

Convolutional Neural Networks Model in Premature Detection of Melanoma

Luis Chávez A., YudelyPalpán F., Ciro Rodriguez

Abstract—The objective of this research is provide to the specialists in skin cancer, a premature, rapid and non-invasive diagnosis of melanoma identification, using an image of the lesion, to apply to the treatment of a patient, the method used is the architecture contrast of Convolutional neural networks proposed by Laura Kocobinski of the University of Boston, against our architecture, which reduce the depth of the convolution filter of the last two convolutional layers to obtain maps of more significant characteristics. The performance of the model was reflected in the accuracy during the validation, considering the best result obtained, which is confirmed with the additional data set. The findings found with the application of this base architecture were improved accuracy from 0.79 to 0.83, with 30 epochs, compared to Kocobinski's AlexNet architecture, it was not possible to improve the accuracy of 0.90, however, the complexity of the network played an important role in the results we obtained, which was able to balance and obtain better results without increasing the epochs, the application of our research is very helpful for doctors, since it will allow them to quickly identify if an injury is melanoma or not and consequently treat it efficiently.

Keywords: neural networks, convolutional neural networks, model, melanoma, premature detection.

I. INTRODUCTION

Skin cancer is one of the biggest problems facing modern medicine, so its preventive diagnosis is of great help to the treatment of the patient. Melanoma is a type of skin cancer which represents 4% of the total cases and is a malignant tumor that has its origin in melanocytes (cells responsible for skin color) [1]. They can be attributed as a common cause of the appearance of this type of injury to sun exposure, although other factors such as skin type, genetics or previous lesions also influence [1].

Image processing is the application of a set of mathematical or statistical techniques and algorithms that change the visual appearance or geometric properties of an image or transform it into another type of data set [2]. According to [2], the reasons why an image can be processed are: improve its contrast, correct its geometry and classify the image into several categories. The convolutional neural networks perform this last type of processing.

The convolutional neural networks are a Deep Learning architecture which is inspired by the mechanism of natural perception of living beings. One of the predecessors of this architecture was the neocognitron proposed by Kunihiko Fukushima, which inspired the discovery in 1959 of Huber

and Wiesel that the cells of the animal's visual cortex were responsible for capturing the light. One of its main characteristics is that it can obtain a good representation of the original images, which makes it possible to recognize the necessary patterns to process the image and make the classification [3].

There have been many studies in which this problem was addressed [4] [5] [6], which mainly pose different architectures of convolutional neural networks, highlighting that in the majority they use a new processing approach: multiple convolutional neural networks channels. These propose segmenting the image, be it taking local samples [4], the color channels of the images [6] and the knowledge of expert dermatologists [7].

In this paper we will use a simpler architecture, which will mainly focus on showing the general characteristics of a convolutional network and how you are working. Then, this architecture will be used to classify a set of images of melanocytic and non-melanocytic lesions, using the Python programming language and TensorFlow and Keras libraries to build the model and validate it.

II. METHODOLOGY

In recent years, many investigations have been published addressing the problem of skin cancer, and with special emphasis on the early detection of a melanocytic lesion.

For the extraction of skin lesions caused by melanoma, with non-demographic images, a two-channel convolutional neural network was used [8]. His main motivation was that the preview of moles and malignant lesions is a great help and priority to be able to decide if surgery will be applied to remove them. One of the most important aspects in the computerized evaluation of the lesion is to determine the region of the lesion and separate it into two classes: lesion and normal skin. His contribution was the development of a two-channel convolutional neural network architecture, which allows to classify the pixels of a photo if they are of the lesion or normal skin. The first step was to perform a preprocessing of the images, in which a smoothing was performed to improve the definition of the edges of the lesion and eliminate other defects in the healthy area (such as spots, slight variations in skin color, hairs, among others). Each of these images was taken by segments: one local and one global; then each of them process them in an independent neural network and then combine the outputs of both to make the prediction of the pixel between whether it belongs to the lesion or not. Then with this information a mask of the lesion and healthy skin is generated. This model

Revised Manuscript Received on 14 September, 2019.

