Optics and Laser Technology 110 (2019) 227–233

Contents lists available at ScienceDirect

Optics and Laser Technology

journal homepage: www.elsevier.com/locate/optlastec

Full length article

Face recognition using total loss function on face database with ID photos

Dongshun Cui [a,b], Guanghao Zhang [b], Kai Hu [c,*], Wei Han [b], Guang-Bin Huang [b]

[a] Energy Research Institute @ NTU (ERI@N), Interdisciplinary Graduate School, Nanyang Technological University, Nanyang Avenue, 639798 Singapore, Singapore
[b] School of Electrical and Electronic Engineering, Nanyang Technological University, Nanyang Avenue, 639798 Singapore, Singapore
[c] College of Information Engineering, Xiangtan University, Xiangtan 411105, China

ARTICLE INFO

Article history:
Received 3 August 2017
Received in revised form 14 October 2017
Accepted 15 October 2017
Available online 2 November 2017

Keywords:
Face recognition
Face recognition benchmark
ID photos
Total loss function
Real-life face recognition system

ABSTRACT

With the development of deep neural networks, researchers have developed lots of algorithms related to face and achieved comparable results to human-level performance on several databases. However, few feature extraction models work well in the real world when the subject which is to be recognized has limited samples, for example, only one ID photo can be obtained before the face recognition task. To our best knowledge, there is no face database which contains ID photos and pictures from the real world for a subject simultaneously. To fill this gap, we collected 100 celebrities' ID photos and their about 1000 stills or life pictures and formed a face database called **FDID**. Besides, we proposed a novel face recognition algorithm and evaluated it with this new database on the real-life videos.

1. Introduction

Tasks related to face like face recognition, as one of the hottest field of computer vision [1,2], have been active research fields for many years. Human-level performance has been achieved by machine-learning based algorithms for face recognition on several existing datasets (e.g., deep neural network [3–5] and extreme learning machines [6]).

The most famous face recognition database is released in 2007 named Labeled Faces in the Wild (LFW [7]), which is designed for unconstrained face recognition. LFW dataset contains more than 13k raw images of faces and over 5k subjects in total. Each person's face image was labeled with his/her name and the eyes' coordinates. This dataset also contains a funneled set which provides aligned face images for each raw image. Lots of work have been done based on this dataset, and amazing results have been reported in [8].

In 2012, Chen et al. [9] built a wide and deep face database called WDRef, which includes about 3000 subjects and more than 1000 subjects associated with over 40 images. However, they only provided low-level features of these images and the official link to these features is no longer available currently.

Cross-Age Celebrity Dataset (CACD) is designed for cross-age face recognition and retrial was released by collecting over 160k face images of 2k celebrities with age ranging from 16 to 62 [10]. However, not only age but also many other factors (e.g., head pose and illumination) affect the performance of a face recognition system in the real applications. So CACD is not a general enough database.

Liu et al. [11] created a Large-scale CelebFace Attributes (CelebA) Database of over 10K celebrities which contains more than 200K images. They labeled five landmarks (centers of the eyes, apex nasi, and the two corners of the mouth) and 40 attributes (smiling, wearing a hat, wavy hair, etc.) for each face image. So this dataset can be employed for many face-related tasks like facial landmarks localization, face attribute recognition, and face recognition/detection. A similar dataset is CASIA WebFace Database [12] which also contains about 10k subjects but extends the total number to 500k.

SFC is created by Facebook by collecting social face images and is reported in [3]. To the best of our knowledge, SFC database contains the most images among all the existing datasets. However, this dataset is private.

FaceScrub database has over 100 k face images of 530 people, and the annotations include the celebrities' name and gender [13]. They provide the links (not always available) of the images

* Corresponding author.
E-mail addresses: dcui002@ntu.edu.sg (D. Cui), gzhang009@ntu.edu.sg (G. Zhang), kaihu@xtu.edu.cn (K. Hu), hanwei@ntu.edu.sg (W. Han), egbhuang@ntu.edu.sg (G.-B. Huang).

https://doi.org/10.1016/j.optlastec.2017.10.016
0030-3992/

and coordinates of faces' bounding boxes (obtained by Haar cascade-based face detector which can only detect faces with limited head poses [14]). The main difference between our database and FaceScrub is we provide ID photo for the celebrities.

