Frontiers in Health Informatics www.healthinformaticsjournal.com
ISSN-Online: 2676-7104
2024; Vol 13: Issue 2 Open Access

Cataract Diseases Prediction Using Deep Learning

[1]Jayashri Bagade, [2]Sanjeevkumar Angadi, [3]Saili Sable, [4]Amol Golhar, [5]Harshad Bankar, [6]Sanket Bendale,

[1]Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, India, jayashree.bagade@viit.ac.in , ORCID: 0000-0002-9709-5530

[2]Department of Computer Science and Engineering, PCET'S & NMVPM'S Nutan College of Engineering and Research, Pune, India, angadi.sanjeevkumar@gmail.com, ORCID:0000-0002-2925-6777

[3]Department of Computer Science and Engineering, PCET'S & NMVPM'S Nutan College of Engineering and Research, Pune, India, sailisable@gmail.com, ORCID:0009-0009-2590-5862

[4]Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, India, amol.22110428@viit.ac.in , ORCID:0009-0002-3833-3268

[5]Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, India, harshad.22111291@viit.ac.in ORCID0009-0004-7076-0547

[6]Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, India, sanket.22110046@viit.ac.in ORCID:0009-0001-9604-8281

Article Info

Article type:

Research

Article History:

Received: 2024-03-20

Revised: 2024-05-06

Accepted: 2024-06-02

Keywords:

Cataract disease prevention, Convolutional neural network , SqueezeNet , Healthcare Innovation

ABSTRACT

This research focuses on cataract disease diagnosis and prediction using machine learning (ML) models. It seeks to solve the difficulties associated with predicting cataract illness by fusing medical necessity with technology breakthroughs. The complex problem statement includes issues including the lack of data, the lack of resources, and the requirement for preventative healthcare actions. By recognizing and addressing these issues, the article seeks to establish the groundwork for a strong and reliable cataract prediction paradigm. The study investigates the use of several machine learning models, such as ensemble approaches, decision trees, and neural networks, and assesses how well they predict the occurrence of cataracts. This methodology advances cataract prediction while also providing a more comprehensive investigation of machine learning applications in healthcare. The application of several convolutional neural network (CNN) models and their analysis. It explores the technical details of the models, such as padding, normalization, kernel sizes, and filter counts. By fusing technology breakthroughs with healthcare requirements, the project aims to contribute to the changing field of healthcare innovation by providing a more nuanced understanding of cataract prevalence and a route toward more effective, focused, and compassionate healthcare interventions. The paper discusses the effectiveness of several convolutional neural network (CNN) models for classifying cataract eye images. A comparison is made between the accuracy of multiple models: KNN, Inception V3, Xception, LeNet-CNN, and the proposed SqueezeNet model. This work emphasizes how important it is to choose the right CNN architecture for cataract eye classification and highlights the possibility for improving diagnostic capabilities in clinical situations. The suggested SqueezeNet model achieved the best classification accuracy for cataract eye pictures, demonstrating its applicability in this specific use case. This highlights the importance of choosing the right CNN architecture for cataract eye categorization as well as the possibility of improving diagnostic capabilities in clinical situations.

1. INTRODUCTION

Cataract disease, characterized by the progressive clouding of the eye's lens, stands as a pervasive global health concern, affecting millions and impairing vision worldwide. Addressing this challenge necessitates a nuanced understanding of its prevalence, spatial distribution, and the application of innovative solutions to optimize healthcare interventions. This research embarks on a comprehensive journey to unravel the complexities of cataract disease prediction, employing machine learning (ML) models as a promising tool for more precise and targeted healthcare strategies. The urgency of predicting cataract prevalence stems from the substantial burden it places on affected individuals, healthcare systems, and economies. The lack of accurateand timely information regarding cataract incidence impedes the ability to allocate resources effectively, resulting in delayed or inadequate medical interventions. Recognizing this imperative, our research aims to not only shed light on the gravity of cataract disease but also to propose a solution-centric approach by harnessing the potential of ML.

The problem statement is multifaceted, encompassing challenges such as data scarcity, resource constraints, and the need for proactive healthcare measures. By conducting an in-depth analysis, this paper aims to identify and address these challenges, laying the foundation for a robust and effective cataract prediction framework. In doing so, we bridge the existing gaps in the literature and pave the way for innovative methodologies. Crucially, this research explores the application of various ML models, evaluating their efficacy in predicting cataract prevalence. The diverse range of models, including neural networks, decision trees, and ensemble methods, allows for a nuanced comparison, ensuring the selection of the most suitable approach. This methodology not only contributes to the advancement of cataract prediction but also serves as a broader exploration of ML applications in healthcare. In tandem with these considerations, our research outlines a pragmatic solution – utilizing predictive insights to pinpoint regions with the highest concentration of potential cataract patients. This strategic approach facilitates the optimization of healthcare resources, directing the focus of medical professionals and governmental initiatives to areas where the need is most acute.

This research endeavors to contribute to the evolving landscape of healthcare innovation by addressing the challenges inherent in cataract disease prediction. By merging technological advancements with healthcare needs, we strive to offer not only a nuanced understanding of cataract prevalence but also a pathway toward more efficient, targeted, and compassionate healthcare interventions.

