

Emotion Detection

Jyotirmay Senapati

jyotirmay.senapati@tum.de

Shayan Ahmad Siddiqui

shayan.siddiqui@tum.de

Abhijeet Parida

Abhijeet.parida@tum.de

Project Proposal

Use state of the art deep learning methods to predict human emotions.

1. Introduction

Facial expressions can be a key to understanding the human emotions. Based on this idea, our aim is to train a neural network which shall be able to predict the emotions by learning from a dataset of images of human facial expressions.

1.1. Related Works

- **Facial Emotion Detection Using Convolutional Neural Networks and Representational Autoencoder Units**

Author: Prudhvi Raj Dachapally.

- **Facial Emotion Recognition in Real Time**

Author: Dan Duncan, Gautam Shine, Chris English.

- **Analysis of Emotion Recognition using Facial Expressions, Speech and Multimodal Information**

Author: Carlos Busso, Zhigang Deng, Serdar Yildirim, Murtaza Bulut, Chul Min Lee, Abe Kazemzadeh, Sungbok Lee, Ulrich Neumann, Shrikanth Narayanan.

2. Dataset

- Dataset is taken from the kaggle challenge;
Challenges in Representation Learning: Facial Expression Recognition Challenge
- The data consists of 48x48 pixel grayscale images of faces. The faces have been automatically registered so that the face is more or less centered and occupies about the same amount of space in each image. The task is to categorize each face based on the emotion shown in the facial expression into one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral).
- Total data consists of 35887 images

- The training set consists of 28,709 examples, validation is 3,589 examples. The final test set, which will be used for testing consists of 3,589 examples.

3. Methodology

The methodology is inspired from the approach taken in the paper **Facenet2Expnet** where a two-stage training algorithm is carefully designed. In the pre-training stage, we will train the convolutional layers of the expression net, regularized by the face net; In the refining stage, we will append fully- connected layers to the pre-trained convolutional layers and train the whole network jointly. Following points present a step by step summary of the proposed approach:

- Implement the loss function proposed in the paper, which uses the output of convolution layer(frozen) of a pretrained Facenet model as a regularization
- Implement a second convolution network(Expnet) for emotion detection, the loss is only backpropagated through this network. We will start with same architecture for Expnet as proposed in the paper, later we will modify it further in order to perform better on our data
- Train the model using the 2 stage training process as presented in the paper on a GPU instance(either Vision Lab or Google cloud).
- Tweaking hyper-parameters to improve the performance further.
- If time permits try an alternate approach and do a comparative analysis.

4. Outcome

The expected outcome is to find out a well classified structure to understand emotions i.e. anger, happiness, sadness, surprised etc from an image input.