Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 44 No. 7
July 2023

Classification of Macular Diseases of Oct Images Using Hybrid Deep Learning Model

[1]S.T.Santhanalakshmi, [2] Nithiyashree.S, [3]Jeyahrini.V.J, [4]JeniferJoyAntony.P, [5] Kavitha Subramani

[1] Assistant Professor, Department of Computer Science and Engineering, Panimalar Engineering College, Chennai,

Tamilnadu , India

[2] [3] [4]Student, Department of Computer Science and Engineering, Panimalar Engineering College, Chennai,

Tamilnadu , India

[5] Professor, Department of Computer Science and Engineering, Panimalar Engineering College, Chennai,

Tamilnadu , India

[1]anandh.shantha@gmail.com [2]snithiyashree22@gmail.com,[3]harini210502@gmail.com,

[4]jeniferantonysp@gmail.com, [5] kavitha.bhskr@yahoo.com

Abstract— Retinal diseases can occur in any part of the eye. A thin layer of tissue on the inside back wall of your eye, can be affected by retinal diseases. These are caused by damage to the blood vessels at the back of the eye, causing fluid to leak. This fluid buildup can damage the retina and cause vision changes. Hypertension and cancer are few conditions that can cause defects in the eye. Convolutional neural networks (CNNs), a cutting-edge technology that enables effective disease detection in deep learning, have seen remarkable success in the classification of numerous eye illnesses. Upon diagnosis, neural networks can capture the colours and textures of lesions specific to respective diseases, which is similar to human decision-making. Convolutional neural networks were tested with different retinal features as input for effective retinal image classification. Understanding issues in the posterior the eyes is very tricky .The diseases that occur in the macular region of the eye due to fluid retention and blood vessel damages. The differences between the diseases are very minute for medical errors to happen .We are deploying a software that classifies the what kind of disease it is or no disease when the user provides their OCT[1] scan image . Thisproject is intended for users who wants to verify their OCT eye scan report to seek second consultation to doctors.

Index Terms—: Disease classification, Deep learning, Convolutional Neural networks, Retinal disease,OCT

INTRODUCTION

Vision and eye health are one of the most crucial things in human life, it needs to be preserved to maintain the life of the individuals. When the retina is damaged and diagnosed at a late stage, there is almost no opportunity to reverse the condition and heal it, meaning that the patient will likely lose some or all of their eyesight. Artificial intelligence (AI) has the potential to revolutionize disease diagnosis and management by performing classification difficult for human experts and by rapidly reviewing immense amounts of images. Despite its potential, clinical interpretability and feasible preparation of AI remains challenging. The development of convolutional neural network layers has allowed for significant gains in the ability to classify images and detect objects in a picture. These are multiple processing layers to which image analysis filters, or convolutions, are applied. The abstracted representation of images within each layer is constructed by systematically convolving multiple filters across the image, producing a feature map that is used as input to the following layer. Eye illnesses are mostly caused by retinal injury. By measuringcoherence tomography (OCT) can perform non-invasive cross- sectional imaging of internal biological tissue structures, enabling ophthalmologists to take a clear look at the back of theeye and identify retinal, macula, and optic nerve damage at an early stage. The purpose of this study is to provide an unique deep learning-based classification model to automatically classify .Utilizing retinal pictures from an Optical Coherence Tomography (OCT) instrument, one can visualise the various retinal disorders.

