Facial Emotion Recognition (FER)

Rui Shi, Yuting Ma, Jiayi Wang, Menglei Wang

FER-2013 dataset

Data Preprocessing and Augmentation

- Standardization: Images are resized to 48x48 pixels and denoised to improve quality.

- Data Augmentation: Augment minority classes with transformation techniques like rotation, flipping, and zooming while downsampling some majority classes.

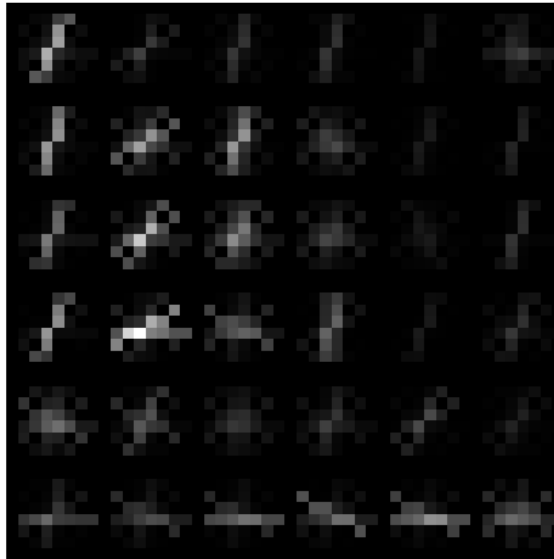# Feature Extraction: HOG + DCT

Step 1: Extract HOG (Histogram of Oriented Gradients) features

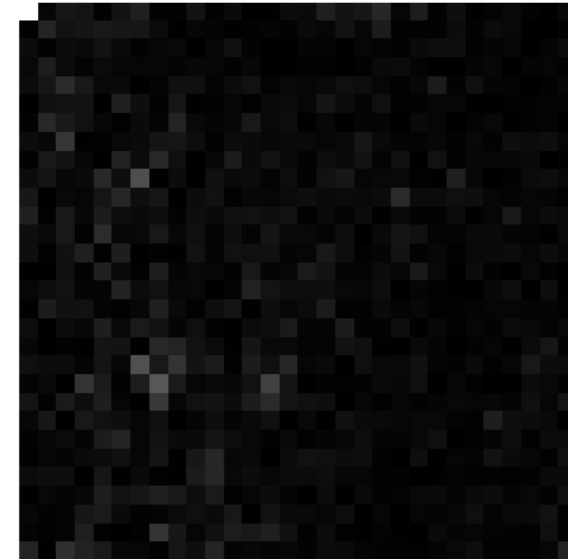Step 2: Apply DCT (Discrete Cosine Transform) to HOG features
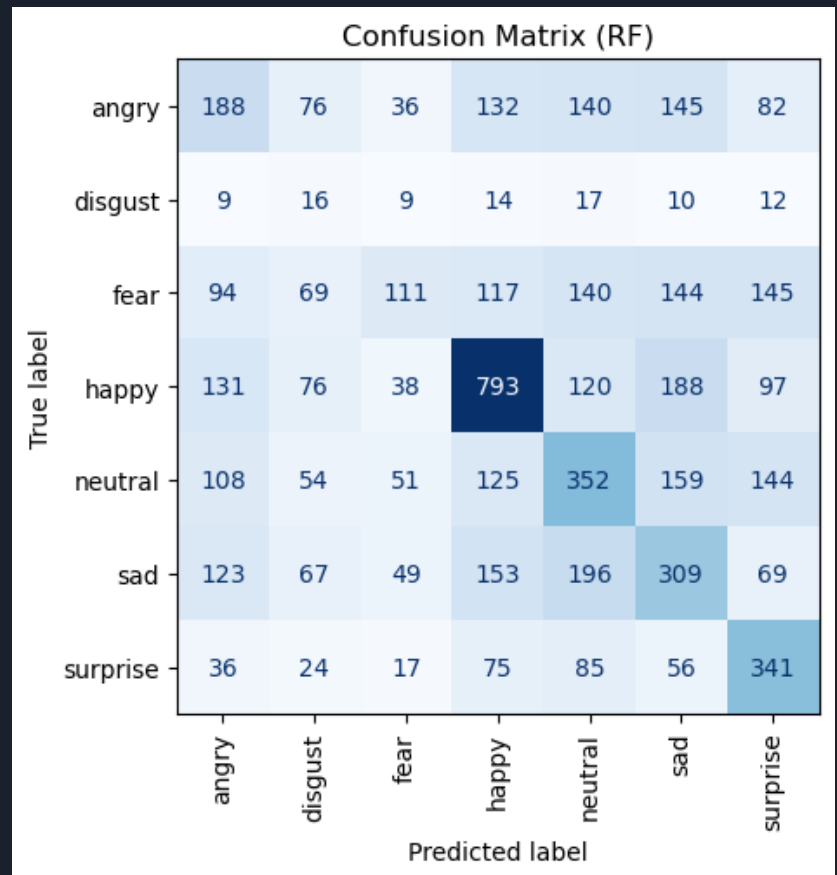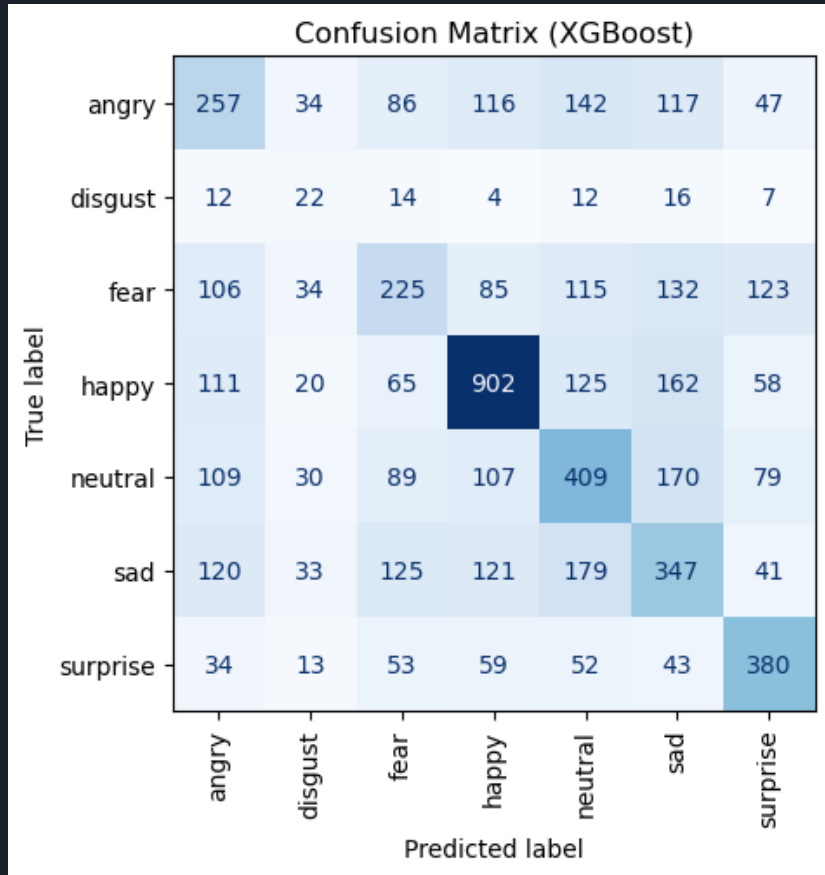

