Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 11, No. 1, November 2022

EYE DISEASE CLASSIFICATION USING DEEP LEARNING TECHNIQUES

D. Guna venkat, K. Lakshman, M. Dwarakesh, K. Ajay Kumar, K. Sandeep Komal, Students
Department of ECE, GMR Institute of Technology
M. V. Nageswara Rao, Corresponding Author Department of ECE, GMR Institute of Technology
nageswararao.mv@gmrit.edu.in

Abstract: *Cataract, Diabetic Retinopathy and Glaucoma are the most commonly observed ocular diseases in the world. In this work, we propose a method to classify the above three diseases for a given fundus image. The proposed method fuses the advantage of Convolutional Neural Networks (CNN) and Transfer Learning (TL) techniques to improve the accuracy. A total of 4000 images are collected from publicly available datasets viz., Ocular Recognition, HRF, DRIVE and IDRiD and validated on 400 images of dataset containing 100 images for each class and achieved an accuracy of 92%.*

Key words: *Transfer Learning, Ocular Recognition, Convolutional Neural Network.*

I. INTRODUCTION

As per the statistics of WHO 2021, the estimated number of people effected with eye diseases is 1.02 billion. Mostly identified eye diseases are Cataract, Glaucoma, Diabetic Retinopathy. Studies have shown that blindness and impaired vision loss is related to age and is categorized into three main categories. They are Cataract, Glaucoma and Diabetic Retinopathy which have 65.2M, 6.9M and 3M cases respectively, which together make up 2.6% of all cases of blindness in the world's population [1]. A cataract occurs when a cloudy layer forms on the clear lens of the eye. Cataracts gradually worsen the quality of vision. As cataract grow, the lens will become opaque, and eventually cataracts will make it difficult to see clearly.

According to current stats by 2025 the number of patients with cataract will reach 40 million. The number of blind cases in the world and the percentage because of cataract is growing because of population growth and growing longevity. As per WHO 285 million people approximately are having visual impairment. Among the patients having visual impairment, almost 39 million have low vision, and 246 million have damaged vision. The result of cataract eye disease will be 70.2% of the people have moderate vision impairment and 66.2% of the people will be blind.

Diabetes can cause a number of different eye conditions together known as diabetic retinopathy. If not identified and treated, this illness can result in significant vision loss or even blindness. Damage to the retina's blood vessels results in diabetic retinopathy. Blood vessels in the retina of some patients with this condition may enlarge and leak fluid or blood. A person may initially detect specks floating in their vision or general eyesight blurriness. Globally, 103.12 million people are predicted to have diabetic retinopathy in 2020, and 160.50 million people may have it by 2045.

Glaucoma can lead to the blindness of a person, which makes this disease one of the most dangerous diseases in the world. In these cases, the optic nerve of the eye will be affected, which results in the blindness of a person. Symptoms observed in Glaucoma are severe headache, eye pain, nausea, blurred vision, halos around lights. After cataracts, glaucoma is the second most common cause of blindness worldwide, according to the World Health Organization. Glaucoma is expected to affect 60 million people worldwide, rising to 79.6 million by 2020 and 111.8 million by 2040. According to research, the frequency of glaucoma ranges from 0.94 percent to 74 percent in numerous Asian countries, with occlusive glaucoma being the most prevalent kind.

Early diagnosis of the above diseases is crucial for effective treatment and can lower the chance of developing permanent blindness. Automatic eye identification and classification typically employ four types of systems/pictures: slit lamp, fundus images, ultrasonography, and retro-illumination. Among these imaging modalities, fundus pictures have piqued the interest of technicians and patients alike since the fundus digital camera is simple to use. [2, 3].

A fundus image is the deep back portion of the eye. It consists of the retina, blood vessels, fovea, macula, etc. There will be some special types of cameras used to capture fundus images, which are called fundus cameras. This camera captures images by pointing the camera towards the pupil. A typical fundus image comprises of Optic Disk, blood vessels, macula, exudate, hemorrhage as shown in the fig.1.

