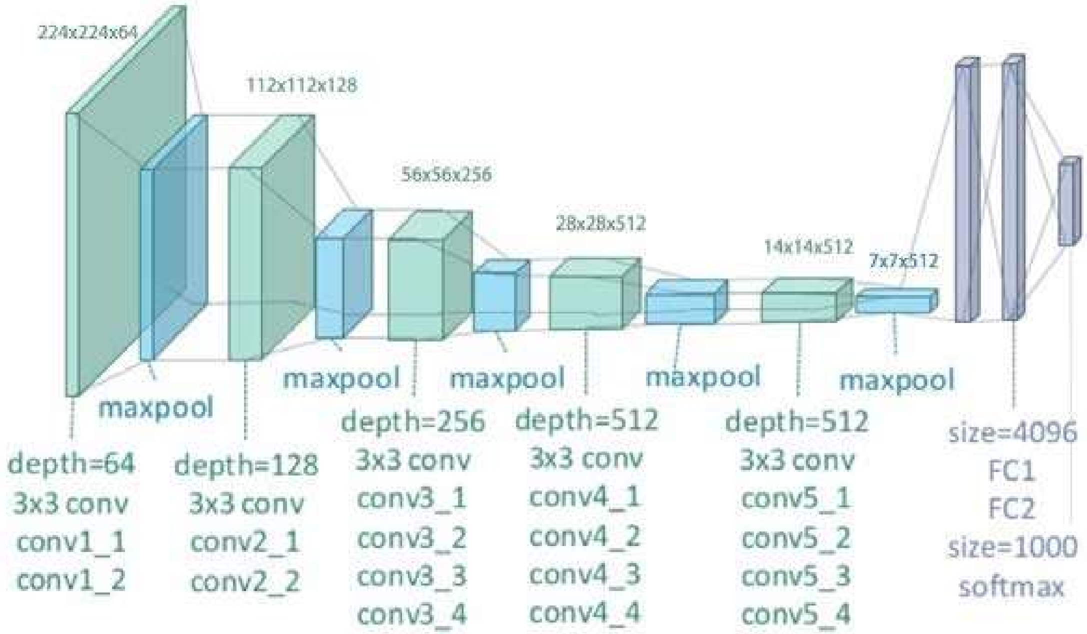

Figure 2. VGG-19 network architecture

4 Experimental Set Up

4.1 Dataset

The 1088 fundus images used in this recommended method to make up the dataset. The Ocular Disease Intelligent Recognition (ODIR) database is an organized ophthalmic dataset that includes data on 5000 individuals, such as their ages, diagnostic keywords provided by their physicians, and colour fundus photographs of their left and right eyes. For our purposes, only cataracts and common fundus images used from the dataset. The patients are categorized into eight different labels, which include Normal (N), Diabetes (D), Glaucoma (G), Cataract (C), AMD (A), Hypertension (H), Myopia (M), and Other diseases/abnormalities (O) also given in Table 1 and Figure 3 and Figure 4.The Python-based deep learning software package Keras is well-liked for its ease of use and adaptability. The interface between TensorFlow and Theano, two machine-learning systems, had provided by this open-source library.

The key benefit of deep learning with Keras is its incredible speed, which makes it possible to create and train a neural network model with very little code. As a result, Keras frequently takes 50% less code than native APIs of deep learning frameworks to create a model. Keras support convolutional and recurrent networks, as well as hybrid networks. The experiment had performed on a computer with the following specifications: an i9-10850K CPU, 64GB RAM, and a NVIDIA Geforce RTX 2080 super GPU operating at 3.60 GHz. The image pre-processing, augmentation, and implementation of the CNN-based model had conducted using Python, Keras, and Tensorflow environments.

Table 1. The ODIR dataset includes a variety of images distributed across different categories or labels.

