Technology and Disease Prediction: Addressing the Challenges in Retinal Disorder Prediction Using Transfer Learning Mechanism

Harshini Lara A K[1], Dr. K. V. Uma[2]

[1]Department of Information Technology, Thiagarajar College of Engineering, Madurai, Tamilnadu, India, harshukuberdevi@gmail.com
[2]Associate Professor, Department of Information Technology, Thiagarajar College of Engineering, Madurai, Tamilnadu, India, kvuit@tce.edu

Analyzing the retina, a crucial part of the eye, is essential for the early diagnosis of major fundus disorders. A "fundus imaging" facility documents anatomical traits and anomalies to help ophthalmologists differentiate between different fundus diseases. Transfer Learning which includes several pretrained Deep CNN architectures speed up the prediction of disorder. This study examines three different architectures – EfficientNetB0, ResNet150, DenseNet121 in-depth in order to determine which is the best fit for the dataset on fundus disorders. The prediction precision is increased by careful fine-tuning that takes into account activation functions, optimizers, learning rates, and accuracy. The paper's main topic is the categorization of retinal disorders using binary and multi-class models and the address of imbalanced dataset problem dominant in multi-class classification than in binary classification. The study highlights the importance of crucial variables and metrics in minimizing the negative effects of unbalanced datasets and demonstrates the potential of deep learning for the categorization of retinal disorders. Ophthalmic medical imaging will advance through the refinement of cutting-edge architectures for more precise diagnoses.

Keywords: retina, fundus disorder, efficientnetb0, resnet150, densenet121, softmax, rmsprop, drop connect rate.

1. Introduction

With the growth of ophthalmology research, retinal images now play a crucial part in locating ocular disorders. Modern medical imaging research has focused on the prediction of retinal disorders, where the combination of ophthalmology and machine learning provides never-before-seen insights into eye health. Within the broad field of academic research that tries to decipher the intricacies of retinal illnesses, investigations frequently traverse a variety of datasets and approaches. Research on retinal illnesses has traditionally looked at a wide range of imaging modalities, with high-resolution retinal images and Optical Coherence Tomography images emerging as the main sources of diagnostic data. Rich visual representations of the retina are offered by these modalities, which provide in-depth investigations for the identification of anomalies ranging from mild abnormalities to obvious diseases. In current scenario, binary and multi-class classification circumstances have received the majority of attention, reflecting the duality that is frequently present in medical decision making. While multi-class classification expands this paradigm to suit the wide range of retinal disorders, binary classification distinguishes between the presence or absence of a specific disorder. These investigations greatly advance the fundamental knowledge of the onset and course of illness.

A retinal image constitutes a delicate tissue lining the back of the eye's retina which gets captured digitally in a retinal image. The need for a tool that can forecast the retinal diseases has become urgent in a world that is frequently characterized by unjustified anxiety over probable eye disorders or disregard of ocular health leading to false predictions. With that in mind the classification of retinal disorders for the following eight categories namely normal, diabetic-retinopathy, glaucoma, age-related macular degeneration (AMD), cataract, hypertension, myopia and others has been taken into account. But that can’t be done in a leisure manner. When it comes to ophthalmology, both proper insistence and assistance by the classifier must be made to avoid discrepancies.

Ocular Disease Intelligent Recognition - ODIR is a dataset that consists of all aforementioned categories of retinal images in JPG format and offers them readily preprocessed as shown in Fig. 1. The quality of these fundus images can also be assessed as done in [8].