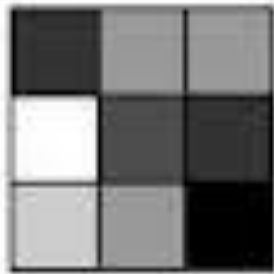Original Image - Label: sad

HOG Features

DCT of HOG Features

Classification with HOG + DCT

Feature Extraction:LBP

3 X 3 Neighborhood

Threshold 50

70	30	30
0	50	60
20	30	90

Binary 00011001

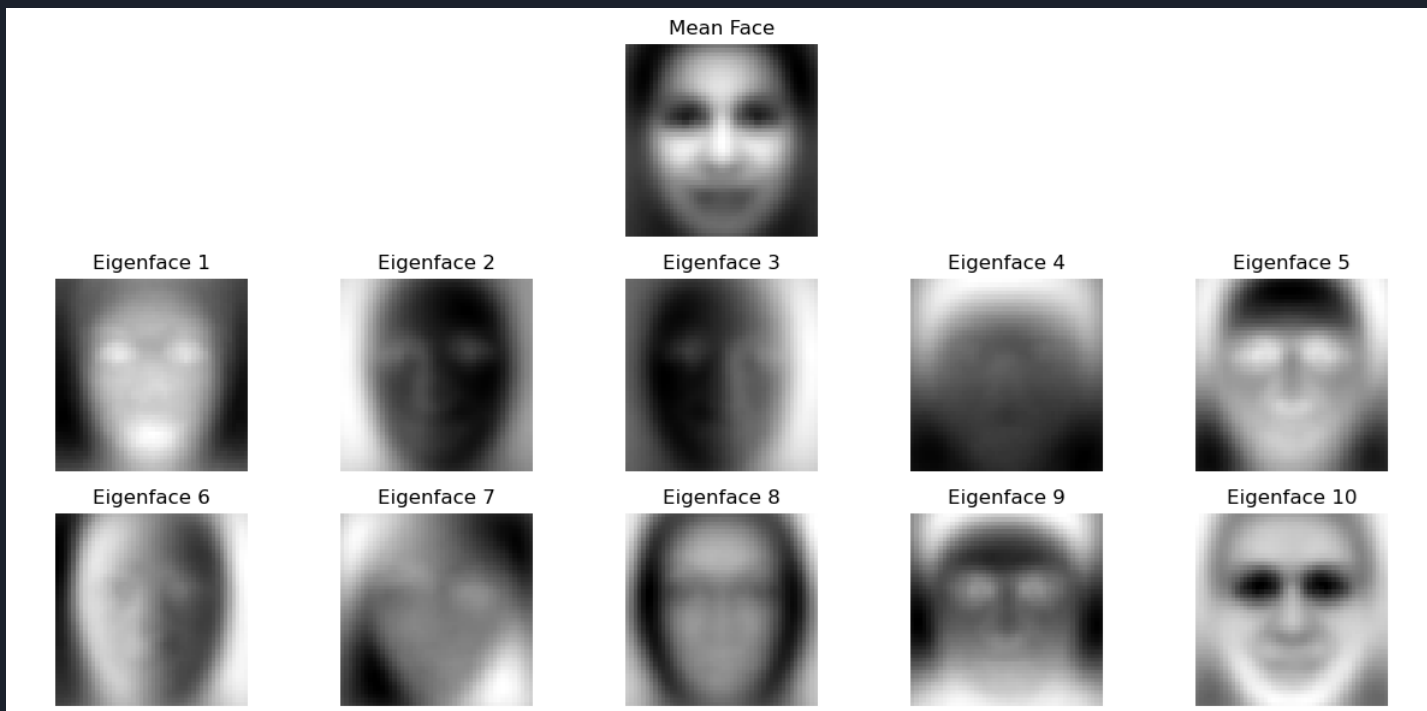| 1 | 0 | 0 |
| 0 | | 1 |
| 0 | 0 | 1 |

Pattern = 00011001
LBP = 1 + 8 + 16 = 25

Feature Extraction:Eigenface Analysis

- Eigenfaces are the "principal components" of a dataset of face images.

- Each eigenface represents a specific direction of variation in facial features, which has a good explanation of the variance.

Feature Extraction:Gabor Filter

- A linear filter used in image processing that's especially good at analyzing textures and edges.

- It essentially looks for specific frequencies and orientations within localized regions of an image for special patterns.

Original Image

Filtered Image
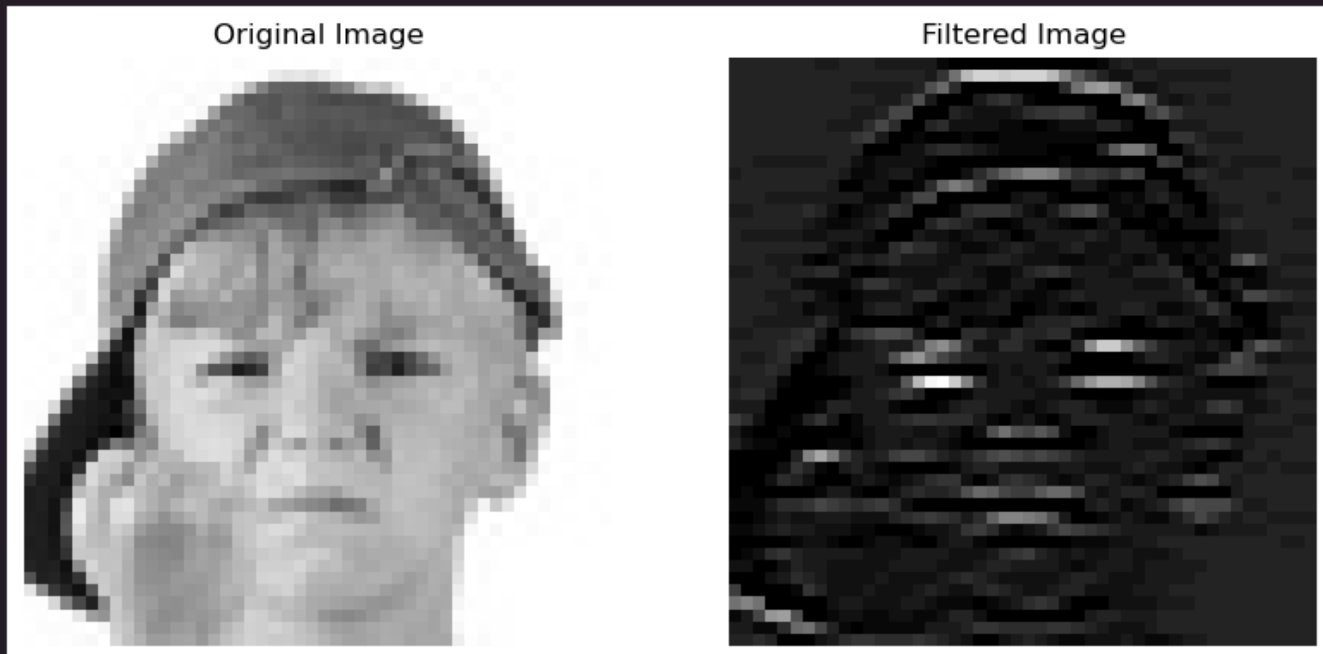
The Performance of Some Feature-based Classifiers

