ISSN 1870-4069

Automatic Cropping of Retinal Fundus Photographs using Convolutional Neural Networks

Gaspar González Briceño[1], Abraham Sánchez[2], E. Ulises Moya Sánchez[2,4], Susana Ortega Cisneros[1], Germán Pinedo[1], Mario S. García Contreras[3], Beatriz Alvarado Castillo[3]

[1] CINVESTAV,
Department of Electrical Engineering,
Mexico

[2] Jalisco Government,
Department of Artificial Intelligence,
Mexico

[3] Centro Médico Nacional de Occidente IMSS,
Department of Ophtalmology,
Mexico

[4] Universidad Autónoma de Guadalajara,
Maestría en Ciencias Computacionales,
Mexico

ggonzalezb@cinvestav.mx,{sortega,gapinedo}@gdl.cinvestav.mx,
{abraham.sanchez,eduardo.moya}@jalisco.gob.mx, drmariosgc@yahoo.com,
eduardo.moya@edu.uag.mx

Abstract. Deep learning (DL) is used in widespread applications including the medical sector. Particularly, Convolutional Neural Networks (CNNs) are architectures commonly used to classify or predicts retinal fundus photographs (RFP). However, the image noise located in the background of the RFP and space location of the retinal Region of Interest (ROI) can affect the performance of these models. One solution to this problem is to make a segmentation of the RFP. In this work we present a Segmentation Model based on CNN (SMCNN) for the cropping process, this model reaches high accuracy levels for test images up to 98%.

Keywords: convolutional neural networks, retinal fundus photographs, segmentation.

1 Introduction

Deep Learning (DL) models are generated by multiple Artificial Neural Networks Layers that can learn the representations of data on multiple levels of abstraction

[6]. Convolutional Neural Networks (CNNs) have proven to be powerful tools for a wide range of tasks, such as image recognition, audio classification, object detection, medical applications, and others [4, 14, 3].

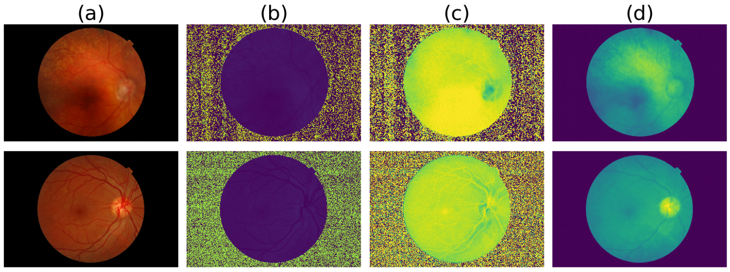

Fig. 1. Background noise representation in two RFPs with apparently good quality. (a) At the top and bottom two RFPs in RGB, (b,c,d) is the representation in HSV of each image respectively. High noise levels can be observed in background once (b) Hue channel and (c) Saturation channel are separated. On the other hand, the (d) Value channel keeps the background values close to 0 (low noise).

Medical applications of CNNs includes clinical and research solutions in the field of eye care, like optical disk segmentation [12], detection of retinal anomalies [7], classification for referable diabetic retinopathy [9], vessel segmentation [13], among others.

Retinal Fundus Photographs (RFPs) are images widely used by physicians to diagnosis an ocular disease. Deficiency in the quality of these images could commit the success of the diagnosis but also the treatment. [1, 8]. In addition, the performance of the DL models could be affected by the difference in the spatial location of the retinal region of interest (ROI) as well as the noise situated in the background, which is not visible to the human eye [15, 8]. These intrinsic problems can be analyzed throw a Hue, Saturation, and Value (HSV) color representation [16], as shown in figure 1(b,c).

In this context, a Segmentation Model based on deep CNNs (SMCNN) is proposed, which predicts and builds the retinal ROI mask of a RFP. This strategy facilitates the retinal cropping process, which helps to avoid the noise of the background and to homogenize the ROI location of each image, all this increasing the image quality used for the CNN models.

The remainder of the paper is structured as follows. The experimental setup, mask generator, SMCNN model, and cropping are described in section 2. Results, can be found in section 3. The discussions are presented in section 4. Finally, conclusions are presented in section 5.

2 Data and Methods

2.1 Classical Mask Generator

In order to build the RFPs' masks used to train, validation and test the SMCNN, we use the method is described in Figure 2. First, the RGB images are turned into the grayscale in order to avoid the background noise of the RFPs.

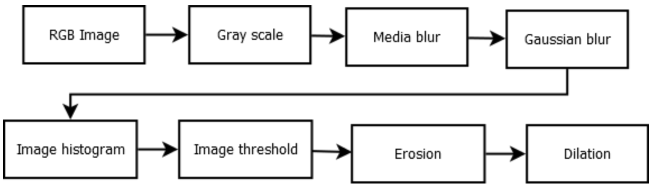

Fig. 2. Flowchart of the classical mask generator for train, validation and test.

