MEDINFO 2017: Precision Healthcare through Informatics
A.V. Gundlapalli et al. (Eds.)

doi:10.3233/978-1-61499-830-3-318

Automatic Identification of Glaucoma Using Deep Learning Methods

Allan Cerentini[a], Daniel Welfer[a], Marcos Cordeiro d'Ornellas[a], Carlos Jesus Pereira Haygert[b], Gustavo Nogara Dotto[b]

[a]*Graduate Program in Computer Science(PPGI), Department of Applied Computing (DCOM), Santa Maria, Rio Grande do Sul, Brazil*
[b]*Department of Clinical Medicine, Santa Maria University Hospital, Federal University of Santa Maria (UFSM), Rio Grande do Sul, Brazil*

Abstract

This paper proposes an automatic classification method to detect glaucoma in fundus images. The method is based on training a neural network using public image databases. The network used in this paper is the GoogLeNet, adapted for this proposal. The methodology was divided into two stages, namely: (1) detection of the region of interest (ROI); (2) image classification. We first used a sliding-window approach combined with the GoogLeNet network. This network was trained using manually extracted ROIs and other fundus image structures. Afterwards, another GoogLeNet model was trained using the previous resulting images. Then those images were used to train another GoogLeNet model to automatically detect glaucoma. To prevent overfitting, data augmentation techniques were used on smaller databases. The results demonstrated that the network had a good accuracy, even with poor quality images found in some databases or generated by the data augmentation algorithm.

Keywords:
Glaucoma; Retina; Neural Network (Computer)

Introduction

Glaucoma is a lesion that occurs inside of the optic nerve, which can cause low vision or even total blindness. Currently, this lesion has no cure, but an early diagnosis with treatment can prevent the disease progression [1] [2]. One of the approaches used to detect disease is called ophthalmoscopy. This examination may generate a fundus image, which is used for this work. Currently, the detection of glaucoma is done manually by retina experts. They use the ratio between optic nerve size and lesion size, to diagnose the presence and severity. It is a process that takes time, and the lesion is not always clear. This work proposes to perform this task automatically. To perform this task, we have to deal with some challenges, namely: very subtle lesions, poor image quality, and illumination problems. There are related works that extract the size of the optic nerve and the lesion to perform this classification; others use machine-learning methods as neural networks for this purpose. This work aims to apply a robust neural network, developed to deal with hard image classification challenges, in order to overcome many of the problems presented previously. This paper presents a summary of some similar works, description of images datasets used, methodology applied, and the results. For automatic classification, two methods were developed: one to find the region of interest (ROI), optic nerve, in the fundus image and another to classify this region between healthy and with glaucoma.

Chen et al. [3] described a method that uses a neural network with four hidden layers. The input layer accepts images with a dimension of 256x256 and three color channels. The datasets used were artificially augmented using random cropping techniques and mirroring. The work performed two experiments using two image bases, (1) The ORIGA-light, which has 650 images and, (2) SCES, which has 1676 images. The first experiment used 99 ORIGA-light images to train the network and 551 for test. The second used all the ORIGA images to train the network and all of the SCES for test. The accuracy of both experiments was 83.12% and 88.7% respectively.

The method described by Sheeba et al. [4] also uses a neural network, but with two hidden layers. There is no clear information about the private dataset provided by Giridhar Eye Institute in Cochin. It consists of 20 images with unknown ground truth, where 28 have glaucoma and 12 are normal. These were processed using erosion and dilation techniques, where the dilated image is used as background and subtracted from the grayscale image obtained. The resulting intensity is adjusted, and then it is converted to binary values by thresholding the optical disc region. The network was trained by using 20 images with unknown label, and tested in the 40 remaining images. They reported that 34 images were classified correctly.

Zhang et al. [5] consists of an online repository of fundus images. The goal is to provide a way to share those images with the public. Researchers also can benchmark their algorithms. A segmentation and classification tool was developed to assist in the construction of the image database. At the date the article was published, the database was composed of 650 images, delimited by specialists. The ratio between the size of the excavation and the size of the lesion was used for classification. The detection of the region of interest was performed automatically. The first stage is the preprocessing of images, like fringe remove. The fringe happens due to patients not putting their eye so close to the capture machine, generating a light entrance by the edges. This light input hinders the ROI's detection, since it is done using the centroid of the brightest part of the image. This method has 96% accuracy according to the article. In cases where ROI is not found by the system, the user can manually select that region.