PCA + XGBoost: overall accuracy: 0.48
accuracy: 0.54

HOG & PCA & Gabor Filter & PCA + XGBoost: overall

Confusion Matrix (Percentages) (PCA + XGBoost)

Confusion Matrix (Percentages) (HOG & PCA & Gabor Filter & PCA + XGBoost)
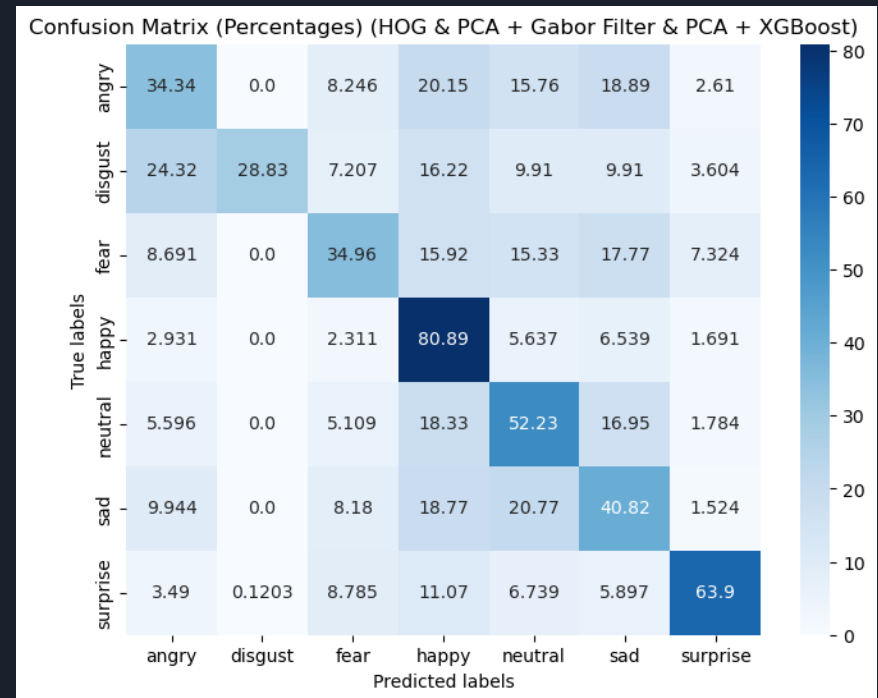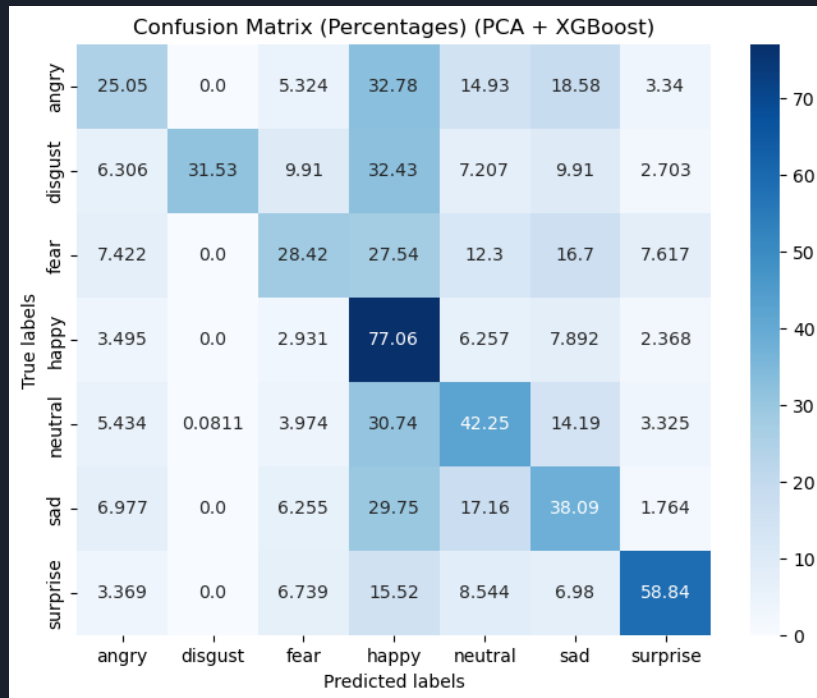
Feature Extraction:MediaPipe

