SOC 2025 end-term report

In this report, I have covered various details regarding my end-term project which was-

Building a Facial Expression Recognition Model Using the FER 2013 Dataset.

1. Introduction

Facial expressions are a key form of non-verbal communication and play a vital role in human interaction. Automated facial expression recognition (FER) can be applied in various domains, such as human-computer interaction, mental health monitoring, and customer satisfaction analysis. This project aims to develop a deep learning model that accurately classifies facial expressions using the FER-2013 dataset.

2. Problem Statement

The objective of this project is to build a convolutional neural network (CNN) that can automatically recognize and classify facial expressions into one of seven categories: Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral. The challenge lies in dealing with grayscale images of low resolution (48x48 pixels) and high intra-class variation.

3. Proposed Solution

To address the problem, we implemented the following approach:

- Preprocessed the dataset by normalizing pixel values.
- Split the dataset into training, validation, and test sets.
- Designed and trained a CNN model using TensorFlow and Keras.
- Applied techniques to avoid overfitting and enhance generalization.
- Evaluated the model using relevant metrics.

4. Dataset Description

The FER-2013 dataset contains 35,887 grayscale images of size 48x48 pixels, each associated with one of the seven expression labels. The data is provided in a CSV format with pixel values stored as strings.

5. Data Preprocessing

- Normalization: Pixel values (0-255) were normalized to the [0,1] range.
- Reshaping: Images were reshaped to 48x48x1 dimensions.
- Label Encoding: The emotion labels were one-hot encoded.
- Data Augmentation: Random rotations, shifts, and horizontal flips were applied to increase data diversity and prevent overfitting.

6. Data Splitting

The dataset was split as follows:

- Training Set: 80%
- Validation Set: 10%
- Test Set: 10%

This ensured a balanced evaluation and effective training monitoring.

7. Model Architecture

The CNN model includes:

- Convolutional Layers: For feature extraction with ReLU activation
- Max Pooling Layers: For spatial downsampling
- Dropout Layers: To prevent overfitting
- Fully Connected Layers: For classification
- Softmax Output Layer: To output class probabilities

8. Model Compilation and Training

- Loss Function: Categorical Crossentropy
- Optimizer: Adam
- Metrics: Accuracy
- Hyperparameters:
    - Learning Rate: 0.001
    - Batch Size: 64
    - Epochs: 50

The model was trained with early stopping and model checkpointing based on validation loss.

9. Results

- Training Accuracy: ~92%
- Validation Accuracy: ~65%
- Test Accuracy: ~64%

The confusion matrix revealed that the model performed best on 'Happy' and 'Neutral' classes, and struggled with 'Disgust' due to fewer samples.

10. Observations and Future Work

- The model shows potential but can be improved with:
    - Deeper architectures (e.g., VGG, ResNet)
    - Transfer learning using pre-trained models
    - Better data balancing and augmentation
- Additional metrics like F1-score and ROC-AUC could provide more insight into class-wise performance.

11. Conclusion

We successfully built a CNN-based facial expression recognition system using the FER-2013 dataset. Despite the dataset's limitations, the model demonstrated reasonable accuracy and serves as a foundation for further improvements. This project highlights the power of deep learning in tackling complex image classification problems.