Diabetic Retinopathy Detection using Deep Learning

Supriya Mishra
Department of Electronics and Communications
Usha Mittal Institute of Technology
Mumbai, India
supriya94mishra@gmail.com

Seema Hanchate
Department of Electronics and Communications
Usha Mittal Institute of Technology
Mumbai, India
smhanchate.umit@gmail.com

Zia Saquib
Sr. Vice-President, Technology
Inovation and Department
Jio Platforms Ltd
Mumbai, India
zsaquib@gmail.com

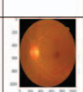
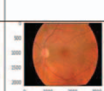
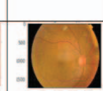
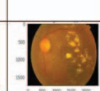
***Abstract*—Diabetic Retinopathy (DR) is human eye illness which occurs in individuals who have diabetics which harms their retina and in the long run, may lead visual deficiency. Till now DR is being screened manually by ophthalmologist which is a very time consuming procedure. And henceforth this task (project) focuses on analysis of different DR stages, which is done with Deep Learning (DL) and it is a subset of Artificial Intelligence (AI). We trained a model called DenseNet on an enormous dataset including around 3662 train images to automatically detect the DR stage and these are classified into high resolution fundus images. The Dataset which are using is available on Kaggle (APTOS). There are five DR stages, which are 0, 1, 2, 3, and 4. In this paper patient's fundus eye images are used as the input parameters. A trained model (DenseNet Architecture) will further extract the feature of fundus images of eye and after that activation function gives the output. This architecture gave an accuracy of 0.9611 (quadratic weighted kappa score of 0.8981) to DR detection. And in the end, we are comparing the two CNN architectures, which are VGG16 architecture and DenseNet121 architecture.**

Keywords*—Deep Learning, Diabetic Retinopathy (DR), DenseNet121 Architecture, VGG16 Architecture, Dataset, Fundus Camera.

I. Introduction

DR is the most debilitating form of diabetes in which serious damage occurs to the retina and causes visual impairments. It harms the veins inside the retinal tissue, making them spill fluid and contort vision. Alongside maladies prompting visual impairment like, waterfalls and glaucoma, DR is one of the most continuous diseases. There are five stages of DR that is 0, 1, 2, 3, and 4.

The below table gives the overall details about DR stages:

Levels	0	1	2	3	4
Diagnosis	No DR	Mild DR	Moderate DR	Severe DR	PDR
Description	No abnormalities	The most early stage, where only microaneurysms can happen	Ability of blood transportation due to their distortion and swelling with the progress of the disease	Blood supply to the retina due to the increased blockage of more blood vessels	The advanced stage, where the growth features secreted by the retina activate proliferation of the new blood vessels
Sample images with their levels					

Each stages has its own symptoms and specific properties, now from normal images doctors can not specify the DR stages. Moreover existing methods for diagnosing are very inefficient because it takes very large time, due to which the treatment may go the wrong way. To detect retino-pathy doctors used fundus camera which takes the picture of veins and nerves which is behind the retina. The initial phase of this disease has no signs of DR, so it turns into a real challenge to recognize it into a starting stage. For early detection we have used the different CNN (Convolutional Neural Network) algorithms, so that doctors can start the treatment at the correct time.

In this paper the dataset which we are using for the project is collected from "Aravind Eye Hospital" and it is available on kaggle that is "APTOS (Asia Pacific Tele-Ophthalmology Society)". We compare the two CNN architecture that is VGG16 architecture and DenseNet121 architecture, and showing the results of these two architectures.

In recent projects and researches, AI models, and in AI specially "Deep Learning" gives the most accurate outputs in finding hidden layers in various AI tasks, particularly in the field of medical image analysis [1]-[3]. Based on the deep learning models which are classify diseases and support medical decision making and can improve the persistent consideration (extra care) [4].

The remaining paper is organized as follows; Section II includes the litrature reviews of the DR image classification. Section III tells all about the dataset information. Section IV includes the Methodology of DL architectures. Section V tell us the main result of this project. Lastly the section VI concludes the paper.

