

Machine Learning Engineer Nanodegree

Capstone Proposal

Mohammed Ashour
December 31st, 2019

Proposal

Domain Background

Emotions are one of the most obvious and important factors to understand humans. Expressing our ideas is made very easier using emotions.

In this project, I'll try to make an emotion intelligent model that can understand human emotions from the facial expressions in images.

By understanding the current emotions of a person we can do ALOT. for example, we can detect the driver situation/state in a car and using this information to reduce car accidents, we can use it in media research,...etc

This topic is very active in research, I'm trying to do my best here to make a well-established contribution as reaching for a good model that can extract the most important features in the images and make a good judgment using the available data.

One of the aspiring companies in this field is affectiva, they had a set of products that using emotions recognition to empower automotive solutions and media research.

Problem Statement

The main objective of this project is classifying human emotions using their facial expressions. Using this solution we can do better AI robotics, more intelligent self-driving cars agents and better media research. The images are in gray-scale and jpg format, we will have a training and testing set. we will have different emotions included in our dataset like

- Happiness
- Surprise
- Neutral
- Disgust

- Fear
- Anger
- Sadness

and the goal of our model is to try to differentiate between them. This is a multi-class classification problem with more than 2 classes. The emotion that will get the most scores of them will determine the model estimation of the image.

Datasets and Inputs

As we Discussed Earlier, our dataset consists of gray-scale images with resolution 350x350 pixels that are classified in these categories

- Happiness
- Surprise
- Neutral
- Disgust
- Fear
- Anger
- Sadness

we will have the CSV file that connects each image with its label, and the images source folder that contains the jpg files The used dataset contains almost 13700 images with labels.

Link: https://github.com/muxspace/facial_expressions

Solution Statement

The proposed solution to this problem is to use a deep learning model that can extract complex features from images, like CNNs. The solution will be a CNN (Mostly I will use a pre-trained model like vgg and resnet) and a preprocessing for the input data to get the most of it. the o/p of the model will be a probability matrix of how likely the image matches each feeling, and we get the bigger one to be the prediction of the model. the summation of the matrix element will be 1.

Benchmark Model

Mostly, the approaches used in this review paper <https://www.mdpi.com/1424-8220/18/2/401>

will be tested and try to get the most useful for our data. My aim to get accuracy more than the avg accuracy mentioned in the mentioned paper (more than 72%)

Evaluation Metrics

For Evaluation, the false positives and negatives score will be used as well as the Traditional Accuracy score. Using the Traditional Accuracy score gives us initial intuition about how many misses our makes with it's guesses. but this is not enough! Using the false negative and positive makes it easier to know better about the dataset, which category is not recognized well..

Project Design

Preprocessing step

Here we will do some processing on the images like resizing all the images to smaller sizes so it's easier to train and make a balanced set of all the categories as possible. this will make it easier for the model to extract the most useful features from them.

Model Design

In this model I will use different tools and libraries like

- Pytorch (for building the model)
- Numpy (for matrix analysis)
- Matplotlib (for visualizations)
- cv2 (for image operation in the preprocessing step)

and for the architecture I will try some architectures like

- Vgg16
- Resnet
- Alexnet and others

and will choose the best model to deploy as a web app.

Training and testing

Here we will train and test our model to check out how it performs on the data it was training on and the data it never saw before. If the model didn't get good result, this maybe be an indicator to change

References

This is some references That I may use in my project

- <https://www.kaggle.com/c/emotion-detection-from-facial-expressions/leaderboard>
- <https://www.mdpi.com/1424-8220/18/2/401>
- <https://arxiv.org/pdf/1902.01019.pdf>

- <https://arxiv.org/pdf/1907.10838.pdf>