Georgia Institute of Technology

ECE6254 - Statistical Signal Processing

Final Project Report

Automatic Diabetic Retinopathy Detection

Arjun **Chakraborty**
Jehoshaph **Chandran**
Laura **Jeyaseelan**
Parav **Nagarsheth**
Varshanjali **Sayyaparaju**

May 2, 2015

1 Summary

Diabetic Retinopathy is a diabetic eye disease which is currently detected manually by doctors. We wish to automate the process to speed up detection, so that treatment can start earlier. First we had to do preprocessing. Then detect the blood vessels and distinguish them from the HMAs and exudate. Finally each image is assigned a label indicating the severity of Diabetic Retinopathy. In the rest of the paper, we will give a more detailed introduction to the problem, detection of Diabetic Retinopathy and expand on our motivation behind our efforts. Then we will provide a brief tutorial to ensure that the reader has enough background to understand our work. Finally, we will present our work and results followed by our understanding of the work and what more can be accomplished.

2 Detailed Description

2.1 Motivation

Diabetic retinopathy is the most common diabetic eye disease and is the leading cause of eye blindness in American adults. It is caused due to the changes in the blood vessels of the retina; the light sensitive tissue at the back of the eye. The US Center for disease control and prevention estimates that 29.1 million people have diabetes. In this regard, there is a pressing need for an automated system that can screen large populations for Diabetic Retinopathy (DR). At the present moment, clinicians manually observe retinal scans for the presence of lesions associated with vascular abnormalities caused by the disease. The project aims to create a comprehensive and automated method of DR screening with realistic clinical potential.

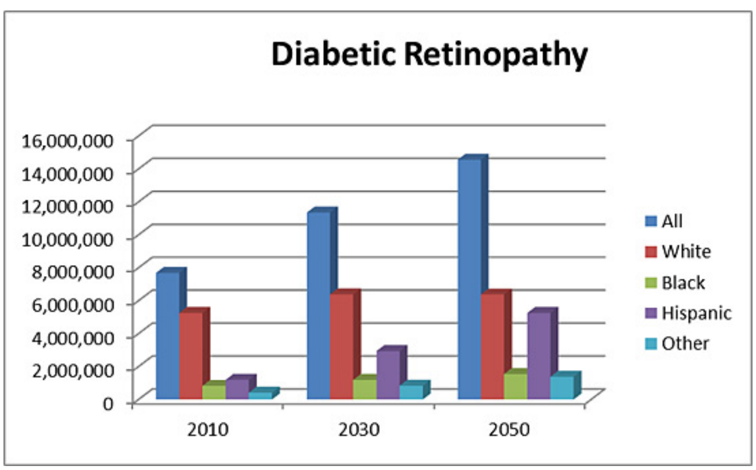

Figure 1: Prevalence of Diabetes

2.2 Background

The project aims at using Deep Learning algorithms, which is part of a broader family of machine learning methods. These algorithms are used to model high level abstractions in

data by using different architectures.The work done in this project aims at uisng a Convolution Neural Network (CNN) for the purpose of classification. A CNN is a type of feed forward neural network that comprises of one or more convolutional layers followed by one or more fully connected layers. It comprises of neurons that have learnable weights and biases. It has a loss function at the end. The input to the convolutional neural network is in the form of images, and this changes its structure with respect to ordinary neural networks.The implementation of the CNN is done using Caffe. Caffe is a deep learning framework developed by Berkley Vision and Learning Center (BVLC). It is an expressive architecture which encourages application and innocation. Models are defined by configuration without hard coding. The switching between CPU and GPU can be made by switching a flag, hence making the framework very efficient.
The images are processed before being fed to the CNN in order to standardize them using size normalization. This is done to reduce their size to the extent that there isn't any significant loss of data while reducing the computational complexity.

