# DIABETIC RETINOPATHY DETECTION

# Abstract

Diabetic retinopathy is a complication of diabetes caused by high blood sugar levels damaging the back of the eye (retina). It can cause blindness if left undiagnosed and untreated. Millions of people suffer from diabetic retinopathy. In fact, diabetic retinopathy is the most prevalent cause of avoidable vision impairment, mainly affecting the world's working-age population. Today, obtaining reliable, inexpensive screening and detection of diabetic retinopathy is still an unattained dream. For people in rural areas, this diagnosis only gets more unattainable.

The recent technological advances in computing power, communication systems, and machine learning techniques have provided us with an opportunity. Our goal is to scale efforts through technology, to leverage CNNs and a dataset that contains numerous images of the retina provided by fundus photography to detect the severity of diabetic retinopathy and diabetic macular edema. ranging from 0 to proliferative diabetic retinopathy and diabetic macular edema. This in turn will help us gain the ability to automatically screen images for the disease and provide quicker and more affordable healthcare.

Thus, we see great potential in this project and its prospective application to change the diagnosis of diabetic retinopathy for the better vastly.

# Table of contents

# Chapter 1: Project details

## Goal

Goal statement: Use CNN's for detection and analysis of severity of diabetic retinopathy to enable more affordable and quicker diagnosis.

The problem: Today diabetic retinopathy is detected through a long and expensive process. First your pupils are dilated by dropping a dilation liquid into your eyes which will blur your vision for 3-4 hours. While your eyes are dilated a doctor will then inspect them for abnormalities. The next process is fluorescein angiography, where, while your eyes are still dilated a dye is injected into a vein in your arm and pictures are taken as the dye circulates through your eyes' blood vessels. The images can pinpoint blood vessels that are closed, broken or leaking. Finally, an optical coherence tomography scan is conducted which provides cross-sectional images of the retina that show the thickness of the retina. This will help determine how much fluid, if any, has leaked into retinal tissue. This diagnosis process is long and can cost upto an astonishing 70000 USD.

How our project helps: Computer-aided disease diagnosis for Diabetic Retinopathy could ease mass screening of the patients with the disease and help clinicians in utilizing their time more efficiently, thus providing faster and more efficient, life-improving service to more patients in need. Our model can

## Data

For our project we were provided with a dataset consisting of images of the retina. This dataset was available as a part of "Diabetic Retinopathy: Segmentation and Grading Challenge" organised in conjunction with IEEE International Symposium on Biomedical Imaging (ISBI-2018), Washington D.C. A clinician rated each image in the dataset for the severity of Diabetic retinopathy and Diabetic Macular Edema.

This dataset was split in a 70-30 split to first train our diabetic retinopathy detection model and then the 30% of the images in the dataset were used to test the model and verify if it has sufficient accuracy. To ensure our model was trained on the highest quality of data to produce maximum reliability and accuracy, the images and labels in the dataset had noise imposed on them, the images may contain artifacts, be underexposed or overexposed or even be out of focus.

# Project flow breakdown

**Importing and tabulating data**

- To import the data, we mounted our google drives on the Google colab notebook, and dropped the columns that contained null values, as can be seen in the image below.

| | Image name | Retinopathy grade | Risk of macular edema | Unnamed: 3 | Unnamed: 4 | Unnamed: 5 | Unnamed: 6 | Unnamed: 7 |
|---|---|---|---|---|---|---|---|---|
| 0 | IDRiD_001 | 3 | 2 | NaN | NaN | NaN | NaN | NaN |
| 1 | IDRiD_002 | 3 | 2 | NaN | NaN | NaN | NaN | NaN |
| 2 | IDRiD_003 | 2 | 2 | NaN | NaN | NaN | NaN | NaN |
| 3 | IDRiD_004 | 3 | 2 | NaN | NaN | NaN | NaN | NaN |
| 4 | IDRiD_005 | 4 | 0 | NaN | NaN | NaN | NaN | NaN |

| | Image name | Retinopathy grade | Risk of macular edema |
|---|---|---|---|
| 0 | IDRiD_001 | 3 | 2 |
| 1 | IDRiD_002 | 3 | 2 |
| 2 | IDRiD_003 | 2 | 2 |
| 3 | IDRiD_004 | 3 | 2 |
| 4 | IDRiD_005 | 4 | 0 |

