Deep Learning for Detection of Diabetic Retinopathy in Retinal Fundus Images

Jose Miguel Arrieta, Nicolas Barrera, Sebastian Ramirez

Machine Learning Final Project

UNIVERSIDAD NACIONAL DE COLOMBIA

ABSTRACT

Diabetic retinopathy (DR) is one of the most severe complications of diabetes, leading cause of blindness in the working-age population of the developed world. This project presents the application and results of different popular Deep Learning architectures that were fine tuned to detect DR in eye fundus images. Previous work from literature Voets et al. (2019) Krause et al. (2018) taken as reference and public datasets are usedfor training, test and benchmark of the models.

INTRODUCTION

Diabetic retinopathy (DR) is one of the most severe complications of diabetes, leading cause of blindness in the working-age population of the developed world, estimated to affect over 93 million people. Diabetic Retinopathy reduces the blood supply to the retina, including lesions that appear on the surface. Clinicians can identify DR by the presence of lesions associated with the vascular abnormalities caused by the disease and can be classified into red lesions such as micro aneurysms and hemorrhages and bright lesions such as exudates and cotton stains.

Scarcity of doctors and lack of availability represents a problem for patients specially in distant regions. The need for a comprehensive and automated method of DR screening has long been recognized, and previous efforts have made good progress using image classification, pattern recognition, and machine learning.

METHOD

The Deep Learning algorithms used consisted of fine tuning for different pretrained architectures of Convolutional Neural Network (CNN)

- **Inception V3:** Is a CNN that is 48 layers deep, used for assisting in image analysis and object detection, and got its start as a module for Googlenet. The network has an image input size of 299-by-299.
- **ResNet-50** Is a CNN that is 50 layers deep, Residual Network or ResNet as the name indicates introduces residual learning. The network has an image input size of 224-by-224.
- **NasNet Mobile** Is a CNN that is 44 layers deep. The architecture of those CNN has been found by Neural Architecture Search (NAS). The network has an image input size of 224-by-224

EXPERIMENTAL SETUP

- **Kaggle (DR) dataset** This dataset consists of high-resolution retina images taken under a variety of imaging conditions. A left and right field is provided for every subject.A clinician rated the presence of diabetic retinopathy in each image on a scale of 0 to 4 (No DR, Mild, Moderate, Severe, Proliferative DR) respectively.
  Original Kaggle dataset is partitioned in Train and Tests dataset, with 35121 images and 42918 images (gradable) respectively of different sizes.
- **Voets Partition** In Voets et al. a different partition of the Kaggle DR dataset is created, replicating the distribution used in Gulshan et al. but with public data. Voets used a pre-processing where fundus images are centered and resized to 299 x 299 pixels.
- **Messindor-2-dataset** This dataset is a collection of Diabetic Retinopathy (DR) examinations, each consisting of two macula-centered eye fundus images (one per eye). Messidor-2 contains 874 examinations (1748 images) with a spreadsheet containing image pairing.
- **Evaluation** A common practice used in literature with DR eye fundus images is to train in a dataset and test in another one. For this project we trained our models using Voets partition with Kaggle images and tested our models on the Messidor-2 dataset.

VIDEO

Presentation of the poster can be seen at https://youtu.be/UUKGujN67lI

PREPROCESING

For algorithm training, input images were scale normalized by detecting the circular mask of the fundus image and resizing the diameter of the fundus to be 299 pixels wide.