Different classes in ODIR	Name of the disease	Training cases
1	Normal (N)	1135
2	Diabetes (D)	1131
3	Glaucoma (G)	207
4	Cataract (C)	211
5	Age-related macular degeneration (A)	171
6	Hypertension (H)	94
7	Pathological myopia (M)	177
8	Other diseases/abnormalities (O)	944

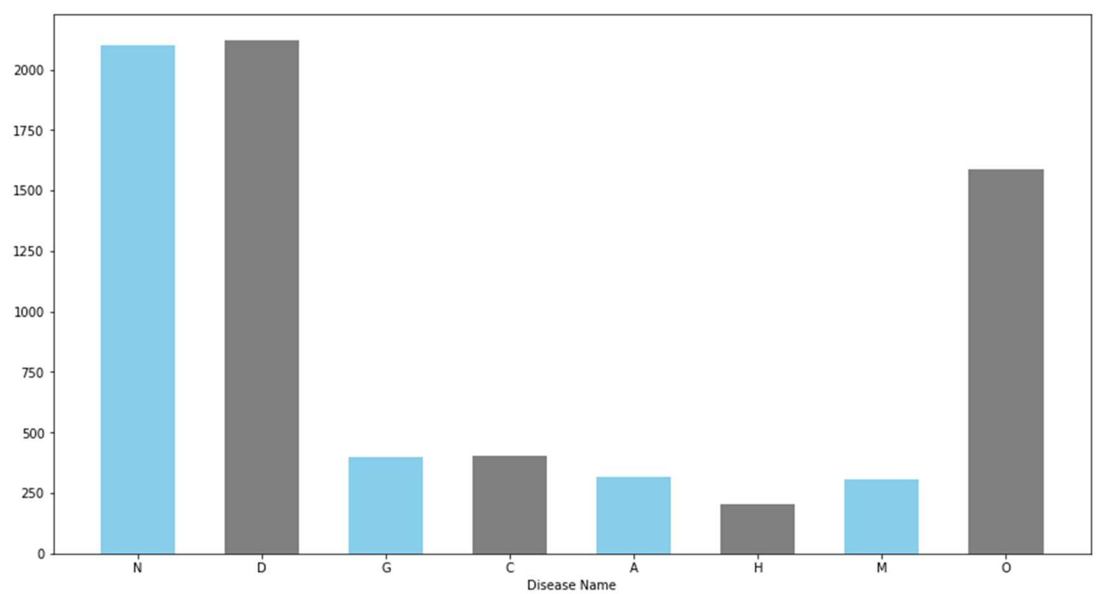

Figure 3. Distribution of Various diseases from the dataset

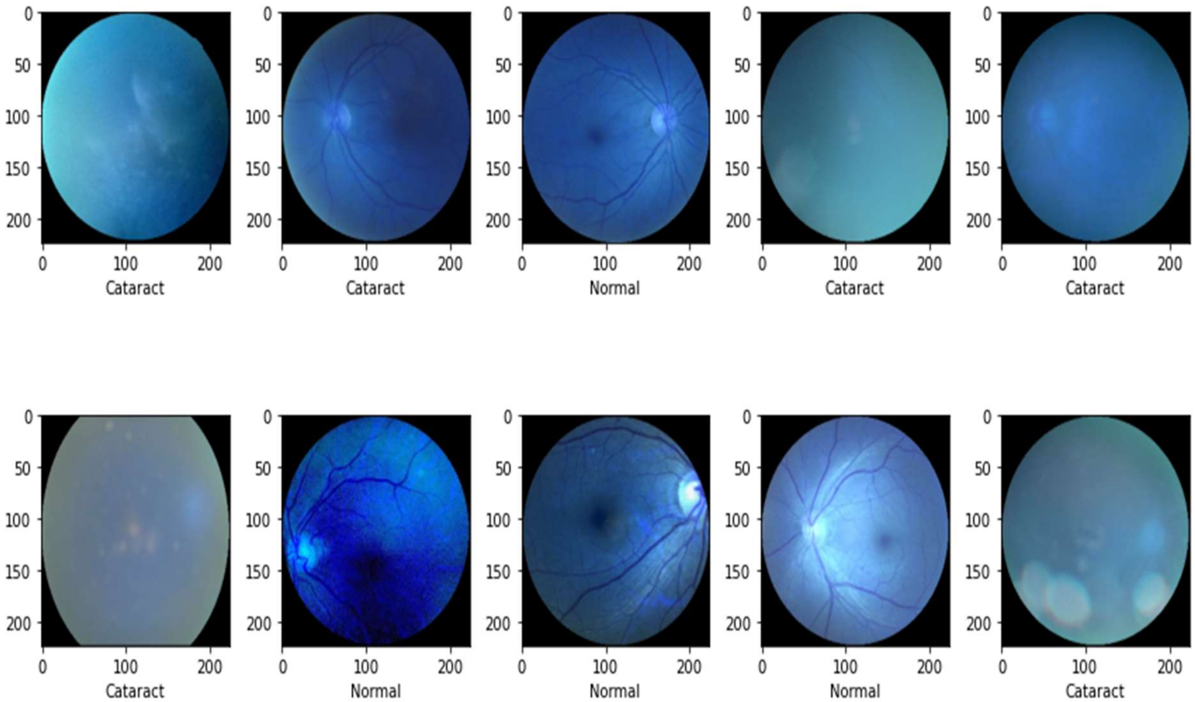

Figure 4.Creating Dataset from images

4. 2 Hyperparameter Evaluation Metrics

In this work, the model had evaluated using the following performance metrics to assess their performance:

4.2.1 Accuracy

Accuracy is a key criterion for evaluating classification models. It refers to the percentage of correct predictions made by the model. Binary classification accuracy had calculated using the following formula, which takes into account both true positives and true negatives:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

4.2.2 Recall

To calculate the recall, we divide the number of True Positives (TP) by the sum of True Positives and False Negatives (FN). The True Positive Rate or Sensitivity also referred to as the number of positive predictions divided by the number of positive class values in the test data.

$$Sensitivity / \mathrm{Re}\,call = \frac{TP}{TP + FN} \tag{4}$$

4.2.3 Precision

Precision, also known as Positive Predictive Value, is a metric that is determined by dividing the sum of True Positives and False Positives (FP) by the total number of True Positives. In simpler terms, it measures the proportion of positive predictions made by a model in relation to the total number of expected positive class values.

$$\Pr ecision = \frac{TP}{TP + FP} \tag{5}$$

4.2.4 Specificity

Specificity refers to the proportion of true negatives that correctly predicted as negatives. In contrast, small percentage of true negatives that mistakenly predicted as positives, resulting in false positives. This ratio is also known as the "false positive rate." The total of specificity and false positive rate is always equivalent to one.

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

4.2.5 F1-Score

