Prediction of Glaucoma through Convolutional Neural Networks

Marco A. Espinoza[1], Germán H. Alférez[1], and Javier Castillo[2]
School of Engineering and Technology[1], Instituto de la Visión del Hospital La Carlota[2]
Universidad de Montemorelos
Apartado 16-5 Montemorelos, N.L. Mexico
E-mails: 1120631@alumno.um.edu.mx[1], harveyalferez@um.edu.mx[1], oftalmo@um.edu.mx[2]

***Abstract*—Glaucoma is a disease that damages the optic nerve and becomes chronic as time progresses. Its early detection is vital for treatment. Although there are tools to carry out the analysis of the optic nerve, they do not automatically detect this illness. For this reason, the objective of this research is to carry out experiments with convolutional neuronal networks to achieve the automatic detection of this disease. The experiments performed obtained an average accuracy of 99%.**

***Index Terms*—Glaucoma, convolutional neuronal networks.**

I. Introduction

One of the most important and complex aspects for glaucoma treatment is an early diagnosis to prevent permanent structural damage and irreversible loss of vision [1], [2]. Moreover, the proper diagnosis of glaucoma is key due to the fact that it occurs without symptoms and can lead to irreversible loss of sight [3].

There are tools for taking pictures of the fundus of the eye, such as Visucam NNM/FA[1] and Kevork[2]. These kind of tools are useful for detecting glaucoma by means of Ocular Coherence Tomography (OCT), Heidelberg Retina Tomography II (HRT II), Optical Coherence Tomography (Stratus OCT), and Scanning Laser Polarimetry with Variablecorneal Compensation (GDx VCC), and Confocal Scanning Laser Ophthalmoscopy. However, these tools lack the automatic early detection of glaucoma.

The contribution of this research work is to use convolutional neural networks implemented in TensorFlow for the automatic early detection of glaucoma by means of analyzing pictures of the fundus of the eye. To this end, pictures were taken at the Instituto de la Visión, an ophthalmologic clinic that belongs to Universidad de Montemorelos, Mexico. Within a period of three months, 22 pictures of the fundus of the eye with confirmed diagnosis of glaucoma and 14 pictures of healthy people were taken. These pictures were used for training a convolutional neuronal network. To this end, TensorFlow was used for extending the pre-trained model Inception v3 with the new pictures. The results of the classification were compared with medical diagnosis.

This paper is organized as follows. The second section presents how glaucoma has become a public health problem. The third session presents the state of art. The fourth section presents the underlying concepts of our approach. The fifth section presents the results. The last section presents conclusions and future work.

II. Glaucoma - A Public Health Problem

According to the National Eye Institute (NEI) [4], glaucoma is a group of diseases that can lead to damage to the eye's optic nerve and ends in blindness. Glaucoma usually has no early symptoms. Moreover, by the time people experience problems with their vision, they usually have lost a significant amount of their sight. According to NEI, increased pressure inside the eye is a key cause of open-angle glaucoma.

Since glaucoma has no symptoms, the only way to detect its progression is by conducting studies based on the internal structure of the eye. The risk of blindness depends on the intraocular pressure, the severity of the disease, the age of glaucoma onset, as well as family history of this disease [5].

Glaucoma is a global health problem. In fact, it is the second leading cause of blindness worldwide, affecting approximately 6.7 million people [6]. Several population-based epidemiological studies have reported that more than 50% of glaucoma cases remain undiagnosed even in developed countries [7]. Most importantly, glaucoma is the main cause of irreversible blindness, being cataract the prime reversible cause. Glaucoma is the main cause of irreversible blindness in people older than 60 years. In fact, blindness due to glaucoma can be prevented if treated early [5]. Although this is true, there are no symptoms of glaucoma. Therefore, the early detection is primarily done by observing structural changes in the exploration of the head of the optic nerve.

Over the years, technology has made great progress in the creation and modification of instruments used for different diagnoses within the medical area. In the field of ophthalmology, there is a wide variety of tools for taking pictures that stand out for their image quality, cost, and easy interaction [8]. Nevertheless, these machines lack the necessary software for the automatic detection of glaucoma.

