Project Plan

Applied Deep Learning

Lukas Prenner - 01630142

Description of Project

The aim of the project is to implement an existing deep learning model which should be able to successfully classify the faces of two persons which have recently been discussed in the Austrian media rather extensively: Sebastian Kurz and Alexander Schallenberg. The application to use the model will be an online application with a simple interface where users will be able to upload an image of either one of the two individuals and get an answer of which of the two it actually is - including the probability of the prediction. (see Figure 1)

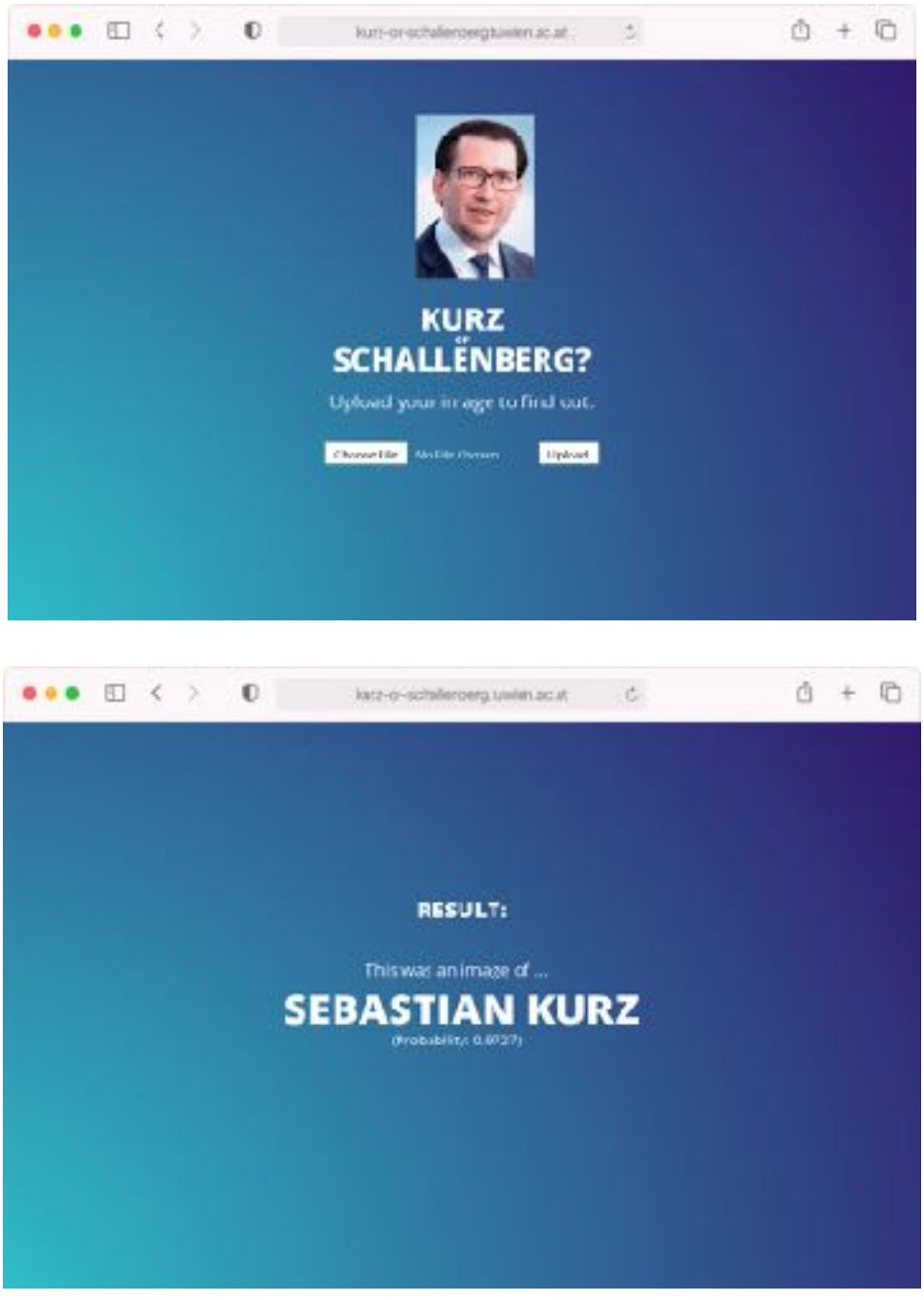

Figure 1: Mockups of the planned web application

Inspired by Balaban, Stephen. (2015)[1], the FaceNet model (proposed by Google in 2015) will be used for the classification of the faces seen on the input images. In order to first locate the images on the given photos, MTCNN (Multi-Task Cascaded Convolutional Neural Network) will be used - as presented in Zhang, Kaipeng et al. (2016)[2].

Dataset
In order to create a raw data set with at least 100 images per class, photographs of Sebastian Kurz and Alexander Schallenberg will be crawled and extracted from the internet - e.g. thumbnails from websites such as GettyImages. As deep learning models usually need a lot of data, the training data will be augmented by pre-processing steps such as image rotation, image scaling, image blurring or image stretching. Depending on the performance of the model using the data, the pre-processing pipeline could be extended using approaches such as eigenface and fisherface as discussed in Tan, Xiaoyang et al. (2006)[3].

Work-Breakdown Structure

TASK	TASK DESCRIPTION	TIME ESTIMATE
Data Collection	Crawling and extracting >100 images of Sebastian Kurz and Alexander Schallenberg. Cropping part of the images manually if more than target person is on the image	3-5 hours
First Implementation	Implementing pipeline including pre-processing, MTCNN, FaceNet, training and testing in Google Colab environment	10-16 hours
Training and Tuning	Running pipeline for collected dataset with different configurations and parameters, evaluating performance indicators such as accuracy and precision; optionally adjusting (hyper-)parameters, pre-processing steps as well as training data	8-16 hours
Online Application	Building online application to present results: design a simple interface, integrate trained model and test results	10-14 hours
Final Report	Planning, writing and layouting final report to summarise approach, learnings, setbacks and final results	3-6 hours
Preparing Presentation	Planning, designing and building slides for final presentation; creating script and notes	3-4 hours
Presentation	Performing final presentation in front of class	0.25 hours

Repository
The code will be stored and published via a public repository on GitHub which can be accessed using the following link: https://github.com/LukasPrenner/applied-deep-learning

[1] Balaban, Stephen. (2015). Deep learning and face recognition: the state of the art. 94570B. 10.1117/12.2181526.

[2] Zhang, Kaipeng & Zhang, Zhanpeng & Li, Zhifeng & Qiao, Yu. (2016). Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. IEEE Signal Processing Letters. 23. 10.1109/LSP.2016.2603342.

[3] Tan, Xiaoyang et al. "Face recognition from a single image per person: A survey." Pattern Recognit. 39 (2006): 1725-1745.