Enhancing Ophthalmic Diagnostics: CNNs in Cataract Detection

Kartik Solanki
Dept. of Computer Science
Graphic Era Hill University, Dehradun
Uttarakhand, India
kartiksolanki67194@gmail.com

Shaurya Pratap Singh
Dept. of Computer Science
Graphic Era Hill University, Dehradun
Uttarakhand, India
shauryaps3121@gmail.com

Aman Yadav
Dept. of Computer Science
Graphic Era Hill University, Dehradun
Uttarakhand, India
amanyadavddm@gmail.com

Vikrant Sharma
Dept. of Computer Science
Graphic Era Hill University, Dehradun
Uttarakhand, India
vsharma@gehu.ac.in

Satvik Vats
Dept. of Computer Science
Graphic Era Hill University, Dehradun
Uttarakhand, India
svats@gehu.ac.in

Abstract—Cataracts have been one of the most prevalent eye disorders, which often can cause significant visual impairment due to the clouding of the eye's lens. This condition can worsen in future, often leading to severe vision problems and even blindness. Consequently, detecting cataracts is paramount in mitigating the associated risks and preventing the onset of blindness. Throughout the years there has been quite good progress in leveraging cutting-edge technology, especially due to machine learning, to improve the perfection and efficacy of cataract detection. Convolutional Neural Networks (CNNs) have made an appearance as a commanding implement for computerized the systematization of eye images in the context of cataract identification. Proposed research was more concerned with fine-tune the process of cataract identification, aiming to increase the accuracy while minimizing the data loss. To accomplish this, we conducted a series of experiments, with a key focus on manipulating a critical parameter: the number of training epochs. The proposed research revealed a compelling relationship between the number of training epochs and the accuracy and loss of data in CNN. As we delved into a spectrum of epoch values, a clear pattern emerged: the higher the number of epochs, the more refined and potent the model became. In this research, a significant milestone was reached when utilizing a generous number of 90 epochs, resulting in an impressive accuracy rate of 97.37%.

Keywords— Cataract detection, Deep learning, efficacy, Convolutional Neural Network (CNNs), Eye disorders, Ophthalmology

I. Introduction

In the initial stages, cataracts might not cause any problem and may only be on a tiny part of the lens, causing blurriness. Cataracts may seem initially as the minor disturbances in the eye lens, often appearing as blurriness due to the limited extent of lens clouding [1]. However, the cataract might grow larger over time, affecting more areas of the lens making it even harder to see. Cataracts don't spread from one eye to another. Cataracts are not contagious or transferable from one eye to another. They usually develop independently in each eye separately. When the transparency of the crystalline lens decreases enough to disturb vision, a clinically significant cataract exists [1]. However, many people get cataracts in both eyes. As the cataract matures, the individual often report a depletion in optical intensity and clarity, prime to difficulties in judicious the punctilious and performing the regular activities such as reading, driving and acknowledgment faces of the people. The world health organization (WHO) have also announced that almost 285 million people across the globe has sight impairment [2]. There are different kind of cataracts, including crumbling-related cataracts, secondary cataracts, traumatic cataracts, or congenital cataracts [3]. Cataracts can be defined into various types based on their origin. These types include age-related cataracts, which are related to the natural aging of the eye, secondary cataracts are associated with underlying diseases, traumatic cataracts can be caused by some eye injuries, and congenital cataracts that can be present from the birth of that person or has been developed during the childhood [3].

The mankind is benefitted a lot with the advancement in medical technologies and associated technologies such as wireless sensor networks, IOT, and Body area networks, etc. But, still early detection of cataract is challenging task [4]–[8].