The intelligence Advanced Research Projects Activity released the IARPA Janus Benchmark A (IJB-A) dataset which contains more than 5k images and 2k videos of 500 subjects [15]. This database can be used both for face recognition and face detection, and it provides the eyes and nose locations.

MegaFace is the first face recognition database which contains million-level images [16]. In the training set, there are 4.7 million photos of over 600k unique identities.

We consider the ID photo of a person is critical for face recognition since the only information we can obtain is ID photo under several specific scenarios before there is a request to recognize some person from the videos or in the real world. For example, it is likely that only ID photos are available before the arrest of criminals. This requires face recognition systems to extract enough useful features from an ID photo and recognize criminals with these representations. Another example is that when offering an intelligent service to some very important persons (e.g., the head of a country) but it is inconvenient to scan them and acquire face samples. The reasons why we choose celebrities as the subjects are:

1. Large quantities. Huge data is the foundation of training a deep neural network [17–19], while the selected celebrities are very famous and large amounts of their stills and life photos (collectively called "non-ID photos") are available on the Internet.
2. Rich scenes. Compared to the ordinary people, celebrities live in much more rich scenes which leads to generating more face images with variable backgrounds and illumination conditions.
3. Diverse appearances. Obviously, the appearances are diversity (e.g., different facial expressions) when celebrities play different roles or attend various activities. In addition, multiple views and age changes increase the variety of facial appearance.
4. Big Difference. There always has a huge gap between the ID photos and stills of the same celebrity due to the techniques of facial make-up and image retouching.

For our dataset, we annotate the coordinates of the tips of noses for the face images. However, we haven't provided the locations of the eyes and the mouth corners because this information is not always available, e.g., when a head turns left at a 90-degree angle, one of the eyes will be invisible.

We have the largest average image quantity (donated as $\sharp$) for each person among all existing database. $\sharp$ is calculated by $\sharp = Q/N$, while Q is the total images of the database and N is the number of subjects. The comparison between the existing datasets and our FIFD dataset is shown in Table 1.

The main contributions of our work are:

- We have leveraged a face benchmark which consists of ID photos and non-ID photos of 100 (Chinese) celebrities for face recognition. The way to build the dataset and detailed analysis of the dataset are explained.
- We have proposed a novel architecture for general face recognition and tested it with the real-world videos to illustrate the quality of our FIFD dataset.

The rest of the paper is organized as follow: Section 2 introduces the related work on existing scientific methods on creating face databases and the state-of-the-art face recognition algorithms. Then we explain the details on how we collect our database and data diversity is shown by using the images from our database. In Section, we propose a novel face recognition architecture and show the results of our methods on the real-life videos by training a model with our database. A conclusion is given, and future work is claimed in Section 5.

2. Related work

2.1. Protocol of building face recognition database

The database is essential for training face recognition algorithms to achieve models and evaluating their performance [20,21]. And there is no doubt that algorithms benefit a lot from a comprehensive and exquisite database which requests a scientific process. We have introduced the size of the existing popular datasets in Section 1, and here the systematic procedure of creating a face recognition database is to be introduced.

The general steps of building a face database are collection, cleaning, and arrangement of face images. Experimental and real-life environment are the two sources for gathering images, and the Internet provides a convenient way to collect real world pictures. Face image cleaning consists of face detection, face alignment, and duplication elimination. All of these fields have been explored for over 20 years, but still provide no perfect solutions [22–24]. Finally, face image annotation (manually or automatically) and stored by the order are done.

2.2. Existing face recognition algorithms

Here we list some state-of-the-art face recognition algorithms. Wang et al. [25] proposed a kernel collaborative face recognition algorithm by considering the non-linear relationship of face images. This method has achieved high performance by adopting non-linear representation-based classifiers. DeepFace is the first wide accepted face recognition algorithm that approaches

Table 1
The comparison between existing datasets.