II. Literature Review

In a particular topic it includes an overview of existing approaches that employed "Deep Learning" for DR automatic early detection.

A. Development and validation of a deep learning algorithm for DR automatic detection

Applied a deep learning to learn an algorithm for automatically detection of DR. Deep learning has ablity to program an algorithm itself because it is a computational methods and learning from a large set of examples that demonstrate the desired behavior. These techniques are uses in clinical imaging. The EyePACS-1 included 963 images from 4997 patients, the Messidor-2 had 1748 images from 874 patients. For the accuracy detection the algorithm had an area under the receiver operating curve of 0.991 (EyePACS-1) and 0.990 (Messidor-2) [5].

The automatic detection of DR is of vital importance, as it is the fundamental cause of irreversible vision loss in the working age or young age of populace in the world. The

978-1-7281-7213-2/20/$31.00 ©2020 IEEE

classification of DR images is very difficult even for trained clinicians. Therefore, using DCNN (Deep Convolutional Neural Network) for the classification of DR with an accuracy of 94.5% [6].

Currently, a novel DCNN, which plays out the beginning time identification by recognizing all microaneurysms (MAs), the first indication of DR, alongside accurately allotting names to retinal fundus pictures which had five classes. The architecture was tested on kaggle dataset and got the output of 0.851 QWK score and 0.844 AUC score. In the early stage recognition, the model showed the sensitivity of 98% and specificity of 94% which shows the effectiveness of technique [7].

An ensuring dataset fidelity by master verification of class labels improves acknowledgement of unobtrusive highlights and found that preprocessing with contrast limited AHE. Transfer learning on models from ImageNet improve accuracies to 74.5%, 68.8%, and 57.2% (2-ary, 3-ary and 4-ary) classification models, respectively [8].

Starting stage of DR can prevent this type of disease with correct tratment. A new feature extraction method that is Modified Xception Architecture has shown in the picture for the diagnosis of DR disease. This method shows that modified deep feature extractor improves DR classification with an accuracy of 83.09% versus 79.59% when compared with the original xception architecture [9].

The target is to automate the discovery of DR and access the seriousness with high efficiency, through a general possible methodology. Explore the utilization of different CNN architectures on pictures from the dataset in the wake of being subjected to suitable image processing techniques. The final results acquired through training. VGG16 gave an accuracy of 71.7% whereas the same for VGG19 gave 76.9% and Inception v3 was 70.2% [10]

Sadly the specific identification of the DR stage is famously precarious and requires expert human understanding of fundus pictures. Right now an automatic deep learning based method for DR stage identification by individual photography of human fundus. The method can be utilized as a method for early stage detection with sensitivity and specificity of 0.99 and QWK score is 0.925466 on APTOS Dataset [11].

III. Dataset

The image data used in this research was taken from dataset. The dataset which we used an open dataset that is this dataset can be used by anyone, which is collected from “Aravind Eye Hospital” which was easily available on Kaggle 4th APTOS (Asia Pacific Tele-Ophthalmology Society) 2019 Blindness Detection. This dataset was largest available on publicly to pre-training our CNNs architecture or model.

The dataset which we are using was provided with a large amount of high resolution retina images taken under a variety of imaging condition. The images which are provide in dataset are recorded from fundus camera which provides color fundus image of DR. A fundus camera is a low power microscope in which camera is attached and designed to take the picture of the interior surface of the eye [13]. The fundus image was used to document the DR condition that is images gave the clear picture for detection.

