Emotion Recognition with Mini-XCEPTION

Project Report

1. Introduction

This project implements a real-time facial emotion recognition system using a lightweight CNN architecture known as **Mini-XCEPTION**. It leverages the FER-2013 dataset and OpenCV for live emotion detection via webcam.

2. Dataset

- **Source:** FER-2013 (Facial Expression Recognition 2013)
- **Classes:** Angry, Disgust, Fear, Happy, Neutral, Sad, Surprise
- **Structure:**
    - fer2013/train: Training images organized by emotion labels
    - fer2013/test: Testing/validation images

3. Data Preprocessing & Augmentation

- **Input Image Size:** 48x48 (Grayscale)
- **Augmentations on Training Set:**
    - Rotation: ±30°
    - Zoom: Up to 30%
    - Width/Height Shift: Up to 30%
    - Shear: 0.2
    - Horizontal Flip
    - Rescaling: 1./255
- **Validation Set:** Rescaled only

4. Model: Mini-XCEPTION

A compact yet effective CNN with separable convolutions and batch normalization.

Architecture Summary:

- Input: (48, 48, 1)
- Layers:
    - Conv2D → SeparableConv2D × 2 → MaxPooling
    - Repeat with increasing filters: 16 → 32 → 64 → 128 → 256
    - GlobalAveragePooling
    - Dropout (0.5)
    - Dense (softmax)
- Parameters: Efficient, fewer than traditional CNNs
- Output: 7 emotion classes

5. Training Details

- **Optimizer:** Adam
- **Loss Function:** Categorical Crossentropy with label smoothing (0.1)
- **Batch Size:** 64
- **Epochs:** 50
- **Callbacks Used:**
    - EarlyStopping: Stops training if no improvement in val_loss (patience=8)
    - ModelCheckpoint: Saves the best model based on val_accuracy
    - ReduceLROnPlateau: Lowers learning rate on plateaus
- **Class Weights:** Computed to address class imbalance

6. Performance Metrics

During training, both accuracy and loss were plotted:

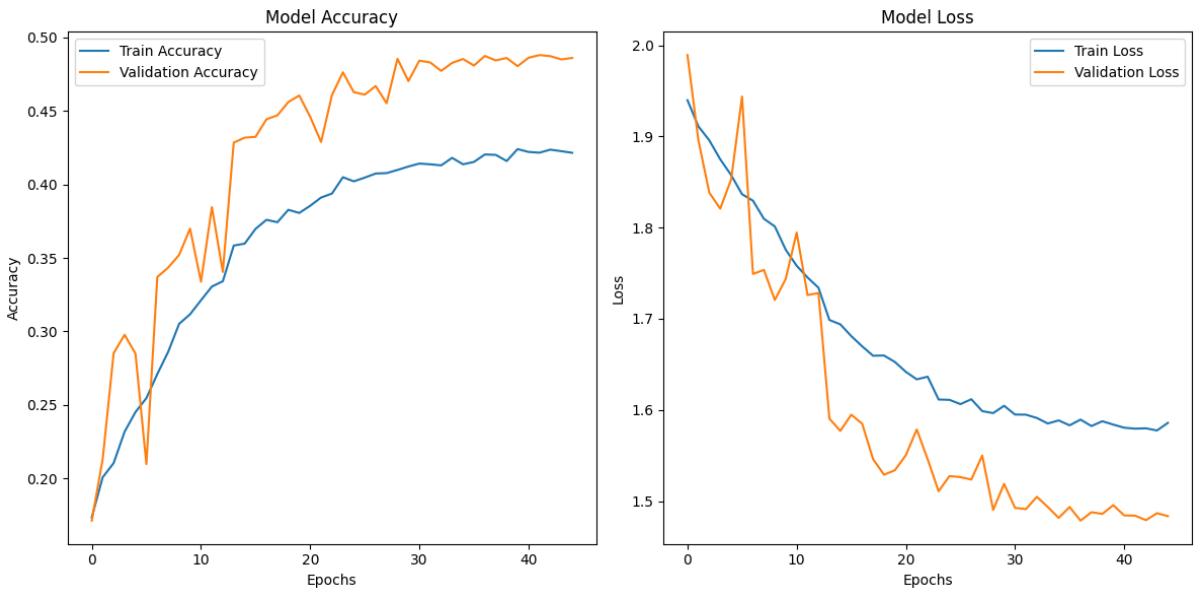
- **Training Accuracy and Loss** showed consistent improvement.
- **Validation Accuracy and Loss** demonstrated generalization without overfitting due to early stopping.