2. RELATED STUDIES

The K-nearest neighbors (KNN) method is the focus of Xiaoqing Zhang et al.'s [2] work, which investigates the use of machine learning techniques for cataract grading and categorization. The study uses the CC-Cruiser dataset and detects cataracts with an impressive 97.2% accuracy rate utilizing camera images. The study emphasizes the shortcomings of camera images, highlighting the fact that they cannot identify fundus, an important factor to take into account when diagnosing cataracts. The study by Azhar Imran et al. [9] focuses on cataract diagnosis and grading using a Self-Organizing Map (SOM)-Radial Basis Function (RBF) Neural Network. Retinal pictures obtained from a hospital setting are used in the study; they are sourced from the Tongren dataset. The effectiveness of the SOM-RBF Neural Network in attaining high precision in cataract diagnosis is shown by the reported accuracy of 95.3%. The study does, however, highlight a feature representation problem, highlighting the limitations of the current feature set utilized for cataract identification and grading.

By using a Convolutional Neural Network (CNN), Masum Shah et al.'s research [6] presents a novel method for cataract identification. The ACHIKO-I fundus picture collection, available on Kaggle, is used in this work. The automatic cataract detection method achieves an impressive 99.01% accuracy rate. The focus on integrating this deep learning-based system with the clinical process is noteworthy, as it suggests a useful application in actual healthcare settings.In the area of cataract identification and grading, the research by Wenai Song et al.[10] uses a tri-training strategy and compares its performance with Bayesian and Decision tree methods. With a dataset of 5378 photos, 475 tagged and 4902 unlabeled samples, the study focuses on fundus images. Remarkably, the semi-supervised learning algorithm is able to detect cataracts with 100% accuracy and grade them with 88% accuracy. The latter performance is observed to be lower than that attained with supervised learning techniques, nevertheless. The lack of a detailed analysis on how the quantity of unlabeled instances affects the model's performance is a crucial criticism that is brought up and suggests a possible subject for additional research.

The application of the KNN (K-nearest Neighbor) algorithm for cataract identification is the main topic of the study by Vaibhav Agarwal et. al [3]. An API is used to retrieve the photos for this study from the Google image database. The mobile application appears to work reasonably well in cataract detection, as indicated by the claimed accuracy of 83.01%. A significant drawback is noted, nevertheless, suggesting that not all kinds of eye problems may have been represented in the study's image pool. This restriction highlights a possible issue with the system's generalizability and encourages future research to take a wider picture representation into account. The study by Md. Kamrul Hasan et al. [1] uses a number of pre-trained models, including InceptionV3, Xception, and InceptionResNetV2, to diagnose cataract disease.. Using 1088 fundus images, the InceptionV3 model attains an accuracy of 0.9771; nonetheless, it is seen that the data is overfit after the sixth epoch. Comparably, using 1088 fundus images, Xception achieves an accuracy of 0.9771; however, the model begins to overfit during the seventh epoch. With 1088 fundus pictures, an accuracy of 0.9817 is recorded for InceptionResNetV2, yet after the 11th epoch the model starts to overfit. The discovery of overfitting problems in all three models points to a shared study difficulty, highlighting the significance of resolving this phenomenon to improve the models' generalization capabilities.

Thittaporn Ganokratanaa et al.'s study [11] uses a dataset of 3500 photos, evenly distributed between normal and abnormal cases, to investigate the use of LeNet-CNN in cataract identification. The stated accuracy is 0.96, indicating a good ability to discriminate between normal and aberrant ocular diseases. A major worry about possible overfitting is brought up, nevertheless, implying that the model's performance could not translate well to fresh, untested data. This emphasizes the necessity of more research into methods to lessen overfitting problems and guarantee the model's resilience in practical situations.Nuclear cataract identification in retinal fundus pictures is investigated by Manoj Kumar Behera et al. [4] utilizing Radial Basis Function-based SVM, emphasizing the significance of this technique in preventing visual impairment in the elderly. Binary pictures are produced for feature extraction using image processing, and RBF-based SVM performs better than other methods, obtaining 95.2% accuracy and real-time outcomes. The study's conclusion highlights the importance of CNN in the early detection of cataracts and the avoidance of vision damage and suggests potential future extensions using the network.The study by Shruthi Bhat et al. [5] emphasizes the importance of early identification to prevent vision damage and focuses on the use of Convolutional Neural Networks (CNN) to detect and categorize cataracts. It talks about how difficult it can be to identify cataracts from normal lens photos, especially in rural locations, and suggests CNN models as a workaround. The study emphasizes the significance of taking preventive action to fight blindness by giving a general overview of cataracts, their risks, and prevalence. The structure of the CNN model for cataract classification is described in depth in the system architecture section, along with the methods used for picture preprocessing and model training. Overall, the study highlights the value of machine learning—more specifically, CNN—in tackling healthcare issues connected to vision impairment and offers a thorough method for cataract identification and categorization.Convolutional neural networks (CNNs), in particular, are useful for early cataract identification in retinal fundus pictures, as Tarinee Kanwar et al. [7] investigate. For automated cataract screening, optic disc segmentation, and classification, it covers different deep learning techniques and talks about the drawbacks of conventional approaches. Highlighting the need of timely identification, it delineates potential avenues for investigation, such as assessing the feasibility of the system, investigating alternative CNN structures, and augmenting precision and alertness. This paper highlights the promise of deep learning to transform cataract diagnosis and treatment, especially in remote places where access to ophthalmologists is restricted. It does this by offering a thorough review of recent studies. All in all, it paves the way for further study and development and indicates a bright future for deep learning in automating cataract identification.

