Face verification based on SVM classifier

Meryem Jabrane

University of Sheffield

Abstract

This report presents a lightweight face verification system designed for resource-constrained and data-limited scenarios. The system employs feature extraction techniques, including Pixel Difference, Principal Component Analysis (PCA), Local Binary Patterns (LBP), and Histogram of Oriented Gradients (HOG). Through experimentation and comparison of different techniques at each step of the pipeline, the proposed approach of using PCA and pixel difference for image processing to be then given to an SVM classifier with an RBF kernel optimised with Grid Search Cross-Validation, was found to outperform the remaining methods.

1. Introduction

Face verification is currently one of the most mainstream methods for identity verification, achieving remarkable accuracies with state-of-the-art techniques, while also being less intrusive compared to fingerprint or iris-based verification. However, while deep learning models such as Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) achieve high accuracy, they are computationally expensive and unsuitable for resource and data-constrained applications.

This report aims to address these challenges by developing a lightweight face verification system that utilizes handcrafted feature extraction methods (e.g., LBP, HOG) combined with a multilayer perceptron, enhancing robustness under conditions of limited data, low quality, and constrained resources. The remainder of this report is structured as follows: Section 2 provides a detailed description of the system, outlining the approach taken to address this implementation. Section 3 presents the experiments conducted to build a performant face verification system. Section 4 discusses the results and their analysis. Finally, Section 5 explores potential alternatives, future works, and concludes the report.

2. System Description

This face verification system aims to classify pairs of face images as either representing the same individual or different people. The system is composed of three major phases: image processing, feature extraction, which includes dimensionality reduction, and classification.

The pre-processing phase begins with normalization, where pixel values are scaled uniformly to ensure consistency. Each image is then flattened, transforming the 2D 62x47 representation into a 1D array, which facilitates subsequent operations.

The choice of the feature extraction technique is highly influential on the system performance. To emphasize on the dissimilarities between two images, in this implementation, a pixel difference feature is computed as the absolute difference between corresponding pixel values of the two flattened images. Mathematically, for two images I1 and I2 , the pixel difference D is given by:

$$D = |I1 - I2| \tag{1}$$

The final feature vector for each image pair is a concatenation of the two flattened images and their pixel-wise difference, to be provided for classification.

Given the high dimensionality of the feature vectors, Principal Component Analysis (PCA) is applied to reduce the feature space while retaining significant variance. PCA projects the original features into a lower-dimensional subspace using the formula: Z = X W where X is the original feature matrix, W represents the top k eigenvectors corresponding to the largest eigenvalues, and Z is the transformed feature matrix. In this implementation, k is set to 300, to allow for a good balance of information retention and lower computational consumption.

In this application, a Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel is employed for classification. The RBF kernel is defined as:

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right) \tag{2}$$

x_i and x_j are input feature vectors and γ is a parameter controlling the influence of individual data points. SVM fits appropriately in this approach for its suitability in handling non-linear relationships, generalisation with limited data while being resource-efficient.

In the process of ensuring model generalization, hyperparameter tuning is critical. Grid Search Cross-Validation minimizes the risk of overfitting by dividing the dataset into multiple training and validation folds during the hyperparameter search. This ensures that each hyperparameter configuration is evaluated across diverse subsets of the data, reducing reliance on any single validation set, particularly for sensitive models like SVMs [3].

3. Experiments

In scenarios characterized by limited data availability and the use of smaller models such as SVM, experimentation through trial and error becomes essential for maximizing performance. This approach will be implemented through systematic experimentation involving data augmentation, the exploration of diverse feature extraction techniques, and hyperparameter tuning.

Given the scarcity of training data, the first step was to augment the dataset using commonly employed techniques for face verification. Initially, a combination of small rotations ($\pm 5^\circ$), scaling ($\pm 10\%$), horizontal flipping, and brightness adjustments was applied to increase data diversity. Subsequently, the values and combinations of these augmentation techniques were varied to assess their impact on performance. However, contrary to expectations, these strategies resulted in unsatisfactory outcomes, as the model exhibited severe overfitting to the training data. This was evidenced by extremely low testing accuracy, suggesting that the augmented data failed to provide meaningful variations for the model to generalize effectively to unseen samples.

For the feature extraction methodology, four techniques were selected based on their suitability for low-quality data and their representation across three distinct categories: global feature extraction, local feature extraction, and pixel-wise feature extraction. The techniques include Pixel Difference, Principal Component Analysis (PCA), Local Binary Patterns (LBP), and Histogram of Oriented Gradients (HOG). Each method was chosen for its ability to capture complementary aspects of facial image data: Pixel Difference represents pixel-level intensity variations between image pairs, PCA captures global variance by reducing dimensionality, HOG extracts strong local edge features, and LBP encodes fine-grained texture patterns.