II.RELATED WORK

Based on the layer structure of OCT images,they[1] have create the ECL, a novel convolutional layer that uses linear kernels with different angles for convolution instead of the common kernel. To improve the model's robustness, we augmented the data by flipping it horizontally. The weights of the network were initialised using the Xavier initialization technique . In this state, instead of 64 convolution weights, there are only 8 trainable. weights in each kernel.At the generator,[2] deconvolution layers are followed by batch normalisation layers. Tanh activation is used on the final convolution layer. The convolution layers are followed on the discriminator side by ReLU activation with a negative regression line of 0.2 on the Layers of batch normalisation. The addition of a mini-batch discrimination layer to the discriminator network prevents mode collapse in the GAN. To prevent overfitting, an optimiseddrop-out factor of 70% is applied to all but the first convolutionlayer of the classification model. [3]To accomplish this goal, the images are patched; the extracted patches are then treated as sequences, and Recurrent Neural Networks are used to classify the images. Four pre-trained models are used and compared, including VGG16, ResNet152V2, NasnetMobile, and Densnet169, as well as a vision transformer model. According to the results, the proposed model achieved 99.38% test accuracy, which is higher than other models.[4]Propose a training pipeline for multi-labeled classification with unequal sample size and difficulty distribution. Directed the initial model's training by weighing the training loss with an inverse- frequency for each class. The model was adjusted for the class weights using the consolidated loss for each class and train for more variants, with the model focusing more on difficult samples and covering the shortcomings of the previous model with each iteration. Finally used Heuristic Stacking algorithm to combine all of the models in order to improve multi-label predictions beyond simple averaging. Our experimental findings show that the proposed method achieves an accuracy score of 88.24%.[5]Intends to introduce and analyse the performance of a Support Vector Machine based hybrid architecture and Transformer Vision to classify Optical Coherence Topography (OCT) Scan results in an attempt to automate the early diagnosis of these ocular deformities.By giving different combination of layers,with multiple flatten layer in between ,effective feature extraction is done. [6]Machine learning, using the KNN algorithm, and deep learning, using two convolutional neural network architectures, Cifar-10 , have been employed in this study. To begin, features are extracted from grayscale images are maximum, minimum, mean, variance, skewness, and kurtosis are fed into KNN for categorisation. The extraction of features in neural networks occurs inside the classifier itself, which works directly on the dataset been collected.

VGG 16architecture[7] was modified by adding more convolu

PROPOSED WORK

The proposed methodology has the following steps:

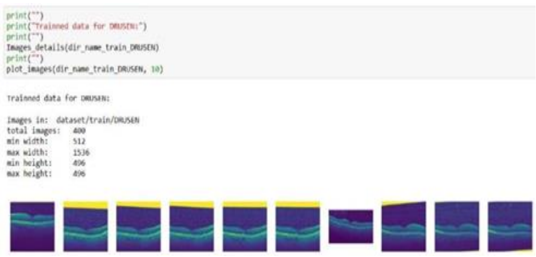
- Dataset Collection and
- Data analysis
- Data Preprocessing
- Data augmentation
- Algorithm Implementation
  - Manual Architecture
    - VGG 16 Architecture

- LeNet Architecture

tion layers and regularisation terms to achieve a result of 93.5.

Deployment Their disadvantage is that they consume a lot of computational

resources Customized pre-processing method to improve image quality.Network optimization, specially designed attention model, which pays more attention to critical regions containing pathological anomalies, is integrated into a typical deep learning network.Changes in the pooling layers such as downsampling and weight sharing were introduced in the proposed work[8] to enhance the efficiency of the model and were demonstrated to substantially reduce the parameters meant for training. Class activation mapping also was carried out, and the resulting image vaguely resembles the retina's actual colour OCT images. When contrasted with the current ResNet-50 model, the proposed system just used 6.9% of the learnable parameters and performed better it in classification. Due to the model's lower complexity and fewer learnable parameters compared to other models, the proposed work reallydoes have the potential to be utilized in practical uses.



Ophthalmologists can obtain cross-section imaging of the eye retina using optical coherence tomographic images.Effective disease detection could be accomplished with the help of digital image analysis methods.There are several methods for extracting features from eye scan report images. The proposed research[9] compares the efficacy of the artificially designed features. The dataset contains 32339 instances divided into four classes.Histogram of Oriented Gradient (HOG),[10-11] ResNet50,Local Binary Pattern (LBP), DenseNet-169, and are the feature extractors. As a result, deep neural network- based methodsoutperformed handcrafted features, with DenseNet and ResNet achieving 88% and 89% accuracy, respectively, compared to 50% and 42% for LBPand HOG .[12-13] Deep neuralnetwork-based methods also performed better on underrepresented classes.For classifying confocal tomography images of macular edema and age related macular degeneration is collected for binary classification. A CNN model[14] that was trained in prior is used to transfer its learnings to a new model.