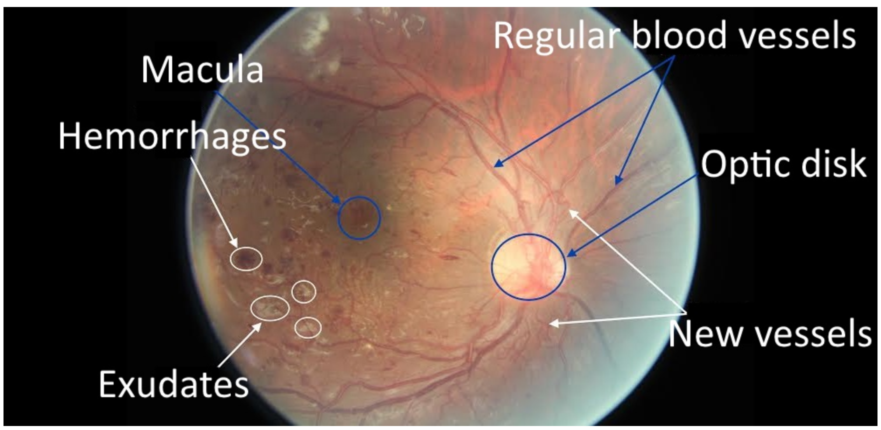

Figure 2: Features of a fundus photograph used to detect retinopathy

2.3 Related Work

There has been a lot of work done previously with respect to analysis of Diabetic Retinopathy classification methods. In [2], Random forests were used for the purpose of classifying Diabetic Retinopathy. Random forests are one of the ensemble methods of classification. In this method, a committee of learners is generated and each one casts a vote for the predicted label. The paper compares the performance of random forests with a conventional statistical approach such as Logisitic regression. The impact of sample sizes on both classifiers is also studied. In [4], high performance pre-processing of the color images was done. This was followed by the application of a recursive region growing segmentation operator. This was used to detect features such as hemorrhages and microanuersysms. [5] used contrast enhancement as a pre-processing technique before using an artificial neural network for training purposes. In [3], Convolution Neural Networks were used for the purpose of training. This was preceded by a fast but loose segmentation process, which helped in producing a set of candidate objects.

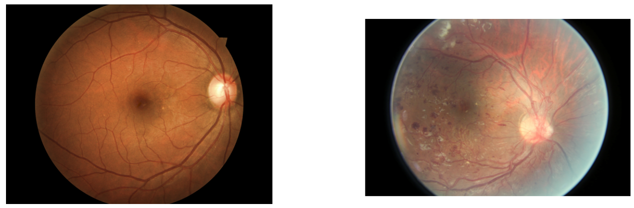

Figure 3: Original (unprocessed) images

3 Data Set

The Kaggle platform provides a large set of high-resolution Fundus images taken under a variety of imaging conditions. This includes a left and right field for every subject. Images, whether they are part of the left or right field, are labeled with a subject ID. A clinician has rated the data set on the presence of diabetic retinopathy in each image on a scale of 0 to 4. It is understood that the images have been taken from a variety of different models and types of cameras. Hence, some of the images may be dark or out of focus. The data set includes 35126 training images and 53576 testing images. 17,000 images have been used for training and testing purposes in this project.

4 Experiment

4.1 Overview

To detect the presence of Diabetic Retinopathy, the following steps are applied: Preprocessing, Segmentation and Classification. Preprocessing is required to ensure that the dataset is consistent and displays only relevant features. This step is necessary to simplify the workload for the following processes. Next,the images are segmentated to differentiate between the normal and abnormal substances. This step improves classification. Then in the final step, the classifier gives a label to each image.

4.2 Pre-processing

4.2.1 Green Channel

Of the three color channels in the image (Red,Green and blue) the contrast between the blood vessels, exudates and hemorrhages is best seen in the green channel and this channel is neither under illuminated nor over saturated like the other two. Hence, we have extracted only the green channel for analysis and classification.

