Deep Learning Project – Project 2

Course Deep Learning

Professor:

Andrea Asperti, Tutor …..

Students:

Lorenzo Visentini, Master Degree in Management Engineering, ID: 0000….

Gianmarco Bolcato, Master Degree in Management Engineering, ID: 0000…

**Hardware**

The hardware we used for this project was:

*Lorenzo base hardware*

Laptop Acer Aspire …

CPU Intel I7-7700U

GPU Intel

RAM 8 GB

SSD Samsung Evo 1 TB

*Gianmarco base hardware*

Laptop …..

CPU ……

GPU ……

RAM …..

HDD ……

*Colab hardware*

In particular, due to the low computational power of our base hardwares we decided to use Google Colab, which notebooks allow you to combine executable code and rich text in a single document.

The strength of this platform is to create your code as notebook, with cells that can be run once a time. At the same time, Colab offers specific settings including the possibility to run the code on a gpu hosted runtime

Specs:

..

.

.

.

**Software**

*Python*

Programming language at high level, very used in the Artificial Intelligence (Machine Learning and Deep Learning) field, thanks to the import of specific libraries.

In our project, we used the *Keras* library for everything related to models and KPI.

*Google Colab*

Describe early in this paper. Employed, as said before, in order to overcome all the hardware difficulties.

*GitHub and GitHub Desktop*

GitHub is a hosting service for software projects. It offers the possibility to work on repository that can be modified by all the members of the group with authorization.

We used GitHub Desktop as the desktop application of GitHub. The advantage is an easy and complete interface.

**Problem**

The problem is divided in two sub-applications that are related to the same starting dataset, very know in the Deep Learning field, called Celeba Dataset.

The Celeba Dataset is composed of more than 200’000 images representing celebrity face profiles. This dataset is great for training and testing models for face detection, particularly for recognizing facial attributes such as finding people with brown hair, are smiling, or wearing glasses. Images cover large pose variations, background clutter, diverse people, supported by a large quantity of images and rich annotations. This data was originally collected by researchers at MMLAB, The Chinese University of Hong Kong.

*Face Orientation*

The first request is to code a classifier which aims to classify images taken as input based on the orientation of the faces (i.e. left, center, right).

*Source of light*

The second request is related to build a classifier which recognize where the source of light is in the photo (i.e. left, center, right).

**Steps adopted**

Firstly, before adopting the strategy “try, observe and refine”, we spent time building up the whole structure.

The approach is the same for both the problems.

* Manual classification of the Celeba Dataset
* Focus on the transfer learning models and then finetune them
* Focus on the ‘From scratch’ model and following finetuning
* Test the best transfer learning model and the from scratch code
* Analyse the results and final comparation

The preparation of the datasets took a lot of time, a very onerous task.

Specifically, we built a python code to put the files in distinct folders. Strictly connected to the requests of the project, the folders were 3, all of these will be use for both the problems as we highlighted that the best composition of the starting dataset is divide manually the images in center, left and right.

With the other group, we decided to divide the task and ended to classify pictures based on source light orientation. 10’000 images for each dataset, orient and source light.

The light, especially in images with medium-low resolution, is not easy to detect as the reflection can create issues about where the source is. Adding to this, lights sometimes provoked misunderstanding due to noise of the background and artificial setup made by photographers.

The importance of the input of a convolutional neural network is very high and consequently it’s indispensable to clean data as well as perform very accurate manual classification. And since detecting the source light is not an exclusive data task – light can come from different position and, as said before, reflections can cause some issues – we adopted rules enhancing the robustness of this stage.

Firstly, the classes decided are only three, avoiding mixed classes such as center-left, center-right. The division is led by the portion of the image – or portion of the face - that has mostly of the light.

Secondly, some transformation to Celeba images has been performed: HSV (Hue, Saturation, Value) color model which describes colors in terms of their shade, pixel-value based threshold pointing out only pixels’ values that are high than 170 – founded by trying – and traditional image.

Example:

Immagine che contiene fotografia, persona, posando, guardando

Descrizione generata automaticamente

Hence, we applied some operations similar to a preprocessing stage.

During the manual classification many pictures have been removed since the human categorization was nearly impossible, despite the efforts.

Transfer Learning

What is it

Advantages and disadvantages

InceptionV3 (why, tree, layers added, plots, tables with accuracy)

ResNet50 (why, tree, layers added, plots, tables with accuracy)

Best TL\_CNN accuracy on whole dataset and csv predictions

FromScratch

Advantages and disadvantages

Tree

Why used some types of layers

Results on train and test with plots and accuracy

Results on whole dataset

Final considerations

Bibliography and sitography