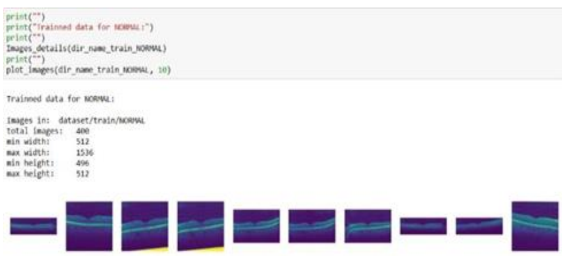
[15]The hyperparameters affect learning speed and quality. . Alexnet, Googlenet,ResNet successfully classified eye scan report images of AMD and DME after transfer learning.

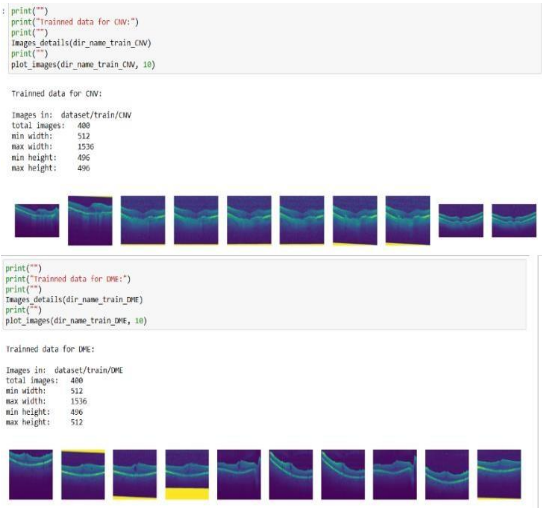
DATASET COLLECTION

Our intial phase of our project is to collect relevant datato train deep learning algorithms. Accuracy does not always increase with the size of the dataset. The datasets was acquired from Kaggle .Datasets are gathered from hospitals for this initiative. The collection includes Optical Coherence Tomographic scan pictures of Drusen, diabetic macular edema, and choroidal neo vascularization. It also includes pictures that are free of any eye conditions. 2000 photos in all have been gathered for this project. The dataset is then split into two parts: training dataset and test dataset. 80% of the data was utilized to teach and train an algorithm. 20% of the testing dataset is utilized to gauge how effectively a training method worked.

A) DATA ANALYSIS

This is a simple but very important step before pre-processsing the image.This procedure is used to analyse and understand thedimensions of the OCT scan images before pre-processing it.





B) DATA PRE-PROCESSING

The features that are extracted for the model to be trained are done in this phase where unwanted noises are removed througheffective extraction layers offered by Keras Library.

Convolutional layer:The main building blocks of convolutional neural networks are convolutional layers. A convolution is the straightforward process of applyinga filter to an input to produce activation. A feature map, produced by repeatedly applying the same filter to input,shows the positions and intensity of any recognised features in an input, such as an image.

Max pooling: The pooling procedure known as "max pooling" selects the most extreme component from the region of the element map that the channel covers. The outcome of the max-pooling layer would then be an element map with the most obvious highlights of the priorcomponent map.

*The Flatten layer:*A three-dimensional tensor of shape (height, width, channels) is the result of a convolutional layer or a pooling layer, where each element stands for a feature map or a channel. This tensor is transformed intoa one-dimensional vector of shape by the Flatten layer, where each element stands for a feature that was taken from the input picture. The Flatten layer is often positioned after the CNN's convolutional phase and before its fully connected phase, which comprises of one or more dense layers and does the final classification or regression operation. The fully connected layers can receive a fixed-length input vector regardless of the size of the input picture by flattening the output of the convolutional layers, simplifying the design and reduce the number of parameters.

