ECS 271 - K-Nearest Neighbors Classification - Explanation and Results

Stephanie Lew

November 19th, 2025

K-Nearest Neighbors Classification

As a baseline for emotion recognition from facial geometry, we use a K-Nearest Neighbors (KNN) classifier. Each frame provides a 1,434-dimensional vector of 3D landmarks, which makes distance calculations noisy and computationally heavy. To stabilize the metric and remove redundancy, we standardize all coordinates and apply Principal Component Analysis (PCA) before running KNN. We also compute entropy scores to measure prediction confidence.

Data Preprocessing and Dimensionality Reduction

The raw landmark vectors are high dimensional so we first apply z-score normalization to place all coordinates on a comparable scale. PCA is then fitted on the training set and used to project each sample into a 50-dimensional subspace. As shown in Figure 1, the first component explains roughly 39% of the variance, the first five together capture about 95%, and by ten components the cumulative variance already exceeds 99%. Reducing to fifty components therefore preserves essentially the full signal while making Euclidean distances more stable and meaningful for KNN.

Entropy-Based Confidence

For each prediction, we compute Shannon entropy using the class probabilities obtained through distance-weighted voting. Low entropy corresponds to strong local agreement among neighbors, while higher values indicate uncertainty or ambiguity in the PCA space. The distribution in Figure 1 shows a large spike near zero entropy—cases where the classifier is highly confident and a noticeable spread extends up to about 1.75, consistent with overlapping class geometry and the mixed clusters shown in the PCA plots.

KNN Method and Model Selection

After PCA reduction, we train a Euclidean-distance KNN classifier with distance-weighted voting,

$$w_j = \frac{1}{d(x, x_j) + \varepsilon}, \qquad p_i = \frac{\sum_{j \in N_i} w_j}{\sum_{j \in N} w_j},$$

where N_i is the set of neighbors from class i. We evaluate a range of neighborhood sizes,

$$k \in \{1, 5, 10, 20, 50, 75, 100, 150, 200\},$$

computing the test accuracy and normalized entropy for each value. Once the best-performing k is identified, the model is retrained and we compute the final confusion matrix and entropy histogram. This allows us to examine both how k affects general performance and how the chosen model distributes its uncertainty across the test set.

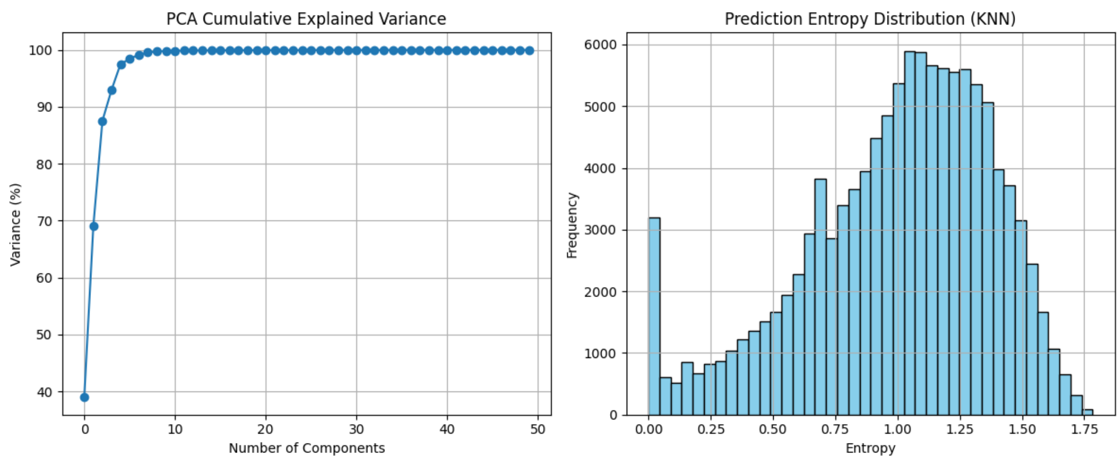

Figure 1: (Left) PCA cumulative explained variance, showing that 50 components retain nearly all variance. (Right) Prediction entropy distribution for the KNN classifier, indicating generally low but non-zero uncertainty.

Model Selection and Final Results

The best performance occurs at k = 50, which indicates that emotion classes are not tightly clustered in the PCA space. With small k, the classifier reacts strongly to local noise and small geometric fluctuations. Larger neighborhoods, however, smooth out these variations and provide a more reliable estimate of the dominant pattern around each point. The fact that performance peaks at a relatively large k suggests that the class structure is diffuse and that meaningful patterns only emerge when averaging over many neighbors.

