

Face Generation using Generative Adversarial Networks

Alejandro Woodward
Genís Floriach

Irina Sánchez
Oriol Vila

alejandro.woodward@alu-etsetb.upc.edu irina.sanchez@alu-etsetb.upc.edu

genis.floriach@alu-etsetb.upc.edu oriol.vila.clara@alu-etsetb.upc.edu

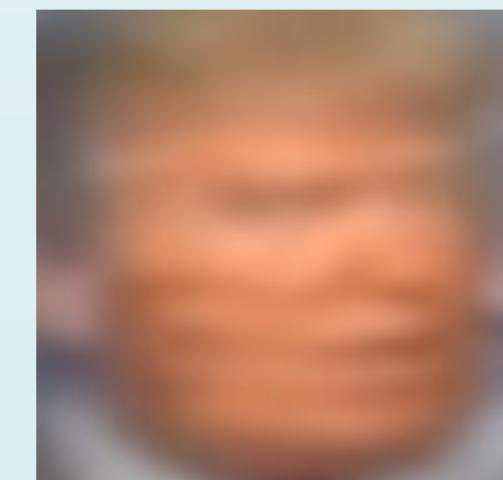
INTRODUCTION

The objective of this project is to train an artificial neural network that is able to generate realistic images of faces (in this case only of Donald Trump) using an architecture based in "BEGAN: Boundary Equilibrium Generative Adversarial Network"[1], that proved to generate very realistic faces. The aim is not only to generate realistic images but condition them to the speech. This means that the mouth should move corresponding to what the person is saying.

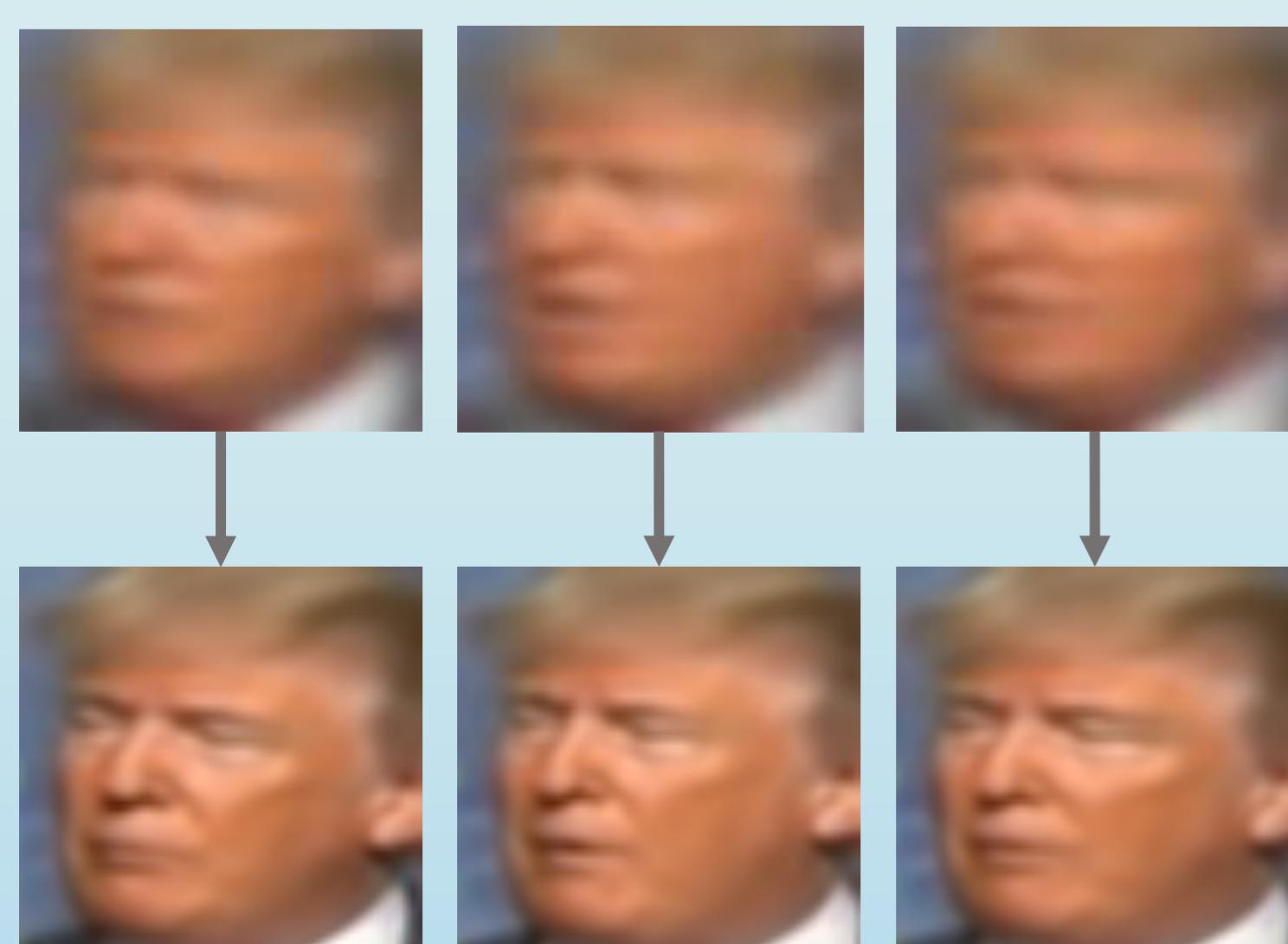
In order to accomplish this goal we generated a dataset using YouTube videos containing Donald Trump speeches. After the processing, the dataset is composed of a list of cropped faces and its corresponding audios of those moments. We tried different approaches using a modified BEGAN architecture, using as input to the generator audio features, but that did not give good results. In our last approach we used a MSE loss (without adversarial loss) to generate the faces depending on the audio, and then we used a network based on BEGAN to lower the blurring effect produced by the MSE loss.

