# Diabetic Retinopathy classification through fundus images using Deep Learning

Thumma Dharani
*Dept. of Electronics and Communication Engineering*
*V R Siddhartha Engineering College*
Vijayawada, India
thummadharani33382@gmail.com

Medikonda Padma Prasamsa
*Dept. of Electronics and Communication Engineering*
*V R Siddhartha Engineering College*
Vijayawada, India
padmaprasamsamedikonda14@gmail.com

Battina Harsha Vardhan
*Dept. of Electronics and Communication Engineering*
*V R Siddhartha Engineering College*
Vijayawada, India
harshabattina18@gmail.com

Jorige Bala Vivek
*Dept. of Electronics and Communication Engineering*
*V R Siddhartha Engineering College*
Vijayawada, India
vivek.jorige04@gmail.com

B. Lakshmi Sirisha
*Dept. of Electronics and Communication Engineering*
*V R Siddhartha Engineering College*
Vijayawada, India
suravarapuls@yahoo.co.in

***Abstract*— One of the most common eye diseases in the people aged between 20-74 years is Diabetic Retinopathy (DR). DR is an eye complication where the patient loses his vision due to an increase in glucose levels in the blood. DR is most prominent in the patients who are diagnosed with diabetic mellitus. Over one-third of the diabetic mellitus patients are diagnosed with DR. For diagnosing DR, the patient has to visit an ophthalmologist for dilated eye examination. However, everyone cannot have this facility. Hence, there is a need for a simple automated software for diagnosing the five stages of DR efficiently. In this paper, a simple model is developed using the Kaggle APTOS Blindness Detection dataset which is publicly available. In the pre-processing step the images are enhanced and the deep learning model ResNet152 architecture is used for the classification step. After training the ReseNet152 model yielded a training and validation loss of 0.073 and 0.107 respectively and validation accuracy of 0.984. Further, a simple Graphical User Interface is developed using tkinter framework in python standard library which classifies the given input fundus image as one of the five stages of DR.**

***Keywords—Diabetic Retinopathy, Deep Learning, ResNet152, Training Loss, Validation Loss, Validation accuracy, Graphical User Interface.***

## 1. INTRODUCTION

Diabetic Retinopathy (DR) is a complication of diabetic mellitus that mainly effects the vision of the patient. It became a very common disease in patients who are previously diagnosed with diabetic mellitus and this condition is rapidly increasing in the adults and now-a-days young people are also becoming a victim of DR. DR is a common complication that occur within 20 years of diagnosing the diabetes. This complication can occur either in Type-1 or Type-2 diabetes effected patients. If DR is not treated early on, it might eventually result in vision loss and blindness in the patient. The main symptoms of DR include blurriness, floaters, and dark spots in the field of vision. In DR, the light-sensitive layer (retina) on the rear of the eye will be seriously affected due to this complication. Further, the new abnormal blood vessels will develop and these blood vessels in the retina start to bleed into the vitreous because of their fragile nature and it eventually leads to severe eye complications like Diabetic Macular Edema (DME), Neovascular glaucoma, Retinal Detachment etc. Early Diabetic Retinopathy identification thus became crucial in the medical field. DR is mainly classified as four stages named as mild non-proliferative diabetic retinopathy, moderate non-proliferative diabetic retinopathy, severe non-proliferative diabetic retinopathy and Proliferative diabetic retinopathy as shown in figure 1.

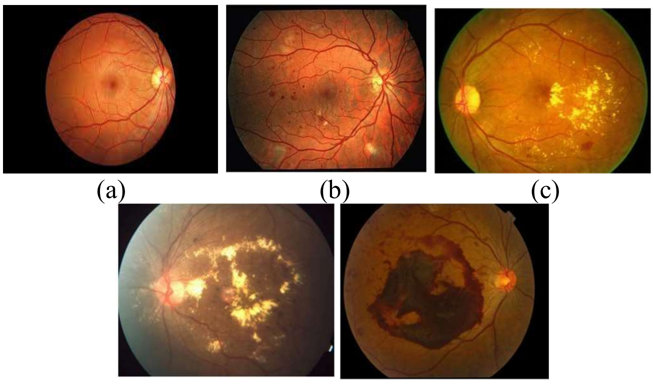


Figure 1. Various Stages of Diabetic Retinopathy (a)Normal Eye (b)Mild Non-proliferative DR (c)Moderate Non-proliferative DR (d)Severe Non-proliferative DR and (e)Proliferative DR

