ARTIFICIAL INTELLIGENCE

GROUP MEMBERS: BCSF22M008 (Mhammad Mujtaba), BCSF22M016 (Saqib Ali Javaid), BCSF22M019 (Faiqa Mustafa), BCSF22M027 (Aqsa Hussain), BCSF22M033 (Ali Husnain)

PROJECT: Facial Emotion Recognition Using CNN

Abstract

This project presents a Facial Emotion Recognition (FER) system developed using a Convolutional Neural Network (CNN) trained on the FER-2013 dataset. The model is designed to classify human facial expressions into seven fundamental emotions: angry, disgust, fear, happy, sad, surprise, and neutral. Pre-processing techniques such as normalization and grayscale conversion were applied to improve model performance. The trained CNN achieved a test accuracy of approximately **65%**, demonstrating a promising capability in recognizing emotional expressions. Evaluation was carried out using both the test split of the dataset and real-time webcam-based predictions. The results highlight the practical applicability of deep learning in emotion analysis from facial cues.

Introduction

Problem Statement

Recognizing human emotions accurately through facial expressions remains a challenging task due to factors such as lighting variations, facial occlusions, and subtle expression differences. Despite advances in computer vision, creating robust emotion recognition systems that perform reliably in real-world scenarios is still a major research area.

Significance of the Project

Facial Emotion Recognition has diverse applications in fields such as human-computer interaction, mental health monitoring, customer experience systems, and surveillance. This project aims to contribute a CNN-based solution to automate the recognition process, enhancing machine understanding of human emotional states.

Methodology

Model Used

- **Convolutional Neural Network (CNN)**

Justification

CNNs are highly effective for image classification tasks due to their ability to capture hierarchical patterns and spatial relationships in pixel data. They are particularly suited for facial expression analysis, as they can detect subtle differences in facial muscle movements and structures associated with different emotions.

Dataset Description

- **Name:** FER-2013 (Facial Expression Recognition 2013)
- **Source:** Kaggle
- **Image Size:** 48x48 pixels
- **Color:** Grayscale
- **Total Images:** 35,887
  - Training Set: ~28,709 images
  - Test Set: ~7,178 images
- **Emotions:** Angry, Disgust, Fear, Happy, Sad, Surprise, Neutral

Pre-processing

- Resizing all images to 48x48
- Normalization of pixel values
- Grayscale conversion
- Optional data augmentation (rotation, flipping, etc.) for model generalization

Results

Overall Test Accuracy: 65.00%

Per-Emotion Accuracy Summary

Emotion	Correct / Total	Accuracy
Angry	469 / 958	48.96%
Disgust	36 / 111	32.43%
Fear	377 / 1024	36.82%
Happy	1451 / 1774	81.79%
Neutral	619 / 1233	50.20%
Sad	646 / 1247	51.80%
Surprise	616 / 831	74.13%

Performance Metrics

- **Accuracy:** 65.00%
- **Precision, Recall, F1-score:** Computed using classification reports
- A CSV file was generated to store actual vs. predicted labels for further analysis

Discussion

Challenges Faced

- Class imbalance, especially in "Disgust" and "Fear" categories
- Confusion between visually similar emotions like "Neutral" and "Sad"
- Generalization to real-time scenarios due to dataset limitations

Insights Gained

- Pre-processing significantly impacts model accuracy
- Grayscale images reduce computational load with minimal accuracy loss
- Simple CNNs can achieve reasonable performance with careful tuning

Comparative Analysis

While this project used a custom CNN model, more advanced architectures like **VGG16**, **ResNet50**, or **YOLOv8** could potentially yield better performance through deeper feature extraction and transfer learning.

Conclusion & Future Work

This project successfully implemented a CNN-based Facial Emotion Recognition system trained on the FER-2013 dataset. The model achieved a test accuracy of **65%**, showing effectiveness in detecting common facial emotions, particularly "Happy" and "Surprise." However, accuracy for certain subtle expressions remains a challenge.

Future Enhancements

- Employing more sophisticated models such as VGG or ResNet
- Applying transfer learning and data augmentation
- Developing a cross-platform mobile or web application for real-time emotion recognition
- Exploring multi-modal emotion recognition (e.g., combining audio and facial data)

References

1. Kaggle FER-2013 Dataset — https://www.kaggle.com/datasets/msambare/fer2013
2. TensorFlow/Keras Documentation — https://www.tensorflow.org/
3. Chollet, F. *Deep Learning with Python*, Manning Publications
4. Goodfellow, I., Bengio, Y., & Courville, A. *Deep Learning*, MIT Press