VISUAL RESULTS

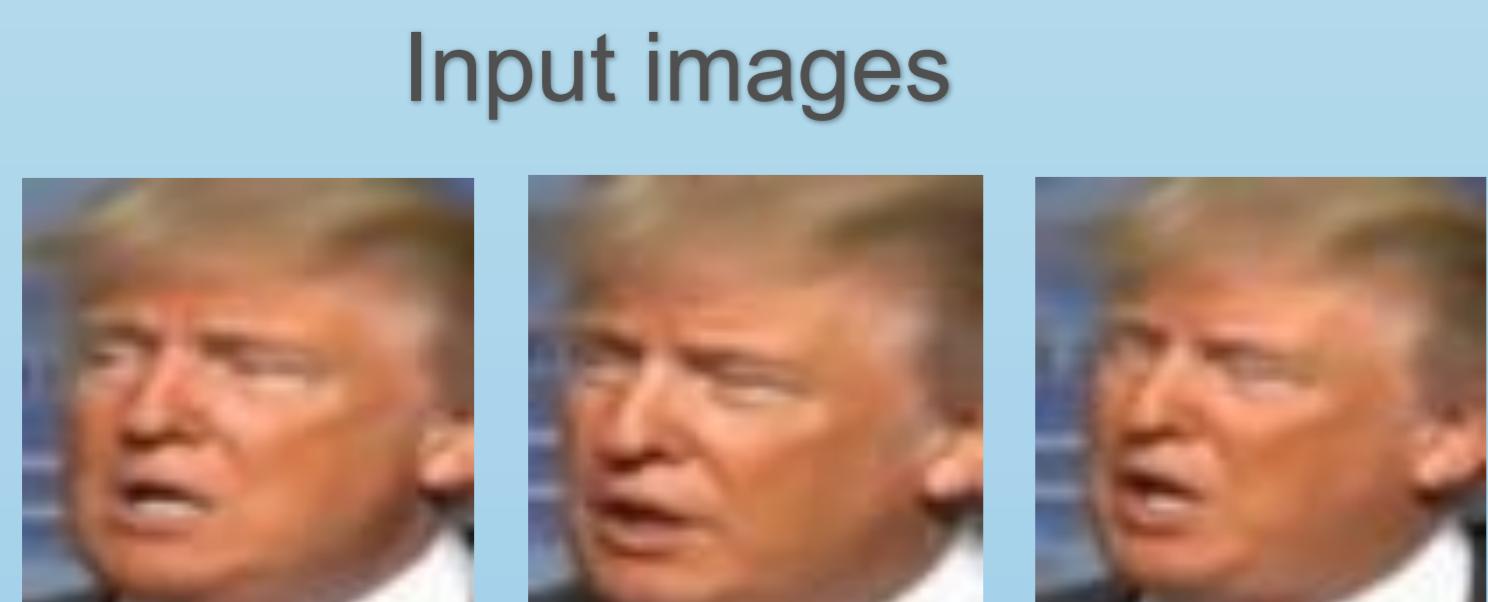
Output of 1. without flipping the faces



