Eye Disease Detection using RESNET

Nihal Bhandary[1], Anish Adnani[2]

[1,2]Student, Department of Computer Engineering, Vivekanand Education Society's Institute of Technology, Mumbai, India

---

Abstract— *Among several eye diseases, cataract is one of the prevalent diseases. An early diagnosis of cataracts can hugely impact n reducing the rates of cataracts worldwide. The proposed system makes use of the RESNET model to do early detection of the same. This system does not need expensive and heavy equipment based on a fundus image. It uses regular eye images to detect cataracts. The model aims to achieve not only accuracy in the prediction but also to ease some problems in the current method of diagnosis and treatment of the disease.*

***Keywords* — Neural networks, cataract, computer vision, CNN (Convolutional Neural Network), RNN (Recurrent Neural Network).**

1. INTRODUCTION

Visual impairment is considered one of the major global health problems[1]. A cataract is found to be one of the leading causes of impairment and even blindness. The system currently in place is too cumbersome and requires a lot of resources to detect and treat cataracts.

The current methods of diagnosis are a visual acuity test, slit-lamp examination, and a retinal exam. A visual acuity test is a test where a chart with symbols is placed at a distance of 6 meters or 20 feet. The person is supposed to respond to the questions about their ability to discern shapes. The chart has symbols of different shapes and sizes. As the person is told to describe them, the ophthalmologist notes these responses. Based on which an individual is provided with a score which is indicative of their visual ability. This test is very helpful in knowing the extent or onset of cataracts. However, it is very difficult to diagnose cataracts at very early stages when a person's vision is not affected by cataracts yet.

Another method of cataract detection is a slit-lamp examination. This method involves the use of a "slit-lamp" which is a combination of a microscope and a bright light source. This is also known as biomicroscopy. An ophthalmologist uses this apparatus to examine the structures inside a person's eye. This helps in determining whether the person's eye has started developing any abnormalities due to cataract or other eye diseases. The doctor checks conjunctiva, cornea, eyelids, iris, pupil, lens, sclera, and retina.

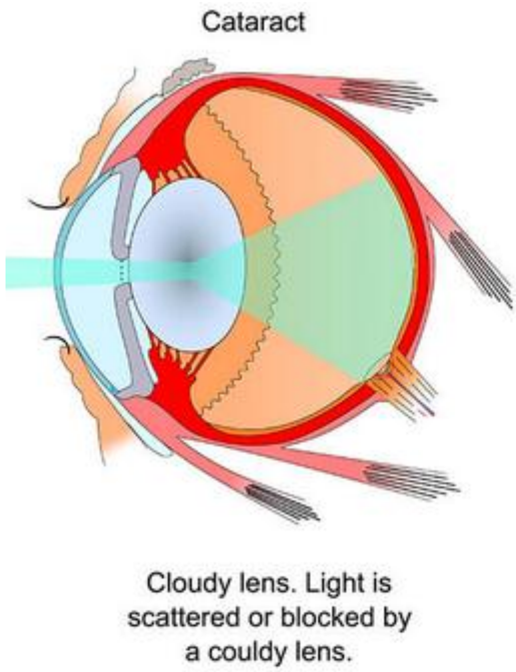

Fig. 1: Cataract eye visualization

Source: [10]

Fig 1 shows a cataract eye. The lens is blocked by the cataract which leads to partial or even complete blindness in some cases.

A retinal exam is another approach employed by ophthalmologists in diagnosing cataracts. The retinal examination mainly focuses on the back of one's eye. It is also known as ophthalmoscopy or funduscopy. The retina, choroid, and optic disk are examined with the help of fundus images.

One major drawback of this method is that it necessitates the use of such costly equipment and a need for an operator for this equipment. Since, in most developing and underdeveloped countries where the majority of cataract cases worldwide are present, the lack of medical equipment and skilled labor has restrained the effectiveness of this method. There needs to be a system or a procedure in place that will take this into the account. With the huge focus on the role of artificial intelligence in medical applications, it has become easier than before to build applications/tools in detecting and diagnosing diseases. With the increased proliferation of smartphone devices worldwide, the proposed system has a greater reach and might be very beneficial to the cause.

