Lecture Notes in Electrical Engineering 621

Kuinam J. Kim
Hye-Young Kim *Editors*

Information Science and Applications

ICISA 2019

Springer

Lecture Notes in Electrical Engineering

Volume 621

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Associate Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Executive Editor (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, SCOPUS, MetaPress, Web of Science and Springerlink ****

More information about this series at http://www.springer.com/series/7818

Kuinam J. Kim · Hye-Young Kim
Editors

Information Science and Applications

ICISA 2019

Springer

Editors
Kuinam J. Kim
Kyonggi University
Suwon-si, Korea (Republic of)

Hye-Young Kim
School of Games
Hongik University
Chungchengnam-do, Korea (Republic of)

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-15-1464-7 ISBN 978-981-15-1465-4 (eBook)
https://doi.org/10.1007/978-981-15-1465-4

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Development of an Adaptive Trait-Aging Invariant Face Recognition System Using Convolutional Neural Networks

Kennedy Okokpujie, Samuel John, Charles Ndujiuba and Etinosa Noma-Osaghae

Abstract Trait-aging causes large intra-subject variations and negatively impacts on the accuracy of face recognition systems. With lapse in time, it returns a false match when identifying and authenticating the same individual. In this paper, an augmented database using pre-processing measures, a feature extracting algorithm, a modified classifier and an adaptive trait-aging invariant face recognition system were developed. The database was populated using data augmentation processes on the subjects got from the Face and Gesture Recognition Network Aging Database (FG-NET AD). The data augmentation process increased the amount of available data, thus making it suitable for deep learning operations. Cross-validation operations were also performed on the augmented dataset and each face image resized to fit into the adopted Convolutional Neutral Network Model (CNN). The CNN model was re-trained with FG-NET AD after it augmented, cross-validated and divided into mini-batches. The re-training was achieved using transfer learning technique. The resulting adaptive face recognition algorithm was validated using the data reserved for testing to ascertain its performance. The performance of the proposed adaptive face recognition system was evaluated using cumulative score curve and Mean Square Error (MSE). It had a 99.90% testing accuracy, a testing loss of 0.003%, Mean Absolute Error (MAE) of 0.0637, and Mean Squared Error (MSE) of 0.0158. Thus, out-performing the best result recorded in the available literature using similar database.

K. Okokpujie (✉) · E. Noma-Osaghae
Covenant University, Ota, Ogun State, Nigeria
e-mail: kennedy.okokpujie@covenantuniversity.edu.ng

E. Noma-Osaghae
e-mail: etinosa.noma-osaghae@covenantuniversity.edu.ng

S. John
Nigerian Defence Academy, Kaduna, Nigeria
e-mail: samuel.john@nda.edu.ng

C. Ndujiuba
Air Force Institute of Technology, Kaduna, Kaduna State, Nigeria
e-mail: charles.ndujiuba@afit.edu.ng

K. J. Kim and H.-Y. Kim (eds.), *Information Science and Applications*,
Lecture Notes in Electrical Engineering 621,
https://doi.org/10.1007/978-981-15-1465-4_41

Keywords Biometric · Transfer learning · Regularization · FG-NET AD · Face recognition · Convolutional neutral network · Inception-ResNet-v2 · Data augmentation

1 Introduction

The need for automated human face recognition cannot be over emphasized. It is required for identification and authentication in various real life applications [1–6]. Biometrics has various applications, examples of these applications include, border control, voting systems, health care, attendance capturing and access control. The variant nature of the face with the passage of time has been found from rigorous research to be responsible for the intra-class variations that make facial recognition systems to return a non-match for genuine users. This factor is called "Trait-Aging" and it makes matching of "query face templates" with stored templates of users' faces in databases unreliable and insecure.

Presently, automated facial recognition systems do not have the capacity to adjust itself to changes in the human facial structure overtime due to trait aging. The facial characteristics of individuals change as they grow older and aging causes a reduction in the elasticity of the skin. The face is usually mapped with representations that are quite permanent but not completely immune to the effect of trait aging. The volume of facial biometric data stored in databases across the world is increasing at an exponential rate and the counter-productive influence of trait aging on the reliability of "static" facial recognition systems is a just cause for alarm.

