Real-Time Facial Categorization using Convolutional Neural Networks

Somnath Sarkar 15CO247

Mehul Sharma 15CO130

Vikram Rathour 15CO252

Problem Statement.

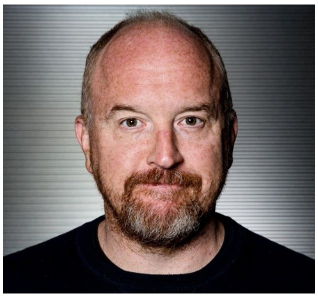
For the the Real-Time Facial Categorization using Convolutional Neural Networks project, we will make a website which uses the webcam interface of a computer to categorize the human face in the stream on the basis of a numerous factors. The faces can be categorized on the basis of gender, age, race, attractiveness and many other factors which are outlined at the end of this abstract. The information related to the categorizations would be shown on the website after the image is processed.

Input.

The project is split into 2 parts, the Front-End and Back-End. The Front-End will be implemented using a website which would be built by us. The website will have a webpage which accepts the input from the user's webcam feed. So, the input to the Front-End would be the video of the user's face. Once this input is recieved by the website, the face detector will extract / crop the face out from the video and discard the rest of the content on the frame. This extracted face would serve as input to our backend. The backend is the face categorizer program, which will categorize the user's face based on numerous attributes. Hence, the input to the Front-End is a video of the user's face, and the input to the Back-End is just the face of the user, cropped out from the frame.

Output.

The output from the Front-End is the cropped out face of the user. This output of the Front-end will serve as the input to the Back-End. The output of the Back-End will be on a webpage, which will categorize the face based on different features. A sample ouput of our program is shown below. This is not the actual output of the program, just how the output would be, fundamentally.



Gender: Male

Age : Middle-Aged

Race : White

Facial Hair : Bald, Moustache

Tools.

Hardware Requirements.

For this project we will be training the model on a high performance computer as we have a very large dataset. We will be training the model on the Graphical Processing Unit. The system requirements are as follows.

- NVIDIA Graphical Processing Unit
- 2GB GDDR5 Video Memory
- Intel/AMD Central Processing Unit
- 6GB of Random Access Memory
- Web camera

Software Requirements.

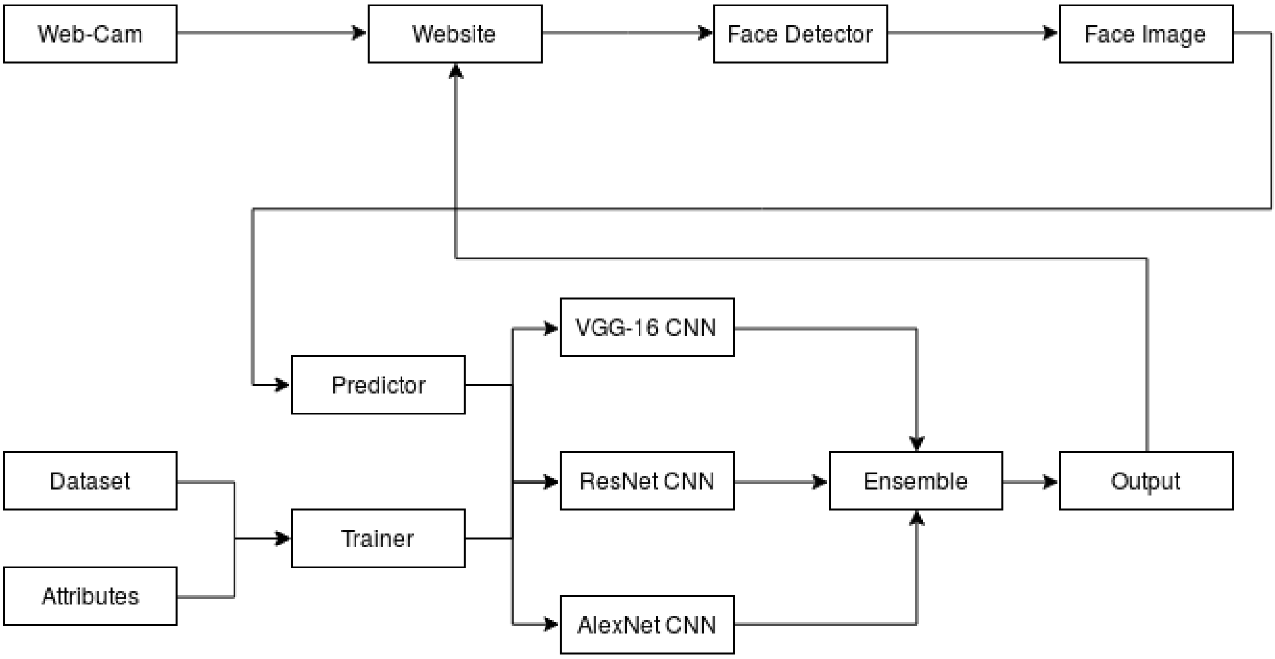
The follwoing tools will be required for our project.

- Scikit-Learn package
- Tensorflow package
- Keras package
- OpenCV package
- Django
- Jupyiter

Implementation

We are planning on training multiple convoluted neural networks on a large data set of human faces. The data set has been crowd-labelled with numerous descriptors which we intend to predict on the provided user image. For training, we will apply several architectures and combine the results to form a strong predictor. On the front-end we will have a web interface which applies computer vision algorithms on a live webcam feed from the user, to isolate the face and pass it to our backend, which consists of a series of compiled trained models generated in the training step. Here, we pass the face images as inputs to each of the networks and return the output to the user back through the interface. We intend to make use of convolutional neural networks as the main model behind our classification system.

Block Diagram.



Installation Instructions

Open windows terminal in administator mode and paste the following command to install the required python packages:

```
pip install numpy scipy tensorflow pandas scikit-learn opencv jupyter django
```

Work Distribution

Somnath - Image Pre-processing, ResNet ConvNet Architecture, Ensemble

Mehul - Web front-end, Video Processing, AlexNet ConvNet Architecture

Vikram - Data Pre-processing, VGG-16 ConvNet Architecture.