Facial Expression Recognition Using CNN

Overview

This project implements a Convolutional Neural Network (CNN) to classify facial expressions into seven emotion categories: Angry, Disgust, Fear, Happy, Neutral, Sad, Surprise. The model is trained on grayscale images resized to 48x48 pixels and validated on a separate test set. A live webcam demo and image testing pipeline are integrated using OpenCV.

Approach

1. Data Preprocessing

- Dataset is structured in directories by emotion labels.
- Grayscale conversion and normalization (rescale=1./255).
- Real-time augmentation on the training set (rotation, shear, zoom, flip).

2. Model Architecture

- 4 Convolutional layers (32, 64, 128, 256 filters) with ReLU activation and max pooling.
- Dropout regularization (10-20%) to prevent overfitting.
- Dense layer with 512 units followed by a softmax output layer (7 classes).
- Compiled with Adam optimizer and categorical crossentropy loss.

3. Training

- Trained for 30 epochs with steps_per_epoch = num_train_imgs // 32.
- Validation performed at each epoch with test images.

4. Evaluation Metrics

- Accuracy and loss plotted for both training and validation sets.
- Classification report and confusion matrix used to evaluate per-class performance.

5. Deployment

- Real-time detection and classification using webcam (test.py).
- Static image testing (testdata.py).

- Face detection via Haar cascades (haarcascade_frontalface_default.xml).

Findings

- Training Accuracy steadily improved and stabilized around 85-90% over 30 epochs.
- Validation Accuracy was slightly lower (~75-80%), indicating generalization with minor overfitting.
- The Confusion Matrix revealed common misclassifications between similar emotions (e.g., Sad vs. Neutral).
- Real-time performance was satisfactory with accurate prediction on frontal face images in good lighting.

Challenges Faced

- Hardware Limitations: Local CPU was insufficient for fast training; offloaded to Google Colab.
- Class Imbalance: Emotions like Disgust had fewer samples, impacting recall.
- Lighting and Angles: Real-time detection accuracy dropped in poor lighting or non-frontal faces.
- Overfitting Risk: Required dropout tuning and data augmentation to improve generalization.
- Inference Delay: Slight latency in webcam prediction due to face preprocessing and model inference time.

Conclusion

The project successfully demonstrates emotion recognition using CNN and integrates it with OpenCV for real-time applications. Future improvements could include better face alignment, using deeper networks like ResNet, and expanding the dataset for underrepresented emotions.