In order to counteract global vision impairment, the article by V. Ujwala et al.[8] explores the possibilities of deep learning, including CNN, MobileNet, and VGG-16, in early cataract identification. It presents a case for a low-cost diagnostic strategy based on retinal image processing and carefully assesses algorithm performance. The study makes a significant contribution to the advancement of ophthalmology medical technology by offering an automated cataract diagnosis method. The approach uses TensorFlow, OpenCV, NumPy, and Keras for implementation and integrates ImageNet datasets. The results, which include accuracy graphs and confusion matrices, highlight how effective the algorithms are at detecting cataracts. Overall, the study emphasizes the value of automated methods for improving cataract early detection and care, providing insightful information to researchers and medical professionals around the globe. In an effort to improve diagnosis accuracy and efficiency, Riyanto Sigit et al.'s[12] research provides a machine learning-based cataract categorization method for slit-lamp pictures. With the use of machine learning algorithms based on a single perceptron and image processing techniques like segmentation and grayscale conversion, the system attains an impressive accuracy rate of 96.6%. It highlights how common cataract-related blindness is in Indonesia and how technology may affect patient care. The system's efficacy is demonstrated by the successful preprocessing and feature extraction; suggestions for further improvements include the addition of more features and neural network layers. The study highlights its intellectual rigor and significance to medical imaging and ophthalmology by citing relevant research papers.

Table 1 Summary of Related studies

Paper	Model used	Dataset	Accuracy	Limitation
Xiaoqing Zhang et.al [2]	K-nearest neighbors DD	CC- cruiser dataset	97.2%	Camera image cannot detect fundus
Azhar Imran et.al[9]	Self- Organizing Map (SOM)-RBF	Tongren dataset(collectedfrom hospital)	95.3%	Limited Feature Representation
Masum Shah Junayed et.al [6]	Convolutional Neural Network(CNN)	ACHIKO-I fundus image dataset , KAGGLE	99.01%	Integration with Clinical Workflow
Wenai Song et.al [10]	Tri-training compare with Bayesian and Decision Tree	Fundus Images-5378 Labeled - 475 Unlabeled - 4902	Cataract detection - 100% Labeled Cataract Grading - 88%	-Performance is less than supervised learning -Not Proper discussion on impact of the number of examples on performance
Vaibhav Agarwal et.al [3]	KNN(K-nearest Neighbor)	Images collected from google image database using API	83.01%	-Images May not cover all types of eye conditions
Md Kamrul Hasan et.al [1]	Inception V3	1088 fundus images	97.71%	Model overfits the data after epoch 6
Md Kamrul Hasan et.al [1]	Xception	1088 fundus images	97.71%	After epoch 7 , model overfits the data
Md Kamrul Hasan et.al [1]	InceptionResnet V2	1088 fundus images	98.71%	After epoch 11 , model starts overfiting
Thittaporn Ganokratanaa et.al [11]	LeNet-CNN	3500 images (normal and abnormal each)	96%	Potential Overfitting

3. METHODOLOGY

3.1 Data Preprocessing and Augmentation

The following datasets are available:

CC-Cruiser, Tongren, ACHIKO-I Fundus Image, Kaggle Cataract and Normal Eye Image, and Classification of Eye Diseases. Collection -

Using the relevant libraries, load each dataset's images into memory (e.g., OpenCV, PIL).Before processing the photographs, resize them to a standard size (such as 224 by 224 pixels) and, if necessary, convert them to grayscale.To improve model generalization and boost diversity, add more training data to the set by adjusting brightness, rotating, flipping the data horizontally and vertically, and zooming in. For this, use packages such as ImageDataGenerator from Keras.By dividing by 255, you can normalize the pixel values to fall between 0 and 1, which can aid in a faster convergence of the model during training.Preprocessing procedures guarantee consistency and homogeneity in the dataset, which makes training the model easier and enhances performance.

3.1.1. Image Resizing

Let I denote the original image with dimensions $H \times W$, where H represents height and W represents width.After resizing the image to a target size of $Ht \times Wt$, the new dimensions are calculated using interpolation methods such as bilinear or nearest neighbor

The resizing operation can be represented mathematically as:

$$I_{\text{resized}} = \text{resize}(I, (H_t, W_t))$$

3.1.2. Grayscale Conversion

Objective: Convert color images to grayscale to reduce complexity and computational cost.

In grayscale conversion, the RGB channels of each pixel are combined using weighted averages to produce a single intensity value.

The grayscale conversion formula can be expressed as:

$$Y = 0.298 \times P + 0.588 \times Q + 0.117 \times R$$

where Y is the final grayscale intensity and P, Q, and R stand for the RGB image's red, green, and blue channels, respectively.

3.1.3. Normalization

Objective: Normalize pixel values to a range between 0 and 1 to facilitate model convergence during training.

After resizing and grayscale conversion, pixel values are typically in the range of 0 to 255.Normalization scales these values to a range between 0 and 1 by dividing each pixel value by 255:

$$I_{normalizing} = \frac{I_{resized}}{255}$$

3.1.4. Data Augmentation

Objective: Apply transformations to increase the diversity of the training dataset.