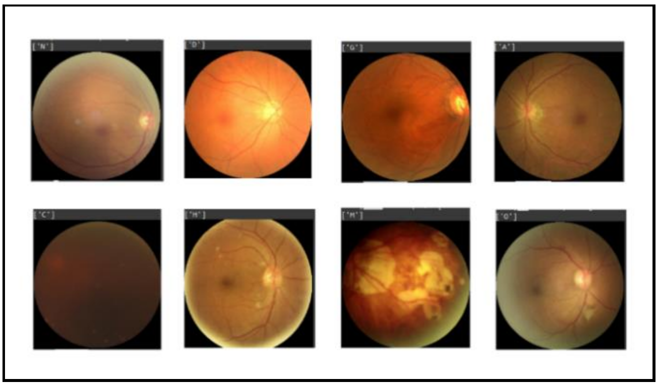

Fig. 1. Different categories of Retinal disorder images available in ODIR dataset

It contains around 6392 retinal images correlating to a specific disorder among the eight. This dataset ODIR is considered to be small when it comes to the training of Deep Learning and Transfer Learning models whose methodologies have presented novel opportunities to transform the predictive power of models. The classification that has been done here involves the contemporary Binary classification and Multi-class classification. To facilitate this task, Transfer Learning comes into play which utilizes a model for a general task, such as image classification, that has been trained on a large dataset. The goal is to apply the knowledge that was acquired in the pre-training phase to the prediction of retinal disorders that uses a smaller dataset. Instead of starting with weights that are randomly initialized, the model starts with weights that include relevant information about common and frequent characteristics found in image data. Further the retinal disease prediction task can be fine-tuned or further trained on the ODIR dataset. Also, this problem can be resolved using pure Deep Learning models with minimal data as discussed in [7] but here we go with Transfer Learning.

In this study, the complex field of retinal disease prediction is explored, emphasizing the difficulties caused by small sample sizes and class disparities using Transfer Learning. In order to make reliable predictions, there is a necessity to customize solutions for the minute characteristics seen in retinal images, as this research carefully examines the peculiarities of dealing with such a dataset. In an effort to foresee and handle the discussed issue, the study considers alternative solutions trying to efficiently lay down the efforts to anticipate and manage it.

2. RELATED WORK

A number of articles have taken attempts to classify Retinal Disorders with different datasets available and different strategies to enhance the performance of the Transfer Learning models in [1], [2], [3], [4], [5], [6], [8] as practiced in this paper.

Digital ocular fundus images have become widely used in ophthalmology practice in recent years. Nevertheless, ophthalmologists have continued to interpret these images, which comes at a high expense. In order to tackle this difficulty, researchers have investigated strong deep learning frameworks for identifying the main eye conditions, including as age-related macular degeneration (AMD), glaucoma (GLC), and diabetic retinopathy (DR). The two main components of the suggested approach are the use of transfer learning techniques employing several convolutional neural network (CNN) models and an initial quality evaluation to weed out low-quality photographs in [1]. Comparing the performance of several models is a noteworthy improvement, as it shows that DenseNet-201 performs better than other CNN models, obtaining a remarkable area under the receiver operating characteristic curve of 0.99. The suggested approach shows promise for autonomously diagnosing various eye disorders because of its high specificity for various situations, such as 100% for AMD and 96.69% for DR.

Building on these developments, researches found the significance of better models for the identification and categorization of retinal diseases. Efforts have been focused on improving the accuracy and dependability of the VGG model in [2]. In comparison to ResNet, AlexNet, and other VGG models, the suggested model has an enhanced VGG training network with densely linked layers and shows better accuracy and dependability.

Furthermore, as evidenced by the Structured Analysis of the Retina (STARE) database used in [3], deep learning convolutional neural networks (CNNs) have been used to the field of retinal illness detection for the automated identification of different retinal disorders. But there are obstacles in the way of reaching the best results as experienced in ODIR database as well, especially when it comes to the classification results, reliance on the number of categories and their decreasing performance with increasing number of categories is inevitable in [3].

In [4] the exploration of random forest transfer learning based on the VGG-19 architecture and ensemble classifiers, have demonstrated promise in improving multi-categorical classification, is one approach being taken to solve these issues. In ten retinal illnesses, the transfer learning strategy with voting and clustering performed the best, offering a possible way to improve the categorization of various retinal complications. Moreover, Furthermore, as color fundus photographs are now routinely used in clinical practice to diagnose ocular abnormalities practiced in [5] as well, research into their potential to predict systemic features has been prompted.