**Understanding the data**

- To understand the data fully, we performed multiple operation on the dataset, to get the count of values in each column, their datatype (Fig1), the mean, standard deviation, minimum, maximum, 1st, 2nd, and 3rd quartiles, for each column(Fig2).

```
0   Image name              413 non-null    object
1   Retinopathy grade       413 non-null    int64
2   Risk of macular edema   413 non-null    int64
```

Fig1

| | Retinopathy grade | Risk of macular edema |
|---|---|---|
| count | 413.000000 | 413.000000 |
| mean | 1.719128 | 1.043584 |
| std | 1.387723 | 0.949215 |
| min | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 |
| 50% | 2.000000 | 1.000000 |
| 75% | 3.000000 | 2.000000 |
| max | 4.000000 | 2.000000 |

Fig2

- To understand the distribution of the Risk of macular edema, and retinopathy grade, two probability distribution graphs were also created (Fig3 and Fig 4). In Fig3 we can infer that in our dataset, the frequency of people having a retinopathy grade as 0 or 2 is maximum, while the RG 1 is the rarest. In Fig 4 as well we can see that the data is sort of unbalanced as images with RME equal to 1 are relatively less than the 0 or 2.

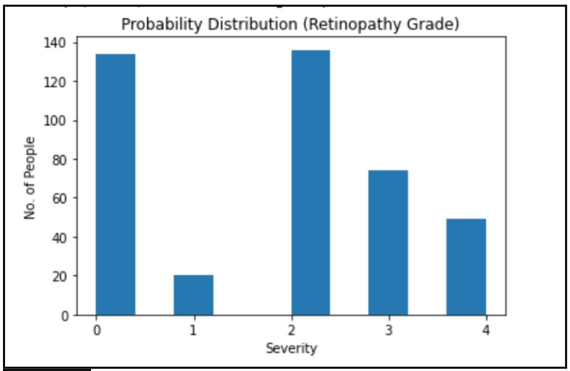


Fig3

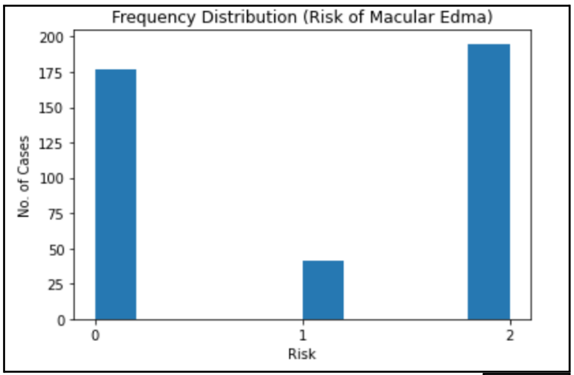


Fig4

## Data Augmentation

- To make sure that the predictions are not affected by slight rotations and noise in our data, we created additional data for our model to train on, which was slightly rotated or horizontally flipped using the ImageDataGenerator. We used preprocess_input from inceptionResnet.

## Contrasting Our Image

- To ensure that the model is able to gather information in an efficient way, we tried to make the image more detailed and clearer using, by contrasting the image using CLAHE. For this we defined a method that converts the original image (Fig5) to the contrasted image which has more prominent veins for instance .(Fig6) (Reference:Stackoverflow for the syntax)

Fig6

Fig5

## Creating 2 models for Retinopathy Grade:

**Model 1: Customised Model Using 18 Different Layers *(Total Trainable Parameters:5,116,467)***

- The model contains 18 layers that includes:
  - 5 pairs of Conv2D and Max Pooling with no. of filters equal to (600, 450,328,164 and64)
  - Following that we have a Flatten layer.
  - 3 layers of Dense and dropouts
  - Lastly, to classify our image in 5 different classes, a dense layer.