Early cataract detection portrays a significant part in the consideration and can decrease the threat of blindness [9]. The world vision reports released by WHO on October 8, 2019, pointed the high costs of accessing eye care, especially for the rural population was the major cause for visual impairments. The equipment's used to detect are quite expensive and requires skills to use [10], However nowadays there are various new methods including deep learning model's which can predict that whether the person is diagnosed with this disease or not. One of the famous deep learning models includes Convolutional Neural Network (CNNs) which can be used for pattern recognition [11]. Among several methods of machine learning the Convolutional Neural Network is very famous among them, due to of its ability to solve problems in the computer vision domain namely detection systems, classification systems. It is a more advance version of multilayer networks. A CNN has various multiple layers including input layers, output layers, and some hidden layers. The hidden layers are named as the convolutional layers, various activation function's such as SoftMax, Sigmoid, ReLU etc.

The suggested method introduces several new section: firstly reducing the model's parameters, inducing more layers and weights, to enhance mensuration efficacity and reduce costs.

II. Literature Review

Extensive research has been done on the implementation of machine learning for the discernment of eye cataract. Other studies have been done based on fundus images [1], [12], that has been plus point for our research, as there is noticeable gap to create a user-friendly model for the general population that can work directly on the image of the eye and can identify the disease as cataract or non-cataract.

While another model (Table 1) proposed by A. Shetty et. al. [5] employed the CNN model with an accuracy rate of 99% that shows the model might be overtrained. Whereas the model [13] has created a transportable and sturdy cataract detection device and a cataract masking of over 200 patients has been done using the device.

Table 1. Features Comparisons of Cataract Detection System with Different System Model

S No.	Key Features	Accuracy	System complexity
Paper[1]	Number of epoch values were 50 which lead to a higher accuracy	95%	Usage of the fundus dataset might not be user-friendly
Paper[2]	Large value of datasets help the model to accurately identify the disease in the eye.	98.17	Less number of epochs might lead to overfitting.
Paper[3]	usage of LeNet-CNN	96%	Less no of epoch values might result in inaccurate results
Paper[4]	Using of the convolutional network, 4746 images for training	99.13%	Usage of the fundus dataset might not be user-friendly
Paper[5]	Usage of RNN(Recurrent Neural Network)	100%	The model has been overtrained which can result into bad decisions .

III. System Model

The proposed research works on a very innovative and advance system model which aims towards enhancing the accuracy and efficiency of the cataract detection. Using the power of machine learning, the proposed system is primarily based on the state-of-the-art CNN. The CNN serves as a major role in identifying the eye cataract efficiently and accurately thus proving a beneficial role in the healthcare sector.

The proposed model consists of various major layers starting with a sequence of Conv2D layers. The Conv2D layers are responsible for extracting essential image features through convolutional operations. The model consists of 4 convolutional layers, each with an increasing depth (32, 64, 128, 256), allowing the model to detect the visible features in the images. There after every convolutional layer, a MaxPooling2D layers are applied. The pooling layers are responsible for reducing the spatial dimensions of the data while getting its most critical features. After the Pooling layers we have used the flatten layers, which is mainly responsible for reshaping the feature maps into a one-dimensional vector. This is a crucial step before feeding the data into the dense layers. The proposed model (Fig. 1) includes two dense layers. The first dense layers have 256 neurons', promoting feature learning and extraction at a higher level of abstraction. These layers use the activation function Relu (rectified linear unit), enhancing the models non-linearity. To prevent overfitting the model has been provided with a dropout layer at a dropout rate of 0.5. Lastly, the final yield layer is a smoggy layer with a single neuron, using the sigmoid activation function. This neuron produces the output, determining whether there is presence or absence of eye cataract in the input image or not.

