International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-09, Issue-04, July 2023

A System for Detecting Retinal Disease Using Deep Learning Techniques

G.Mahalakshmi, Teaching Fellow, Department of Information Science and Technology, CEG Campus Anna University Chennai & India, mlakshmig27@gmail.com

V.G Dharsana Dharani, M.Tech- IT Student, Department of Information Science and Technology, CEG Campus Anna University Chennai & India, harsanavg2000@gmail.com

Abstract **The unfortunate truth is that 75% of India's 15 million blind individuals once had treatable blindness. For every doctor in India, there are 10,000 patients. A number of conditions can cause vision impairment, including cataracts, glaucoma, and trachoma, among others. Any area of the retina might get damaged by retinal disorders. Untreated retinal conditions can cause significant vision loss and perhaps blindness. Some retinal illnesses are treatable with early diagnosis, while others can be slowed down or controlled to protect or even restore eyesight. Early-stage disorders that go untreated are the main causes of blindness in India, according to studies. Only with early, accurate diagnosis can the progression of many eye conditions be stopped. To get a precise diagnosis, it is required to analyze a variety of symptoms. Large neural networks with neurons connected to one another are used in deep learning and have the capacity to adjust their hyper-parameters whenever new, updated data is received. Because of this technology, computers can now learn new things on their own without any direct human programming. We suggested a system to implement this procedure in order to effectively and early diagnose the condition. The suggested model distinguishes between a healthy eye and a diseased eye and uses a straightforward user interface to notify the user about eye-related issues and illnesses. The suggested paradigm would be able to direct people toward understanding their state of vision. To get their eyes checked, the model gives a thorough report of the ailment to the local eye physicians.**

Keywords — Deep Learning, Cataracts, Glaucoma, Trachoma, CNN

I. Introduction

Every single individual depends on their eyes to see and perceive the world around them, making them a vital component of human life. Sight, which makes up 80% of the information we take in, is one of the most important senses. One of the main reasons why people in their 80s and 90s go blind is age-related macular degeneration (AMD). Due to AMD, 8.7% of people in their 60s and older are totally blind[2]. By taking good care of our eyes, we may reduce our risk of going blind and losing our eyesight while also keeping an eye out for any eye disorders that may be developing, such as glaucoma and cataracts. The majority of people eventually have eye problems. While some eye problems are mild, easily treatable at home, and go away on their own, some serious eye problems necessitate the aid of skilled medical professionals. When these eye conditions are promptly and thoroughly identified. Electroencephalogram (EEG) analysis and auxiliary analysis of nervous system illnesses require accurate eye blink artifact identification, particularly when frontal epileptic form discharges are present [7]. A medical dataset contains a sizable quantity of information about diseases, prescription medications, hospitalizations, patient insurance, etc. Diabetes mellitus, hypertension, ageing, cardiovascular disease, and a family history of CKD are all contributing factors to CKD [13].

II. Literature Survey

Vision issues can be brought on by a number of eye conditions, including trachoma, cataracts, and corneal ulcers. Only when these eye diseases are found early on can they be properly treated. All of these eye disorders' symptoms are visible to the naked eye. It is vital to examine a number of symptoms in order to accurately diagnose eye problems. In light of this, Hu, Xiaoyan, and colleagues [1] presented the "GLIM-Net: chronic glaucoma forecast transformer for irregularly sampled sequential fundus images. The usefulness of the two modules we provide is also supported by the ablation tests, which can serve as a useful benchmark for the improvement of the Transformer model. Describe how to most accurately train a classifier for glaucoma (the second major eye issue) using GLIM-Net. 3000 base photos are used to train the classifier in this model. "Force and Velocity Based Puncture Detection in Robot-Assisted Retinal Vein Cannulation: In-Vivo Study,"

 DOI : 10.35291/2454-9150.2023.0144

as proposed by Alamdar, Alireza, et al [3], is a technology that will advance in the following years. The first in-vivo retinal vein cannulation trial on rabbit eyes employing sensorized metal needles is described in this study, along with an investigation into puncture detection. Using a sensing tool, a classifier can distinguish retinal veins from other eye parts. Find a cure for the second most prevalent retinal vascular disease, which affects 16.4 million individuals globally. This classification sensing method is not universal for treating other veins or other eye conditions.