In the mild stage of DR, small bulges develop in the retina's blood vessels, these are termed as microaneurysms. These microaneurysms doesn't lead to any vision problems but it can lead to the progress of the disease to the next stages. In the moderate stage the blood vessels swell and the fluid spreads into the retina which leads to the vision loss and it further leads to the conditions like DME. In the severe stage, the rupturing of the blood vessels severely increases and more fluid spreads into the retina which results in the obstruction of retina's blood circulation. This condition leads to the creation of contemporary blood vessels which are fragile in nature in this stage. In final stage i.e., in the proliferative stage, these new blood vessels rupture due to its fragile nature and leads to retinal detachment and permanent vision loss.

Many patients are unaware about this condition and patients are only consulting the specialists when they are in proliferative stage. The hospitals need an ophthalmologist for DR diagnosis which can be done by using dilated eye examination. However, in rural areas it is hard to have an access for a specialist. Hence, for early detection of the DR easily and for detecting the different stages of DR, several automated softwares are developed using various techniques like image processing, Artificial Intelligence, Machine Learning and Deep Learning. Deep Learning is the latest technology in neural networks which offers a wide range of uses and one such use is in healthcare. To diagnose various diseases automatically, deep learning architectures provide a greater use. The deep learning model is mainly trained on the dataset that defines the disease, then the model extracts the features from the input dataset and utilize these features for learning purpose. The model will train itself based on these parameters and transform the model to be able to classify the disease. There are various architectures that are invented to build the model. Some of them are Alexnet, VGG16, Resnet, Densenet, Inception models, Efficientnet models etc. These models include successive convolution layers which are used for training.

## 2. Literature Survey

L. Qiao et al., [1] proposes the algorithm for early detection of diabetic retinopathy. Lesions and microaneurysms are the main signs of Diabetic Retinopathy. Hence, convolutional neural network algorithm with deep learning as the main component using GPU as the hardware accelerator is used for analysing whether microaneurysms are present in the fundus image by performing image detection and image segmentation. Semantic segmentation is mainly used to classify normal eye and DR infected eye which is done to determine the characteristics of a microaneurysm by dividing the image pixels according to their shared semantics. Also, through this approach lesions are detected accurately regardless of texture, form and scale of it. Hence, this algorithm can effectively identify Non-Proliferative Diabetic Retinopathy (NPDR) stage.

M. M. Abdelsalam et al., [2] In this study, they considered images of optical coherence tomography angiography (OCTA) type instead of normal retinal fundus images for various features analysis. An original MATLAB software was used to process the OCTA images. Here, generalised dimensions, lacunarity, and singularity spectrum are discovered using this geometry. For the automation of diagnosis process and for resultant accuracy improvement which is done by classification and regression, a Support Vector Machine (SVM) algorithm of a supervised learning method is used. This methodology achieved 98.5% accuracy, 100% sensitivity, 97.3% specificity, and 96.8% precision.

T. Araújo et al., [3] concentrates on the identification of Proliferative DR which is the last stage of DR. In order to compensate for the paucity of Proliferative DR cases in DR-labelled datasets, a data augmentation approach based on the heuristics according to the synthesis of neo vascular (NV)-like structures is presented. The use of the neo vessel generation algorithm, which mostly depends on common knowledge of the position and geometry of these structures is made. In order to expand the training sets for deep neural networks, NVs are created and introduced in retinal images that already exist. The creation of prototype NVs using the semi-random graph-based method is the first step in the NV-based data augmentation. The type of vessel and the location insertion are considered while processing the produced graphs.

S. Majumder et al., [4] provides a multitasking deep learning model to find all five DR phases. A classification model and a regression model make up the two models that make up this multitasking model. The characteristics are extracted, concatenated, and fed to a multilayer perceptron network to classify the five stages of DR, and a DenseNet architecture is used to identify the stages of DR. These two models are trained independently. Applying the created multitask model to the two significant Kaggle datasets, APTOS and EyePACS, yielded kappa scores of 0.90 and 0.88 for each dataset respectively. The micro average and macro average areas of the ROC curve are 0.96 and 0.93 respectively.