Feedforward layer: Each neuron or unit in the feed-forward layer is linked to every other neuron in the layer above, and each connection has a weight parameter that is learnt throughout the training process. A vector of activations representing a high-level abstraction of the input is the result of a feed forward layer. For classification or regression tasks, the feed forward layer is often utilised as the last layer in a neural network design. The size of the output vector and the model's complexity are both influenced by the density of neurons in the Dense layer. While a lesser number of neurons might restrict the model's ability to capture complicated information, a higher number of neurons can make the model more expressive but also raise the danger of overfitting.To reduce the gap between the anticipated output and the actual output, the weights of the Feedforward layer are changed during training using an optimisation approach such stochastic gradient descent (SGD). The job will determine the loss function, such as mean squared error for regression or binary cross-entropy for binary classification.The Feedforward layer, in general, is an essential part of many neural network architectures and has been successfully used in a variety of applications, including recommender systems, natural language processing, and image recognition.

Batch normalization:A technique for creating extremely deep neural networks called batch normalisation normalises thecontributions to a layer for each tiny batch. As a result, fewer training sessions are required to adequately prepare large companies, which has the impact of increasing learning opportunities

C) DATA AUGMENTATION

Data augmentation is a technique used in machine learning and computer vision to artificially increase the size of a dataset by creating new variations of the existing data. The goal is to improve the generalization ability of the trained model by exposing it to a wider range of input variations.In the context of image classification, data augmentation involves applying a set of image transformations, such as scaling,, flipping, cropping ,rotation, and color shifting, to the original images to generate new images with different variations. For example, an image of a cat can be rotated by a few degrees, flipped horizontally, or scaled up or down to create new images that are variations of the original cat image .Tominimize overfitting and enhance the precision and resilience of the trained model, data augmentation can be employed to expand the amount and variety of a training dataset. As a result, the model's capacity to generalize to new pictures is enhanced. It can also help the model learn to become more invariant to variations in the input images, such as changes in lighting, perspective, and orientation.

Table 5.2 Output shape parameter details of VGG-16 architecture

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d 2 (Conv2D)	(None, 224, 224, 64)	1792
conv2d_3 (Conv2D)	(None, 224, 224, 64)	36928
max_pooling2d (MaxPooling2D)	(None, 112, 112, 64)	0
conv2d_4 (Conv2D)	(None, 112, 112, 128)	73856
conv2d_5 (Conv2D)	(None, 112, 112, 128)	147584
max_pooling2d_1 (MaxPooling 2D)	(None, 56, 56, 128)	0
conv2d_6 (Conv2D)	(None, 56, 56, 256)	295168
conv2d_7 (Conv2D)	(None, 56, 56, 256)	590080
conv2d 8 (Conv2D)	(None, 56, 56, 256)	590080
max_pooling2d_2 (MaxPooling 2D)	(None, 28, 28, 256)	0
conv2d_9 (Conv2D)	(None, 28, 28, 512)	1180160
conv2d_10 (Conv2D)	(None, 28, 28, 512)	2359808
conv2d 11 (Conv2D)	(None, 28, 28, 512)	2359808
max_pooling2d_3 (MaxPooling 2D)	(None, 14, 14, 512)	0
conv2d_12 (Conv2D)	(None, 14, 14, 512)	2359808
conv2d_13 (Conv2D)	(None, 14, 14, 512)	2359808
conv2d 14 (Conv2D)	(None, 14, 14, 512)	2359808
max pooling2d 4 (MaxPooling 2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 256)	6422784
dense_1 (Dense)	(None, 128)	32896
dense_2 (Dense)	(None, 4)	516

Total params: 21,170,884
Trainable params: 21,170,884
Non-trainable params: 0

I. ARCHITECTURE IMPLEMENTATION

A. Manual Architecture

We designed an architecture manually to see how the dataset performs.We used Keras library from Tensorflow to build this Sequential model for image classification using convolutional neural networks (CNNs). The model obtained an accuracyof 83.6%.Here is a brief summary of the model architecture.

The first layer is a 2D convolutional layer with 58 filters of size 3x3, using ReLU activation function, and input_shape of

(496,496), which corresponds to an RGB image of size 496x496 pixels.