In this the author proposed a model to fill the gap between regular and UWF fundus and provide more UWF fundus images for training, author suggest the usage of a modified cycle generative adversarial network (CycleGAN) model. To enhance and govern the quality of the generated data, a consistency regularization term is recommended in the losses of the GAN. The proposed strategy makes data gathering easier because it doesn't necessary that photos from the two domains be paired or even that their semantic labels match [4].

Ngo, Lua, Jaepyeong Cha, and Jae-Ho Han[5] suggested Deep Neural Network Regression for Automated Retinal Layer Segmentation in Optical Coherence Tomography Images. For typical retinal pictures, this suggested ANIS algorithm performs fairly accurately and averagely. According to the examination of computational complexity provided here, reformulating the segmentation as a regression issue eliminates the requirement for a sizable dataset and greatly reduces the difficulty. In the method's evaluation, which involved 114 photos, the training step for each boundary line only required 30 s, and the processing time for detecting eight borders was roughly 10.596 seconds per image.JLi, Liu, and colleagues[10] suggested "A large-scale database and a CNN model for attention-based glaucoma detection".

In this paper [6] author proposed a method investigate implicit pairwise links between OCT and VF information in both global and regional contexts, a novel deep reasoning mechanism is proposed. In order to improve the representation with complementary information for each modal, a carefully thought-out deep transformer method is built using pairwise relations. The global relation module, the guided regional relation module, and the interaction transformer module, in that order, are meant to extract and collect crucial data for glaucoma diagnosis based on reasoning and transformer mechanisms.

III. SYSTEM IMPLEMENTATION

Image processing can be classified into a number of categories, such as "image compression," "image upgrade," and "reclamation and measurement extraction." It helps to reduce the amount of memory needed to store an intricate image.The photos may be discarded due to issues with the digitising process and other factors.In this study, we combined our own unique VGG-19 architecture with the ability to analyse images. History of VGG-19: Several advances in picture classification have been made possible with the help of deep neural networks. These sophisticated models have also helped numerous other visual recognition tasks. As a result, we often enhance our understanding of problems, find solutions to them more effectively, and increase our accuracy. But when we delve further into neural networks, accuracy suffers and training becomes more difficult. These problems are resolved by using VGG-19. Instead of having a large number of hyperparameters, the VGG-19 is a CNN-based model that uses 3 3 filters with a single stride and consistently uses the same padding and max pooling layers of 2 2 filters with a stride of 2. The convolution and max pooling layers are arranged similarly in the architecture There are two FC layers in the model. There are more than 138 million trainable parameters in this VGG-19 network, which is a large network. The figure depicts the VGG-19 network architecture.