The use of artificial neural networks in recent years has given solutions to a wide variety of pattern recognition problems, such as computer vision, image recognition, and speech recognition [9]. Currently, the most used tool in the

[1]http://www.medicalexpo.es/prod/carl-zeiss-meditec/product-67959-801991.html
[2]http://kevork-instruments.com/productos/nexy

construction of convolutional neural networks is TensorFlow[3]. TensorFlow is an open source software specialized in speech, text, and image recognition. Its flexible architecture allows to implement algorithms in one or more CPUs or GPUs. TensorFlow was developed by Google. Its great impact on artificial intelligence has led companies such as Intel, Nvidia, and Twitter to use this tool [10]. Through the flexibility offered by TensorFlow, models based on previously trained convolutional neural networks can be extended such in the case of the Inception v3 model. Inception v3 has been trained with images from ImageNet [11]. This model can be retrained in order to save computing time in the generation of a new classification model. We believe that by extending the Inception v3 model with pictures of the fundus of the eye, the glaucoma detection process could be automated.

III. State of the Art

Ophthalmologists would like to achieve the detection and diagnosis of glaucoma in the shortest possible time [12]. To this end, in [13] different machine learning algorithms were compared to evaluate the performance of each one of them for detecting glaucoma. In that research work, the authors used multilayer perceptron, support vector machines, among other algorithms. These algorithms had a high performance during classification [13].

In [14] the results of comparing different classifiers for the diagnosis of glaucoma were presented. In that research work, artificial neural networks and other image processing methods were used to compare them with the Stratus OCT that is frequently used for the detection and diagnosis of glaucoma or other eye diseases. The result was successful with a sample of 89 images of eyes of patients with glaucoma using feed-forward backpropagation.

Other research works also used feed-forward backprogation. For example, in [15], pictures of fundus of eye were collected and used to train a feed-forward backpropagation neural network. To this end, Matlab was utilized to process the images and thus obtain the features to be used in training. In [16], the gradient method was used for detecting the presence of glaucoma using an artificial neural network that is trained with the processed images. As a result it was possible to detect the different stages of glaucoma using backpropagation. In [17], the gradient method was used to detect the presence of glaucoma using artificial neural networks. In that research work, photo editing methods were used to eliminate the noise of each image. The artificial neural network was based on feed-forward backpropagation and the mathematical morphology was used for the analysis of the extraction of the features necessary for training.

In [16], an approach based on support vector machines was proposed to detect glaucoma. The objective of that research work was to use pictures of the fundus of the eye processed by techniques such as noise elimination, contrast enhancement, and principal component analysis. By means of analyzing 40 images, an accuracy of 85% was obtained.

[3]https://www.tensorflow.org/about/uses

Although the approaches abover based on feed-forward backpropagation or other classical classifiers are interesting, recent advances in computer vision indicate that the trend has turned to the use of convolutional neural networks for the recognition of patterns in images [18].

IV. Underpinning of Our Approach

In this section the underlying concepts of our approach are presented. The relationship of these concepts is shown in Figure 1.

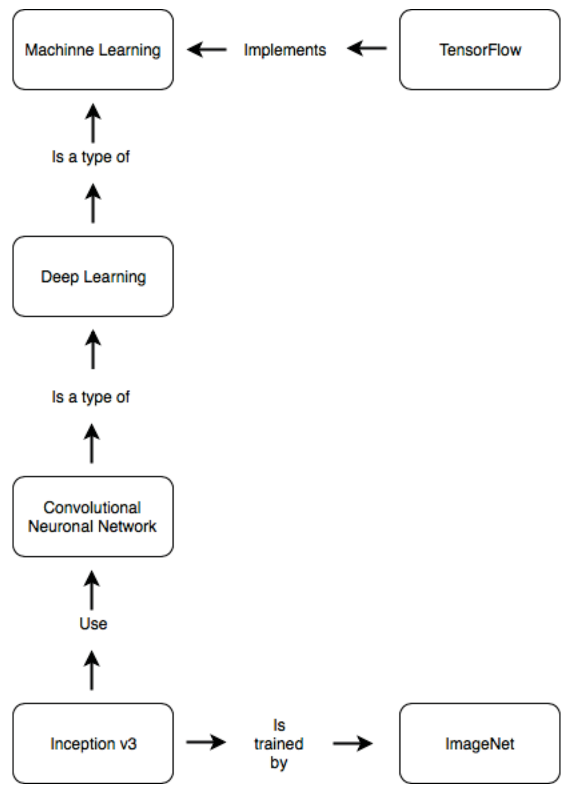


Fig. 1. Conceptual framework.