Dataset	Year [a]	N	Q	$\sharp$	ID Photo	Pose variation	Availability
LFW	2007	5749	13233	2.30	N	Limited	Public
WDRef	2012	2995	99773	33.31	N	Limited	Public (link is invalid)
CACD	2014	2000	163446	81.72	N	Limited	Public (partial annotated)
CASIA WebFace	2014	10575	494414	46.75	N	Full	Public
FaceScrub	2014	530	106863	201.63	N	Limited	Public
SFC	2014	4030	4400000	1091.81	N	Limited	Private
CelebA	2015	10177	202599	19.91	N	Limited	Public
IJB-A	2015	500	5712	11.42	N	Full	Public
Megaface	2016	672057	1027060	1.49	N	Full	Public
FIFD (Ours)	2017	100	112839	1128.39	Y	Full	Public [b]

[a] The time of publication of the corresponding papers.
[b] Please visit http://www.escience.cn/people/cuidongshunEN/dataset.html to download the proposed database.

human-level performance on LFW (97.35%) [3]. It follows the general process pipeline of face detection, face alignment, face representation and face verification. DeepFace trains an efficient nine-layer deep neural network (DNN) and tune 120 million parameters with no weight sharing between local connections. Later, DeepID, DeepID2, DeepID2+, and DeepID3 are proposed by modifying the structure of the network (for example, DeepID adopts Convolutional Network.) and increase the accuracy further [26,27,4,28].

FaceNet is proposed by using triplet loss (each pair of triplet consists of an anchor sample x^a, a positive sample x^p, and a negative sample x^n) and increase the accuracy to 99.63% . The loss function is

$$\sum_{i}^{N}\left[\|x_i^a - x_i^p\|_2^2 + \alpha - \|x_i^a - x_i^n\|_2^2\right]. \tag{1}$$

Here, α is a constant which guarantees the distance between positive and negative pairs.

3. Our database

There are massive image data on the Internet, and we can find a enormous amount of stills and life photos for most celebrities. In our task, only the celebrities of whom we can obtain a clear ID photo will be put into our dataset.

3.1. Collection rules and flowchart

The target of or work is to build a face recognition database which contains one ID photo and as more non-ID photos as possible. These non-ID photos should be diverse enough to simulate the real-world scenario. With the help of search engines (like Google, Baidu), we have obtained the ID photos of 100 Chinese celebrities so far. Folders are created and named after the celebrities, and ID photos are stored in the corresponding folders.

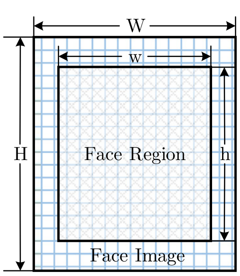

Fig. 1. Sketch of the ID photo.

According to the standards developed by International Civil Aviation Organization (ICAO) on the size of the passport photo [29], the crown-to-chin portion of the facial image shall be 70–80 percent of the height of the picture. We take 75 percent as the reference weight ratio and height ratio and crop each face image to suit these ratios approximately. Assume the width and height of the face region and the whole face image and are (w, h) and (W, H) respectively, then we have $w \approx 0.75 \times W$ and $h \approx 0.75 \times H$, shown in Fig. 1. We have checked the size of passport photos in many countries and got the conclusion that the width of a photo in the passport is about 35 mm (≈1.378 in. ~413 pixels) and the height is about 45 mm (≈1.772 in. ~531 pixels) [30].

Based on our database including ID photos and non-ID photos, we re-design the architecture of the face recognition as shown in Fig. 2.

In the **training** stage, we first apply a data preprocessing for each image in the database which consists of ID photos and non-ID photos. This module includes image denoising, face detection, face segmentation, and histogram equalization (to decrease the effect of non-uniform illumination). We don't implement facial landmarks detection because some landmarks are not visible when the angle of head pose is large. We obtain a clean face after data preprocessing and then we extract features of the face image. Feature extraction methods are usually one of the core modules of the whole face recognition systems, and lots of traditional and fashion feature extraction methods introduced in Section 2 can be adopted here. Another core module is evaluation which consists of similarity/dissimilarity computation and outputs an accuracy. With the highest accuracy, we can select the best feature extraction model and evaluation method. Simultaneously, we output the features of each face image and store them into a face-feature database with their corresponding indexes.