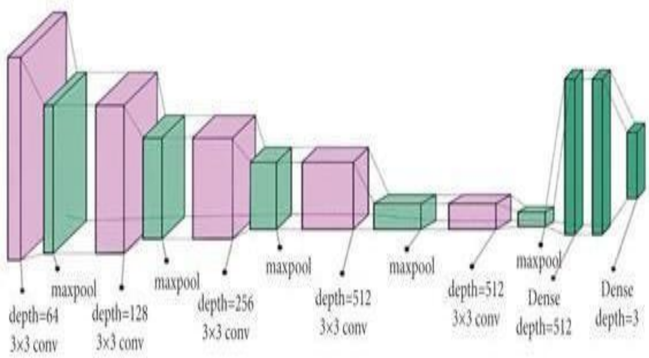


Fig-1:VGG-19 ARCHITECTURE

Six layers make up the foundation architecture, upon which we have developed our additional model. Our first layer, Conv2D, is used to mimic a 2-dimensional convolutional neural network because CNN is our foundation architecture. ReLu has been employed by us as our activation function. We have a MaxPoll2D layer placed after the Conv2D layer, which is meant to mimic max pooling in TensorFlow. We chose a stride length of two for the maximum pooling. After the fundamental structure has been put in place, we must flatten our output and then use an activation function to obtain the final probability. We evaluated our technique using more than 5,000 facial photos and discovered that our suggested eye detector was reliable and efficient. To obtain noticeably high efficiency and a good classification rate, we merged feature points with cascaded CNNs [8]. To train more potent eye recognition models in our upcoming work, we intend to gather additional facial photos. For some applications, such as eye tracking systems, and eye illness diagnosis, it's also crucial to determine the locations of the eye's regions, classify the left and right eyes, and place the eye's center.

 DOI : 10.35291/2454-9150.2023.0144

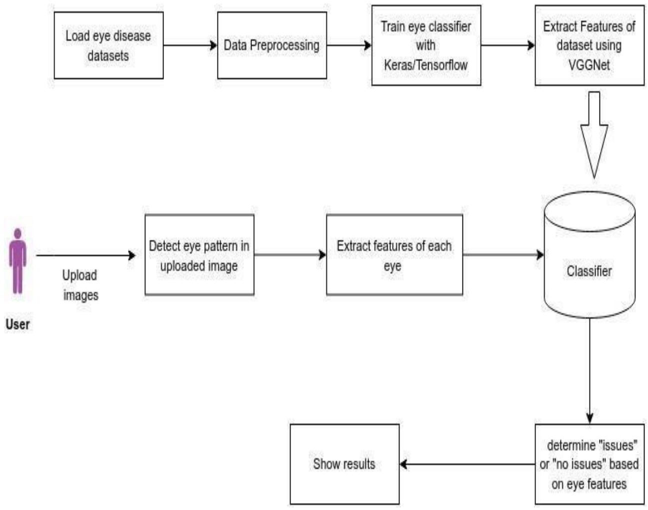


FIG-2: SYSTEM ARCHITECTURE

DATASET

Ocular Disease Intelligent Recognition (ODIR) is a structured ophthalmic database with information on 5,000 patients, including their age, the color of their fundus photos in both eyes, and their doctors' diagnostic keywords. This dataset is intended to reflect a "real-life" set of patient data that Shanggong Medical Technology Co., Ltd. has gathered from various hospitals and medical facilities in China. These institutions use a variety of cameras available on the market, including Canon, Zeiss, and Kowa, to acquire fundus images, producing images with different image resolutions.

- Normal (N)
- Diabetes (D)
- Glaucoma (G)
- Cataract (C)

MODULES

1. Data preprocessing
2. Training
3. Validation
4. UI Design

DATA PREPROCESSING

The four eye conditions that will be covered in this essay are normal, diabetes, glaucoma, and cataracts. The data set for the disorder's cataract and Normal eyes was compiled using the Kaggle website and a small fraction of the internet. Online databases have also been developed for diabetes and glaucoma, albeit with the help of a nearby optometrist. High performance of eye disease detection systems needs solving several significant challenges, including building a database and standardizing image dimensions. The method that will be used to resize images is described in the section that follows. Our data was divided into a training set, which contained the photos used to train the classifier. Images from the test set are used to verify our classifier.

FIG 3:SAMPLE PREPROCESSING DATA

DATA PREPARATION

When we originally started gathering the pictures, they were all in different sizes. Our informational index ranges in height, width, and size. In any case, our profound neural classifier needs a corresponding informative index in order to build and test the informational index. Thus, 200 × 200 pixels were chosen for the pixels.

TRAINING

We construct models in this module using a variety of layer types, including Conv2D, Maxpooling2D, Flatten, Dropout, and Dense. All of the features extracted from these layers will be applied to our categorized dataset. Following several iterations, accurately extracted eye features were saved inside a model file. All the distinctive features are included in this model file. To guarantee the accuracy of proposed system, these model files were retrained numerous times.

