Blindness Detection Using Deep Learning

Minakshi Mohanty

June 2020

Introduction:

Diabetic retinopathy (DR) is a condition that can cause damage to blood vessels in the retina. This condition can occur in anyone with either type I or type II diabetes. According to a study published by Diabetes Canada, Diabetic Retinopathy is the most common cause of blindness in working age people. The Eye Diseases Prevalence Research Group determined that in the US, the crude prevalence rate of retinopathy in adults with diabetes is 40.3%; sight-threatening retinopathy occurred at a rate of 8.2% [1].

Early stages of diabetic retinopathy may not present any symptoms. However, as the disease progresses, the following symptoms may occur:

- Spots or dark strings floating in your vision (floaters)
- Blurred vision
- Fluctuating vision
- Impaired color vision
- Dark or empty areas in your vision
- Vision loss

A comprehensive dilated eye examination is conducted to diagnose diabetic retinopathy. In this exam, eye drops are used to dilate the pupils and fundus images are taken to assess blood vessels, optic nerve head and retina. The image below shows the difference in the fundus images of a healthy and affected eye. In early stages of DR, the walls of the retinal blood vessels begin to weaken. Tiny bulges protrude out of the blood vessels and may sometimes leak blood into retina. There may be swelling in the retinal nerve fibers, causing white spots to form in the retina. As the disease progresses, new blood vessels may grow to and burst causing damage to vision.

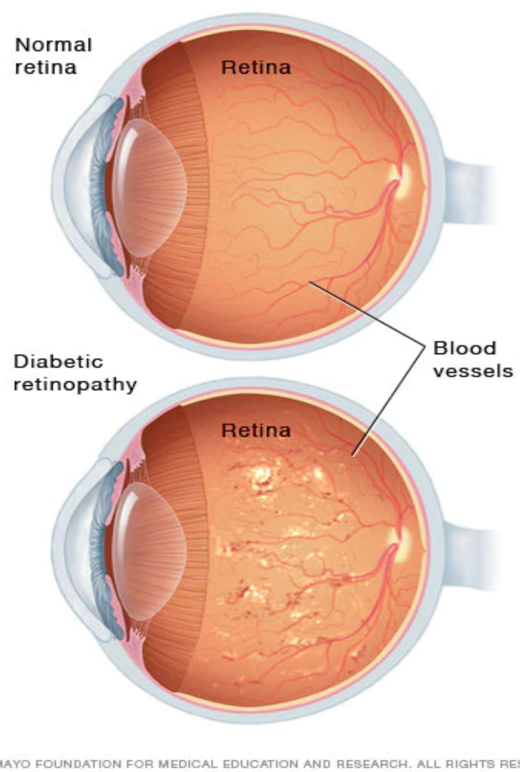


Image of a health and diabetic retinopathy affected eye (2)

Objective: The goal of the project is to develop a new algorithm that can effectively identify the severity of diabetic retinopathy from retina images and classify it into the following categories:

