

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Deep-Facial Feature-Based Person Reidentification for Authentication in Surveillance Applications

Yogameena Balasubramanian, Nagavani Chandrasekaran, Sangeetha Asokan and Saravana Sri Subramanian

Abstract

Person reidentification (Re-ID) has been a problem recently faced in computer vision. Most of the existing methods focus on body features which are captured in the scene with high-end surveillance system. However, it is unhelpful for authentication. The technology came up empty in surveillance scenario such as in London's subway bomb blast, and Bangalore ATM brutal attack cases, even though the suspected images exist in official databases. Hence, the prime objective of this chapter is to develop an efficient facial feature-based person reidentification framework for controlled scenario to authenticate a person. Initially, faces are detected by faster region-based convolutional neural network (Faster R-CNN). Subsequently, landmark points are obtained using supervised descent method (SDM) algorithm, and the face is recognized, by the joint Bayesian model. Each image is given an ID in the training database. Based on their similarity with the query image, it is ranked with the Re-ID index. The proposed framework overcomes the challenges such as pose variations, low resolution, and partial occlusions (mask and goggles). The experimental results (accuracy) on benchmark dataset demonstrate the effectiveness of the proposed method which is inferred from the observation of receiver operating characteristic (ROC) curve and cumulative matching characteristics (CMC) curve.

Keywords: video surveillance, person reidentification, facial feature-based reidentification, Faster R-CNN, SDM

1. Introduction

Nowadays, a large network of cameras is predominantly used in public places like airports, railway stations, bus stands, and office buildings. These networks of cameras provide enormous video data, which are monitored manually and may be utilized only when the need arises to ascertain the fact. Fascinatingly, an automated analysis of such huge video data can improve the quality of surveillance by processing the video faster. Above all, it is more useful for high-level surveillance tasks like suspicious activity detection or undesirable event prediction for timely

alerts. Especially, the person Re-ID task is one of the current attentions in computer vision research. Establishing the correspondence between the image sequences of a person, across multiple camera views or in same camera at different time intervals, is known as person Re-ID. Simply, it implies that a person, seen previously, is identified in his/her next appearance using a unique descriptor of the person. Humans do it all the time without much effort. Our eyes and brains are trained to detect, localize, identify, and later reidentify the objects and people in the real world. Humans are able to extract such a descriptor based on the person's face, height and structure, attire, hair color, hair style, walking pattern, etc. However, a person's face is the most unique and reliable feature that human uses to identify the people [1]. Therefore, facial feature-based Re-ID is used to verify and recognize either the person seen in the camera is the same person spotted earlier in the same camera at a different time. Especially, it is applicable in controlled environment where the face database is available.

1.1 Facial feature-based person reidentification

In earlier days, it was stated that “reidentification cannot be done by face due to immature camera capturing technology” [2]. Nowadays due to remarkable growth of VLSI-based fabrication techniques, a person's face-capturing ability of camera has increased even in low illumination condition [3]. Therefore, facial feature Re-ID booms, and it is a well-authenticated one. Facial feature-based reidentification is a process of identifying a person using his/her face under consistent labeling across multiple cameras or even with the same camera to reestablish different tracks. Since the face is a biometric feature that cannot be replicated easily, it is used for human reidentification [4]. Also the face is the most natural and unique hallmark widely used as a person's identifier [5]. In reality, reidentification cannot be applied to find similarity among people after several days due to likely alterations in their visual appearance like attire, gait, etc. Li et al. [6] say that the face is also helpful in person reidentification and deserves attention. Li et al. [7] says the feature extracted from neck and above is an important clue for person reidentification. Biometric recognition features like the face, iris, and fingerprint can overcome these constraints by working on highly discriminative and stable features. Unlike the iris and fingerprint, to identify and recognize a person's “face” are successfully captured in the scene with improved camera technology. Beyond face recognition techniques, face reidentification techniques improve the system's metric learning and provide the best assurance to person's presence in the captured environment [8]. This proposed framework focuses on facial feature-based Re-ID for indoor surveillance such as IT sectors, government agencies, and ATM centers. The emergence of the facial feature-based person Re-ID task can be attributed to the increasing demand of public safety and the widespread huge camera networks in theme parks, university campuses, streets, IT sectors, etc. However, it is extremely expensive to rely solely on brute-force human labor to accurately and efficiently spot a person-of-interest or to track a person across cameras [9, 10]. Automation of the facial feature-based person Re-ID is quite difficult to be accomplished without human intervention. It is still a challenging topic, due to the fact that the appearance of the same face looks dramatically different in controlled or uncontrolled environments with pose variations, different expressions, illumination conditions, low resolutions, and partial occlusions specifically, in the abovementioned scenarios.

The rest of the chapter is organized as follows. In Section 2, prior research works on person reidentification including non-facial feature-based and facial feature-based Re-ID are summarized. Section 3 includes problem formulation, objective, and the key contribution toward this work. Section 4 elucidates the detailed

description of the proposed Re-ID framework. Section 5 presents the experimental results and discussion on face detection and Re-ID with challenging face detection benchmark datasets and TCE dataset. The step-by-step process of the proposed facial feature-based Re-ID framework's result for TCE dataset is also explored in Section 5. Finally, conclusions and the future research scope are presented in Sections 6 and 7, respectively.

1.2 Motivation

Three incidents in surveillance scenario motivate the research work toward person Re-ID. The first, being the London's subway bomb blast on July 7, 2005, where 52 persons were killed and 784 persons injured. It took thousands of investigators and several weeks to parse the city's CCTV footage after the attacks. The second, being the Boston Marathon bombing on April 15, 2013, where 3 persons were killed and 264 persons injured. Investigators had gone through hundreds of hours of video, looking for people "doing things that are different from what everybody else is doing." The work was painstaking and mind-numbing. One agent watched the same segment of video 400 times [11]. The third incident was the Bangalore ATM brutal attack on November 19, 2013, where one woman was seriously injured. The police commissioner of Bangalore expressed that in spite of all their sincere efforts, no arrest was made in the ATM attack case. However, they could identify the assailant only through CCTV footage. In all these three cases, the technology came up empty, even though the suspected images especially faces exist in official databases.

1.3 Applications

Facial feature-based person reidentification has various applications. It is applied in tracking a particular person across multiple nonoverlapping cameras and detecting the trajectory of a person for surveillance, forensic, and security applications. Further, in government offices and IT parks, the access card-based entry system can be replaced by facial feature-based Re-ID system to improve security and authentication.

1.4 Challenges

Facial feature-based person Re-ID as a task has many challenges such as varying poses, low resolution, illumination variations, different expressions, different hair-styles, wearing goggles, and occlusions. These challenges create intricacy in face detection and verification. In this chapter, the major challenges such as pose variations, partial occlusions, and wearing goggles are focused.

2. Related works

The person reidentification research started along with multi-camera tracking in the year 2005 [12]. Several important Re-ID directions have been addressed since then; some of them are based on camera setting, sample set, appearance-based, nonappearance-based, and body model as shown in **Figure 1**. Comparison of recent facial feature-based reidentification techniques are shown in **Table 1**.