These techniques were chosen to ensure a comprehensive representation of facial features, as existing research has demonstrated their effectiveness in face recognition and verification tasks [3]. Experiments were conducted to evaluate these techniques both individually and in pairwise combinations. Specifically, HOG and LBP, as statistical approaches for local feature extraction, were used interchangeably with PCA and Pixel Difference.

The experimental results revealed that the combination of Pixel Difference with PCA retaining over 300 components, emerged as the most effective approach for achieving satisfying model generalization.

Among the various techniques available for this task, two were selected for experimentation: Random Search, a cost- and time-efficient method, and Grid Search, a systematic and high-performing approach. Both methods were employed in conjunction with cross-validation to optimize the hyperparameters.

The hyperparameters targeted for tuning are as follows: C , the misclassification penalty, which regulates the trade-off between model complexity and classification errors, and γ, the RBF kernel coefficient, which controls the flexibility of the kernel in fitting the training data.

Using Grid Search Cross-Validation, an exhaustive search was conducted over the ranges $C = \{1, 10, 100$ and $\gamma = \{0.001, 0.01, 0.1\}$, identifying the optimal parameters as $C = 10$ and $\gamma = 0.01$. A similar exploration was carried out using Random Search Cross-Validation with three trials. However, while Random Search is computationally economical, it demonstrated lower accuracy compared to Grid Search while the computational efficiency gains were negligible due to the relatively small dataset size, making Grid Search the preferred choice for this task.

4. Results and Analysis

The results demonstrate that the combination of Pixel Difference and PCA with an SVM classifier using an RBF kernel provided the best model generalization. Hyperparameter tuning through Grid Search Cross-Validation revealed that C and γ achieved the highest validation accuracy of 68

The confusion matrix in Figure 1 highlights the model's performance on the test set. The classifier achieved good differentiation between "same" and "different" classes, with 370 correctly classified "different" pairs and 310 correctly classified "same" pairs. However, misclassifications were observed in 130 "different" pairs and 190 "same" pairs, suggesting that the model occasionally struggles with edge cases.

These results indicate that the selected feature extraction techniques, combined with this hyperparameter tuning, effectively balanced model complexity and generalization. However, the observed misclassifications could be attributed to the limited size and quality of the training data.

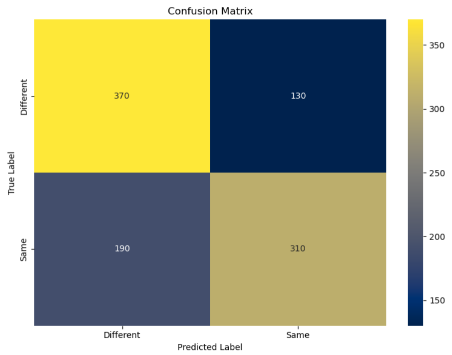

Figure 1: *Confusion Matrix of the proposed approach.*

5. Discussion and Conclusions

While the performance achieved demonstrates the model's capability under data and resource constraints, further improvements are possible. Thus, including more training data could significantly enhance performance. But, beyond data limitations, the use of alternative feature extraction methods not attempted in this experimentation may provide better resutls. Techniques such as Gabor filters, which capture spatial frequency and orientation, or lightweight deep learning approaches could improve performance without excessive computational overhead.

In addition, experimenting with ensemble methods, such as combining multiple classifiers, may improve robustness, as well as further experimentation with other Hyperparameter tuning techniques.

In conclusion, while the current system achieves reasonable accuracy using SVM classifier paired with PCA and pixel difference, enhancing training data, exploring advanced feature extraction techniques, and leveraging ensemble models could further improve its performance and generalization.

Table 1: *The classification accuracy per different values of C and γ .*

C	γ	Classification Accuracy
1	0.001	66%
10	**0.01**	**68%**
100	0.1	53%

[1] [2] [3]

6. References

[1] J. Olivares-Mercado, G. Aguilar-Torres, K. Toscano-Medina, M. Nakano-Miyatake, and H. Perez-Meana, "Gmm vs svm for face recognition and face verification," *Reviews, Refinements and New Ideas in Face Recognition*, pp. 1–338, 2011.

[2] P. Lameski, E. Zdravevski, R. Mingov, and A. Kulakov, "Svm parameter tuning with grid search and its impact on reduction of model over-fitting," in *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing: 15th International Conference, RSFDGrC 2015, Tianjin, China, November 20-23, 2015, Proceedings*. Springer, 2015, pp. 464–474.

[3] S. B. Chaabane, M. Hijji, R. Harrabi, and H. Seddik, "Face recognition based on statistical features and svm classifier," *Multimedia Tools and Applications*, vol. 81, no. 6, pp. 8767–8784, 2022.