The clinicians are divided these DR into five classes which shows the stages of DR :

- No DR (class 0)
- Mild DR (class 1)
- Moderate DR (class 2)
- Sever DR (class 4)
- PDR (Proliferative DR) (class 5)

This dataset contains many folders like train.csv, test.csv, train_images, test_images, and sample_submission.csv. The below figure shows the information of folders:

```
['test_images', 'test.csv', 'sample_submission.csv', 'train_images', 'train.csv']
```

Fig. 1: List of folders in dataset

CSV (Comma Separated Values) file gives all the information of image and it is in excel sheet. Train.cvs contains the fundus eye image name and its severity level (class) and test.csv includes only the eye image name because it is going to be test after training the CNN architecture. Now the below picture is the sample image of fundus camera and it is the sample from dataset :

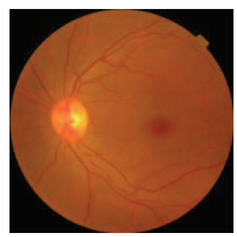


Fig.2: Sample image

The above figure shows all the nerves which is behind the eye. In our dataset all the image have 224X224 pixels and 3 channels that is RGB channel and divided into five classes. Dataset includes 3662 train images and 1928 test images (in below figure).

```
(3662, 224, 224, 3)
(3662, 5)
(1928, 224, 224, 3)
```

Fig.3: Number of train and test images

Again the fig.4 includes the counting’s of all the classes. Class 0 has 1805 images (number of people), class 1 has 370 images (number of people), class 2 has 999 images (number of people), class 4 has 295 images and class 3 has 193 images.

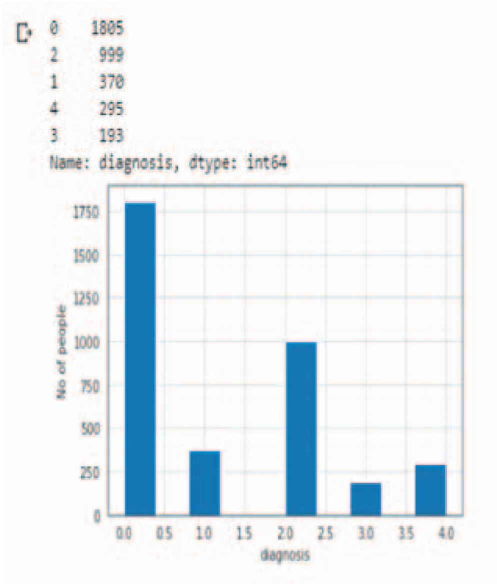


Fig.4: Number of images in each class

A. ImageNet

Our CNN architecture is pre-trained with ImageNet dataset. The ImageNet dataset improves the accuracy of CNNs model in our case it improve the accuracy of DenseNet121 architecture.

The ImageNet dataset is a very large set of photographs designed for developing the algorithms or models like computer vision, AI (Artificial Intelligence), ML (Machine learning) and DL (Deep learning). The Challenges, models and algorithms etc, uses the subsets means that images which we want to train from the ImageNet dataset when they have annual competition.

Based on the statics about the dataset recorded on the ImageNet there are 14 million different images linke animals, medical images, plant data, etc in the dataset. The goal of developing the dataset was to provide a resource to promote the research and development of improved methods for computer vision, AI, machine learning and deep learning.

IV. Methdology

As we know that DR detection problem is a primary cause of blindness. To overcome from this problem early detection is the first concern. So for early detection we are using the deep learning architecture called "DenseNet 121Architecture".

A. Deep learning framework for DR

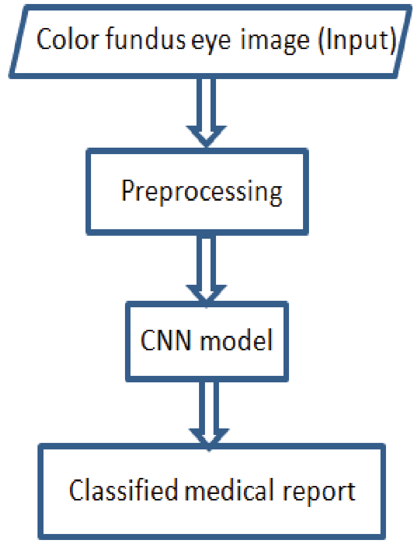


Fig.5: Deep learning framework

The above fig.5 is all about the framework of deep learning for DR.