Hence, The results support further attempts to improve the accuracy and utility of these models in therapeutic situations by demonstrating the relevance of fundus images in revealing essential data about a patient's systemic features.

3. METHODOLOGY USED

A. Performing Binary Classification

1) Binary Classifier

An input data point is classified into one of two potential classes or categories. It is a particular kind of model that works well for jobs where the result may be classified as either positive or negative, yes or no, or 1 or 0.

In this case, the output is either “Normal” or “Diseased”. Binary classification is here carried out for two main classes among the eight. For instance “cataract” and “myopia” are considered.

The significance of selecting both of these classes is the good size of the image data available under both of these categories in this dataset.

In this paper, the binary classification between “normal” and “cataract” is highlighted. The data label rendered as an output will be either 1 for cataract or 0 for normal case as shown in Fig. 2.

2) Process flow

In order to collect "cataract" labeled images, the first step in data assimilation entails extracting filenames from the attributes "Left-Diagnostic Keyword" and "Right-Diagnostic Keyword." This results in both left and right image arrays which are then concatenated into a single array. The "normal" keyword is applied to normal images in a similar manner. As a result, both "normal" and "cataract" data are represented by two distinct arrays.

An equal number of images are chosen from the "normal" and "cataract" categories to create a balanced dataset. The "normal" image array's total original size, which includes the left and right image data, is 5675. However, only 594 photographs from the "normal" category are picked in order to maintain balance and match the number of "cataract" images that are readily available. This method balances the distribution of the dataset. This subset is created for training the model and consists of an array of image data and the accompanying class labels (1 for "cataract" and 0 for "normal"). Fair representation is ensured throughout model training by using this balanced subset.

The "ResNet50" Deep CNN model is applied in this case. 224 pixels is designated as the image size. On the ImageNet dataset, this ResNet50 model underwent pre-training for image classification, taking advantage of its discovered characteristics. This architecture's 50 layers are all still in their pre-trained weights and biases left frozen, with no changes made to them.

This ResNet50 model serves as a feature extractor on top of the Sequential model's fundamental structure. To restructure the output, a flattening layer is placed after the ResNet50. The final step is the addition of a dense layer with a single unit and sigmoid activation function. A batch size of 32 is used to train the model over a period of 50 epochs. A callback that tracks validation loss and halts training if no improvement is noticed for 7 consecutive epochs is used to implement early stopping. The training set constitutes the created dataset and the testing set constitutes the binary values depicting the results relative to training inputs with a test size of 25%.

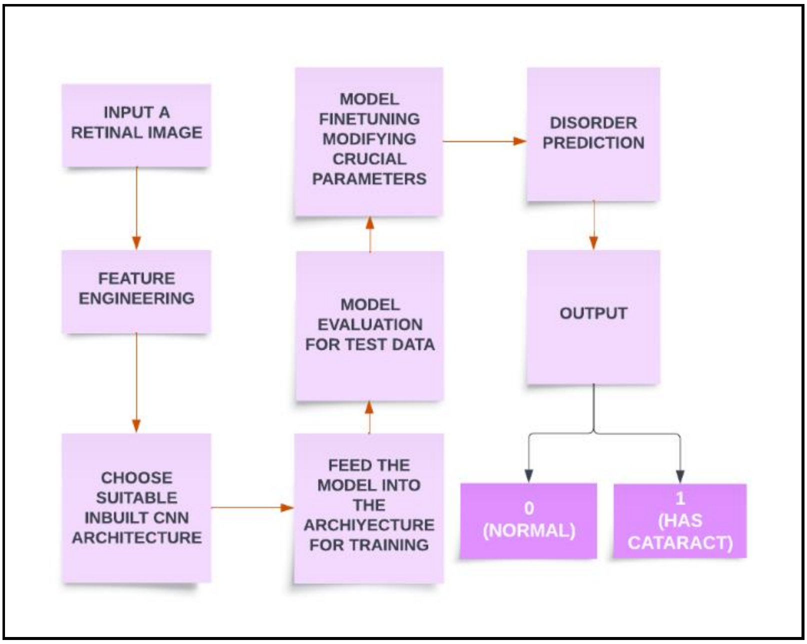

Fig 2. Binary Classifier module

A. Performing Multi-class Classification

1) Multi-class Classifier

It functions as a central pivot, precisely assigning each individual data point to a specific category from a diverse array of classes.