- Extract 478 3D facial landmark points per image.
- Output: Landmark-based geometric features representing facial structure.
- Remark: PCA with 15 components captures the full variance of the MediaPipe features
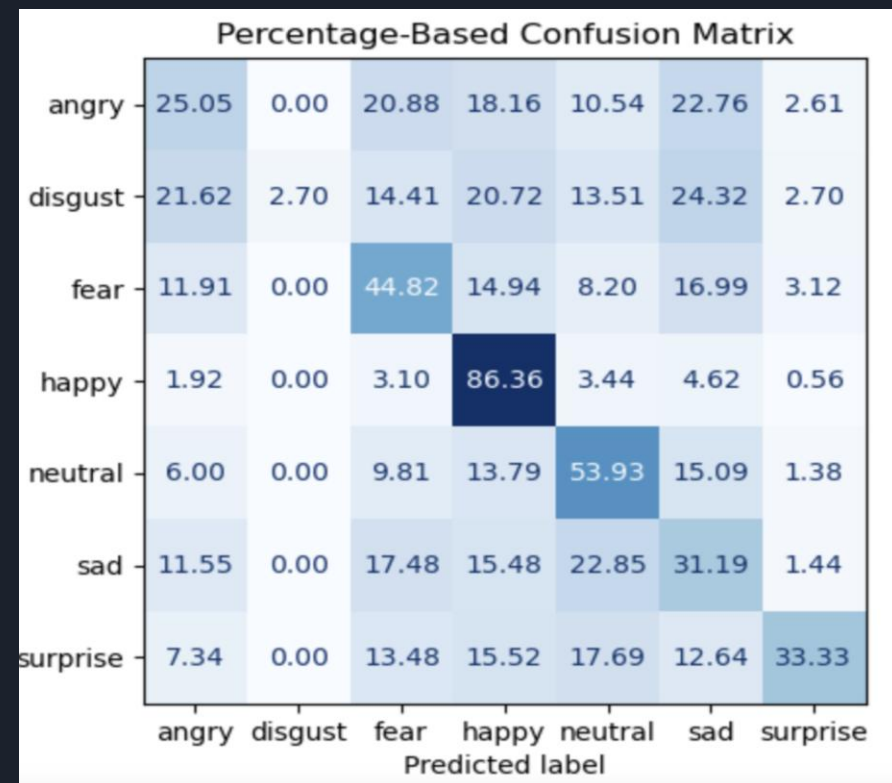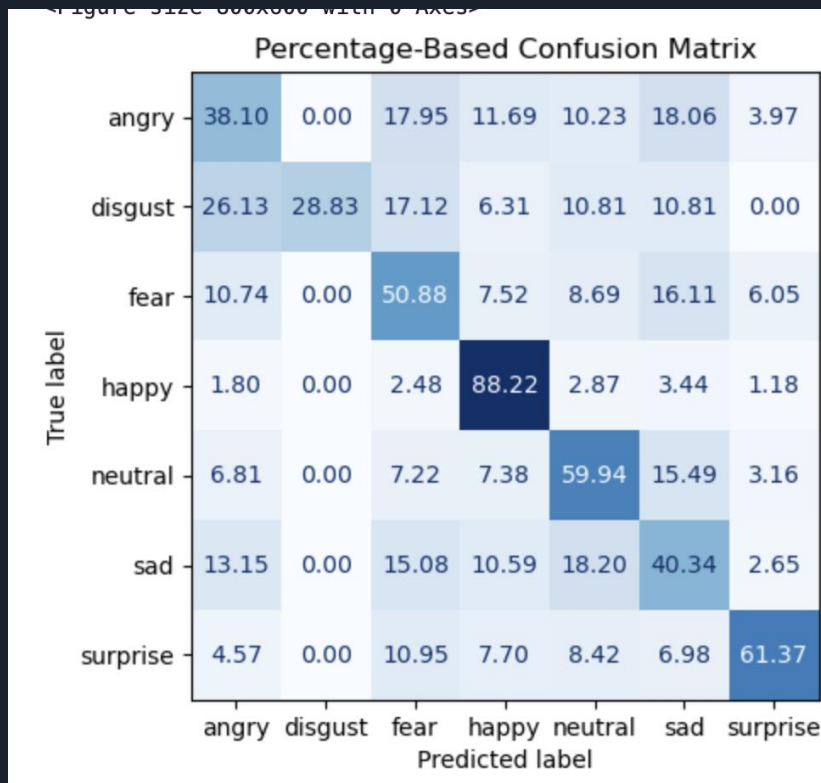