Ratana pakorn T et al., [11] proposed an automated software using top down programming. The severity classification of DR and different features like neovascularization, vessels, fovea and macula localisation are extracted to detect pathologies of DR and were carried out using the MATLAB image processing toolkit and MATLAB R2015a. The MATLAB GUI toolkit is then used to construct a graphical user interface. This software's DR detection accuracy, sensitivity, and specificity were each 96.25%, 98%, and 67%, respectively. This software has a detection accuracy of 66.58% for both proliferative and non-proliferative DR. One fundus image typically takes 7 seconds to process.

Abhishek Samanta et al., [12] suggested a convolutional neural network (CNN) architecture based on transfer learning for colour fundus photography using a significantly smaller dataset of skewed class consisting of training images of 3050 and validation images of 419 for classifying DR. For stage identification, various characteristics including blood vessels, texture, and hard exudates are employed. The dataset underwent four stages of DR training on Google Colab. The training was done on five different layers using DenseNet model which is effective when compared to other models. Finally, Cohen's kappa obtained for validation data is 0.8836 and for training data is 0.9809. The screening time is 0.99 seconds which makes it robust.

It was observed that the deep learning models of CNN (Convolution Neural Networks) like ResNet, DenseNet and EfficientNets gave better results based on literature survey. The primary objective of this survey is to propose a deep learning model with greater accuracy than the proposed models for the classification of five stages of DR including No DR and developing a GUI for diagnosing these stages easily without the need for an ophthalmologist.

## 3. Proposed Method

In the CNN model, the deep neural networks consist of an input layer, output layer and a set of successive convolution layers in between them for the extraction of features from the dataset images and use these features for learning and classification of the five stages of DR. In this paper, the ResNet152 architecture containing 152 layers is used for training the model after enhancing the fundus images using several pre-processing techniques and the model's performance is calculated based on the parameter "accuracy".

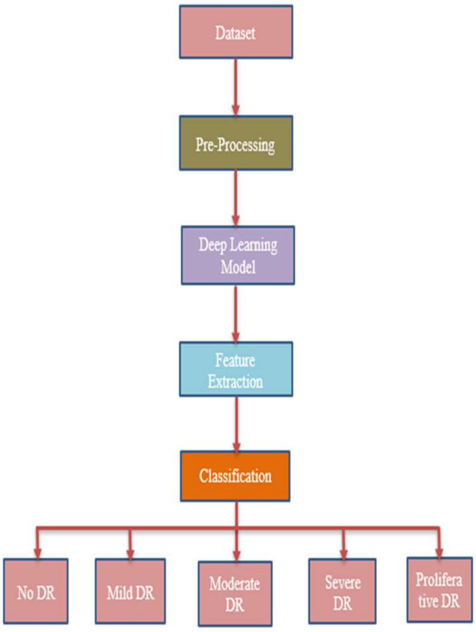


Figure 2. Block Diagram

Figure 2 represents the block diagram of the proposed method which describes the flow of the model by considering different stages like pre-processing, deep learning model and classification.

### *3.1. Dataset and Software*

The Dataset for our model is collected form the publicly available Kaggle APTOS 2019 blindness detection dataset. The dataset consists of 5590 fundus images in which 3662 images are segregated for training and 1928 images for testing. The training images are labelled from 0 to 4 according to the stages of DR as No DR, Mild stage of DR, Moderate stage of DR, Severe stage of DR and Proliferative stage of DR respectively. The number of images for the classes from 0 to 4 are 1805, 370, 999, 193 and 295 images respectively which are displayed in the figure 3. Google Colab software accelerated with GPU is utilized for pre-processing the images and training the model using the deep learning model ResNet152.