Apart from facial feature-based person reidentification algorithms which suffer from noisy samples with background clutter and partial occlusion, it is problematic to differentiate an individual. Very few deep learning algorithms on "facial

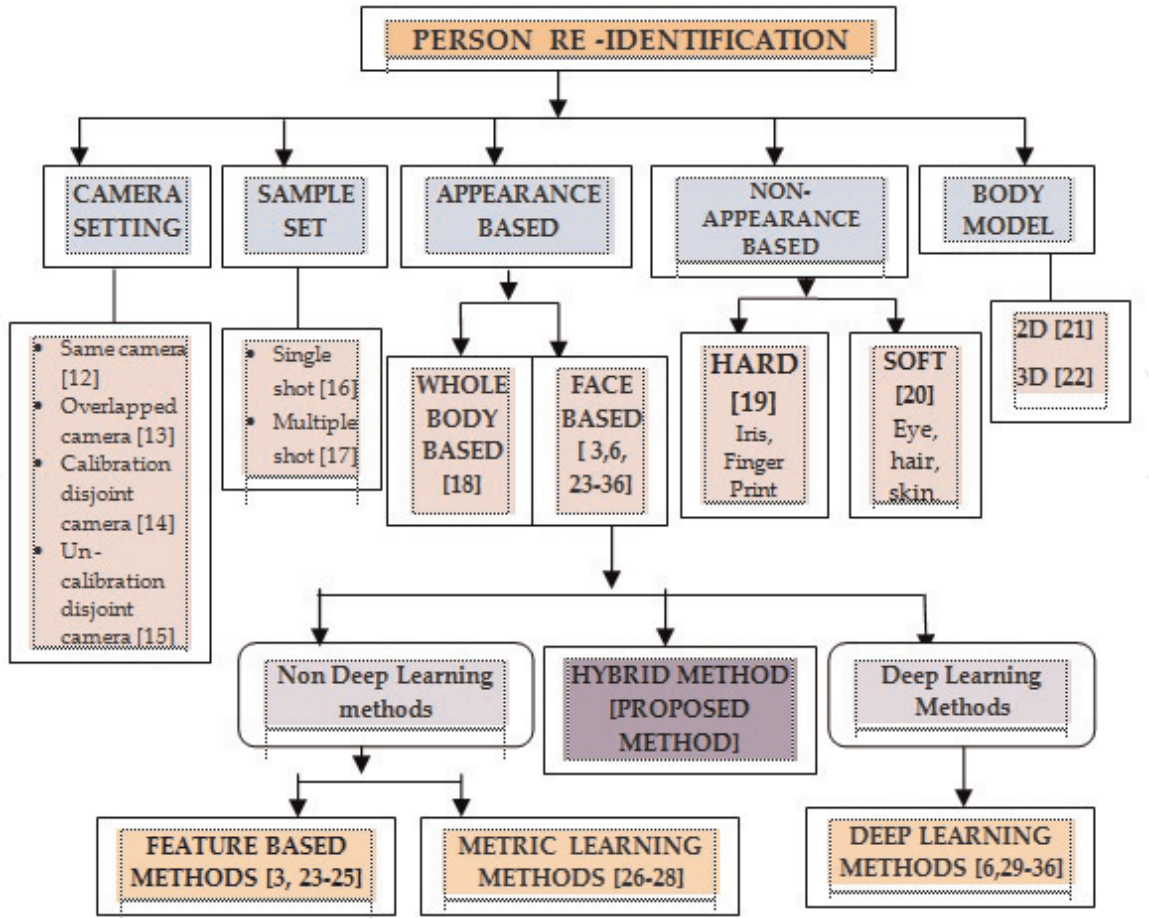


Figure 1.
Categorization of person reidentification algorithms [3, 6, 12–36].

feature-based” person reidentification are found in literature. However, deep learning features are heavily dependent on large-scale labeling of samples, they deal only with frontal and profile faces, and they fail under various illumination conditions, pose variations, and partial occlusions.

2.1 Observation and inference

From the existing related works, it can be concluded that very few works focus on deep learning methods for facial feature-based person reidentification. These works do not concentrate on the real-world challenges such as low image resolution, pose variations, and partial occlusions. Nevertheless, when we consider a controlled environment, such as authenticated laboratories and IT parks, face recognition-based person reidentification is possible which is vague currently. From the above discussion and analysis, a deeply trained facial feature-based person Re-ID framework is proposed which includes face detection by Faster R-CNN, joint Bayesian face-verification approach, and face reidentification. The scope of this chapter incorporates the challenges in the real-world environment like pose variation, low resolution, illumination changes, partial occlusion, and even goggle-wearing conditions.

3. Problem formulation

Existing works, related to the person Re-ID, deal only with the gait-based Re-ID for a short period, and very few works focus on long period reidentification of an individual. Research has been in progress toward long-term Re-ID (i.e., video is

TITLE	NON DEEP LEARNING METHODS	MERITS AND DEMERITS	TITLE	DEEP LEARNING METHODS	MERITS AND DEMERITS
	FEATURE BASED		[30]	Face Net	Online triplet method-face aligned matching Fails for crowded scenario.
[24]	HAAR+LBP(LAB) CS-LBP, XCS-LBP	Fast and Accurate face detection. When detection rate slightly increases, false positive also increases.	[31]	Deep transfer metric learning	Cross dataset verification Does not work in crowded scenario.
[25]	Active shape Model(ASM), Extended Active shape model	Ability to recognize 2D & 3D face images. This statistical model, deformed while fitting the target image.	[32]	Deep Face	Strong Abilities of independent learning & feature extraction. Complex and long process
[26]	Adaptive Ensemble	Efficient ranking system. Only highest KL divergence samples are preserved in memory for future adaptations.	[6]	Thermal face Net	Ability to detect spoofing attacks. Deal only thermal frontal images and not suitable for profile faces.
[3]	3D face recognition	Good success rate. Restricted only to KINECT bounded environment.	[33]	Hyper plane-VGG Net	Efficient for profile faces and pose variations. Fails with partial occluded faces.
	METRIC LEARNING BASED		[34]	Deep self paced Learning	Learning stable discriminative features Ranking performance is poor.
[27]	LMNN	Learn with Multiple features Do not recognize occluded Faces	[35]	Siamese DNN	Efficient for low illumination images Fails under pose variations.
[28]	ITML & LUPI	Learn an additional metric for privileged information. Learning Metrics are complex.	[36]	GAN	Best for pose variation and large viewpoint variation Huge training samples are required to improve the accuracy, which is challenging in case of surveillance dataset.
[29]	Multimodal distance metric learning methods	Efficient learning metrics Demands large computational time.	[37]	Multi-stream features distance fusion method	Excellent Extraction using body partition network Fails under partial occlusions.

Table 1.
Comparison of recent face reidentification techniques [3, 6, 24–37].

recorded for a month using a single camera), but at the same time, it is the need of the hour problem for authentication as well as for public safety. Here, facial feature-based Re-ID is the authenticated one, and other feature-based Re-ID is the suspicious one. Hence, there is a need to develop facial feature-based Re-ID using deep learning algorithm which handles low resolution, illumination variation, pose variation, and partial occlusion.

3.1 Objective

The main objective of the proposed framework is to develop facial feature-based person reidentification algorithm, using deep learning technology that works well

for long-term Re-ID even in low illumination, pose variation, partial occlusion condition (Goggles, Mask, etc.) for a controlled environment.

3.2 Contribution face-based: hybrid Re-ID method

The existing person reidentification is entirely based on global appearances or gait features. The prevailing algorithms have been developed so far to reidentify a person, based on his/her facial features that identify a person and do not address the experimentation on the challenging conditions such as low resolution, varying illumination, pose variations, and partial occlusion. This chapter proposes a hybrid combination of deep learning method Faster R-CNN for face detection and uses traditional method like joint Bayesian with SDM approach for reidentification which takes the advantages of both methods.

Moreover, another key contribution is the strong experimentation with benchmark datasets and TCE dataset captured under varying illumination conditions, with pose variations, various resolutions, and partial occlusion such as mask (green, blue, black shawl), specs, and goggles.

4. Methodology

The proposed facial feature-based person reidentification framework for surveillance applications in a controlled environment is portrayed in **Figure 2**.