1) Preprocessing : There are few steps which we have to follow during the preprocessing :

a) Take an image as an input.

b) Apply preprocessing technique to highlight the important features.

c) Cropping and resizing of image.

d) Proper data cleaning and removing black images.

e) Rotation and mirroring ofimages to balance the dataset, if the dataset is imbalance.

f) Conversion to numpy array.

g) Now use for traing or testing.

2) CNN model : After preprocessing, next step is train our CNN model or architecture. There are many CNN models or architetures are available in deep learning methods to train the network.

3) Medical report : Once we train our model, now we will get the final report that is output of input image. It means if we put any unseen image as a testing it will give the report of that unseen image.

B. Flowchart of our project :

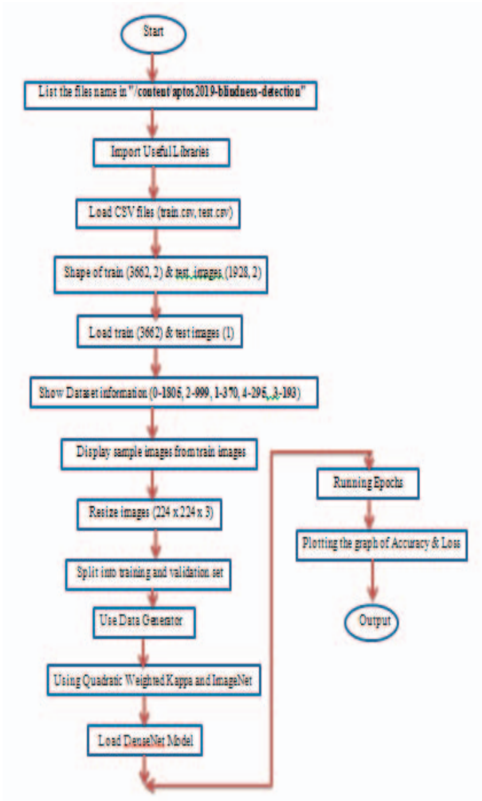


Fig.6: Flowchart

The fig.6 is full flowchart of our project, which uses the ImageNet for better accuracy with DenseNet architecture. For VGG16 architecture we don't use ImageNet and we will see the difference between VGG16 architecture and DenseNet 121 architecture. As we saw in above figure the flowchart is self-explanatory it includes preprocessing step, show the dataset information, display the shape of the image, using of quadratic weighted kappa and ImageNet and at the end running the epochs and got the output.

C. DenseNet 121 Architecture: The below figure shows the block diagram of DenseNet 121 architecture

DenseNets are increasing the depth or layer of DCNN. DenseNets exploit the potential of the network by reusing the feature. For DenseNet121 Architecture, there is no need to learn feature maps and requires fewer or lesser maps.

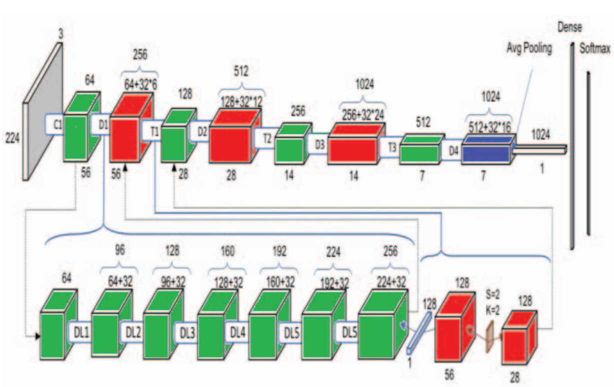


Fig.7: DenseNet 121 architecture

DenseNet architecture is an advance version of ResNet architecture. This architecture do not summation or add the outcome of the features of the layer with the incoming features but concatenate them.

DenseNet121 are broke into DenseBlocks, where the dimension of the featurs remains constant or unchange within a block, but the number of filters changes between the blocks, these layers are called transition layer.