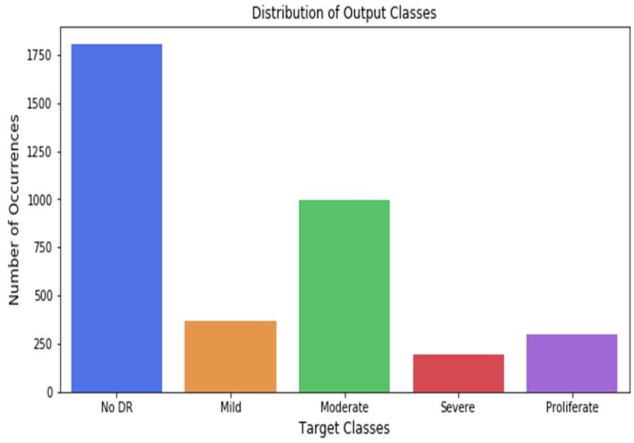


Figure 3. Distribution of five Training Classes

### *3.2. Pre- Processing*

The next step after collecting data is to apply pre-processing techniques as shown in the figure 2 in order to improve the features of the data. For pre-processing, random images in the dataset are presented from the labels 0 to 4 in the figure 4. In order to define the image's region of interest (ROI) for these images, a binary mask is applied after converting the RGB image to gray scale image. Then, the image is resized to an optimal size of 456X456 pixels by padding with zeroes or by cropping the extra pixels in order to preserve the features without any shrinkage of image. To these masked and resized fundus images, gaussian noise is added to the training set of images in order to improve the variability and robustness of the images such that the overfitting problem which mainly occurs in the deep learning procedures can be avoided. After applying the pre-processing techniques, the features are improved and become prominent as shown in figure 5 and these pre-processed images are considered for training purpose.

Hence, by applying these pre-processing methods to the training set of fundus images in the dataset, the data can be augmented and the tiny features are more highlighted which improves the learning of the model and leads to better results in the training since the features are more prominent in the obtained pre-processed images.


Figure 4. Random Input images from the labels 0 to 5

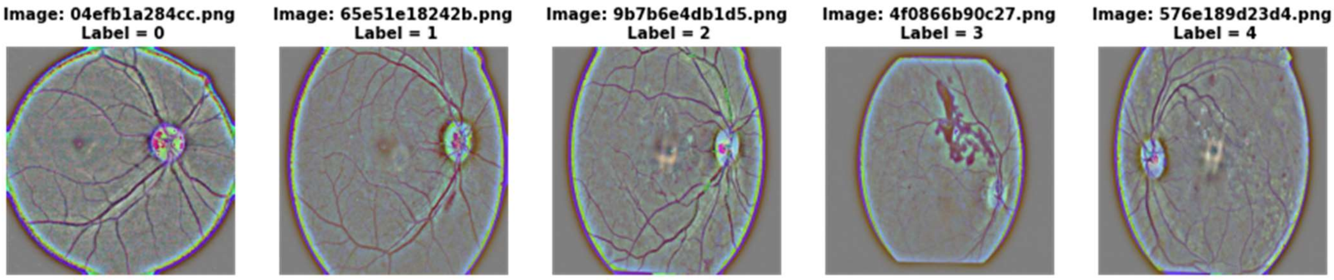


Figure 5. Image after subjecting to pre-processing techniques

### *3.3. Deep Learning Model*

Deep learning models mimics the human biological nervous system to learn the different features of images, audio and video. The Neural Networks in Deep Learning are inspired from the Human Nervous System. They consist of layers for learning purpose. The layers are together interconnected with each other nodes successively. The connections are assigned with some weights. The first layer acts as an input layer which takes the data as input and the final layer classifies the stages. The layers in-between are hidden layers and are used for training the images by extracting the features from the input images. There are several architectures in deep learning. Each architecture differs in the number of layers and differ in the interconnection between the layers. Here, we considered the architecture ResNet152 for training our model.

#### *3.3.1. ResNet152 Architecture*

The Residual Network is a deep neural network. The ResNet here used is of 152 layers. In a plain neural network, the connection between the layers is successive one by one whereas in the deep neural network like ResNet the layers are connected successively and also connected through skip connections. The skip connections are used for identity mapping. Thus, the usage of skip connections gives the deep layers for the ResNet architecture. The skip connection translates the input of one convolution layer to the output of the following convolution layer as shown in figure 6. The issue of disappearing gradients is minimised by ResNet's use of skip connections and it makes sure that the layer doesn't perform worse than the previous layers and that the highest layer outperforms the lowest. Also, the ResNet is of bottle neck design in which a 3x3 convolution layer is preceded and succeeded by a 1x1 convolution layer. Hence, in this way the layers in this architecture are increased. The increased number of layers leads to the increase in training, thereby increasing the accuracy with lesser number of trainable parameters.