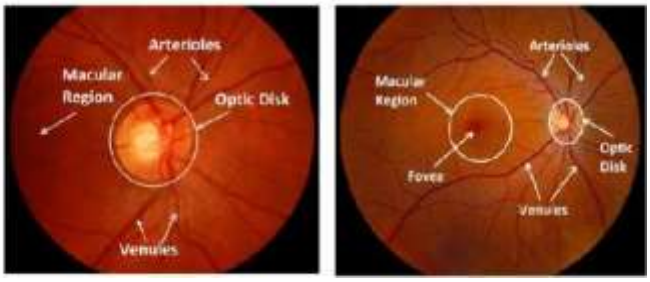

Fig. 1. Retinal fundus image. The image is taken from [25]

Fundus images for eye diagnostics provide a number of obstacles. To begin, many previous investigations of eye detecting, and grading have relied entirely on subjective feature extraction [4,5,6,7,8]. That is difficult and also time-consuming challenge. Second, the retinal datasets are extremely tiny and uneven. Third, the fundus image diagnosis has to be validated by specialist ophthalmologists with extensive experience. To alleviate the difficulties, an automated approach for fundus image analysis is necessary.

Deep learning (DL) has recently emerged as a new subject of computer vision and image processing [9]. Deep learning allows us to resolve issues that have been formerly insolvable by machine learning. Deep learning makes use of neural networks to raise the computational tasks and offer correct results.

CNNs are deep learning algorithms that are useful for recognition and classification of images and objects. In contrast to typical ML algorithms, CNN does not involve subjective feature extraction. The availability of labelled data, the complexity of retinal properties and vessels, which may affect other vessels and depend on image quality, and the subjectivity of manually classified fundus images all pose challenges to the use of CNN. Transfer learning is used as a result to overcome those challenges. In recent years, transfer learning has advanced significantly and attracted a lot of interest. By simply changing the hyperparameters, transfer learning (TL) techniques can be applied to any type of assignment after being trained on a larger dataset (like ImageNet) [10].

II. MATERIALS & METHODS

Richard and Yunendah [12] have proposed a CNN model for cataract detection using fundus photography. They proposed a CNN model that can detect cataract only. They used the dataset collected from Eye hospitals of North Sumatra. The dataset consists of 73 normal images and 326 cataract images. Initially, they used transfer learning techniques for making predictions on cataract. They used GoogleNet, MobileNet and ResNet pre-trained models. They found that the training time is too long and performance is not up to the mark. Finally, they developed a simple CNN model having 5 hidden layers and output layer has SoftMax activation. There proposed model achieved an accuracy of 92%.

Zubair Khan and Fiaz Gul Khan [13] has proposed a non-linear scale invariant deep model called VGG-NiN for determining severity of Diabetic Retinopathy (DB). In this work, the authors stacked the VGG16 pre-trained model with a Spatial Pyramid Pooling (SPP) and network-in-network (NiN) layers. They have used dataset from the Kaggle website which is organized by EyePacs. The dataset has 35126 labelled images. The authors mentioned that to avoid overfitting the learning rate is kept

adaptive. They used batch size of 8, initial learning rate set to 0.01 and trained the model for 30 epochs. There proposed model achieved an AUC of 83.8 on average.
In [14] the authors proposed 4 machine learning models for predicting glaucoma from fundus image. They used Corvis ST dataset that consists of 103 samples of Glaucoma. Since feature extraction must be done before feeding data to a machine learning model, the authors employed variance to choose the dataset's most pertinent characteristics. The authors used Gradient Boosting Classifier (GBC), Random Forest (RF), Stochastic Gradient Descent Classifier (SDGC) and K-Nearest Neighbour (KNN) machine learning models whose performance in terms of accuracy are 76%, 73%, 75% and 71% respectively.