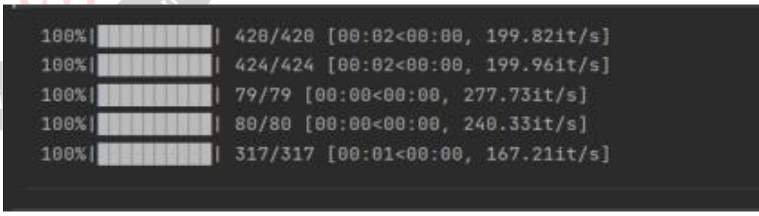


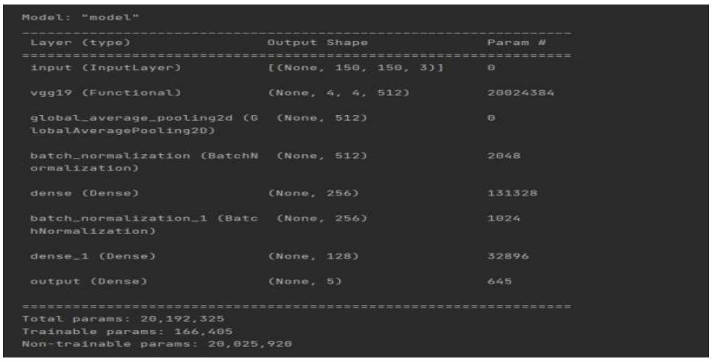
FIG 4: TRAINING SET

VALIDATION

The Adam optimizer is used to build our model. Our training dataset, which makes up 90% of the dataset, is utilized for training, and the remaining 20% is used for validation. Six thousand images make up our training dataset. Therefore, we can say that there are 840 images in the validation sets and 960 images in the training sets. 50 epochs were used to train the model, and 50 batches were used to train our classifier.

 DOI : 10.35291/2454-9150.2023.0144

ISSN : 2454-9150 Vol-09, Issue-04, July 2023



```
Model: "model"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input (InputLayer)          [(None, 150, 150, 3)]     0

 vgg19 (Functional)          (None, 4, 4, 512)         20024384

 global_average_pooling2d (G  (None, 512)              0
 lobalAveragePooling2D)

 batch_normalization (BatchN  (None, 512)              2048
 ormalization)

 dense (Dense)               (None, 256)               131328

 batch_normalization_1 (Batc  (None, 256)              1024
 hNormalization)

 dense_1 (Dense)             (None, 128)               32896

 output (Dense)              (None, 5)                 645

=================================================================
Total params: 20,192,325
Trainable params: 166,405
Non-trainable params: 20,025,920
```

FIG 5: MODELLING DATA

CONFUSION MATRIX

The confusion matrix for a certain dataset has been used to evaluate the effectiveness of classification methods. It could only be established if the true values of the test data were known. There are two dimensions to a two-dimensional matrix that separates the projected and actual values as well as the overall number of forecasts. A projected value is determined by the model, whereas an actual value is determined by observational data. In Figure 9, the N versus H class is depicted for different classification confusion matrices. It has been demonstrated that the VGG-19 model that has been put into use can correctly categorize 54 true positive (TP) images and 75 true negative (TN) images. The model predicted 12 false positives (FPs) and 4 false negatives (FNs), and it also incorrectly categorized several photos.