Luis Chávez A., Systems Engineering Faculty at National University Mayor de San Marcos, Peru. (E-mail: luis.chavez24@unmsm.edu.pe)

YudelyPalpán F. Systems Engineering Faculty at National University Mayor de San Marcos, Peru. (E-mail: yudely.palpan@unmsm.edu.pe)

Ciro Rodriguez, Systems Engineering Faculty at National University Mayor de San Marcos, Peru. (E-mail: crodriguezro@unmsm.edu.pe)

generated a better separation of these areas than other proposals in the state of the art, resulting in a better segmentation of the area of the lesion.

Another model proposed is to use the HSV channels to be able to determine these regions which are healthy skin and damaged skin, in which each channel is processed with a neural network [6]. Their motivation is to attend an important step in the computer-assisted diagnosis of melanoma, which is difficult since the lesions are different in each of the patients. His contribution was the development of a three-channel architecture, which processes HSV channels in an independent convolutional neural network, but using the same architecture for each channel. As a result, they obtained a good approximation of the segmentation of the lesion, compared with that performed by a dermatologist reaching an accuracy between pixels of 95.3% in validation and 93.4% in tests.

This image processing can also improve the diagnoses of skin diseases, combining convolutional neural networks with human knowledge [5]. Its main contribution was to create a simple model that classifies different diseases, to which human knowledge is added to facilitate decision making. This was achieved by grouping the diagnostic scenarios based on the knowledge of the experts and representing it in a hierarchical structure. First, a review of the labeling of the training images was carried out with a group of up to three experts to reach a consensus on the classification, then a multi-layer convolutional neural network was constructed and finally to represent the decisions based on the knowledge of the experts (who usually use more information than just the photo). The results were a considerable improvement in accuracy by 88.24% using the semantic information of said knowledge.

III. CONVOLUTIONAL NEURAL NETWORKS

Because the problem requires that an image be classified captures a melanocytic lesion or not, it is proposed to use this type of neural network architecture. The main advantage gives us the use of these is to be able to extract the characteristics directly from a photo, which is called a feature map, and making use of these characteristics to distinguish between one class or the other. According to [3], we can distinguish three types of layers in this architecture: Convolutional layer, pooling layer and fully connected layer.

3.1 Convolutional layer

This layer is responsible for learning the representations of the characteristics of the photos. A convolutional layer is composed of several convolution cores or kernels which are used to calculate different feature or activation map. Each neuron of this activation map is connected to neurons within its neighborhood in the anterior layer, and that neighborhood is referred to as the receptive field of the neuron in the anterior layer. A new activation map can be calculated first by performing a convolution at the input with the kernel and using a nonlinear activation function (such as tanh, ReLU and sigmoid) in the results of the convolutional.

The forward or forward pass of this layer consists in carrying out the convolution operation of the input matrix (bytes of an image) with the filter matrix. This operation

consists of traversing the input matrix with the filter from left to right and multiplying in each time the sub-input matrix with the transposition of the filter matrix, which will generate a new matrix which we will call a activation map.

It can be represented as in (1) by the following mathematical expression:

$$U \cong Conv(X, F)$$

Where:

U : Matrix representing the activation map.

X : Matrix representing the bytes of the image.

F : Filter matrix.

$Conv(x, y)$: Convolution operation, y over x .

In addition, when using an activation function as in (2) on the U matrix we have to:

$$O = f(U)$$

In addition, when using an activation function as in (2) on the U matrix we have to:

Where:

O : Output matrix of the convolutional layer.

f : Activation function of the convolutional layer

For retro propagation or backpropagation, it should be analyzed as soon as the error of the output varies respect to the input matrix and respect to the filter matrix.

