Udacity Machine Learning Nanodegree 2019

Capstone Proposal

Diabetic Retinopathy Detection

Identify signs of diabetic retinopathy in eye images

Domain Background

Diabetic Retinopathy is the leading cause of blindness in the working-age population of the developed world. It is estimated to affect over 93 million people.

The US Center for Disease Control and Prevention estimates that 29.1 million people in the US have diabetes and the World Health Organization estimates that 347 million people have the disease worldwide. Diabetic Retinopathy (DR) is an eye disease associated with long-standing diabetes. Around 40% to 45% of Americans with diabetes have some stage of the disease. Progression to vision impairment can be slowed or averted if DR is detected in time, however this can be difficult as the disease often shows few symptoms until it is too late to provide effective treatment.

Currently, detecting DR is a time-consuming and manual process that requires a trained clinician to examine and evaluate digital color fundus photographs of the retina. By the time human readers submit their reviews, often a day or two later, the delayed results lead to lost follow up, miscommunication, and delayed treatment.

Clinicians can identify DR by the presence of lesions associated with the vascular abnormalities caused by the disease. While this approach is effective, its resource demands are high. The expertise and equipment required are often lacking in areas where the rate of diabetes in local populations is high and DR detection is most needed. As the number of individuals with diabetes continues to grow, the infrastructure needed to prevent blindness due to DR will become even more insufficient.

After doing the Dog Breed classifier I started loving this Computer Vision area specifically, it’s quite intriguing how these amazing Deep Neural Network can learn so much with just forward feed, gradient descent and back propagation. I clearly see this as my area of interest and would also like to grow in this particular field of computer vision.

Problem Statement

In recent times, most of the problems of computer vision have been solved with greater accuracy with the help modern deep learning algorithms, Convolutional Neural Networks (CNNs) being an example. CNNs have been proven to be revolutionary in different fields of computer vision such as object detection and tracking, image and medical disease classification and localization, pedestrian detection, action recognition, etc. The key attribute of the CNN is that it extracted features in task dependent and automated way.

The main objective of this project will be to use Deep Learning technique like CNN to automatically grade the severity of Diabetic Retinopathy by classifying images based on disease pathologies from four severity levels.

When given an input as Image of the retina the model should be able to classify it as one of the four levels of Diabetic Retinopathy

Datasets and Inputs

Dataset is obtained from Kaggle which has a large set of high-resolution retina images taken under a variety of imaging conditions. A left and right field is provided for every subject. Images are labeled with a subject id as well as either left or right (e.g. 1_left.jpeg is the left eye of patient id 1).

A clinician has rated the presence of diabetic retinopathy 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

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). There are generally two ways to tell if an image is inverted:

- 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) then it's not inverted. If there is no notch, it's inverted.

Like any real-world data set, you will encounter noise in both the images and labels. Images may contain artifacts, be out of focus, underexposed, or overexposed.

There are around 16843 images, images are equally distributed 50% on each left and right eye.

These are high resolution images of typical resolution of 3456 x 2304, 2560 x 1920, 3888 x 2592 etc. Basically the images are not of fixed size and these images are of mixed high resolutions for both eyes.

Total size of images is around 17GB and each image is roughly around 1MB to 2MB.

There is a huge class imbalance in the available dataset, following is the class level data for the distribution

Class	Available data per class
0 - No DR	11367
1 - Mild	1118
2 - Moderate	2405
3 - Severe	382
4 - Proliferative DR	306

Solution Statement

The Proposed solution to this problem is to apply Deep learning techniques that have proved to be highly successful in the field of image classification.

The next step would be to create a model architecture using transfer learning and modifying last couple of layers and for this we can experiment and choose from available COCO models in TensorFlow.

This model then would be trained with these dataset and make predictions. I believe this would be effective in finding pattern within the DR dataset much better using the Transfer learning than writing the model from scratch.

We will use the evaluation metrics described later to compare the performance of these solutions against the benchmark models in the next section.

Benchmark Model

Methods of detecting microaneurysms and grading DR involving k-NN[1], support vector machines[2], and ensemble-based methods[3] have yielded sensitivities and specificities within the 90% range using various feature extraction techniques and preprocessing algorithms.

Previous CNN studies[4] for DR fundus images achieved sensitivities and specificities in the range of 90% for binary classification categories of normal or mild vs moderate or severe on much larger private datasets of 80,000 to 120,000 images.