Output of 1. with flipped faces



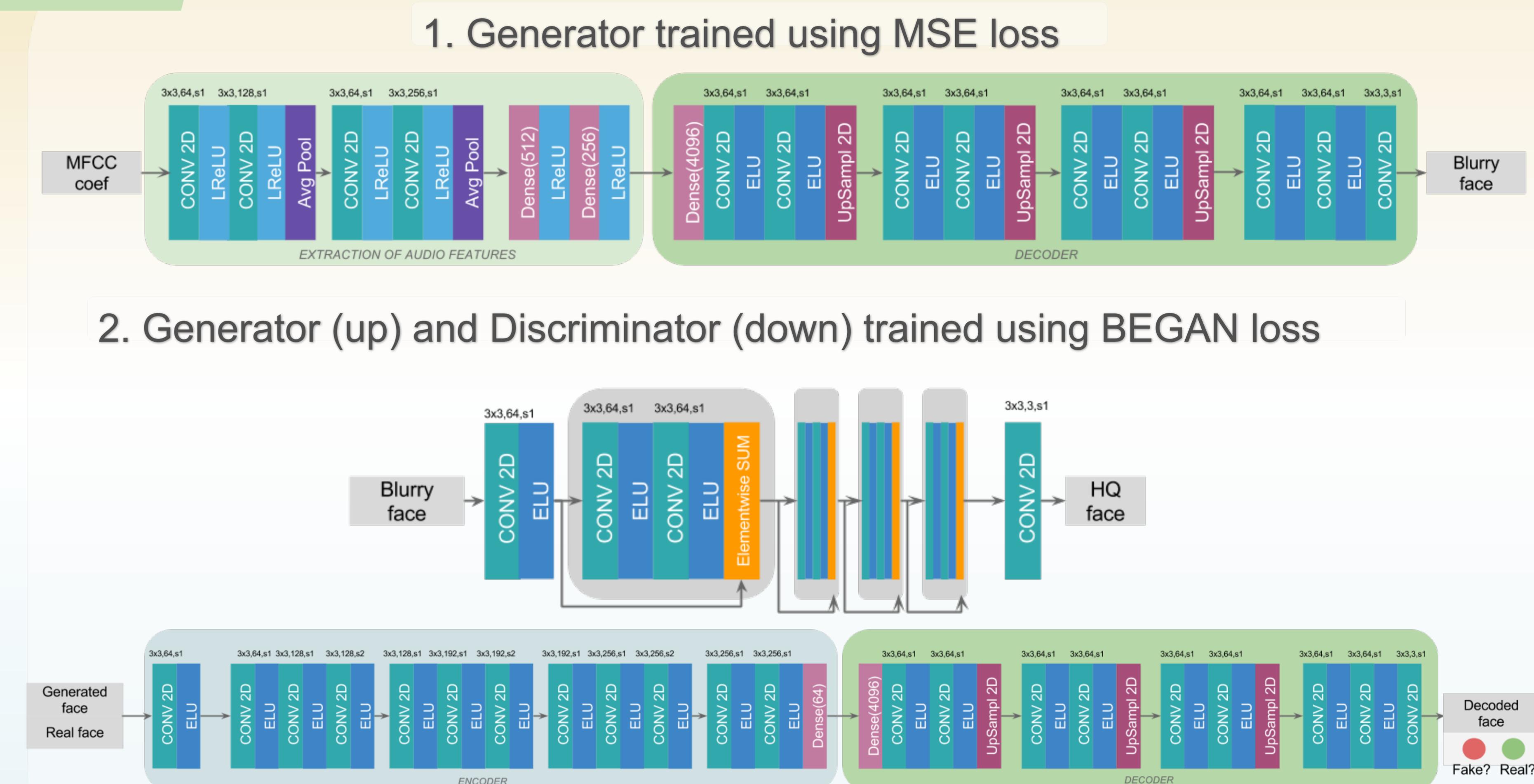
Final output after 2.



