Diabetic Retinopathy Detection using Deep Convolutional Networks

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Mukesh B R
Department of Electrical Engineering
Indian Institute of Science, Bengaluru

Supervisor
Dr. Chandra Shekhar Seelamanthula

In partial fulfillment of the requirements for the degree of
Summer Research Intern

July 8, 2019

Acknowledgements

I would like to sincerely thank Dr.Chandra Sekhar Seelamantula Sir and Spectrum Lab for the valuable guidance given to me during the duration of the internship and also for providing all the necessary facilities and hardware resources to perform computation on data of this magnitude in an academic environment for the research work which helped in completion of the project in the stipulated duration.

I would also like to thank Aniketh M, Harsha Shridhar, Dwiref Oza, Mihir Deshhmukh, Kevin Raj for the support and insight which provided a much detailed understanding of the project.

Finally I would like to express my gratitude to all the members of Spectrum Lab who have helped directly or indirectly into the progress of the project and supported us throughout the period of the internship.

Mukesh B R

Abstract

- Diabetic retinopathy is the leading cause of blindness in the working-age population of the developed world. 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. 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.
- In developing countries like India with scarce medical facilities, it takes about a week to diagnose the extent of DR. With increasing advancement in Image processing and Deep Learning techniques, we sought to use it in order to speed up the DR grading process.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Pathology

Retinal pathologies are responsible for millions of cases of blindness around the world. Glaucoma (4:5 million cases), age related macular degeneration (3:5 million cases) and diabetic retinopathy (2 million cases) are the lead- ing cause of blindness in adults. Early diagnosis is pivotal in slowing down the progression of these diseases and therefore preventing the occurrence of blindness.

1.2 Data

The data originates from a 2015 Kaggle competition:'Diabetic Retinopathy Detection'. Images are taken from under varies lighting conditions from different cameras by different doctors due to which data is extremely noisy and requires multiple preprocessing steps to get all images to a usable format.

Different datasets like Messidor and iDRiD was used for cross-validation to ensure the robustness of the model.

The training data is comprised of 35,126 images, which are graded into 5 classes as follows:

- 0 - No DR
- 1 - Mild DR
- 2 - Moderate DR
- 3 - Severe DR
- 4 - Proliferative DR

The dataset had the following class distribution :

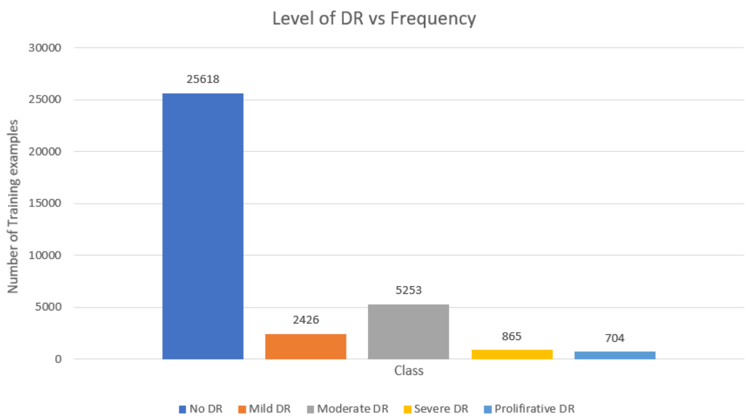


Figure 1.1: Class imbalance of the EyePacs training set

During the training of the model, one if the biggest challenges we faced was the class imbalance as shown in the aforementioned graph. During the course of the internship, different methords were tried out. The best approach has been mentioned in this report.

Chapter 2

Preprocessing

The original kaggle dataset consisted of fundus images of different sizes and orientation. A sample image is as follows:

After inspection we found the green channel of the image to contain maximum visual contextual information along with maximum contrast.

Also in order to highlight neo-Vascularization we had to enhance contrast. To achieve this contrast enhancement, we chose Contrast Limited Adaptive Histogram Equilization(CLAHE) with clipLimit as 2.0 and tile size as 8x8.(As seen in fig2.)

Also to highlight the high frequency exudates we choose the L channel in LAB image space.