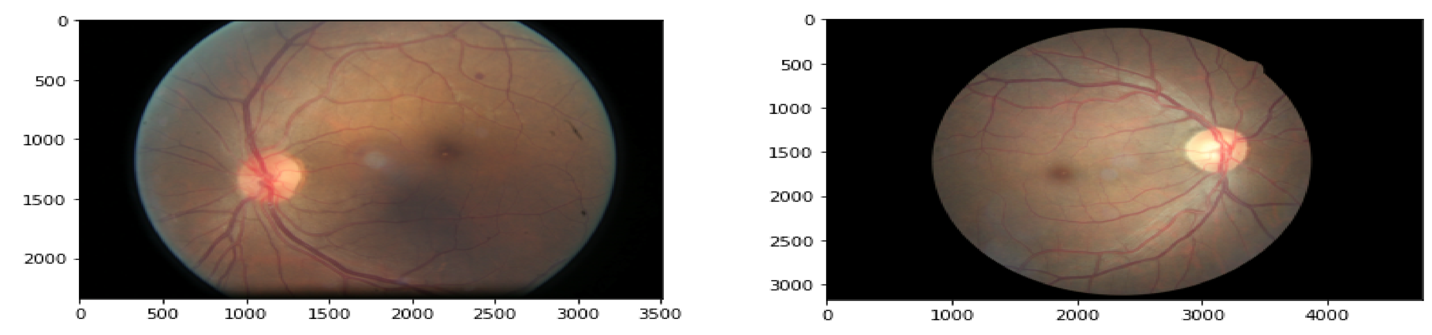

Figura 1: Original Images

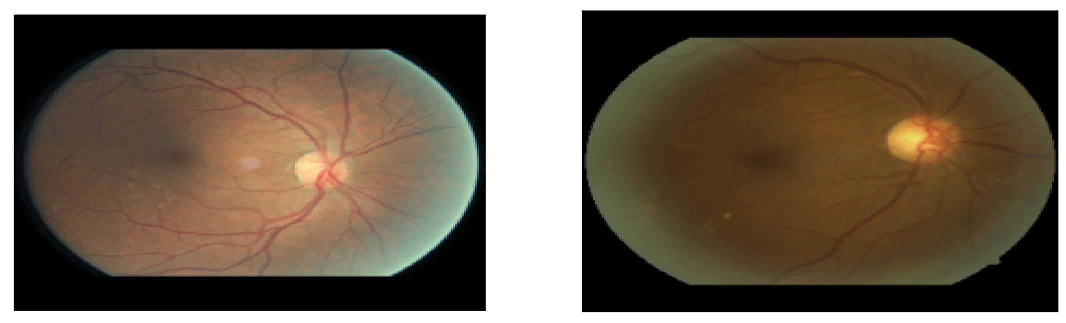

Figura 2: Preprocesed Images

RESULTS

The results of the experimentation are resumed in the following table:

Taula 1: Area Under Curve

Algorithm	AUC		
	Train	Test	Messidor2
InceptionV3	0.93	**0.96**	**0.88**
ResNet	0.86	0.95	0.85
NASNetMobile	0.79	0.95	0.80

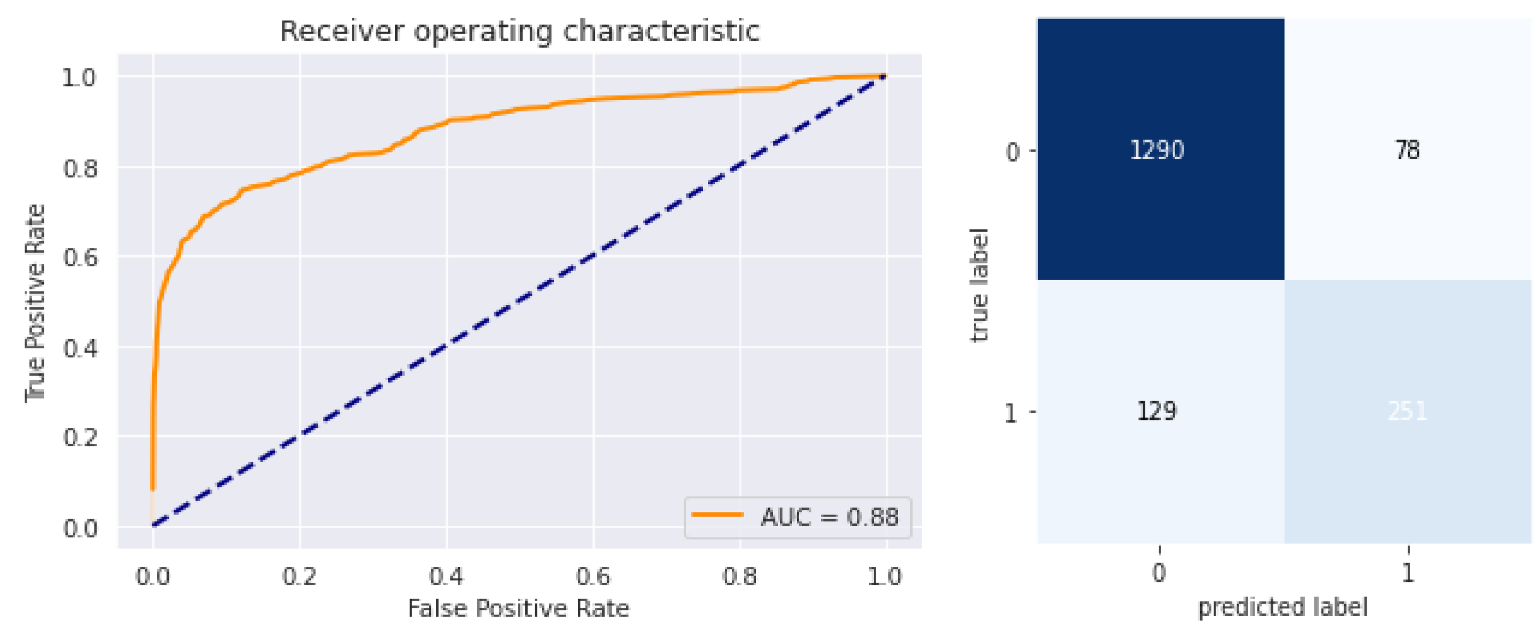

Figura 3: ROC curve and confusion Matrix of Inceptionv3 model on Messidor-2 dataset

CONCLUSIONS

- Using pre-trained Deep models and fine tune with a dataset from a very different domain is a very good approach and present good results.
- The results are very sensitive to the labels and the partition of the dataset so having good ground truth labels of the training and testing of the data is important.
- Futuro work like TPU using TFRecords could reduces the time it takes to train with several batches of the data and increase experiments.

BASIC REFERENCES

Krause, J., Gulshan, V., Rahimy, E., Karth, P., Widner, K., Corrado, G. S., . . . Webster, D. R. (2018). Grader Variability and the Importance of Reference Standards for Evaluating Machine Learning Models for Diabetic Retinopathy. *Ophthalmology*, *125*(8), 1264–1272. doi: 10.1016/j.ophtha.2018.01.034

Voets, M., Møllersen, K., & Bongo, L. A. (2019). Reproduction study using public data of: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *PLoS ONE*, *14*(6), 1–11. doi: 10.1371/journal.pone.0217541

[jmarrietar, nibarrerabo, jusramirezru]@unal.edu.co