Innovation & Entrepreneurship Hub for Educated Rural Youth (SURE Trust – IERY)

Emoji Recommendation based on Facial Expression Recognition

The domain of the Project:

Computer Vision & Artificial Intelligence

By

Chigurupati Varsha - pursuing B.Tech , 4th year

Period of the project

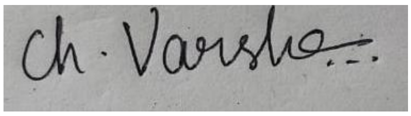
July 2024 to April 2025

Innovation & Entrepreneurship Hub for Educated Rural Youth (SURE Trust – IERY)

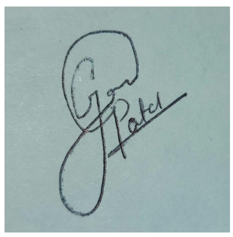
Declaration

The project titled “Emotion Recommendation based on Facial Expression Recognition” has been mentored by Gaurav patel, organised by SURE Trust, from July 2024 to April 2025, for the benefit of the educated unemployed rural youth for gaining hands-on experience in working on industry relevant projects that would take them closer to the prospective employer. I declare that to the best of my knowledge the members of the team mentioned below, have worked on it successfully and enhanced their practical knowledge in the domain.

Chigurupati Varsha — Signature



Mentor
Gaurav Patel — Signature



Seal & Signature

Prof.Radhakumari

Executi ve Director & Founder

SURE Trust

Table of contents

Executive Summary

This project explores the development of a Facial Emoji Recommendation System using deep learning and facial expression recognition. The system is designed to detect human emotions from facial images and recommend appropriate emojis, such as a happy face emoji for a happy expression. The project utilizes publicly available facial emotion datasets, applies advanced preprocessing and data augmentation techniques, and implements both a custom Convolutional Neural Network (CNN) and Transfer Learning using ResNet50. The final model which achieved around 67% accuracy is deployed using Streamlit, offering an interactive user experience. The project demonstrates the practical use of AI in enhancing communication and expression in digital platforms.

Introduction

Facial expressions are a fundamental aspect of human communication. With the advancement of artificial intelligence and computer vision, recognizing and interpreting facial expressions has become increasingly feasible and valuable. This project aims to leverage facial expression recognition to recommend emojis that best match the detected emotion, enhancing user interaction in digital platforms. The use of deep learning, particularly CNNs, provides robust techniques to accurately classify emotions such as happiness, sadness, anger, and surprise.

This project report details the development of a facial emotion recognition model utilizing the ResNet50 architecture within the TensorFlow framework. The central focus of this project is to create a system capable of automatically identifying emotions from images of faces. This is achieved through the implementation of a complete workflow that leverages deep learning techniques.

Project Objectives

The core objectives of the project include:

- Developing a robust deep learning model to detect facial emotions from static images.
- Recommending appropriate emojis based on the detected emotion.
- Handling class imbalance in datasets through augmentati on.
- Deploying the model using Streamlit for user interaction.
- Exploring transfer learning models for improved performance.

The project also sought to experiment with multiple modeling techniques. A custom CNN architecture was built from scratch to understand the model's learning capability and performance. In parallel, transfer learning models such as ResNet50 were tested to leverage pre-trained knowledge for better generalization.

Finally, the project intended to deploy the model using Streamlit to make it accessible to end-users. The objective was to create a simple interface where users could upload an image and receive an emoji that best matches the detected emotion. This brings practicality to the model and makes it useful for real-time applications in social media, mental health monitoring, and human-computer interaction.

Dataset Description & Preprocessing

Datasets Explored:

1. **FER 2013** – 28,709 training images, 7,178 test images, grayscale 48x48.
2. **Face Expression Recognition Dataset** – 28,821 training images, 7,066 test images.
3. **CK+ Dataset** – Kaggle version with around 920 images.
4. **Facial Emotion Recognition Dataset** – Combined with others to get over 63,000 images.

Final Dataset:

- Chosen: Face Expression Recognition Dataset (Kaggle)
- Total images after augmentation: 50,148
- 7 classes: Angry, Disgust, Fear, Happy, Neutral, Sad, Surprise

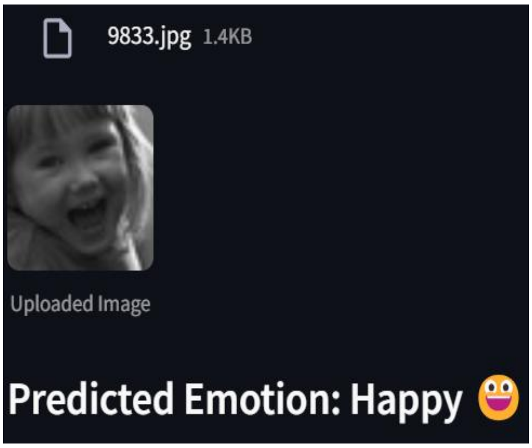
Preprocessing:

- **Data Augmentation:** Applied rotation, shifting, zooming, flipping.
- **Class Balancing:** Augmented underrepresented classes (e.g., Disgust).
- **Normalization:** Pixel values scaled for training.