2. LITERATURE SURVEY

Some of the systems that used the same method in detecting cataract were studied. It can be concluded that while many of them succeeded in detecting cataracts they either fail in employing it on large scale or had higher requirements to be met.

The system[2] uses an object detection algorithm to detect and classify normal eye images. It has good accuracy and low false negatives for the images. However, it required the images to be of higher resolution than an image normally captured by a phone camera. This would hamper the reach of this model while distributing it on large scale.

The paper[3] did feature extraction on 2000 of normal eye images and using those features employed an RNN model to detect cataract. It has an accuracy rate of around 92% and a lower error rate. This model was not trained on a diverse dataset, so the results might vary when used on a larger and diverse dataset. Also, the model is bulky and might not be able to be run on a smartphone.

The model has applied another approach in detection and classification[5]. It uses a normal eye image and personal information to assess and take into account the other factors involved in cataract. The model has an over accuracy rate of 75% and this is mostly because of the use of a smaller neural network.

The paper mentions a diagnosis system using only normal medium resolution eye images[7]. It has a high accuracy of 98%. But this accuracy might be due to the overfitting of the dataset. The model has to overfit the training dataset. The actual implementation may yield a lower accuracy rate.

In conclusion, the existing systems had problems with either the model, dataset, or their inability to be used on a public scale. The proposed system aims to address these issues and would be able to be employed on a large scale basis.

3. PROPOSED METHODOLOGY

The proposed system would make use of regular eye images with resolution comparable to an average smartphone user. The images would then undergo several transformations and processing before the results are calculated by the neural network model.

3.1. System architecture

The detection model is depicted in fig.2 to define the flow of the proposed system.

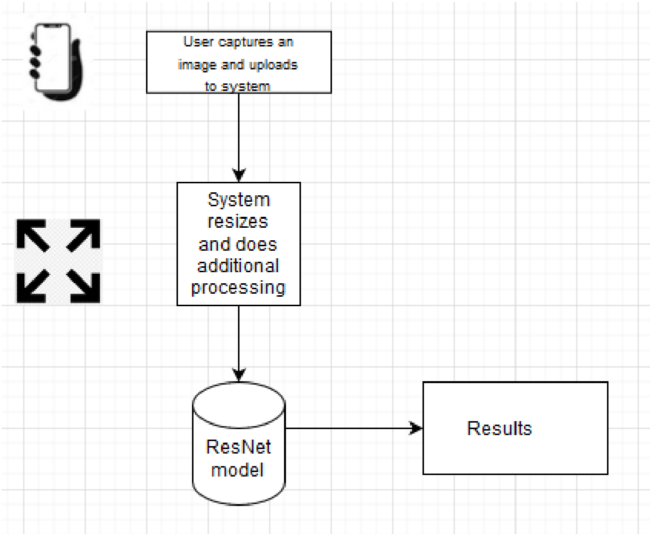

Fig.3: System architecture of the proposed system

The central aspect of the system, the neural network model is a TensorFlow model that can be employed on either smartphones or other computer devices. It is a ResNet model which has several convolutional layers to process the data and output a tangible result. The user has to capture a regular eye image and then feed it into the system. The system would feed this image to the model and calculate the result.

The system can be thought of as three chief parts – a pre-processing part, the model, and the user interface. The pre-processing part deals with the resizing, enhancing certain details of the eye image. The model does the heavy work of calculating the results and the user interface allows the end-user to interact with the system. This part allows a layman to use the system without having to bother with the specifics of either ophthalmology or machine learning.

3.2. METHODOLOGY

The standard flow of the proposed work is described below. Each step is necessary for the flow of the system. It involves loading an image, pre-processing, model, and decision making.

3.2.1. Load Image

The user captures an eye image and loads it into the system with the help of the user interface. The image capturing and loading marks the first step in the process.

3.2.2. Pre-processing

Pre-processing an image is an important step. It cleans out any irregularities which might create a problem with the model. The images are resized according to the needs of the model. The features which are more influential in decision making are enhanced so that they can be easily detected and used by the model.

3.2.3. Processed image into the trained model

The pre-processed image is then fed into the lower layer of the model. From there on, the image makes its way through successive layers of convolution and pooling. It ends up at the last layer of the model where the decision making takes place.