Inception V3 model on Kaggle[5] having precision 68% and recall 87%

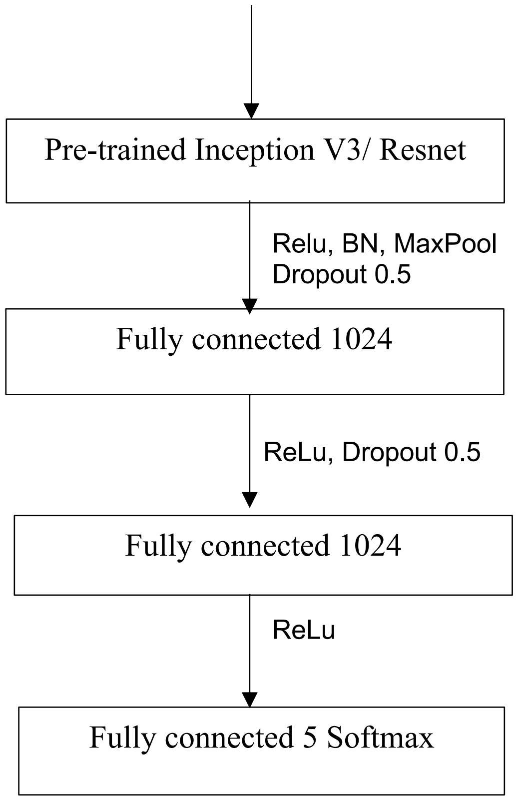
Evaluation Metrics

The evaluation metric for this problem would be Precision, Recall and AUC-ROC curve.

Project Design

I will use Transfer learning based approaches with pretrained models Inception V3/Resnet . The last fully connected layer will be removed, then a transfer learning scenario with 2 fully connected network and the last layer would be fully connected layer with 5 output and Softmax as activation function. The transfer learning retains initial pretrained model weights and extracts image features via a final network layer.

Proposed Architecture



Data Preprocessing

First we need to pre-process the data as the original fundus images may contain a lot of noise, a lot of exposure and lighting variation and would be of different resolutions.

Data Augmentation

The performance of deep neural network is strongly correlated with the size of available training data. Possible data augmentation that could be performed.

- Rotation : Rotate Images randomly between 0° to 360°
- Shearing :- Sheared randomly with some angle
- Flip : Flip Images both horizontally and vertically
- Zoom : Stretched Images randomly
- Crop :- Crop Images randomly to 85 − 95% of the original size
- Translation : Randomly shift images

Data Splitting

Split the data into a training set and validation set with an 80-20 split.

Model training and evaluation

I will start with the simple model architecture first before training and evaluating it. Then iterate this process trying different architectures and hyper-parameters to reach an accuracy score we are happy with.

For this five class problem we define specificity as the number of patients correctly identified as not having DR out of the true total amount not having DR and sensitivity as the number of patients correctly identified as having DR out of the true total amount with DR. We define accuracy as the amount of patients with a correct classification. Looking at the confusion matrix will give a clear picture about the model performance, where specificity playing a vital role for this usecase.

The classifications in the network were de-fined numerically as: 0 - No DR 1 - Mild DR 2 - ModerateDR 3 - Severe DR 4 - Proliferative DR

Reference

1] Quellec G., Lamard M., Josselin P. M., Cazuguel G., Cochener B., Roux C. Optimal wavelet transform for the detection of microaneurysms in retina photographs. IEEE Transactions on Medical Imaging, 2008;27(9):1230–1241. [PMC free article] [PubMed] [Google Scholar]

2] UR A. Decision support system for diabetic retinopathy using discrete wavelet transform. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 2013;227(3):251–261. [PubMed] [Google Scholar]

3] Antal B., Hajdu A. An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. IEEE transactions on biomedical engineering, 2012;59(6):1720–1726. [PubMed] [Google Scholar]

4] Gulshan V., Peng L., Coram M., Stumpe M. C., Wu D., Narayanaswamy A., Venugopalan S., Widner K., Madams T., Cuadros J., et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA, 2016;316(22):2402–2410. [PubMed] [Google Scholar]

5] https://www.kaggle.com/kmader/inceptionv3-for-retinopathy-gpu-hr#Evaluate-the-results

6] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5961805/#r14-2840838