A lot of postulations, models and proposals have been made to put the challenge of trait aging in facial recognition systems in check with very little success rate. The trait-age invariant phenomenon has been considered and studied in human perception, biology and biometrics. The trait-age variant characteristic of the face is considered to be multifaceted. Trait-age variations cause the face to be altered to different extents in its shape and texture due to time lapse. The shape of the face changes dramatically from childhood to adulthood and its texture (wrinkles) changes markedly from adulthood through old age [7, 8]. Face recognition is affected by external factors such as sensor limitation (distance from camera), spatial resolution, acquisition spectrum (visible vs. infrared), frame rate (2D vs. 3D), variations in user interaction (expression and pose), lifestyle and changes in the environment (Occlusion Illumination, Make-up, Accessories and Background scene). These factors are considered as noise to the recognition system.

Face recognition is adversely affected by trait aging. This challenge is enormous because there could be some level of disparity between the appearance of the face of a human subject at the point of training or enrollment and when the human subject's face is being verified or authenticated. The development of algorithms for trait-age invariant face recognition have been suggested as a means of checkmating the adverse effect of trait aging. A database, called the Face and Gesture Recognition Network (FG-NET) have been adopted by many researchers in the quest to develop trait-age invariant algorithms for face recognition. The database was released in 2004 and is used by researchers to study the effect of aging on the face [9] and therefore was selected for this experimental research work.

Convolutional Neural Networks (CNNs) have shown great promise for such tasks as image classification, face recognition and object detection. CNNs are able to extract features from complex datasets using several pooling and convolutional layers. Classification is performed by linear layers (in this case, Softmax). CNNs can solve a variety of complex recognition tasks using its powerful feature extraction and learning ability but not without some drawbacks such as its exorbitant computational cost, enormous training data, etc. In order to mitigate against these drawbacks, many researchers have proposed very effective solutions, such as data augmentation, regularization (dropout, batch normalization and data augmentation) and various activation functions. The growth of the network architecture has been very rapid, it started with LeNet-5 [10] followed by AlexNet [11] and went on to GoogLeNet [12], VGG-Net [13], ResNet [14], Inception network [15] and lastly, Inception- ResNet [16]. The numerous advantages of CNNs have been highlighted by feedbacks from both researchers and industrialists.

2 Related Works

Kamarajugadda and Polipalli proposed the use of the features of the periocular region [17]. The researchers submitted that the features of the periocular region were trait-age invariant. The features of the periocular region were extracted using Scale Invariant Feature Transform (SIFT) and Speeded Up Robust Features (SUFT). The extracted feature vector space was given a low dimensionality using Enhanced Principal Component Analysis (EPCA). The extracted features were fed into an ANN based classifier. An accuracy of 99.97% was achieved using the FG-NET database.

Zhou and Lam proposed an identity-inference model that enabled age-subspace learning using appearance age labels [18]. The human identity was modelled along with its aging variables using Probabilistic Linear Discriminant Analysis (PLDA). The age-subspace was determined iteratively using the Expectation-Maximum (EM) algorithm and joined using Canonical Correlation Analysis (CCA) for maximum correlation. An accuracy of 89.8 and 93.12% were recorded for the FG-NET and AG-IIM face image database.

Duong et al. proposed a novel generative probabilistic model that was able to impressively follow the progression of the aging process of the face [19]. The researchers used a Temporal Non-Volume Preserving (TNVP) transformation to model the effect of face aging at each stage. The suggested model was able to show non-linear age related variance and give a smooth synthesis of the aging process. The proposed model was evaluated for accuracy of synthesis in age-processed faces and cross-age face verification. Its consistency was tested on well-known face aging databases like MORPH and FG-NET. The model had a Rank-1 identification accuracy result of 47.72% with one million distractors (MegaFace Challenge 1—FGNET) when it was protocol trained with cross-age faces.

Osman Ali et al. suggested the combination of texture and shape features for face recognition [20]. The authors proposed the use of phase congruency features on a face edge map for shape analysis. The face edge map helped with tracking cranio-facial growth patterns that are pronounced in children as they grow into adults. A Local Binary Pattern (LBP) texture descriptor was used to handle texture variations in subjects. The authors submitted that the fusion of facial features yielded more accurate recognition results than single features. The proposed system was used on the FG-NET database. A recognition accuracy of over 93% was recorded.