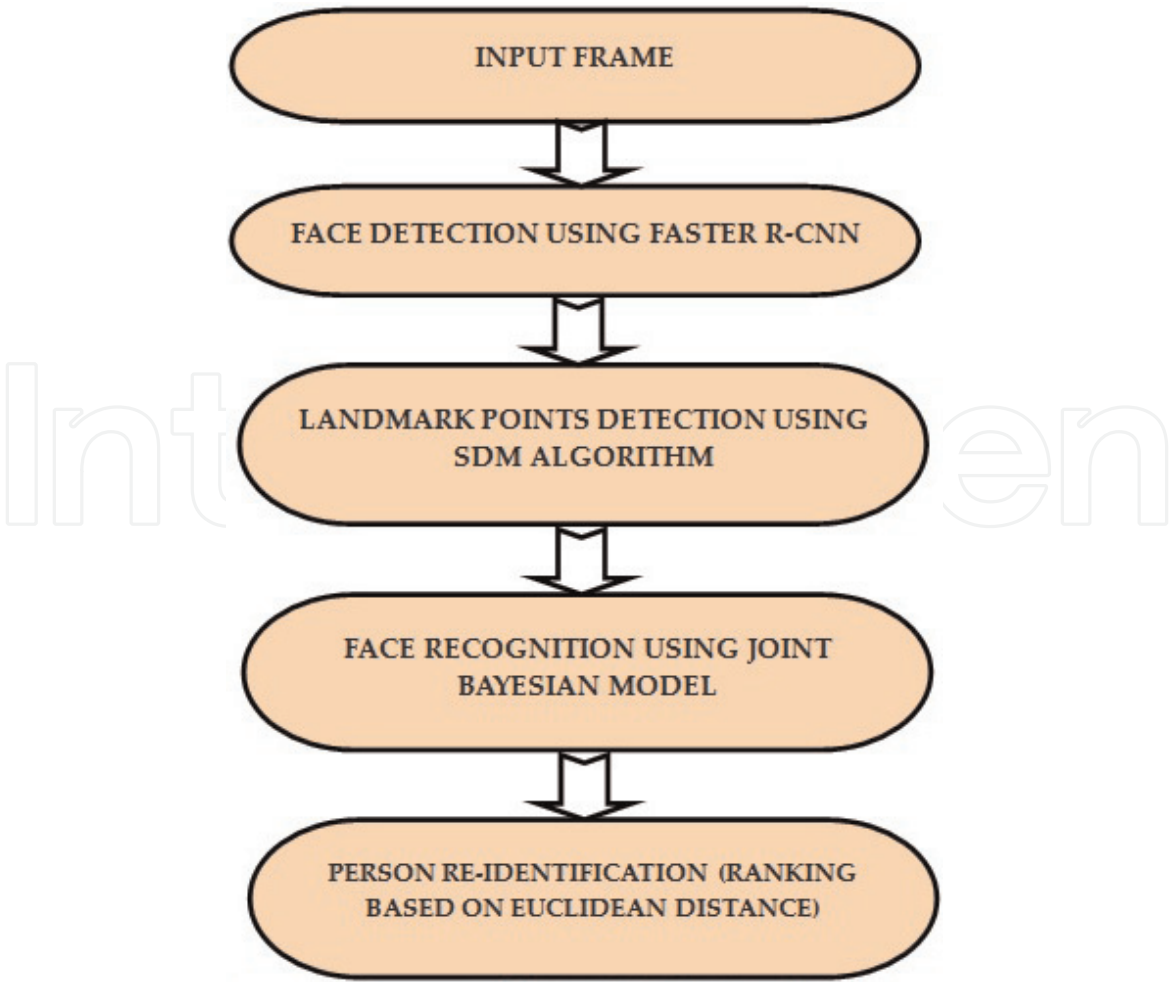


Figure 2.
Overview of the proposed deep-facial feature-based person Re-ID framework.

Here, the face detection module is implemented, by means of the deep learning-based approach (Faster R-CNN), where several convolutional and pooling layers are employed to extract deep features. Face recognition is performed, using the joint Bayesian model. Finally, the ranking is done, based on the similarity measure between the query image and the images in the database to provide a Re-ID. Finally, the ranking is done, based on the similarity measure between the query image and the images in the database to provide a Re-ID.

4.1 Overview of deep learning algorithms for face detection

After the remarkable success of a deep CNN in image classification on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012, Ross Girshick and his peers concluded that for a given complicated image, CNNs can be used to identify different objects and their boundaries in the image. Ross et al. [38] introduced a region-based CNN (R-CNN) for object detection. The pipeline consists of two stages. First, R-CNN creates bounding boxes, or region proposals, using a process called selective search. The selective search process identifies the object selecting the image area through the windows of different sizes, and for each size, it tries to group together the adjacent pixels by texture, color, or intensity. Once the proposals are created, R-CNN warps the region to a standard square size (e.g., 227×227) and passes it through to a modified version of AlexNet. On the final layer of the CNN, R-CNN adds a classifier that simply classifies whether this is an object, and if so, identifies the type of the object. The final step of R-CNN is to tighten the bounding box to fit the true dimension of the object. This is done, by using a simple linear regressor on the region proposal. The significance of the R-CNN is that it brings high accuracy by CNNs on classification tasks for the object detection problem. Its success is largely due to the act of transferring the supervised pretrained object representation for image classification. The R-CNN used different models to extract CNN-based image features, classify, and tighten bounding boxes. This makes the pipeline extremely hard to train these models. Ross Girshick, the first author of R-CNN, solved these problems, leading to the second algorithm—the Fast R-CNN [39]. Fast R-CNN uses a technique known as RoI Pool (region of interest pooling), which shares the forward pass of a CNN for an image across its subregions. For each region, the CNN features are obtained by selecting a respective region from the CNN's feature map. In addition, the Fast R-CNN jointly trains the CNN, classifier, and bounding box regressor in a single model. The R-CNN used different models to extract CNN-based image features, classify, and tighten bounding boxes, whereas Fast R-CNN used a single network to compute all these three. **Figure 3a** shows sample face detection results along with the confidence score using R-CNN. Even with all these advancements, there was still one remaining clog in the Fast R-CNN process, the region proposer. In the Fast R-CNN, these were done, using a slow process selective search, which was found to be the hindrance of the overall process. In [40], Ross Girshick and his team found a way to solve this problem and named it Faster R-CNN. The Faster R-CNN works to combat the complex training pipeline that both R-CNN and Fast R-CNN get exhibited. The slowest part in the Fast R-CNN was the selective search.

4.2 Face detection using Faster R-CNN

This chapter trains the Faster R-CNN on the existing benchmark datasets and in our TCE dataset for face detection. The input frames are resized based on the ratio $1024/\max(w, h)$ in order to fit it in the GPU memory, where w and h are the width and height of the image, respectively. The Faster R-CNN is designed to



Figure 3. (a) Face detection result using R-CNN for TCE dataset, (b) detected landmark points using SDM algorithm, and (c) ranking list of the TCE gallery set with similarity.

extract the visual features hierarchically, from local low-level features to global high-level ones, by using convolution and pooling operations. Region proposal network (RPN) is used to generate region proposals for faces in an image. In the RPN, the convolution layers of a pretrained network are succeeded by a 3×3 convolutional layer. This corresponds to map a large spatial window or receptive field (e.g., 227×227 for AlexNet) in the input image to a low-dimensional feature vector at a center stride. Two 1×1 convolutional layers are then added for classification and regression branches for all spatial windows. Here, the regions are positive if the sample is >0.5 (denoted as $L = 1$), when the region has an intersection over union (IOU) overlap with the ground truth and the regions are negative if sample is <0.35 (denoted as $L = 0$). The remaining regions are ignored [41].

Softmax loss function given by Eq. (1) is used for training the face detection task:

$$Loss = -(1 - L) \cdot \log(1 - p) - L \cdot \log(p) \quad (1)$$

In the aforementioned equation, p is the probability of occurrence of the candidate region, which is a required facial feature. The probability values p and $1 - p$ are obtained from the final fully connected CNN layer for the detection task.

