# Diabetic Retinopathy Detection Using Machine Learning

*By Pratye Aggarwal*

## Introduction

Diabetic Retinopathy is a well-known eye problem that results in vision loss and, in adverse cases, complete blindness. The complication is caused by high blood sugar levels and is associated with long-standing diabetes. Estimates done by the World Health Organisation (WHO) showed that over 347 million people have this disease worldwide, and these numbers will rise by 11-20% in the upcoming years.

A few of the most common symptoms of Diabetic Retinopathy include blurred vision, difficulty in perceiving colours and having dark areas in vision. Detection and the proper treatment of the disease in the early stages are crucial to prevent long-term vision impairment issues.

## Proposed Solution

Detection of Diabetic Retinopathy requires highly experienced doctors. Standard tests like pupil dilation and visual acuity test often take time in the evaluation and detection process. The unavoidable delays lead to unnoticed changes that occur during that time. Moreover, the mass screening of the population with this disease cause more delays and difficulty in achieving reliable detections.

Machine Learning could help ease this issue. With the help of a model, we can easily detect the disease and its severity. The model will provide more accurate detections and will reduce the tedious job of evaluation. This will help the doctors in giving more accurate and quick diagnoses to their patients. The time and money that the patients need to invest each month will be drastically reduced along with their suffering.

## Dataset And Augmentation

### Dataset

I used the publicly available Kaggle Dataset of Diabetic Retinopathy Detection (link: https://www.kaggle.com/c/diabetic-retinopathy-detection) to train my Machine Learning Model. The database contains a large set of high-resolution retina images that were taken under different imaging conditions. The images are labelled into five categories: (1) No DR; (2) Mild; (3) Moderate; (4) Severe; (5) Proliferative DR. The images were labelled by a clinician, which makes the dataset ideal for training a model that is to replace the work of conventional evaluations. In total, there were 35126 labelled images in the dataset. The breakdown of the number of images in each class is as follows:

| S. No. | Label | Number of Images |
|---|---|---|
| 1 | No DR | 25,810 |
| 2 | Moderate | 5,292 |
| 3 | Mild | 2,443 |
| 4 | Severe | 873 |
| 5 | Proliferative DR | 708 |

## Augmentation and Preprocessing

The dataset provided a good amount of images for training. Yet, due to the high-Class Imbalance issue, the dataset was not ideal for training a good model. Class Imbalance directly affects the performance of our model. Due to the significant difference in-class examples, the model tends to gain biases with more training examples. For training a good model, it is necessary to maintain class balance.

To encounter this problem, I made a couple of changes to my dataset by randomly taking the same number of training examples like the one with the least training examples. By doing so, I ended up with 3,540 images, with 708 images in each class.

The process of class balancing lead to a tremendous amount of data loss. For accurate prediction of the different severities of Diabetic Retinopathy required more data. For this, I performed a couple of augmentation techniques on the images to create more training samples. I used an online tool called Roboflow (website: https://roboflow.com), which provides different options for augmentation and preprocessing. I resized my images into 1080x720 and added the following augmentations on the images:

| S. No. | Augmentation | Type and Amount |
|---|---|---|
| 1 | Flip | Horizontal and Vertical |
| 2 | Hue | Between -1° and +1° |
| 3 | Saturation | Between -5% and +5% |
| 4 | Brightness | Between -15% and +15% |
| 5 | Exposure | Between -2% and +2% |
| 6 | Blur | Up to 0.25px |
| 7 | Noise | Up to 1% of pixels |

The augmentations outputted three images for every training example. This produced a total of 10,620 images, with 2,124 images in each class. The data was split into train, test and validation by AutoML Vision. Following is its breakdown:

| Data Set | Total Number of Images | Total Number of Images Per Class |
|---|---|---|
| Train | 8495 | 1699 |
| Validation | 1065 | 213 |
| Test | 1060 | 212 |

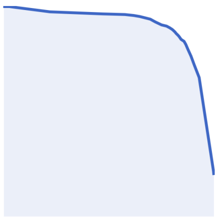
# Model

The model was initially set for training for 240 node hours, as recommended by AutoML Vision. But due to some issues with the Free-Trial account, my model trained for over 200 node hours only. Yet, it could still achieve an average precision of 90.9 % (0.909 as shown on the console) and a recall of 83.21%.

# Graphs and Stats

## Overall Stats:

EyePACS_2



**Average precision**

0.909

| | |
|---|---|
| Precision* | 86.13% |
| Recall* | 83.21% |

* Using a score threshold of 0.5

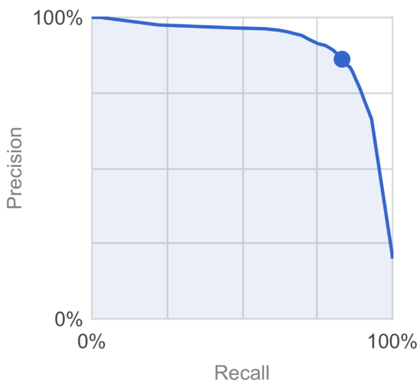
| | |
|---|---|
| Model ID | ICN1634470214177390592 |
| Created | 11 Nov 2021, 20:19:05 |
| Base model | None |
| Data | 10,620 images |
| Model type | Cloud |
| Train cost | 198.993 node hours |
| Deployment state | Not deployed |

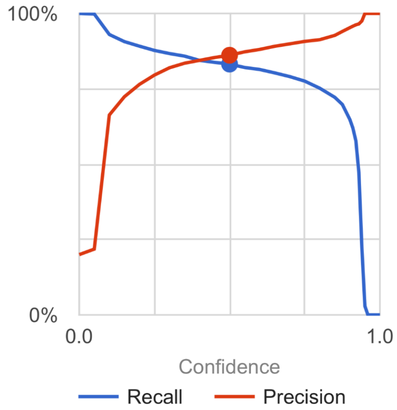
## Precision and Recall:

| | |
|---|---|
| **Total images** | 9,560 |
| **Test items** | 1,060 |
| **Precision** | 86.13% |
| **Recall** | 83.21% |

| | |
|---|---|
| All labels | 0.90941 |
| Mild | 0.86764 |
| Moderate | 0.88387 |
| No DR | 0.89130 |
| Proliferative DR | 0.97478 |
| Severe | 0.93308 |

# Precision and Recall Graphs:




# Confusion Matrix:

| True Label \ Predicted Label | Proliferative DR | No DR | Mild | Severe | Moderate |
|---|---|---|---|---|---|
| Proliferative DR | 93% | - | 0% | 5% | 1% |
| No DR | 2% | 82% | 11% | 0% | 5% |
| Mild | 1% | 8% | 80% | 1% | 9% |
| Severe | 5% | 2% | 1% | 85% | 6% |
| Moderate | 3% | 3% | 7% | 4% | 83% |

| True Label \ Predicted Label | Proliferative DR | No DR | Mild | Severe | Moderate |
|---|---|---|---|---|---|
| Proliferative DR | 198 | - | 1 | 10 | 3 |
| No DR | 4 | 173 | 24 | 1 | 10 |
| Mild | 2 | 17 | 170 | 3 | 20 |
| Severe | 11 | 5 | 3 | 180 | 13 |
| Moderate | 6 | 7 | 15 | 9 | 175 |