As shown in the above figure, the measurement of each volume represents the sizes of the 2D that is its depth and width, whereas the numbers on the top which provides the features dimension. Here 32 is the growth rate of model. The volume of each block of denseblocks increases by the growth rate multiply by the number of dense layers within that denseblock. Every layer is adding to the previous of these 32 growth rate which is the new feature adding to it. By doing all this,layers are increasing from 64 to 256 after 6 layers. Furthermore transition block performed as 1 X 1 convolution with 128 filters . 2 X 2 pooling with a stride of 2, resulting on seperating the size of the volume and the number of features on half.

D. VGG16 architecture : The below figure shows the VGG16 architecture. We do not use the ImageNet in this architecture.

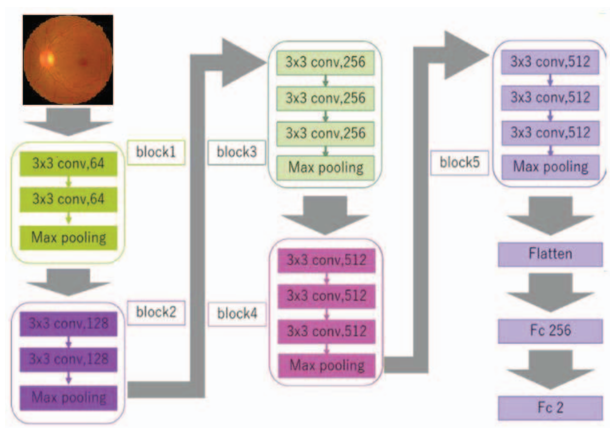


Fig.8: VGG16 achitecture

The input of conv1 layer is of same size (224 X 224), wherever we see the input it is of same size and it is a RGB image. The image is gone through a layers (multiple layers) of convolutional layers, where the filters were used. The padding of convolutinal layer input is the spatial resolution is preserved after convolution that is the padding is one pixel for 3 X 3 conv layers. Spatial pooling is perform by five max-pooling layers, which follow some of the conv layers. Max-pooling had over a 2 X 2 pixel window and it had stride 2. Fully connected (FC) layers which is almost the last layer, follows a stack of convolution layers (which has a different depth in different architecture). The FC layers have 4096 channels each.

E. Quadratic Weighted Kappa:

The quadratic weighted kappa is very useful when codes are ordered. Three matrices are included the matrix of observed score, the matrix of expected scores based on chance agreement, and the weight matrix. There are few steps to calculate the QWK, which is as follows :

Step 1: Create a multiclass confusion matrix (confusion_matrix) 0 between predicted and actual values.

Step 2: In step 2 each element is weighted. Predictions that are further away from actuals are marked harshly than predictions that are closer to the actuals (construct the weighted matrix which calculates the weight between the actual and predicted values).

Step 3: Create two vectors, one for preds and another for actuals, which provides how many values of each rating exist in each vectors (calculate value_counts() for each rating in preds and actuals).

Step 4: E is the Expected Matrix which is exterior product of the two vectors calculated in step 3 (calculate E, which is the outer product of two value_count vectors).

Step 5: Normalize both matrices to have same sum. Normalize E and 0 matrix.

Step 6: Calculate numerator and denominator of wighted kappa and return the weighted kappa matrix as 1-(num / den).

V. Results and Analysis

After done with the experiments, we got the experiment results in which we show the accuracy of our project. We used two architectures for same dataset and see the accuracies of each.

Architecture	Dataset	QWK	Loss	Accuracy
VGG16	Kaggle	Not used	0.7874	0.7326
DenseNet121	Kaggle	0.8981	0.1197	0.9611

As we seen clearly in the above table VGG16 is used without ImageNet and QWK and DenseNet is used with ImageNet and QWK. So without ImageNet VGG16 gives the less accuracy and with ImageNet DenseNet gives better accuracy than VGG16. Now will see the accuracy and loss graph of VGG16 and DenseNet respectively.

