

SIMPSONS CHARACTER IDENTIFIER CONVOLUTIONAL NEURAL NETWORK

Luciano Orlandini – 460952

ETI3v.IA

Applied Computer Science

Tutor: Mr Dawid Zalewski

December 2020

Content

1 Introduction	3
2 Project Set Up	4
Data-Set	4
Software	5
3 The Work	5
4 Conclusion	6

1 Introduction

For this Machine Learning assignment, I have chosen to create an applied project, using the combined knowledge gathered over the first three assignments.

I was genuinely interested in the examples followed, particularly on the image identification. The example of identifying traffic signs was eye opening as it is very useful to be able to teach a computer to read signs and then take actions with the results.

I was hoping to be able to expand on that, hence I went to www.kaggle.com to look for databases and came across the one used in this project.

A data-set containing over 20 thousand images of The Simpsons characters sparked my interest as I grew up watching the show, hence I created a Convolutional Neural Network model that would be able to identify varied characters from the show.

The original data-set can be found on the link below for reference:

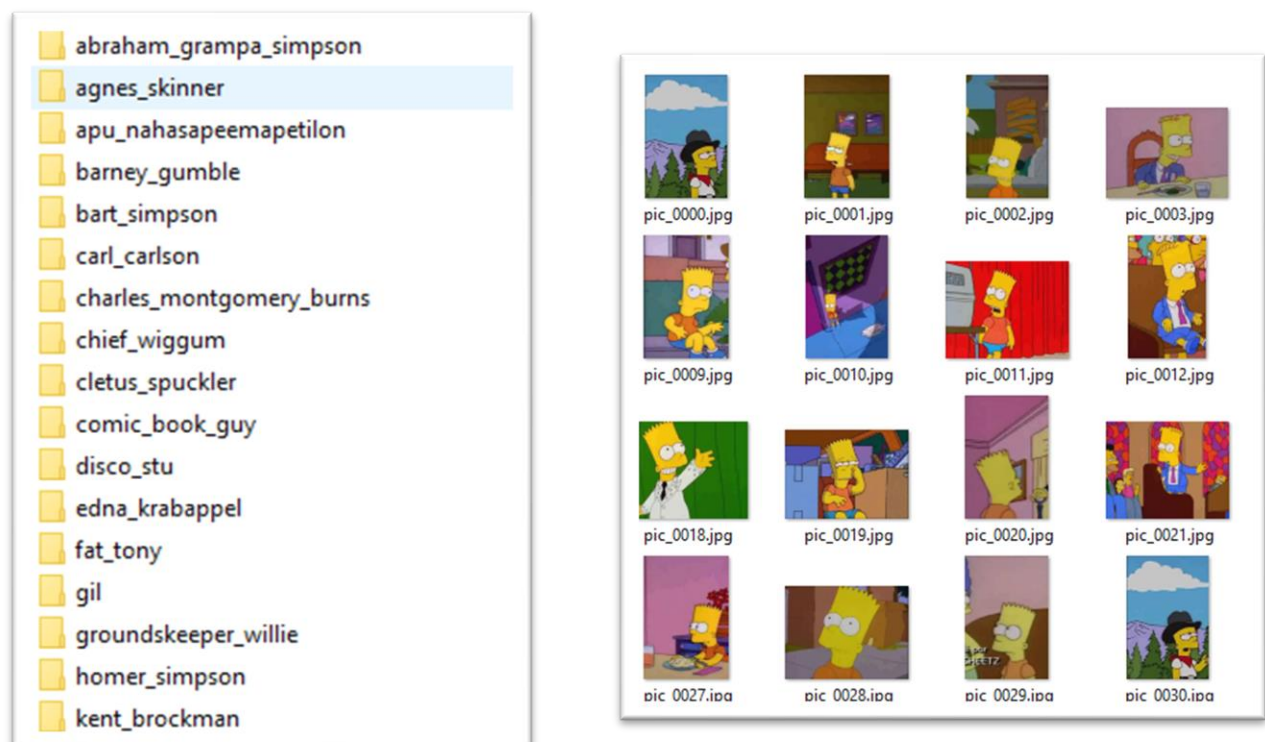
<https://www.kaggle.com/alexattia/the-simpsons-characters-dataset>

2 Project Set Up

Data-Set

The Simpsons character data-set was divided in the following way:

- A training directory containing 42 folders, each named after a different character from the show. Within each folder a large selection of images of such character. In total there were over 20 thousand images of slightly different dimensions. The images below show examples of the directory structure.



- A directory with 990 mixed images of the different characters of the show.






Software

For this project I've created a virtual environment in my Windows 10 via the command prompt. In this environment I've installed Python 3, the required libraries for the project and Jupyter Notebook.

For TensorFlow I managed to configure it to make use of my nVidia GPU, which has drastically improved the computational speed when comparing to using the CPU instead.

3 The Work

I have submitted my Jupyter Notebook file including the saved model and weights. Within the Jupyter Notebook file there are plenty of comments and explanations on the how the programme works, hence I will not include any extra explanation in this report.

 `model_keras.h5`
 `model_weights.h5`
 `Simpsons.ipynb`

4 Conclusion

This project was very enjoyable to work with and I navigated through all previous assignments in this course as well as through the internet on various sites.

I've added my own twist on how the programme is presented and looked up all different functions used so that I could tweak the parameters to further improve it.

I ended up running this many times in order to trial different parameters and managed to get an accuracy between 90 and 95%.

At the end, where you can see the images that were identified, I noticed that the programme sometimes gets confused with certain characters as they are actually quite similarly drawn, as for example Nelson Muntz and Barney Gumble, or Marge and her sisters, but overall, I was quite happy with the result.