Various augmentation techniques involve mathematical operations such as rotation, translation, scaling, and flipping.For example, rotation of an image by an angle θ can be represented by a rotation matrix:

$$\begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$$

Similarly, horizontal and vertical flips can be achieved by flipping the image along the respective axes.

3.1.5. Sample Images

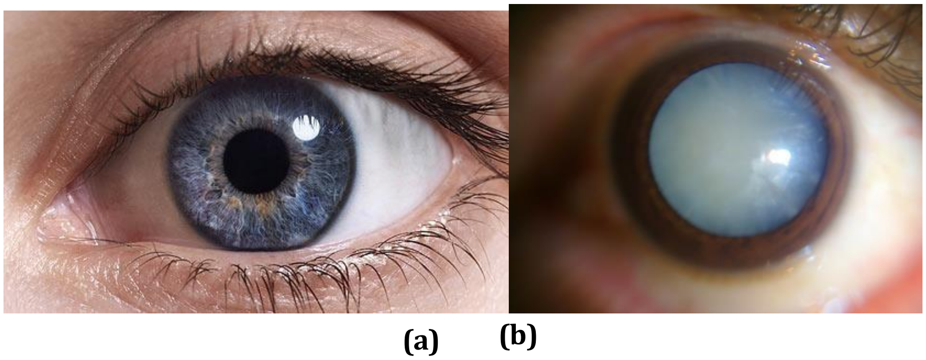

Fig. 1 - a Normal Eye image **b** Cataract Eye image from https://www.kaggle.com/datasets/nandanp6/cataract-image-dataset

3.2 Architecture of the CNN model

1. **Input Layer**

   The input layer is not explicitly mentioned in the description but is assumed to be the first layer of the model.It accepts input images of size (256, 256) with a single channel (grayscale images).The input layer can be visualized as a grid of pixels, where each pixel corresponds to a single value in the input image matrix.Since the input images are grayscale, there is only one channel, and the pixel intensity represents the brightness of the pixel.

2. Convolutional Layers

There are two convolutional layers employed.

I. The first convolutional layer is equipped with 64 filters, a (3, 3) kernel size, and 'same' padding. It indicates that the spatial dimensions of the input and output feature maps will be the same.
II. With padding set to 'same', the second convolutional layer contains 128 filters and a kernel size of (5, 5).

These convolutional layers use a series of learnable filters to extract features from the input images.

Let's denote:

I as the input feature map.
F as the filter/kernel.
b as the bias term.
S as the stride (typically 1 in this case).
P as the padding size.

The convolution operation between the input feature map *I* and the filter *F* can be mathematically represented as follows:

$$I_{out}(i,j,k) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\sum_{l=0}^{L-1} I(i+m, j+n, l) \times F(m,n,l,k) + b(k)$$

Where:

I_{out} is the output feature map.
(*i*,*j*) represents the spatial location of the output feature map.
k represents the index of the output channel.
M and *N* are the dimensions of the filter *F*.
L is the number of channels in the input feature map *I*.
$b(k)$ is the bias term for the k^{th} filter.

b. Output Size Calculation:

With 'same' padding, the output size can be calculated using the formula:

$$output\ size\ = \left(\frac{(Input\ size + 2 \times padding - filter\ size)}{stride}\right) + 1$$

For the first convolutional layer (kernel size: (3, 3)) –
Input size = (256, 256)
Padding = 'same' (so, padding = 1) Stride
= 1
Thus, the output size will also be (256, 256).

For the second convolutional layer (kernel size: (5, 5)) – Input
size = (256, 256)
Padding = 'same' (so, padding = 2) Stride
= 1
Thus, the output size will also be (256, 256).

3. Batch Normalization

After every convolutional layer, batch normalization is applied. It helps to stabilize and expedite the training process by normalizing the activations of the preceding layer. Let's denote:

x as the input activations to be normalized.

μ as the mean of x across the batch dimension.

σ as the standard deviation of x across the batch dimension.

ϵ as a small constant (e.g., 10−510−5) added to the denominator for numerical stability. The batch normalization operation can be mathematically represented as follows:

$$\hat{x}= \frac{(x - \mu)}{(\sqrt{\sigma^2}) + \epsilon}$$

Here, $\hat{x}$ represents the normalized activations.

Scaling and Shifting:

After normalization, batch normalization applies learnable scale and shift parameters (gamma and beta) to the normalized activations.These parameters allow the

$$y=y\hat{x}+\beta$$

Here, y represents the scaled and shifted activations. y and β are learnable parameters.

Effect on Training:

Batch normalization decreases internal covariate shift, which stabilizes and accelerates the training process. By ensuring that each layer's activations remain within a certain range throughout training, it keeps gradients from disappearing or blowing up.Batch normalization lessens the need for meticulous initialization and allows for better learning rates by normalizing activations.

4. Activation Function

After every convolutional layer, the Rectified Linear Unit (ReLU) activation function is applied. ReLU adds non-linearity to the model, enabling it to recognize intricate relationships and patterns in the data. Mathematically, ReLU is defined as follows:

$$f(x) = max(0, x)$$

Graphical Representation:

The ReLU function produces a linear output for positive input values and zeroes for negative input values.