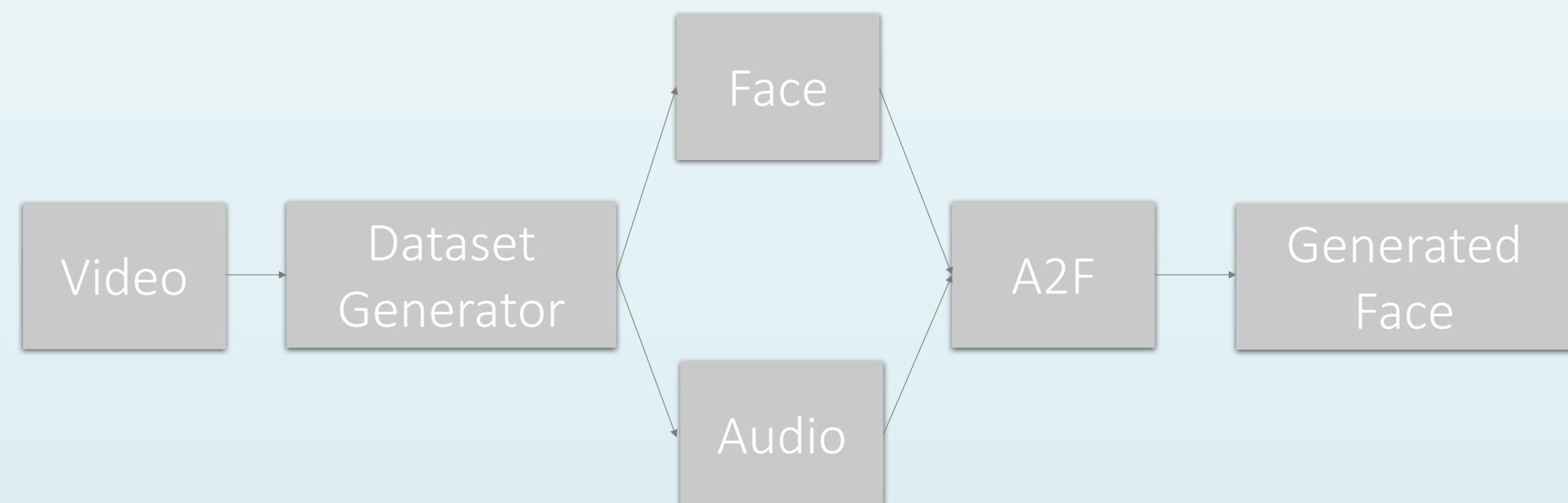
Input images



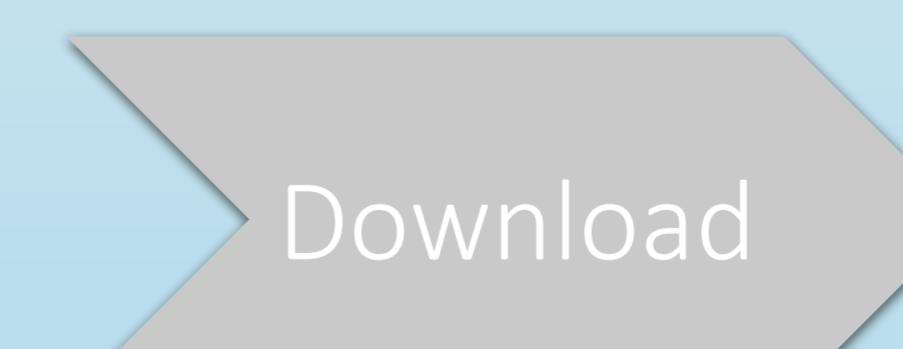