The F1 score is a widely used metric for evaluating classification tasks, particularly when both accuracy and recall are important. It is used to assess the performance of binary classification algorithms, which classify items into positive or negative categories. The primary objective is to achieve accurate predictions, regardless of whether they are positive or negative. False predictions, on the other hand, been minimized. The F1 score is a measure of a model's accuracy and recall, and it computed by the following formula:

$$F1score = \frac{2 \times \Pr ecision \times \operatorname{Re} call}{\Pr ecision + \operatorname{Re} call} \tag{7}$$

4.2.6 Confusion matrix

It represents the chart, which is the most beneficial for understanding the model's performance details. Confusion matrix, used to compute significant predictive metrics including recall, specificity, accuracy, and precision. Since it make it simple to compare variables such True Negatives, False Negatives (FN), True Positives, and False Positives, confusion matrix are advantageous. This became important in this research because researchers wanted to guarantee accuracy and memory. We wanted to identify contaminated images with little to no misclassification, and the InceptionResnetV2 model allowed us to do this. As a result, researchers are capable of determining the specificity and sensitivity/recall of the total model performance.

5 Results and Discussions

In this section, we evaluate the effectiveness of the proposed VGG 19 model with a Random Forest classifier in detecting cataracts using the dataset shown in Figure 5. We compare our experimental results with other state-of-the-art methods for cataract detection. The confusion matrix of the VGG19, is shown in Figure 6. The matrix displays the predicted and true labels of the images, where each row and column corresponds to a true and predicted label, respectively. The confusion matrix indicates that there were only five misclassifications in the total of 218 images in our test set. These misclassifications were due to healthy eyes were incorrectly labelled as cataracts and vice versa.