The variation of the error respect to the input matrix as (3) will allow us to retro propograde the error to the previous layers (if any), because these inputs represent the output of the previous layer (which may be an additional convolutional operation or simply network entry). We can express it with the following mathematical expression:

$$\frac{\partial E}{\partial X} = \frac{\partial E}{\partial O} * \frac{\partial O}{\partial U} * \frac{\partial U}{\partial X}$$

The variation of the error with respect to the filter matrix as (4) will allow to update the matrix, since, these act as the weights of an ordinary multilayer perceptron and will allow the network to extract maps of significant characteristics to be able to perform the subsequent classification. So, it can be expressed with the following mathematical expression:

$$\frac{\partial E}{\partial F} = \frac{\partial E}{\partial O} * \frac{\partial O}{\partial U} * \frac{\partial U}{\partial F}$$

3.2. Pooling layer

Also called the reduction layer and its main purpose is to reduce the change of invariance by applying a reduction in the resolution of the activation map and it is generally in the middle of two convolutional layers. The amount of this reduction is fixed according to a reduction size or pool size.

The forward or forward pass of this layer takes the size of reduction, the input matrix being partitioned in proportion to that size. After these partitions you can obtain the maximum value (Max Pool) or the average (Average Pool). This causes the dimensions of the input matrix to be reduced relative to the size applied.

In the case of the Max Pool operation, the error matrix of the previous layer is taken and the index are found where the highest value for the forward pass was obtained, then a new matrix is created with the same dimensions of the matrix of entries of the layer and it is placed in each one of the

elements of the matrix of errors of the previous layer, according to the positions of the elements that were selected as greater when applying the size of reduction of the layer.

In the case of the Average Pool operation, the derivative of the average function is taken with respect to each of the partitions which were averaged and using the chain rule to find the local errors of the activation map.

3.3. Fully connected layer

This layer is responsible for performing high-level reasoning, after several layers of convolutions and sampling reduction. For this purpose, each of the neurons of previous layer is taken and connected to the neurons of the current layer to generate general information of a semantic nature.

Generally, the first total connection layer receives a matrix of size $m \times n$ as input (from convolution and sample reduction operations), which is transformed into a one dimensional array of dimensions $m \times n$.

The steps forward and the backpropagation of the error are performed as in a conventional multilayer perceptron, by we can see how this layer performs the task of classifying the activation maps.

IV. ARCHITECTURE & RESULTS

The problem of premature melanoma detection can be posed as a classification problem, where it is necessary to determine whether or not a demographic image represents a melanocytic lesion.

The process of implementation, training, validation and analysis of results was performed using the Keras library, which is built on TensorFlow. This library allows us to pre-process the data and create, train and validate our neural network models in a very fast and simple way.

4.1. Dataset

For this research we have used the photos (Figure 1.) from the Melanoma Project of the International Skin Imaging Collaboration ISIC, that is developing proposed digital imaging standards and creating a public archive of clinical and dermoscopic images of skin lesions. This project is focused on facilitating the application of digital images of the skin, in order to reduce melanoma mortality



Figure 1. Photographs of a non-melanocytic lesion (a) vs. a melanocytic lesion (b), both used for network training. (Source: International Skin Imaging Collaboration)

In the medical realms, the use of images for education, diagnosis, telediagnosis, patient triage, mole monitoring, documentation, and coordination of care benefit in three categories of standards that help to insure Quality, Privacy, and Interoperability:

The dataset is in the ISIC repository and it is available by the author, to download and use. It has 2148 images between benign and melanoma lesions. The data were divided into three sets: Training, testing and validation, with the amounts and percentage of all images as showed in Table I.

Table 1. Dataset Division

Sets	Amount	Percentage
Training	1505	70%
Validation	429	20%
Testing	214	10%
Total	2148	100%

4.2. Preprocessing of images

The images were loaded into memory with the "ImageDataGenerator" class, which has many useful functions to extract, normalize and increase images in general. Since our data set was in a directory, we used the "flow_from_directory" method to bring to memory and make the magnifications to the image, in order to reduce the overfitting of the model.