In the **test** stage, we apply the same data preprocessing for the test image and obtain the face regions (if there exists at least a face). We use the best feature extraction model to learn the representations of each detected face and compute the similarities/dissimilarities between these representations and features from the

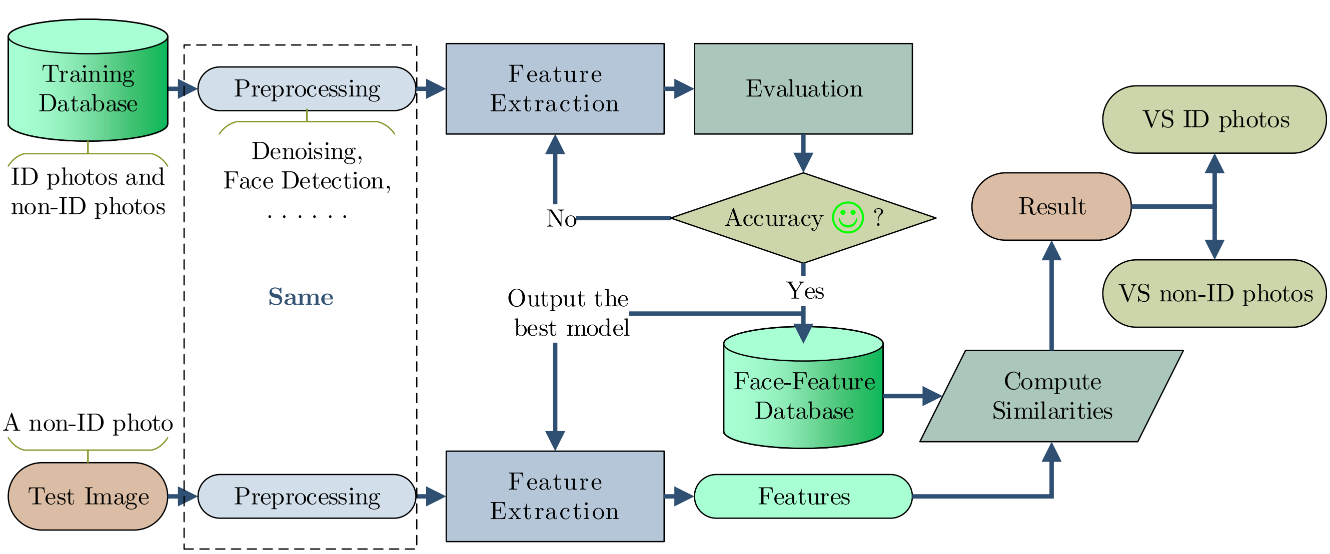

Fig. 2. The proposed architecture of the face recognition.

face-feature database. A final result is given for the test image after comparing with the existing ID photos and non-ID photos.

3.2. Resources of the ID and non-ID photos

There are lots of pictures on the Internet for the celebrities we can collect, and the first step is to find celebrities' ID photos. So far, we have gathered 100 celebrities' ID photos and stored these ID photos separately. With these celebrities' name and the full consideration of data diversities, we continue to search their stills and life photos mainly from the popular movie databases (e.g., IMDb, Mtime), famous image search engines (e.g., Google Images and Baidu Image), and some social networks (e.g., Baidu Tieba, Sina Weibo).

3.3. Remove the duplicates

An essential thing when we collect the non-ID photos is removing the duplicates since they haven't offered any new information. Considering a significant amount of our database, we use a software named Duplicate Cleaner (free version) to detect the duplicates loosely. Then we check each group of the candidate duplicates one by one and remain only one of the duplicates when any of the following conditions are met: they have the exact same contents; they have the same contents except one has some watermarks/text; one can be accurately obtained by flipping/rotating/rescaling the other.

3.4. Face segmentation

We adopt the face detection method proposed by Zhang et al. [31] using multitask cascaded convolutional networks. The reasons why we select this face detector are: it outperforms all the compared approaches in benchmarks of FDDB and WIDER FACE by a large margin; its speed is 99 FPS on GPU which is faster than many other detectors.

A final manual check is done for the outputs of the automatic face detector according to the original images one by one, and the false positive results (not face regions) and the outliers (other persons' face regions) are removed. In addition, we add the false-negative results by cropping the face regions from the raw images manually. Refer to the collection Rules in Section 3.1, we crop and re-size all the right face regions from the original images into 531×413. Each final face image is put into the corresponding folder after the celebrity's ID-photo and named sequentially.

3.5. Insight of the proposed database

We have created the face database (FDID) with the ID photos of the celebrities from the internet and as many of their corresponding non-ID photos as possible. We have randomly selected 36 celebrities (18 males and 18 females) and shown their ID photos in Fig. 3.

The reason why we pay more attention to the ID photo is that it usually provides one person's most information among all his face images. Here we have listed some properties of an ID photo in the passport [32]:

1. No head pose.
2. Neutral expression.
3. No occlusions (e.g., glasses and marks) on the face.
4. High contrast and light/gray background.
5. Clear and unaltered by image processing software.