In Choi et al. [15], for the purpose of diagnosing various retinal illnesses from fundus images, the authors suggested a CNN model. They used STARE dataset. The authors used VGG19 pre-trained model and achieved an accuracy of 37% approximately which is quite low. The authors mentioned this is due to fact that the model is trained to classify a large number of retinal diseases.

ArunKumar et al. [16] has put forth an unsupervised technique for categorizing retinal illnesses. They used a Restricted Boltzmann Machine (RBM) layer in their Generalized Regression Neural Network (GRNN), which can extract complex information. They have a 79.32% accuracy rate.

In [17], the authors put forth a model using InceptionV3 for detecting Diabetic Retinopathy (DB) and edema. They used the pre-trained model weights from ImageNet and updated them using Stochastic Gradient Descent (SGD) optimizer. They achieved specificity and sensitivity of 90.3% and 90.1% respectively on EyePACS dataset.

Li et al. [18] used InceptionV3 architecture for glaucoma detection and classification. For preprocessing of the image, they applied average color in local space to make all images to be equally bright.

Juan et al. [19] employed 5 different CNN architectures to detect glaucoma in a fundus image (CNN, VGG19, ResNet, GoogleNet, and DENet). They employed one confidential dataset obtained from Hospital de la Esperanza in Barcelona, Spain, as well as two publicly available datasets (DRISHTI-GS and RIM-ONE). Initially, as a preprocessing step they resized images by centering optical disk. The above task is done using morphology operations on the digital images and manual segmentation of images. In their work they mentioned the VGG19 architecture have achieved highest accuracy, sensitivity and specificity compared to other architecture models.

Ram et al. [20] approach is about classification of normal, cataract, AMD and myopia. They found that increasing number of FC layers depends on dataset that has been used. They investigated the relationship between the quantity of classes and fully connected (FC) layer numbers. Slow learning of the neural network is caused by increasing the FC layer count along with rising dataset complexity.

Tariqul Islam et al. [21] developed a technique for employing a CNN to find eight distinct retinal disorders. Contrast Limited Adaptive Histogram Equalization (CLAHE) was used to preprocess the images before they were fed into the CNN model. Their AUC value was 80.5%, their F1 score was about 85%, and their Kappa score was 31%.

A CNN is a neural network that uses convolutional, pooling, non-linear activation functions, etc. to extract complicated characteristics from images [22][23]. The advantage of using CNN is that it automatically extracts complex features i.e., manual feature extraction is not required. This feature of CNN makes it a popular model to deal with images. A new model is proposed in this paper to identify and classify Cataract, Diabetic Retinopathy, Glaucoma diseases from a fundus image. The proposed architecture's flowchart is shown in the fig. 2. The dataset that has been used is collected from Kaggle Ocular Recognition [11].

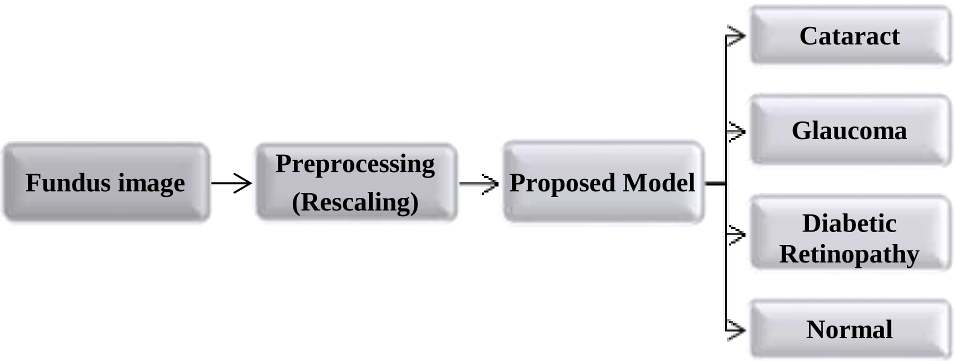

Fig. 2. Flow chart

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 efficientnetb3 (Functional)  (None, 1536)             10783535