Here, there exist 8 different classes in total comprising of normal, diabetic-retinopathy, glaucoma, age-related macular degeneration(AMD), cataract, hypertension, myopia and other.

At a stretch all the 8 classes are considered for performance and accuracy analysis..

2) Process flow

As a primer step, images are sorted into discrete lists according to the dataset's "label" attribute. This division makes up the sub dataset meant for several target categories integrated, resembling the approach used in binary classification.

The dimensions of all the images in both the training and test sets are uniformly fixed to 224 pixels. Class names are transformed into numerical representations via one-hot encoding as they are categorical in the test set. This translation makes it easier to interpret and analyze numerical data.

The sequential model establishes a base for the architecture. A layer of EfficientNetB0 is then added on top of this foundation, enclosing the top layer and using a drop connect rate of 0.4 while inheriting weights from ImageNet. Sequentiallly, a layer that flattens out is followed by a dense layer of 1000 neurons that uses Rectified Linear Unit (ReLU) activation. The concluding layer involves using an output dense layer utilizing the softmax function for categorizing the retinal images into 8 different categories.

The model is built up by compiling it with specific settings which plays a vital role here. It uses the RMSprop optimizer with a learning rate of 0.0001. Categorical cross-entropy is chosen as the loss function for multi-class classification. The early stopping callback is implemented to improve training, much like it is in binary classification. To foster the learning process, the model is trained using a batch size of 32 for 50 epochs. And 25% of the data taken from the output dataset is encompassed in the test set.

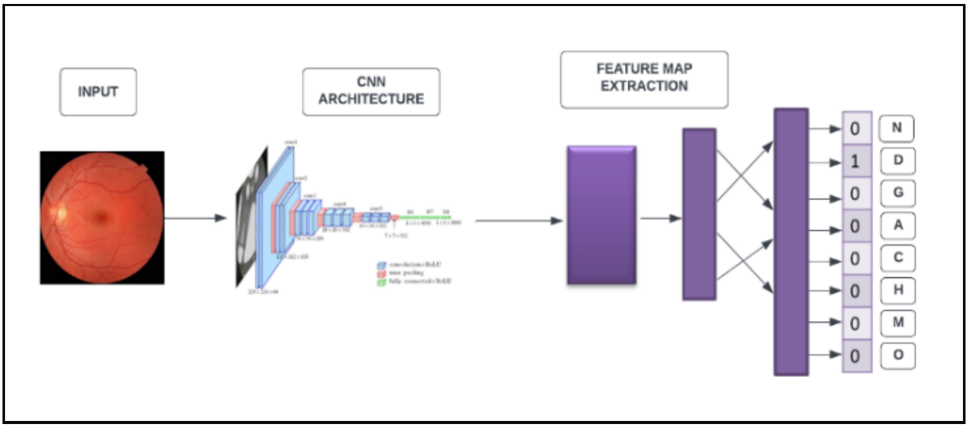

Fig 3. Multi-class classifier module

B. Output

The output is generated in the form of a web application designed for binary classification of retinal disorders. This application focuses on determining if an input retinal image exhibits cataract or not whose System Architecture is portrayed in Fig. 4.

The application uses its predictive ability to identify the presence or absence of cataract after being fed with the photograph of retina either of left or right eye as an input. Results are then given, together with comprehensive details on any accompanying symptoms and the corresponding preventive steps.

