Assignment #3: Face Recognition

Marko Medved
IBB 2024/25 , FRI, UL
mm12755@student.uni-lj.si

***Abstract*—In this assignment, a face recognition pipeline was developed, starting with face detection using the Viola-Jones algorithm, which achieved an IoU of 0.689. Three feature extraction methods—LBP, HOG, and Dense SIFT—were then implemented and compared using Rank-1 accuracy, Rank-5 accuracy, and CMC curves. Among these, HOG demonstrated the best performance, achieving Rank-1 and Rank-5 accuracies of 45.4% and 64.6%, respectively. Notably, these results were obtained when the face detection step was omitted, suggesting that bypassing this step retained additional features beneficial for recognition.**

I. Introduction

The objective of this assignment was to implement a face recognition pipeline. The CelebA-HQ dataset was used for this purpose. The dataset was split into two parts: a training set for parameter optimization and a test set for evaluating the pipeline's performance.

II. Methodology

Face Detection:
The Viola-Jones algorithm was used to detect faces in images. If no faces were detected, the majority of the image was manually selected as a face. In cases where multiple faces were detected, the largest one was chosen. The parameters of the algorithm were optimized on the training split using grid search, with the Intersection over Union (IoU) as the evaluation metric. The IoU was then evaluated on the test split to assess performance.

Feature Extraction:
Three feature extraction methods were employed: Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), and Dense SIFT. Feature vectors were compared using the Chi-Squared distance and the distances were gathered into similarity matrices for each method. Then the matrices were grouped on 1 axis, by the identity of the person on the image, and by taking the minimum distance from the current compared image, among all images corresponding to the same identity. To continue, the pipeline's performance was assessed using the Cumulative Match Characteristic (CMC) curve in addition to Rank-1 and Rank-5 accuracies. Optimization of parameters for all three feature extraction methods was performed on the training split using grid search and the parameters were chosen based on the highest Rank-1 accuracy. The final evaluation was conducted on the test split by plotting the CMC curve and calculating the accuracies.To compare the impact of face detection, the entire process was executed twice: once including face detection in the pipeline and once omitting it

III. Experiments

The dataset consists of 887 images of 100 individuals, split into 475 images for training and 412 for testing. Each split contains 50 different individuals.
After optimizing the parameters of the Viola-Jones algorithm on the training set, the best parameters were determined as follows:

- **Scale factor**: 1.4
- **Minimum number of neighbors**: 4
- **Minimum face size**: Did not affect the IoU, as the largest detected box was always chosen when multiple faces were detected. To reduce computational load, a large minimum size of (500, 500) was used.
- **Maximum face size**: Not set, as faces occupied a significant portion of the images.

For LBP feature extraction, both with and without face detection, the optimal parameters were:

- **Radius**: 2
- **Number of points**: 16 (around the observed pixel)
- **Grid size**: 24×24

For the HOG method, both with and without face detection:

- **Number of orientation bins**: 9
- **Sizer of the block in cells**: (3,3)
- **Size of the cell in pixels**: (4,4)

And lastly, for the Dense SIFT method, the optimal parameters, also for both with and without face detection, were:

- **Step size:** 4
- **Overlap:** No overlap

IV. Results and Discussion

A. Results

Firstly, the performance of the Viola-Jones algorithm achieved an IoU of 0.689 during the testing stage.

Secondly, in Table I, we can compare the Rank-1 and Rank-5 performances of different feature extraction methods. The comparison is made for two configurations: first, including face detection in the pipeline, and second, omitting it.

On Figure 1 we can observe the comparison of CMC curves calculated for all 3 different feature extractors, when also including face detection in the pipeline. On Figure 2 we can similarly compare CMC curves however without including the face detection step in the pipeline.

B. Discussion

Table I demonstrates that the HOG method outperformed both dense SIFT and LBP in terms of Rank-1 and Rank-5 accuracies across both pipelines. Notably, the performance of LBP improved slightly with the inclusion of face detection, while the Rank-1 accuracies for HOG and dense SIFT showed a slight decline. This decline may be attributed to the characteristics of the dataset, where a significant portion of each image is already occupied by the face. As a result, applying face detection could inadvertently exclude peripheral regions that contain valuable features for the feature extraction process.

Figure 1 illustrates that when face detection is included, LBP performs the worst among the feature extractors at lower ranks. However, its performance improves at higher ranks, eventually surpassing both HOG and dense SIFT. A similar trend can be observed in Figure 2, although in this case, HOG maintains superior performance for a larger range of ranks, with LBP overtaking it only at the final few ranks.

Face Detection	Feature Extraction	Rank 1 [%]	Rank 5 [%]
yes	LBP	37.6	60.4
	HOG	44.4	66.2
	Dense SIFT	38.8	62.9
no	LBP	37.4	59.5
	HOG	45.4	64.6
	Dense SIFT	41.7	62.1

TABLE I
Comparison of Rank 1 and Rank 5 accuracies for different feature extraction methods, with and without the inclusion of face detection

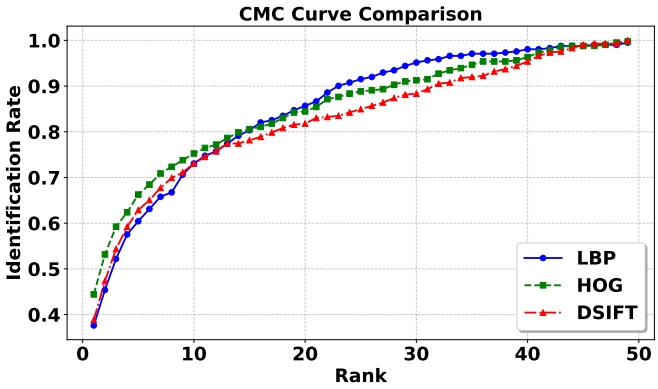

Fig. 1. Comparison of CMC curves when including face detection in the pipeline

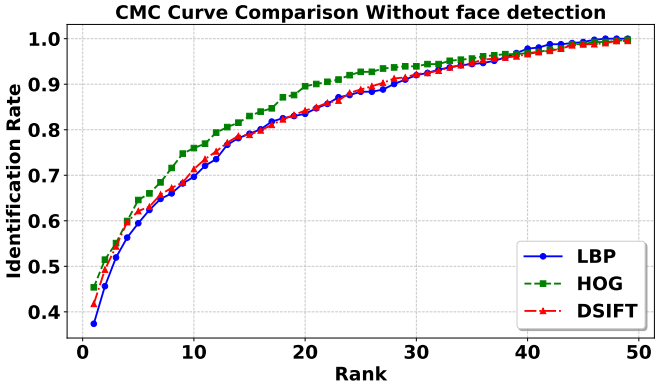

Fig. 2. Comparison of CMC curves, without including face detection in the pipeline

V. Conclusion

In this assignment, the Viola-Jones method was implemented and optimized for the CelebA-HQ database, achieving an Intersection over Union (IoU) score of 0.689. Following this, three feature extractors (LBP, HOG, and dense SIFT) were implemented, optimized, and evaluated using Rank-1 accuracy, Rank-5 accuracy, and cumulative match characteristic (CMC) curves. The performance of each feature extractor was compared both with and without face detection in the face identification pipeline. The results revealed that HOG outperformed the other feature extractors, with its Rank 1 accuracy being higher when face detection was omitted from the pipeline.