 batch_normalization_4 (Batc  (None, 1536)             6144
 hNormalization)

 dense_3 (Dense)             (None, 256)               393472

 dropout_2 (Dropout)         (None, 256)               0

 dense_4 (Dense)             (None, 4)                 1028

=================================================================
Total params: 11,184,179
Trainable params: 11,093,804
Non-trainable params: 90,375
_________________________________________________________________
```

Fig. 3. Proposed model architecture

The architecture of the model is as shown in fig.3.The model takes a RGB image of size 224x224 as an input. The model uses pre-trained EfficientNetB3 model having 10,783,535 parameters and these are set to be trainable when training the model. The output from EfficientNetB3 is obtained by setting the parameter pooling to max and then is fed to a Dense layer having 256 neurons whose activation is ReLU, after passing it through a Batch Normalization layer. Dropout rate of 0.4 was used to prevent the model from overfitting. To further reduce overfitting, the L1 and L2 regularizers are used as bias regularizer and kernel regularizer respectively in the Dense layer. Instead of using standard values of l1 and l2, we have the parameters in l1 and l2 as 0.006 and 0.016 respectively. Then, the output from dropout layer is fed to Dense layer having 4 neurons that represents our four classes viz., cataract, diabetic retinopathy, glaucoma and normal. The model predicts the probabilities of every class, and the class with the maximum probability receives the output. SoftMax activation function is used in the output dense layer (1).

$$S(Z_i) = \frac{e^{Z_i}}{\sum_{j=1}^{K} Z_j} \quad (1)$$

The model has a total of 11,184,179 parameters out of which 11,093,804 are trainable and 90,375 are non-trainable parameters. Finally, in compilation step, the model is compiled using Adamax optimizer whose learning rate is specified at 0.001, and the loss is categorical cross entropy.

III. RESULTS

The performance of the proposed EfficientNetB3 is discussed in this section. Using the same dataset, this model and two additional pre-trained algorithms were assessed for cataract, diabetic retinopathy, and glaucoma diagnosis. The experimental Results are compared to cutting-edge approaches cataracts, diabetic retinopathy, and glaucoma.

Methods	**Accuracy (%)**			**Precision (%)**			**Recall (%)**			**F1-Score (%)**		
	SGD	AD-AM	ADA-MAX	SGD	AD-AM	ADA-MAX	SGD	ADA-M	ADA-MAX	SGD	AD-AM	ADA-MAX
VGG19	80	88	88.9	86.4	90.5	91.3	72.3	85.9	85.4	81.6	88	88.8
ResNet50	89.4	92.5	91.9	91.6	93.6	93.1	88.7	91.3	89.8	86.9	89.4	89.5
Efficient NetB3	87.5	90	90.8	90.2	91.9	92.8	84.6	88.4	88.1	84.4	90	89.4
Efficient NetB3 with Fine Tuning	92.1	92.4	**93.8**	92.6	94.1	**95**	91.3	91.4	**93.5**	89.7	91.9	**93.3**

Table 1. Comparison between the performance of the proposed three pre-trained models