The second layer is a max pooling layer with pool size of 3x3, stride of 2x2 and padding='same', which reduces the spatial dimensions of the output by half.

The third layer is another 2D convolutional layer with 26 filters of size 3x3, using ReLU activation function.

The fourth layer is another max pooling layer with pool size of 3x3 and default stride and padding values, which again reducesthe spatial dimensions of the output by half.

The fifth layer is a flatten layer, which flattens the output from the previous layer into a 1D vector.

The sixth layer is a fully connected (dense) layer with 30 units and ReLU activation function.

The seventh layer is another dense layer with 12 units and ReLU activation function.

The final layer is a dense layer with 4 units and softmax activation function, which produces the classificationprobabilities for each of the 4 output classes.

The model is compiled with the Adam optimizer, categorical cross-entropy loss function and accuracy metric.

The model summary provides a more detailed overview of the layers and their output shapes improved upon and modified for improved results when training a VGG16 model for picture categorization.

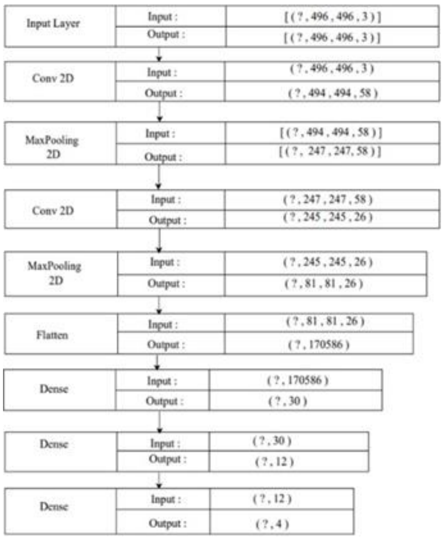


Figure of Proposed Manual Architecture

B. VGG-16 Architecture

C. LeNet Architecture

For the LeNet convolutional neural network architecture for image categorization.Two sets of convolutional and average pooling layers make up the LeNet architecture, which is then followed by a flattening layer, two fully connected layers, and an output layer with softmax activation.The training photos are scaled to values between 0 and 1, and random transformations like shearing, zooming, and horizontal flipping are applied in the code using the Keras Image Data Generator class. Just resizing is done to the test photos. This categories cross-entropyloss function, the RMSprop optimizer, and accuracy are used to

II. ACCURACY COMPARISON

The architecture designed manually yielded 83.59%,The VGG-16 architecture yielded 30% accuracy and LeNet architecture yielded 97.75% accuracy.

Screenshot of Manual architecture

```
334
Epoch 30/30
16/16 [==============================] - ETA: 
Epoch 30: accuracy improved from 0.79492 to 0.
```

Screenshot of VGG-16 architecture

build the model. The best model's weights are saved using the model checkpoint callback nce the training process has This VGG16 model was created with TensorFlow's Keras API. An ImageDataGenerator is used to create augmented pictures while the model is being trained on a dataset of photos dividedinto four classes.Several Conv2D layers with ReLU activation are followed by MaxPooling layers in the model design. The probabilities for each class are produced by completely linked dense layers with softmax activation. The categorical crossentropy loss function and the accuracy metric are used in the model's construction.The highest accuracy model weights on the validation set are saved during training using the ModelCheckpoint callback. Following training, two graphs displaying the training and validation accuracy/loss curves throughout the epochs are created.Overall, this code may be Completed 40 epochs with a batch size of 32.