A. Machine Learning

Machine learning emerged in the 60's as an idea of artificial intelligence but it was until the 80's that it gained importance [19]. Machine learning refers to the practice of training computers through algorithms and mathematical formulas giving the computer the ability to learn and remember from historical data and thus suggest a possible future scenario [20].

The goal of machine learning is to create a model that allows to solve a task. In machine learning, it is desirable that the model is trained using a large amount of data. Through operations and algorithms, the model learns from this data and is able to make predictions. As a result, large companies such as Amazon, Baidu, Google, IBM, Microsoft, and others, use this tool for their business.

There are two types of machine learning approaches: supervised learning and unsupervised learning. On one hand, supervised learning is about training a model from a set of input data with the expected outputs. On the other hand, in unsupervised learning, training is performed with a set of unlabeled input data [21]. This research work is focused on supervised learning.

B. Deep Learning

Deep learning, or better known as deep neural networks, is a branch of artificial intelligence. It emulates the learning approach that human beings use to obtain certain types of knowledge [22]. Deep learning became popular recently when the media headlines showed that Google's AlphaGo beat the world champion of Go, a complex game with thousands of possible combinations [23].

Deep learning allows the processing of computational models composed of multiple layers to learn representations of data with different levels of abstraction. These methods have dramatically improved techniques in image and audio recognition [24].

Deep learning carries out the automatic learning process by means of using an artificial neural network that is composed of a number of hierarchical levels. At the first layer of the hierarchy, the network learns something simple and then sends that information to the next layer. On the second layer, for example, edges are combined by constructing simple shapes, such as a diagonal line or a right angle. In the third layer, the simple forms of the previous layer are combined and more complex objects are obtained like ovals or rectangles. In the next layer, ovals and rectangles could be combined, giving shape to figures in the image [23].

C. Convolutional Neural Networks

Convolutional neural networks are a category of artificial neural networks that have proven to be very efficient in areas such as image recognition, classification, and speech recognition [25]. Convolutional neuronal networks use several identical copies of the same neuron [26]. LeNet was one of the first convolutional neuronal networks that helped drive the field of deep learning since 1990 [27].

CNNs have four main operations [28]: 1) convolution, 2) non-linearity (Relu), 3) pooling or sub-sampling, and 4) classification. These operations act on images that can be represented as a matrix with pixel values [29]. These four operations are explained below:

- *Convolution:* Convolution is a continuous nonlinear function that transforms an input signal into a new output signal. The main purpose of the convolution in the case of the convolutional neuronal network is to extract features from the input image [30]. For example, in Figure IV-C a 5x5 image whose values are only 0s and 1s is considered. Another 3x3 matrix is the filter or kernel. The filter slides over the original matrix from the first pixel, and for each position the multiplication and sum of the elements between the two matrices is calculated. The result is placed in a new matrix. .
  By modifying the filter values, different maps of the same input image are obtained. Figure 3 shows the effects of convolution when applying different filters to an image. Through these filters it is possible to perform operations

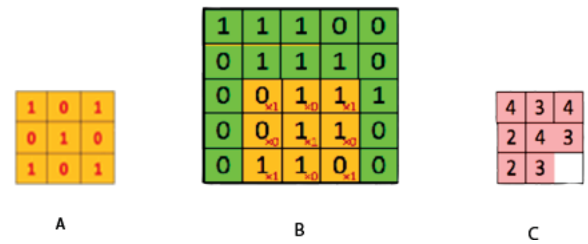


Fig. 2. Filter or kernel (A). Matrix of the image (B). Convolved feature (C) [28].

such as edge detection, sharpness, and blurring. In these operations, the values of the filter matrix are combined before the convolution. This means that different filters can detect different features of an image, for example, edges, curves, lines, etc. [28].