Considering the previous works, we noticed that literature already presents works using neural networks, to detect the presence of glaucoma. However, the results with good accuracy are based on good quality images. The purpose of our work is to perform this detection even in images with low contrast, high amount of noise and low resolution. And, to detecting the ROI of those images using a robust neural network.

Materials and Methods

Image Databases

We used four public image databases that were used for training and a validation of the network. The first database, High Resolution Backgrounds (HRF) [6], was composed of 45 fundus images, divided into 3 groups of 15 images. A group of images had glaucoma, while the other groups had healthy eyes and, the last, diabetic retinopathy. For the test and validation, only the groups with glaucoma and normal images were selected.

The remaining three databases, RIM-ONE r1, RIM-ONE r2 and RIM-ONE r3, were part of the work done by Fumero *et al.* [7]. RIM-ONE r1 and RIM-ONE r2 were composed of ROI images and RIM-ONE r3 of stereo fundus images. The first database had 40 images with some degree of glaucoma and 118 healthy images, while the second had 200 images with glaucoma and 225 healthy images. The third had 74 images with glaucoma and 85 healthy images. Each image has a ground truth delimited by a specialist. For the third database, the stereo images were divided in two parts and only the diagnosed part was used in the experiment.

Deep Learning

Artificial neural networks are part of a set of techniques in the Machine Learning area. This technique was inspired by the learning process of a biological brain, and is constructed using a fully connected neural network. Briefly, each neuron receives input signals from several sources, as well as from other neurons. Each entry is multiplied by a value, called weight. All results are then summed, and verified in an activation function that decides whether the neuron should send a signal ahead. The learning occurs by adjusting those weights. When a set of inputs gives a wrong output value, the weights are adjusted to make that output correct. Trying not to disturb the result of another set of inputs [8]. The depth of an artificial neural network is measured according to the number of layers between the input and the output of a network, they are called hidden layers. An input signal passes through the network until the last layer, called output, is reached. The last layer gives the prediction of that input signal. A larger number of hidden layers allow for the classification of more complex data, such as images [9].

Deep learning is the process of training multi-layered neural networks. To facilitate this process, we will use an open source framework developed by Google called Tensorflow [10]. It is the second generation of a large-scale machine learning system developed by Google. The system originated from a project called Google Brain started in 2011, where they built the DistBelief, which was the first generation. Tensorflow uses tensors, a multi-dimensional array, to represent data as images. Graphs are employed to represent the flow of operations, where each node represents an operation with zero or more tensors. This framework has implementations of several types of algorithms, mathematical models and specific functions optimized for the training of neural networks. It also provides a practical way to use a graphics processing unit (GPU) to accelerate the training process.

An existing neural network model, called GoogLeNet[11], was used. This model was developed by Google and competed in the ImageNet challenge from 2014, whose objective was to classify about 1.3 million of images in one of the 1000 different classes. The goal of Google was to create a computationally efficient network that could be run even with low computation resources. Some of techniques employed like batch normalization, residual connection and factorization allowed for increased accuracy while maintaining performance. The network can be acquired previously trained, with the pre-defined weights. For this work the network was modified to train only a number of classes needed to solve some problems. The weights were maintained to use the principle of learning transfer [12].

Region of Interest (ROI) Detection

Finding manually the ROI of each image, from a large image database, is a hard process. An algorithm was developed to automatically detect this region. This is based on object detection works using neural networks from Malisiewicz et al. [13], Sermanet et al. [14] and an ROI detection work from Xu et al. [15]. A neural network was trained using the GoogLeNet model to classify images into two categories: region of interest or background. For this process, 107 images from ROI and 4693 from other fundus regions were taken from HRF database. An algorithm was made to make those crops, and they were manually classified. It was not a difficult task, because the crops were sequential. The network was trained and performed at about 99% accuracy, distinguishing ROI images from other structures from fundus. This good detection accuracy is due to the distinct characteristics presented in the images.

Figure 1 – The first step of the algorithm is to verify which are the possible ROI in the fundus image, resulting in a) Overlapping sliding windows. Then a suppression method is applied, resulting in only one location b) Result after suppression algorithm

To find the region of interest within the image, an algorithm was developed that uses sliding windows to scan the image in search of that region. Each window has size proportional to the size of the image. For each window it is checked on the previously described network if it is over the region of interest. If it is then the coordinates of this window are saved. If there is more than one case where this happens then a suppression method based on Malisiewicz et al. [13] and Xu et al. [15] is used, to avoid cases like that of Figure 1, where each window detects a part of the region of interest. This leads to loss of accuracy, since we need the whole region for classification.