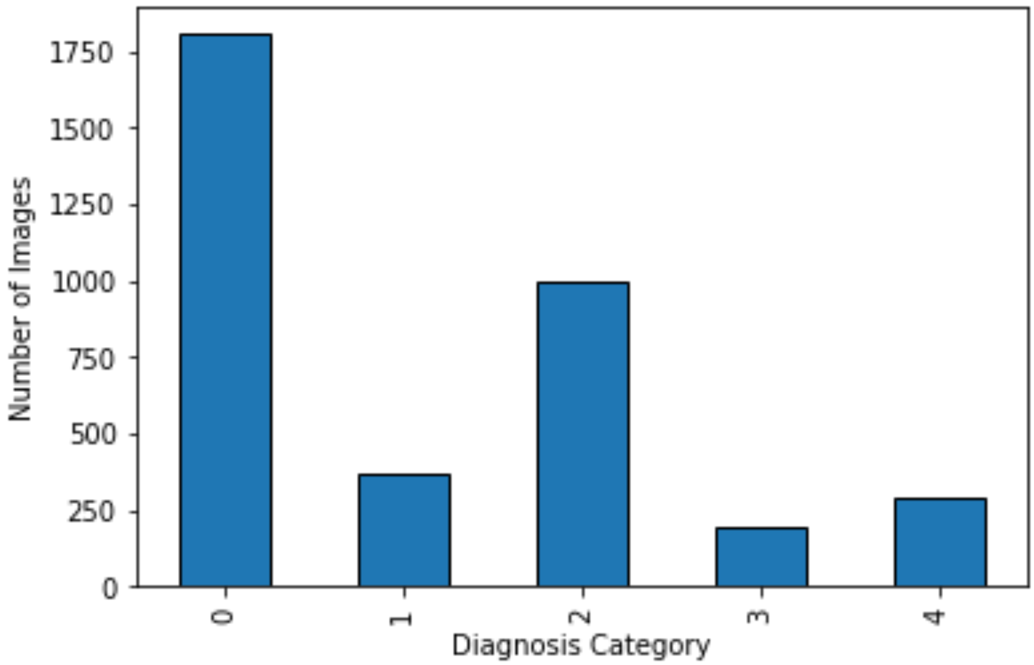
0 - No Diabetic Retinopathy (DR)
1 - Mild
2 - Moderate
3 - Severe
4 - Proliferative DR

Dataset: Asia Pacific Tele-Ophthalmology Society (APTOS) and Aravind Eye Hospital, India provides the dataset, through their competition hosted on Kaggle. The link to the competition is: https://www.kaggle.com/c/aptos2019-blindness-detection/overview

Data Wrangling: The dataset consists of images acquired by Aravind Eye Hospital technicians from multiple clinics, using different cameras. The dataset consists of 3622 images that have been labelled based on their severity. We first load the csv file that contains the unique ids of the images and their labels into a data frame.

	id_code	diagnosis
0	000c1434d8d7	2
1	001639a390f0	4
2	0024cdab0c1e	1
3	002c21358ce6	0
4	005b95c28852	0
5	0083ee8054ee	4
6	0097f532ac9f	0
7	00a8624548a9	2
8	00b74780d31d	2
9	00cb6555d108	1