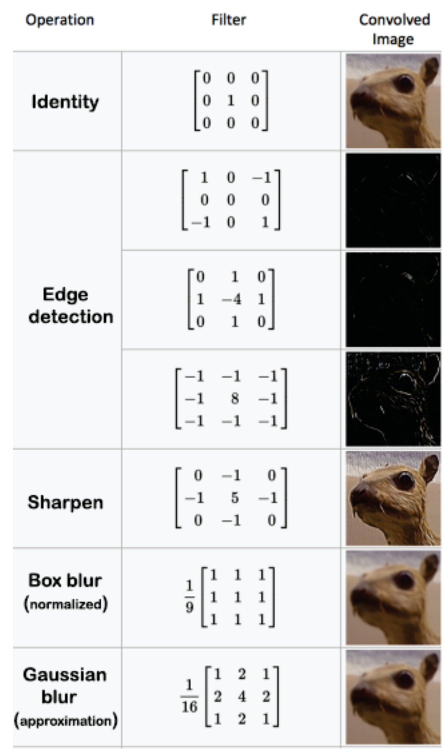


Fig. 3. Different filters [28].

- *Non-linearity (Relu):* Non-linearity is an operation that is applied to each pixel and replaces all negative values of pixels to zeros. The objective of Relu is to introduce non-linearity in the convolution since most of the data that the artificial neuronal network wishes to learn is non-linear [31]. The Relu operation can be seen in Figure 4.
- *Grouping or sub-sampling:* Grouping reduces the dimension of each feature map while retaining the most important information. The grouping can be of different types: maximum value, average value, sum, etc. These functions allow to detect objects inside the image no matter where they are [32]. For example, in the case of max pooling, the largest element of the map features is taken. Figure 5 shows an example with the max pooling

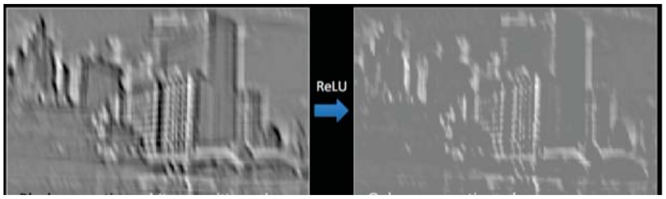


Fig. 4. Operation Relu [28].

operation in a function map [33].

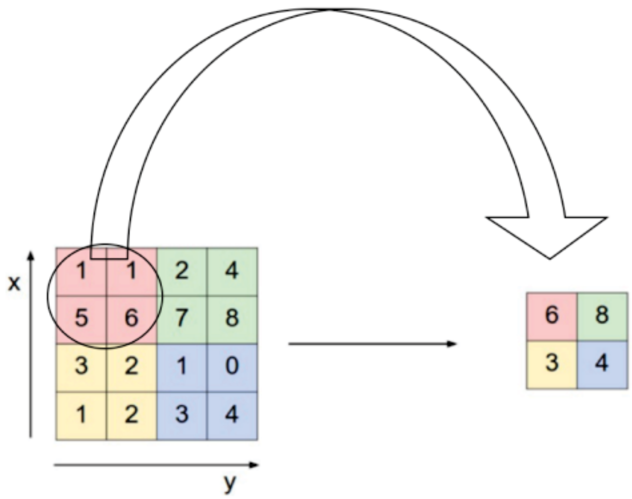


Fig. 5. Max Pooling [28].

- *Fully Connected Layer:* As a whole, the stages of convolution, Relu, and grouping or sub-sampling stages extract useful features, introduce non-linearity in the network, and reduce the dimensions of features to make them smaller and more manageable. Finally, the output of the grouping or sub-sampling layer acts as an input layer to the Fully Connected layer [32]. The objective of this layer is to use the features obtained in previous layers for the classification of the input image [28].

D. TensorFLow

TensorFlow is used in a large number of applications or tools that we use today, such as Google's speech recognition [34]. The TensorFlow API was launched as an open source package in November 2015. TensorFlow uses numerical calculations using data flow graphs. The nodes of the graph represent mathematical operations, while the tensors of the graph handle multidimensional data matrixes to achieve communication between nodes [10], [35]. The value of the tensor can be calculated using Session.run when the graph is executed [36].

E. ImageNet

ImageNet is an image bank organized according to the hierarchy of WordNet. It is a free and accessible resource for researchers and educators around the world [37].

F. Inception v3

Inception-v3 is trained for the ImageNet Large Visual Recognition Challenge using the data from 2012. This is a standard task in computer vision, where models try to classify images into 1,000 classes.
[11].

V. RESULTS

This section presents the methodology that was used in this research work and the results obtained.s

A. Methodology

The activities that were carried out in each of the stages of this research work are described as follows.