The algorithm makes predictions on the test set after training, computes the classification report and confusion matrix, and outputs the findings. Moreover, it displays the accuracy and lossover epochs for the training history.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 75, 75, 32)	896
max_pooling2d (MaxPooling2D)	(None, 37, 37, 32)	0
conv2d_1 (Conv2D)	(None, 12, 12, 128)	36992
max_pooling2d_1 (MaxPooling2D)	(None, 6, 6, 128)	0
flatten (Flatten)	(None, 4608)	0
dense (Dense)	(None, 256)	1179904
dense_1 (Dense)	(None, 4)	1028

Total params: 1,218,820
Trainable params: 1,218,820
Non-trainable params: 0

Table 5.3 Output shape parameter details of LeNet architecture

```
2656
Epoch 20/20
50/50 [==============================] - ETA: 0s -
Epoch 20: accuracy improved from 0.29688 to 0.3000
50/50 [==============================] - 154s 3s/s
```

Screenshot of Lenet architecture

```
0.8409
Epoch 99/100
50/50 [==============================] - ETA: 0s - loss: 0.0930 - accuracy: 0.9675
Epoch 99: accuracy did not improve from 0.97812
50/50 [==============================] - 27s 544ms/step - loss: 0.0930 - accuracy: 0.9675 - val_loss: 0.6
0.8344
Epoch 100/100
50/50 [==============================] - ETA: 0s - loss: 0.0637 - accuracy: 0.9775
Epoch 100: accuracy did not improve from 0.97812
50/50 [==============================] - 24s 477ms/step - loss: 0.0637 - accuracy: 0.9775 - val_loss: 0.6
0.8469
```

III. CLASSIFICATION REPORT

Screenshot of Manual architecture

	precision	recall	f1-score	support
CNV	0.86	0.84	0.85	101
DME	0.78	0.55	0.64	84
DRUSEN	0.71	0.49	0.58	65
NORMAL	0.62	0.99	0.76	78
accuracy			0.73	328
macro avg	0.74	0.72	0.71	328
weighted avg	0.75	0.73	0.72	328

Screenshot of VGG-16 architecture

Screenshot of LeNet architecture

	precision	recall	f1-score	support
CNV	1.00	0.94	0.97	53
DME	0.75	1.00	0.85	53
DRUSEN	1.00	0.94	0.97	53
NORMAL	1.00	0.81	0.89	62
accuracy			0.92	221
macro avg	0.94	0.92	0.92	221
weighted avg	0.94	0.92	0.92	221

	precision	recall	f1-score	support
CNV	0.33	0.33	0.33	101
DME	0.34	0.33	0.34	84
DRUSEN	0.21	0.15	0.17	67
NORMAL	0.28	0.35	0.31	80
accuracy			0.30	332
macro avg	0.29	0.29	0.29	332
weighted avg	0.29	0.30	0.29	332

CONFUSION MATRIX

Figure 1 is the confusion matrix for manual architecture Figure 2 is the confusion matrix for VGG-16 architecture Figure 3 is the confusion matrix for LeNet architecture