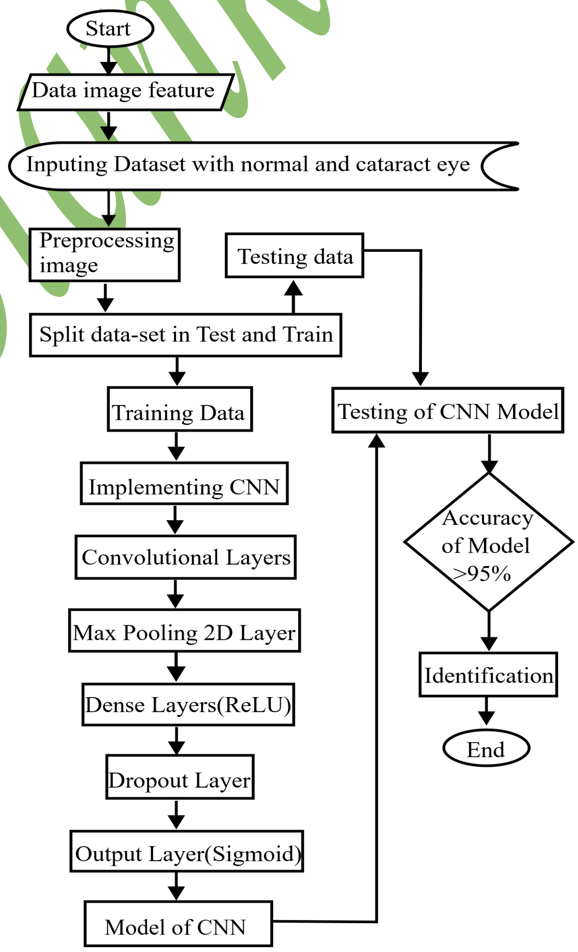

Fig. 1.Diagram to show the various layers in the System Model(flowchart)

A. Dataset

The dataset used for the proposed work is taken from Kaggle [14]. The datasets contain basically two directories named as

Test and Train. The Test folder has been used for the testing purpose whereas our Train directory has been used for the training purpose only. Both test and train folder contain two folders of the normal eye, and the cataract eye.
Our dataset theory is further based on image augmentation techniques to enhance the quality and diversity of the dataset of cataract so that the model has vast range of images to understand. Image augmentation is a fundamental technique employed in our dataset. This technique further allows us to increase the diversity of our training dataset. The augmentation has further been divided into steps:

- Rescaling: Images have been rescaled to ensure the consistency of the pixels (1.0/255.0).
- Rotation: Images have been rotated to 20 degrees, helping to increase the diversity of the images.
- Shifts: Both horizontal and verticals have been implemented, allowing in the variations in the position of the cataract within the image.

Further we have used image data generators class from the tensor-flow library to implement image augmentation. Two separate data generators have been used, one for the training dataset and another for the testing dataset.

B. *Pre-processing*

The image dataset has been provided by Kaggle. We have 491 images belonging to the training dataset in which we have two directories (Fig. 2) of the normal and the cataract eye. We have another 118 images belonging to the test dataset containing two directories of the normal and the cataract eye.

```
from tensorflow.keras.preprocessing.image import ImageDataGenerator


train_data_dir = "/Users/kartiksolanki/Desktop/eye_cataract/dataset/train"
test_data_dir = "/Users/kartiksolanki/Desktop/eye_cataract/dataset/test"


train_datagen = ImageDataGenerator(rescale=1./255,
                                   rotation_range=20,
                                   width_shift_range=0.2,
                                   height_shift_range=0.2,
                                   shear_range=0.2,
                                   zoom_range=0.2,
                                   horizontal_flip=True,
                                   fill_mode='nearest')

test_datagen = ImageDataGenerator(rescale = 1.0/255.)

train_generator = train_datagen.flow_from_directory(train_data_dir,
                                                    target_size=(224, 224),
                                                    batch_size=32,
                                                    class_mode='binary')


test_generator = test_datagen.flow_from_directory(test_data_dir,
                                                  target_size=(224, 224),
                                                  batch_size=32,
                                                  class_mode='binary')


Found 491 images belonging to 2 classes.
Found 118 images belonging to 2 classes.
```

Fig. 2. Partition and Labelling

In the pre-processing step, it is determined that how many epochs are used for the training dataset. In pre-processing a process named augmentation has been used for the training images. Augmentation is the process in which the images go through a step of processes such that the model will detect that the changed image is a different image. Augmentation can be helpful for increasing the accuracy of the model. The Images have been rescaled to ensure the consistency of the pixels (1.0/255.0). Also, the Images have been rotated to 20 degrees, helping to increase the diversity of the images.