Final Accuracy

$$\text{Accuracy} = 0.3712$$

For a six-class problem (chance ≈ 0.167), this is a measurable improvement over guessing and shows that facial geometry carries some usable information. However, the modest accuracy also reflects the difficulty of separating expressions when only shape-based features are used.

Confusion Matrix

$$\text{CM} = \begin{bmatrix} 6460 & 2586 & 3341 & 1434 & 2814 & 3881 \\ 2584 & 8133 & 4247 & 1694 & 2184 & 4223 \\ 2429 & 3108 & 6231 & 922 & 2619 & 4185 \\ 1595 & 1702 & 1435 & 8528 & 2355 & 1849 \\ 1050 & 1242 & 2173 & 1454 & 6594 & 3055 \\ 2603 & 2198 & 3936 & 778 & 3008 & 6954 \end{bmatrix}$$

Class order: Angry, Disgust, Fear, Happy, Neutral, Sad.

The confusion matrix shows a lot of cross-emotion mistakes. Disgust is often misclassified as Sad, and Fear also gets pushed toward Sad a lot, which suggests the model treats those emotions as visually similar.

On the other hand, some pairs barely get confused: Fear misclassified as Happy and Sad misclassified as Happy have the lowest off-diagonal counts. Overall, the matrix shows that the model struggles mainly with Disgust, Fear, Sad, and Happy while the other classes have somewhat cleaner rows but still plenty of leakage.

Entropy Analysis

The mean normalized prediction entropy is

$$\text{Entropy} = 0.3858.$$

The entropy histogram shows that most predictions have entropy values between roughly 0.5 and 1.4, rather than being concentrated near zero. This indicates that the KNN model is not extremely confident for most samples. Instead, it frequently assigns similar probability mass across multiple emotion classes, meaning the decision boundaries between emotions are often ambiguous.

PCA Visualization

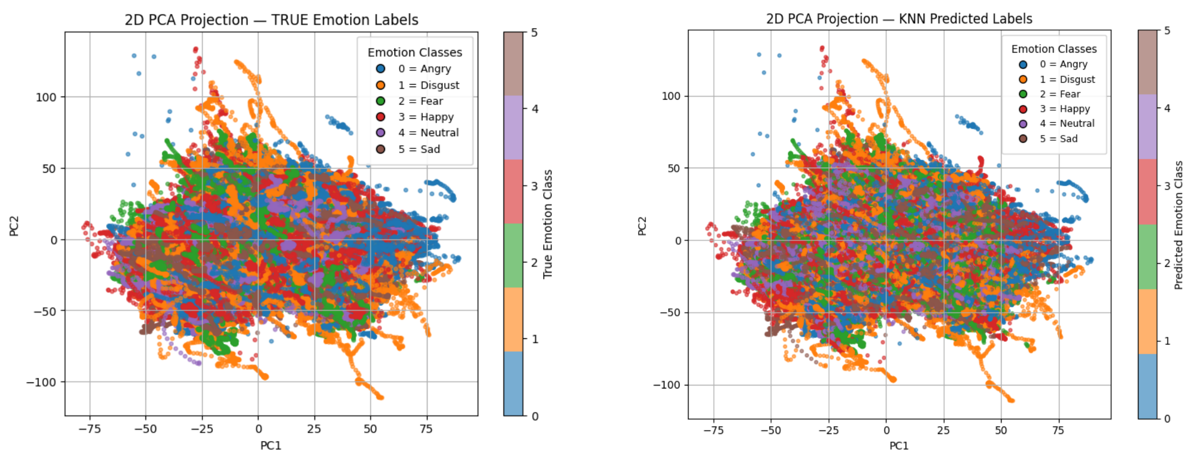

Figure 2: 2D PCA projection showing (left) true emotion labels and (right) KNN predictions. Strong class overlap explains limited classification performance.

The PCA plots show that none of the emotions separate cleanly. In the true-label plot, some emotions form loose clusters, but they still overlap heavily. The predicted-label plot is even more scattered—each emotion breaks into smaller blobs rather than forming any consistent region. This matches how KNN relies on local density rather than global structure.

Conclusions

PCA combined with KNN gives an interpretable baseline for seeing how much emotional structure appears in the 3D landmark data, but it isn't a strong classifier on its own. The model reaches 37.12% accuracy, which shows that posture and facial shape do contain some emotion-related cues, yet the PCA plots make it clear that most emotions overlap heavily. Geometry alone misses the subtle muscle movements and detailed shape changes needed to separate similar expressions.

Because KNN relies entirely on Euclidean distance in PCA space, it has trouble when classes are uneven or overlapping—which is exactly what we observe. The method works, but it's limited. Using a different distance metric or a classifier that adapts better to the structure of the data would likely perform much better. Overall, PCA+KNN is a useful exploratory tool, but not the best choice for a reliable emotion classifier.