Percentage-Based Confusion Matrix

MediaPipe Facial Landmarks
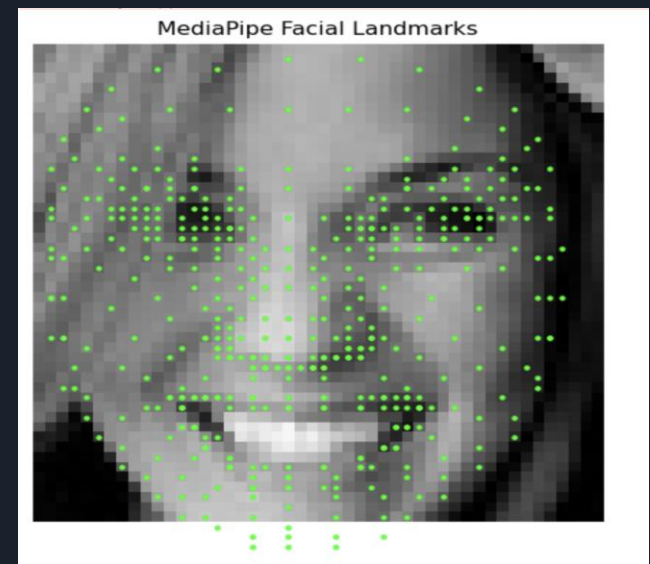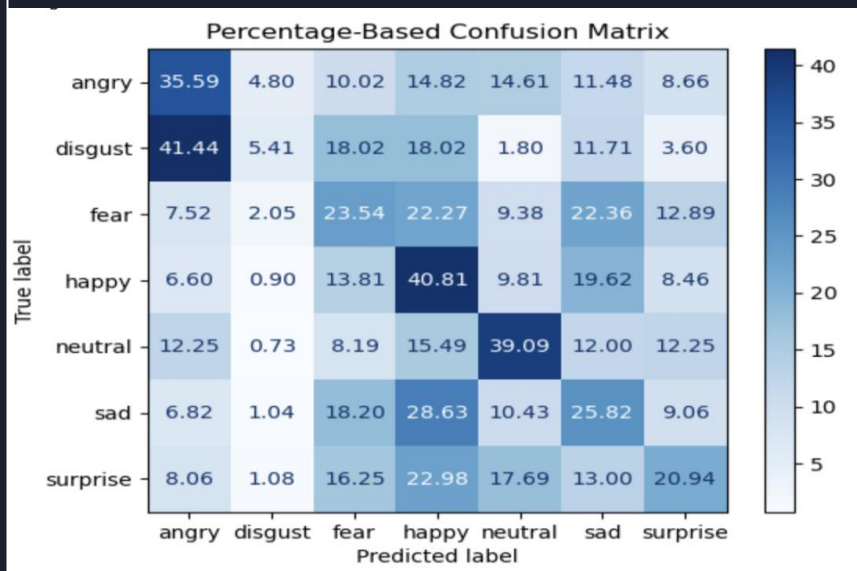
BaseLine Model: MediaPipe+PCA+KNN, overall accuracy = 34%