C. *Partition and Image Label*

The images have been already divided into 2 categories basically named Test and Train. Based on Fig. 2 we can see that there are 491 images used as training data, 118 images as testing data. Each of these directories have been divided into two classes cataract and normal.

IV. Working

Our model explains the practicalities of the convolutional Neural Network (CNN) designed for the cataract detection. The model has been structures such that it processes eye images and make binary predictions regarding the presence or absence of cataracts. Further we have provided a detailed description of the how our model operates.

A. *Model Architecture*

The model consists of several layers, each with a unique role in processing and classifying input images:

- Input Layer (Conv2D): The input layer has 32 filters, each with a 3*3 kernel, and has been provide by the Rectified Linear Unit (ReLU) activation function. The image size that has been provide is 224*224 pixels with three colour channels.
- MaxPooling2D Layer: After every convolutional layer, a MaxPooling2D layer has been used with a admix of 2*2. This Layer is used for reducing the structural dimensions of the feature maps while retaining the important features.
- Convolutional Layers: The model further includes three additional convolutional layers, each with an increased number of filters (64, 128, and 256). These layers capture the important image patterns after that the images are passed through the ReLU activation function which introduces indiscriminate, enhancing the capacity to learn.
- Flatten layer: After the convolution and pooling layers, the flatten layers reshapes the 3D feature map into a one-dimensional vector.
- Dense Layers: This model includes two layers. The first one has 256 neurons', using the 'ReLU' activation function. This layer is used for feature extraction and helps in higher level feature learning.
- Dropout Layer: To mitigate the risk of overfitting, a dropout layer with a dropout rate of 0.5 is place after the first dense layer. It is used to deactivate a fraction of neuron's, thus helping in enhancing the robustness and generalisation of the model.
- Output Layer: The output layer consists of the dense layer with a single neuron. It consists of the 'sigmoid'

activation function, producing a binary output. The final layer or the output layer predicts whether there is cataract present in the given image or not.

B. *Model Operation:*

The model processes input images through a series of convolutional layers and pooling layers. The convolutional layers detect distinctive patterns, that are crucial for the cataract detection. The pooling layers lowers the structural dimensions of the feature maps while retaining the prime attribute. The flattened feature vectors are further passed to the dense layers, which extracts the higher-level features and relationships between the detected patterns. The dropout layer prevents overfitting during training. In the output layer our model makes binary prediction that whether there is cataract present or not.

C. *Model Activation Formulas*

In context of Convolutional Neural Network, activation function plays a salient role in shaping the working of the neural network. They are basically mathematical operations applied to the output signal of the neuron's, helping them to determine whether they should be activated based on their weighted sum of the inputs.

Our model has used to major activation functions namely ReLU, and sigmoid:

1) *ReLU(Rectified Linear Unit):-*

The ReLU activation functions (Equation 1)are one of the most widely used activation function in CNNs. It determines the output value of the neuron as follows:

- If the input value is less than 0, then the output would be set to 0.
- If the input value is greater than 0, the value of the output is directly equal to the value of the activation functions.

$$f(x) = \max(0, x) \quad (1)$$

Where function f(x) returns 0 if it gets a negative input, but for the positive value of x it returns the value back.

2) *Sigmoid*

Sigmoid function (Equation 2) is one of the most fundamental activation functions of CNN, known for its ability to introduce the non-linearity in the model (see Fig.3). This activation function is denoted by S(x), here it takes any real-valued input x, and maps it to a binary value ranging from 0 to 1. Sigmoid function resembles an elongated “s” curve. It is characterised by a smooth transition from 0 to 1, making it useful for binary classification, where it can represent an event that is occurring.

$$s(x) = 1/\left(1 + e^{-x}\right) \quad (2)$$

Where function S(x) is the sigmoid function, and it returns the value 0 and 1 for the given input x.