1) Preparation of the training environment: In order to train the classification model, 25 pictures of the fundus of the eye were taken from people who have a confirmed diagnosis of glaucoma. Also, 19 pictures of the fundus of the eye were taken from patients who do not show symptoms of glaucoma. These pictures were taken in a period of 3 months at the Instituto de la Visión. All the pictures used in the experiments are available online[4].

To take the pictures of the fundus of the eye of patients with glaucoma, the following procedure was performed:

1) An ophthalmologist confirms that a patient has a clinical diagnosis of glaucoma. To this end, tomographic and campimetria diagnoses were carried out previously.
2) The patient is informed of the importance of taking the picture of his or her optic nerve.
3) The patient signs a medical consent in which he or she accepts that a picture of the fundus of his or her eyes will be taken for research. At this point, the patient is informed about the procedure and future use of the pictures.
4) To take a picture of the fundus of the eye, the pupil of the patient is dilated with phenylephrine eye drops.
5) With the pupil dilated, the picture is taken with the VISUCAM NM / FA machine.

The pictures of the fundus of the eye of patients who do not have a diagnosis of glaucoma were also taken with the VISUCAM NM / FA machine. This machine already had in storage several pictures of healthy patients. As a result, the aforementioned steps did not have to be taken.

Figure V-A1 shows two pictures of the fundus of the eyes of two different patients. On one hand, Section A of this figure shows a picture of the fundus of the eye of a patient with glaucoma. On the other hand, section B shows a picture of the fundus of the eye of a healthy patient. With respect to nerve fibers, the eye at the left is deteriorated.

[4]https://goo.gl/ufjRqP

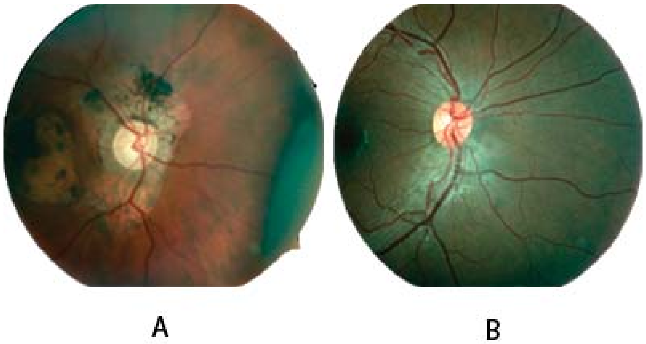


Fig. 6. Pictures of the fundus of the eye: A) with glaucoma and B) without glaucoma.

2) Re-training the Inception v3 convolutional neuronal network : This step consists of re-training the Inception v3 convolutional neuronal network. To this end, a MacBook Air computer with MacOS high Sierra V. 10.13.3 operating system was used. This machine has a 1.6 GHz Intel Core i5 processor and 8 GB 1600 MHz DDR3 of RAM memory. Also, TensorFlow version 1.5.0 was installed and executed. Specifically, the following tools were installed and the following downloads were made:

1) Python version 2.7 was installed from a terminal with the following instruction: "sudo pip install python".
2) TensorFlow was installed through the terminal with the following instruction: "pip install tensorflow".
3) The code label_image.py was downloaded. This code uses the Inception model to classify images that are passed to the model from the command line[5].
4) The code retrain.py was downloaded. This code allows to adapt the pretrained Inception network for other classification problems. A detailed overview of this script can be found online[6].

The command line used to retrained the Inception v3 model is the folloiwng:

sudo python retrain.py --bottleneck_dir /tf_glaucoma/bottlenecks --model_dir /tf_glaucoma/inception --output_graph /tf_glaucoma/retrained_graph.pb --output_labels /tf_glaucoma/retrained_labels.txt --image_dir ~/DataSet/glaucoma

This instruction indicates the script that carries out the retraining (retrain.py) on the Inception model. Also, it indicates the output graph and the labels that are used for training, glaucoma and no glaucoma. The path where the pictures are found is the following: ~/DataSet/glaucoma. In this path, there are two directories. One with the pictures of the fundus of the eye of patients with glaucoma and another with pictures of healthy patients.

