Deep Learning Based Person Identification Using Facial Images

Hamidur Rahman(✉), Mobyen Uddin Ahmed, and Shahina Begum

School of Innovation, Design and Engineering, Mälardalen University,
72123 Västerås, Sweden
{hamidur.rahman, mobyenUddin.ahmed,
shahina.begum}@mdh.se

Abstract. Person identification is an important task for many applications for example in security. A person can be identified using finger print, vocal sound, facial image or even by DNA test. However, Person identification using facial images is one of the most popular technique which is non-contact and easy to implement and a research hotspot in the field of pattern recognition and machine vision. In this paper, a deep learning based Person identification system is proposed using facial images which shows higher accuracy than another traditional machine learning, i.e. Support Vector Machine.

Keywords: Face recognition · Person identification · Deep learning

1 Introduction

Person identification using facial images is one of the most popular biometric topics in computer vision community because of its broad applications such as human-computer interaction (HCI), homeland security, entertainment etc. Identifying a person using their facial images includes image processing and a classification based on machine learning can be defined as face recognition. However, this still remains a challenging problem because of intra-subject variations due to head pose, illumination, facial expression, occlusion due to the person themselves, other objects or accessories, facial hair, aging and so on [1].

A facial recognition technique is proposed by Chennamma et al. using SIFT features [2] where frontal facial images are used for training and manipulated facial images are used for testing. Even though the method shows higher accuracy than popular eigen face approach, the method is slower due to searching large number of matching points. Another novel face recognition algorithm method based on adaptive 3-D Local Binary Pattern and Singular Value Decomposition method is proposed in [3]. The experimental simulation results show good feature extraction effect and face recognition performance comparing other state-of-the-art methodologies. Another facial recognition method is proposed using generalized mean deep learning neural network [4]. The algorithm provides fast convergence of the feature set and the performance of the proposed algorithm is better in terms of identification accuracy. Deep learning models achieve higher accuracy in the field of object detection, text classification, image classification,

M. U. Ahmed et al. (Eds.): HealthyIoT 2017, LNICST 225, pp. 115–119, 2018.
https://doi.org/10.1007/978-3-319-76213-5_17

facial recognition, gender classification, scene-classification, digits and traffic signs recognition etc. where it uses large architectures with numerous features [5].

In this paper, a person identification system using facial images i.e. facial recognition using deep learning is proposed. A deep learning model called Alexnet [6] is used for facial recognition which has 23 layers including five convolution layers, max-pooling layers, dropout layers and three fully connected layers. Alexnet is a specific type of deep learning algorithm that can both perform classification and extract features from raw images.

2 Materials and Methods

2.1 Data Collection

For this experiment both a public online image database and local image database (i.e. created internally by us) are used. Online public database AT&T[1] is a widely used face database for face recognition problem. The database has 400 gray scale images of 40 distinct subjects each for 10 different images. For some subjects, the images are captured at different situations such as times, lighting, facial expressions etc. All the images are captured against a dark homogeneous background with the subjects in an upright, frontal position. Images are saved in PGM format and the size of each image is 92 × 112 pixels, with 256 grey levels per pixel. Another image database called 'IDT' containing 50 color images of five subject is created locally by us. Ten different color images of different face angle are captured for each subject and saved all the images in JPG format in the 'IDT' database. In this database, each image size is 640 × 480 though later image size is changed into 227 × 227 pixels. During this data collection, environmental illumination is not constant and background is also not homogeneous.

2.2 Methods

In this paper work, facial recognition is performed using both traditional machine learning algorithm i.e. Support Vector Machine (SVM) and deep learning algorithm which are later compared for evaluation. Using both the databases 70% of the images are randomly selected for training and 30% of the images are selected for testing. In machine learning approach, i.e. SVM, HOG (Histogram of Oriented Gradients) features are extracted using MATLAB function 'extractHOGFeatures' [7] both for training and test images. For example, SVM is used in other domain such as stress classification [8] or physical activity classification [9]. To build a model, supervised machine learning model i.e. multiclass Support Vector Machine (SVM) [10] is used which is actually K*(K−1)/2 binary SVM model where K is the number of unique class labels. In the model one-versus-one coding design are considered and linear kernel function (i.e. dot product) is used. To identify an input image using the learned model, first face detection is performed using Viola-and-Jones algorithm [11] and HOG features are

[1] "AT&T Face Database: http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html."

extracted from the registered facial image. SVM classifier is used for matching database feature vector with query feature vector. It finds best matching faces from the database and gives ID of best matching face image as a recognition output. An overview of the model is presented in Fig. 1.