Second, medium blur and high-value Gaussian blur filters are applied with the goal of reducing the variance between the pixel values, this produces a uniform distribution throughout the image. Then, the contrast histogram is generated, which works to know the threshold between the minimum and maximum values, identifying thus the upper and lower limits for the image binarization. Finally, erosion and dilatation filters are added to smooth the edges of the retinal ROI. As a result, we obtain some preliminary masks and these ROI masks were selected and corrected by hand.

2.2 Mask Dataset

A total of 500 binary masks were generated with the classical mask generator(aided by human) from the Kaggle Diabetic Retinopathy Detection dataset [2]. The 500 images were used to train and validate the SMCNN, which predicts the masks of new RFPs. The quantity of images used is shown in Table 1. Binary masks obtained from the classical mask generator are resized with a 1024x1024 resolution, and then they are correlated with their grayscale images counterpart, which in turn proceeds from the RFPs in RGB.

Table 1. Dataset used for the SMCNN, based on Kaggle dataset.

Dataset	No. of Images	Train	Validation	Test
Kaggle	500	280	120	100

2.3 SMCNN Training

In Figure 3, the training an validation process are presented. The configuration of the U-Net architecture [10] is based on a back-propagation algorithm[11] and Adam optimizer [5]. We use a GCloud instance with a V-100 GPU. Then, a cropping process was generating to obtain the same spacial proportion, see figure 4 as an example.

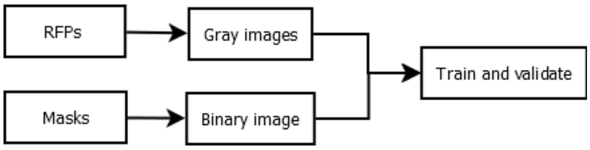

Fig. 3. SMCNN training flowchart process.

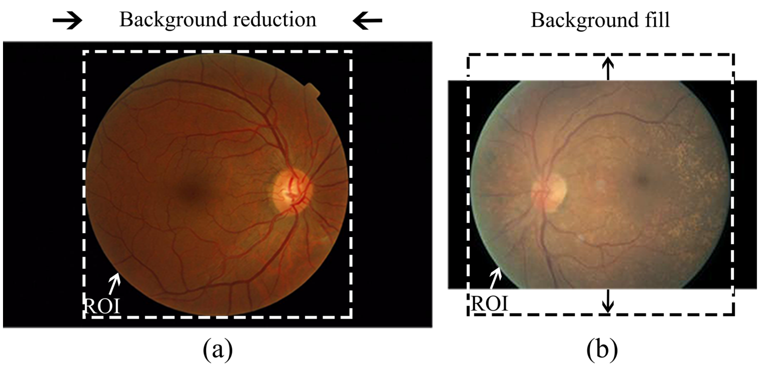

Fig. 4. Cropping process, (a) background reduction, maintain ROI's aspect ratio, (b) background reduction and fill of ROI.

3 Results

3.1 Dataset Creation With Classical Mask Generator

Examples of the dataset are shown in Figure 5. Note that, only the results with the best approximation of the shape of the retina were selected manually.

3.2 SMCNN Performance

The performance of the SMCNN model made to predict the masks of RFPs is presented at table 2. High levels obtained in the accuracy for training, validation,

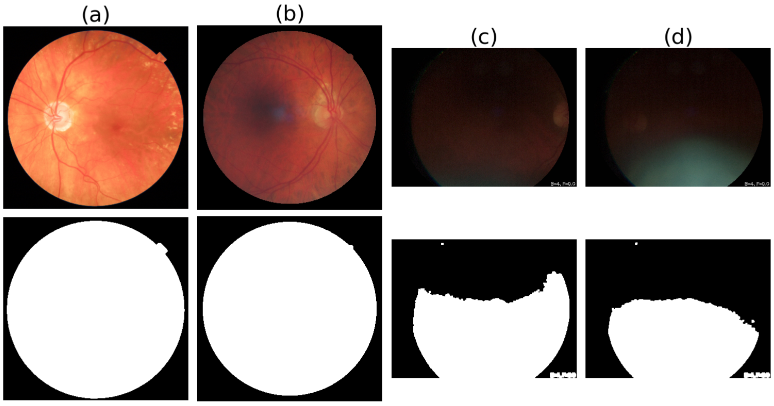

Fig. 5. Example of the results of our personalized technique for the creation of masks. Figures (a, b) are some of the successful results therefore they were selected for the training, validation and testing process. Meanwhile figures (c, d) are unsatisfactory so they were discarded.