To make use of this function, the data set had to be structured in such a way that each set contains one file per class. Three files as (Figure 2.) were created: Training, Validation and Testing, and each of these folders also contained two other folders: Melanoma and Non-Melanoma, which contained the images belonging to their classes.

For processing through the neural network, the images were also resized to 128 x 128 pixels, in order to reduce the computational cost involved in processing larger images.

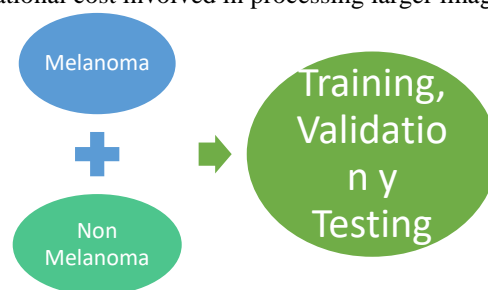


Figure 2. Dataset structure

4.3. Topology

The architecture proposed is a variation of the proposal by Laura Kocobinski of the University of Boston (https://github.com/laurakoco/melanoma_detection_ml), which we have named as "LKNet". The variation reduce the depth of the convolution filter of the last two convolutional layers, in order not to saturate the model with smaller maps

to reach a point of balance in the extraction of the characteristics as in Figure 3.

The network has as input an image of 128 x 128 pixels in RGB format and as output a two-dimensional arrangement, where the first index the arrangement indicates the

probability of belonging to the class "Non Melanoma" and the second, that of belonging to the "Melanoma" class. The complete topology of the network used can be seen in Table 2.

Table 2. Proposed topology with Convolutional neural network

Layer	Dimension	Activation function
Convolution	(32, 3, 3)	ReLU
Max Pooling	(2, 2)	-
Convolution	(64, 3, 3)	ReLU
Max Pooling	(2, 2)	-
Convolution	(128, 3, 3)	ReLU
Max Pooling	(2, 2)	-
Convolution	(256, 3, 3)	ReLU
Max Pooling	(2, 2)	-
Fullyconnected	(256)	ReLU
Dropout	(256)	-
Fullyconnected	(2)	Sigmoid

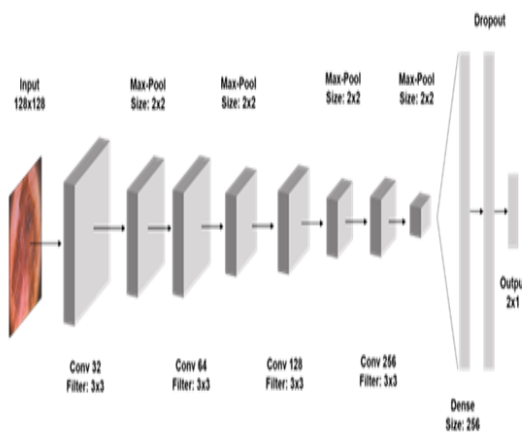


Figure 3. Architecture of the Convolutional Neural Network applied to the case study

4.4. Implementation

function on an input tensor.

The Python programming language was used to implement the model, which is widely used in various works related to Deep Learning and TensorFlow and Keras libraries.

To implement the model, the class "keras.models.Sequential" was used, which allows us to create a sequential model to which different types of layers can be added (Convolutions, Pooling or Fully Connected).

The layers used to build the model were the following:

- *keras.layers.Conv2D*: Creates a convolution kernel, executes the two-dimensional convolution operation and generates an output tensor.
- *keras.layers.MaxPooling2D*: Performs a Max-Pooling operation for an input tensor.
- *keras.layers.Activation*: Executes an activation

- *keras.layers.Dropout*: Dropout operation to an input tensor, randomly choosing a portion of the connections that will be sent to the next layer to prevent overfitting [9].
- *keras.layers.Dense*: Represents a regular total connection layer, with a pre-established number of connections.
- *keras.layers.Flatten*: Performs an operation that converts a tensor into a one-dimensional array between a convolutional or pooling layer and a total connection layer.