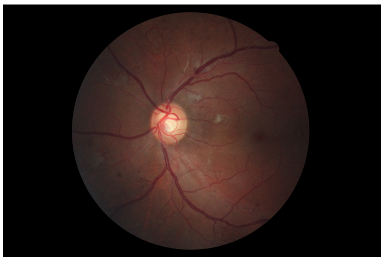
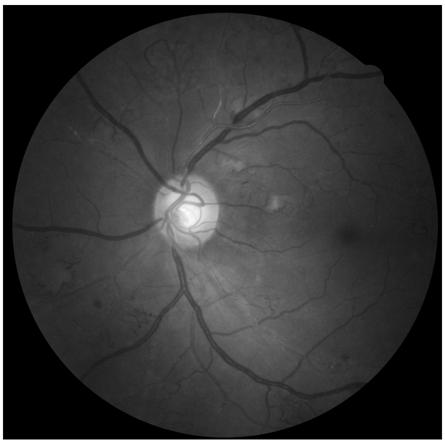


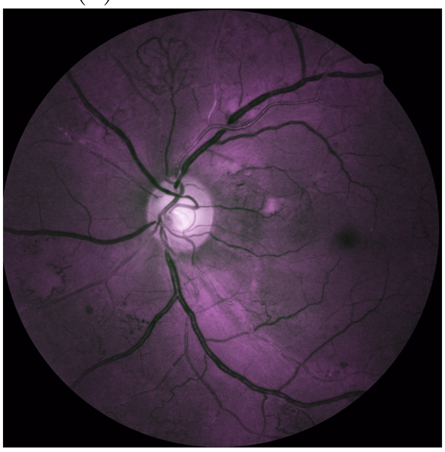
Figure 2.1: A sample training image

(a) Green Channel

(b) Contrast Enhanced



(c) Luminosity



(d) Processed image

Table 2.1: Steps for preprocessing

Since the classe 3 and 4 were under-represented, these classes were augmented using flips and rotations to oversample these images by 4 times.

Chapter 3

Model Architecture and Training

Model Architecture

We propose a VGG + FCNN based architecture for the coarse classification. The model consists of a VGG16 based convolutional layers for feature extraction and the FCNN for feature classification.

The VGG with its 3x3 filters provides a better localized learning whereas the FCNN 5x5 filters after pooling provides contextual data. All convolutional segemnts in the VGG were followed by MaxPool layer and all convolutional segments in the FCNN layers were followed by a batch normalization layer to prevent

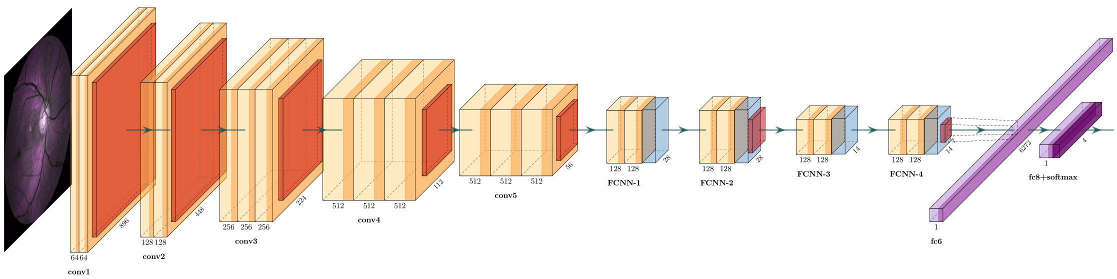


Figure 3.1: Proposed model Architecture

overfitting (high variance). All the layers in the model use a 'ReLU' activation except the last layer which uses a softmax activation.

3.1 Hyperparameters

- **Activation Function**

  The ReLU activation was used throught the experiment since the ReLU activaton function is computationally inexpensive. Also the ReLU activation function doesnt saturate as quickly like the sigmoid activation function.

- **Optimizer**

  The Stochastic Gradient Descent (SGD) with Nestrov momentum was used as the optimizer for initial convergence. SGD is faster than batch Gradient Descent algorithm since it updates weights over one single example instead of calculating the gradient over all training examples to take just one take just one decent step. But the SGD is too agresive an optimizer for finding a global optimum. Therfore AdaDelta was used for further convergence.

- **Loss Function**

  The DR grading problem can be treated as a multiclass problem, suffering from a massive class imbalance. To overcome the class imbalance weighted categorical crossentropy was used for intial convergence followed by weighted focal loss.