4.3 Face recognition using SDM and joint Bayesian approach

After detecting the face and extracting the facial feature, the next task is recognition of face, i.e., the given face is verified with the class of faces (face verification) and certified with face identity (face identification). Face verification means verifying whether the given two faces belong to the same person or not. Face identification means an identity number is assigned to the probe person face with respect to the gallery. The conventional face recognition pipeline uses the facial features for face alignment and face verification. To detect facial landmark points SDM is used. SDM learns in a supervised manner generic descent directions and is able to overcome many drawbacks of second-order optimization schemes, such as non-differentiability and expensive computation of the Jacobians and Hessians. Moreover, it is extremely fast and accurate. This method improves the minimization of analytic functions that overcomes the problem of facial feature detection and tracking. SDM solves nonlinear least squares (NLS) and accurate in facial feature detection and tracking in challenging databases. SDM algorithm [42] detects facial landmarks as shown in **Figure 3b**. By detecting the landmarks, face images are globally aligned by similarity transformation. Further based on the extracted features, the face is recognized by joint Bayesian model [43]. The joint probability of two faces of the same or different persons is calculated, by using joint Bayesian model. The feature representation of a face is given as a combination of inter- and intrapersonal variations, or $f = \sum (\mu, \varepsilon)$, where both μ and ε are estimated from the training data and represented in terms of Gaussian distributions. Face recognition is achieved through log-likelihood ratio test, as given in Eq. (2):

$$\text{Log} \frac{p(f1, f2 | H_{\text{inter}})}{p(f1, f2 | H_{\text{intra}})} \quad (2)$$

Here, the numerator and denominator are the joint probabilities of two faces ($f1$ and $f2$), when given the inter- or intrapersonal variation hypothesis (H_{inter} and H_{intra}), respectively.

4.4 Euclidean distance-based reidentification process

Let us consider a probe person image p and a gallery set $G = \{g_i \mid i = 1, 2, \dots, n\}$, where n is the size of the gallery. Through computing their L2 (Euclidean) distances $d(p, g_i)$, the query result can be obtained as $R_p(G) = \{g_1^0, g_2^0, \dots, g_n^0\}$ where g_i^0 represents i -th image in the rank list and the distances between p and g_i^0 satisfy $d(p, g_1^0) < d(p, g_2^0) < \dots < d(p, g_n^0)$. Here a score $S(p, g_i^0)$ is used to define the similarity between p and g_i^0 , and it is equal to the rank index of g_i^0 . Based on the similarity score, a smaller distance indicates that the two images are more similar. Finally, all gallery images are ranked in ascendant order, by matching their L2 distances with the probe image to find out, which top n images can perform the corrected matches. **Figure 3c** shows the order in which the gallery images are ranked based on their similarity with the query image. The first image on the left corner has a higher similarity or a lower distance.

5. Experimental results

5.1 Dataset description

The HALLWAY, the WIDER FACE, Fddb, SPEVI (surveillance performance evaluation initiative) datasets are the benchmark datasets, used for face detection in this experiment. The HALLWAY dataset is used to evaluate person-to-person interaction recognition module. The WIDER FACE dataset is an effective training source for face detection. The WIDER FACE dataset is 10 times larger than existing dataset. The Fddb is designed for studying the problem of unconstrained face detection. It contains annotations for 5171 faces in a set of 2845 images taken from wild dataset. The SPEVI dataset is used for testing and evaluating target tracking algorithms for surveillance-related applications. Apart from these benchmark datasets, real-time TCE dataset is also used in this experiment. Sample frames of various benchmark datasets and TCE dataset is depicted in **Figure 4**. It consists of face images of various persons, captured under varying illumination conditions, with pose variations, various resolutions, and partial occlusion such as mask (green, blue, black shawl), specs, and black goggles. In TCE dataset, each row in figure corresponds to the same person, but the variations exist due to the difference in pose, viewpoint, illumination, image quality, and occlusion. Their corresponding specifications are given in **Table 2**.

5.2 Evaluation using benchmark and TCE dataset

This chapter considers a single-size training mode. **Figure 5a–c** brings out the sample detection results on the WIDER FACE, Fddb, and HALLWAY dataset, where the red color bounding boxes are ground-truth annotations and the yellow color bounding boxes are the detection results, using Faster R-CNN. Finally, more



Figure 4. Sample frames with challenging conditions (a) HALLWAY, (b) and (c) WIDER FACE, (d) Fddb, (e) SPEVI, and (f) TCE dataset.

DATASETS	IMAGE/ VIDEO	SCENARIO	FRAME SIZE	NUMBER OF FACES	FRAME FORMAT
Hallway	Video(17 FPS)	Indoor	320*240	Interaction faces	.avi
WIDER FACE (Press Conference)	Image	Indoor & Outdoor	576*1024 to 1544*1024	393,703 labeled faces	.jpg
WIDER FACE (handshaking)	Image	Indoor and Outdoor	295*1024 to 1249*1024	90	.jpg
Fddb	Image	Indoor and Outdoor	450*320 (approx.)	5,171 Faces	.jpg
SPEVI 1	Image	Indoor	21*53 to 176*326	3081 faces	.jpg
TCE	Image	Indoor	576*576	300 faces	.jpg

Table 2. Specifications of various benchmark datasets and TCE dataset.

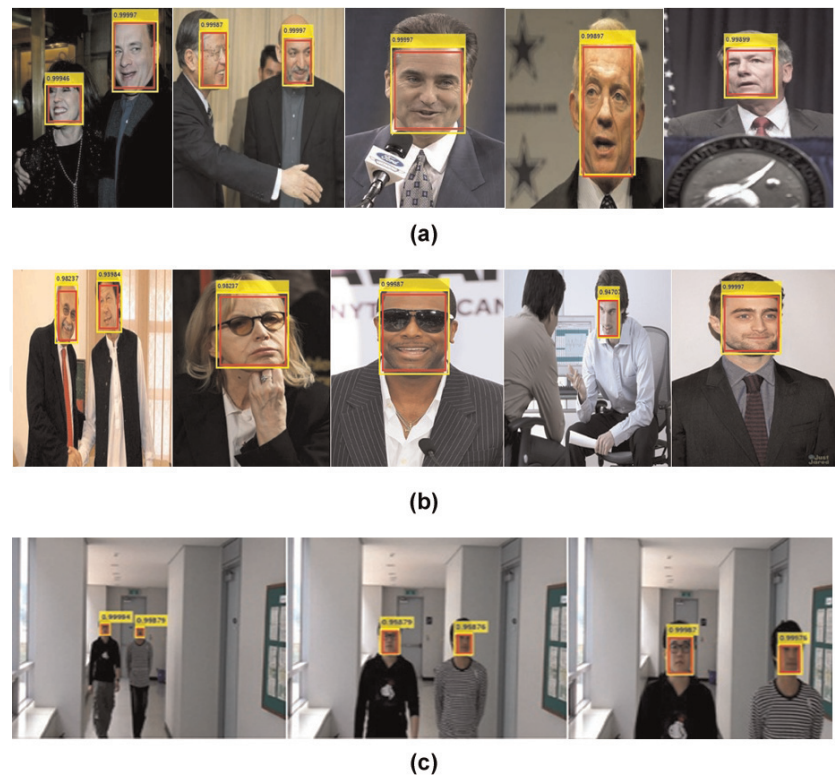


Figure 5. Sample detection results on the various dataset, where red color bounding boxes are ground-truth annotations and yellow color bounding boxes are detection results using Faster R-CNN sample detection results using Faster R-CNN, (a) WIDER FACE dataset, (b) FDDB dataset, and (c) HALLWAY dataset.