Figure1

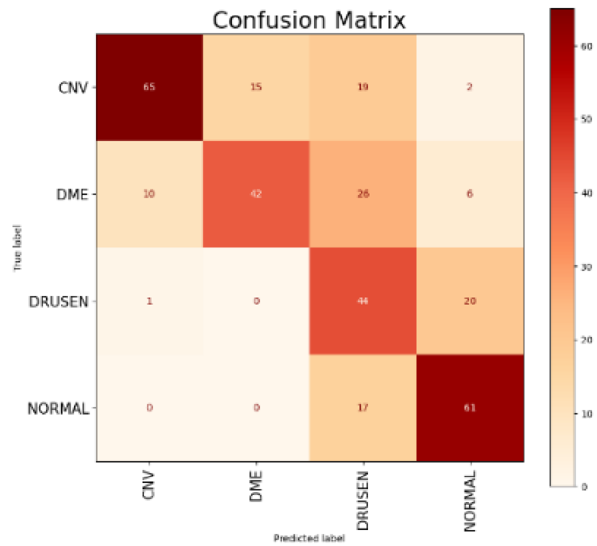


Figure2

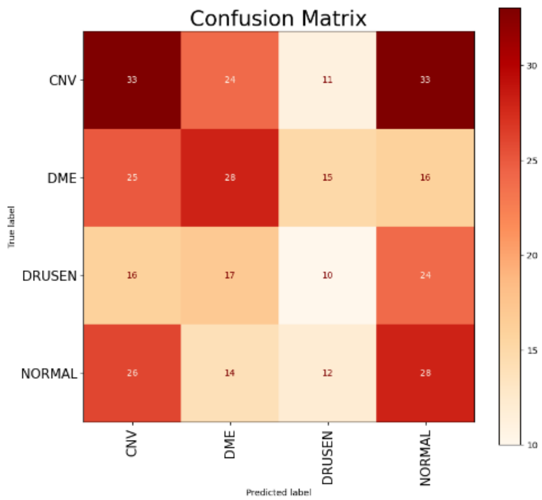
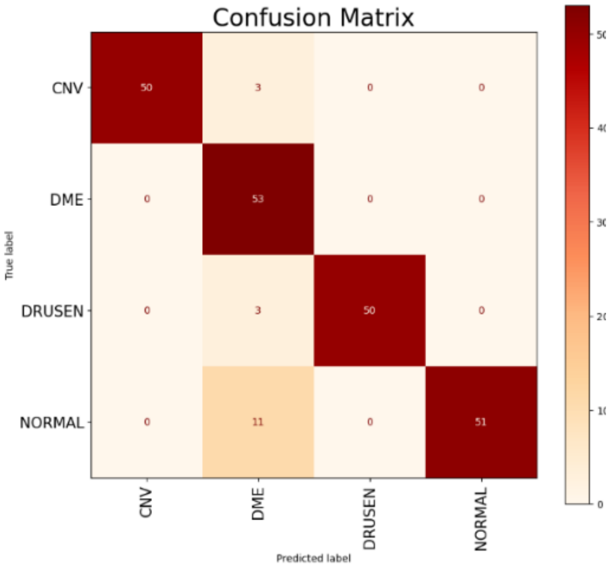
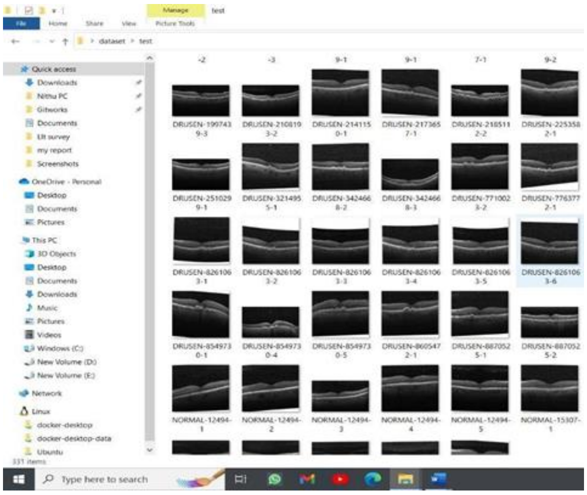
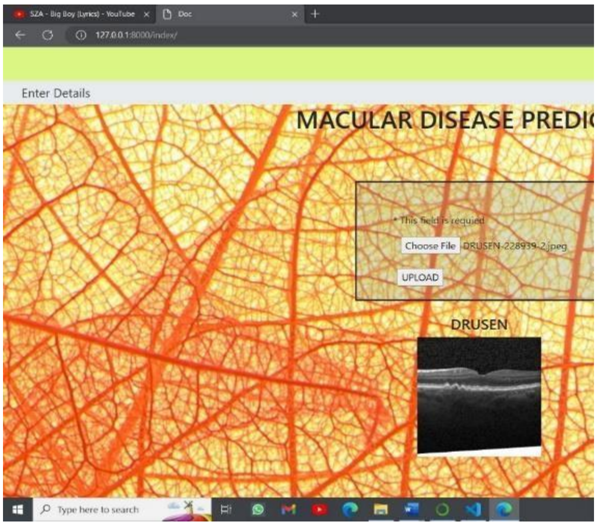


Figure3







Correct Classification of Disease

Test Images Of All 4 Classes

IV. CONCLUSION

Medical errors are very prevalent these days because of the tight schedule for the doctors to address patients. Since the difference between the doctor to patient ratio is very high, there is a need for certain things to be automated. Our proposed work will help the patients to verify the illness they have been diagnosed over the internet. By feeding their eye scan report to our system will help them classify their disease which is equivalent to seeking second consultation. Our system has accuracy about 97.75%. Since the defects that happen in the eye look very similar but require different treatments, wrong diagnose will detoriate the eye health.