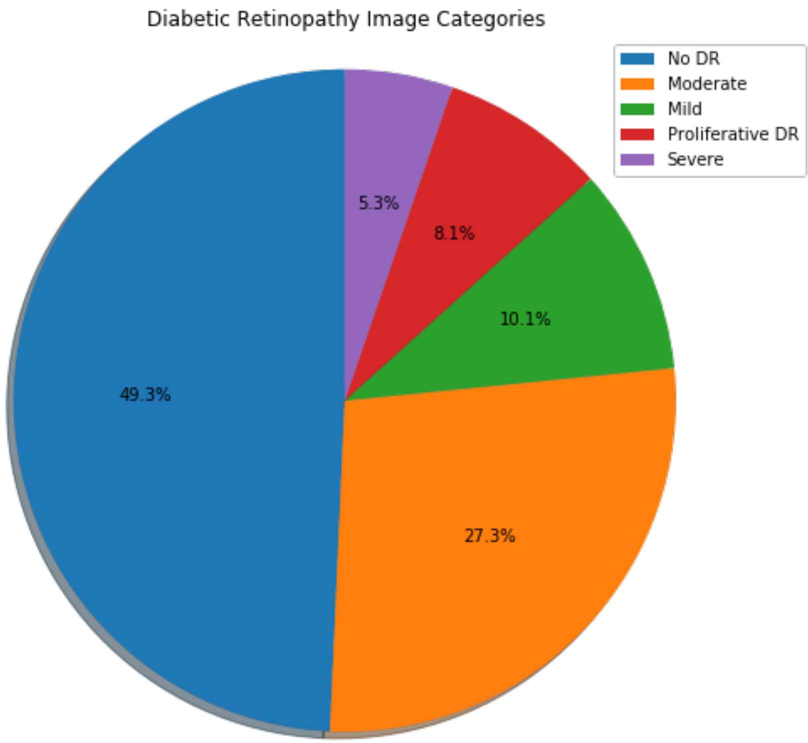
Dataset overview



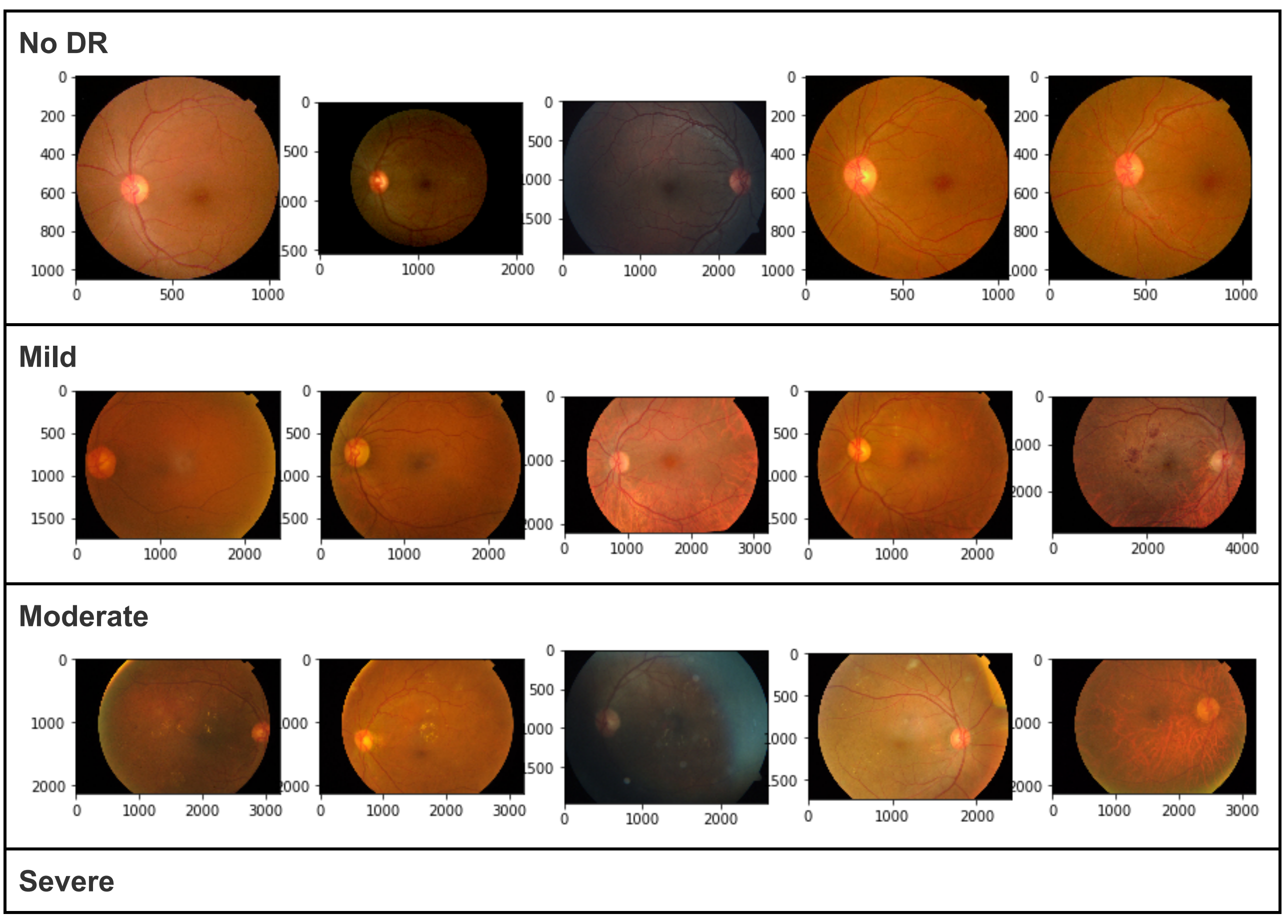
Histogram showing distribution of data

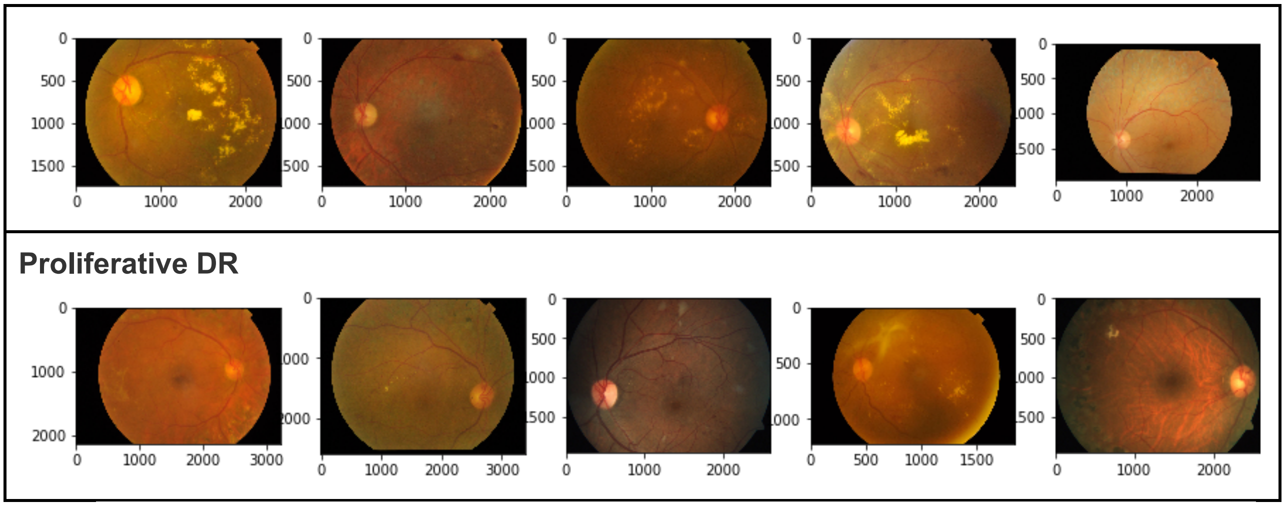
Total number of images	3622
No DR	1805
Mild	370
Moderate	999
Severe	193
Proliferative DR	295

The dataset is very imbalanced. Majority of the data belongs to the No DR category, followed by Moderate, Mild, Proliferative and Severe categories respectively. The pie chart below, shows a good representation of the data distribution.



Pie chart showing distribution of data





Fundus images from every category in the dataset

Building the dataset: To build deep learning models, we require training and test data. We divide the dataset into 70:30 split for training and testing, respectively. 10% of the training data will be used for validation during model training.

Number of train images= 2563
Number of test images= 1099

These images were generated from multiple clinics. The data collection process is not completely standardized. Thus, variation in image size is expected. We load the training set of images and get some summary statistics on the size of the training set.

Min Dimensions: [480, 640, 3]
Max Dimensions: [2848, 4288, 3]
Average Dimensions: [1528.82676551, 2018.98556379,   3]
Median Dimensions: [1536, 2144, 3]