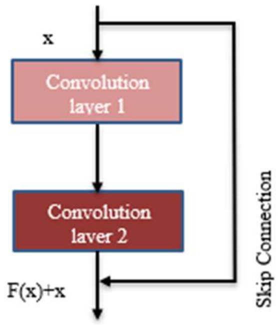


Figure 6. Skip connection in ResNet Architecture

In the ResNet layers; first, there is a max pooling layer and then the convolution layers succeeded by average pooling layer. Then a fully connected layer is presented. In the convolution layers the batch normalization function is used for the improvisation of the learning rate and a ReLU activation function is presented for preventing the activation of neurons at the same time.

The architecture of ResNet152 is clearly demonstrated in figure 7. The first convolution layer of the ResNet152's one-fifty-two-layer design has kernels that are 7x7 in size, and there are 64 total kernels with a stride size of 2 which takes the dataset images as input. Next, there is a layer called maxpooling with a stride size of 2. There is a 3x3,64 kernel in the following convolution layer that is padded with a 1x1,64 kernel above and a 1x1,256 kernel below. These layers are then repeated three times, giving us a total of 9 layers in this stage. There is a 3x3,128 kernel in the following step that is padded with a 1x1,128 kernel above it and a 1x1,512 kernel below it. These layers are then repeated eight times, giving us a total of 24 layers in this phase. Then, we have a 3x3,256 kernel in the following convolution layer, which is padded with a 1x1,256 kernel above it and a 1x1,1024 kernel below it. These levels are then repeated thirty-six times, giving us a total of 108 layers in this stage. In the subsequent stage, a 3x3,512 kernel is present that is padded with 1x1,512 kernels above and 1x1,2048 kernels below. These layers are then repeated three times, giving us a total of 9 layers. Following that, there is an average pool layer, followed by a fully connected layer with a total of 1000 nodes, before the softmax function concludes the layer hierarchy and gives us one layer.

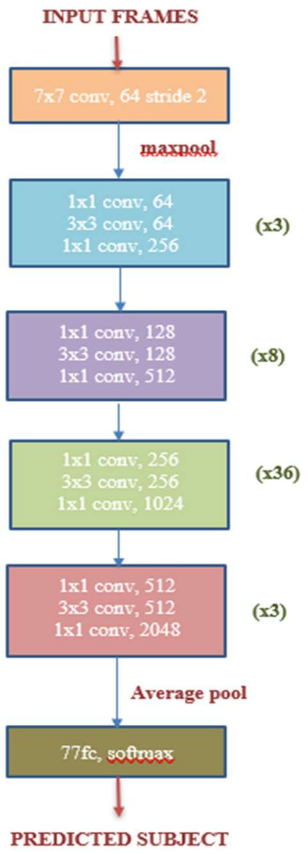


Figure 7: ResNet 152-layer Architrecture

Hence, the ResNet architecture is considered because of its deeply connected layers through skip connections and a greater number of layers can result in better accuracy with a smaller number of parameters. The network helps the model to train effectively by repeatedly training the parameters cause through the skip connection the input is again added to the output throughout the network layers. The model is totally trained on the 58970117 trainable parameters and on five epochs.

### *3.4.GUI Development*

The development of the graphical user interface is done by using the tkinter framework in the python GUI toolkit. The model is loaded into the IDLE shell and the easy to use interface is developed by importing several modules. The GUI developed takes the fundus image as input and displays the label, predicted class and the severity of the disease in the IDLE.

## 4. RESULTS

The model is trained on 5 epochs. After the completion of 5 epochs the training loss and validation loss are 0.073 and 0.107 respectively, the validation accuracy is 0.984 which are respectively demonstrated in figure 8 and figure 9. As the number of Epochs increases the loss values are decreased as in figure 8 and the accuracy values are increased as in figure 9 for both training and validation dataset. The model exhibits optimal fitting nature on further training, hence the training is terminated at these epochs.