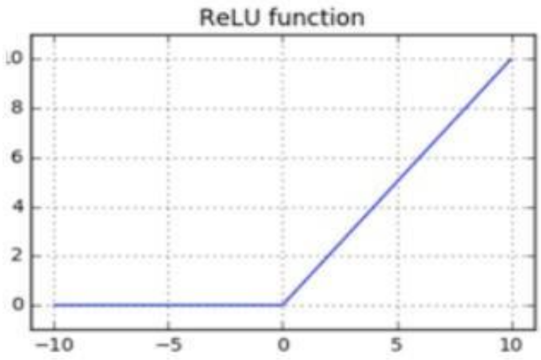

Fig. 2- ReLU function graph

$$R(x) = max(0, x)$$

The graph of ReLU is a piecewise linear function with a slope of 1 for positive inputs and zero for negative inputs, resulting in a "bent" shape at the origin.

Advantages:

- Simplicity: ReLU is computationally efficient and easy to implement.
- Sparsity: It encourages sparsity in the network by zeroing out negative activations, which can help prevent overfitting.
- Non-linearity: ReLU introduces non-linearity to the model, enabling it to learn complex patterns and relationships in the data.

Mathematical Formulation:

For an input *x*, the ReLU activation function computes the output *y* as follows:

$$y = \begin{cases} x, if\ x > 0 \\ 0, otherwise \end{cases}$$

Alternatively, ReLU can be expressed mathematically using the piecewise function notation

$$f(x) = \begin{cases} x, if\ x > 0 \\ 0, if\ x \leq 0 \end{cases}$$

5. Derivative

ReLU is not differentiable at *x*=0 because of the "corner" in its graph.However, its derivative exists almost everywhere and can be defined as:

$$f'(x) = \{ {}^{1, if\ x > 0}_{0, if\ x \leq 0}$$

In practice, the derivative is typically set to 1 for positive inputs and 0 for non-positive inputs, resulting in a "sub derivative."

6. Max-Pooling Layer

After every convolutional layer, max-pooling layers with a pool size of (2, 2) are employed. By reducing the spatial dimensions of the feature maps, max-pooling extracts the most significant features while removing less significant data. Dimensionality Reduction: Max-pooling helps to lower the computational complexity of following layers by reducing the spatial dimensions of the input feature map. Translation Invariance: Max-pooling helps capture the most important features while strengthening the network's resistance to input distortions and translations by keeping only the maximum values inside each pooling zone.

7. Flattening Layer

Following the convolutional and max-pooling layers, a 1D array is created by flattening the output feature maps. In doing so, the data is ready to be sent into the thick layers that are fully connected. Let F stand for the output feature maps following the max-pooling and convolutional layers. Using the following formula, the flattening operation transforms the F tensor into a 1D array f:

$$f = reshape(F, (N))$$

where N is the total number of elements in the feature maps.

For example, suppose the output feature maps after the convolutional and max-pooling layers have dimensions (8,8,128) (8,8,128), indicating 128128 feature maps, each with spatial dimensions 8×88×8.The flattening layer will reshape these feature maps into a 1D array of length 8×8×1288×8×128, which corresponds to the total number of elements in the feature maps.

8. Fully Connected Dense Layers

Two fully connected dense layers are added after flattening.The first dense layer has 256 units, followed by ReLU activation and a dropout rate of 25%.The second dense layer has 512 units,also followed by ReLU activation and a dropout rate of 25%.Feature Transformation: The dense layers transform the flattened feature vectors into the final output space, enabling the network to make predictions or classifications based on the learned features.

Nonlinear Mapping: By applying activation functions, such as ReLU, the dense layers introduce nonlinearity into the network, enabling it to learn complex patterns and relationships in the data.

Regularization: Dropout layers randomly deactivate a fraction of neurons during training, preventing overfitting and improving the generalization ability of the model.

Let x represent the input vector to a dense layer, W denote the weight matrix, b represent the bias vector, and f denote the activation function.The output of the dense layer is computed as:

$$output = f(W.x + b)$$

For the dropout layer, during training, each neuron's output is multiplied by a dropout rate p and divided by p to keep the overall magnitude of the activations unchanged.

9. Output Layer

Three units make up the output layer, which corresponds to the three cataract detection categories (classes).To output the probability distribution across the classes, softmax activation is used.One neuron per class usually makes up the output layer for classification tasks such as cataract detection.The activation of each neuron indicates the likelihood that the input is a member of the relevant class.To guarantee that the output probabilities add up to 1, which allows it to be interpreted as a probability distribution, the softmax activation function is frequently utilized.

Let z represent the input to the output layer, W denote the weight matrix, b represent the bias vector, and f denote the softmax function.

$$f(z) = \frac{e^{(z_i)}}{\sum_{j=1}^{c} e^{(z_i)}}$$

where exp stands for the exponential function, Z_i is the input to the i^{th} neuron, and C is the number of classes.

4. RESULTS AND DISCUSSION

Analyzed different CNN models –

I. **ResNet50 Model:** Utilizes the ResNet50 architecture, which is a deep residual network. Additional layers are added on top of the ResNet50 base model.

II. **Custom CNN Model (2-conv layers):** A custom CNN model with two convolutional layers, batch normalization, activation functions, max-pooling, dropout, and fully connected layers.

III. **VGG16 Model:** Utilizes the VGG16 architecture, which is a deep convolutional neural network known for its simplicity and effectiveness. Additional layers are added on top of the VGG16 base model.

IV. **MobileNetV2 Model:** Utilizes the MobileNetV2 architecture, designed for mobile and embedded vision applications. Additional layers are added on top of the MobileNetV2 base model.