From the 44 pictures taken at the Instituto de la Visión, 6 were separated for testing. Specifically, 3 pictures were separated from the 19 pictures without glaucoma and 3 pictures were separated from the 25 pictures with glaucoma. The retraining was done in 8 minutes.

[5]https://goo.gl/hDYKnG
[6]https://goo.gl/b7EcBe

B. Presentation of the Results

The results are shown in Tables V-B and V-B. Table V-B shows the results obtained by entering the 3 pictures of the fundus of the eye with glaucoma to the computer. Each image shows a percentage of accuracy. The evaluations of the three pictures shown to the computer were correctly labeled with an average accuracy of 99%. Although these results are promising, we believe that a better predictive model could be generated by retraining the algorithm using more fundus images.

Table 1. Evaluation of pictures of the fundus of the eye with glaucoma.

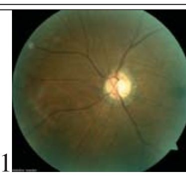
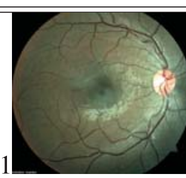
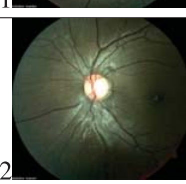
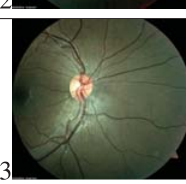
Evaluations	No glaucoma	Glaucoma
1 	0.00246	0..99754
2 	0.00011	0.99989
3 	0.00869	0.99624
Average	0.00375	0.99624

Table V-B shows the results of the three pictures of the fundus of the eye that do not have glaucoma. In these cases, the classification model has an average accuracy of 99%.

Table 2. Evaluation of pictures of the fundus of the eye without glaucoma.

Evaluations	No glaucoma	Glaucoma
1 	0.99762	0.00238
2 	0.99444	0.00556
3 	0.9948	0.00052
Average	0.99718	0.00846

VI. CONCLUSION AND FUTURE WORK

This research work was conducted to prove that convolutional neural networks implemented in TensorFlow can be used for the early automatic detection of glaucoma. To this end, the Inception v3 model was extended with 44 photographs taken at the Instituto de la Visión at Universidad de Montemorelos, Mexico. An evaluation of the classification was carried out with 6 pictures. As a result, the trained model obtained in average a 99% accuracy for classifying glaucoma and no glaucoma.

As future work, although the results in this research work were promising, the sample of pictures of the fundus of the eye of healthy and sick patients will be extended in order to improve the classification. Likewise, it is expected to extend the predictive model so that it does not only detect glaucoma, but also to assess stages of its severity. Also, we expect to extend the model for the early detection of other diseases such as retinal pathology.

REFERENCES

[1] S. Pamela, "Diagnostic tools for glaucoma detection and management," November 1, 2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2643302/

[2] W. Havener, "Importance of early detection of glaucoma," January 16, 2018. [Online]. Available: https://jamanetwork.com/journals/jama/article-abstract/302280?redirect=true

[3] D. Calkins, "Glaucoma," Tech. Rep. 88-207450, 2016. [Online]. Available: https://www.glaucoma.org/GRF_Understanding_Glaucoma_ES.pdf

[4] The National Institutes of Health, "Results–Ocular Hypertension Treatment Study (OHTS)," n.d. [Online]. Available: https://nei.nih.gov/glaucomaeyedrops/ohts_backg

[5] K. Boyd, "What is glaucoma," August 08, 2017. [Online]. Available: https://www.aao.org/salud-ocular/enfermedades/que-es-la-glaucoma

[6] C. Traverso, "Glaucoma the importance of early detection and early treatment," *European Ophthalmic Review*, vol. 13, no. 4, pp. 13–14, 2011. [Online]. Available: http://www.touchophthalmology.com/articles/glaucoma-importance-early-detection-and-early-treatment

[7] Q. Ha, "Number of people with glaucoma worldwide," *The British Journal of Ophthalmology*, vol. 80, no. 5, pp. 389–393, May 1996. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/8695555

[8] C. Zeiss, "Visucam 524 fundus imaging," 2016. [Online]. Available: https://www.zeiss.com/meditec/us/products/ophthalmology-optometry/retina/diagnostics/fundus-imaging/visucam-fundus-imaging.html

[9] C. Shallue, "Inception V3 model," 2018. [Online]. Available: https://www.kaggle.com/google-brain/inception-v3

[10] Google, "TensorFlow," January 27, 2018. [Online]. Available: https://www.tensorflow.org/

[11] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," *CoRR*, vol. abs/1512.00567, 2015. [Online]. Available: http://arxiv.org/abs/1512.00567

[12] E. Oh, "Artificial neural network approach for differentiating open-angle glaucoma from glaucoma suspect without a visual field test," vol. 26, March 14 de 2016.