In order to facilitate this process, two accessory elements are integrated into the application to improve user accessibility and comprehension. First, a translation module renders the symptom descriptions and preventative recommendations into Tamil, guaranteeing that native Tamilians with limited English proficiency can also understand them. Secondly, a text-to-speech conversion mechanism speaks the Tamil information aloud, making it understandable through both spoken and written language. This all-encompassing strategy seeks to offer educational and approachable aid for the diagnosis and prevention of retinal disorders.

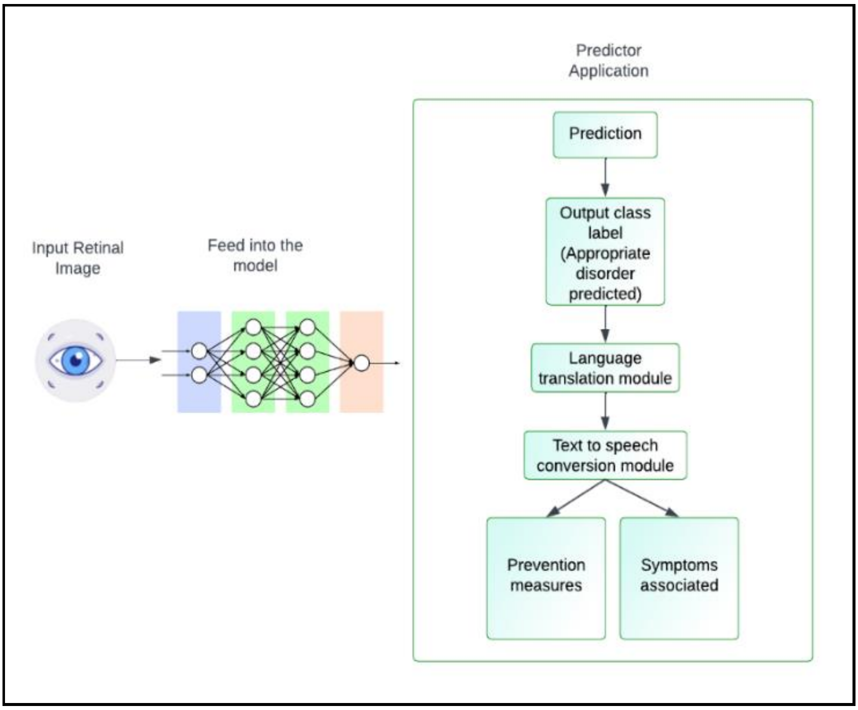

Fig. 4. System Architecture

4. EXPERIMENT RESULTS

Comparing both the classifications done for the dataset, the severe drift in the validation accuracy in multi-class classification is inevitable, moreover it is obvious that the overfitting issue – low bias and high variance is encountered in case of multi-class classification. The following figures Fig. 5 and Fig. 6 representing Confusion Matrices for Binary Classification and Multi-class Classification respectively are displayed.