Apparently, these rules echo the main challenges of real-world face recognition applications since non-ID photos always don't fulfill one or several of these rules. To make an intuitive comparison, we randomly select four celebrities (two males and two females) to show the diversity contents of our database.

Examples of face images with large-angle head pose in our database are shown in Fig. 4. Normally, face alignment using facial landmarks is included in the procedure of building face databases (e.g., IJB-A, CASIA). However, in most practical cases, not all the facial landmarks are visible when there exists a large-angle head pose. Based on such observation and to perverse face images at the aboriginality, we do not align the face images in our database.

We know that the same people have different facial expressions will look different. There are six basic facial expressions (anger,

Fig. 3. Some of the ID photos in our collected database. Note that these are Chinese ID-card photos, and people can have slight smiles when they take photos.

Fig. 4. Examples of face images with head pose in our FIFD database. Photos in the same column from top to down show the same actor's face images with head poses of **yaw**, **pitch**, and **roll** respectively.

happiness, fear, surprise, disgust and sadness) in our daily lives [33,34]. The most usual three among them are surprise (characterized by eyebrow pull up, mouth hangs open, etc.), sadness (characterized by eyelids loose, etc.) and happiness (characterized by mouth corners raised, cheeks raised, etc.) as shown in Fig. 5.

Occlusions (e.g., sunglasses, masks, and the hair which are shown in Fig. 6) render a low-rank error image in contrast to the full-rank original image [35]. That is, they remove some necessary information that can be used for face recognition. Even though some researchers have proposed several algorithms to generate the missing part by using Variational Autoencoder (VAE) and Generative Adversarial Networks (GAN) recently [36–38]. However, they cannot guarantee that the generated occlusion-free face image still belongs to the same person.

Fig. 6. Examples of partially occluded face images in our FIFD database. Photos in the same column are the same actor. From left to right, the occlusions in the first row are glasses while in the second row are a phone, a mask, the hair, and a gun respectively.

Uneven illumination has a similar effect on the face region to occlusions that weaken the useful information for face recognition. Too strong or too weak lighting conditions will compress the surface texture characteristics. Examples of face images under natural, dark and bright illumination are shown in Fig. 7.

Celebrities always need toilette when playing the roles or appearing on formal occasions, and they look slight or very different with their ID photos, shown in Fig. 8. Besides, image/video processing software will also increase this kind of difference.

In addition, age change is also an important factor that affects a person's face appearance, and one person's face images across age can be very different [39]. Examples of cross-age face images are shown in Fig. 9.

4. Proposed method and results

In this section, we will introduce a novel and general face recognition loss function for our database. We point two kinds of intra-person variances based on the assumptions of the same person's non-ID photo should be similar to his/her ID photo and the same person should look similar to his/her other non-ID photo.

Fig. 5. Examples of expressional face images in our FIFD database. Photos in the same column from top to down show the same actor's surprise, unhappy and happy expressions respectively.

Fig. 7. Examples of face images under different illumination conditions in our FIFD database. Photos in the same column from top to down show the same actor's face images under natural, dark and bright illuminations respectively.

Fig. 8. Examples of make-up face images in our FIFD database. Photos in the same column are the same actor. Images in different rows show varying degrees of the toilette.

Fig. 9. Examples of cross-age face images in our FIFD database. Photos in the same column from top to down present the same actor's change from young to old.

4.1. Total loss function

We use x_{ij} indicates the j-th photo of the i-th person in the dataset. To distinguish the ID photo and non-ID photos, we assume $j = 0$ denotes the former while $j \in [1, m]$ indexes the latter. Assume $r(x)$ as the representation of sample x, it can be manually descriptors (e.g., LBP, Gabor, and Eigenvector) or learned by a neural network (e.g., AlexNet, PCANet, and CNN). Dissimilarity function is expressed as $d(\cdot)$. These two kinds of intra-person variances yield the following two loss functions.

The first intra-person loss function is:

$$\mathcal{L}_1 = \sum_{i\in[1,n]} \sum_{j\in[1,m]} d_1(r(x_{i0}), r(x_{ij})). \tag{2}$$

The second intra-person loss function is:

$$\mathcal{L}_2 = \sum_{i\in[1,n]} \sum_{j_1,j_2\in[1,m]} d_2(r(x_{ij_1}), r(x_{ij_2})). \tag{3}$$

Intra-person loss functions constraint the similarity between the photos of the same person, and for face recognition, we also need to constraint the dissimilarity between the photos of the different persons.

