Department of Economics, Management and Quantitative Methods
Università degli Studi di Milano
Data Science and Economics

ALGORITHMS FOR MASSIVE DATA MODULE

Comics faces

Shaikyp Ulan (944462)

September 18, 2022

Abstract

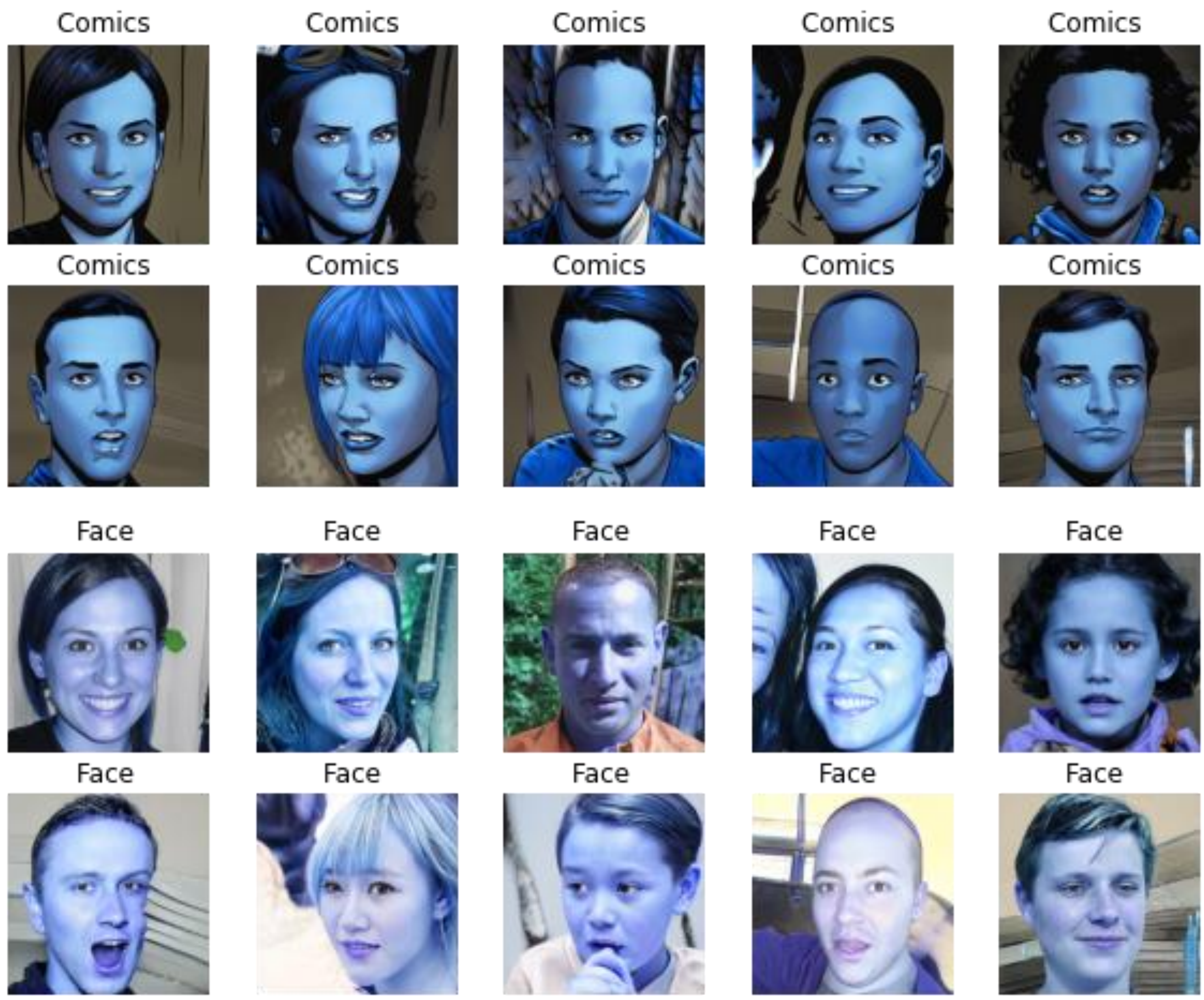
The goal of this project is to develop a deep-learning-based system to discriminate between real faces and comics, using the "Comics faces" dataset available on Kaggle. For this task I explore the Convolutional Neural Networks (CNN), which are largely used for image classification.

Introduction

The performance of well-known convolutional neural networks (CNNs) for categorizing images into two classes—Real Face Pictures and Comic Pictures—is empirically examined in this research. This experimental project's main objective is to train a computer to recognize image categories using Deep Learning methods.

Dataset

The "Comic faces" dataset is published on Kaggle and released under the CC-BY 4.0 license, with attribution required. It contains pictures organized in two folders. One containing 10'000 pictures of real images. The other one containing 10'000 corresponding drawings of the real images, that are the comics. For a total of 20'000 images in 1024x1024 format.



