Diabetic Retinopathy Detection

A P Vaideeswaran

IIT Hyderabad

27 October 2023

Domain/Area of the Project

Computer Vision in Healthcare

Motivation/Relevance of the Project

Diabetic retinopathy (DR) is the leading cause of blindness in the working-age population. This project aims to detect and classify DR into four stages if the disease exists. This might help people catch the condition and give a reasonable estimation of the severity even before a visit to the doctor. This is especially important as the disease barely presents any symptoms in the early stages. Several people can seek early medical attention, thereby preventing loss of eyesight or slowing down the disease. Doctors can also use this model to keep track of the patient's condition without even a visit. DR detection is time-consuming and requires a lot of expertise and infrastructure. Therefore, this project would help doctors and patients alike, giving the health sector time to improve the infrastructure required to cater to the needs of an increasing diabetic population, especially in rural areas.

Description of The Dataset

The dataset includes many high-resolution retina images taken under various imaging conditions. A left and right field is provided for every subject. Images are labeled with a subject ID and either left or right. For instance, `1_left.jpeg` is the left eye of patient ID 1.
In this labeled dataset, a clinician has rated the presence of the disease in each image on a scale of 0 to 4, according to the following scale: 0 - No DR 1 - Mild 2 - Moderate 3 - Severe 4 - Proliferative DR.

Description of The Dataset

The images in the dataset come from different models and types of cameras, which can affect the visual appearance of left vs. right. Some images are shown as one would see the retina anatomically (macula on the left, optic nerve on the right for the right eye). Others are shown as one would see through a microscope condensing lens (i.e., inverted, as one sees in a typical live eye exam). An image can be told to be inverted using the following methods: It is inverted if the macula (the small dark central area) is slightly higher than the midline through the optic nerve. If the macula is lower than the midline of the optic nerve, it's not inverted. If there is a notch on the side of the image (square, triangle, or circle), it's not inverted. If there is no notch, it's inverted. The dataset can be downloaded from the following link: `https://www.kaggle.com/competitions/diabetic-retinopathy-detection/data`

Specific Objectives/Goals of the Project

Given the image of the eye, the task is to classify it amongst five classes (stage 0 - NO disease to stage 4 - proliferative DR). The model should be resilient to noise, artifacts, lack of focus, and inappropriate exposure.

Challenges anticipated

- Effects of noise and artifacts, lack of focus, and inappropriate exposures.
- We have a large dataset comprising several large images (upwards of 5000x5000). The aspect ratios of the images are different as the images might have been taken using different instruments.
- Pre-processing of each image, possibly involving downsampling, must be done. However, this should be done so that the key indicators of the condition, like microaneurysms, are NOT lost.
- In this case, there is a notion of distance between classes. For instance, say the model incorrectly classifies stage 1 as stage 2; it is an incorrect classification. However, if stage 1 is classified as stage 4, it is also incorrect but, in some sense, MORE incorrect. This can be quantified using the quadratic weighted Kappa metric.
- Class imbalance is also expected as there are several images of the early stages of the disease compared to stage 4.

How do you expect to meet those challenges?

- While pre-processing, methods like histogram stretching and Gaussian blurring might improve image quality.
- FC-DenseNet can be used to identify the microaneurysms. The cross-entropy loss function of DenseNet can be replaced with the focal loss to weigh the aneurysms more than the background itself.
- A loss function that considers the distance between the classes has to be chosen.
- To deal with class imbalance (especially for stage 4), data augmentation, self-supervised learning, and focal loss can be employed.
- Multiple cores should be employed to deal with huge datasets efficiently. The following article was used for some initial insights on doing pre-processing: `https://www.nature.com/articles/s41598-021-04750-2`

Initial Approach to the Project

- Adaptive cropping with a certain threshold has been performed on the images. In other words, if a column has more than 95
- This process took 0.5 seconds/image. Cropping of all the images utilizing four cores and the multiprocessing Python Library took around 1.5 hours.
- Subsequently, amongst the state-of-the-art classification models like VGG16, Resnet-50, and EfficientNet 3 and 4, the best model for this project is to be identified.
- The ideation phase involves retaining the CNNs for all the models while removing the ANNs. We try a single hidden layer ANN with 16 nodes for each model. The models shall then be trained on the 200x200 images for around 50 epochs. If the performance is satisfactory, we try 32 nodes and possibly even 64 nodes.
- The top two models shall further experiment with a second ANN layer, and the performance shall be re-evaluated on the 200x200 images. This would help identify the golden bullet model for this project.

Initial Approach to the Project

- This model shall then be trained on 1000x1000 images. The idea behind using 200x200 images during the ideation phase is to save lots of time. Any model would take significantly more time on the 1000x1000 images than the 200x200 images, slowing down the ideation phase significantly. Spending lots of time on a model without guaranteed performance would be wasteful.
- If the model performance is sufficiently good, more advanced methods can be used that effectively find the microaneurysms. In other words, we could fine-tune the model to boost its performance.
- On the other hand, hyper-parameter tuning can be performed if the model performance is NOT sufficiently good. The model could be trained on a larger number of epochs, or the width and depth of the ANN could be increased.