GitHub: Project Link

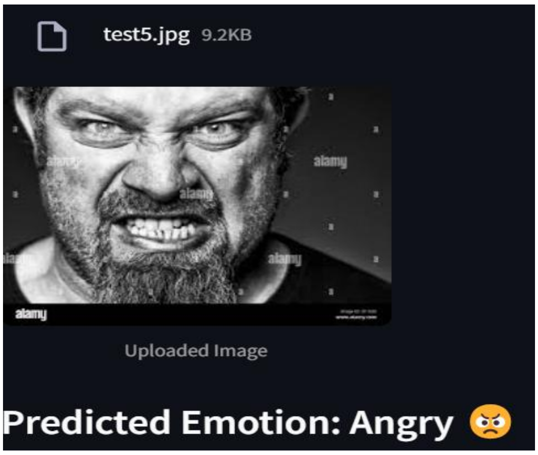
Results


9460.jpg 1.7KB
Uploaded Image
Predicted Emotion: Disgust


9833.jpg 1.4KB
Uploaded Image
Predicted Emotion: Happy


9993.jpg 1.7KB
Uploaded Image
Predicted Emotion: Neutral


9205.jpg 1.5KB
Uploaded Image
Predicted Emotion: Fear


test5.jpg 9.2KB
Uploaded Image
Predicted Emotion: Angry

Methodology

Workflow :

The workflow of the project began with data loading, where the selected dataset was imported and organized for training and validation. Following this, data augmentation techniques were applied to enhance the diversity of the dataset and balance the number of images in each emotion category. Preprocessing steps, including normalization and resizing, were then conducted to prepare the data for model input.

After preprocessing, Exploratory Data Analysis (EDA) was carried out to understand the distribution of classes and visual characteristics of the images. This helped in identifying any biases or inconsistencies in the data. The next phase involved building and training the deep learning models. Both a custom Convolutional Neural Network and a ResNet50-based model were implemented and trained to recognize facial expressions.

Finally, the trained model was deployed using Streamlit, allowing users to interact with the system by uploading images and receiving emoji recommendations based on detected emotions.

Models Used:

The project involved two main approaches for model training. The first approach was the development of a custom Convolutional Neural Network (CNN). This model was built from scratch to learn the features of facial expressions directly from the dataset. The CNN achieved a training accuracy of 77.06% and a validation accuracy of 51.10%. Training was conducted over 50 epochs but stopped early at the 22nd epoch due to EarlyStopping criteria that prevent overfitting by halting training when the model performance on validation data plateaus.

The second approach employed transfer learning using the ResNet50 architecture. This model leveraged pre-trained weights from large datasets to improve learning efficiency and performance on the emotion dataset. The ResNet50 model achieved a training accuracy of 63.24% and a validation accuracy of 65.28%, showing its superior generalization capabilities compared to the custom CNN. It trained for a maximum of 50 epochs and was stopped early at the 39th epoch.

In both models, the Adam optimizer was used to minimize the loss function, which was categorical crossentropy, suitable for multiclass classification tasks.

Deployment Approach

The trained model was deployed using Streamlit, an open-source Python library for creating interactive web apps. Users can upload a facial image, and the system returns the detected emotion and a corresponding emoji.

Issues Faced:

During deployment, the project faced several technical challenges. One major issue was kernel crashes that occurred during the training phase, especially while using transfer learning models like ResNet50. These crashes were primarily due to high memory usage that exceeded the limits of available computational resources. Additionally, while training in Google Colab, the runtime environment frequently disconnected, interrupting the model training process.

To overcome these limitations, the project adopted a strategy of model simplification. Batch sizes were adjusted, and some architectural elements were optimized to make the training process more resource-efficient. This ensured smoother training and successful deployment of the model on the Streamlit platform.

Challenges & Solutions

The project faced several data and technical challenges that had to be addressed for successful model development. One significant data challenge was the class imbalance in the dataset. Emotions such as 'disgust' and 'fear' had fewer image samples compared to emotions like 'happy' or 'neutral'. This imbalance led to biased model predictions. The solution involved implementing data augmentation techniques to synthetically increase the number of samples in underrepresented classes, thereby achieving a more balanced dataset.

Another challenge was the low quality and resolution of some images, which made it difficult for the model to accurately detect subtle facial features. Preprocessing steps such as normalization and resizing were applied to improve image quality and consistency across the dataset.

From a technical standpoint, memory constraints posed a major issue. Transfer learning models, which are typically large, caused memory overflows and kernel crashes. These were mitigated by simplifying model structures and using smaller batch sizes during training. Additionally, distinguishing between similar emotions, such as fear and surprise, proved difficult due to overlapping facial features. Further improvements in the dataset and model complexity are planned to address this challenge more effectively.

Social and Industry Relevance

This project has strong implications in both social and industrial domains:

- **Social Media:** Enhancing user experience with auto-suggested emojis.
- **Mental Health:** Recognizing emotional states to offer timely support.
- **Customer Feedback Analysis:** Understanding sentiment from customer images.
- **Education:** Monitoring student engagement and emotional response.

Learnings & Reflections

- Developed deep understanding of CNNs and transfer learning.
- Gained practical experience with image augmentation and preprocessing.
- Learned to tackle real-world challenges like class imbalance and deployment limitations.
- Improved debugging and model optimization skills through iterative training.

There are several potential enhancements that could expand the functionality and performance of the Facial Emoji Recommendation System. One promising direction is the integration of real-time webcam support, allowing the system to detect and classify emotions from live video feeds. This would make the application more interactive and suitable for use in communication platforms or mental health monitoring tools.

Another area for improvement is the expansion of the emoji library. Including a wider variety of emojis that correspond to nuanced emotional expressions could provide more personalized and expressive feedback to users.

In terms of technical improvements, deploying lightweight versions of the model would enable its usage on mobile and embedded devices. This could make the system more accessible and practical for a broader range of applications. Additionally, experimenting with ensemble techniques, where multiple models contribute to the final prediction, may help improve overall accuracy and robustness.

Conclusion:

The Facial Emoji Recommendation System successfully demonstrates how deep learning can be used to detect facial emotions and suggest emojis accordingly. Despite computational limitations and dataset challenges, the project achieved a functional and deployable solution with real-world application potential. This lays the groundwork for future research and development in emotion-aware applications.