Data Augmentation Process

The data augmentation process consists of adding deformations and noise to a data set, with the purpose of increasing the amount of data available for training. Providing a better accuracy to predict data from other datasets [16]. This technique is often used in small datasets, in order to prevent the network from eventually learning characteristics that are not relevant to classification, such as: bright images being classified as glaucoma and low contrast images as normal. It is a process that also prevents overfitting, which is a process where the network begins to decorate the images instead of learning about them. This also causes the network to have a good accuracy in the trained dataset and bad accuracy in others. The work proposed by Wu et al. [17] demonstrates the efficacy of this process to eliminate such problems presented previously, where several transformations in images are used for the learning processes focus on key features in the images.

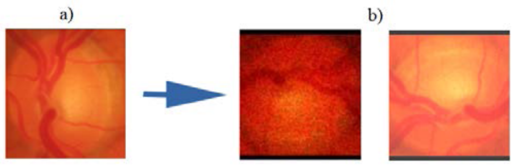

Figure 2 – Image resulting from ROI extraction algorithm a) Extracted ROI from HRF database. Due to the small amount of images available in this database, transformations were applied. b) Results from data augmentation process.

In this work, a rotation of 90, 180 or 270 degrees is randomly applied on each image, as well size rescaling, gamma variation and addition of Gaussian noise. We can apply this process several times in an image. For each iterationm we add an image to the training database. Figure 2 demonstrates an example of an image being transformed in two extra images for the training.

Image Classification

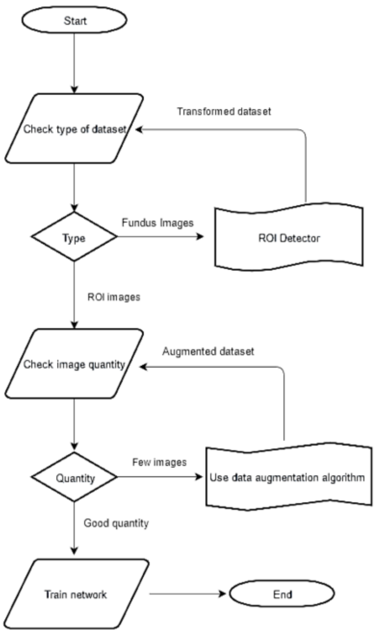

Figure 3 – Dataset training workflow

For the glaucoma classification, a neural network also based on the GoogLeNet, was trained. The training process is presented in Figure 3 and was applied to all image databases. First, we have to verify if it is a fundus or ROI database. The images must include the ROI for the classification. If it has fundus images, then all its images go through the process of ROI detection, described earlier. This process transforms a set of fundus images into a set of images of ROI images. Then we can verify if this dataset contains the minimum amount of images, to avoid problems like overfitting. Databases with less than 100 images per category were classified as "few images". In this case the image augmentation algorithm is used to expand the quantity of this category into an arbitrary value greater than the minimum value. When the quantity is good, the database is trained in the network. The processing was applied to HRF, Rim-one-r1 and Rim-one-r3 databases. The HRF database went through the data augmentation process, turning its 30 images into 330. The Rim-one-r1 database glaucoma imaging class also went through this process, the number of images increased from 40 to 120. The Rim-one-r3 database went through ROI detection process.

Experimental results

The results obtained in this experiment are shown in Table 1. They were obtained by using 10% of dataset images to validate the accuracy of network. These images were randomly selected and were not used for neural network training. This helps us to evaluate a network more precisely, by simulating images from other image databases. The training process used 5000 training steps, the number of times a set of images from the data set passes through the network. And with a learning rate, sensitivity adjustment in weights, at 0.01 per update. Training each dataset took about 7 minutes using a GeForce Gtx 1070. For the final test, involving all the image databases, the artificially generated images, generated by the data augmentation algorithm were removed.

Table 1 – Success rate in the glaucoma detection using the proposed method

Image Database	Accuracy
HRF	90,0%
RIM-ONE r1	94,2 %
RIM-ONE r2	86,2 %
RIM-ONE r3	86.4 %
HRF + RIM-ONE r1 + RIM-ONE r1 + RIM-ONE r3	87,6%

We obtained a result close to the experiment described in the work done by Chen et al. [3], but with different databases. It is a satisfactory result, considering the amount of images considered difficult for this classification, such as: large variation in brightness, much noise and ROI barely visible. Some images are considered problematic are presented in Figure 1. There are no post-processing techniques to improve the quality of images. As well neither the common practice of removing blood vessels from those images.

