

Abstract

Face recognition is a widely used technology in various fields, including security, user authentication, and video analysis. This study evaluates the comparative performance of classification algorithms—Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Decision Tree—in facial detection tasks. To enhance processing efficiency, Principal Component Analysis (PCA) is employed as a feature extraction and dimensionality reduction method, as several studies indicate that PCA can help reduce the dimensionality of large datasets, thereby improving the efficiency and accuracy of classification models.

The dataset used consists of 77,658 images, categorized into two classes: owner and strangers. The models are trained using hyperparameter tuning and evaluated through k-fold cross-validation, repeated 10 times to ensure a robust evaluation by averaging the results across the main performance metrics. Performance is measured using key metrics such as accuracy, precision, sensitivity, and f1-score.

Experimental results show that the Decision Tree algorithm outperforms SVM and KNN, achieving the highest accuracy of 97.81% on the third fold. For the owner class, this algorithm reached 91.25% precision, 68.05% sensitivity, and 77.94% f1-score. Meanwhile, for the strangers class, it achieved 98.11% precision, 99.60% sensitivity, and 98.85% f1-score.

This research provides valuable insights into the effectiveness of various classification methods in facial detection systems, and highlights the Decision Tree algorithm as the most effective in handling facial recognition tasks efficiently—although it still faces challenges in predicting data from the owner class. It is expected that future studies can further develop this research by utilizing larger and more diverse datasets, including variations in facial expressions, viewpoints, and lighting conditions, as well as improving owner face identification and implementing real-time face recognition systems.

Keywords: face recognition, support vector machine, k-nearest neighbors, decision tree, machine learning.