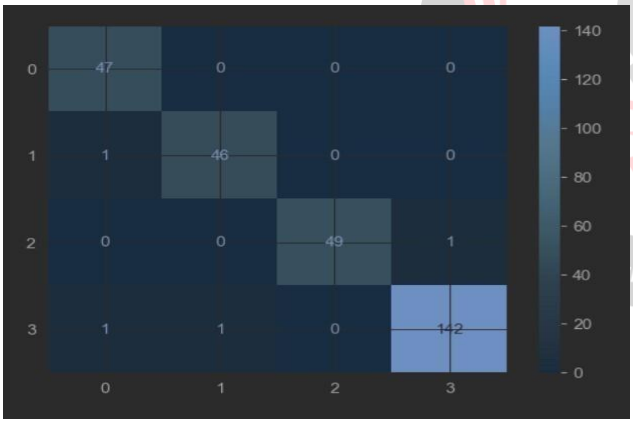


FIG 6: MATRIX TABLE

Model comparison

The primary distinction is that although typical work trains all classes simultaneously, we train each class separately. The table above demonstrates that VGG architecture is primarily employed in eye illness recognition tasks where they achieve accuracy lower than 90%. This work is considered to be more accurate than the others, with a VGG-19 accuracy of 97.94%. We employed the same VGG-19 in our research, and it produced accuracy results that were 98.10%, somewhat higher than in the earlier study. Researchers also used the EfficientNet and DenseNet topologies, however, their accuracy rates were only about 80%–87%. Among all of the aforementioned studies, our conclusion demonstrates a more accurate and precise classification of ocular illnesses.


```
1 {
2     "percentage": "0.9684832692146301",
3     "prediction": "\"0\""
4 }
```


FIG 7: COMPARISON AND PREDICTION

UI DESIGN

This module uses the Flask framework, HTML, CSS, and Bootstrap to create a user interface. This interface will collect user data and eye pictures. These photos are sent to the Python model as a request, and it then gathers the return data and produces a report.HTML CSS Bootstrap is used to generate the patient's authentication sites, and front-end languages and frameworks like React and Node.js are used to create the related pages.

IV. PERFORMANCE EVALUATION

The majority of the currently used techniques are challenging to put into practice and depend on the face detector's accuracy. In an experimental setting, the locations of the eyes' regions and centers were determined in about 9 milliseconds, with about 2 milliseconds going towards proposing candidate feature points for the eyes and another 2 milliseconds going towards determining the precise locations of the eyes' left or right regions (1st set of CNNs), and 30 milliseconds going towards determining the locations of the eyes' centers (2nd CNNs). The majority of eye detection jobs operate at a frame rate of 30 to 60 fps. This demonstrates that our suggested strategy is capable of addressing issues that emerge in genuine circumstances in real-time. Fig 8 shows the accuracy of the proposed model and Fig 9 shows the loss in the model.