On the same dataset, we create two pre-trained models (with good classification results) in order to assess how well EfficientNetB3 performs. These models are (i) VGG-19 [22], and (ii) ResNet50[23]. In Table 1, the effectiveness of our EfficientNetB3 is measured in terms of Accuracy, Precision, Recall, and F1-Score in comparison to various pre-trained models. The best performance is highlighted in bold. In all evaluation metrics, our suggested EfficientNetB3 takes first place and performs at the highest level. In comparison to the remaining pre-trained models, EfficientNetB3 exceeds them with an accuracy of 87.3%. Fig. 4 shows the EfficientNetB3's confusion matrix. A predicted label and an actual label, respectively, are assigned to each row and column of the table. Out of 633 valid images 48 are wrongly classified. From fig. 4 we can see 10 cataract images are wrongly classified as Glaucoma and normal, 2 Diabetic Retinopathy are wrongly classified as normal, 21 Glaucoma images are wrongly classified as Cataract and normal, and 15 normal images are wrongly classified as Cataract, Diabetic Retinopathy and normal. In fig. 5, illustrate the validation loss, accuracy, precision and recall curves of the EfficientNetB3 model respectively. The number of epochs is represented on the X-axis, and the loss, accuracy, precision, and recall values are on the Y-axis. These graphics demonstrate the validation and training loss curves have merge together after 6th epoch whereas, accuracy, precision and recall of validation and training curves become parallel after 6th epoch.

IV.DISCUSSIONS

Table 3 summarises the reported studies on deep learning-based cataract, diabetic retinopathy, and glaucoma detection and provides the studies' accuracy, precision, recall, and f1-score. High values for recall, accuracy, precision, and f1-score have been bolded. Our model has the high accuracy, precision, recall, and f1-score compared to the previous works of [12] and [13]. Our model is performing poorly on glaucoma and has less accuracy when compared to the [14]. However, when compared to other models [14], this model outperforms them in terms of accuracy, recall, and f1-score. [12] used Deep learning-based algorithms to identify cataract with accuracy of 92%. [13] used Deep learning-based algorithms were used to identify detect diabetic retinopathy with about 85% accuracy. [14] Used Machine learning-based algorithms to identify Glaucoma with accuracy of 81%. Table 4 summarizes the predictions of the proposed model on a new dataset [25].

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 11, No. 1, November 2022

Fig. 4. The proposed model Confusion Matrix

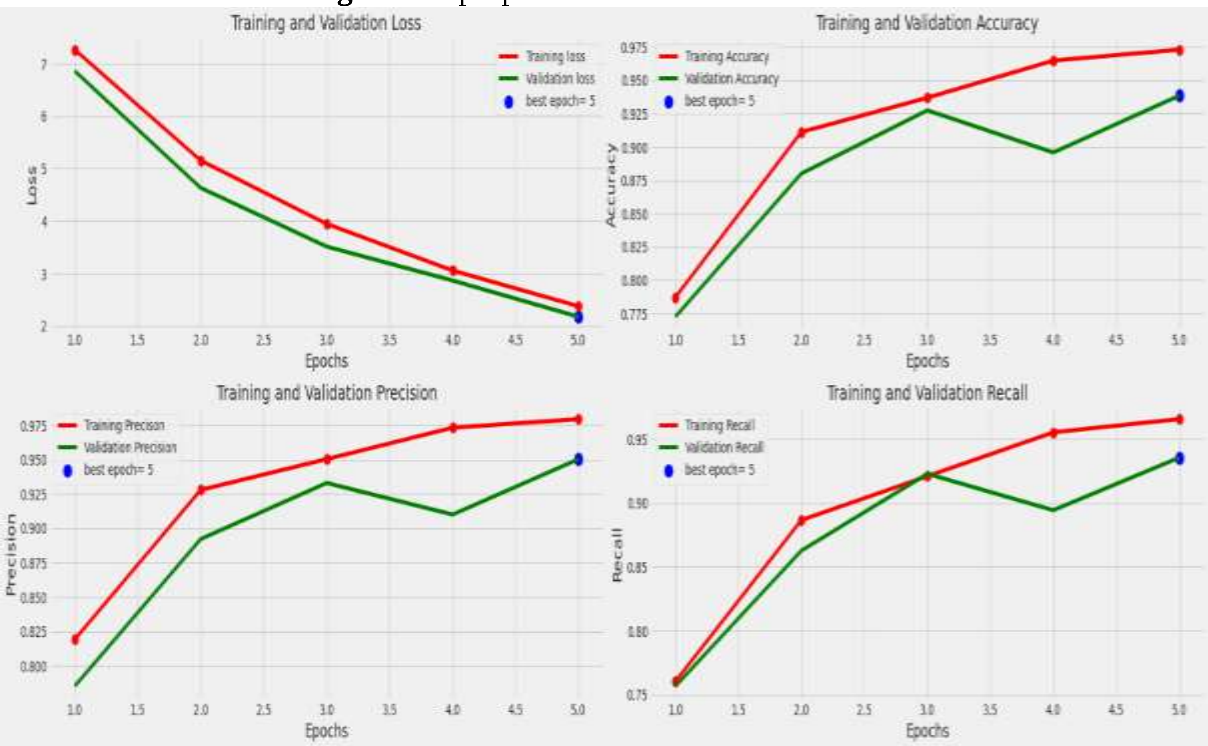

Fig. 5. Loss, Accuracy, Precision and Recall

Author	Ocular Disease	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
[12]	Cataract	92	90	89.6	90
[13]	Diabetic Retinopathy	85	67	55.6	59.6
[14]	Glaucoma	81	73	65	69

This Work	Cataract, Diabetic Retinopathy, Glaucoma	**92.3** **100** **89.4**	**97.3** **99.4** **88.2**	**92.3** **100** **89.4**	**94.7** **99.7** **88.8**
This Work	Average	**93.4**	**93.4**	**93.3**	**93.4**

Table. 2. Comparison between our model and other approaches in the available literature