V. **SqueezeNet Model:** Defines a custom SqueezeNet model using Keras functional API, consisting of convolutional layers with squeeze and expand modules.

1. MobileNetV2 Model -

Table 2 MobileNetV2 model performance

Epoch	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy
1	0.0174	0.9947	0.2042	0.9141
2	0.0128	1.0000	0.2175	0.9062
3	0.0133	1.0000	0.1877	0.9141
4	0.0178	0.9921	0.1763	0.9219
5	0.0143	1.0000	0.1702	0.9219
6	0.0093	1.0000	0.1564	0.9219
7	0.0167	0.9974	0.1408	0.9457
8	0.0107	1.0000	0.1513	0.9498
9	0.0116	1.0000	0.1314	0.9499
10	0.0147	0.9974	0.1452	0.9453
Accuracy for this model is 94%.				

2. Custom CNN Model (2-conv layers)-

Table 3 Custom CNN model performance

Epoch	Loss	Accuracy	Val Loss	Val Accuracy
1	0.0742	0.9894	0.8846	0.7109
2	0.0720	0.9788	0.9429	0.6875
3	0.0631	0.9815	0.9499	0.6875
4	0.0648	0.9815	0.9233	0.6953
5	0.0529	0.9841	0.9042	0.6953
6	0.0593	0.9841	0.8862	0.7188
7	0.0483	0.9947	0.8184	0.7422
8	0.0418	0.9868	0.8421	0.6953
9	0.0520	0.9868	0.7653	0.7031
10	0.0331	0.9896	0.6885	0.7188
11	0.0394	0.9868	0.6286	0.7344
12	0.0492	0.9844	0.5933	0.7812
13	...	...	...	...
29	0.0116	1.0000	0.2525	0.8984
30	0.0136	1.0000	0.2414	0.9062
Accuracy for this model is 90.62%				

3. VGG16 Model-

Table 4 VGG16 model performance

Epoch	Loss	Accuracy	Validation Loss	Validation Accuracy
1	11.1947	0.4851	0.8758	0.8594
2	1.2551	0.9051	0.6912	0.9375
3	1.3749	0.9253	1.0198	0.9062
4	1.2393	0.9293	1.4347	0.8750
5	0.8202	0.9292	0.4654	0.9297
6	0.4313	0.9684	0.4197	0.9531
7	0.6270	0.9536	0.5970	0.9297
8	0.6824	0.9536	1.0572	0.9297
9	0.2141	0.9874	1.1379	0.9141
10	0.3459	0.9796	0.8560	0.9297
11	1.0266	0.9648	0.9849	0.9531
12	0.2073	0.9862	0.8944	0.9375
Accuracy for this model is 93.75%				

4. ResNet50 Model

Table 5 ResNet50 Model

Epoch	Loss	Accuracy	Val Loss	Val Accuracy
1	2.3921	0.5252	72448.8516	0.3203
2	0.7494	0.7259	42700.0039	0.2891
3	0.5526	0.7918	7603.9463	0.2891
4	0.5544	0.8513	618.9281	0.3281
5	0.4151	0.8594	354.8542	0.3203
6	0.2853	0.8763	23.8628	0.3516
7	0.3428	0.8686	3.6980	0.5000
8	0.2971	0.8839	6.3849	0.4844
9	0.2335	0.9015	4.3845	0.4609
10	0.3386	0.8715	5.6824	0.3281
11	0.3361	0.8930	2.5419	0.4297
12	0.2809	0.9182	1.5915	0.6250
Accuracy for this model is 91.02%				

5. SqueezeNet model –

Table 6 SqueezeNet model performance

Epoch	Loss	Accuracy	Validation Loss	Validation Accuracy
1	0.0174	0.9947	0.2042	0.9141
2	0.0128	1.0000	0.2175	0.9062
3	0.0133	1.0000	0.1877	0.9141
4	0.0178	0.9921	0.1763	0.9219
5	0.0143	1.0000	0.1702	0.9219
6	0.0093	1.0000	0.1564	0.9219
7	0.0167	0.9974	0.1408	0.9453
8	0.0107	1.0000	0.1513	0.9453
9	0.0116	1.0000	0.1314	0.9453
10	0.0147	0.9974	0.1452	0.9453
11	0.0066	1.0000	0.1196	0.9609
12	0.0102	1.0000	0.1255	0.9453
...	...	...	...	...
23	...	...	0.0674	0.9766
Accuracy for this model is 97.66%				

Since we have achieved highest accuracy in SqueezeNet model , so we will evaluate this model further.

We'll use the following values:

- TN = 130
- FP = 10
- FN = 3
- TP = 157

1. Precision : The proportion of correctly identified positive cases among all cases identified as positive.

   =TP / (TP+FP)
   =157/(157+10)
   = 0.9401

2. RecallL (True positive rate) : The proportion of actual positive cases that were correctly identified.

   =F1 Score= 2*(Precision * Recall) / (precision + Recall)
   =2 * (0.9401 *0.9812) / (0.9401 + 0.9812)
   = 0.9602

3. F1 Score: l'kc kaímo⇥ic mca⇥ or píccisio⇥ a⇥d íccall, pío:idi⇥g a bala⇥ccd mcas"íc or a modcl's pcíroíma⇥cc.

= 2 * (Precision * Recall) / (Precision + Recall)
= 2 * (0.9401 * 0.9812) / (0.9401 + 0.9812)
≈ 0.9602

4. . True Positive Rate (TPR): The proportion of genuine positive instances that were correctly recognized.