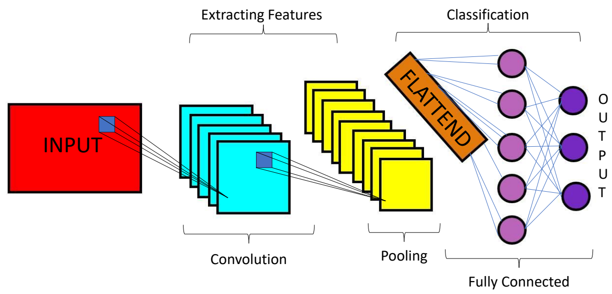

Fig.3. Representation of CNN Model

V. RESULT AND DISCUSSION

Learning Rate and Epoch value Optimization

During the research various modification were made to the model, and one of the major modifications was of the epoch value. During the research it was noticed as higher the epoch value was, the more the accuracy of the model was. Considering this we needed the optimized value of the model to avoid overtraining. In searching for the right value of the epochs several tests were performed at different epoch values. Each iteration of the epoch value noticed a rise in the accuracy of the model, and a major decrease in the loss of the model. Further shown in the following Fig. 4 it showed that the increase in the epoch value we got a good level of accuracy.

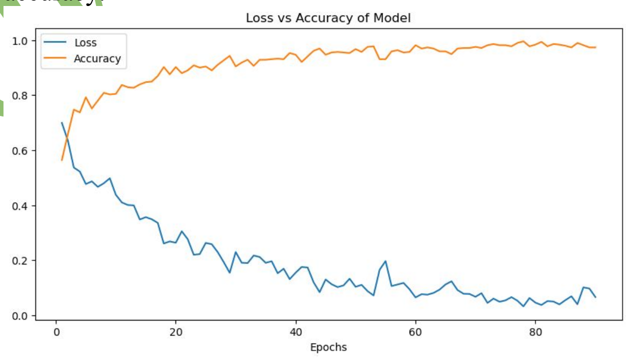

Fig. 4.Accuracy vs. loss curve (Epoch 90)

After using a various epochs value the accuracy and loss can be seen in Table 2 as follows:

TABLE 2. ACCURACY AND LOSS ON DIFFERENT EPOCHS

Epochs	Accuracy	Loss
10	0.7943	0.4812
20	0.9002	0.2756
50	0.9613	0.1154
90	0.9735	0.0657

VI. Result and Discussion

After the CNN modelling, it is tested against the images to test the accuracy of the model. Testing is carried out by using 2 sets of images each representing different classes where set-1contains eye images with cataract and set-2 contains the images of the normal eye. The model is optimised to such an extent that it predicts the right value for the input image which can be seen in Fig. 7 and Fig. 7. The confusion matrix representing the effectiveness of the proposed model is shown in Fig. 5.

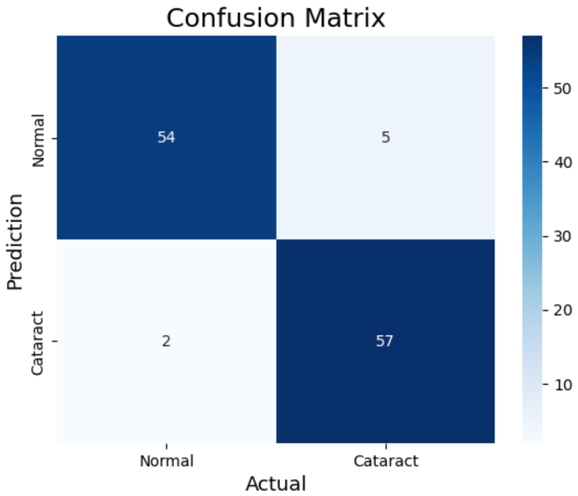

Fig. 5. Confusion matrix

```
img = image.load_img(test_image_path, target_size=(224, 224))
img_array = image.img_to_array(img)
img_array = np.expand_dims(img_array, axis=0)
img_array /= 255.0

prediction = model.predict(img_array)

binary_prediction = 1 if prediction[0] > 0.5 else 0

label_string = 'Normal' if binary_prediction == 1 else 'Cataract'
print("Actual: Cataract")
print(f"Predicted: {label_string}")

1/1 [==============================] - 0s 43ms/step
Actual: Cataract
Predicted: Cataract
```