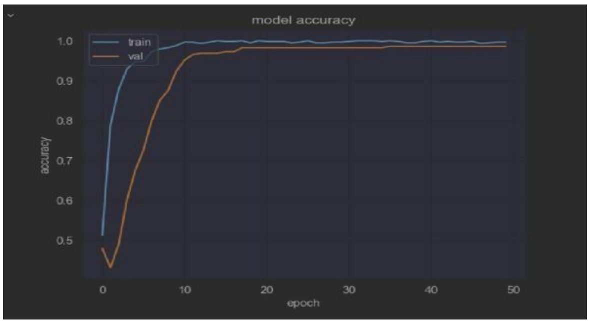


FIG 8: Accuracy of the Proposed Model

DOI : 10.35291/2454-9150.2023.0144

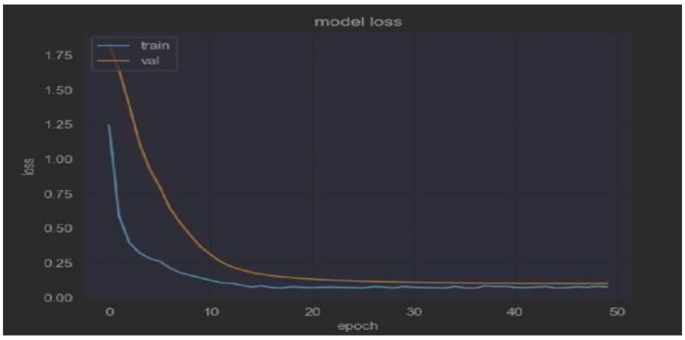


FIG 9: Model Loss

V. RESULTS AND DISCUSSIONS

Based on our investigation, we determined that taking the same number of photographs would fix the class imbalance issue. The ODIR dataset showed a significant discrepancy in the classes. When there are fewer photos, we can increase our accuracy by using this approach greatly. The performance evaluation approaches' relative metrics, accuracy loss graphs, and other comparable indicators were then looked at and graphically displayed. We evaluated the capability of our model to correctly anticipate a certain condition using the VGG-19 architecture. We demonstrated the validity of our model's prediction using the confusion matrix on the test. Fig 10 shows the login page and Fig 11 shows the Blog page.

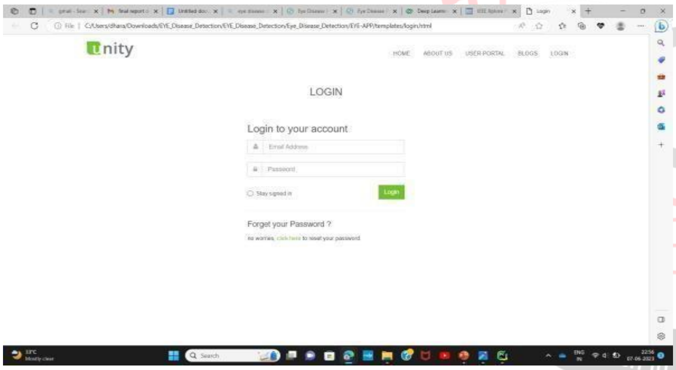


FIG 10: Login Page

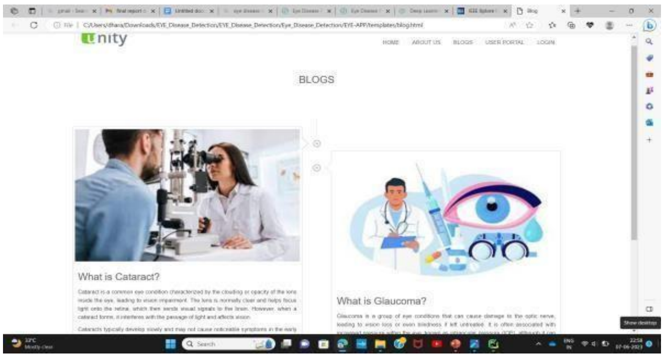
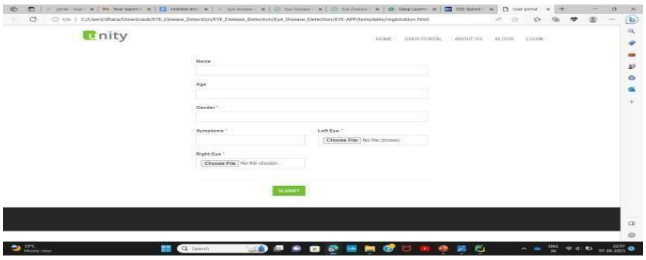


FIG 11: Blog Page

Fig 11 shows the Index page and Fig 12 shows the Report generated by the model.



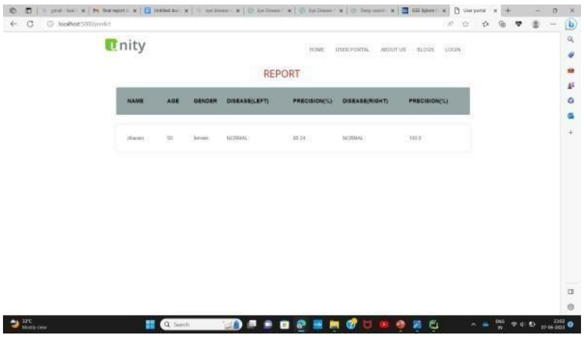


FIG 11: Index Page

VI. CONCLUSION

The model VGG-19 was used in this work to categorize the major ocular disorders and predict whether an eye has any diseases or a healthy fundus. Additionally, the performance much exceeds expectations. The accuracy for the normal vs myopia class in this experiment was 98.10%, and the accuracy for the normal versus cataract class was 94.03%. Additionally, we obtained a 90.94% accuracy rate when comparing the normal and glaucoma classes. The suggested method outperforms current CNN-based models for categorizing eye diseases while needing less latency. It is also simple to modify for other medical image-based classifications of illnesses.