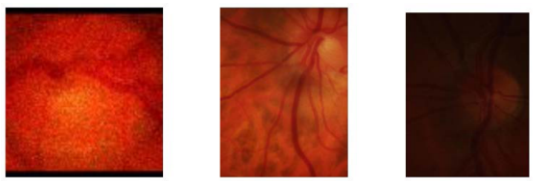

Figure 4 – Examples of challenging images, due to low contrast or much noise.

Conclusion

In this paper a deep learning method was used, to detect the presence of glaucoma in the fundus images. The GoogLeNet neural network model from Google was used to accomplish this classification. It was able to detect the presence of glaucoma even in images where it appears only subtly, and in images with very low quality, generated by data augmentation. We obtained 90% accuracy on HRF database, RIM-ONE r1 with 94,2% accuracy, RIM-ONE r2 with 86,2% accuracy, and RIM-ONE r3 with 86,4% accuracy. The combination of all databases resulted in 87,6% accuracy. To increase the amount of data of some networks, helping the training, a data augmentation algorithm was used. For the extraction of ROI from fundus image databases an algorithm was developed to detect this area. For future work a pre-processing can be done on the images of the databases, as well as expanding the number of databases of images for testing.

References

[1] D. B. Henson and R. Thampy, "Preventing blindness from glaucoma: Better screening with existing tests should be the priority," BMJ Br. Med. J., vol. 331, no. 7509, pp. 120–121, Jul. 2005.

[2] L. Rossetti *et al.*, "Blindness and Glaucoma: A Multicenter Data Review from 7 Academic Eye Clinics," *PLoS One*, vol. 10, no. 8, p. e0136632, Aug. 2015.

[3] X. Chen, Y. Xu, D. W. Kee Wong, T. Y. Wong and J. Liu, "Glaucoma detection based on deep convolutional neural network," 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, 2015, pp. 715-718.

[4] O. Sheeba, J. George, P. K. Rajin, T. Nisha, G. Sherin, "Glaucoma Detection Using Artificial Neural Network", International Journal of Engineering and Technology, vol. 6, no. 2, pp. 158-161, 2014.

[5] Z. Zhang et al., "ORIGA-light: An online retinal fundus image database for glaucoma analysis and research," 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, 2010, pp. 3065-3068.

[6] J. Odstrcilik, R. Kolar, T. Kubena, P. Cernosek, A. Budai, J. Hornegger, J. Gazarek, O. Svoboda, J. Jan and E. Angelopoulou, "Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database", IET Image Processing, vol. 7, no. 4, pp. 373-383, 2013.

[7] F. Fumero, S. Alayon, J. Sanchez, J. Sigut and M. Gonzalez-Hernandez, "RIM-ONE: An open retinal image dataase for optic nerve evaluation", 2011 24th International Symposium on Computer-Based Medical Systems (CBMS), 2011.

[8] J. Schmidhuber, "Deep learning in neural networks: An overview", Neural Networks, vol. 61, pp. 85-117, 2015.

[9] Y. Bengio, A. Courville and P. Vincent, "Representation Learning: A Review and New Perspectives", IEEE Tranctions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1798-1828, 2013.

[10] M. Abadi et al., "TensorFlow: A System for Large-Scale Machine Learning," USENIX Symp. Oper. Syst. Des. Implement. vol. abs/1605.08695, 2016.

[11] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," arXiv:1512.00567, 2015.

[12] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," Adv. Neural Inf. Process. Syst. 27 (Proceedings NIPS), vol. 27, pp. 1–9, 2014.

[13] T. Malisiewicz, A. Gupta, and A. A. Efros, "Ensemble of exemplar-SVMs for object detection and beyond," in Proceedings of the IEEE International Conference on Computer Vision, 2011, pp. 89–96.

[14] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks," arXiv Prepr. arXiv, p. 1312.6229, 2013.

[15] Y. Xu, D. Xu, S. Lin, J. Liu, J. Cheng, C. Cheung, T. Aung and T. Wong, "Sliding Window and Regression Based Cup Detection in Digital Fundus Images for Glaucoma Diagnosis", Lecture Notes in Computer Science, pp. 1-8, 2011.

[16] J. Salamon and J. P. Bello, "Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification," arXiv Prepr. ArXiv1608.04363, 2016.

[17] R. Wu, S. Yan, Y. Shan, Q. Dang, and G. Sun, "Deep Image: Scaling up Image Recognition," Arxiv, p. 12, 2015.

Addresses for correspondence

Allan Cerentini, acerentini@inf.ufsm

Daniel Welfer, welfer@gmail.com

Marcos Cordeiro d'Ornellas, marcosdornellas@gmail.com