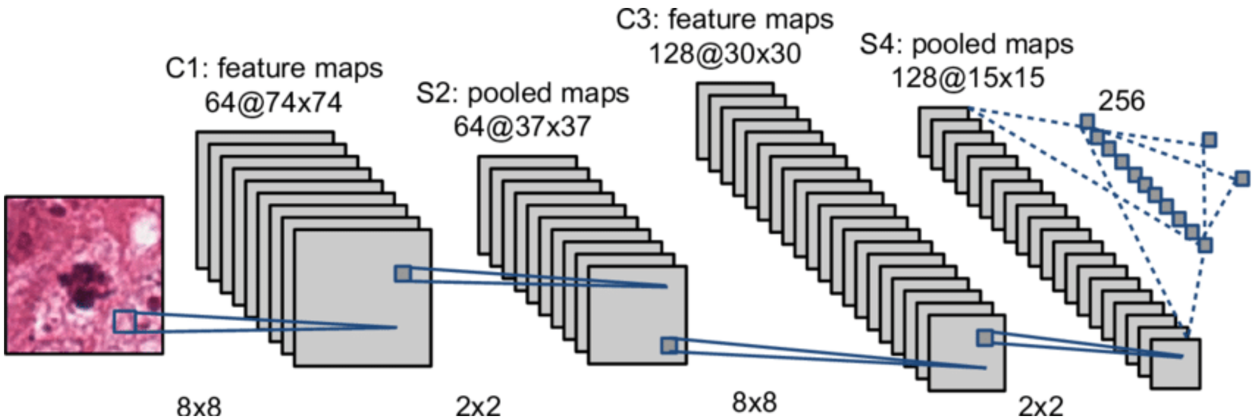
We resize all images (train and test) to standard size [224, 224,3]. Resizing images to these standard dimensions make it easy to be used for our deep learning network and also any pre-trained network that we will use during transfer learning. We then normalize the dataset by dividing it by 255.

Deep Learning Modeling and Optimization:

In this section, we will build three deep learning models and train them using our training dataset. As mentioned before, 10% of the training dataset will be used for validation. These models will then be evaluated using our test dataset.

Model 1: Convolutional Neural Network

A convolutional neural network (CNN) consists of a series of convolution layers, and sub-sampling layers. They are most commonly used to analyze 2D data, such as images. Convolution layers learn patterns from data, such as edges, corners. Each layer learns larger patterns based on the output of the previous layer. CNNs therefore help automate feature engineering. Pooling layers help with down sampling and dimension reduction. A typical CNN architecture can be seen in the image below.

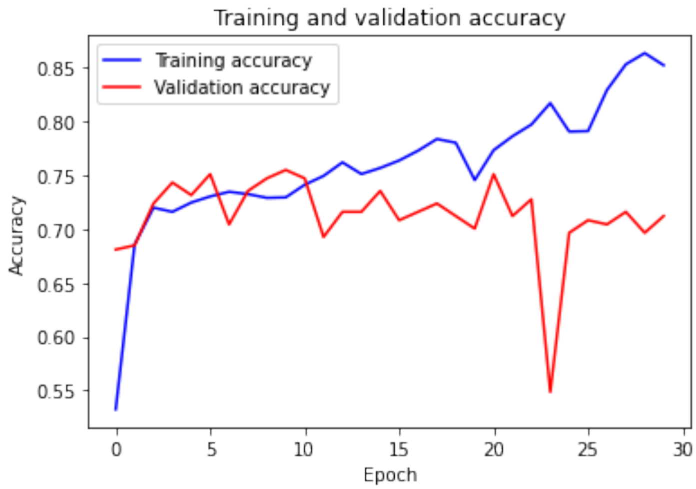


A typical CNN Architecture (3)

The first model that we will implement here is a CNN from scratch. The architecture of our model is shown below. This model contains 3 convolution layers and 3 pooling layers.

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_4 (Conv2D)            (None, 224, 224, 32)      896
_________________________________________________________________
leaky_re_lu_4 (LeakyReLU)    (None, 224, 224, 32)      0
_________________________________________________________________
max_pooling2d_4 (MaxPooling2 (None, 112, 112, 32)      0
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 112, 112, 64)      18496
_________________________________________________________________
leaky_re_lu_5 (LeakyReLU)    (None, 112, 112, 64)      0
_________________________________________________________________
max_pooling2d_5 (MaxPooling2 (None, 56, 56, 64)        0
_________________________________________________________________
conv2d_6 (Conv2D)            (None, 56, 56, 128)       73856
_________________________________________________________________
leaky_re_lu_6 (LeakyReLU)    (None, 56, 56, 128)       0
_________________________________________________________________
max_pooling2d_6 (MaxPooling2 (None, 28, 28, 128)       0
_________________________________________________________________
flatten_2 (Flatten)          (None, 100352)            0
_________________________________________________________________
dense_4 (Dense)              (None, 512)               51380736
_________________________________________________________________
dropout_3 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_5 (Dense)              (None, 512)               262656
_________________________________________________________________
dropout_4 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_6 (Dense)              (None, 5)                 2565
=================================================================
```

