Deep learning models for diabetic retinophaty screening program

Abraham Sánchez-Pérez[1], E. Ulises Moya-Sánchez[1], Raúl Nanclares-da Veiga[1], Jorge Martínez-Ortega[1], Alexander Quevedo-Chacón[1] and Gaspar González-Briceño[2]

[1]Coordinación General de Innovación, Gobierno del Estado de Jalisco
[2]Departamento de Diseño Electrónico, CINVESTAV Guadalajara

Abstract

Diabetic Retinopathy (DR) in Mexico is especially challenging due to i) high prevalence of diabetes in the country, ii) low rate of ophthalmologists and iii) the lack of public policies to address the DR screening. In this context, two Mexican institutions and one international financial support a three years project, to implement a DR screening program to harness the power of Artificial Intelligence (AI). In this work, we present our preliminary Deep Learning (DL) models to cropping and classifying Retinal Fundus Images (RFI) from three public datasets and one private dataset. Some of our models can achieve 93% in test accuracy and up to 98% of sensitivity. We are going to perform transfer learning with a new local dataset. We expect to improve the user-experience based on AI and reduce the DR detection time.

1 Introduction

One of the biggest challenges in the public health Mexico is the screening of DR, mainly due to high rates of diabetes, the few specialized centers and physicians as well as the lack of public policies to follow up on this problem. Jalisco government, Mexican university and an international financial institution, propose to use AI through DL to reduce the time it takes the specialists to evaluate the RFI for DR screening. We expect that DL models coupled with specialized equipment and public policies, will help to DR early detect and prevent blindness, reducing the adverse effects of this disease. We have used pre-trained models based on Inception V3 [7], ResNet50 [4], VGG-19[6] for RFI classification, and UNet [5] for RFI cropping. Some models achieve 93% in test accuracy and we think that we can use those pre-trainded models to apply transfer learning with a new local dataset.

2 Data and Methods

We used 3 public datasets [3, 1, 2] for training and testing and one private dataset d_4 only for testing. The RFI are rated from grades R0 (not DR) to R4 (maximum DR grade) [8]. We label 5 datasets: i) d_0 used for cropping process (all RD grades), ii) d_1 grade R0 versus {R3-R4} iii) d_2 grade R0 versus {R2-R4} iv) d_3 grade R0 versus R2 and {R3-R4}, v) d_4 grade R0 versus {R3-R4}, see Table 1.

We performed image cropping trough a segmentation technique by using UNet. Cropping process helps to reduce irrelevant information like the background of RFI, also it performs image resize which result in square-size image. For classification, we made nine experiments that consist of a cross-test between d_1, d_2, d_3 and three well known CNNs, Inception V3, ResNet50 and VGG-19, pretrained with imagenet weights. The input layer was changed to the image size (512x512), in addition, an extra flatten layer was appended to adapt the number of classes.

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

Table 1: Composition of the datasets used for cropping and classification on RFI.

Data	Crop	Classification			
	d_0	d_1	d_2	d_3	d_4
Train	1800	2606	8778	3909	-
Validation	1800	1156	3900	1734	-
Test	2500	868	2924	1302	121

3 Results and Conclusions

The validation loss value reported in the cropping model was 0.0025. The classification results are shown in the Table 2 and confusion matrices in Table 3. In Addition, we use the ResNet50 trained with d_1 model to evaluate the local dataset (d_4) which provided a sensitivity value of 98% (Table 3).

Table 2: CNN architectures and classification datasets comparison.

Dataset	Model	Acc	Loss	Val acc	Val loss	Eval acc	Eval loss
d_1	Inception V3	0.9996	9.1210e-04	0.9599	0.3118	0.93	0.439
	ResNet50	**0.9977**	**0.0050**	**0.9608**	**0.2575**	**0.93**	**0.298**
	VGG-19	0.4879	0.6932	0.5142	0.6931	0.5	0.69
d_2	**Inception V3**	**0.9967**	**0.0075**	**0.8888**	**0.7802**	**0.879**	**0.808**
	ResNet50	0.9965	0.0096	0.8811	0.7408	0.868	0.8187
	VGG-19	0.5006	0.6932	0.5041	0.6931	0.5	0.693
d_3	**Inception V3**	**0.9964**	**0.0115**	**0.7620**	**1.3168**	**0.758**	**1.33**
	ResNet50	0.9913	0.0270	0.7274	1.2731	0.727	1.31
	VGG-19	0.3216	1.0989	0.3396	1.0986	0.33	1.098

Table 3: Confusion matrix of DR classification of the best results of the datasets.

	d_1		d_2		d_4	
	No DR	DR	No DR	DR	No DR	DR
No DR	412	22	1364	98	13	44
DR	37	397	254	1208	1	63

	d_3		
	R0	R2	R3,4
R0	369	54	11
R2	96	259	79
R3,4	15	60	359

We have presented one pre-preprocessing model and 3 pre-trained models for RD classification. We expect to use those models to improve the user-experience and reduce the DR detection time.

References

[1] Eye Clinic. Diabetic Retinopathy Segmentation and Grading Challenge. `https://idrid.grand-challenge.org/Data/`, 2012. [accessed 04-September-2019].

[2] Etienne Decencière, Xiwei Zhang, Guy Cazuguel, Bruno Lay, Béatrice Cochener, Caroline Trone, Philippe Gain, Richard Ordonez, Pascale Massin, Ali Erginay, Béatrice Charton, and Jean-Claude Klein. Feedback on a publicly distributed database: the messidor database. *Image Analysis & Stereology*, 33(3):231–234, August 2014.

[3] EyePACS. Diabetic Retinopathy Detection competition. `https://www.kaggle.com/c/diabetic-retinopathy-detection/`, 2015. [accessed 04-September-2019].

[4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.

[5] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015.

[6] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv e-prints*, September 2014.

[7] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. *CoRR*, abs/1512.00567, 2015.

[8] Sonia Zachariah, William Wykes, and David Robert Yorston. Grading diabetic retinopathy (dr) using the scottish grading protocol. In *Community eye health*, 2015.