Approaches	**Dataset**	**Accuracy**	**F1-score**
This work	Kaggle [25]	90.75	90.78

Table. 3. Predictions on new data

V. CONCLUSIONS

In this study, we introduced Efficient net B3, a deep learning-based automated system for the identification of glaucoma, diabetic retinopathy, and cataracts. The dataset used to feed the deep network was initially improved by pre-processing, rearranging, and augmenting a dataset of fundus images. The algorithm Efficient Net B3 was designed with the primary purpose of investigating various layers, activation functions, loss functions, and optimization techniques for decreasing computing cost while maintaining model validity. In comparison, the Efficient net B3 beat two of the three pre-trained CNN models, VGG19 and ResNet50. In terms of accuracy (91.50%), precision (92.69%), recall (91.50%), and f1-score (91.56%), our model beat state-of-the-art cataract, diabetic retinopathy, and glaucoma detection techniques. The ophthalmologists could quickly and more accurately diagnose cataract, diabetic retinopathy, and glaucoma disease utilising Efficient net B3 because to its excellent accuracy, cost- and time-efficiency. The disease like myopia, AMD, etc. cannot be detected by our model, but it can detect cataract, glaucoma, and diabetic retinopathy. The severity of the sickness may also be implemented in the future.

VI. ACKNOWLEDGMENTS

It gives us an immense pleasure to express deep sense of gratitude to our guide Dr. M. V. Nageswara Rao, Associate Dean (Academics), GMRIT for his valuable guidance and help in this work.

VII. REFERENCES

[1]. R R Bourne, S R Flaxman, T Braithwaite, M V Cicinelli, A Das, J B Jonas, J Keeffe, J H Kempen, J Leasher, & H Limburg (2017), “Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis”, The Lancet Global Health, vol. 5, no. 9, pp. e888– e897.
[2]. Guven (2013), “Automatic detection of age-related macular degeneration pathologies in retinal fundus images”, Comput. Methods Biomech. Biomed. Eng. 16(4), 425–434.
[3]. B Raju, N S D Raju, J Akkara, & A Pathengay (2016), “Do it yourself smartphone fundus camera DIYretCAM”, Indian J. Ophthalmol., vol. 64, no. 9, p. 663.
[4]. Dong Y, Wang Q, Zhang Q, & Yang J (2016), “Classification of cataract fundus image based on retinal vascular information”, In: International Conference on Smart Health, Springer, pp. 166–173.
[5]. Fan W, Shen R, Zhang Q, & Yang J J (2015), “Principal component analysis-based cataract grading and classification”, International Conference on E-health Networking, Application and Services (HealthCom), IEEE, pp. 459–462.
[6]. Manchalwar M, & Warhade K (2017), “Detection of cataract and conjunctivitis disease using histogram of oriented gradient”, Int. J. Eng. Technol. (IJET).

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 11, No. 1, November 2022

[7]. Qiao Z, Zhang Q, Dong Y, & Yang J J (2017), "Application of SVM based on genetic algorithm in classification of cataract fundus images", IEEE International Conference on Imaging Systems and Techniques (IST), IEEE, pp. 1–5.