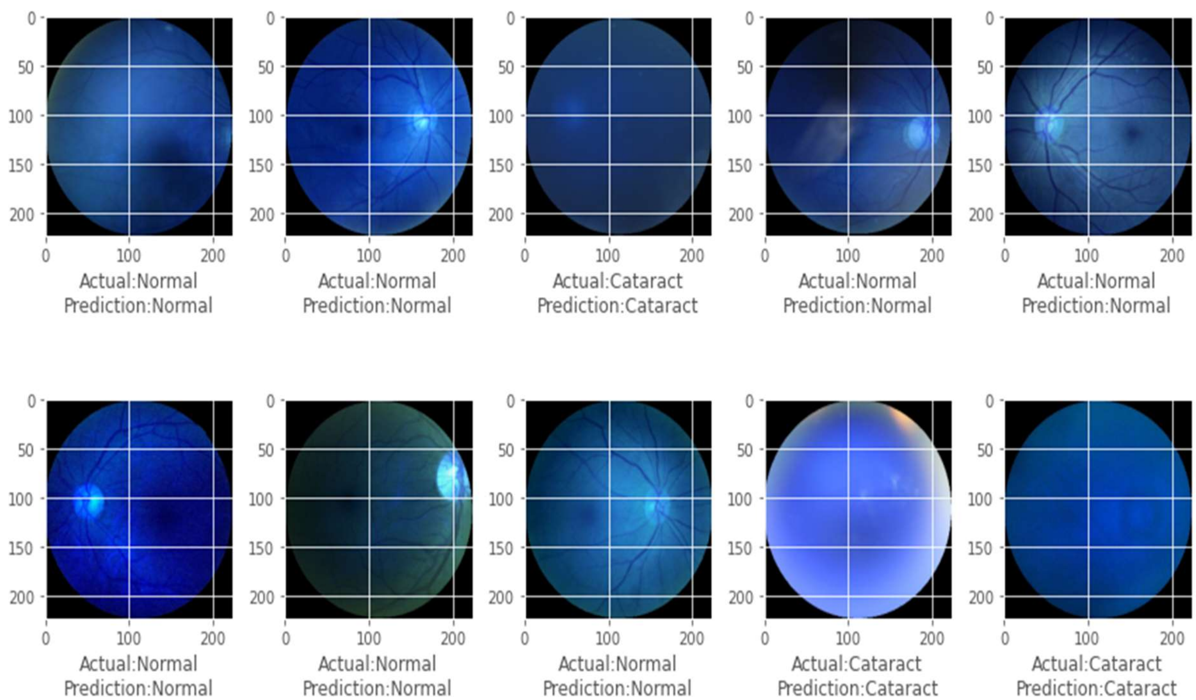

Figure 5. Prediction of normal (N) and cataract (C) classes with VGG-19 model

Dividing dataset into x (features) & y (target) and Creating VGG19 Model and obtained the high outcomes which is shown in table below