[13] M. Goldbaum, "Comparison of machine learning and traditional classifiers in glaucoma diagnosis," *Transactions on Biomedical Engineering*, vol. 49, no. 9, 2002.

[14] H. Ling, "Development and comparison of automated classifiers for glaucoma diagnosis using stratus optical coherence tomography," *Investigative ophthalmology & visual science*, vol. 46, no. 11, pp. 4121–4129, 2005.

[15] O. Sheeba, *Glaucoma Detection Using Artificial Neural Network*, IACSIT, Ed. IACSIT, April 2014, vol. 6, no. 2.

[16] D. Abhishek, "Automated glaucoma detection using support vector machine classification method," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 3, no. 1, 2015.

[17] P. Sahithya Sri, P. Chitra, and K. Mahalakshmi, "A survey based on glaucoma detection," *Imperial Journal of Interdisciplinary Research (IJIR)*, vol. 6, pp. 1605–1611, 2015.

[18] H. Eckerson, "Deep learning past, present, and future," may 2017. [Online]. Available: https://www.kdnuggets.com/2017/05/deep-learning-big-deal.html

[19] C. Robert, *Machine learning, a probabilistic perspective*. Taylor & Francis, 2014.

[20] G. Wiederhold, J. McCarthy, and E. Feigenbaum, "Arthur Samuel pioneer in machine learning," *Communications of the ACM*, vol. 33, no. 11, pp. 137–139, 1990.

[21] R. Salas, "Redes neuronales artificiales," *Universidad de Valparaıso. Departamento de Computación*, vol. 1, 2004.

[22] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*. MIT press Cambridge, 2016, vol. 1.

[23] C. Moreno, "Qué es deep learning y para qué sirve?" 2018. [Online]. Available: https://www.indracompany.com/es/blogneo/deep-learning-sirve

[24] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.

[25] A.-H. Ossama, "Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 4277–4280.

[26] C. Olah, "Conv nets a modular perspective," p. 1, July 8, 2014. [Online]. Available: https://colah.github.io/posts/2014-07-Conv-Nets-Modular/

[27] E. Culurciello, "Neural network architectures," Mar. 2017. [Online]. Available: https://towardsdatascience.com/neural-network-architectures-156e5bad51ba

[28] U. Ujjwalkarn, "An intuitive explanation of convolutional neural networks," August 11, 2016. [Online]. Available: https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

[29] A. Geitgey, "Deep learning and convolutional neural networks," Jun 13, 2016. [Online]. Available: https://medium.com/@ageitgey/machine-learning

[30] P. Méndez and J. Ibarra, "Implementación de una red neuronal de convolución para el reconocimiento de poses en imágenes de rostros," *ACI Avances en Ciencias e Ingenierías*, vol. 6, no. 2, 2014.

[31] R. Fergus, "Neural networks," MLSS 2015. [Online]. Available: http://mlss.tuebingen.mpg.de/2015/slides/fergus/Fergus_1.pdf

[32] K. Alex and Sutskever, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[33] J. Jhonson, "Convolutional neural networks for visual recognition," 2017. [Online]. Available: http://cs231n.github.io/convolutional-networks/

[34] S. Pichai, "TensorFlow smarter machine learning, for everyone," Nov. 2015. [Online]. Available: https://googleblog.blogspot.mx/2015/11/tensorflow-smarter-machine-learning-for.html

[35] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, and C. Citro, et al., "TensorFlow large scale machine learning on heterogeneous distributed systems," Google, Tech. Rep., Nov. 2015. [Online]. Available: http://download.tensorflow.org/paper/whitepaper2015.pdf

[36] C. Fredrik, "tf.tensor," January 27, 2018. [Online]. Available: https://www.tensorflow.org/api_docs/python/tf/Tensor

[37] O. Russakovsky, "ImageNet," 2016. [Online]. Available: http://www.image-net.org/about-overview