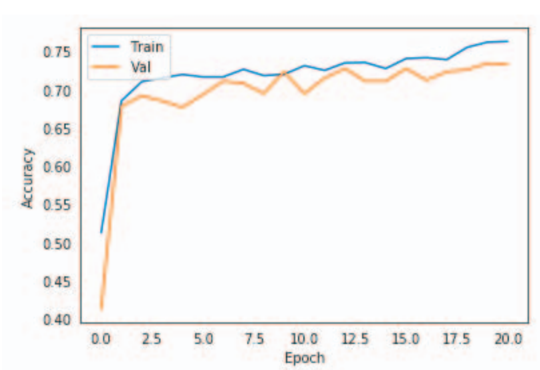


Fig.9: (a)

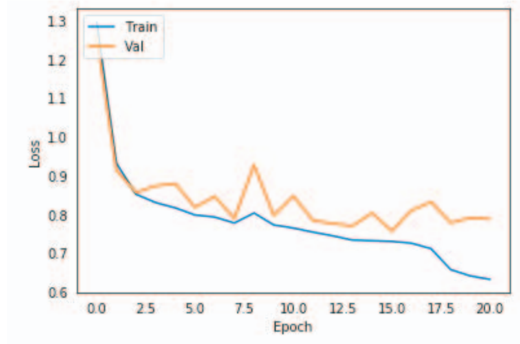


Fig.9: (b)

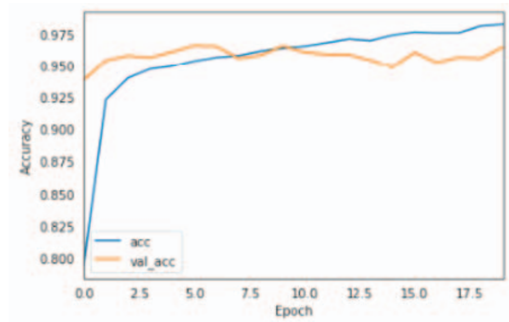


Fig.10: (a)

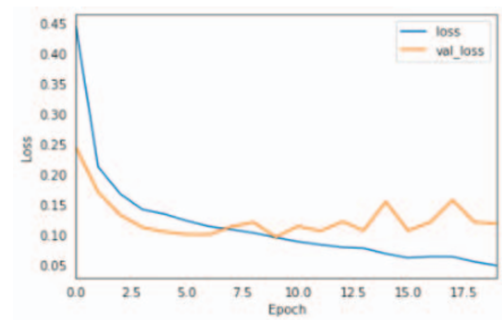


Fig.10: (b)

The above figures fig.9(a), fig.9(b) and fig.10(a), fig.10(b) are shown the accuracies and losses of VGG16 and DenseNet121 architectures respectively where VGG16 architectures do not used ImageNet and DenseNet architecture used ImageNet.

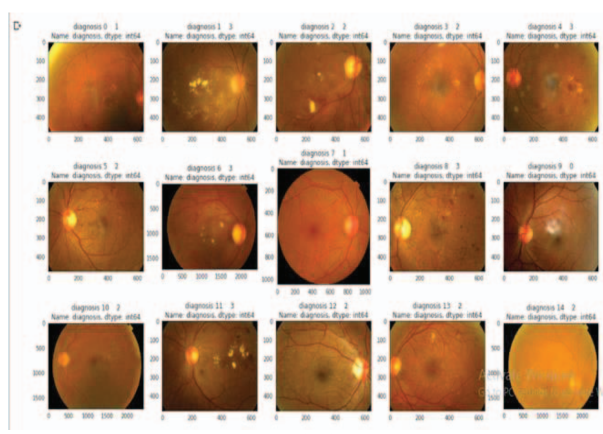


Fig.11: Multiple test images

Now if we want to detect the DR severity for at a time multiple images then it is possible to do. The above fig.11 shows the multiple images DR detection. The fig 12 shows the single test image which give the output of one image also.