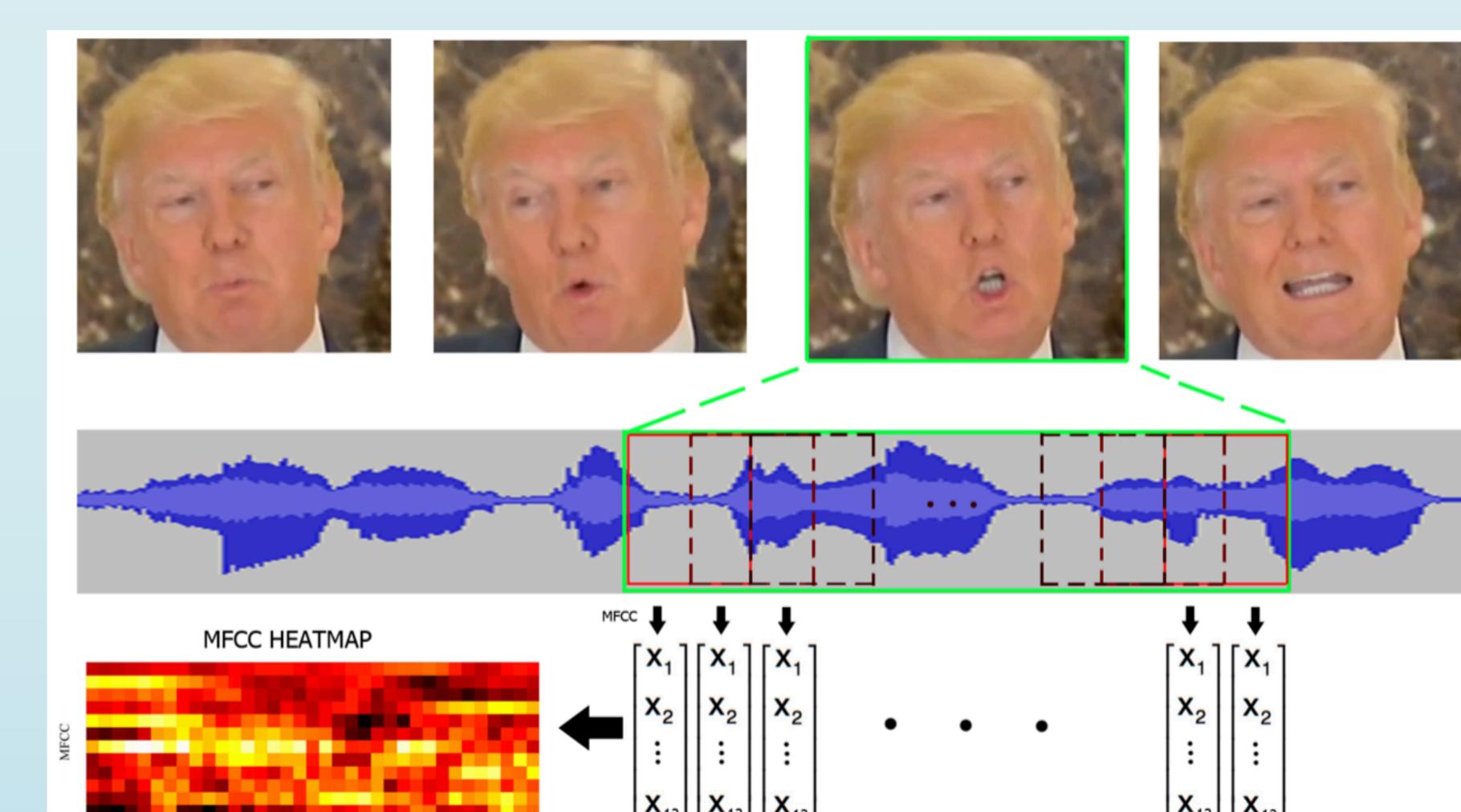
APPROACH – Audio2Face