3.2.4. Decision making

The last layer is the decision-making layer. It is a fully-connected layer in the model. It uses the data from the previous layer to make a conclusion on the image fed to the model.

The ResNet-152 is a convolutional neural network model with 34 different layers. The model is not too heavy or bulky for a smartphone device with low processing capabilities.

3.3 The ResNet Model

Recently proposed residual networks (ResNets) get state-of-the-art performance on the ILSVRC2015 classification task and allow training of extremely deep networks up to more than 1000 layers [11]. The ResNet model is used for the classification of images. It consists of multiple 3×3 kernel-sized filters one after another. This is in contrast to the AlexNet with large kernel-sized filters (11 and 5 in the first and second convolutional layer, respectively). The reason for using the ResNet model is that AlexNet and VGG models' layers might be massive and prone to overfitting. The problem is greatly reduced in the ResNet model without going deep in layers. A dense layer represents a matrix-vector multiplication of the weights and inputs from the previous layer. The values in the matrix are the temporary parameters that get updated during backpropagation.

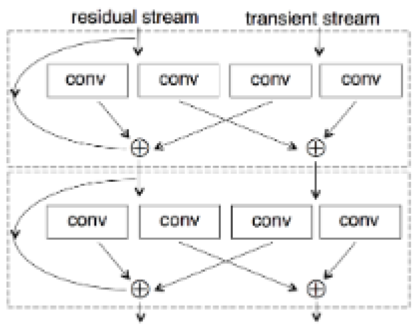

Fig. 6: ResNet layer

Source: [11]

Fig 3 represents the different layers present in the ResNet-152 model. This model consists of convolutional layers, Max Pooling layers, Activation layers, Fully connected layers. The convolutional layer is made of 3x3 size whereas the Max Pooling layer is made of 2x2 sized filters.

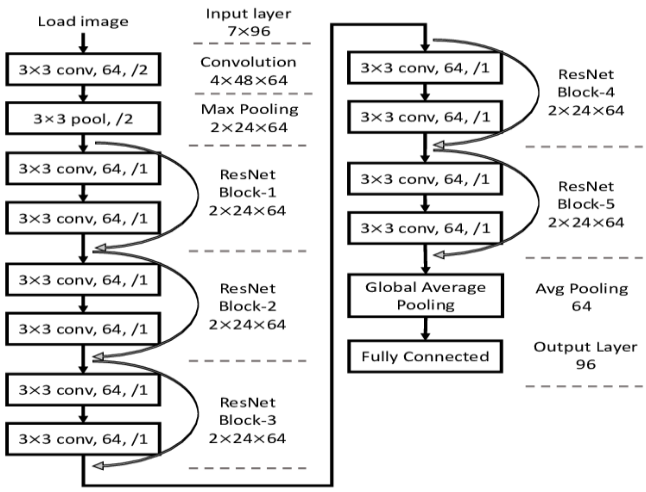

Fig.7: ResNet Architecture [11]

Fig 4 represents the architecture of the ResNet model. The diagram displays the number of filters and the size of inputs of each layer. There are 64 filters in 1st Convolution layer, the following 10 layers also contain the same amount of filters.

4. IMPLEMENTATION

The Convolutional Neural Network model is written in python3 and implemented in jupyter notebooks. We used CPU instead of GPU for processing. Tensorflow, Keras, and NumPy modules were used in making the model. Initially, the weights were randomized. During the course of the training, the weights were being successfully changed to the final values. The images were obtained from the open sources available for research use on kaggle.com.

5. RESULT AND ANALYSIS

After training our model on the training dataset, we observed the accuracy improved by increasing the epochs. But too many epochs created the problem of overfitting where the accuracy of the classification of new eye images tends to be poor. We found the optimal epoch value to be 30, where our testing and training losses were the lowest and hence model predictions were the most accurate.