  Weighted Cross-entropy is defined as

  $\mathrm{H}_{y'}(y) := -\sum_i (y'_i \log(y_i) + (1 - y'_i)\log(1 - y_i))$

  Focal loss for binary classification is defined as

  $\mathrm{FL}(p_\mathrm{t}) = -\alpha_\mathrm{t}(1 - p_\mathrm{t})^\gamma \log(p_\mathrm{t})$

3.2 Training

The classes 3 and 4 of the EyePacs Kaggle training Dataset was supplemented with images from the testing dataset to reduce the class imbalance. The images were then preprocessed with the procedure mentioned above.

The model was loaded with imagenet pretrained weights for VGG-16 feature extractor and Xavier initialization for the FCNN layers. The model was trained for 30 epochs on SGD(lr=0.01 and momentum = 0.9) as the optimizer and weighted crossentropy as the loss function to obtain a coarse classification. Then Adadelta(lr=0.005 and Adadelta decay = 0.95) and weighted focal loss as the loss function for another 50 epochs for a fine classification.

Chapter 4

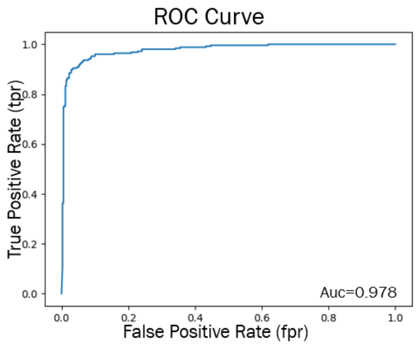
Results

4.1 Referable DR detection

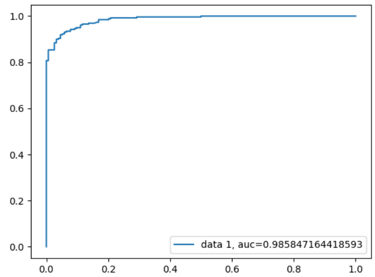
Referable DR is classified as the stages 0,1 versus stages 2,3,4. Referable DR refers to the patient suffering from prominent vision loss and possible lasting damage.

	Kaggle	Messidor	iDRiD
Sensitivity	0.8172	0.9084	0.9653
Specificity	0.9583	0.9484	0.8701
Error_rate	0.1135	0.0683	0.0702
Accuracy	0.8865	0.9417	0.9298
Precision	0.9529	0.9268	0.9259
Dice_Coeff	0.8798	0.9175	0.9452
Jaccard	0.7885	0.8476	0.8961
AUC	0.9621	0.978	0.9856

Table 4.1: Table of Results for Referable DR
The results suggest that we are suffering with a case of lower specificity.



(a) Messidor ROC



(b) iDrID- ROC

Table 4.2: Reciever Operator Characteristics curve

4.2 Severity Grading

Due to the downsampling of images, Micro-aneurisms were lost. In order to resolve this issue we used a bucket of [0,1] and classify this bucket with the other classes. This 4 class classification has exudates and Neo-vascularisation as its key features and hence the downsampling of images has relatively lesser impact.

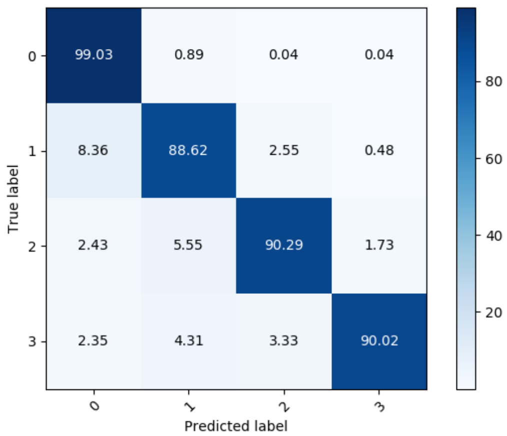
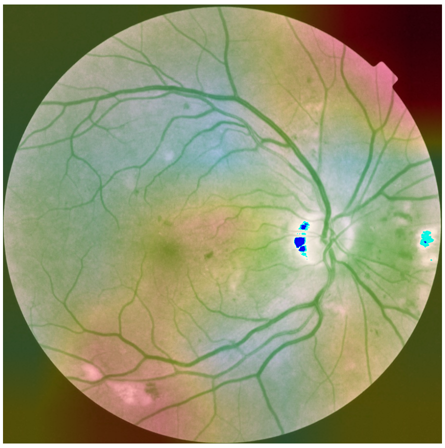