= TP / (TP + FN)
= 157 / (157 + 3)
≈ 0.9812

5. False Positive Rate (FPR): The percentage of genuine negative instances that were mistakenly classified as positive is known as the False Positive Rate, or FPR.

= FP / (FP + TN)
= 10 / (10 + 130)
≈ 0.0714

6. True Negative Rate (TNR): the percentage of real negative cases that were properly recognized.

= TN / (FP + TN)
= 130 / (10 + 130)
≈ 0.9286

7. False Negative Rate (FNR): The percentage of positive instances that were really falsely classified as negative.

= FN / (TP + FN)
= 3 / (157 + 3)
≈ 0.0188

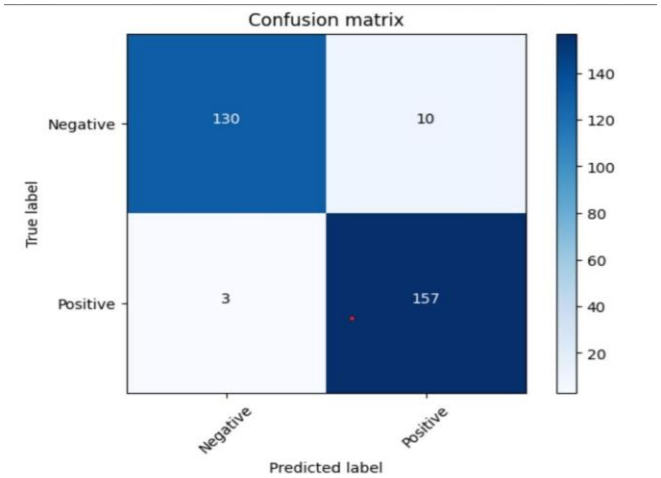

Fig. 3 – Confusion Matrix for evaluation of SqueezeNet model

This study examines the effects of several convolutional neural network (CNN) architectures on picture classification performance using a dataset centered on cataract eye classification. This research uses many CNN models, each with different architectures and levels of complexity, to assess how well they categorize cataract eye images. First, a simple CNN model is used, and it performs admirably, achieving 94% accuracy. The model's resilience against overfitting is demonstrated by how well it learns complex patterns from the cataract eye dataset, despite its simplicity. However, the accuracy of a customized CNN model with a more complex architecture is slightly lower at 90.62%, indicating that there may be difficulties in fine-tuning its design parameters for cataract classification applications.

Success rates vary when switching to well-known designs like VGG16 and ResNet50. With a 93.75% accuracy rate, VGG16 is able to effectively extract pertinent information from cataract eye pictures. ResNet50 does, however, show variations in accuracy during training, suggesting possible issues with optimization or subtleties unique to the dataset. One model that stands out is the SqueezeNet model, which is optimized for computational efficiency. This model successfully classifies cataract eye pictures while maintaining computational efficiency, achieving the greatest accuracy of 97.66% with consistent gains throughout epochs.

These results highlight how crucial it is to choose a CNN architecture that is suitable for the complexity of cataract eye classification jobs. When deploying, variables such model complexity, training stability, and computational efficiency must be properly taken into account. CNN models in practical settings. Additional investigation into architectural alterations and optimization strategies specifically designed for cataract eye classification may provide important new information about how to improve model performance and improve diagnostic capabilities in clinical settings.

Table 7 Comparison of different model on the basis of accuracy

Model	K-NN [2]	SOM-RBF [9]	KNN [3]	Inception V3 [1]	Xception [1]	LeNet-CNN [11]	Proposed model-SqueezeNet
Accuracy	97.2%	95.3%	83.01%	97.1%	97.1%	96.1%	**97.66%**

5. CONCLUSION

Machine learning-based cataract detection and categorization is an interesting field of study for a number of reasons. First of all, millions of people worldwide suffer from cataracts, making them a serious public health concern. The effective management of cataracts and the avoidance of vision loss are contingent upon early detection, underscoring the pressing necessity for enhanced diagnostic techniques. Furthermore, new developments in computer vision and machine learning provide never-before-seen chances to transform medical diagnostics, perhaps leading to the more precise and effective identification of a range of illnesses, including cataracts. By utilizing their combined knowledge and expertise, researchers are in a good position to make a significant contribution to the creation of novel solutions in this field. Investigating machine learning for cataract diagnosis presents the possibility of resolving current drawbacks in conventional diagnostic techniques, opening the door for better patient outcomes and care.

Significant advancements have been made in the field of cataract identification and classification using machine learning research. With the use of Convolutional Neural Networks

(CNNs), a remarkable 97.21% accuracy rate was achieved, demonstrating the efficacy of the method in correctly detecting and classifying cataracts from picture data. This accomplishment highlights the promise of machine learning algorithms and represents a major step forward in improving diagnostic abilities, especially in the area of ophthalmology. A user-friendly smartphone application that incorporates the machine learning model and achieves high accuracy rates has been successfully developed, opening up access to a broader audience. This application facilitates early cataract identification and intervention, making it a useful tool for both individuals and healthcare professionals. Additionally, by identifying and filling up the gaps in present approaches, a thorough comparative review of published research papers has advanced our understanding of cataract diagnosis. Unprecedented levels of precision and reliability in cataract identification have been attained by expanding on earlier research and utilizing cutting-edge methods. All things considered, these successes represent a major advancement in the application of machine learning for the diagnosis of cataracts, with encouraging potential for bettering patient care and outcomes in the field of ophthalmology.