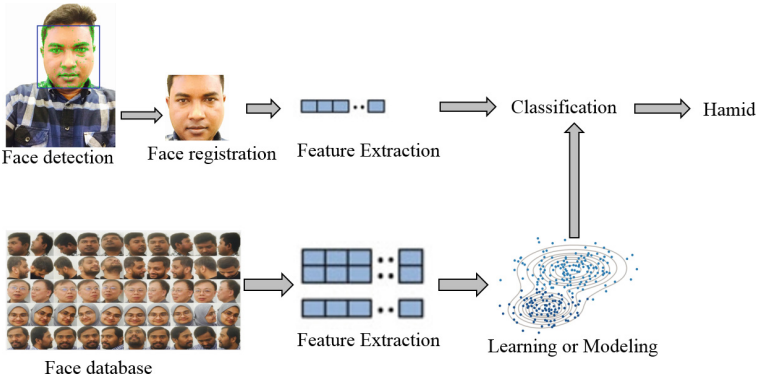

Fig. 1. Face recognition overview

Deep learning algorithms can learn features, representations, and tasks directly from images which eliminate the hassle of manual feature selection. In this study, a deep learning model called Alexnet [6] is used. Alexnet has 25 layers including one input layers, five convolution layers, seven Relu (Rectified Linear Unit) layers, three maximum pooling layers, two cross channel normalization layers, two dropout layers, three fully connected layers, one softmax layer and one output layers. Since Alexnet network was trained on 227 × 227-pixel images, all of our images are resized into the same resolution. The first layer of the Alexnet is input layer where the resolution of the images is assigned which are 227 × 227 × 3 and 227 × 227 for IDT and ATT database respectively. The last layer is output layer which uses probabilities returned by softmax activation function for each input to assign it to one of the mutually exclusive classes. An overview of deep learning based facial recognition is presented in Fig. 2.

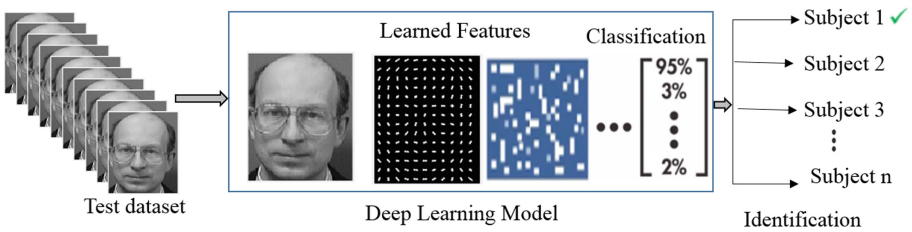

Fig. 2. Deep learning face recognition

3 Results and Evaluation

Both SVM and deep learning (DL) approaches are run 10 times for both the database. Every time a random 70% of the images are used for training and the rest of the images are used for testing. Training and test accuracy are calculated and then statistical parameters such as maximum (MAX), minimum (MIN), average (AVG) and standard deviation (STD) are calculated for 10 runs. The summary of the result of face recognition test is presented in Table 1. For the training, MAX, MIN and AVG accuracy of SVM are 65%, 40% and 54.4% for IDT database and 91%, 86% and 88.8% for ATT database whereas MAX, MIN and AVG accuracy for deep learning are 100%, 94% and 96.2% for IDT database and 100%, 95%, 98% for ATT database. STD of deep learning for IDT and ATT database are 1.8 and 1.6 and STD for SVM for IDT and ATT database are 7.6 and 1.4.

