Diabetic Retinopathy using Computer Vision and Deep Learning

# Overview

The overall intent of this project to automate detection of Diabetic Retinopathy using computer vision and deep learning. Diabetic retinopathy is one of the major cause of blindness in adults. In this project, I plan to exhibit the usage of state of the art Convolutional Neural Network on fundus images for recognising diabetic retinopathy stages.

# Dataset Used

Coloured Retinal images captured using Fundoscope were used. Those images are freely hosted at Kaggle and were provided by eyePACS, a free platform for retinopathy screening. A total of approximately 35,000 images are being used for training, testing and validation purpose. Images are divided into four classes namely ‘normal’, ‘mild’, ‘moderate’, ‘severe’, and ‘end stage’.

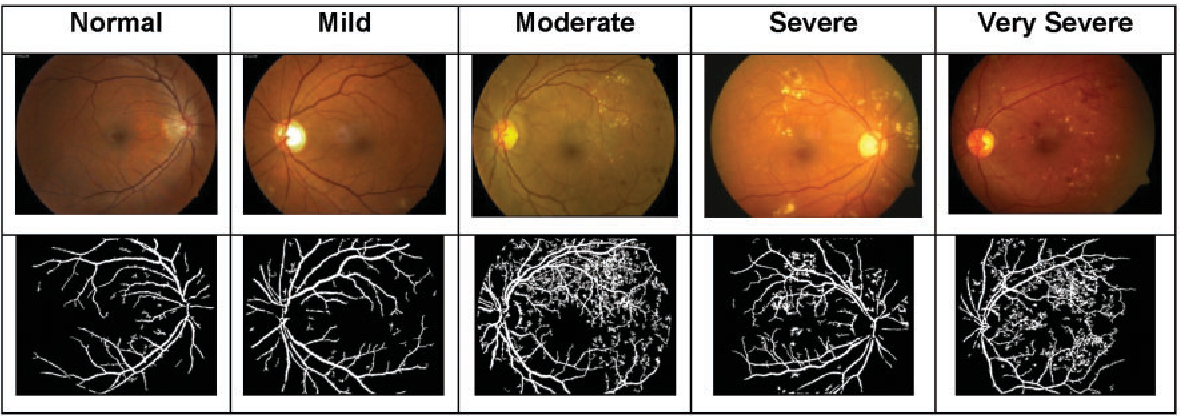


Figure Different stages of Diabetic Retinopathy

# Proposed Architecture

In order to find the most suitable deep learning architecture, various CNNs including the state of the art CNNs will be explored with major focus on VGGNet and ResNET

## VGGNet

VGGNet is made up of 16 convolutional layers. It is one of the most preferred choice for extracting features from images. Weight matrix of VGGNet is publicaly available and has been used in many other applications as a baseline feature extractor. VGGNet consists of 138 millin parameters, which is a bit tricky to handle



Figure VGGNet architecture

## RESNet

Residual Neural Network also known as RESNet was introduced by Kaiming He et al. It consists of skip connections and heavy batch normalisation. Such skip connections are called gated units or gated recurrent units resembling strongly to RNNs. RESNet consists of heavy 152 convolutional layers while still less complexity than VGGNet.

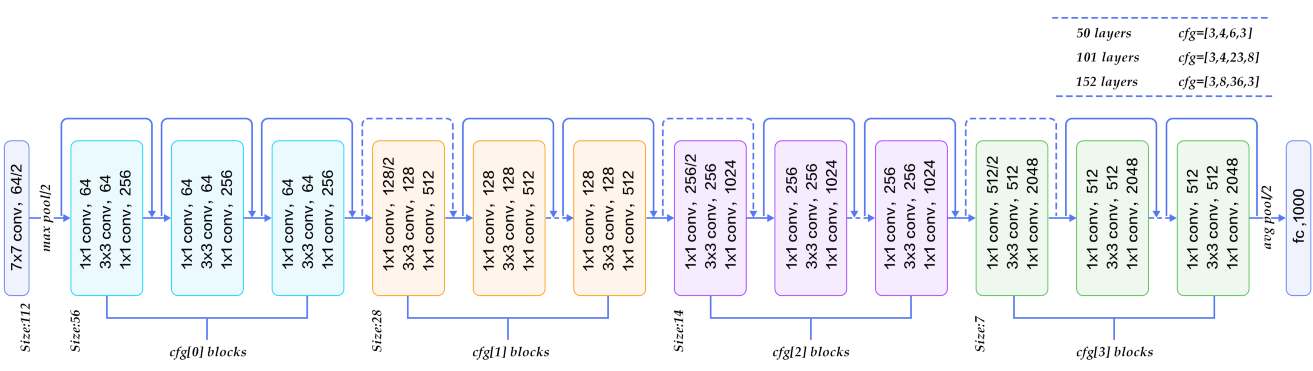


Figure RESNet architecture

# Transfer Learning

While using state of the art deep learning architectures, transfer learning will be tried out. For that, the last fully connected layer would be removed, then a transfer learning scenario would be followed by treating the remaining network components as a fixed feature extractor for the new dataset. The transfer learning retains initial pre-trained model weights and uses final network layer to extract image features.

# Data Augmentation

Data set which is planned to be used is highly imbalanced. To correct it, data augmentation techniques will be used to generate additional data for the classes that lack sufficient number of images. Ideally after data augmentation is performed, all the classes will have equal number of images approximately. For data augmentation Cropping, Horizontal, Vertical flips, Translation and Rotation will be used.

# Results Evaluation

Since this is a classification based problem so Confusion Matrix will be used to generate the results. As the dataset is not fully balanced so instead of accuracy, precision, recall and f-score will be calculated for comparing different architectures.