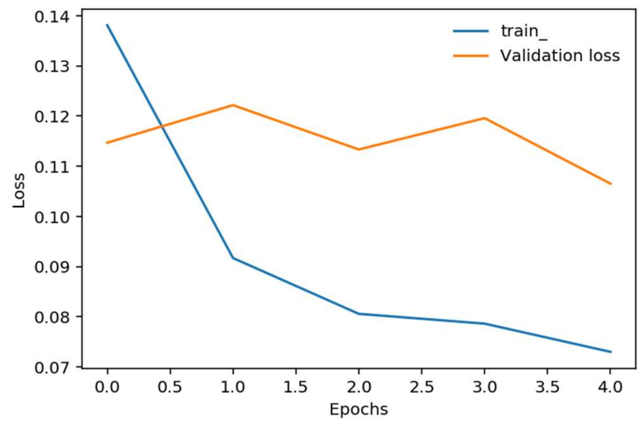


Figure 8: Training and Validation loss

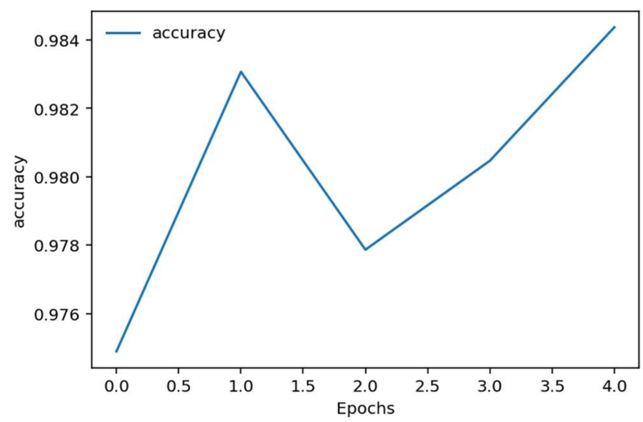


Figure 9: Validation Accuracy

```
E:/PROJECT_DR/NON PRO DR2.jpg
Transforming your image...
Passing your image to the model....
Predicted Severity Value:  2
class is:  Moderate
Predicted Label is 2
Predicted Class is Moderate
E:/PROJECT_DR/PRO DR.jpg
Transforming your image...
Passing your image to the model....
Predicted Severity Value:  4
class is:  Proliferative DR
Predicted Label is 4
Predicted Class is Proliferative DR
```

Figure 10: Output in python IDLE shell

The GUI for classification of stages takes the fundus image of .png format as input and gives the output in idle shell which describes the stage of DR as shown in figure 10.

## 5. Conclusion

Hence, the ResNet model yielded an accuracy of 0.984 and the interface can be effectively used to detect the five stages. However, the model is unable to detect the severe stage of DR accurately because of the small number of fundus images in the dataset. This stage can be effectively detected by increasing the number of images for this stage. Also, apart from ResNet deep neural network, many deep neural network architectures can yield greater accuracies which can be done as future work.

## References


**[1]** L. Qiao, Y. Zhu and H. Zhou, "Diabetic Retinopathy Detection Using Prognosis of Microaneurysm and Early Diagnosis System for Non-Proliferative Diabetic Retinopathy Based on Deep Learning Algorithms," in IEEE Access, vol. 8, pp. 104292-104302, 2020.

**[2]** M. M. Abdelsalam and M. A. Zahran, "A Novel Approach of Diabetic Retinopathy Early Detection Based on Multifractal Geometry Analysis for OCTA Macular Images Using Support Vector Machine," in IEEE Access, vol. 9, pp. 22844-22858, 2021.

**[3]** T. Araújo et al., "Data Augmentation for Improving Proliferative Diabetic Retinopathy Detection in Eye Fundus Images," in IEEE Access, vol. 8, pp. 182462-182474, 2020.

**[4]** S. Majumder and N. Kehtarnavaz, "Multitasking Deep Learning Model for Detection of Five Stages of Diabetic Retinopathy," in IEEE Access, vol. 9, pp. 123220-123230, 2021.

**[5]** A. Momeni Pour, H. Seyedarabi, S. H. Abbasi Jahromi and A. Javadzadeh, "Automatic Detection and Monitoring of Diabetic Retinopathy Using Efficient Convolutional Neural Networks and Contrast Limited Adaptive Histogram Equalization," in IEEE Access, vol. 8, pp. 136668-136673, 2020.