4.2.2 Contrast Enhancement

To further enhance the features of the image Contrast Limited Adaptive Histogram Equalization is performed. The image is divided into smaller blocks and histogram equalization

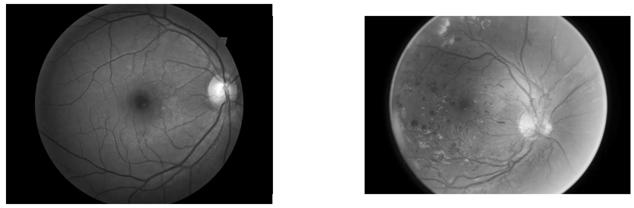

Figure 4: Separation of green channel of images

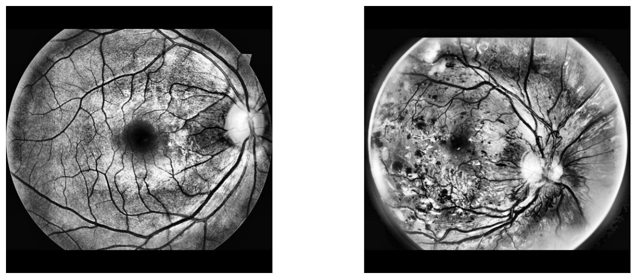

Figure 5: Images after contrast enhancement and cropping to field of view

is applied to each block. The contrast, especially in homogeneous area, is limited to avoid over-amplifying any noise that is present.

4.2.3 Cropping and Resizing

Since the original images vary widely in size and some images were chopped at the top and bottom, they had to be standardized before being passed to the CNN. Since the Field of View (FOV), the section of the retina that can be seen in the image, is circular the image is first cropped to a square of side equal to the diameter of the FOV (As some images don't have the top and bottom segments, a back patch is added to the images containing these segments to make them uniform). This new square image is then down sampled to the size 512×512 pixels.

4.2.4 Data Augmentation

Data Augmentation helps an over-fitting net perform better by training on a larger dataset. Generally, if the Convolutional Neural Network does not over- fit, it is recommended to increase the size of the dataset. Data augmentation helps in the increase of training dataset by adding transformation such as adding noise or flipping the images. It is much more economic than just adding more examples by hand. In the case of our project, the number of images labelled 0 was much larger than any other class (around 70 percent). On the other hand, the number of images labelled 4 was about 2 percent. Therefore,to make the data uniformly represented, each class is randomly sampled with replacement and from this sampled set, 2432 images of each class are selected.

4.3 Segmentation

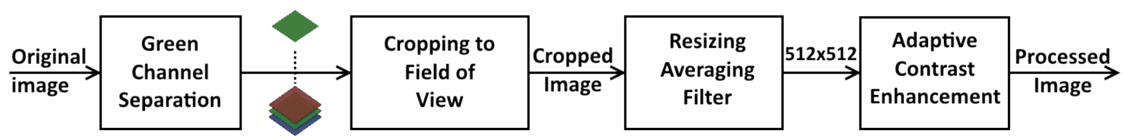

Figure 6: Steps carried out in segmentation

4.3.1 Blood Vessel Detection

The 512×512 image is used to obtain 3 copies of the same image in different sizes: 512×512, 256×256 and 128×128. The smallest image captures only the larger features, while the original sized image is able to capture the finer details. The three images are then pushed through an anisotropic diffusion filter which performs blurring of the image while retaining edges. They are then passed through a Frangi Filter [1] which detects vessels by using parameters from the Hessian of each image. This is followed by a simple thresholding process to provide a black and white image. The three images are then up sampled to their original size of 512×512 and combined to form a single black and white image showing the presence of blood vessels.

4.4 Classification

4.4.1 CNN Structure

The pipeline of the convolutional neural network being used in this project consists of 2 convolutional layers, 2 pooling layers and 3 fully connected layers. The input image to the network is a 512×512 image. The first convolutional layer uses 64 filters and a receptive field of size 4×4. The stride and the zero padding are kept at 4 and zero respectively to have 128 neurons in the first convolutional layer. The pooling layer is generally a 2×2