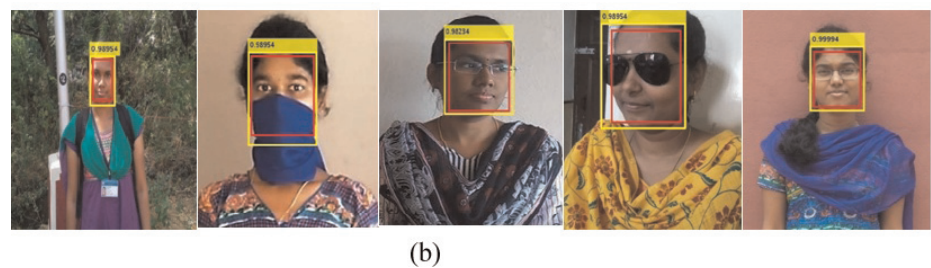
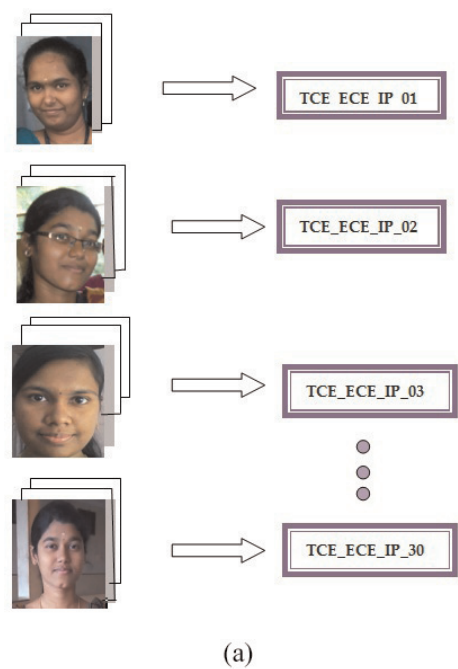


Figure 6. (a) TCE dataset gallery—persons with ID and (b) sample detection results using Faster R-CNN-TCE dataset.

number of faces are trained and learned, and the experiments prove that Faster R-CNN achieves highly triggering results against the other state-of-the-art face detection methods.

Apart from the above benchmark datasets, our approach is evaluated on TCE dataset. It is captured to test all the challenges in one single dataset which is absent as benchmark. The gallery of the TCE dataset consists of the images of 30 students, under varying pose conditions, illumination variations, and occlusion conditions. For each student, at least 300 images are tested under those conditions. Moreover, an ID is provided for each student in the database such as TCE_ECE_IP_01, TCE_ECE_IP_02, TCE_ECE_IP_03... TCE_ECE_IP_30 (as shown in **Figure 6a**). Once a student enters the lab, her face is detected using Faster R-CNN. **Figure 6b** shows some of the sample detection results on the real-time TCE dataset, where the red color bounding boxes are ground-truth annotations and the yellow color bounding boxes are detection results, using Faster R-CNN.

The detected face is recognized, using the joint Bayesian model after finding facial landmarks, by means of the SDM algorithm. Afterward, the images in the

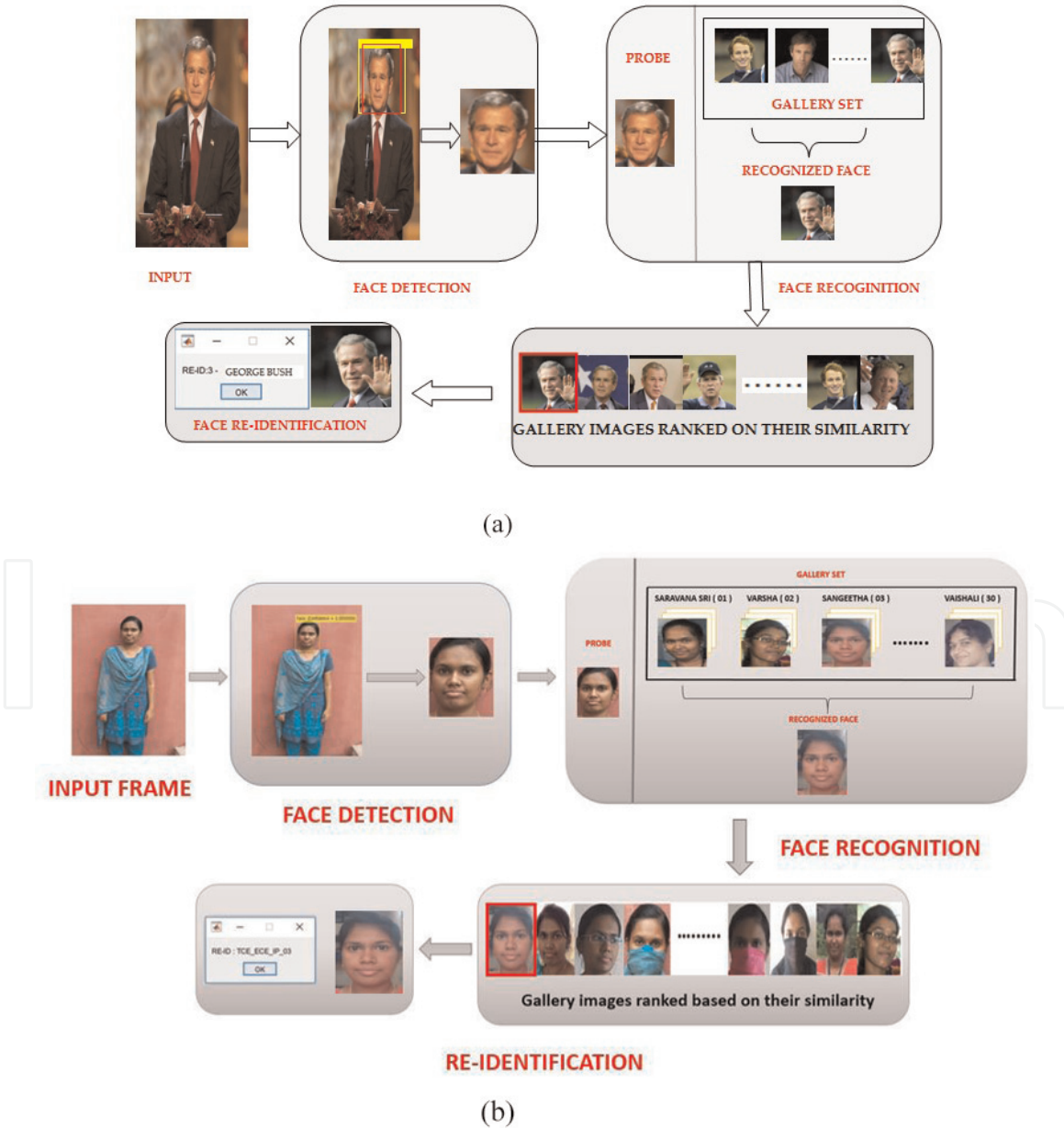


Figure 7.
The proposed facial feature-based Re-ID results for LFW and TCE dataset.

gallery set are arranged, based on their similarity. Finally, from the ranking list, the image with lower distance (rank 1) or with higher similarity score is displayed along with the Re-ID. The overall schematic representation of the proposed framework's result for a sampled query frame is shown in **Figure 7**.