**Model 2: Transfer Learning Model Using InceptionResnet**

- Appending 4 layers in the end, to make the model more suitable for us:
  - GlobalAveragePooling2D layer
  - 3 Dense Layers with 1024, 512 and 5 neurons respectively

**<u>Creating 2 models to find the Risk of Macular Edema:</u>**

**Model 1: Customised Model Using 18 Different Layers**

- The model contains 18 layers that includes:
  - 5 pairs of Conv2D and Max Pooling with no. of filters equal to (600,450,328, 164 and 64)
  - Following that we have a Flatten layer.
  - 3 layers of Dense and dropouts
  - Lastly, to classify our image in 3 different classes, a dense layer.

**Model 2: Transfer Learning Model Using InceptionResnet**

- Appending 4 layers in the end, to make the model more suitable for us:
  - GlobalAveragePooling2D layer
  - 3 Dense Layers with 1024, 512 and 3 neurons respectively

**<u>Accuracy and Loss</u>**

We got a roughly 50% accuracy for our custom model and 70-75% accuracy with Inception Resnet with loss ranging from 0.5-1.8

# Conclusions

We used two models for each of the learning, one a custom model as defined above, this is less complex, as it has less layers, and less trainable parameters and the other was a transfer learning model that used InceptionResnet as its base model. Upon training the model, our accuracy for the customised model varied between 45-55% whereas with the transfer learning model it varied from 70% to 80%. The accuracy increased when we used a transfer learning model, and we assume it happened so because it is more complex, it has more layers and trainable parameters.

We experimented with our learning rates and the number of epochs for each of our models, to see what gives us the maximum accuracy, and we have come to the conclusion that for our custom model a learning rate of 0.0001 and epoch equal to 30 worked the best whereas for the Transfer learning model, out of 0.001, 0.0001, and 0.0002, 0.0001 gave us the maximum accuracy at 35-40 epochs.

# Further Enhancements

1. **Improving the contrasting of the data**

The contrasting that we did on the data to understand our data better, could have been applied to each image in our dataset, so that our model would be able to learn the different features of each image in a better way.

2. **Increasing the number of layers in our customised model**

By increasing the number of layers in the customised model, larger and more complex combinations of patterns can be captured. This allows our CNN to learn compositional functions that it was not able to learn before. Hence the vector would have accumulated more data from each pixel of the image. More data gathered would improve the model.

Experiment with different base models for our transfer learning model.

3. **Passing the data through more augmentations.**

Image data augmentation is used to expand the training dataset in order to improve the performance and ability of the model to generalize. Image data augmentation is supported in the Keras deep learning library via the ImageDataGenerator class. Data augmentation helps improve the quality of the dataset by augmenting pre-existing data in the dataset such that the model recognises the data as new data hence while training the model the diversified dataset improves robustness and the quality of the model, enabling greater accuracy.

4. **Testing the model with Adversarial Attacks.**

An adversarial attack consists of subtly modifying an original image in such a way that the changes are almost undetectable to the human eye. The modified image is called an adversarial image, and when submitted to a classifier is misclassified, while the original one is correctly classified. Adversarial attacks improve the model by making the model more “robust” hece allowing it to have higher accuracy and confidence in outputting the correct outputs and classification when supplied by a varied dataset making it more useful and applicable in the real world.

5. **Experimenting with different base models for our transfer learning model**

For our transfer learning model we have used InceptionResnet as our base model, using which we have done predictions. However, to increase the accuracy and reduce the loss, we could have created a few more models that used other base models such as MobileNet50, Resnet50 etc.

6. **Saliency Map**

To understand how our model works, we can use saliency maps to see what part of the image the model focussed on in order to get a better accuracy.

# References


https://bimedis.com/search/search-items/ophthalmic-equipment-optical-coherence-tomography

https://www.mayoclinic.org/diseases-conditions/diabetic-retinopathy/diagnosis-treatment/drc-20371617