The inter-person loss function is:

$$\mathcal{L}_3 = \sum_{i_1,i_2\in[1,n]} \sum_{j_1,j_2\in[1,m]} d_3(r(x_{i_1j_1}), r(x_{i_2j_2})). \tag{4}$$

So, the total loss function $\mathcal{L}_{total}$ for the set of face images is:

$$\mathcal{L}_{total} = \lambda_1\mathcal{L}_1 + \lambda_2\mathcal{L}_2 - \lambda_3\mathcal{L}_3, \tag{5}$$

where $\lambda_1, \lambda_2, \lambda_3$ are the weights of $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3$ respectively. Our target is to minimize this loss function by learning an effective model to represent the face images, which is expressed as:

$$\mathcal{T}(r, \lambda_i) = min(\mathcal{L}_{total}). \tag{6}$$

Currently, most face recognition models are the instances of Eq. (6) by setting $\lambda_1 = \lambda_2$.

4.2. Results and discussion

We train a model with our dataset by adopting the similar network structure of DeepFace ([3], one of the state-of-the-art algorithms on the most popular face database LFW). We replace the softmax loss function with our proposed total loss function and test it with our database and some real-life videos. ID and non-ID photos of 80 of our collected celebrities are randomly selected for training and ten celebrities' photos for verifying, and the best feature model (mainly tuning the parameters of the networks) and dissimilarity computation algorithm are achieved. The remaining 10 celebrities' non-ID photos are used to test, and the testing accuracy is 79.64%, which is about five percent higher than the accuracy of DeepFace on FDID.

Two conclusions could be drawn: first, the proposed loss function is better than softmax function and outperforms DeepFace.

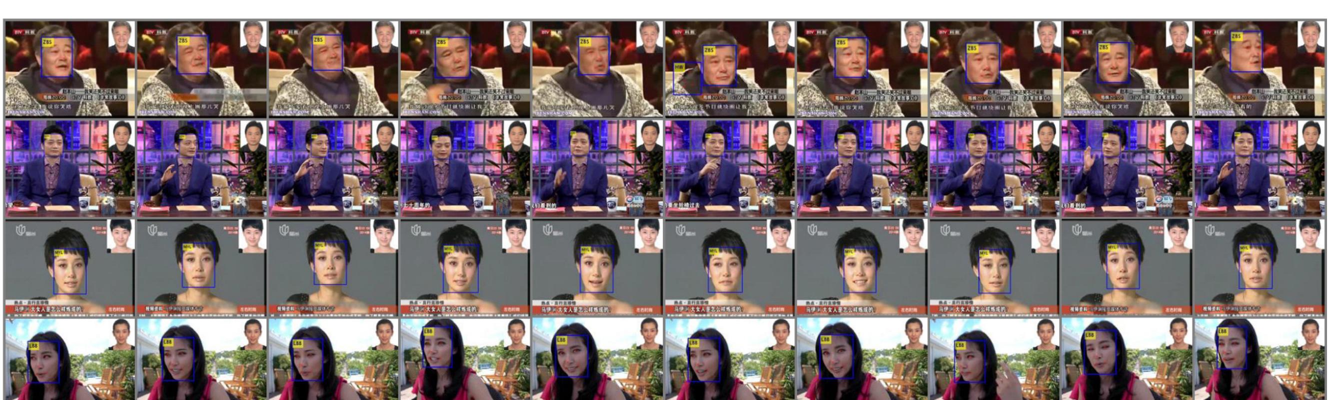

Fig. 10. Results of the proposed method on real-life videos (frame sequences) by training on our database. Each row shows the face recognition result of the same celebrity.

Second, our database is suitable for practical application than LFW database since many face recognition algorithms (including DeepFace) have achieved nearly 100% percent on it. Noted that the results mentioned above maybe be improved further with optimization.

Besides, we collect these 10 celebrities' videos from YouTube and transfer them into frames. Then we implement the same procedures of preprocessing (e.g., face detection), and extract features with the optimized model. Finally, we compute the dissimilarity between features of the input face images and all the candidate ID photos and output the celebrity's name which corresponded to the minimum dissimilarity value. After having an ID photo. we extract features from it, and then for the input videos, we do the face detection, feature extraction, and dissimilarity computation. Finally,we output the face identification results which are demonstrated in Fig. 10. We haven't computed the precise accuracy of the algorithm on the videos since face images in videos have strong correlations. But from an intuitive point of view, the accuracy rate is higher than 85% for the test videos.

