Human Face Emotion recognition using Deep Learning

Riyanshu Jain
B20AI060
jain.59@iitj.ac.in

Ruthvik K
B20AI037
ruthvik.2@iitj.ac.in

Shubh Soni
B20BB039
soni.11@iitj.ac.in

Abstract

Facial expression-based automatic emotion recognition is a fascinating research area that has significant applications in various fields such as safety, health (detection of mental illness, understanding of human behavior, and psychological profiling), and human-machine interfaces (biometrics). Hence, Facial emotion recognition (FER) is important for human-computer interactions like behavioural description. In recent years, researchers have been exploring different techniques to interpret and code facial expressions using deep learning architectures to improve prediction accuracy. In this project, we build an automatic facial emotion recognition (FER) via deep learning and analyse various architectures, databases used in these contributions. Overall, this project serves as a resource to FER using deep learning techniques as we are going to deploy it.

1. Introduction

Emotions play a vital role in our daily lives, and detecting and understanding them can be essential in various fields such as psychology, healthcare, and marketing. The advancement of deep learning techniques has enabled researchers to build emotion detection systems that can identify human emotions from various sources such as facial expressions, voice, and text. Emotion detection using deep learning techniques involves training a neural network with large amounts of labeled data, enabling it to learn the patterns and features that represent different emotions accurately. In this project, we will explore the application of deep learning techniques for emotion detection, their benefits, limitations, and future prospects. We further plan to deploy a web application that can detect human emotion out of an image.

2. Related Work

Accurately classifying the human face emotion is a challenging task due to the heterogeneity of human faces and variations in images [1]. The paper [1] presents using VGGNet architecture to achieve a high performance without using extra training data on the FER2013 dataset. FER 2013 dataset, introduced in ICML 2013, has became a benchmark in comparing model performance in emotion recognition. Many other CNN variants have achieved an accuracy in the range of 67-76 percentage. For instance, the paper [2], presents LHC-Net to achieve SOTA performance on the FER-2013 dataset but uses extra training data for its purpose.

3. Methodology

We plan to give users an option to select the model they want to use among the SOTA performing models like ResNet, WideResNet and a few more. The user will be allowed to give an input image which will be used to predict the emotions on the face. We also plan to provide grad-cam visualisation to highlight the important areas in the image that the model used to predict the face emotion.
Framework to be used Python, PyTorch, Pandas, NumPy, Streamlit
Dataset to be used

- **Facial Expression Recognition 2013 Dataset**
  Fer2013 (link) contains approximately 30,000 facial RGB images of different expressions with size restricted to 48×48, and the main labels of it can be divided into 7 types: 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral.

- **AffectNet**
  AffectNet (link) is a large facial expression dataset with around 0.4 million images manually labeled for the presence of eight (neutral, happy, angry, sad, fear, surprise, disgust, contempt) facial expressions.

References

[1] Yousif Khaireddin and Zhuofa Chen. Facial emotion recognition: State of the art performance on fer2013, 2021. 1

[2] Roberto Pecoraro, Valerio Basile, and Viviana Bono. Local multi-head channel self-attention for facial expression recognition. *Information*, 13(9), 2022. 1