[8]. Xiong L, Li H, & Xu L (2017), "An approach to evaluate blurriness in retinal images with vitreous opacity for cataract diagnosis", J. Healthc. Eng.
[9]. Krizhevsky A, Sutskever I, & Hinton G E (2012), "Imagenet classification with deep convolutional neural networks". In: Advances in Neural Information Processing Systems, pp. 1097–1105**.**
[10]. Pan S J, & Yang Q (2009), "A survey on transfer learning. IEEE Trans. Knowl", Data Eng. 22(10), 1345–1359.
[11]. https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k
[12]. Richard Bina Jadi Simanjuntak, Yunendah Fu'adah, Rita Magdalena, & Sofia Saidah (2022), "Cataract Classification Based on Fundus Images Using Convolutional Neural Network", JOIV, pp.33-38.
[13]. Zubiar Khan, Fiaz Gul Khan, Ahmad Khan, & Zia Ur Rehman (2021), "Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture", IEEE, vol.9, pp.61408-61416.
[14]. Danilo Leite, Maria Campelos, Ana Fernandes, & Pedro Batista (2021), "Machine Learning automatic assessment for glaucoma and myopia based on Corvis ST data", ELSEVIER, pp.454-460.
[15]. J Y Choi, T K Yoo, J G Seo, J Kwak, T T Um, & T H Rim (2017), "multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database", PloS one, vol. 12, no. 11, p. e0187336.
[16]. R Arunkumar & P Karthigaikumar (2017), "multi-retinal disease classification by reduced deep learning features", Neural Computing and Applications, vol. 28, no. 2, pp. 329–334.
[17]. V Gulshan, L Peng, M Coram, M C Stumpe, D Wu, A Narayanaswamy, S Venugopalan, K Widner, T Madams, & J Cuadros (2016), "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs", Jama, vol. 316, no. 22, pp. 2402–2410.
[18]. Z Li, Y He, S Keel, W Meng, R T Chang, & M He (2018), "Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs", Ophthalmology, vol. 125, no. 8, pp. 1199–1206.
[19]. J J G´omez-Valverde, A Ant´on, G Fatti, B Liefers, A Herranz, A Santos, C I Sanchez, & M J Ledesma-Carbayo (2019), "Automatic glaucoma classification using color fundus images based on convolutional neural networks and transfer learning", Biomedical optics express, vol. 10, no. 2, pp. 892–913.
[20]. A Ram & D C C Reyes-Aldasoro (2020), "The relationship between fully connected layers and number of classes for the analysis of retinal images", arXiv preprint arXiv:2004.03624.
[21]. M T Islam, S A Imran, A Arefeen, M Hasan, & C Shahnaz (2019), "Source and camera independent ophthalmic disease recognition from fundus image using neural network", IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON). IEEE, pp. 59–63.
[22]. M Simon, E Rodner, & J Denzler (2016), "ImageNet pre-trained models with batch normalization", arXiv:1612.01452.
[23]. K He, X Zhang, S Ren, & J Sun (2016), "Deep residual learning for image recognition", in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 770–778.
[24]. M Badar, M Haris, & A Fatima (2020), "Application of deep learning for retinal image analysis: A review", Computer Science Review, vol. 35, p. 100203.
[25]. https://www.kaggle.com/datasets/jr2ngb/cataractdataset