3 Methodology

3.1 Database Specifications and Complexities

The FG-NET database contains 1002 images from 82 different subjects with age ranging from newborns to 69 years. However, ages up to 40 years are the most populated in the database. The quality of the images depends on the skill of the photographer, the quality of the photographic paper and capturing equipment used. This was so because most of the images used were acquired from the personal collections of the subjects save the more recent ones which were digital. The images exhibited considerable variability in resolution, image sharpness, illumination, background, view-point and facial expression. This made it a highly challenging data for age-invariant face recognition. Furthermore, occlusions in the form of spectacles, facial hair and hats were also present in a number of images.

3.2 Deep Learning Model

Inception-ResNet-v2 was trained with more than one million images from the ImageNet database (used for the ImageNet large-scale visual recognition challenge). It is a well-known convolutional neural network. This CNN is 164 layers deep and it is able to classify images into a thousand unique object categories. It also has an image input size of 299×299. In this study, a pre-trained Inception-ResNet-v2 network that has already learned to extract powerful and informative features from natural images was used as a starting point to learn representative features from the FG-NET database using transfer learning for age-invariant face recognition. The network architecture of Inception-ResNet-v2 is shown in Fig. 1. In order to use Inception-ResNet-v2 network, its application was installed in MATLAB R2018b. The host computer had 64 GB of RAM and a Core i7 Intel processor.

Fig. 1. An overview scheme for training and testing of the inception-ResNet-v2 using FG-Net AD augmented database for the development of the proposed model

3.3 Pre-processing the FG-NET Database for Deep Learning

Deep learning models need a very large amount of data for training. Since the FG-NET database contains around 10–15 face images of one subject at different ages, it is a very small dataset for deep learning. The data was pre-processed to improve the images and increase the number of images for better feature extraction using deep learning. Each image was read from the dataset one after the other and the following operations were performed:

1. If it was a grayscale image, it was converted into an RGB image so that all images in the database have the same number of channels.
2. The faces of the subjects were detected and cropped using Viola Jones Face Detector to remove the background and allow the deep learning model to extract more rich and relevant features.
3. Five different versions of the cropped image were made by adding different types of noises to it. The noises are listed below:
   - No noise (original cropped image preserved)
   - Gaussian Noise
   - Poisson Noise
   - Salt and Pepper Noise
   - Speckle Noise.

Addition of noise increases the number of images in the database without over-fitting the network and extracts better feature representation. The technique of adding noises to images in order to increase the numbers is called data augmentation [21]. Data augmentation is a well-established concept both in literature and real-world [22, 23]. As a result, we obtained a pre-processed FG-NET database with a total of 4800 face images and around 60 face images per subject.

3.4 Training the Deep Learning Model

Transfer learning was used to train the Inception-ResNet-v2 network on the pre-processed FG-NET database for age-invariant face recognition. The training process is summarized step-by-step below:

a. Load the pre-processed FG-NET database in MATLAB using ImageDatastore object.
b. Split the database into train set (80% images) and validation set (20% images).
c. Resize all images in the train and test sets to 299 × 299 to make them compatible with Inception-ResNet-v2 network.
d. Load Inception-ResNet-v2 network in MATLAB.
e. Specify training options for transfer learning.
f. Re-train Inception-ResNet-v2 network on the train set of pre-processed FG-NET.
g. Test the trained network on the validation set.
h. Compute the network accuracy.

The training was successfully completed on a CPU with Core i7 processor and 64 GB RAM in 10 h. The training progress plots showing the accuracy and loss of the network at every iteration are shown in Fig. 1. The figure shows a smooth training progress and convergence to the optimum solution.

3.5 Using the Trained Deep Learning Model for Face Recognition in FG-NET

The trained deep learning model was used for testing on images from the FG-NET dataset by following these steps:

a. Read an image from FG-NET database in MATLAB.
b. Convert the image into RGB format if it is a grayscale image.
c. Detect and crop the face of the subject in the image using Viola Jones Face Detector.
d. Resize the image to 299 × 299 to make it compatible with Inception-ResNet-v2.
e. Load the re-trained Inception-ResNet-v2 network.
f. Pass the image from step (d) to the re-trained network for class prediction.
g. Observe the predicted result and compare it to the ground truth.