Fig. 6.Cataract eye

```
test_image_path = 'normal.png'
img = image.load_img(test_image_path, target_size=(224, 224))
img_array = image.img_to_array(img)
img_array = np.expand_dims(img_array, axis=0)
img_array /= 255.0

prediction = model.predict(img_array)

binary_prediction = 1 if prediction[0] > 0.5 else 0

label_string = 'Normal' if binary_prediction == 1 else 'Cataract'
print("Actual: Normal")
print(f"Predicted: {label_string}")

1/1 [==============================] - 0s 25ms/step
Actual: Normal
Predicted: Normal
```

Fig. 7. Normal eye

As you can see in Fig. 8. The data has been labelled as normal and cataract. The normal eye is represented by the binary digit 1 whereas the cataract eye is represented by binary digit 0.

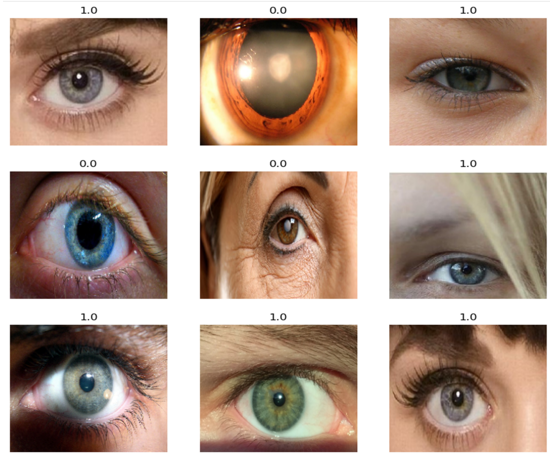

Fig. 8. Data being labelled by the model as cataract or normal

The Fig. 9 shows the different layers and the weight of the final model.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 222, 222, 32)      896

 max_pooling2d (MaxPooling2  (None, 111, 111, 32)      0
 D)

 conv2d_1 (Conv2D)           (None, 109, 109, 64)      18496

 max_pooling2d_1 (MaxPoolin  (None, 54, 54, 64)        0
 g2D)

 conv2d_2 (Conv2D)           (None, 52, 52, 128)       73856

 max_pooling2d_2 (MaxPoolin  (None, 26, 26, 128)       0
 g2D)

 conv2d_3 (Conv2D)           (None, 24, 24, 256)       295168

 max_pooling2d_3 (MaxPoolin  (None, 12, 12, 256)       0
 g2D)

 flatten (Flatten)           (None, 36864)             0

 dense (Dense)               (None, 256)               9437440

 dropout (Dropout)           (None, 256)               0

 dense_1 (Dense)             (None, 1)                 257

=================================================================
Total params: 9826113 (37.48 MB)
Trainable params: 9826113 (37.48 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Fig. 9. Weight parameters in each layer of CNN model

VII. Conclusion

As we can observe from the graph, we almost got close to 97% accuracy in cataract detection using the Convolution Neural Network. When the values of epoch's were increased to 20, we got a very good accuracy initially which is useful for initial diagnosis, and another thing that can be observed is that the loss decreases and accuracy increases till 20 epoch's after that curvature of increment of accuracy, and curvature of decrement of loss decreases and gives us a plateau. Hence data in this research is sufficient to conclude that the Convolution Neural Network created is very much capable for diagnosis of cataract with an exceptional precision.

References

[1] I. Weni, P. E. P. Utomo, B. F. Hutabarat, and M. Alfalah, "Detection of Cataract Based on Image Features Using Convolutional Neural Networks," *IJCCS (Indonesian Journal of Computing and Cybernetics*

Systems), vol. 15, no. 1, p. 75, Jan. 2021, doi: 10.22146/IJCCS.61882.