4.5. Training

Using the classes and functions defined above we proceeded to train the model.

Epochs: Determine a midpoint of times in which the evaluated algorithms conducted their training and through a trial and error process determine a number of epochs in which the model converge. The amount of times to train the model was 30.

Optimizer: Used as a variation of the Stochastic Gradient Descent algorithm called “Adam”, proposed by Diederik P. Kingma and Jimmy Lei Ba [10]. According to [10], this algorithm has a very low computational cost, as shown in Figure 4(a).

Cost function: Used to train the model was called “Categorical Cross Entropy”, which is a generalization of the cost function “Cross Entropy” as Figure 4(b).

Training environment: The model was trained using the Google Colab platform that allows to create Jupyter Notebooks that run on Google Cloud Engine and provides us a CPU and GPU. The training period required by the model was 1 hour

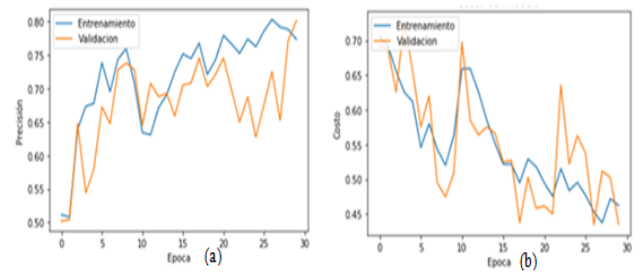


Figure 4. Graph of variation of accuracy (a) with respect to training and validation times, and graph of cost variation (b) with respect to training and validation times

4.6. Validation

For the validation stage, three important aspects of the model were analyzed: The evolution of precision and cost, the confusion matrix and the classification report.

Evolution of Precision and Cost: During the training of the model, progress was made to improve the accuracy of validation and the reduction of the cost function. The graph summarizes both data. Confusion matrix and classification report: In order to evaluate the model, the best result was saved, for the “val_acc” metric, also to evaluate validation accuracy of the “ModelCheckpoint” class, which allows us monitoring the metrics of the best saved model results. To obtain the confusion matrix, as in Table 3., and the classification report, Table 4., we will load the saved model to predict with the “fit_generator” method, using the dataset of 214 test images

Table 3. Confusion matrix

Predicted values	Real values	
	Non melanoma	Melanoma

(a)

(b)

Non melanoma	99	7
Melanoma	32	76

Table 4. Classification report

	Accuracy	Thoroughness	F value	Support
Non melanoma	0.76	0.93	0.84	106
Melanoma	0.92	0.70	0.80	108
Precision			0.82	214

V. ANALYSIS OF RESULTS

The model compared to the base model obtained a better accuracy from 0.79 to 0.83. It is also important to note that when testing our proposed model, more training times were required to achieve this accuracy, which is due to the fact

that the data set was not complete. In the literature review the implementation of AlexNet, shows it is a network architecture that is widely used for the classification of



images in n classes, which obtains an accuracy of 90.2 % with 20 epochs, as shown in Table 5

Table 5. Result comparison of evaluated models

Model	Accuracy	Epochs
LKNet	0.788	30
AlexNet	0.902	20
Proposed	0.834	30

VI. CONCLUSIONS

This research applies the concepts of convolutional neural networks and to use tools and datasets for Python with TensorFlow and Keras, which provide high-level APIs, in addition, the model contributes to the solution of diverse kinds of problems.

VII. FURTHER WORKS

For further works, we will use the proposed architecture model to improve it, so that it provides greater accuracy at the time of classification, we will apply the CNN architecture in other problems involving image processing, such as the determination of the age of an ecchymosis or in the detection of factors that causes malaria.