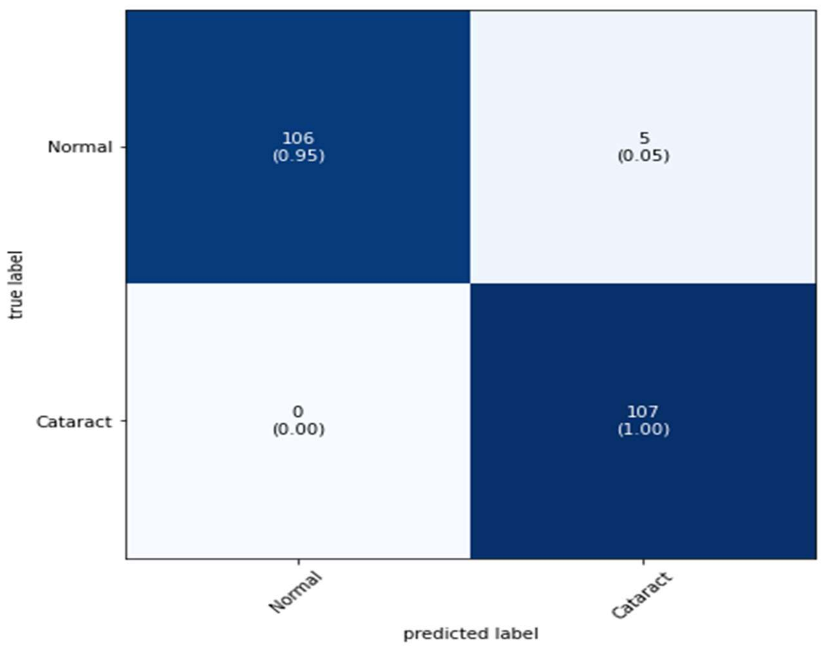

Figure 6. Confusion Matrix for VGG19

Figure 7 displays the training and validation accuracy and loss graphs of the VGG 19 model. The X-axis represents the epochs, and the Y-axis represents the accuracy or loss values. This figure demonstrate that, after 9 iterations, the validation accuracy and loss level out and become parallel to training, which is sufficient to recreate the pattern with slight irregularities. This figure shows that after 9 epochs, the validation accuracy and loss stabilize and become parallel to the training data.

This suggests that the model has learned the pattern of the data, with some slight irregularities. We have created three pre-trained Convolutional Neural Network (CNN) models, which have demonstrated excellent classification accuracy. We will use these models, namely ResNet 50, DenseNet 121, and VGG-19, to assess the effectiveness of our proposed Model on the same dataset. Table 2 displays a comparison between the performance of our model and the three pre-trained models (ResNet 50, DenseNet 121, and VGG-19) based on several evaluation metrics, including Precision, Recall (Sensitivity), Specificity, Accuracy, and F1-Score. In addition, Table 3 shows the performance of VGG-19, while Table 4 compares our proposed work with existing methods.

Table 2 Performance comparison of the models

Models	Accuracy (%)	Precision (%)	Recall (Sensitivity) (%)	F1-Score (%)
ResNet 50	87	85	89	87
DenseNet 121	93	91	95	92
VGG-19	**98**	**96**	**100**	**98**

Table 3 Performance of VGG19

	Precision (%)	Recall (%)	F1 Score (%)	Support
Normal	100	95	98	111
Cataract	96	100	98	107
Micro Average	98	98	98	218
Weighted Average	98	98	98	218
Accuracy (%)			98	218

Table 4 Comparison of Existing work with the proposed work

Compared with other work	**Models**	**Accuracy (%)**
Singh Gill et.al (2022), [26]	VGG16 EfficientNetB5	81.76 87.25
Gill et.al (2022), [27]	EfficientNetB6 DenseNet169	86.52 86.76
Rajalakshmi et.al (2022) ,[24]	VGG16	86.28
Proposed work	**VGG19**	**98**