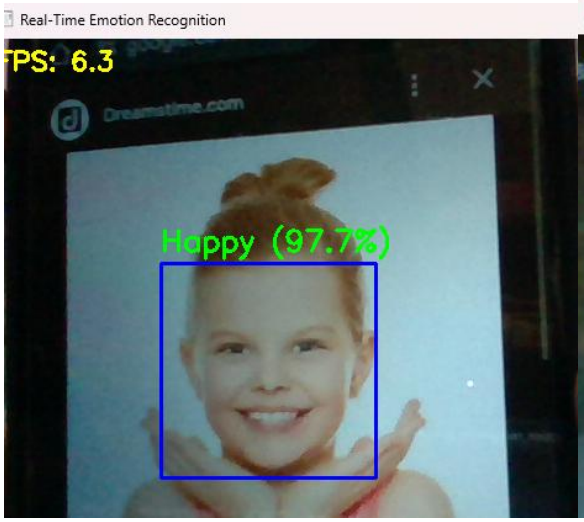
7. Real-Time Emotion Recognition

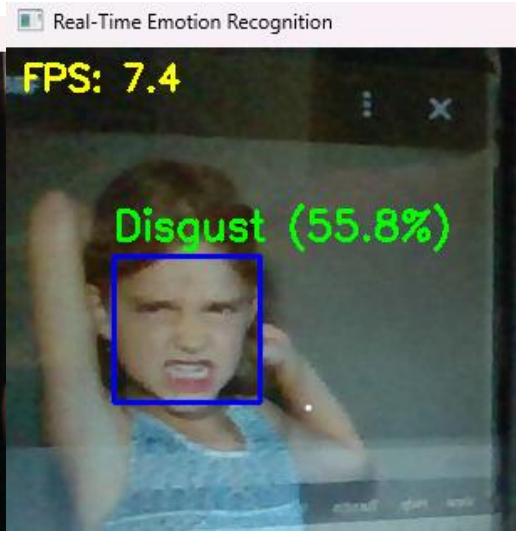
- **Face Detection:** OpenCV's Haar cascade (haarcascade_frontalface_default.xml)
- **Inference Pipeline:**
    - Capture from webcam
    - Convert to grayscale
    - Detect faces and crop
    - Resize to 48x48, normalize, reshape
    - Predict with Mini-XCEPTION model
    - Annotate frame with predicted emotion and confidence
- **FPS Counter:** Displays real-time processing speed
- **Exit Key:** Press 'q' to quit

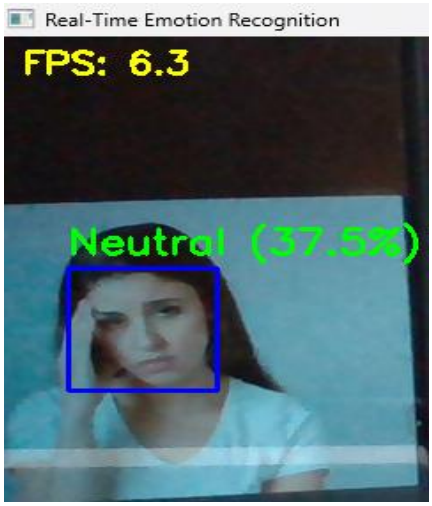
8. Results

- **Real-time Predictions:** Responsive with good accuracy on most common expressions
- **Lightweight Model:** Achieves real-time performance (~15–30 FPS) on most systems
- **Confidence Scores:** Displayed for each emotion prediction




Real-Time Emotion Recognition
FPS: 6.3
Happy (97.7%)


Real-Time Emotion Recognition
FPS: 7.4
Disgust (55.8%)


Real-Time Emotion Recognition
FPS: 6.3
Neutral (37.5%)