Emotion Recognition using PCA and SVM with Unlabeled Data Augmentation

Pathri Vidya Praveen

June 12, 2025

Introduction

Facial emotion recognition is a fundamental task in Emotion AI with applications in human-computer interaction, mental health assessment, and surveillance systems. However, building robust emotion recognition models often requires large amounts of labeled data, which is expensive and time-consuming to collect. This project addresses the challenge of improving emotion classification performance with limited labeled data by leveraging unlabeled face images from the LFW dataset.

Our high-level approach involves three main phases:

1. Preprocessing both labeled and unlabeled facial images to a consistent grayscale format.
2. Performing Principal Component Analysis (PCA) on the unlabeled data to learn a robust face subspace.
3. Using this PCA basis to project labeled data and train a Support Vector Machine (SVM) for emotion classification.

Methodology

Data Preprocessing

All face images, including the 48x48 grayscale labeled CK+ samples and color LFW images, were resized to 64x64 and converted to grayscale. Each image was then flattened and normalized to [0,1]. This ensured uniformity in representation for PCA and SVM.

PCA Feature Engineering

To reduce dimensionality and improve generalization:

- We first evaluated different numbers of PCA components (from 20 to 200) by training SVM classifiers on labeled data using a cross-validation strategy.

- The optimal number of components was selected based on highest validation accuracy.
- Then, PCA was re-trained solely on the **unlabeled** LFW dataset (after preprocessing) to learn a generalized face representation.
- Labeled data was projected onto this face subspace to obtain lower-dimensional features.

SVM Classification

We employed Support Vector Machines as the classifier due to its effectiveness in high-dimensional spaces. We performed an exhaustive grid search over the following hyperparameters:

- Kernels: `linear`, `rbf`, `poly`
- Regularization parameter `C`: 0.1, 1, 10
- `gamma`: `scale`, `auto`
- Polynomial degree (only for `poly` kernel): 2, 3

After finding the best model through 5-fold cross-validation, we evaluated it on a held-out 30% test set.

Results

Best Parameters

- Kernel: linear
- C: 1
- Gamma: scale
- Degree: 2

Training set performance

- **Accuracy**: 0.777

Classification Report:

```
Classification Report on Training Data:
          precision    recall  f1-score   support

   Anger       0.50      0.31      0.38        45
Contempt       0.00      0.00      0.00        18
```

Disgust	0.79	0.56	0.65	59
Fear	0.82	0.36	0.50	25
Happiness	0.97	0.91	0.94	69
Neutral	0.83	0.98	0.90	593
Sadness	0.00	0.00	0.00	28
Surprise	0.97	0.87	0.92	83
accuracy			0.84	920
macro avg	0.61	0.50	0.54	920
weighted avg	0.80	0.84	0.81	920

Test Set Performance

- **Accuracy:** 0.7754

Classification Report:

	precision	recall	f1-score	support
Anger	0.40	0.15	0.22	13
Contempt	0.00	0.00	0.00	5
Disgust	0.55	0.61	0.58	18
Fear	0.00	0.00	0.00	8
Happiness	0.87	0.62	0.72	21
Neutral	0.80	0.97	0.88	178
Sadness	0.00	0.00	0.00	8
Surprise	0.89	0.64	0.74	25
accuracy			0.78	276
macro avg	0.44	0.37	0.39	276
weighted avg	0.72	0.78	0.74	276

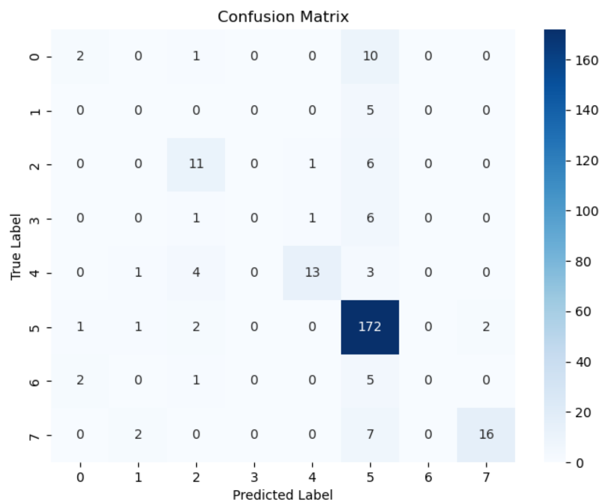

Figure 1: Confusion Matrix of Test Predictions

Discussion and Insights

Insights from SVD on Unlabeled Data

Singular Value Decomposition (SVD) revealed that a small number of principal components could represent a large portion of variance in unlabeled faces. This confirmed the viability of leveraging unsupervised data to learn a generic face manifold, which helped reduce noise and redundancy in labeled samples.

Effectiveness of Unlabeled PCA Projection

Projecting labeled data onto the PCA basis learned from unlabeled LFW images improved generalization. The unlabelled face space captured diverse variations in facial structure, aiding the SVM in learning discriminative patterns for emotion recognition.

Challenges Faced

- Many LFW images had irregular formats, requiring custom error handling in preprocessing.
- Class imbalance in the labeled dataset caused reduced precision in minority classes.
- SVM training was computationally expensive during grid search.

Limitations

- The labeled dataset was small and imbalanced.

- PCA assumes linearity; nonlinear embeddings (e.g., t-SNE, UMAP) might preserve structure better.
- Using PCA from LFW may not align perfectly with expression-focused variation.

Future Work and Improvements

- Explore deep learning-based autoencoders or pretrained CNNs for better representation.
- Investigate data augmentation techniques to balance class distribution.
- Replace SVM with ensemble classifiers (Random Forests, Gradient Boosted Trees) for comparison.
- Try to make the dataset consistent by taking training data such that all classes in the labeled dataset have same number of samples
- Try to improve more upon hyperparameters