V. REFERENCES

1. Automatic Classification of Macular Diseases from OCT Images Using CNN Guided with Edge Convolutional Layer Ebrahim Nasr Esfahani, Parisa Ghaderi Daneshmand, Hossein Rabbani,Gerlind Plonka
2. 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) Scottish Event Campus, Glasgow, UK, July 11-15, 2022
3. A Data-Efficient Approach for Automated Classification of OCT Images Using Generative Adversarial Network Vineeta Das , Samarendra Dandapat∗, and Prabin Kumar Bora∗, IEEE Sensor CouncilVOL. 4, NO. 1, JANUARY 2020
4. Kalaivani, K., Kshirsagarr, P.R., Sirisha Devi, J., Bandela, S.R., Colak, I., Nageswara Rao, J. and Rajaram, A., 2023. Prediction of biomedical signals using deep learning techniques. *Journal of Intelligent & Fuzzy Systems*, (Preprint), pp.1-14

5. A Multiclass Retinal Diseases Classification Algorithm using Deep Learning Methods
6. 2022 10th RSI International Conference on Robotics and Mechatronics (ICRoM)R. Behbahani NejadJ. KhoramdelA. GhanbarzadehM. SharbatdarE. Najafi . 24 November 2022
7. Ilakkiya, N. and Rajaram, A., 2023. Blockchain-assisted Secure Routing Protocol for Cluster-based Mobile-ad Hoc Networks. *INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL*, *18*(2).
8. Multi-Expert Deep Networks for Multi-Disease Detection in Retinal Fundus ImagesLinquan Lyu; Imad Eddine Toubal; K. Palaniappan 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)
9. Indira, D.N.V.S.L.S., Ganiya, R.K., Ashok Babu, P., Xavier, A., Kavisankar, L., Hemalatha, S., Senthilkumar, V., Kavitha, T., Rajaram, A., Annam, K. and Yeshitla, A., 2022. Improved artificial neural network with state order dataset estimation for brain cancer cell diagnosis. *BioMed Research International*, *2022*.
10. Retinal Malady Classification Using AI: A novel ViT-SVM combination architecture. 2022 6th International Conference on Computing Methodologies and Communication (ICCMC) 13 April 2022
11. Artificial Intelligence applied to the classification of retinal diseases in Optical Coherence Tomography images. Beatriz Silva Brasil; Auzuir Ripardo de Alexandria; Glendo de Freitas Guimarães. 2022 Fifth International Conference on Vocational Education and Electrical Engineering (ICVEE) November 2022
12. Rathish, C.R. and Rajaram, A., Hierarchical Load Balanced Multipath Routing Protocol for Wireless Sensor Networks.
13. Deep learning for identifying and classifying retinal diseasesby Mohamed Berrimi; Abdelouaheb Moussaoui 2nd International Conference on Computer and Information Sciences (ICCIS) 2020
14. OctNET: A Lightweight CNN for Retinal Disease Classification from Optical Coherence Tomography Images Neurocomputing,ElsevierVolume 462, 28 October 2021, Pages 491-505
15. DeepRetino: Ophthalmic Disease Classification from Retinal Images using Deep LearningFatima Zahra Belharar; Nabila Zrira .2022 IEEE 9th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT)
16. Harikrishnan, G. and Rajaram, A., 2017. Improved throughput based recognition connection denies for aggressive node in wireless sensor network. *Journal of Computational and Theoretical Nanoscience*, *14*(12), pp.5748-5755.
17. Classification of age-related macular degeneration using convolutional-neural-network-based transfer learningYao-Mei Chen, Wei-Tai Huang,Wen-Hsien Ho & Jinn-Tsong TsaiVolume 22 Supplement 5 Proceedings of the International Conference on Biomedical Engineering Innovation (ICBEI) 2021