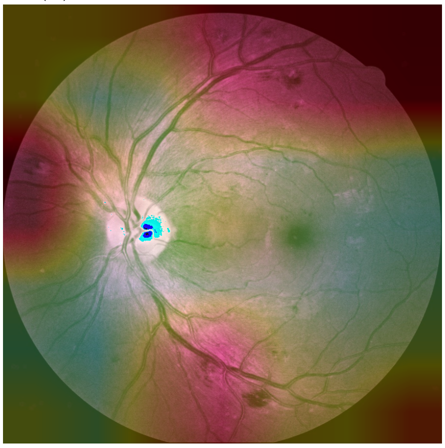
Figure 4.1: Confusion matrix on the EyePacs Training set(Validation Subset)

4.3 Heatmaps and Visualisations

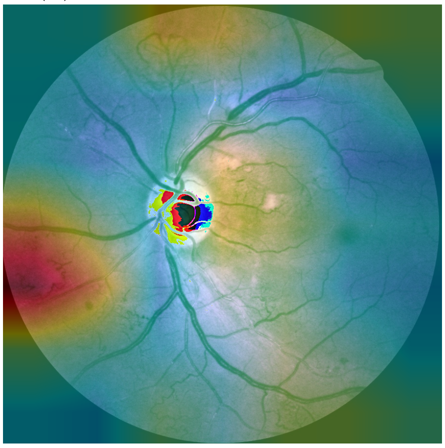
(a) Class 2 confidence of 0.87



(b) Class 2 confidence of 0.96



(c) Class 3 confidence of 0.91



(d) Class 4 confidence of 0.98

Table 4.3: Gradient Activation Maps of a few training examples

As shown above the model is able to pick up on the classic traits of different stages of DR such soft exudates of class 2, Exudates for Class 3 and Exudates+Neo-Vascularisation for class 4. Hence we can conclude the model hasnt picked up on Micro-Aneurisms .

Chapter 5

Challeneges and Future work

5.1 Challenges

The Eyepacs dataset was heavily imbalanced with respect to classes and hence presented us with a problem with class 3 and 4 detection. Also due to the size of Micro aneurisms, they were lost by any downsampling of any kind and hence class 1 was not effectively being detected. The quality of a this dataset also presented a problem due to the amount of noise and adversarial examples being presented.

The grading of this disease is quite difficult for some of the most trained opthamologists and hence consistency across different datasets poses a significant challenge.

Due to the difference between the lighting conditions, imaging equipment and resolution of images across datasets, Domain adaptation across datasets was not applicable. Hence cross validation posed a few adversarial examples bringing down the AUC.

5.2 Future Work

Due to the fact that micro-aneurisms were lost, we were not able to achieve a 0-1 classification of Normal Diabetic Retinopathy. Using a segmentation map for micro-aneurisms as a pre-processing would alleviate this issue.

Using different pre-trained architecture for feature extraction would help achieve a higher accuracy.

Preprocessing techniques to filter out adversarial examples from the EyePacs datasets with lower contrast would help enrich data quality.

Using Deformable Convolutions would increase spatial awarness of the model.

Chapter 6

Conclusions

Diabetic Retinopathy being a leading cause of blindness has led to a lot of research on it. We observe that using a segmentation based network leads to considerable increase in performance. Also we note that increasing the image size ensures more contextual and textural data. After visualising the heatmaps and filters we can conculde the model performs as expected since its able to pick up on classical traits of each notion and grade of DR.Also since the model was able to adapt to adversarial examples we can conclude the model to be robust across different datasets.

References

[1] Y. Zhou, X. He, L. Huang, L. Liu, F. Zhu, S. Cui, and L. Shao, "Collaborative learning of semi-supervised segmentation and classification for medical images," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[2] Z. Wang, Y. Yin, J. Shi, W. Fang, H. Li, and X. Wang, "Zoom-in-net: Deep mining lesions for diabetic retinopathy detection," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 267–275, Springer, 2017.

[3] Y.-W. Chen, T.-Y. Wu, W.-H. Wong, and C.-Y. Lee, "Diabetic retinopathy detection based on deep convolutional neural networks," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1030–1034, IEEE, 2018.

[4] R. Pires, S. Avila, H. F. Jelinek, J. Wainer, E. Valle, and A. Rocha, "Beyond lesion-based diabetic retinopathy: a direct approach for referral," *IEEE journal of biomedical and health informatics*, vol. 21, no. 1, pp. 193–200, 2015.