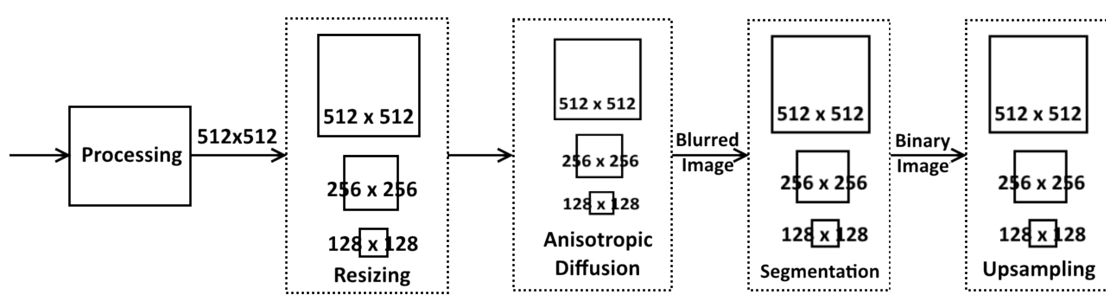

Figure 7: Steps carried out for vein detection

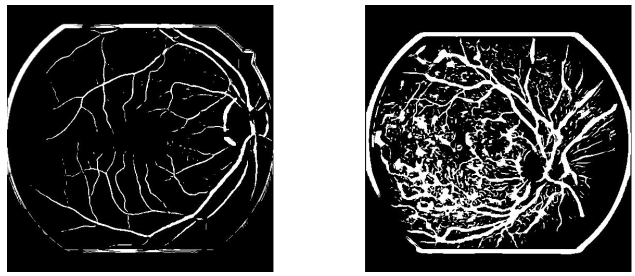

Figure 8: Images after vein detection

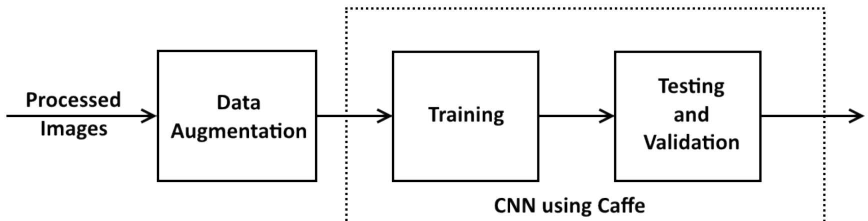

Figure 9: Steps carried out in classification using a Convolution Neural Network (CNN)

down-sampling layer, which follows the convolutional layer. The second convolutional layer receives a 256×256 down sampled image. This layer also consists of 64 filters, but they are placed at a stride of 2 and there is one layer of zero padding. A dropout layer is placed at the end of the first fully connected layer. The motivation behind using the dropout layer is to reduce over-fitting associated with coadaptation of feature detectors.It prevents this co-adaptation by setting some of the unit activation values in a given layer to zero. The ReLU activation layer is placed after the dropout layer which is a layer of neurons that increase the non linear properties of the decision function.

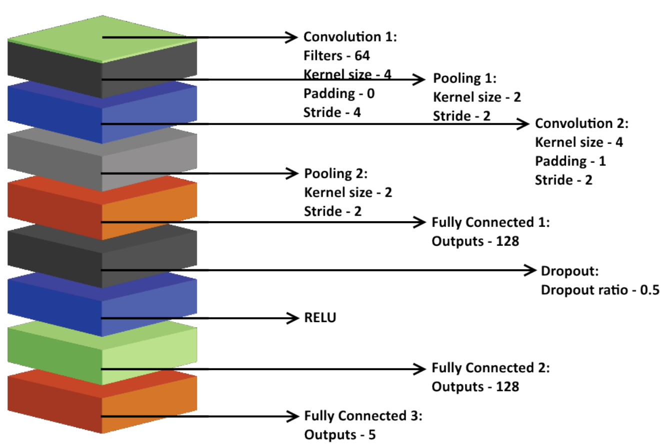

Figure 10: Structure of the CNN and parameters of each layer

4.5 Class Imbalance

The dataset that we used had significant class imbalance. For example classes numbered 0 and 1 outnumber the other three by an order of magnitude. As a result, the CNN was not able to learn the features of the less represented classes. To overcome this class imbalance, we used a combination of oversampling and undersampling. The dataset was brought to an approximately balanced representation by undersampling (randomly removing) images from over-represented classes and oversampling (duplicating) images from underrepresented classes. This resulted in around 3500 images per class. However, as we shall see in the results, oversampling leads to overfitting and led to a minor improvement in the per-class accuracy at a huge penalty to the overall accuracy.