For the test, MAX, MIN and AVG accuracy of SVM are 80%, 40% and 58% for IDT database and 98%, 93% and 95.1% for ATT database whereas MAX, MIN and AVG accuracy for deep learning are 100%, 80% and 88.3% for IDT database and 100%, 97%, 98.7% for ATT database. STD of deep learning for IDT and ATT database are 6 and 0.9 and STD for SVM for IDT and ATT database are 14 and 1.5.

It is observed that traditional machine learning algorithm, i.e. SVM does not work well when images are not captured in controlled environment (i.e. homogeneous background constant illumination, equal distance between camera and face etc.) while DL algorithm works better for any kind of images because of its rich feature extraction capability. STD of DL and SVM are 0.9 and 1.5 for ATT database and 6 and 14 for IDT database which indicates that the fluctuation of accuracy for each run for ATT database is less than IDT database due to higher number of data sets.

Table 1. Statistical measurement of face recognition accuracy

Training/Testing	Statistical parameters	DL accuracy		SVM accuracy	
		IDT database	ATT database	IDT database	ATT database
Training	MAX	100	100	65	91
	MIN	94	95	40	86
	AVG	**96.2**	**98**	**54.4**	**88.8**
	STD	1.83	1.6	7.6	1.4
Testing	MAX	100	100	80	98
	MIN	80	97	40	93
	AVG	**88.3**	**98.7**	**58**	**95.1**
	STD	6.0	0.9	14	1.5

4 Conclusion

Person identification using facial images i.e. facial recognition faces challenges due to several complex situations such as intra-subject variations due to head pose, illumination, facial expression, occlusion due to other objects or accessories, facial hair,

aging and so on. Both the traditional machine learning algorithm i.e. SVM and deep learning algorithm are used for facial recognition and results show that deep learning based approach provide higher accuracy level than traditional machine learning approach. Results can be improved using larger dataset for better training model in internal database. A real time facial recognition using deep learning is considered as future work.

Acknowledgement. The authors would like to acknowledge the Swedish Knowledge Foundation (KKS), Hök instrument AB, Volvo Car Corporation (VCC), The Swedish National Road and Transport Research Institute (VTI), Autoliv AB, Prevas AB Sweden, and all the test subjects for their support of the research projects in this area.

References

1. Olszewska, J.I.: Automated face recognition: challenges and solutions. In: Ramakrishnan, S. (ed.) Pattern Recognition - Analysis and Applications, Chap. 4. InTech, Rijeka (2016)
2. Chennamma, H.R., Rangarajan, L., Veerabhadrappa: Face identification from manipulated facial images using SIFT. In: 2010 3rd International Conference on Emerging Trends in Engineering and Technology, pp. 192–195 (2010)
3. Li, Y.: Novel face recognition algorithm based on adaptive 3D local binary pattern features and improved Singular Value Decomposition method. In: 2016 International Conference on Inventive Computation Technologies (ICICT), pp. 1–7 (2016)
4. Sharma, P., Yadav, R.N., Arya, K.V.: Face recognition from video using generalized mean deep learning neural network. In: 2016 4th International Symposium on Computational and Business Intelligence (ISCBI), pp. 195–199 (2016)
5. Islam, S.M.S., Rahman, S., Rahman, M.M., Dey, E.K., Shoyaib, M.: Application of deep learning to computer vision: a comprehensive study. In: 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV), pp. 592–597 (2016)
6. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems (2012)
7. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), vol. 1, pp. 886–893 (2005)
8. Barua, S., Begum, S., Ahmed, M.U.: Supervised machine learning algorithms to diagnose stress for vehicle drivers based on physiological sensor signals. In: 12th International Conference on Wearable Micro and Nano Technologies for Personalized Health (2015)
9. Ahmed, M.U., Loutfi, A.: Physical activity identification using supervised machine learning and based on pulse rate. Int. J. Adv. Comput. Sci. Appl. (IJACSA) **4**(7), 209 (2013)
10. Rajesh, K.M., Naveenkumar, M.: A robust method for face recognition and face emotion detection system using support vector machines. In: 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEEC-COT), pp. 1–5 (2016)
11. Viola, P., Jones, M.: Robust real-time face detection. In: Proceedings, Eighth IEEE International Conference on Computer Vision, ICCV 2001, p. 747 (2001)