UNIVERSITÀ
di VERONA

Department of computer Science in Artificial Intelligence

MACHINE LEARNING PROJECT
Face & Non-Face Classification

KIDUS DAGNAW BELLETE
Kidusdagnaw.bellete@studenti.univr.it
ABEL ABEBE BZUAYENE
abelabebe.bzuayene@studenti.univr.it

Machine Learning Project Face and Non-Face Classification

Project Report

Objective:

The objective of this project was to explore different dimensionality reduction techniques and classification algorithms for face recognition.

Data Preprocessing:

- The dataset comprised images of faces and non-faces.
- Images were loaded and resized to a standard size.
- Data was shuffled to ensure randomness.
- Min-max scaling was applied to normalize the data.

Dimensionality Reduction:

- Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) were implemented.
- PCA retained 85% of the variance in the data, resulting in a reduced feature space.
- LDA aimed to maximize class separability.

Classification Algorithms:

- Three classifiers were evaluated: Logistic Regression, Support Vector Classifier (SVC), and K-Nearest Neighbors (KNN).
- Each classifier was trained and tested using three different datasets: original data with min-max normalization, PCA-transformed data, and LDA-transformed data.
- Confusion matrices are generated to visualize the classification results.

Results:

Classifier	Feature Extraction	Testing Score
Logistic Regression	Min-Max Normalization	95.47%
	PCA	95.88%
	LDA	76.95%
Support Vector Classifier	Min-Max Normalization	97.94%
	PCA	98.35%
	LDA	80.65%
K-Nearest Neighbors (KNN)	Min-Max Normalization	89.30%
	PCA	90.94%
	LDA	78.77%

Conclusion:

- Overall, PCA showed good performance across all classifiers, with minimal loss in accuracy compared to the original data.
- LDA, although effective in reducing dimensionality, resulted in lower accuracy compared to PCA and min-max normalization for logistic regression and KNN.
- SVC consistently outperformed other classifiers, achieving the highest accuracy across all datasets, indicating its robustness in handling both feature engineering techniques and classification tasks.

Recommendations:

- Based on the results, SVM with PCA-transformed data is recommended for face recognition tasks due to its superior performance.
- Further experimentation with different classification algorithms and tuning of hyperparameters may yield even better results.
- Exploring advanced techniques such as convolutional neural networks (CNNs) could potentially enhance accuracy further.

Limitations:

- The choice of parameters and techniques for dimensionality reduction and classification could influence results.
- There is a limitation of using high dimensional data in order to train and test the model (hardware limitation of GPU).