5. Conclusion

We have created a new database which contains ID photos of 100 celebrities with their over 1000 stills and life photos. Details on how to build collect this database are introduced, and the data diversity on faces with different head poses, facial expressions, occlusions, illuminations, degrees of toilette, and ages are shown. Besides, we proposed a novel architecture on face recognition when the ID photos are available, and new total loss function which contains two intra-person loss functions and one inter-person loss function are presented. Models are trained with the proposed database using our target function, and video-level experiments are performed to demonstrate the meaning of our database and the effectiveness of the proposed method.

We will publish the database with annotations and make it free to download so that face-related tasks could benefit from it. It is noted that this dataset just keeps on growing. In the future, we will extend the number of the subject, and celebrities from other countries will be included. Besides, video sequences will be provided for each subject in the dataset.

Acknowledgements

The authors would like to thank the support from collaborators for this research.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.optlastec.2017.10.016.

References

[1] H. Lu, Y. Li, M. Chen, H. Kim, S. Serikawa, Brain intelligence: go beyond artificial intelligence, Mobile Networks Appl. (2017) 1–8.
[2] H. Lu, Y. Li, S. Mu, D. Wang, H. Kim, S. Serikawa, Motor anomaly detection for unmanned aerial vehicles using reinforcement learning, IEEE Internet Things J. (99) (2017) 1–8.
[3] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, Deepface: closing the gap to human-level performance in face verification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1701–1708.
[4] Y. Sun, X. Wang, X. Tang, Deeply learned face representations are sparse, selective, and robust, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 2892–2900.
[5] F. Schroff, D. Kalenichenko, J. Philbin, Facenet: a unified embedding for face recognition and clustering, Proc. IEEE Conf. Comput. Vision Pattern Recogn. (2015) 815–823.
[6] J. Tang, C. Deng, G.-B. Huang, Extreme learning machine for multilayer perceptron, IEEE Trans. Neural Networks Learn. Syst. 27 (4) (2016) 809–821.
[7] G.B. Huang, M. Ramesh, T. Berg, E. Learned-Miller, Labeled faces in the wild: a database for studying face recognition in unconstrained environments, Technical Report 07-49, University of Massachusetts, Amherst 1 (2) (2007) 1–11.
[8] E. Learned-Miller, G.B. Huang, A. RoyChowdhury, H. Li, G. Hua, Labeled faces in the wild: a survey, Adv. Face Detect. Facial Image Anal. (2016) 189–248.
[9] D. Chen, X. Cao, L. Wang, F. Wen, J. Sun, Bayesian face revisited: a joint formulation, Eur. Conf. Comput. Vision (2012) 566–579.
[10] B.-C. Chen, C.-S. Chen, W.H. Hsu, Cross-age reference coding for age-invariant face recognition and retrieval, Eur. Conf. Comput. Vision (2014) 768–783.
[11] Z. Liu, P. Luo, X. Wang, X. Tang, Deep learning face attributes in the wild, in: 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 3730–3738.
[12] D. Yi, Z. Lei, S. Liao, S.Z. Li, Learning face representation from scratch, CoRR (2014) 1–9. abs/1411.7923 .
[13] H.-W. Ng, S. Winkler, A data-driven approach to cleaning large face datasets, in: 2014 IEEE International Conference on Image Processing (ICIP), 2014, pp. 343–347.
[14] R. Ranjan, V.M. Patel, R. Chellappa, Hyperface: a deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition, CoRR (2016) 1–13 (abs/1603.01249).
[15] B.F. Klare, B. Klein, E. Taborsky, A. Blanton, J. Cheney, K. Allen, P. Grother, A. Mah, A.K. Jain, Pushing the frontiers of unconstrained face detection and recognition: IARPA Janus benchmark A, Proc. IEEE Conf. Comput. Vision Pattern Recogn. (2015) 1931–1939.