5.3 Comparative analysis

The performance of face detection is measured in terms of recall and intersection over union (IoU). Each detection is considered as positive, if the IoU ratio is >0.5 , matched with ground-truth annotation. The threshold of the detected scores is varied to generate a set of true positives and false positives. Finally, ROC curve is plotted. The larger the threshold is, the fewer the proposals that are considered to be true objects. **Figure 8a** and **b** illustrates the quantitative comparisons of using 300–2000 proposals. RPN is compared with other approaches including selective search (SS) and edge box (EB), and the N proposals are the top N-ranked ones, based on the confidence generated by these methods. The recall of SS and EB drops

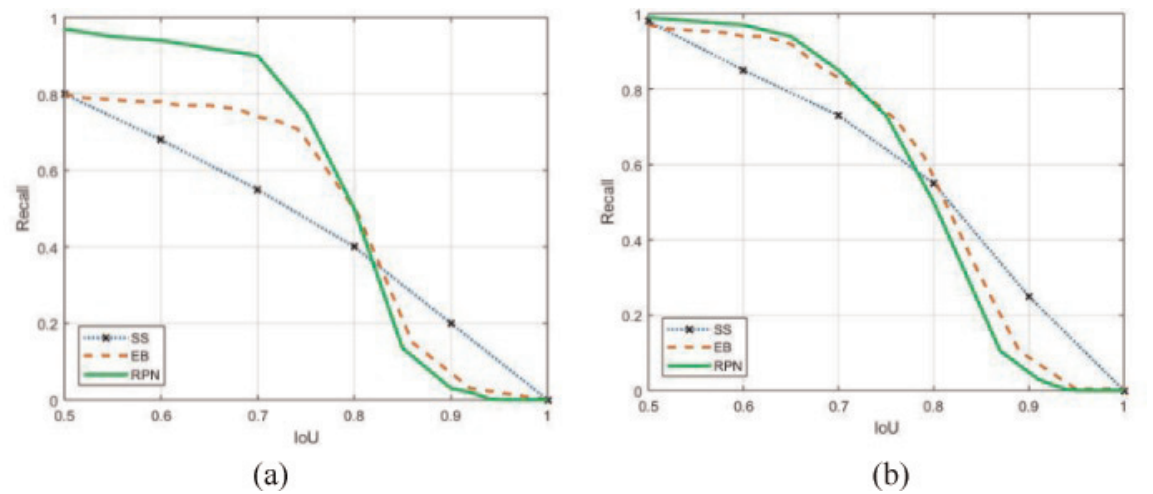


Figure 8.
(a) Recall vs. IoU overlap ratio with 300 proposals and (b) recall vs. IoU overlap ratio 2000 proposals.

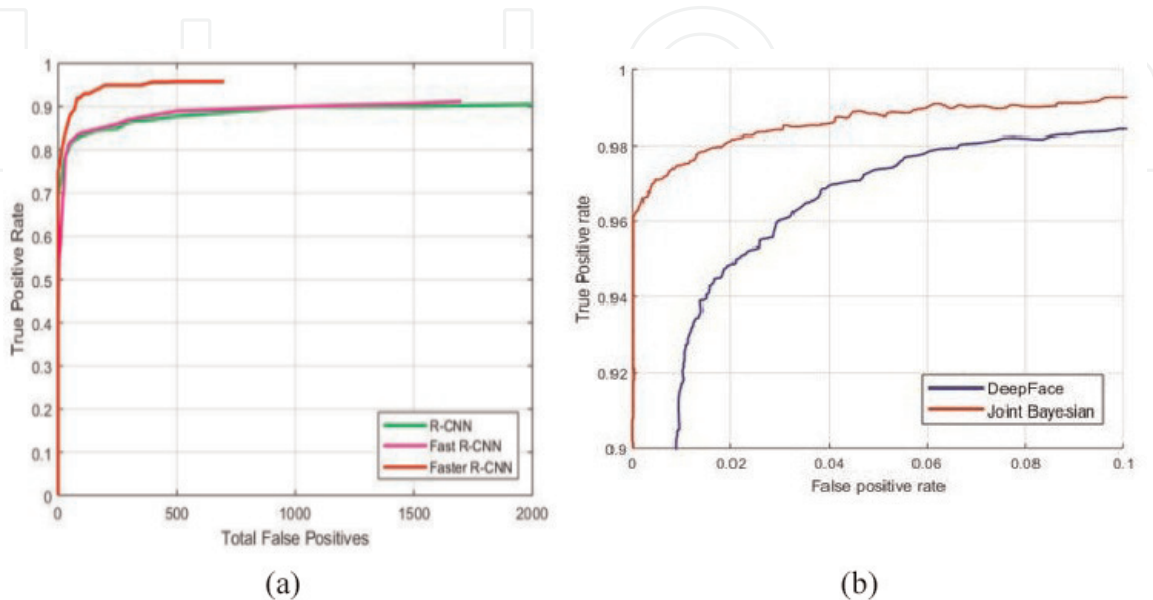


Figure 9.
(a) Comparisons of R-CNN, Fast R-CNN, and Faster R-CNN face detection methods on TCE dataset and (b) ROC comparison with the deep face method.

more quickly than RPN for fewer proposals. The plots show that using RPN yields a much faster detection system than using either SS or EB, when the number of proposals drops from 2000 to 300.

In addition the face detection performance of the R-CNN is compared with the Fast R-CNN and the Faster R-CNN on TCE dataset. As observed from **Figure 9a**, the Faster R-CNN significantly outperforms the other two. Deeply trained network

METHOD	ACCURACY (%)
Deep Face	91.4±1.9
Joint Bayesian	98.3±1.1

Table 3.
Accuracy comparison on TCE dataset.

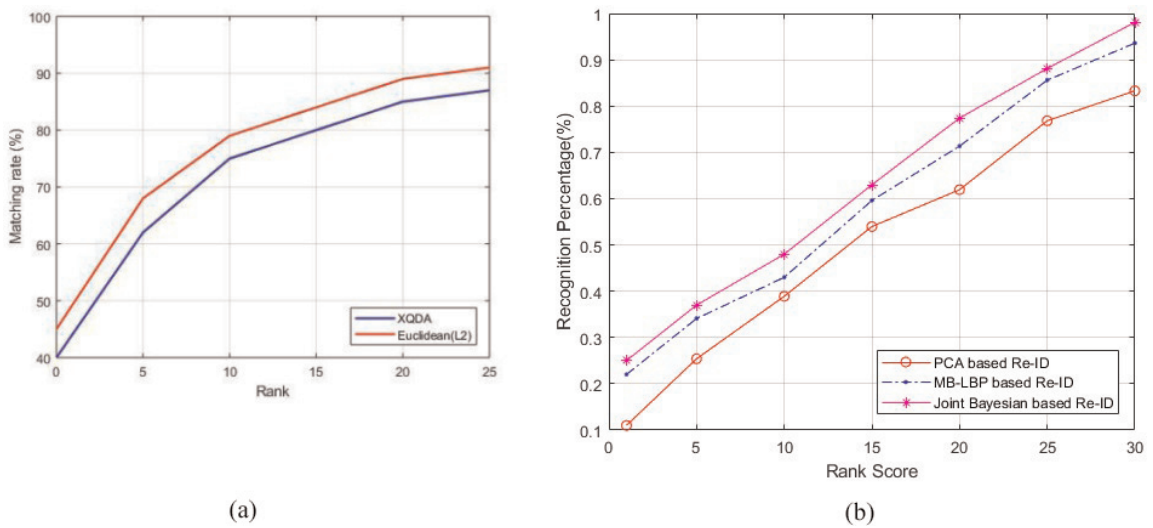


Figure 10.
(a) CMC curve for different ranking methods and (b) CMC curve for various face recognition methods.

