

Assignment #3: Face Recognition Pipeline

Žan Pušenjak
IBB 2024/25 , FRI, UL
zp4613@student.uni-lj.si

Abstract—Like in the previous assignment we used different feature extraction methods to detect identities of people based on their portraits. This time we used deep learning based model and compared the results with the ones we got with algorithm based approaches.

I. INTRODUCTION

We were provided with a database of portraits of 100 different people. We implemented a face recognition pipeline that will identify different people.

The pipeline uses various different methods to extract features from the image. Comparing the features will yield the identity of the person on the image, This time around we will be using pre-trained neural network feature extractors and detectors.

II. METHODOLOGY

We used two different models to detect the bounding boxes of the faces. The first was YOLO¹ and the second one was InsightFace's² face detector. InsightFace also provides other deep face analysis models like face swap and recognition. The later we used in the second part of the assignment. We determined the best detection by comparing their Intersection over Union (IoU) metrics on detected regions.

Next we used two models to extract features from the images. We used the feature extractors on the entire image and than only on the part that was detected by the face detection models.

For the feature extractors we used, already mentioned InsightFace's face recognition model and originally named FaceRecognition³ model.

We compared the feature vectors using cosine similarity to identify the person on the image. To better visualize the comparison between algorithm based and deep learning methods we again calculated different rank metrics and plotted a Cumulative Match Characteristic (CMC) cruve.

Rank-N metric represents the percentage of images, that are correctly classified in the N-th try, meaning the method will

determine their true identity to be in the top N identities most closely resembling the one on the image.

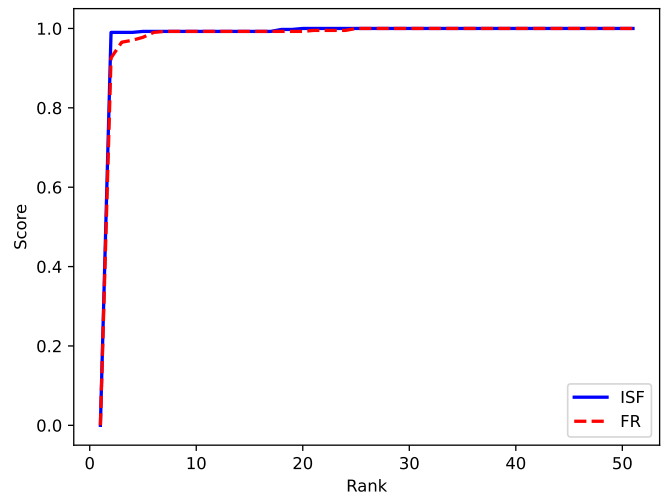


Fig. 1. The CMC curves of different models ran on the entire image

III. EXPERIMENTS

To make comparison between the algorithm based methods and deep learning models we only included the test set of images even if no training was required.

The pre-trained face detection models had problems detecting some faces. When inspecting those images, we concluded that if something was blocking the persons face, for example hands or the face was widely hidden behind hair, the model could not detect the face. In previous assignment we handled those exceptions by simply assigning a max rank detection. That gave us a weird spike at the end of the CMC curve. This time around we omitted such examples and got better results.

The face recognition models were used to extract the features, meaning we did not use the last layer of the network that also classifies the image. As mentioned we handled the classification or identification ourselves by looking at the cosine similarity.

¹<https://www.v7labs.com/blog/yolo-object-detection>

²<https://insightface.ai/>

³https://github.com/ageitgey/face_recognition

IV. RESULTS

Comparing the IoU metrics (Table I) of the detectors we saw that InsightFace's model performed better, so we used that model's face detections in later steps.

TABLE I
TABLE OF IOU METRICS OF DIFFERENT DETECTION MODELS

	IoU
YOLO	0.854
InsightFace	0.861

Since we used pre-trained models, no training was needed. Looking at the results, we can see that the Rank-1 metrics are already way above the algorithm based methods, some reaching 99% and all Rank-5 being above 99% (Table II). Comparing that to the previous results (Figure 3) shows the superiority of deep learning models.

TABLE II
TABLE OF RANK METRICS OF DIFFERENT MODELS

	Rank-1	Rank-5
InsightFace	0.99	0.99
FaceRecognition	0.926	0.99
InsightFace (on detections)	0.96	0.99
FaceRecognition (on detections)	0.93	0.99

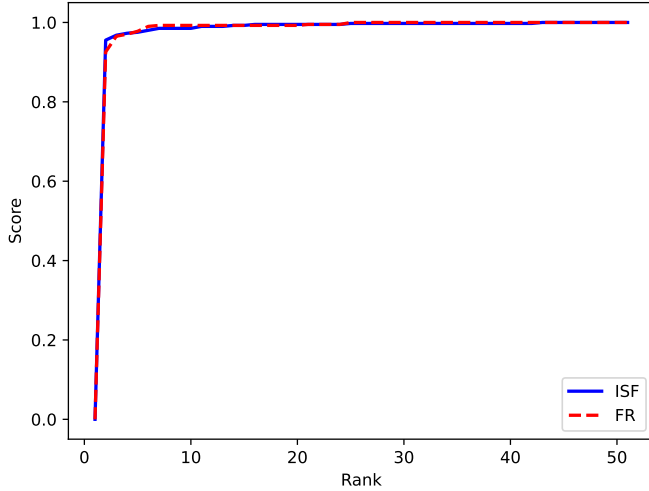


Fig. 2. The CMC curves of different models ran on the detected faces

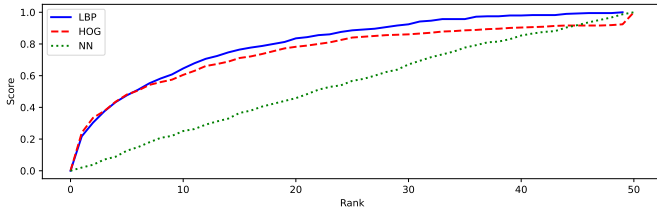


Fig. 3. The CMC curves of algorithm based methods ran only on the detected part of the image

V. CONCLUSION

When comparing the algorithm based methods and deep learning models we clearly see the difference in the results. Deep learning is in a league of its own and this comparison explains the rise in popularity of such models over the last ten or so years.

We can observe that the recognition worked a little better when using the entire image opposed to only the detected part. This is surprising and could be attributed to the fact that some models may also detect the face by themselves (even tho they are advertised to only extract features), so the first part of detecting the faces may be redundant and only made the feature extraction worse.