Classification Ensembles for MediaPipe

- PCA of 249 features with Gabor Filter captures 99.5% Variance
- Concatenate total 269 features from Gabor Filter and MediaPipe
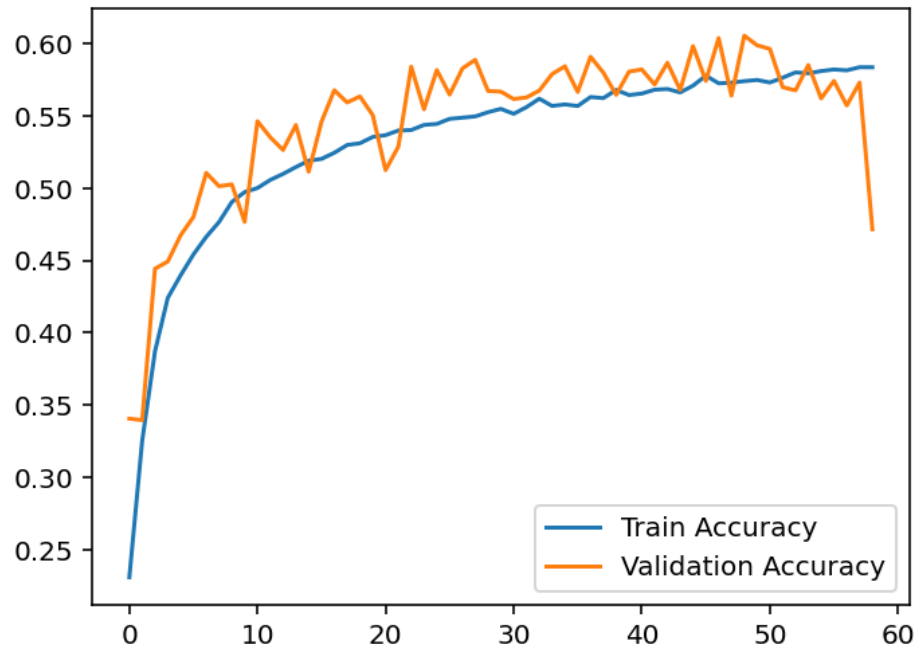- Apply Random-Forest and XGBoost as classifiers
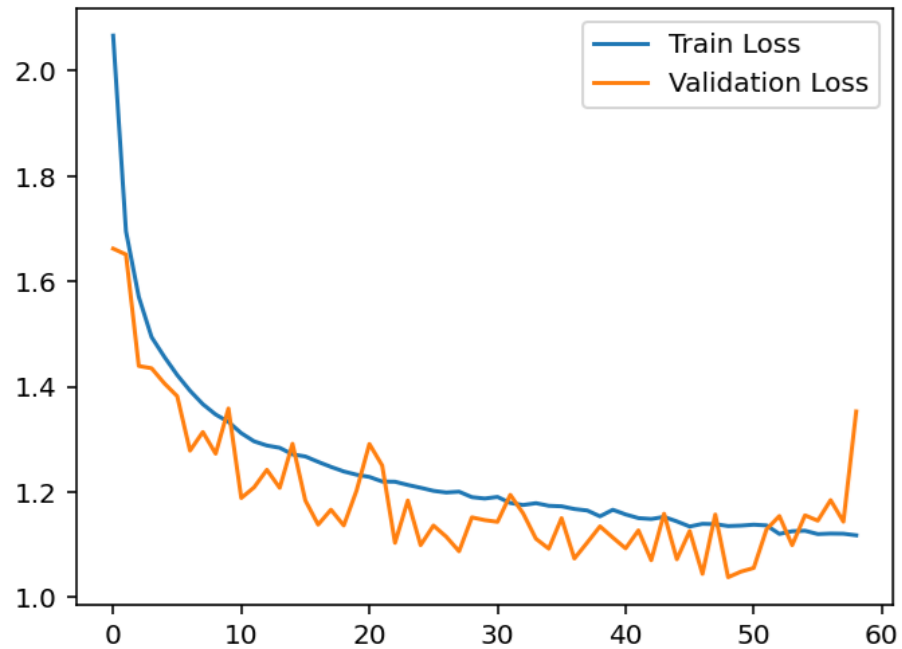
Deep Learning Approach: CNN

- - Training Strategy:

- * Loss function: Categorical cross-entropy.