```
Total params: 51,739,205
Trainable params: 51,739,205
Non-trainable params: 0
```



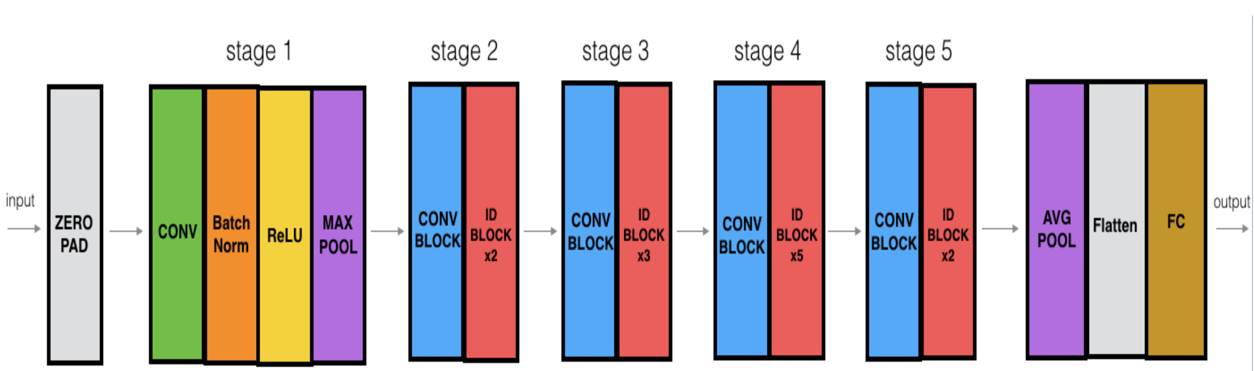
Learning Curves for CNN model

Transfer Learning

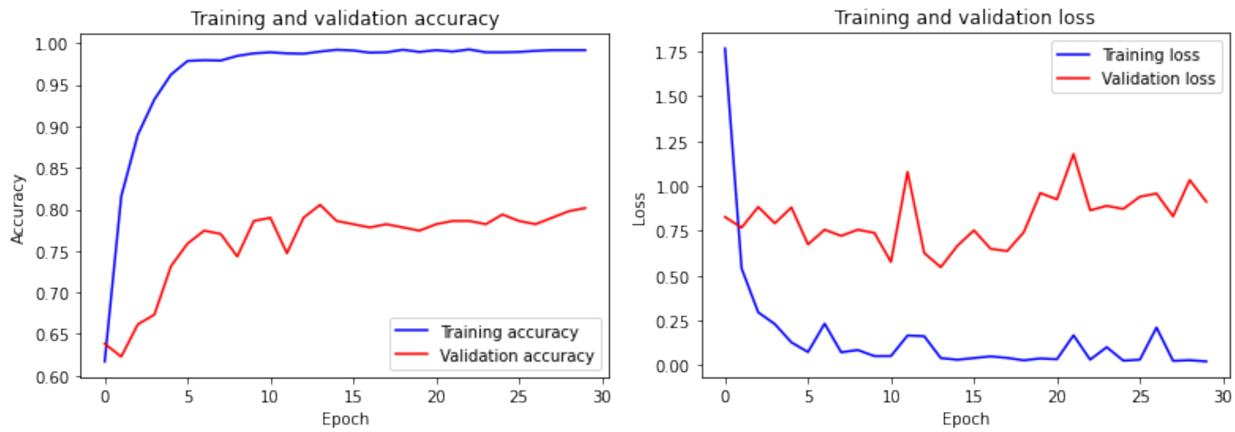
Transfer learning is a process in which a model trained on a particular problem is repurposed to solve a different problem. In this case, we use models that have been trained on the ImageNet data to solve our problem. The ImageNet is a database, hosted by WordNet hierarchy, that contains more than 14million images. In this section, we leverage models that have been pre-trained on this ImageNet dataset for our purpose.

Model 2: ResNet50

ResNet50 stands for Residual Network. The ResNet-50 model consists of 5 stages each with a convolution and Identity block. Each convolution block has 3 convolution layers and each identity block also has 3 convolution layers. The ResNet-50 has over 23 million trainable parameters. We use the keras open source library to load this model. We plug in additional dense layers to perform the classification task.



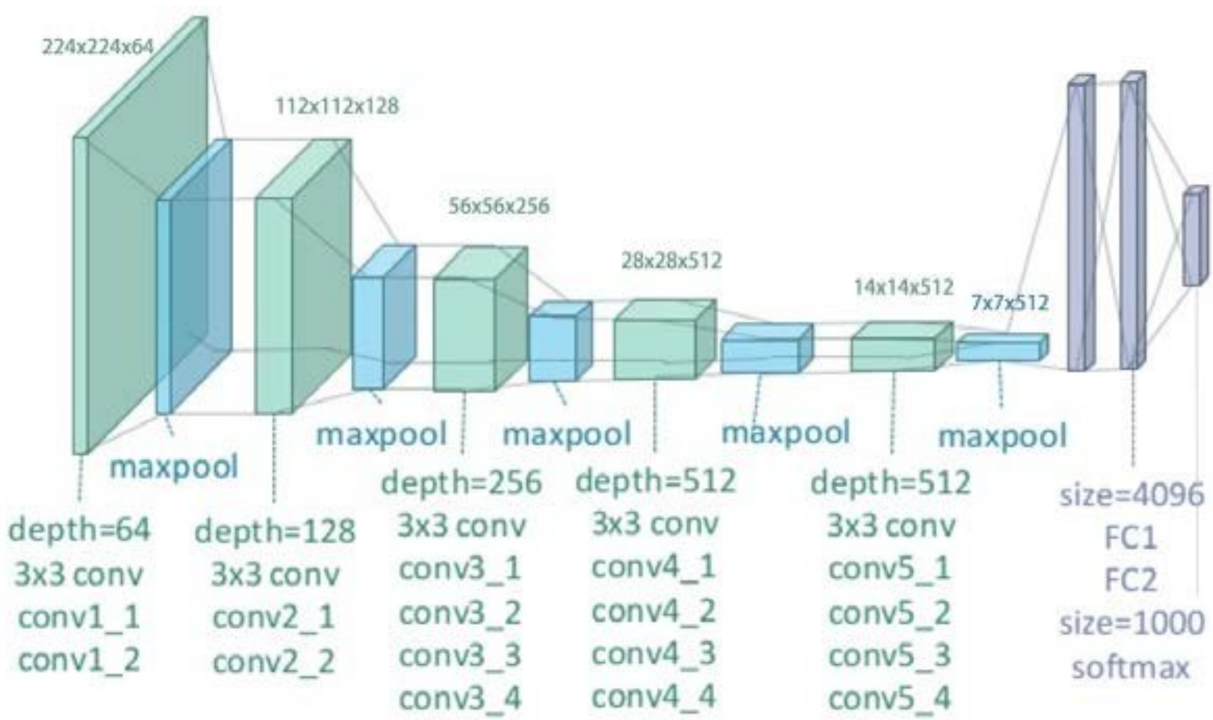
ResNet-50 Model (4)



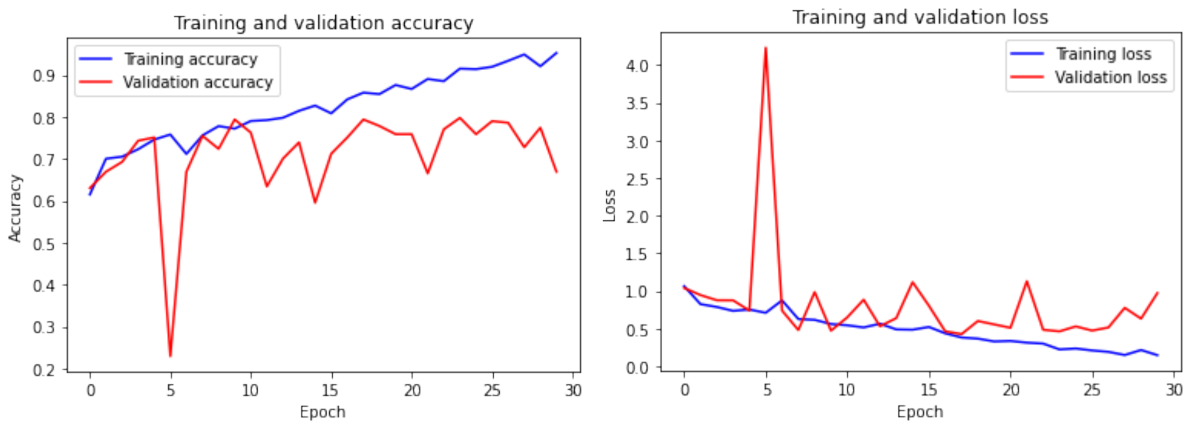
Learning Curves for Resnet-50 model

Model 3: VGG-19