3.6 The Propose "Age-Invariant" Face Recognition Model

Figure 1 gives an overview of the proposed age-invariant face recognition model. The input represents the preprocessed and augmented images. The stem represents the part of the convolutional neural network called the residual network. The stem does the initial convolution operations on the input images. The Inception-Restnet-A represents the first Inception module that performs the next convolutional operations. This operation is repeated five times. Reduction block A performs the first dimension reduction through max-pooling and passes its output to the next Inception module called Inception-ResNet-B. Inception-ResNet-B performs the convolution operation and the operation is repeated ten times. The Reduction-B block also carries out dimension reduction using max-pooling. The final Inception module is called Inception-ResNet-C which does the final convolution and assigns the final weights/biases on the images. Dropout is performed and a final dimension reduction is done by average pooling. From the stem to the average pooling layer are used to learn and extract the image features. While the final layer called Softmax is used to classify the images from the features learnt (for this application, only eighty-two (82) classes are needed). Through the instruments of transfer learning and loss function, the proposed model is able to classify face images correctly even with the interference of "trait-aging", after several epochs. Thus, making the proposed model "age-invariant".

4 Result Analysis

Figure 2 shows the accuracy and loss (error) of the proposed age invariant face recognition model. After the model parameters were learned and fixed through training and transfer learning, it was discovered that the training accuracy was 99.97%. When the test images were fed into the CNN model, the comparison carried out with the ground truth established during testing yielded an accuracy of 99.90%. This means nearly all the test images fed into the convolutional neural network were classified correctly.

The loss (error), which was derived using the negative log-likelihood was low. The model was able to satisfactorily interprete the training and testing data. The summation of errors made during training and testing kept decreasing as the epochs increased. After fifty (50) epochs at iteration of 5000, the learning model had well minimized loss function values on both training and testing.

Table 1 presents a summary of the key performance indices of the proposed age invariant face recognition model. The proposed model had a training loss of 0.008% and a testing loss of 0.003%. The training accuracy of the proposed model was 99.97% and the testing accuracy was 99.90%. The training was stopped after 50 epochs at 5000 iteration. The calculated experimental Mean Squared Error (MSE) and Mean Absolute Error (MAE) were 0.0158 and 0.0637 respectively.

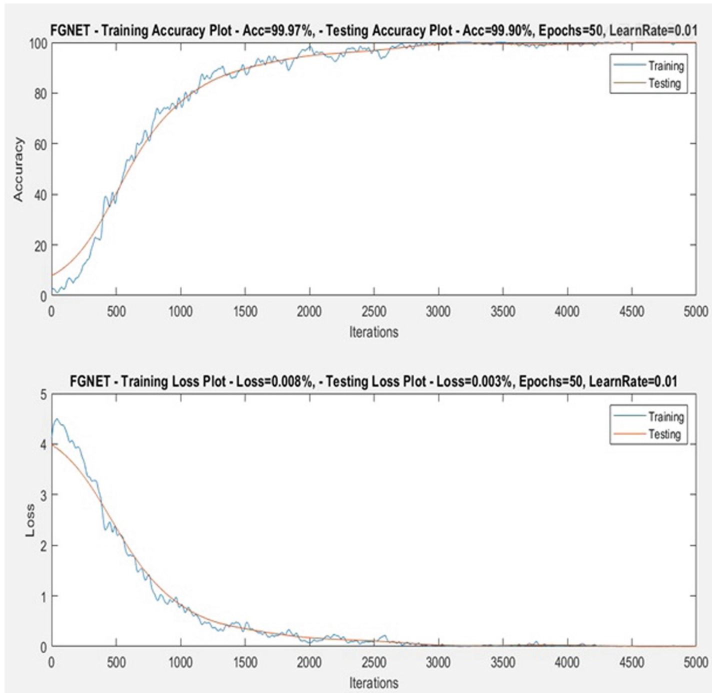

Fig. 2. **a** Percentage accuracy and **b** Loss (Error) of training and testing cumulative score curve of Fg-Net database using Re-trained inception-ResNetv2 and transfer learning technique

Table 1. Table of results

Metric	Experimental result
Training accuracy	99.97%
Testing accuracy	99.90%
Training loss	0.008%
Testing loss	0.003%
Maximum epochs setting	50
Mean Squared Error (MSE)	0.0158
Mean Absolute Error (MAE)	0.0637
Training finished	Reached final iteration at 5000

5 Conclusion

In this paper, a trait-aging invariant face recognition system using convolutional neutral network is proposed. The Face and Gesture Recognition Network Ageing Database (FG-NET AD) was augmented by introducing various noises, thereby enlarging the database and making it suitable for deep learning operation. Transfer learning, regularization and various operations such as training, testing etc. were performed. To the best of our knowledge, the result of our proposed method outperforms the best result recorded in the literature using similar database.