- * Optimizer: Adam with learning rate scheduling.
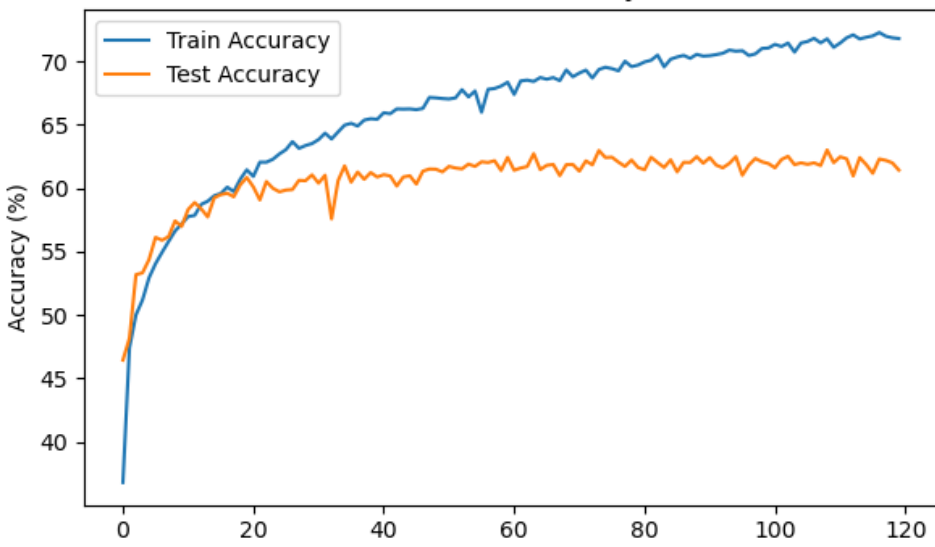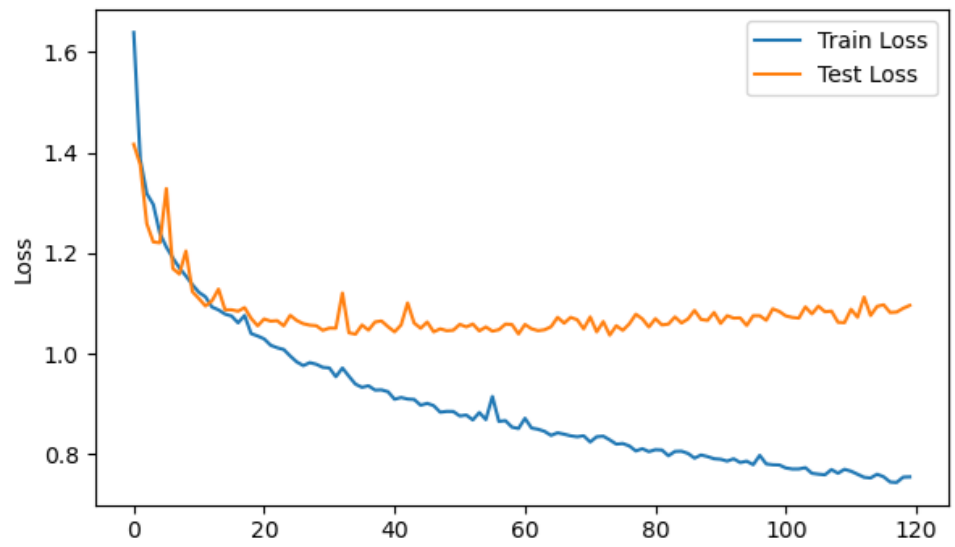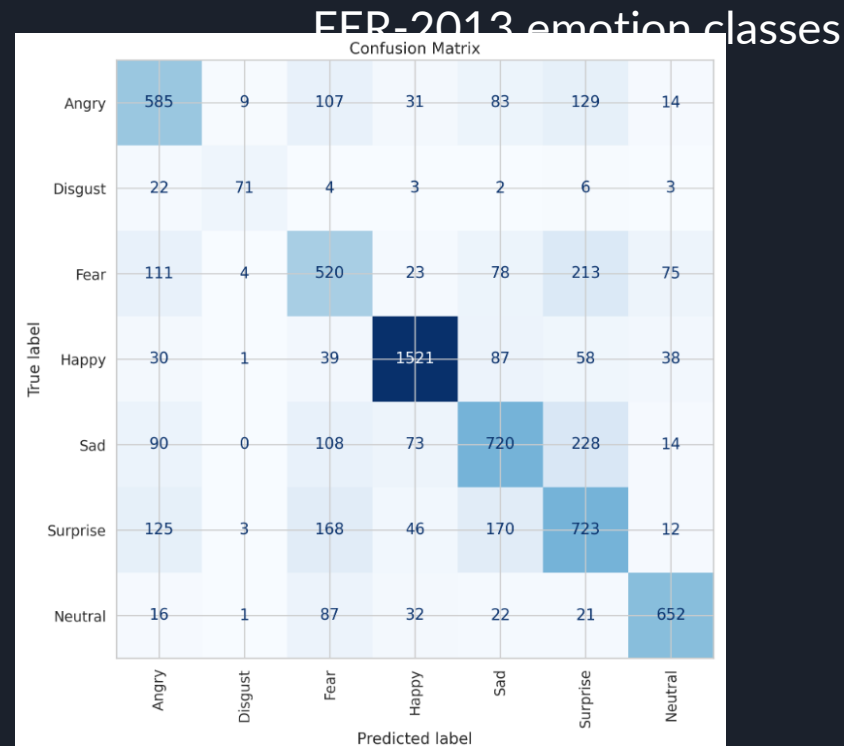
Deep Learning Approach: CNN

Transfer Learning Approach: ResNet50

ResNet50: a 50-layer deep convolutional neural network pre-trained on ImageNet

Fine-tuning for FER-2013: replace the fully connected layer to output the number of

FER-2013 emotion classes

Future Work

- Consider preprocessing techniques that account for low-resolution images

- Explore alternative methods such as EfficientNet or Vision Transformers

- Explore explainability techniques, such as Grad-CAM or SHAP, to better understand model predictions

- Investigate the integration of emotion recognition into real-world systems, such as interactive VR or mental health monitoring tools