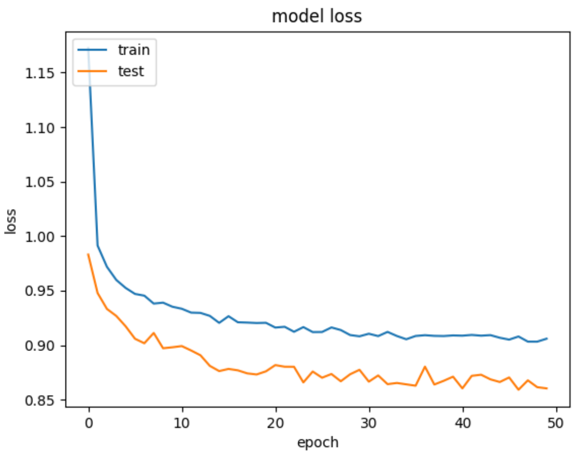

Fig 8: Model Loss with epochs

Figure 8 represents the model loss graph, the bifurcation of train and test is represented in the same figure.

```
model.fit(x_train, y_train,
          batch_size=batch_size,
          epochs=epochs,
          verbose=1,
          validation_data=(x_test, y_test))
score = model.evaluate(x_test, y_test, verbose=0)
print('Test loss:', score[0])
print('Test accuracy:', score[1])

x_train shape: (60000, 28, 28, 1)
60000 train samples
10000 test samples
Train on 60000 samples, validate on 10000 samples
Epoch 1/2
60000/60000 [==============================] - 6s 102us/sample - loss: 2.3050 - acc: 0.0848 - val_loss: 2.3045 -
val_acc: 0.0841
Epoch 2/2
60000/60000 [==============================] - 6s 103us/sample - loss: 2.3036 - acc: 0.0946 - val_loss: 2.3032 -
val_acc: 0.0925
Test loss: 2.3032085037231447
Test accuracy: 0.0925
```

72.3 % detected   0% detected

Fig. 10: Eye images with a percentage of cataract detected.

Figure 10 is the output of the system. The eye images are classified and then the percentage of surety is presented. Figure 11 shows the type and weights of the layer in the final model.

		image
	Input	array (size: 3×224×224)
conv1	ConvolutionLayer	array (size: 64×112×112)
bn_conv1	BatchNormalizationLayer	array (size: 64×112×112)
conv1_relu	Ramp	array (size: 64×112×112)
pool1_pad	PaddingLayer	array (size: 64×113×113)
pool1	PoolingLayer	array (size: 64×56×56)
2a	NetGraph (12 nodes)	array (size: 256×56×56)
2b	NetGraph (10 nodes)	array (size: 256×56×56)
2c	NetGraph (10 nodes)	array (size: 256×56×56)
3a	NetGraph (12 nodes)	array (size: 512×28×28)
3b1	NetGraph (10 nodes)	array (size: 512×28×28)
3b2	NetGraph (10 nodes)	array (size: 512×28×28)
3b3	NetGraph (10 nodes)	array (size: 512×28×28)
4a	NetGraph (12 nodes)	array (size: 1024×14×14)
4b1	NetGraph (10 nodes)	array (size: 1024×14×14)
4b2	NetGraph (10 nodes)	array (size: 1024×14×14)
4b3	NetGraph (10 nodes)	array (size: 1024×14×14)
4b4	NetGraph (10 nodes)	array (size: 1024×14×14)
4b5	NetGraph (10 nodes)	array (size: 1024×14×14)
4b6	NetGraph (10 nodes)	array (size: 1024×14×14)
4b7	NetGraph (10 nodes)	array (size: 1024×14×14)
4b8	NetGraph (10 nodes)	array (size: 1024×14×14)
4b9	NetGraph (10 nodes)	array (size: 1024×14×14)
4b10	NetGraph (10 nodes)	array (size: 1024×14×14)
4b11	NetGraph (10 nodes)	array (size: 1024×14×14)
4b12	NetGraph (10 nodes)	array (size: 1024×14×14)
4b13	NetGraph (10 nodes)	array (size: 1024×14×14)
4b14	NetGraph (10 nodes)	array (size: 1024×14×14)
4b15	NetGraph (10 nodes)	array (size: 1024×14×14)
4b16	NetGraph (10 nodes)	array (size: 1024×14×14)
4b17	NetGraph (10 nodes)	array (size: 1024×14×14)
4b18	NetGraph (10 nodes)	array (size: 1024×14×14)
4b19	NetGraph (10 nodes)	array (size: 1024×14×14)
4b20	NetGraph (10 nodes)	array (size: 1024×14×14)
4b21	NetGraph (10 nodes)	array (size: 1024×14×14)
4b22	NetGraph (10 nodes)	array (size: 1024×14×14)
5a	NetGraph (12 nodes)	array (size: 2048×7×7)
5b	NetGraph (10 nodes)	array (size: 2048×7×7)
5c	NetGraph (10 nodes)	array (size: 2048×7×7)
pool5	PoolingLayer	array (size: 2048×1×1)
flatten_0	FlattenLayer	vector (size: 2048)
fc1000	LinearLayer	vector (size: 1000)
prob	SoftmaxLayer	vector (size: 1000)
	Output	class