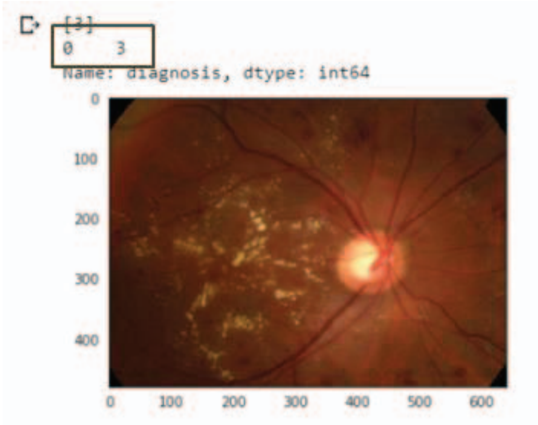


Fig 12: Single test image

VI. Conclusion

As we know that the DR (Diabetic Retinopathy) is primary concern for the diabetes patients, and manually it took a long time to detect DR. So we developed a architecture for automatic detection of DR, here we took two architectures to compare them that which architecture is best at what condition. The two architectures are VGG16 and DenseNet121 and the accuracies are 0.7326 and 0.9611 respectively. The QWK helped us to give the confidence of accuracy which we got from DenseNet architecture.

Acknowledgment

We wish to express our deepest gratitude to “Dr. Zia Saquib” who helped us a lot for this project and lastly thanked to our college to co-operate with us for completion of this project.

References

[1] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, ““Breast cancer diagnosis with transfer learning and global pooling,” arXiv preprint arXiv:1909.11839, 2019.

[2] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, R. Deters et al, “A hybrid deep learning architecture for leukemic blymphoblast classification,” arXiv preprint arXiv:1909.11866, 2019.

[3] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, , “Classification of histopathological biopsy images using ensemble of deep learning networks,” arXiv preprint arXiv:1909.11870, 2019.

[4] Xiaomin Zhou, Chen Li, Md Mamunur Rahaman, Yudong Yao et al. "A Comprehensive Review for Breast Histopathology Image Analysis Using Classical and Deep Neural Networks", IEEE Access, 2020

[5] Varun Gulshan, Subhashini Venugopalan, Rajiv Raman, “Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs,” JAMA. 2016;316(22):24022410. doi:10.1001/jama.2016.17216.

[6] Kele Xu, Dawei Feng, and Haibo Mi, “Deep Convolutional Neural Network-Based Early Automated Detection of Diabetic Retinopathy Using Fundus Image,” Received: 10 November 2017; Accepted: 22 November 2017; Published: 23 November 2017.

[7] Sheikh Muhammad Saiful Islam, Md Mahedi Hasan, and Sohaib Abdullah, “Deep Learning based Early Detection and Grading of Diabetic Retinopathy Using Retinal Fundus Images,” arXiv:1812.10595v1 [cs.CV] 27 Dec 2018.

[8] Lam C, Yi D, Guo M, Lindsey T., “Automated Detection of Diabetic Retinopathy using Deep Learning,” AMIA Jt Summits Transl Sci Proc. 2018 May 18;2017:147-155. PMID: 29888061; PMCID: PMC5961805.

[9] Sara Hosseinzadeh Kassani, Peyman Hosseinzadeh Kassani, Reza Khazaeinezhad, Michal J. Wesolowski et al. "Diabetic Retinopathy Classification Using a Modified Xception Architecture", 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2019

[10] Anuj Jain, Arnav Jalui, Jahanvi Jasani, Yash Lahoti, Ruhina Karani. "Deep Learning for Detection and Severity Classification of Diabetic Retinopathy", 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), 2019

[11] R Borys Tymchenko, Philip Marchenko and Dmitry Spodarets, “Deep Learning Approach to Diabetic Retinopathy Detection”.

[12] Weiguo Fan, Edward A. Chandan K. Reddy, “A Deep Learning Based Pipeline for Image Grading of Diabetic Retinopathy”.

[13] Eswar Kumar Kilari, Swathi Putta. " Delayed progression of diabetic cataractogenesis and retinopathy by in STZ-induced diabetic rats ", Cutaneous and Ocular Toxicology, 2016

[14] N. Yalin, S. Alver and N. Uluhatun, ”Classification of retinal images with deep learning for early detection of diabetic retinopathy disease,” 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp.