Computer Vision and Deep Learning
Object detection and instance segmentation with Mask R-CNN.

Sagaf Youssouf

Université Paris-Est Créteil (UPEC)
International Master of Biometrics and Intelligent Vision
https://www.international-master-biometrics-intelligent-vision.org/

1 Introduction

The objective of this lab is to play with the deep neural network Mask R-CNN to solve instance segmentation problems in machine learning. The framework will separate an object from the background. For us, we chose the face mask as the object. We will give it an image and the output is an object bounding box, classes, and masks. For this, we used the original implementation : https://github.com/matterport/Mask_RCNN

2 Dataset preparation

The dataset used is composed of 60 images for training and 12 images for validation. When we started, we downloaded a dataset composed of thousands of images in Kaggle: face-mask-detection. Some images were chosen to constitute our own dataset. Unfortunately, the first training results were not encouraging. The loss started at 4. We thought that this was due to the great diversity of images and different masks used. To have good performances, we decided to build our own dataset. Our dataset is made up of images of politicians wearing masks (Emmanuel Macron, Angela Merkel, Olivier Veran, Edouard Philippe, Jean Castex and Sibethe Ndiaye). For each person, there are 10 images for training and 2 images for validation. To avoid bias, we put a black personality and women. The annotation was done using the open source VGG image annotation tool (via). Some images of the dataset are given in the following figure:

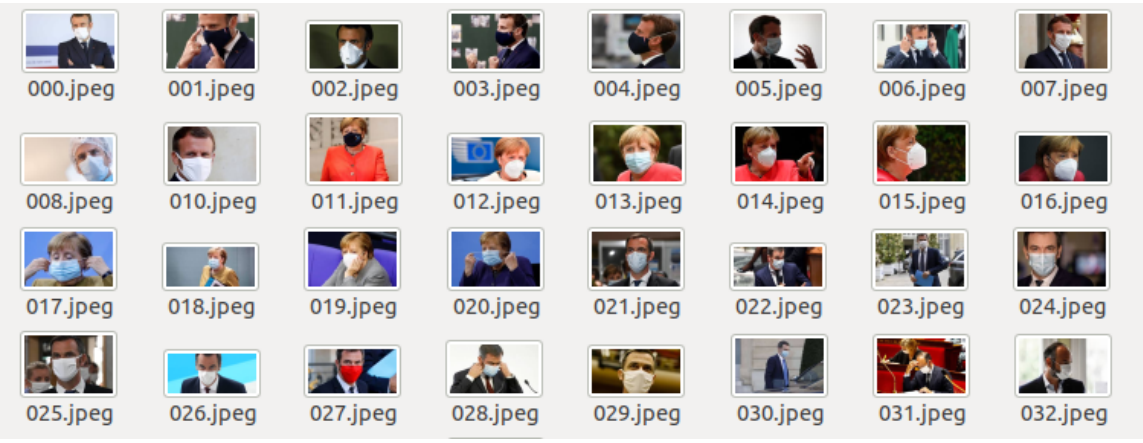

Fig. 1: Examples of images used for training.

3 Results et discussions

To understand the behavior of the model, we chose to keep the default parameters of the original code first. The number of epochs was 30, and the variable STEPS_PER_EPOCH equaled to 100. The loss starts at 0.47 and kept decreasing. This step took at least 5 hours. There was no sign of overfitting. In google colab, we are limited to 12 hours of free session. So we decided to let the training continue until 50 epochs. The loss curve is given by the following figure 3. We notice that the curve indicates the ideal shape by approaching an asymptote. Final loss at 50 epoch is 0.0035. The next two images are the predictions results obtained from images that the model has never seen. The model can easily detect the masks with high precision and the shape of the mask obtained covers the entire mask. However, on the second image, the face mask of Macron could not be detected. This is due to the fact that, we chose to skip the detections with a confidence less than 90% and the dataset doesn't contain a lot of images with masks on the side part as in the image.

Fig. 2: Results of Mask R-CNN on test images, using ResNet-101-FPN

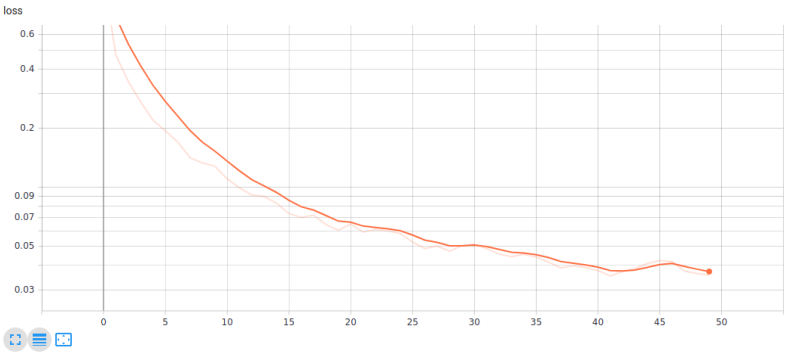

Fig. 3: Graph of loss function

4 Conclusion

We trained a model for face mask detection and segmentation. The good performances obtained are largely due to the work done in the preparation of the dataset. In this period of sanitary crisis, we can imagine many applications of face mask detection using this technique like in clasroom or in railway station.