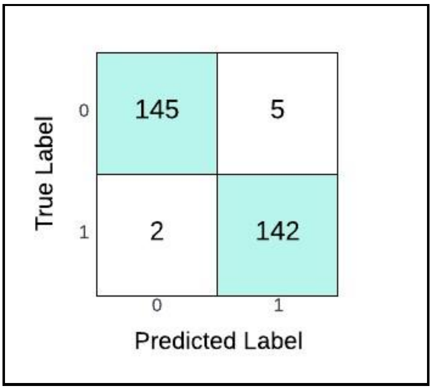

Fig. 5. Confusion Matrix for Binary Classification

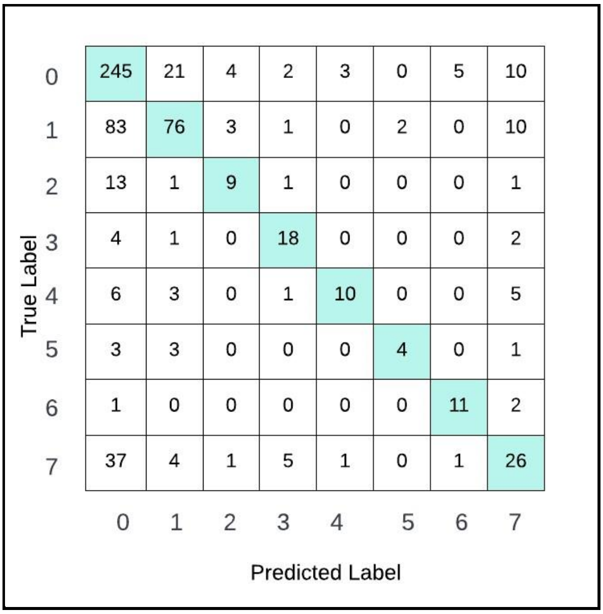

Fig. 6. Confusion Matrix for Multi-Class Classification

It is apparent from the Confusion Matrix of multi-class classification as shown in Fig. 6 that, most of the retinal images are misclassified.

There are several reasons for the same. One among those reasons which is dominant is the class imbalance as well as the smaller size of the dataset, which becomes the major issue here. Also the images under each category are uneven - normal(2873), diabetic-retinopathy(1608), glaucoma(284), age-related macular degeneration(266), cataract(293) hypertension(128), myopia(232), others(708).

Moreover in binary classification the images are extracted from two attributes namely left diagnostic keyword and right diagnostic keyword which accounts in preferably good number of images and so overfitting issue is not discovered in binary classification but when it comes it multi-class classification where they are extracted using the label of the images, only low number of images are obtained. This is evident in case of cataract whose size stands as 594 and 293 in binary and multi class respectively. As well as the overall size of the dataset is also not appreciable which encounters only around 6000.

There are a variety of potential solutions to this issue, with the application of model hyperparameter tuning, data augmentation, GAN (Generative Adversarial Network) model use, regularization techniques, ensemble model implementation, , and manual dataset expansion emerging as the most crucial options.

A. Embarking on Hyperparameter Tuning

In order to find the model with the highest performance, intensive hyperparameter adjustment is required, which shall allow the model for experimentation with many suitable pre-trained architectures. The model's architecture, learning rate, number of epochs, batch size, activation function, dropout rate, optimizer, and loss function are just a few examples of the parameters covered by the hyperparameter grid used.

TABLE I. PARAMETERS VS. RESULT

S.No	Architecture	Optimizer	LR	Drop Connect Rate	Batch Size	Epochs Taken	VA	VL	TA	TL
1	EfficientNetB0	RMSprop	0.001	0.4	32	12	0.6380	2.5514	0.8285	0.4892
2	EfficientNetB0	RMSprop	0.0001	0.4	32	8	0.5966	4.4157	0.9875	0.044
3	EfficientNetB0	Adam	0.001	0.4	32	17	0.6034	1.52	0.82	0.51
4	EfficientNetB0	Adam	0.0001	-	32	16	0.6271	1.2209	0.8267	0.5332
5	ResNet50	Adam	0.0001	-	32	16	0.5160	1.4972	0.8193	0.5550
6	DenseNet121	Adam	0.001	-	32	16	0.5481	1.3564	0.7880	0.6258
7	DenseNet121	Adam	0.001	-	32	22	0.5481	1.5034	0.7682	0.6718
8	EfficientNetB0(Data augmented)	Adam	0.0001	0.4	32	28	0.6388	1.1252	0.7745	0.6416
9	ResNet50(Data augmented)	Adam	0.0001	-	32	9	0.6028	1.2116	0.7209	0.7244
10	EfficientNetB0(Data augmented)	Adam	0.0001	0.4	32	10	0.6226	1.1422	0.7385	0.6871

TABLE I displays the varied outcomes against various parameters.

"EfficientNetB0" here stands out among the other Transfer Learning models especially in the context of multi-class classification for this dataset. The model performs well and produces validation accuracy of around 64%. During training, it encounters all of the following parameters when trained with EfficientNetB0 architecture: Optimizer as RMSprop, learning rate as 0.0001, with a drop connect rate of 0.4, batch size of 32 and with

28 epochs.