**[6]** E. Hacisoftaoglu, MahmutKarakaya, Ahmed B. Sallam, Deep learning frameworks for diabetic retinopathy detection with smartphone-based retinal imaging systems, Pattern Recognition Letters, Volume 135, 2020, Pages 409-417.

**[7]** Shiva Shankar Reddy, Nilambar Sethi, R. Rajender, Gadiraju Mahesh, Extensive analysis of machine learning algorithms to early detection of diabetic retinopathy, Materials Today: Proceedings, 2020.

**[8]** M. Sakthi sree devi, S. Ramkumar, S. Vinurajkumar, G. Sasi, Detection of diabetic retinopathy using OCT image, Materials Today: Proceedings, Volume 47, Part 1, 2021, Pages 185-190.

**[9]** Ayaka Sugeno, Yasuyuki Ishikawa, Toshio Ohshima, Rieko Muramatsu, Simple methods for the lesion detection and severity grading of diabetic retinopathy by image processing and transfer learning, Computers in Biology and Medicine, Volume 137, 2021, 104795.

**[10]** Vaibhav V. Kamble, Rajendra D. Kokate, Automated diabetic retinopathy detection using radial basis function, Procedia Computer Science, Volume 167, 2020, Pages 799-808.

**[11]** Ratanapakorn T, Daengphoonphol A, Eua-Anant N, Yospaiboon Y, Yospaiboon Y, Digital image processing software for diagnosing diabetic retinopathy from fundus photograph, Clin Ophthalmol. 2019; 13:641-648.

**[12]** Abhishek Samanta, AheliSaha, Suresh Chandra Satapathy, Steven Lawrence Fernandes, Yu-Dong Zhang, Automated detection of diabetic retinopathy using convolutional neural networks on a small dataset, Pattern Recognition Letters, Volume 135, 2020, Pages 293-298.

**[13]** Pustokhin, Denis & Pustokhina, Irina & Dinh, Phuoc & Phan, Son & Nhu, Nguyen & Joshi, Gyanendra Prasad& K., Shankar. (2020). An effective deep residual network-based class attention layer with bidirectional LSTM for diagnosis and classification of COVID-19. Journal of Applied Statistics. 1-18

**[14]** Ložnjak, Stjepan & Kramberger, Tin & Cesar, Ivan & Kovačević, Renata. (2020). Automobile ClassificationUsing Transfer Learning on ResNet Neural Network Architecture. 10.19279/TVZ.PD.2020-8-1-18.

**[15]** Kohner, Eva M., Vinod Patel, and Salwan MB Rassam. "Role of blood flow and impaired autoregulation in the pathogenesis of diabetic retinopathy." *Diabetes* 44.6 (1995): 603-608.

**[16]** Gargeya, Rishab, and Theodore Leng. "Automated identification of diabetic retinopathy using deep learning." *Ophthalmology* 124.7 (2017): 962-969.

**[17]** Gulshan, Varun, et al. "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs." *Jama* 316.22 (2016): 2402-2410.

**[18]** Abràmoff, Michael David, et al. "Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning." *Investigative ophthalmology & visual science* 57.13 (2016): 5200-5206.

**[19]** Takahashi, Hidenori, et al. "Applying artificial intelligence to disease staging: Deep learning for improved staging of diabetic retinopathy." *PloS one* 12.6 (2017): e0179790.

**[20]** Bellemo, Valentina, et al. "Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study." *The Lancet Digital Health* 1.1 (2019): e35-e44.

**[21]** Sayres, Rory, et al. "Using a deep learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy." *Ophthalmology* 126.4 (2019): 552-564.

**[22]** Hacisoftaoglu, Recep E., Mahmut Karakaya, and Ahmed B. Sallam. "Deep learning frameworks for diabetic retinopathy detection with smartphone-based retinal imaging systems." *Pattern recognition letters* 135 (2020): 409-417.


## Web Sources


1. Dataset - https://kaggle.com/competitions/aptos2019-blindness-detection

2. https://towardsdatascience.com/review-resnet-winner-of-ilsvrc-2015-image-classification-localization-detection-e39402bfa5d8

3. https://towardsdatascience.com/creating-densenet-121-with-tensorflow-edbc08a956d8