The VGG-19 is a 19 layer neural network developed by Karen Simonyan and Andrew Zisserman from Visual Geometry Group, University of Oxford. We use Keras to load the VGG-19 model and freeze the convolution blocks, so that it can be used as an image feature extractor. We then add our dense layers at the end to perform classification.



VGG-19 Model Architecture [5]



Learning Curves for VGG-19 model

Model Evaluation:

In this section, we evaluate the performance of the model on the test dataset. The models were built on the training dataset.

Models	**Accuracy**	**Loss**	**Precision (weighted average)**	**Recall (weighted average)**
CNN	0.70	1.08	0.66	0.70
ResNet50	0.72	2.03	0.66	0.72
VGG-19	0.62	1.55	0.71	0.62

ResNet50 performs the best on the test dataset. It gives an accuracy of 72%. This is followed by the CNN model at 70% and VGG-19 at 62% accuracy. In all the three models the precision was highest in the No DR class. It could be because there were more images in that category. For future work, utilizing advanced preprocessing techniques, such as image augmentation, can help enhance the performance of these model.

Conclusion:

Diabetic Retinopathy detection is not an easy process. It requires trained personnel and regular screening. Availability of trained professionals and screening equipment are a serious concern. However, it is very encouraging to see the potential of deep learning models to achieve some automation in the detection process. Collaboration of clinicians and deep learning experts would be crucial to take this project to the next level and help elevate the standard of care for patients.

References:

1. Diabetes Canada Clinical Practice Guidelines Expert Committee. Diabetes Canada 2018 Clinical Practice Guidelines for the Prevention and Management of Diabetes in Canada. Can J Diabetes. 2018;42(Suppl 1):S1-S325. (http://guidelines.diabetes.ca/cpg/chapter30#sec1)
2. Diabetic Retinopathy, Mayo Clinic, 2018 (https://www.mayoclinic.org/diseases-conditions/diabetic-retinopathy/symptoms-causes/syc-20371611#dialogId8139015)
3. Wang, Haibo & Cruz-Roa, Angel & Basavanhally, Ajay & Gilmore, Hannah & Shih, Natalie & Feldman, Mike & Tomaszewski, John & González, Fabio & Madabhushi, Anant. (2014). Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. Journal of Medical Imaging. 1. 1-8. 10.1117/1.JMI.1.3.034003.
4. Understanding and coding a ResNet in Keras, Priya Diwedi, Towards Data Science
5. Zheng, Yufeng & Yang, Clifford & Merkulov, Aleksey. (2018). Breast cancer screening using convolutional neural network and follow-up digital mammography. 4. 10.1117/12.2304564.)