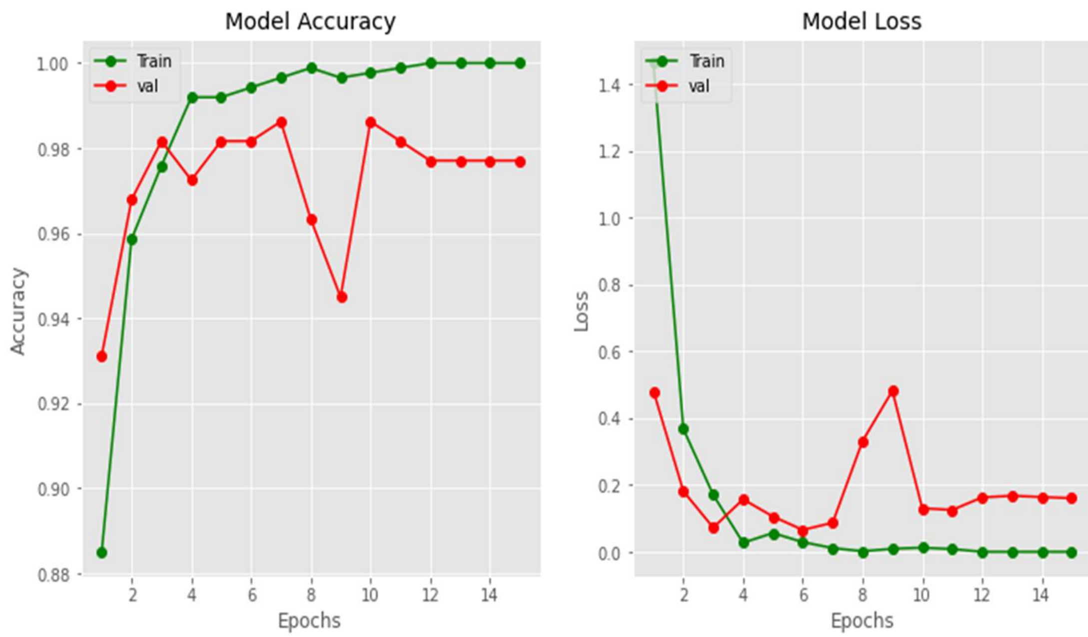

Figure 7. The training and validation accuracy and loss graphs of the VGG 19

6 Conclusion

The proposed method had performed using CNN. The model used is VGG19. Performance of this model in detecting cataract is 98%. Here Normal or cataract had detected from the images present in the database. The performance of the proposed model had verified with other existing models. Various analysis has given to prove the superiority of the proposed model.

ACKNOWLEDGMENT

Funding not Applicable

References

1. Allen Foster, F. F. (2000) ‘Vision 2020: the cataract challenge’, Community Eye Health, vol. 13, no. 34, pp. 17–19.

2. Caixinha, M., Amaro, J., Santos, M., Perdigão, F., Gomes, M. and Santos J. (2016) ‘In-vivo automatic nuclear cataract detection and classification in an animal model by ultrasounds’, *IEEE Trans. Biomed. Eng.*, vol. 63, no. 11,pp. 2326-2335.

3. Cao, L., Li, H., Zhang, Y., Zhang, L. and Xu, L. (2020) ‘Hierarchical method for cataract grading based on retinal images using improved Haar wavelet’, *Inf. Fusion*, vol. 53, pp. 196-208.

4. Fan, W., Shen, R., Zhang, Q., Yang, J. and Li, J. (2015) ‘Principal component analysis based cataract grading and classification’, in *Proc. 17th Int. Conf.E-Health Netw., Appl. Services (HealthCom)*, pp. 459-462.

5. Flaxman, S. R., Bourne, R. R. A., Resnikoff, S., Ackland, P., Braithwaite,T., Cicinelli, M. V., Das, A., Jonas, J. B., Keeffe, J., Kempen, J. H., Leasher, J., Limburg, H., Naidoo, K., Pesudovs, K., Silvester, A., Stevens, G. A., Tahhan, N., Wong, T. Y. and Taylor, H. R. (2017) ‘Global causes of blindness and distance vision impairment 1990-2020:Asystematic review and metaanalysis’, *Lancet Global Health*, vol. 5, no. 12, pp. e1221-e1234.

6. Gao, X., Li, H., Lim, J. H. and Wong, T. Y. (2011) ‘Computer-aided cataract detection using enhanced texture features on retro-illumination lens images’, in *Proc. 18th IEEE Int. Conf. Image Process.*, pp. 1565-1568.

7. Gao, X., Lin, S. and Wong, T. Y. (2015) ‘Automatic feature learning to grade nuclear cataracts based on deep learning’, *IEEE Trans. Biomed. Eng.*, vol. 62, no. 11, pp. 2693-2701.

8. Gao, X., Wong, D. W. K., Ng, T., Cheung, C. Y. L., Cheng, Y. and Wong, T. Y. (2012) ‘Automatic grading of cortical and PSC cataracts using retro illumination lens images’, in *Proc. Asian Conf. Comput. Vis.* Berlin, Germany: Springer, pp. 256-267.
9. Guo, L., Yang, J., Peng, L., Li, J. and Liang, Q. (2015) ‘A computer aided healthcare system for cataract classification and grading based on fundus image analysis’, *Comput. Ind.*, vol. 69, pp. 72-80.

10. Harini, V. and Bhanumathi, V. (2016) ‘Automatic cataract classification system’, in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, pp. 0815-0819.

11. Hu, S., Wang, X., Wu, H., Luan, X., Qi, P., Lin, Y., He, X. and He, W. (2020) ‘Unified diagnosis framework for automated nuclear cataract grading based on smartphone slit-lamp images,” *IEEE Access*, vol. 8, pp. 174169-174178.