Table 2. Train, validation and testing results.

Train		**Validation**		**Test**	
Accuracy	Loss	Accuracy	Loss	Accuracy	Loss
0.99	0.0030	0.98	0.0031	0.98	0.01898

and test are interesting results, due that they represent the convergence rate between the original RFP and the one predicted by our SMCNN model.

Figure 6 shows an example of a result of the cropping process with SMCNN, using low contrast images, in addition, is compared with a classical mask generator.

4 Discussion

Techniques described in this work represent a solution to get the retinal ROI of diversity images obtained with different methodologies. Nonetheless, a mask obtained by this methodology could be incomplete if the image quality is over or under certain quality values. The quantification and demonstration of these quality levels are proposed as future work.

5 Conclusions

In this work, a retinal segmentation model based on CNN was introduced. ROI mask obtained of this model facilitates the cropping process, avoiding the noise located in the background of the RFPs and homogenizing the location of the

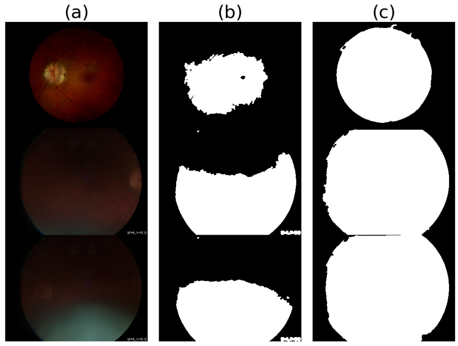

Fig. 6. Example of the comparison between a classical mask generator and DL segmentation process. (a) RFP in RGB. (b) Result using classical technique like Figure 5. (c) Result of segmentation process based on DL.

retinal ROI in the image space. Results show an accuracy of up to 98% on the test set. We hope to use transfer learning on another dataset in order to generate more segmented samples form other datasets.

References

1. Armanious, K., Jiang, C., Fischer, M., Küstner, T., Hepp, T., Nikolaou, K., Gatidis, S., Yang, B.: Medgan: Medical image translation using gans. Computerized Medical Imaging and Graphics 79, 101684 (2020)
2. EyePACS: Diabetic retinopathy detection competition. https://www.kaggle.com/c/diabetic-retinopathy-detection/ (2015)
3. Greenspan, H., Van Ginneken, B., Summers, R.M.: Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. vol. 35, pp. 1153–1159 (2016)
4. Ji, S., Xu, W., Yang, M., Yu, K.: 3d convolutional neural networks for human action recognition. vol. 35, pp. 221–231 (2013)
5. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. CoRR 1412.6980 (2014)
6. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature 521(7553), 436 (2015)
7. Li, Q., Fan, S., Chen, C.: An intelligent segmentation and diagnosis method for diabetic retinopathy based on improved u-net network. Journal of Medical Systems 43(9), 304 (2019)
8. Moya-Sánchez, E.U., Sánchez, A., Zapata, M., Moreno, J., Garcia-Gasulla, D., Parrés, F., Ayguadé, E., Labarta, J., Cortés, U.: Data augmentation for deep learning of non-mydriatic screening retinal fundus images. In: International Conference on Supercomputing in Mexico. pp. 188–199. Springer (2018)
9. Pires, R., Avila, S., Wainer, J., Valle, E., Abramoff, M.D., Rocha, A.: A data-driven approach to referable diabetic retinopathy detection. Artificial Intelligence in Medicine 96, 93–106 (2019)
10. Ronneberger, O., P.Fischer, Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). LNCS, vol. 9351, pp. 234–241. Springer (2015)

11. Rumelhart, D.E., Durbin, R., Golden, R., Chauvin, Y.: Backpropagation: The Basic Theory, p. 1–34. L. Erlbaum Associates Inc., USA (1995)
12. Sevastopolsky, A.: Optic disc and cup segmentation methods for glaucoma detection with modification of u-net convolutional neural network. Pattern Recognition and Image Analysis 27(3), 618–624 (2017)
13. Son, J., Park, S.J., Jung, K.H.: Retinal vessel segmentation in fundoscopic images with generative adversarial networks (2017)
14. Topol, E.J.: High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine 25(1), 44 (2019)
15. Wang, J., Bai, Y., Xia, B.: Feasibility of diagnosing both severity and features of diabetic retinopathy in fundus photography. IEEE Access 7, 102589–102597 (2019)
16. Wu, Y., Zheng, J., Song, W., Liu, F.: Low light image enhancement based on non-uniform illumination prior model. IET Image Processing 13(13), 2448–2456 (2019)