The VGG-19 model can also be used to develop a method for categorizing real ocular illnesses that affect consumers. The method's ability to be quickly adapted to different kinds of disease classification based on medical images is what makes it so interesting. Ocular image segmentation could also be used in this study. In recent years, numerous models for detecting eye diseases have been developed. We employed deep neural networks in this paper, along with some essential libraries including OpenCV, Keras, Tensor Flow, Pandas, and NumPy.

REFERENCES

[1] Hu, Xiaoyan, et al. "GLIM-Net: chronic glaucoma forecast transformer for irregularly sampled sequential fundus images." IEEE Transactions on Medical Imaging (2023).

[2] Shen, Yuhe, et al. "Graph Attention U-Net for Retinal Layer Surface Detection and Choroid Neovascularization Segmentation in OCT Images." IEEE Transactions onMedical Imaging (2023).

 DOI : 10.35291/2454-9150.2023.0144

[3] Alamdar, Alireza, et al. "Force and Velocity Based Puncture Detection in Robot Assisted RetinalVein Cannulation: In-Vivo Study." IEEE Transactions on Biomedical Engineering 69.3 (2021):1123-1132.

[4] Ju, Lie, et al. "Leveraging regular fundus images for training UWF fundus diagnosis models via adversarial learning and pseudo- labeling." IEEE Transactions on Medical Imaging 40.10 (2021):2911- 2925.

[5] Ngo, Lua, Jaepyeong Cha, and Jae- Ho Han. "Deep neural network regression for automated retinal layer segmentation in optical coherence tomography images." IEEE transactions on image processing 29 (2019): 303-312.

[6] Song, Diping, et al. "Deep relation transformer for diagnosing glaucoma with optical coherence tomography and visual field function." IEEE Transactions on Medical Imaging 40.9 (2021):2392-2402.

[7] Wang, Jianhui, et al. "Eye blink artifact detection with novel optimized multi-dimensional electroencephalogram features." IEEE Transactions on NeuralSystems and Rehabilitation Engineering 29 (2021): 1494-1503.

[8] Jiang, Yuming, et al. "JointRCNN: a region-based convolutional neural network for optic disc and cup segmentation." IEEE Transactions on Biomedical Engineering 67.2 (2019): 335-343.

[9] Fu, Huazhu, et al. "Angle-closure detection in anterior segment OCT based on multilevel deep network." IEEE transactions on cybernetics 50.7 (2019): 3358-3366.

[10] JLi, Liu, et al. "A large-scale database and a CNN model for attention-based glaucoma detection." IEEE transactions on medical imaging 39.2 (2019): 413-424.

[11] Aujih, Ahmad Bukhari, et al. "EDR- Net: Lightweight Deep Neural Network Architecture for Detecting Referable Diabetic Retinopathy." *IEEE Transactions on Biomedical Circuits and Systems* 16.3 (2022): 467-478.

[12] Pan, Lingjiao, et al. "OCTRexpert: a feature-based 3D registration method for retinal OCT images." *IEEE Transactions on Image Processing* 29 (2020): 3885-3897.

[13] G.Mahalakshmi, ShimaaliRiyasudeen, Sairam R, Hari Sanjeevi R and B. Raghupathy, “A Survey: Effective Machine Learning Based Classification Algorithm for Medical Dataset”, International Research Journal on Advanced Science Hub (IRJASH), Volume 03 Issue 09S September 2021.

 DOI : 10.35291/2454-9150.2023.0144