5 Results

Figure 11 shows the confusion matrix obtained from the CNN without adjusting for class imbalance. It is clearly seen that many classes were predicted to be either 0 or 1, the majority classes. After adjusting for class imbalance, figure 12 shows that the per-class accuracy was improved for some of the minority classes. However, this increase comes at a cost of reduced overall accuracy and overfitting. We feel this problem occurs because of oversampling the minority classes. Indeed, the CNN is not learning features that differentiate minority classes from the majority ones. Instead, it is overfitting to those features and is unable to classify the test set accurately.

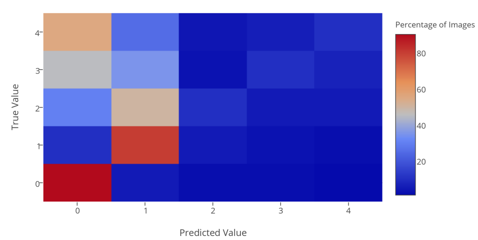

Figure 11: Confusion Matrix without correcting for Class Imbalance

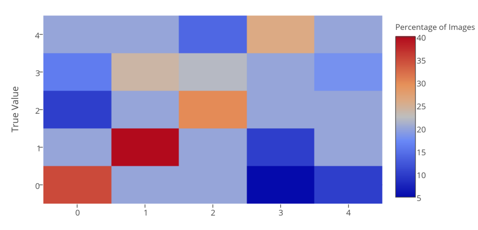

Figure 12: Confusion Matrix after correcting for Class Imbalance

6 Conclusion and Future Work

In this work, algorithms were used to process retinal digital images, localize major retinal landmarks and classify the required diabetic pathologies. In this work, a variety of techniques have been described for the purpose of image pre-processing. These include contrast enhancement, cropping and resizing. Classification has been done using a nine layer Convolutional Neural Network. The results have been showcased in the section above. In order to improve the results, data augmentation has been done. However, our data augmentation techniques (oversampling and undersampling) proved to be insufficient to overcome the challenges of this dataset. This has been showcased through the confusion matrix displayed in the section above. However, the results do present encouraging results in identification of diabetic retinopathy. We feel we can improve results by using more advanced data augmentation techniques and better loss functions. We also aim to improve the pre-processing results by using blob detection to identify the exudates. In case of classification, there is the possibility of implementing the GoogLenet CNN. Due to complex structure of the GoogLenet, and GPU constraints, we were unable to implement it. The idea behind using the GoogLenet is to use the pre-trained parameters which have been successfully used for object recognition tasks. Tuning would involve decreasing the overall learning rate, but increasing the learning rate of the last layer. Intuitively, the lower convolution layers in the GoogLenet would have performed generic object recognition tasks such as edge detection; it is the higher layers, which would have to learn faster with respect to the level of abstraction.

7 Work load distribution

Arjun Chakraborty	Classification using CNN,Caffe.
Jehoshaph Chandran	Pre-processing
Laura Jeyaseelan	Pre-processing including blood vessel detection
Parav Nagarsheth	Classification, setting up on Jinx
Varshanjali Sayyaparaju	Group leader, contrast enhancement, CNN classification

References

[1] Attila Budai, Rüdiger Bock, Andreas Maier, Joachim Hornegger, and Georg Michelson. Robust vessel segmentation in fundus images. *International journal of biomedical imaging*, 2013, 2013.

[2] Ramon Casanova, Santiago Saldana, Emily Y Chew, Ronald P Danis, Craig M Greven, and Walter T Ambrosius. Application of random forests methods to diabetic retinopathy classification analyses. *PloS one*, 9(6):e98587, 2014.

[3] Gilbert Lim, Mong Li Lee, Wynne Hsu, and Tien Yin Wong. Transformed representations for convolutional neural networks in diabetic retinopathy screening. In *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014.

[4] Chanjira Sinthanayothin, JF Boyce, TH Williamson, HL Cook, E Mensah, S Lal, and D Usher. Automated detection of diabetic retinopathy on digital fundus images. *Diabetic medicine*, 19(2):105–112, 2002.

[5] D Usher, M Dumskyj, M Himaga, TH Williamson, Sl Nussey, and J Boyce. Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening. *Diabetic Medicine*, 21(1):84–90, 2004.