[2] M. K. Hasan *et al.*, "Cataract Disease Detection by Using Transfer Learning-Based Intelligent Methods," *Comput Math Methods Med*, vol. 2021, 2021, doi: 10.1155/2021/7666365.

[3] T. Ganokratanaa, M. Ketcham, and P. Pramkeaw, "Advancements in Cataract Detection: The Systematic Development of LeNet-Convolutional Neural Network Models," *Journal of Imaging 2023, Vol. 9, Page 197*, vol. 9, no. 10, p. 197, Sep. 2023, doi: 10.3390/JIMAGING9100197.

[4] R. Salama, F. Al-Turjman, M. Aeri, and S. P. Yadav, "Intelligent Hardware Solutions for COVID -19 and Alike Diagnosis - A survey," *2023 International Conference on Computational Intelligence, Communication Technology and Networking, CICTN 2023*, pp. 796–800, 2023, doi: 10.1109/CICTN57981.2023.10140850.

[5] R. Salama, F. Al-Turjman, D. Bordoloi, and S. P. Yadav, "Wireless Sensor Networks and Green Networking for 6G communication- An Overview," *2023 International Conference on Computational Intelligence, Communication Technology and Networking, CICTN 2023*, pp. 830–834, 2023, doi: 10.1109/CICTN57981.2023.10141262.

[6] R. Salama, F. Al-Turjman, S. Bhatia, and S. P. Yadav, "Social engineering attack types and prevention techniques- A survey," *2023 International Conference on Computational Intelligence, Communication Technology and Networking, CICTN 2023*, pp. 817–820, 2023, doi: 10.1109/CICTN57981.2023.10140957.

[7] R. Salama, F. Al-Turjman, P. Chaudhary, and S. P. Yadav, "Benefits of Internet of Things (IoT) Applications in Health care - An Overview," *2023 International Conference on Computational Intelligence, Communication Technology and Networking, CICTN 2023*, pp. 778–784, 2023, doi: 10.1109/CICTN57981.2023.10141452.

[8] R. Salama, F. Al-Turjman, M. Aeri, and S. P. Yadav, "Internet of Intelligent Things (IoT) - An Overview," *2023 International Conference on Computational Intelligence, Communication Technology and Networking, CICTN 2023*, pp. 801–805, 2023, doi: 10.1109/CICTN57981.2023.10141157.

[9] M. S. Junayed, M. B. Islam, A. Sadeghzadeh, and S. Rahman, "CataractNet: An automated cataract detection system using deep learning for fundus images," *IEEE Access*, vol. 9, pp. 128799–128808, 2021, doi: 10.1109/ACCESS.2021.3112938.

[10] A. Shetty, K. Bathija, and R. L. Priya, "Cataract Eye Detection Using Machine Learning Models," *Lecture Notes in Networks and Systems*, vol. 164, pp. 919–929, 2021, doi: 10.1007/978-981-15-9774-9_84/COVER.

[11] S. Kumar, S. Pathak, and B. Kumar, "Automated detection of eye related diseases using digital image processing," *Handbook of Multimedia Information Security: Techniques and Applications*, pp. 513–544, Jul. 2019, doi: 10.1007/978-3-030-15887-3_25/COVER.

[12] "Cataract Detection Using Deep Learning," Jul. 2023, doi: 10.21203/RS.3.RS-3178940/V1.

[13] S. Pathak, R. Raj, K. Singh, P. K. Verma, and B. Kumar, "Development of portable and robust cataract detection and grading system by analyzing multiple texture features for Tele-Ophthalmology," *Multimed Tools Appl*, vol. 81, no. 16, pp. 23355–23371, Jul. 2022, doi: 10.1007/S11042-022-12544-5/FIGURES/12.

[14] "Cataract dataset." Accessed: Dec. 15, 2023. [Online]. Available: https://www.kaggle.com/datasets/nandanp6/cataract-image-dataset