Acknowledgements This paper was sponsored by Covenant University, Ota, Ogun State, Nigeria.

References

1. Okokpujie K, Noma-Osaghae E, John S, Grace KA, Okokpujie I (2017) A face recognition attendance system with GSM notification. In: IEEE 3rd international conference on electro-technology for national development (NIGERCON), IEEE 2017, pp 239–244
2. Okokpujie K, Noma-Osaghae E, John S, Oputa R (2017) Development of a facial recognition system with email identification message relay mechanism. In: International conference on computing networking and informatics (ICCNI), IEEE 2017, pp 1–6
3. Charity A, Okokpujie K, Etinosa NO (2017) A bimodal biometric student attendance system. In: IEEE 3rd international conference on electro-technology for national development (NIGERCON), IEEE 2017, pp 464–471
4. Noma-Osaghae E, Robert O, Okereke C, Okesola OJ, Okokpujie K (2017) Design and implementation of an iris biometric door access control system. In: International conference on computational science and computational intelligence (CSCI), IEEE 2017, pp 590–593
5. Okokpujie K et al (2018) Fingerprint biometric authentication based point of sale terminal. In: International conference on information science and applications, pp 229–237. Springer
6. Okokpujie K et al (2018) Integration of iris biometrics in automated teller machines for enhanced user authentication. In: International conference on information science and applications, pp 219–228. Springer
7. Narayanan R, Rama C (2016) Modeling age progression in young faces modeling age progression in young faces
8. Fu Y, Guo G, Huang TS (2010) Age synthesis and estimation via faces: a survey. IEEE Trans Pattern Anal Mach Intell 32(11):1955–1976
9. Yan S, Wang H, Tang X, Huang TS (2007) Learning auto-structured regressor from uncertain nonnegative labels. In: 2007 IEEE 11th international conference on computer vision
10. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):1–57
11. Krizhevsky A, Sutskever I, Hinton G (2012) ImageNet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 1097–1105
12. Szegedy C, Reed S, Sermanet P, Vanhoucke V, Rabinovich A (2014) Going deeper with convolutions. arXiv1409.4842v1 17 Sep 2014, pp 1–12
13. Karen S, Andrew Z (2015) Very deep convolutional networks for large-scale image recognition. Public as a conference Pap. ICLR. arXiv:1409.1556, pp 1–14
14. Kaiming H, Zhang X, Shaoqing R, Jian S (2016) Deep residual learning for image recognition. Proc IEEE Conf Comput Vis Pattern Recognit 2016:770–778
15. Szegedy C, Vanhoucke V, Shlens J, Wojna Z (2015) Rethinking the inception architecture for computer vision. arXiv1512.00567v3 [cs.CV] 11 Dec 2015 Available @ IEEE Xplore, pp. 1–10
16. Szegedy C, Ioffe S, Vanhoucke V (2016) Inception—V 4, Inception -ResNet and the impact of residual connections on L earn. In: Work. track—ICLR 2016, no 2015, pp 2015–2017
17. Kamarajugadda KK, Polipalli TR (2019) Age-invariant face recognition using multiple descriptors along with modified dimensionality reduction approach. Multimed Tools Appl
18. Zhou H, Lam KM (2018) Age-invariant face recognition based on identity inference from appearance age. Pattern Recognit 76:191–202

19. Duong CN, Quach KG, Luu K, Le THN, Savvides M (2017) Temporal non-volume preserving approach to facial age-progression and age-invariant face recognition. Proc IEEE Int Conf Comput Vis, pp 3755–3763
20. Osman Ali AS, Sagayan V, Saeed AM, Ameen H, Aziz A (2015) Age-invariant face recognition system using combined shape and texture features. IET Biom 4(2):98–115
21. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data 6(60):1–48
22. Nelson MRA, Matheus G, Leandro TH, Heitor SL (2017) The effect of data augmentation on the performance of convolutional neural networks. In: Brazilian society of computational intelligence, Niteroi, rio, Janeiro, pp 1–12
23. Nanni L, Brahnam S, Maguolo G (2019) Data augmentation for building an ensemble of convolutional neural networks. LNCS, Springer, Singapore