Also the model is updated involving experimentation with the other techniques like data augmentation setting up with suitable values for suitable parameters like Rotation, Width, Height, Zoom Range, Horizontal Flip and Brightness and Contrast Adjustments. Added to that, drop connect rate is introduced as a part of dropout regularization technique which is depicted in the above table as well. Similarly weight loss function is also added separately to handle class imbalance problem. It involves computing class weight and it assigns higher weights to the underrepresented classes so as to provide more importance to the minority classes. This may ensure that the model prioritizes minimizing error in the underrepresented classes and hence resulting in a more balanced performance. But all of these techniques resulted in not much significant difference in the validation accuracy due to the nature.

On the other hand, binary classification for the class label "cataract" for this dataset is achieved at around 98% with RMSprop optimizer and a batch size of 32 for 12 epochs and it goes well with ReNet50 architecture which also resulted as the best model in [6].

5. CONCLUSION

We can conclude that this dataset ODIR appears to be more suitable for performing binary classification rather than multi-class classification as multi-class classification underperforms when all the available eight classes in the dataset are as a whole used for classification. In summary, restricted dataset size and imbalanced classes remain to be major obstacles in retinal disease prediction, despite persistent attempts to improve class imbalances, fine-tune model parameters, and employ numerous additional strategies. Additionally, traditional methods lose some of their effectiveness Whenever dealing with datasets containing images with intricate details, conventional solutions may become less effective . This issue may be resolved via enhancing the model, prioritizing ensemble learning strategies to boost model robustness, expanding the size of dataset preferably through acquisition, and exploring the potential of Generative Adversarial Networks (GANs) for the creation of synthetic data.

References

1. Pin, K., Chang, J. H., & Nam, Y. (2022). Comparative Study of Transfer Learning Models for Retinal Disease Diagnosis from Fundus Images. Computers, Materials & Continua, 10.32604/cmc.2022.021943. [DOI:10.32604/cmc.2022.021943.
2. Tyasi, T. L. et al. (2021). Improved Model of Eye Disease Recognition Based on VGG Model. Intelligent Automation and Soft Computing,10.32604/iasc.2021.016569. [DOI:10.32604/iasc.2021.016569].
3. Choi, J. Y. et al. (2017). Multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database. PLOS ONE, https://doi.org/10.1371/journal.pone.0187336. [DOI:10.1371/journal.pone.0187336].
4. Das, A. et al. (2019). Classification of Retinal Diseases Using Transfer Learning Approach. In Proceedings of the International Conference on Computer and Emerging Sciences (ICCES), July 2019, https://doi.org/10.1109/ICCES45898.2019.9002415.

[DOI:10.1109/ICCES45898.2019.9002415].

5. Sheeta, S. S. M. et al. (2021). Retinal Disease Identification Using Upgraded CLAHE Filter and Transfer Convolution Neural Network. Journal of Information and Communication Technology, https://doi.org/10.1016/j.icte.2021.05.002. [DOI:10.1016/j.icte.2021.05.002].
6. Elsharif, A. A. E. F., & Abu Naser, S. S. (2022). Retina Diseases Diagnosis Using Deep Learning. International Journal of Academic Engineering Research (IJAER), Vol. 6, Issue 2, February 2022, Pages: 11-37. ISSN: 2643-9085. (www.ijeais.org/ijaer).
7. Karthikeyan, S. et al. (2019). Detection of Multi-Class Retinal Diseases Using Artificial Intelligence: An Expeditious Learning Using Deep CNN with Minimal Data. Biomedical & Pharmacology Journal, September 2019, Vol. 12(3), p. 1577-1586. Published by Oriental Scientific Publishing Company. DOI: 10.13005/bpj/1788.
8. Abdel-Hamid, L. (2021). Retinal Image Quality Assessment Using Transfer Learning: Spatial Images vs. Wavelet Detail Subbands. Ain Shams Engineering Journal, https://doi.org/10.1016/j.asej.2021.02.010. [DOI:10.1016/j.asej.2021.02.010].