Assignment #3: Face Recognition Pipeline

Ana Poklukar
IBB 2024/25 , FRI, UL
ap3956@student.uni-lj.si

I. Introduction

This report presents a face recognition pipeline evaluated through three stages: face detection, image-based recognition, and refined recognition on detected faces.

II. Methodology

In the first stage, we implemented the Viola-Jones algorithm for face detection. To enhance its performance, we optimized the algorithm's parameters using the training set. The detection accuracy was assessed using the Intersection over Union (IoU) metric, which measures the overlap between predicted and ground-truth bounding boxes. IoU values were reported for both the training and test sets to evaluate the detector's performance.

The second stage involved feature extraction for face recognition. We implemented three distinct methods to capture facial features: Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), and Dense Scale-Invariant Feature Transform (SIFT). LBP encodes texture information by thresholding pixel neighborhoods, creating a compact representation of local patterns. HOG captures shape and appearance by analyzing the distribution of gradient orientations in localized image regions. Dense SIFT extracts robust keypoint descriptors over a dense grid.

In the final stage, we evaluated recognition performance in two scenarios. First, recognition was performed on the entire images using each of the three feature extraction methods, with results visualized through Cumulative Match Characteristic (CMC) curves. Second, the complete pipeline was evaluated, where faces were first detected using the optimized Viola-Jones detector, resized, and then processed through the three feature extraction methods. The recognition performance of the full pipeline was also assessed and compared using CMC curves.

III. Experiments

The experiments for our face recognition pipeline were conducted using a subset of the CelebA-HQ dataset [1], comprising 857 images with predefined entities, face bounding boxes, and train-test splits. Of these, 475 images were allocated for training and 412 for testing. The pipeline, evaluated across face detection, feature extraction, and recognition stages, employed carefully tuned parameters for optimal performance. In the face detection stage, the Viola-Jones algorithm was optimized with a scale factor of 1.03, minimum neighbors of 5, and a minimum size of (550, 550) after testing various configurations. For feature extraction, LBP were configured with a radius of 1 and 8 points to effectively capture local texture patterns. HOG utilized $pixels\_per\_cell = (8, 8)$ and $cells\_per\_block = (2, 2)$. Dense SIFT employed a step size of 8.

IV. Results and Discussion

This section presents the evaluation results of our face recognition pipeline.

A. Results

The results of the face detection stage, using the Viola-Jones algorithm, are summarized in Table I. The table displays the Intersection over Union (IoU) values for both the training and test sets.

TABLE I
Intersection over Union (IoU) values for face detection performance on the training and test sets.

Dataset	Average IoU [%]
Training set	68.9
Test set	68.8

Figures 1 and 2 illustrate the Cumulative Match Characteristic (CMC) curves for face recognition performance under two different experimental setups.

The following tables show the Rank-1 and Rank-5 accuracy for different feature extraction methods, with and without the use of the Viola-Jones face detection algorithm. Table II presents the performance of the feature extraction methods when applied directly to the full image, without any face detection. Table III shows the performance after first applying the Viola-Jones algorithm to detect faces, followed by feature extraction. These evaluations were conducted on the test set consisting of 50 unique entities.

B. Discussion

The results highlight the importance of face detection and feature extraction methods in face recognition performance. The IoU values for both the training and test datasets

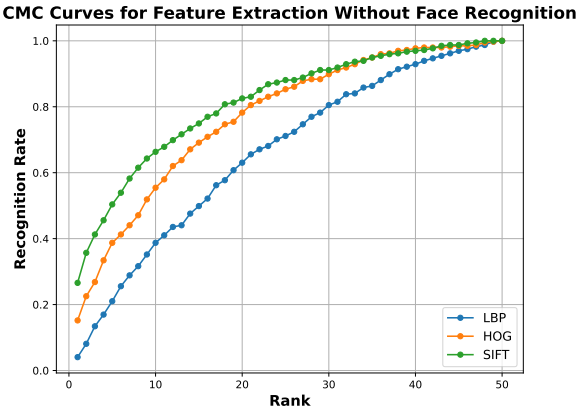

Fig. 1. CMC curve showing face recognition when it was performed directly on the full image, without any face detection. The curves demonstrate the recognition accuracy as a function of the rank, with the performance improving as more candidates are considered.

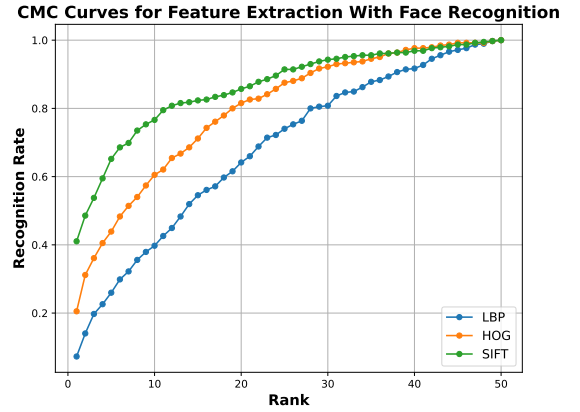

Fig. 2. CMC curve showing face detection when it was incorporated into the pipeline using the Viola-Jones algorithm, followed by feature extraction.

were relatively low (68.9% and 68.8%), suggesting room for improvement in face localization.

The CMC curves show that SIFT performed best, especially at lower ranks, followed by HOG and LBP. Without face detection, SIFT achieved the highest Rank-1 accuracy (26.58%) and Rank-5 accuracy (50.38%). When face detection via the Viola-Jones algorithm was applied, recognition accuracy improved across all methods. SIFT showed the greatest improvement, reaching 41.04% for Rank-1 and 65.19% for Rank-5, followed by HOG and LBP.

V. Conclusion

Our results demonstrate that incorporating face detection using the Viola-Jones algorithm significantly improves recognition accuracy across all feature extraction methods. While SIFT achieved the highest performance, followed by HOG and LBP, the overall accuracy was notably better

TABLE II
Rank-1 and Rank-5 accuracy for different feature extraction methods, without applying the Viola-Jones face detection algorithm.

Feature Extraction	Rank-1 Accuracy [%]	Rank-5 Accuracy [%]
LBP	4.05	21.01
HOG	15.19	38.73
SIFT	26.58	50.38

TABLE III
Rank-1 and Rank-5 accuracy for different feature extraction methods, after applying the Viola-Jones face detection algorithm.

Feature Extraction	Rank-1 Accuracy [%]	Rank-5 Accuracy [%]
LBP	7.27	25.97
HOG	20.52	43.90
SIFT	41.04	65.19

when face detection was applied. The CMC curves revealed that SIFT performed best at lower ranks, showing a clear advantage over HOG and LBP in both Rank-1 and Rank-5 accuracy metrics.

References

[1] "Celeba-hq-small dataset subset," https://tinyurl.com/celebahqsmall, accessed: 2024-11-29.