[16] I. Kemelmacher-Shlizerman, S.M. Seitz, D. Miller, E. Brossard, The megaface benchmark: 1 million faces for recognition at scale, Proc. IEEE Conf. Comput. Vision Pattern Recogn. (2016) 4873–4882.
[17] Y. Li, H. Lu, J. Li, X. Li, Y. Li, S. Serikawa, Underwater image de-scattering and classification by deep neural network, Comput. Electr. Eng. 54 (2016) 68–77.
[18] H. Lu, Y. Li, S. Nakashima, H. Kim, S. Serikawa, Underwater image super-resolution by descattering and fusion, IEEE Access 5 (2017) 670–679.
[19] H. Lu, Y. Li, T. Uemura, Z. Ge, X. Xu, L. He, S. Serikawa, H. Kim, FDCNet: filtering deep convolutional network for marine organism classification, Multimedia Tools Appl. (2017) 1–14.
[20] K. Hu, X. Gao, F. Li, Detection of suspicious lesions by adaptive thresholding based on multiresolution analysis in mammograms, IEEE Trans. Instrum. Meas. 60 (2) (2011) 462–472.
[21] K. Hu, W. Yang, X. Gao, Microcalcification diagnosis in digital mammography using extreme learning machine based on hidden Markov tree model of dual-tree complex wavelet transform, Exp. Syst. Appl. 86 (2017) 135–144.
[22] M. Kawulok, M.E. Celebi, B. Smolka, Advances in Face Detection and Facial Image Analysis, Springer, 2016.
[23] Q. Liu, J. Deng, D. Tao, Dual sparse constrained cascade regression for robust face alignment, IEEE Trans. Image Process. 25 (2) (2016) 700–712.
[24] J. Tang, Z. Li, M. Wang, R. Zhao, Neighborhood discriminant hashing for large-scale image retrieval, IEEE Trans. Image Process. 24 (9) (2015) 2827–2840.
[25] D. Wang, H. Lu, M.-H. Yang, Kernel collaborative face recognition, Pattern Recogn. 48 (10) (2015) 3025–3037.
[26] Y. Sun, X. Wang, X. Tang, Deep learning face representation from predicting 10,000 classes, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1891–1898.
[27] Y. Sun, Y. Chen, X. Wang, X. Tang, Deep learning face representation by joint identification-verification, Adv. Neural Inform. Process. Syst. (2014) 1988–1996.
[28] Y. Sun, D. Liang, X. Wang, X. Tang, Deepid3: face recognition with very deep neural networks, CoRR (2015) 1–5. abs/1502.00873 .
[29] I.C.A. Organization, Doc 9303: Machine Readable Travel Documents, Part 3: Specifications Common to all MRTDs, 2015.
[30] Passport photo size and format, <https://www.persofoto.com/lexicon/passport-photo/size/> (accessed: 2016-09-14).
[31] K. Zhang, Z. Zhang, Z. Li, Y. Qiao, Joint face detection and alignment using multitask cascaded convolutional networks, IEEE Signal Process. Lett. 23 (10) (2016) 1499–1503.
[32] Rules for passport photos. <https://www.gov.uk/photos-for-passpor-ts/photo-requirements> (accessed: 2017-02-27).
[33] M. Batty, M.J. Taylor, Early processing of the six basic facial emotional expressions, Cogn. Brain Res. 17 (3) (2003) 613–620.
[34] D. Cui, G.-B. Huang, T. Liu, Smile detection using pair-wise distance vector and extreme learning machine, in: 2016 International Joint Conference on Neural Networks (IJCNN), 2016, pp. 2298–2305.
[35] J. Yang, L. Luo, J. Qian, Y. Tai, F. Zhang, Y. Xu, Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes, IEEE Trans. Pattern Anal. Mach. Intell. 39 (1) (2017) 156–171.
[36] A.B.L. Larsen, S.K. Sønderby, H. Larochelle, O. Winther, Autoencoding beyond pixels using a learned similarity metric, CoRR (2015) 1–8. abs/1512.09300 .
[37] A. Dosovitskiy, T. Brox, Generating images with perceptual similarity metrics based on deep networks, Adv. Neural Inform. Process. Syst. (2016) 658–666.
[38] X. Wang, A. Gupta, Generative image modeling using style and structure adversarial networks, Eur. Conf. Comput. Vis. (2016) 318–335.
[39] S. Liao, Y. Hu, X. Zhu, S.Z. Li, Person re-identification by local maximal occurrence representation and metric learning, Proc. IEEE Conf. Comput. Vis. Pattern Recogn. (2015) 2197–2206.