12. Kaggle Ocular Disease Recognition dataset

13. Kim, D., Jun, T. J., Eom, Y., Kim, C. and Kim, D. (2019) ‘Tournament based ranking CNN for the cataract grading’, in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, pp. 1630-1636.

14. Khan, M. S. M., Ahmed, M., Rasel, R. Z. and Khan, M. M. (2021) ‘Cataract detection using convolutional neural network with VGG-19 model’, in *Proc. IEEE World AI IoT Congr. (AIIoT)*, pp. 0209-0212.

15. Martin Joel Rathnam. and Jobin Christ. (2022) '"A Novel Method for Cataract detection and segmentation Using Nakagami Distribution', *Journal of Medical Imaging and Health Informatics*, Vol. 12, no. 1, pp. 45-51.

16. MayoClinicStaff,cataractshttps://www.mayoclinic.org/diseases conditions/cataracts/symptoms-causes/syc-20353790.

17. Niya, C. P. and Jayakumar, T. V. (2015) 'Analysis of different automatic cataract detection and classification methods', in *Proc. IEEE Int. Advance Comput.Conf. (IACC)*, pp. 696-700.

18. Oda, M., Yamaguchi, T., Fukuoka, H., Ueno, Y. and Mori, K. 'Automated Eye Disease Classification Method from Anterior Eye Image Using Anatomical Structure Focused Image Classification Technique', https://arxiv.org/abs/2005. 01433.

19. Pascolini, D. and Mariotti, S. (2012) 'Global estimates of visual impairment: 2010', *Brit. J. Ophthalmol.*, vol. 96, no. 5, pp. 614-618.

20. Pratap, T. and Kokil, P., (2019) 'Computer-aided diagnosis of cataract using deep transfer learning', *Biomed. Signal Process. Control*, vol. 53, Art. no. 101533.

21. Pratap, T. and Kokil, P. (2021) 'Efficient network selection for computer-aided cataract diagnosis under noisy environment', *Comput. Methods Programs Biomed.*, vol. 200, Art. no. 105927.

22. Ramani, G. and Menakadevi, T. (2022) 'Detection of Diabetic Retinopathy using Discrete Wavelet Transform with Discrete Meyer in Retinal Images', *Journal of Medical Imaging and Health Informatics*, Vol. 12, no. 1, pp. 62-67.

23. Ran, J., Niu, K., He, Z., Zhang, H. and Song, H. (2018) 'Cataract detection and grading based on combination of deep convolutional neural network and random forests', in *Proc. Int. Conf. Netw. Infrastruct. Digit. Content (IC-NIDC)*, pp. 155-159.

24. Rajalakshmi, M., Saravanan, V., Arunprasad, V., Khalaf, O. and Karthik C. (2022) 'Machine learning for modeling and control of industrial clarifier process', *Intelligent Automation & Soft Computing*, vol. 32, no. 1, pp. 339–359.

25. Sigit, R., Triyana, E. and Rochmad, M. (2019) 'Cataract detection using single layer perceptron based on smartphone', in *Proc. 3rd Int. Conf. Informat. Comput. Sci. (ICICoS)*, pp. 1-6.

26. Singh Gill, H., Ibrahim Khalaf, O., Alotaibi, Y., Alghamdi, S. and Alassery, F.(2022) 'Fruit image classification using deep learning', *Computers, Materials & Continua*, vol. 71, no. 3, pp. 5135–5150.

27. Singh Gill, H., Ibrahim Khalaf, O., Alotaibi, Y., Alghamdi, S. and Alassery, F. (2022) 'Multi-model CNN-RNN-LSTM based fruit recognition and classification', *Intelligent Automation & Soft Computing*, vol. 33, no. 1, pp. 637–650.

28. Xiong, L., Li, H. and Xu, L. (2017) ‘An approach to evaluate blurriness in retinal images with vitreous opacity for cataract diagnosis’, *J. Healthcare Eng.*, vol. 2017, pp. 1-16.

29. Yang, M., Yang, J., Zhang Q., Niu, Y. and Li, J. (2013) ‘Classification of retinal image for automatic cataract detection’, in *Proc. IEEE 15th Int. Conf. e-Health Netw., Appl. Services (Healthcom)*, pp. 674-679.