Convolutional Neural Network

Ana M. Velosa Orduz
Universidad de los Andes
am.velosa@uniandes.edu.co

Abstract

1. Introduction

Nowadays you can find neural networks to solve many current problems such as automatic machine translations, object calcification in images, segmentation of objects in video or images, among others. The algorithms that best solve the problems that were previously mentioned are considered.[1]

These neural networks are divided into different layers, which have different characteristics that enrich the model. One of them is the convolutional layer, this layer contains the learnable kernels that are transferred from layer to layer. Another layer that can be found in neural networks is a non-linear layer, this layer consists of activation functions and they are faster than linear ones. The pooling layers reduce the spatial size of the representation that comes from the convolutional layers by means of an average or maximization. [1]

Finally there are the fully connected layers, these connect all the neurons of one of the layers with the neuroses of the next to be able to incorporate all the possible information.[1]
Especially important, Convolutional Neural Network (CNN), is a Deep Learning algorithm that take input images to which it assigns learnable weights and biases to aspects and objects that later on will define the difference between one image and the other. [3] This algorithm also combine the pre-processing of the image with the algorithm that it wants to implement depending on the input images that are placed on the convolutional neuronal network. In the pre-processing the Network learns the filters and the characteristics that are needed to have a better performance in terms of image classification or segmentation.[3]

2. Methodology

The CelebA database was used, which consists of more than 200K celebrity images with annotations of 40 attributes each. In this particular case we only worked with 10 of those attributes, which correspond to, 'Eyeglasses', 'Bangs', 'Black Hair', 'Blond Hair', 'Brown Hair', 'Gray Hair', 'Male', 'Pale Skin' , 'Smiling', 'Young' as we can see 1.[2]
For the resolution of this problem the images had to be organized which were all in a folder and their correspondence to the grian trian, val and test, as their labels were in files in txt format, two different architectures of CNN were implemented. Both maded a transformation which consisted in transforming the image into a gray scale.One of the CNN made an espacial tranformacion in orther to keep the proportion of the input of the arquitecture of the CNN. As for the architectures that were made one corresponds to the one taken from AleNet and the second one with a completely new architecture. This new architecture consists of two convolutional layers divided by a pooling layer and finally 3 fully connected layered.

Figure 1.

Once you have these layers, you continued to train the model and test the accuracy with which both architectures were behaving.

3. Results and Discussion

Once both architectures were tested it was found that a very detailed organization of the data had to be done so that they could be read and also that even if the data were in a

fromato torch they could not be entered into the network to be trained .

For this reason we tried to perform the same test with another MNIST database which is in the torch library but presents problems with dimensions which are not consistent with the calculations of dimesionaldiades made before building the network that does not correspond to AlexNet. For this reason it was decided to make the change to gray levels of all the images but then there is a pooling error to which no solution was found because in the docuemtacion that pooling function is found but the error is presented. After this, the architect of AlexNet was tested, hoping that the problems related to pooling if problems of dimensionality were encountered, what was tried to be solved by means of a transformation to ensure that the images entered with dimensions of 224x224. Followed by this, the gray-scale transformation was eliminated to observe if there was any change, but the dimsions remained unclassified.

4. Conclusions

It is important to emphasize that although networks are of great help when it comes to solving complex problems, they require great mathematical knowledge to be implemented. In addition to the above, it is necessary to have a wide computational knowledge to be able to solve the problems that are observed throughout the construction of the networks.

References

[1] A. Bhandare, M. Bhide, P. Gokhale, and R. Chandavarkar. Applications of convolutional neural networks. *International Journal of Computer Science and Information Technologies*, 7(5):2206–2215, 2016. 1

[2] Z. Liu, P. Luo, X. Wang, and X. Tang. Deep learning face attributes in the wild. In *Proceedings of International Conference on Computer Vision (ICCV)*, 2015. 1

[3] S. Saha. A comprehensive guide to convolutional neural networks—the eli5 way, 2018. 1