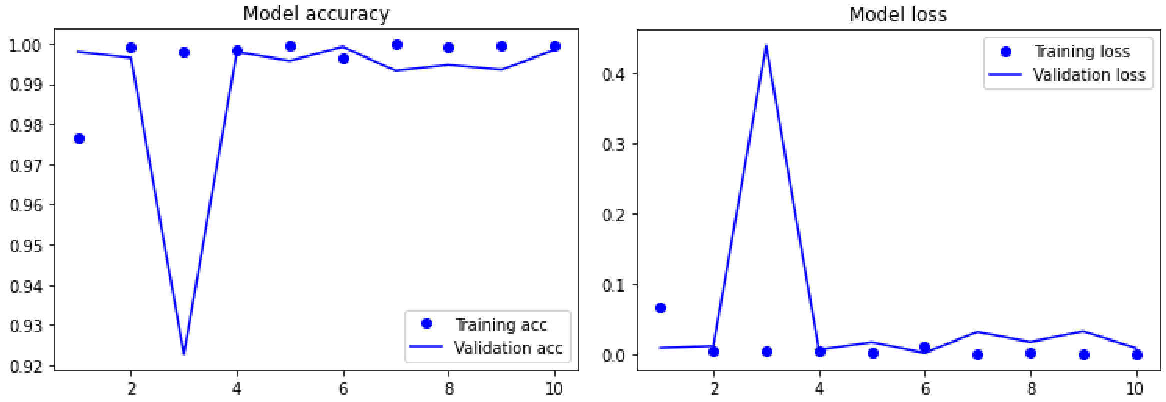
Methodology

A Convolutional Neural Network is a Deep Learning algorithm which can take an image as input, detect its main characteristics and distinguish it from others. Hence, the algorithm reduces the images without losing critical features for a good prediction. A layer in the CNN is composed by a “convolutional layer” and a “pooling layer”. Where the number of layers defines the level of details captured. The higher is this number, the higher is the complexity and the more computational power is required. The convolutional operation takes place in the Convolutional Layer and involves the “filter” element. This shifts through the image until it is considered in its entirety. The number of times this filter shifts depends on the “stride length”. The objective is to extract high-level features from the image. Actually, the first convolutional layer captures the low-level features. The high-level features are captured adding more layers, which define the network architecture. On the other hand, the Pooling Layer reduces the dimensionality of the convolved features, helping in decrease the computational power required. Also, it extracts the dominant features. It is a “max pooling” if it returns the maximum value of the filter’s window. Or it is an “average pooling” if it returns the average of the values of the filter’s window. Generally, max pooling performs better. Once the model can understand the features, the image must be flattened into a column vector in order to be passed to a “Dense layer”, also called “fully-connected layer”. Each neuron of this layer receives input from all the neurons of previous layer. This is given as input to a neural network able to perform the classification using a Softmax function. This technique has two interesting results: the obtained values sum up to 1 and they are all non-negative. This means that we can interpretate them as a probability distribution over the classes. In our case, the probability of being a real face image or a comics image.

Model.

Convolution Neural Networks (CNNs) are a class of Neural Networks that have excelled in the areas of image recognition, processing, and classification, and I have employed them in my experimental project. In order to determine training weights for the CNN model and to evaluate its performance, training data is necessary. As shown in the following figure, convolution layers with kernels are used to process each input image: The Convolution + MaxPooling layers extract information from the input image, while a fully connected layer serves as a classifier. The network in the example above predicts which class a picture belongs to when it receives it as input and gives it the highest probability. Additionally, "same padding" is the sort of padding employed here. Furthermore, "same padding" is the sort of padding employed here, which essentially means that the output image is identical to the input image. Max pooling, which only extracts features with the highest value, was used for the pooling layer. Finally, in Fully Connected Layer, the sigmoid function is used for the commuting class of input picture, and each node is connected to every other node in the next layer. In conclusion, the three convolution blocks (tf.keras.layers) that make up our sequential model. Conv2D), each of which contains a max pooling layer (tf.keras.layers.MaxPooling2D). There is a layer that is entirely connected (tf.keras.layers. Dense), which is activated by a ReLU activation function ('relu'), has 128 units on top of it. Furthermore, before training the model with augmented photos another strategy to avoid overfitting 'Dropout' regularization was used to the network. When we apply Dropout to a layer, it randomly removes some of the layer's output units during training by setting the activation to zero. Dropout accepts fractional numbers in the form of 0.1, 0.2, 0.4, etc. as its input value. By doing this, 10%, 20%, or 40% of the output units from the applied layer are dropped at random.

Results.
We assess performance using both the training and validation sets as a result (evaluation metric is accuracy). The results show that the validation accuracy is 99,4% and the training accuracy is 99,6%. As seen in the line chart above, for the first 10 epochs, our accuracy was nearly identical, with the exception of a 0.015% discrepancy between training and validation accuracy in the fourth epoch.



When we predicted using random comic and real face images our model forecasted categories with high accuracy of 100%.

Declaration

I declare that this material, which I now submit for assessment, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work. I understand that plagiarism, collusion, and copying are grave and serious offences in the university and accept the penalties that would be imposed should I engage in plagiarism, collusion or copying. This assignment, or any part of it, has not been previously submitted by me or any other person for assessment on this or any other course of study.