VIII. ACKNOWLEDGEMENTS

We thank our families and class friends who encourage us to be better every day as well as the professors of our professional training, which allows us to research such as the one contained in this paper

REFERENCES

1. B. Martínez-AmoresMartínez, F. Vicente Martín, M. DuránPoveda y R. Molina Villaverde, «Melanoma» *Medicine - Programa de FormaciónMédicaContinuadaAcreditado*, vol. 12, n° 33, pp. 1980-1989, 2017.
2. J. S. Aber, I. Marzloff y J. B. Ries, *Image Processing and Analysis*, Elsevier, 2010, pp. 159-181.
3. J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai y T. Chen, «Recent advances in convolutional neural networks, » *Pattern Recognition*, vol. 77, pp. 354-377, 2018.
4. M. H. Jafari, E. Nasr-Esfahani, N. Karimi, S. M. R. Soroushmehr, S. Samavi y K. Najarian, «Extraction of skin lesions from non-dermoscopic images for surgical excision of melanoma,» *International Journal of Computer Assisted Radiology and Surgery*, pp. 1021-1030, 2017.
5. X. Zhang, S. Wang, J. Liu y C. Tao, «Towards improving diagnosis of skin diseases by combining deep neural network and human knowledge» *BMC Medical Informatics and Decision Making*, vol. 18, n° 2, 2018.
6. Y. Yuan y Y.-C. Lo, «Improving Dermoscopic Image Segmentation With Enhanced Convolutional-Deconvolutional Networks» *IEEE Journal of Biomedical and Health Informatics*, vol. 23, n° 2, pp. 519-526, 2019.
7. Walker, B.N., J. Rehg, A. Kalra, R. Winters, P. Drews, J. Dascalu, E. David y A. Dascalu, «Dermoscopy diagnosis of

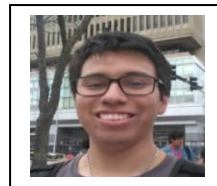
cancerous lesions utilizing dual deep learning algorithms via visual and audio (sonification) outputs: *Laboratory and prospective observational studies» EBioMedicine*, vol. 40, pp. 176-183, 2019.

8. M. H. Jafari, E. Nasr-Esfahani, N. Karimi, S. M. R. Soroushmehr, S. Samavi y K. Najarian, «Extraction of skin lesions from non-dermoscopic images for surgical excision of melanoma» *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, n° 6, pp. 1021-1030, 2017.
9. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever y R. Salakhutdinov, «Dropout: A Simple Way to Prevent Neural Networks» *Journal of Machine Learning Research*, vol. 15, pp. 1929-1958, 2014.
10. D. B. J. Kingma, «Adam: A Method for Stochastic Optimization» *ICLR*, 2015.
11. A. Hosking, B. J. Coakley, D. Chang, F. Talebi-Liasi, S. Lish, S. W. Lee y A. M. Zong, «Hyperspectral imaging in automated digital dermoscopy screening for melanoma» *Lasers in Surgery and Medicine*, vol. 51, n° 3, pp. 214-222, 2019

AUTHORS

Luis Ricardo Chávez Aliaga

Member of the research group of the Faculty of Systems Engineering at the National University Mayor de San Marcos, in the areas of Artificial Intelligence, Data Science, Microservices, Cloud Computing and Convolutional Neural Networks.



Yudely Guadalupe Palpán Flores

Member of the research group of the Faculty of Systems Engineering at the National University of San Marcos, in the areas of Artificial Intelligence, Image Processing, Data Science, UI / UX Design and Convolutional Neural Networks



Ciro Rodriguez Rodriguez

Professor at the School of Software Engineering at the National University Mayor de San Marcos, and at the Computer Science School and Graduate School of the National University Federico Villarreal, with science studies at the Abdus Salam International Center



for Theoretical Physics (ICTP) and the United States Particle Accelerator School (USPAS).