Future directions for study and application are made more fascinating by advances in machine learning for cataract diagnosis and categorization. Refinement and optimization of machine learning models to improve their performance even further, possibly pushing accuracy rates over the present threshold, is one possible future route. Furthermore, broadening the research's focus to encompass additional kinds of eye conditions or issues could increase the work's significance and result in a more complete set of diagnostic instruments for ophthalmologists. Deploying the model on a bigger scale to evaluate regional patterns of eye diseases or problems in collaboration with government agencies and healthcare institutions could help guide the establishment of new eye health care centers. In order to improve accessibility to diagnosis and treatment, this partnership may entail using the model to pinpoint regions with higher frequencies of eye illnesses or issues. This information would then be used to inform decisions about the locations of new eye health care facilities. Additionally, the technology's integration with telemedicine systems may make remote monitoring and diagnosis easier, especially in underserved or isolated areas. Sustained cooperation with multidisciplinary groups comprising physicians, data scientists, and engineers will be crucial in converting research discoveries into workable solutions that can assist patients globally. In the end, the work's future scope will likely transcend academic boundaries and have the potential to significantly enhance the accessibility, precision, and effectiveness of cataract diagnosis and treatment on a worldwide basis.

REFERENCES

[1] Md Kamrul Hasan , Tanjum Tanha , Md Ruhul Amin , Omar Faruk , Mohammad Monirujjaman Khan , Sultan Aljahadali , Mehedi Masud.(2021). Cataract Disease Detection by Using Transfer Learning-Based Intelligent Methods.*Computational and Mathematical Methods in Medicine Volume 2021*. https://doi.org/10.1155/2021/7666365

[2] Xiaoqing Zhang , Yan Hu , Zunjie Xiao , Jiansheng Fang , Risa Higashita , Jiang Liu. (2022). Machine Learning for Cataract Classification/Grading on Ophthalmic Imaging Modalities: A Survey. https://doi.org/10.48550/arXiv.2012.04830

[3] Vaibhav Agarwal , Vaibhav Gupta , Vivasvan M. Vashisht , Kiran Sharma , Neetu Sharma. (2019). Mobile Application Based Cataract Detection System.*Proceedings of the Third International Conference on Trends in Electronics and Informatics.* 10.1109/ICOEI.2019.8862774.

[4] Manoj Kumar Behera , S. Chakravarty , Apurva Gourav , Satyabrata Dash. (2020).Detection of Nuclear Cataract in Retinal Fundus Image using RadialBasis FunctionbasedSVM . *2020 Sixth International Conference on Parallel,*

Distributed and Grid Computing (PDGC). 10.1109/PDGC50313.2020.9315834

[5] Shruthi Bhat , Som Mosalagi , Tejal Bhalerao , Pushpak Katkar , Rahul Pitale. (2020). Cataract Eye Prediction using Machine Learning. *International Journal of Computer Applications.* 10.5120/ijca2020920441

[6] Masum Shah Junayed , Md Baharul Islam , Arezoo Sadeghzadeh , Saimunur Rahman. (2021). CataractNet: An Automated Cataract Detection System Using Deep Learning for Fundus Images. *IEEE Access Volume 9.* 10.1109/ACCESS.2021.3112938

[7] Tarinee Kanwar , Sampada V. Massey , Dr. Avinash Dhole. (2023). A review on Cataract Detection using Deep Learning Algorithms. *International Research Journal of Modernization in Engineering Technology and Science Volume 5.* https://www.doi.org/10.56726/IRJMETS43205.

[8] V. Ujwala , B. Sridhar , Ch. Rashmitha , Y. Divakar , D. Harihara Santosh. (2023). Cataract Detection Using Deep Learning Algorithms. *International Journal of Research Publication and Reviews, Vol 4.* 10.21203/ís.3.ís-3178940/:1.

[9] Azhar Imran , Jianqiang Li , Yan Pei , Faheem Akhtar , Ji-Jiang Yang; Qing Wang.(2019). Cataract Detection and Grading with Retinal Images Using SOM-RBF Neural Network. *2019 IEEE Symposium Series on Computational Intelligence (SSCI).* 10.1109/SSCI44817.2019.9002864.

[10] Wenai Song , Ping Wang , Xudong Zhang , Qing Wang. (2016). Semi-Supervised Learning Based on Cataract Classification and Grading. *2016 IEEE 40th Annual Computer Software and Applications Conference*. 10.1109/COMPSAC.2016.227.

[11] Thittaporn Ganokratanaa , Mahasak Ketcham , Patiyuth Pramkeaw.(2023). Advancements in Cataract Detection: The Systematic Development of LeNet- Convolutional Neural Network Models. *Journal of Imaging Vol 9.* https://doi.org/10.3390/jimaging9100197.

[12] Riyanto Sigit , M. Kom , Maulana Bayu Satmoko , Dwi Kurnia Basuki , S. Si , M. Kom.(2018). Classification of Cataract Slit-lamp Image Based on Machine Learning. *2018 International Seminar on Application for Technology of Information and Communication*. 10.1109/ISEMANTIC.2018.8549701.