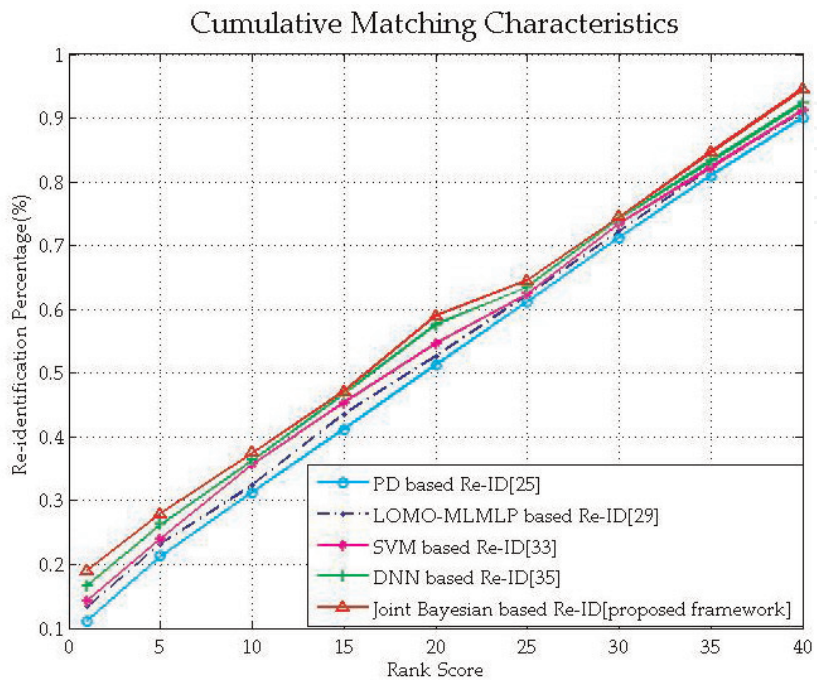


Figure 11.
CMC curve for various state-of-the-art facial feature-based Re-ID methods.



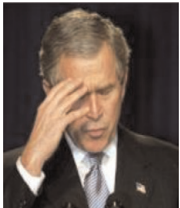







	Success cases			Failure cases	
Query					
Image with rank 1 (first image in the ranking list)					
Distance with the query image	0.5460	0.6989	0.60341	1.6792	1.710

Table 4.
Success and failure cases of the proposed frame work.

such as RPN boosts the performance of Faster R-CNN. Also, the Faster R-CNN has high computational speed than R-CNN and Fast R-CNN.

The comparison of the joint Bayesian method with the recent state-of-the-art deep face method in terms of the mean accuracy and ROC curves are presented in **Table 3** and **Figure 9b**, respectively. It can be observed that the joint Bayesian method advances the state-of-the-art deep face method, closely approaching human performance in face recognition. An accuracy of about $98.3 \pm 1.1\%$ in face recognition is achieved on TCE dataset.

The most widely used evaluation methodology for Re-ID is the cumulative matching characteristics curve, also known as CMC curve. This performance metric is adopted since Re-ID is intuitively posed as a ranking problem, where each element in the gallery is ranked, based on its comparison to the probe face. **Figure 10a** represents the comparison of rank vs. matching rate of Euclidean (L2) method with the XQDA method. It is evident from the plot that Euclidean (L2) method achieves better Re-ID matching rate than XQDA method on TCE dataset.

Recognition rate indicates probabilities of recognizing an individual, depending on how similar their measurements are to other individuals measurements in the gallery set and compared with performance of a biometric system, operating in the closed-set identification task. The probability of the equivalent match is ranked, and the value has been plotted against the size of the gallery set. **Figure 10b** represents the comparison of the recognition rate of joint Bayesian with the PCA-based eigenface approach algorithm. This shows PCA algorithm fails in some low-resolution images, wearing goggles, and different hairstyles. **Figure 11** represents the comparison of the reidentification rate of joint Bayesian method with other recent methods. **Table 4** shows the success and failure cases of the proposed framework on TCE dataset and LFW dataset.

6. Conclusion

This chapter has presented an approach to robustly detect human facial regions from image sequences collected under various challenging conditions, such as partial occlusions, low resolutions, varying face poses, illumination variations, etc., and to reidentify a person even under those conditions. The well-established Faster R-CNN method is adopted to confirm whether the detected region proposals are human faces. Although the Faster R-CNN is designed for generic object detection, it manifests the impressive face detection performance, when attempted on a suitable face detection training set. The approach is tested on challenging benchmark datasets such as the WIDER FACE dataset, the FDDB, HALLWAY, and on own TCE dataset as well. The experimental results and various performance measures depict that the facial feature-based Re-ID results achieved are competitive and exclusive approach even in the presence of partial occlusions and other challenging conditions as mentioned above.

7. Future work

Till now, the scope of the algorithm (as shown in **Table 5**) is limited for frontal and profile face verifications, handling partial occlusions in a sparse crowd. Future work focuses on person Re-ID in a high-dense crowd under severe occlusions.

SCOPE	Algorithm works for multi-view faces, illumination changes and partial occlusions (wearing mask, specs and goggles)
CONSTRAINT	Completely occluded faces throughout the video cannot be re-identified. However, if partial face is available even in few frames, it is possible for the re-id process.

Table 5.
Scope and constraint of the proposed frame work.

Acknowledgements

This work has been supported under Video Analytics and Development System (VADS) project sponsored by IISC Bangalore under DST.

Author details

Yogameena Balasubramanian^{1*}, Nagavani Chandrasekaran², Sangeetha Asokan¹ and Saravana Sri Subramanian¹

1 Department of Electronics and Communication Engineering, Thiagarajar College of Engineering, Madurai, India

2 Department of Electronics and Communication Engineering, Kamaraj College of Engineering and Technology, Madurai, India

*Address all correspondence to: ymece@tce.edu

IntechOpen

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Bedagkar-Gala A, Shah SK. A survey of approaches and trends in person re-identification. *Image and Vision Computing*. 2014;**32**:270286. DOI: 10.1016/j.imavis.2014.02.001
- [2] Bazzani L, Cristani M, Murino V. Symmetry-driven accumulation of local features for human characterization and re-identification. *Computer Vision and Image Understanding*. 2013;**117**:131-144. DOI: 10.1016/j.cviu.2012.10.008
- [3] Liangliang R, Jiwen L, Jianjiang F, Jie Z. Multi-modal uniform deep learning for RGB-D person re-identification. *Pattern Recognition*. 2017;**72**:446-457. DOI: 10.1016/j.patcog.2017.06.037
- [4] Sarattha K, Worapan KR. Human identification using mean shape analysis of face images. In: *Proceedings of the 2017 IEEE Region 10 Conference (TENCON)*; Penang: Malaysia; 5-8 November 2017. pp. 901-905
- [5] Artur G, Marcin K, Norbert P. Face re-identification in thermal infrared spectrum based on thermal facenet neural network. In: *Proceedings of 2018 22nd International Microwave and Radar Conference (MIKON)*; Warsaw Univ. of Technology: Poland; 2018. pp. 179-180
- [6] Li P, Prieto ML, Patrick JF, Mery D. Learning face similarity for re-identification from real surveillance video: A deep metric solution. In: *Proceedings of the Joint Conference on Biometrics (IJCB)*; Denver: CO, USA; 1-4 October 2017. pp. 243-252
- [7] Li P, Joel B, Patrick JF. toward facial re-identification: Experiments with data from an operational surveillance camera plant. In: *Proceedings of the 2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*; Niagara Falls: NY, USA; September 2016
- [8] De-la-Torre M, Granger E, Sabourin R, Gorodnichy DO. Individual specific management of reference data in adaptive ensembles for face reidentification. *IET Computer Vision*. 2015;**9**:732-740. DOI: 10.1049/iet-cvi.2014.0375
- [9] Zheng L, Yang Y, Hauptmann AG. Person Re-identification: Past, present and future. *Journal of Latex Class Files*. 2016;**14**:1-20. DOI: arxiv.org/abs/1610.02984
- [10] Mazzeo PL, Spagnolo P, D'Orazio T. Object tracking by non-overlapping distributed camera network. In: *Blanc-Talon J, Philips W, Popescu D, Scheunders P, editors. Advanced Concepts for Intelligent Vision Systems*. Vol. 5807. Berlin, Heidelberg: Springer; 2009. pp. 516-527. DOI: 10.1007/978-3-642-04697-1.ch48. ACIVS 2009. *Lecture Notes in Computer Science*
- [11] Masi I, Lisanti G, Bartoli F, Del Bimo A. Person Re-Identification: Theory and Best Practice. 2015. Available from: <http://www.micc.unifi.it/reid-tutorial> [Accessed: September 02, 2015]
- [12] Vezzani R, Baltieri D, Cucchiara R. People Re-identification in surveillance and forensics: A survey. *ACM Computing Surveys*. 2013;**46**:1-37. DOI: 10.1145/2543581.2543596
- [13] Brendel W, Amer M, Todorovic S. Multi object tracking as maximum weight independent set. In: *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*; Colorado Springs: CO, USA; 20-25 June 2011. pp. 1273-1280
- [14] Madrigal F, Hayet JB. Multiple view, multiple target tracking with principal axis-based data association. In: *Proceedings of the IEEE International Conference on Advanced Video and Signal-Based Surveillance*; Klagensfurt:

Austria; 30 August-2 September 2011. pp. 185-109

[15] Dantcheva A, Dugelay JL. Frontal-to-side face re-identification based on hair, skin and clothes patches. In: Proceedings of the IEEE International Conference on Advanced Video and Signal-Based Surveillance; Klagenfurt: Austria; 30 August-2 September 2011. pp. 309-313

[16] Albiol A, Albiol A, Oliver J, Mossi J. Who is who at different cameras: People re-identification using depth cameras. *IET Computer Vision*. 2012;**6**:378-387. DOI: 10.1049/iet-cvi.2011.0140

[17] Bak S, Corvee E, Bremond F, Thonnat M. Person re-identification using spatial covariance regions of human body parts. In: Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance 29 August-1 September 2010; Boston, MA: USA: IEEE; 2010. pp. 435-440

[18] Bazzani L, Cristani M, Perina A, Murino V. Multiple-shot person re-identification by chromatic and epitomic analyses. *Pattern Recognition Letters*. 2012;**33**:898-903. DOI: 10.1016/j.patrec. 2011.11.016

[19] Chen L, Chen H, Li S, Wang Y. Person Re-identification by color distribution fields. *Journal of Chinese Computer System*. 2017;**38**:1404-1408. DOI: Xwxt.sict.ac.cn/EN/Y2017/V38/I6/1404

[20] Miyazawa K, Ito K, Aoki T, Kobayashi K, Nakajima H. An effective approach for iris recognition using phase-based image matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2007;**30**: 1741-1756. DOI: 10.1109/TPAMI. 2007.70833

[21] Cheng DS, Cristani M, Stoppa M, Bazzani L, Murino V. Custom pictorial structures for re-identification. In:

Proceedings of the British Machine Vision Conference (BMVC'11); 29 August-2 September 2011. pp. 1-11

[22] Fischer M, Ekenel H, Stiefelhagen R. Person re-identification in TV series using robust face recognition and user feedback. *Multimedia Tools and Applications*. 2011;**55**:83-104. DOI: 10.1007/s11042-010-0603-2

[23] Baltieri D, Vezzani R, Cucchiara R. SARC3D: A new 3D body model for people tracking and re-identification. In: Proceedings of the IEEE International Conference on Image Analysis and Process; September 14-16; Ravenna: Italy; 2011. pp. 197-206

[24] Caroline S, Thierry B, Carl F. AneXtended center-symmetric local binary pattern for background modelling and subtraction in videos. In: Proceedings of the International Joint Conference Computer Vision, Imaging and Computer Graphics Theory and Applications; VISAPP Berlin: Germany; March 2015. pp. 1-9

[25] Milborrow S, Nicolls F. Locating facial features with an extended active shape model. In: Proceedings of European conference on computer vision. Lecture Notes in Computer Science; Springer: Berlin, Heidelberg; 2008. pp. 504-513

[26] Miguel D, Eric G, Robert S, Dmitry OG. Individual-specific management of reference data in adaptive ensembles for face re-identification. *IET Computer Vision*. 2015;**9**:732-740

[27] Xu X, Li W, Xu D. Distance metric learning using privileged information for face verification and person Re-identification. *IEEE Transactions on Neural Networks and Learning Systems* December 2015;**26**: 3150-3162. DOI: 10.1109/TNNLS. 2015.2405574

- [28] Cui Z, Li W, Xu D, Shan S, Chen X. Fusing robust face region descriptors via multiple metric learning for face recognition in the wild. In: Proceedings of IEEE Conference Computer Vision Pattern Recognition; Portland: USA; Jun 2013. pp. 3554-3561
- [29] Xie P, Xing EP. Multi-modal distance metric learning. In: Proceedings of 23rd International Joint Conference Artificial Intelligence; Beijing: China; August 2013. pp. 1806-1812
- [30] Schroff F, Kalenichenko D, Philbin JF. A unified embedding for face recognition and clustering. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015. pp. 815-823
- [31] Hu J, Lu J, Tan Y, Zhou J. Deep transfer metric learning. IEEE Transactions on Image Processing. 2016; 25:5576-5588. DOI: 10.1109/TIP.2016.2612827
- [32] Mai G, Cao K, Pong CY. On the reconstruction of face images from deep face templates. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2018;99:1-15. DOI: 10.1109/TPAMI.2018.2827389
- [33] Kobri H, Jones M. Improving face verification and person re-identification accuracy using hyper plane similarity. In: Proceedings of International Conference Computer Vision Workshops; Venice: Italy; October. 2017. pp. 1555-1563
- [34] Sanping Z, Jinjun W, Deyu M. Deep self-paced learning for person Re-identification. Pattern Recognition. 2017;76:739-751. DOI: 10.1016/j.patcog.2017.10.005
- [35] Varior RR, Haloi M, Wang G. Gated Siamese convolutional neural network architecture for human re-identification. In: Proceedings of European Conference on Computer Vision; Amsterdam: The Netherlands; 2016. pp. 791-808
- [36] Borgia A, Hua Y, Kodirov E, Robertson N. GAN-Based pose-aware regulation for video-based person re-identification. In: Proceedings of IEEE Winter Conference on Applications of Computer Vision (WACV); Waikoloa Village, HI: USA; 2019. pp. 1175-1184
- [37] Huang Z et al. Contribution-based multi-stream feature distance fusion method with k-distribution Re-ranking for person Re-identification. IEEE Access. 2019;7:35631-35644. DOI: 10.1109/ACCESS.2019.2904278
- [38] Ross G, Jeff D, Trevor D, et al. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of International Conference Computer Vision and Pattern Recognition; Columbus, OH: USA; June 2014. pp. 580-587
- [39] Ross G. Fast R-CNN. In: Proceedings of International Conference on Computer Vision; Santiago: Chile; December 2015. pp. 1440-1448
- [40] Huaizu J, Miller EL. Face detection with the faster R-CNN. In: Proceedings of International Conference on Automatic Face and Gesture Recognition; Washington, DC: USA; June 2017. pp. 650-657
- [41] Ranjan R, Sankaranarayanan S, Castillo CD, Chellappa R. An all-in-one convolutional neural network for face analysis. In: Proceedings of 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017); Washington: DC; 2017. pp. 17-24
- [42] Xiong X, De la Torre F. Supervised descent method and its applications to face alignment. In: Proceedings of 2013 IEEE Conference on Computer Vision

and Pattern Recognition; Portland: OR;
2013. pp. 532-539

[43] Chen D, Cao X, Wipf D, Wen F, Sun J. An efficient joint formulation for Bayesian face verification. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017;**39**:32-46.
DOI: 10.1109/TPAMI.2016.2533383

IntechOpen

IntechOpen