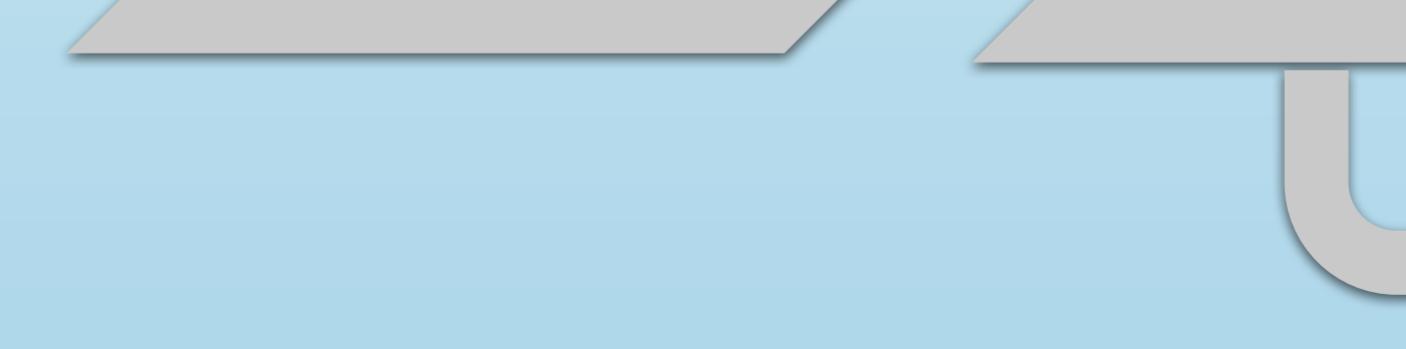
PIPELINE



DATASET GENERATOR



Download



Process

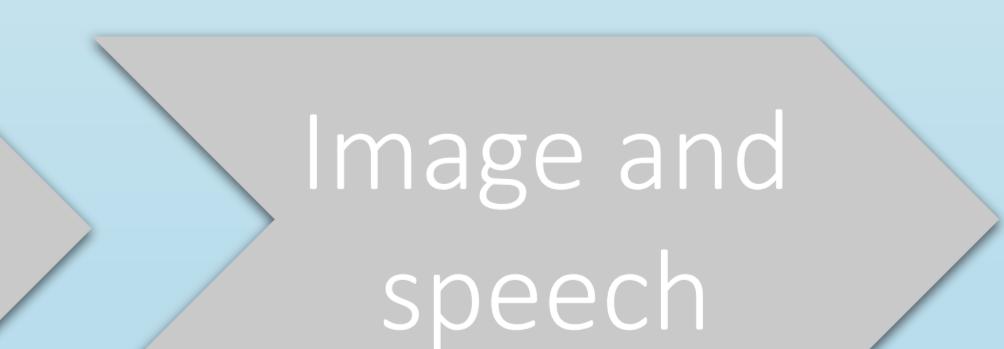


Image and speech

Face detection,
alignment and flip.
MFCC coefficients

REFERENCES:

- [1] David Berthelot, Thomas Schumm, Luke Metz. "BEGAN: Boundary Equilibrium Generative Adversarial Networks " (2017)