Fig. 11: Weight parameters in each layer of the model

6. CONCLUSION

A convolutional neural network model is thus developed using weights obtained by training the model on a large dataset of cataract and non-cataract eye images.

7. REFERENCES

[1] Felix Grassmann, Judith Mengelkamp, Caroline Brandl, Sebastian Harsch, Martina E. Zimmermann, Birgit Linkohr, Annette Peters, Iris M. Heid, Christoph Palm, Bernhard H.F. Weber, "A Deep Learning Algorithm for Prediction of Age-Related Eye Disease Study Severity Scale for Age-Related Macular Degeneration from Color Fundus Photography", Ophthalmology,Volume 125, Issue 9,2018,Pages 1410-1420,ISSN 0161-6420,https://doi.org/10.1016/j.ophtha.2018.02.037.

[2] Malik, S.; Kanwal, N.; Asghar, M.N.; Sadiq, M.A.A.; Karamat, I.; Fleury, M. Data Driven Approach for Eye Disease Classification with Machine Learning. Appl. Sci. **2019**, 9, 2789

[3] Turimerla Pratap, Priyanka Kokil, "Computer-aided diagnosis of cataract using deep transfer learning",Biomedical Signal Processing and Control,Volume 53,2019,101533,ISSN 1746-8094,https://doi.org/10.1016/j.bspc.2019.04.010.

[4] Brant, Arthur R. BS; Hinkle, John MD; Shi, Siyu BS; Hess, Olivia BS; Zubair, Talhah BS; Pershing, Suzann MD, MS; Tabin, Geoffrey C. MD Artificial Intelligence in Global Ophthalmology, Journal of Cataract & Refractive Surgery: September 03, 2020 - Volume Publish Ahead of Print - Issue - doi: 0.1097/j.jcrs.0000000000000407

[5] Akkara John Davis, Kuriakose Anju , " Role of artificial intelligence and machine learning in ophthalmology" Kerala Journal of Opthalmology Year : 2019,Volume: 31

[6] N. Hnoohom and A. Jitpattanakul, "Comparison of Ensemble Learning Algorithms for Cataract Detection from Fundus Images," 2017 21st International Computer Science and Engineering Conference (ICSEC), Bangkok, 2017, pp. 1-5, doi: 10.1109/ICSEC.2017.8443900.

[7] Martínez-Velasco, A.; Martínez-Villaseñor, L.; Miralles-Pechuán, L.; Perez-Ortiz, A.C.; Zenteno, J.C.; Estrada-Mena, F.J. The Relevance of Cataract as a Risk Factor for Age-Related Macular Degeneration: A Machine Learning Approach. Appl. Sci. **2019**, 9, 5550.

[8] Ting DSJ, Ang M, Mehta JS, et al

Artificial intelligence-assisted telemedicine platform for cataract screening and management: a potential model of care for global eye health

British Journal of Ophthalmology 2019;103:1537-1538.

[9] Yu F, Silva Croso G, Kim TS, et al. Assessment of Automated Identification of Phases in Videos of Cataract Surgery Using Machine Learning and Deep Learning Techniques. JAMA Netw Open. 2019;2(4):e191860. doi:10.1001/jamanetworkopen.2019.1860

[10] Elliot DB. Evaluating visual function in cataract Optometry and Vision Science : Official Publication of the American Academy of Optometry. 1993 Nov: 70(11): 896-902. DOI: 10.1097/00006324-199311000-00006.

[11] Resnet in Resnet: Generalizing Residual Architectures by Sasha Targ, Diogo Almeida, Kevin Lyman.