KINGDOM OF SAUDI ARABIA
MINISTRY OF EDUCATION
ALBAHA UNIVERSITY

January – June 2020 Volume 4 Issue 1

ALBAHA UNIVERSITY JOURNAL OF
BASIC AND APPLIED
SCIENCES

Labeled Faces in the Wild Recognition using Enhanced Convolution Neural Network Deep Learning

Mohammad E. Alzahrani

Department of Computer Science,
College of Computer and Information Technology,
Albaha University, Albaha, Saudi Arabia

p-ISSN:1658-7529
e-ISSN:1658-7537

A Refereed Academic Journal Published by Albaha University

ALBAHA UNIVERSITY JOURNAL OF

BASIC AND APPLIED SCIENCES

January- June 2020 Volume 4 Issue 1

CONTENTS

Editorial

Review Article

Research

Author Guidelines

https://portal.bu.edu.sa/web/jbas
Published by BUJBAS, Albaha University, 65451 Albaha, Kingdom of Saudi Arabia

All scientific articles in this issue are refereed.

SCOPE

Albaha University Journal of Basic and Applied Sciences (BUJBAS) publishes English language, peer-reviewed papers focused on the integration of all areas of sciences and their application. Supporting the concept of interdisciplinary BUJBAS welcomes submissions in various academic areas such as medicine, dentistry, pharmacy, biology, agriculture, veterinary medicine, chemistry, mathematics, physics, engineering, computer sciences and geology.

BUJBAS publishes original articles, short communications, review articles, and case reports.

The absolute criteria of acceptance for all papers are the quality and originality of the research.

Papers for publication should be addressed to the Editor, via the website:
https://portal.bu.edu.sa/web/jbas
E-mail: bujs@bu.edu.sa

ONLINE SUPPORT

BUJBAS is published by Albaha University. For queries related to the journal, please contact
https://portal.bu.edu.sa/web/jbas
E-mail: bujs@bu.edu.sa

L.D. No: 1438/2732

p-ISSN: 1658-7529

e-ISSN: 1658-7537

Albaha University Journal of Basic and Applied Sciences 4(1) (2020) 7–14

Article available at Albaha University Journal of Basic and Applied Sciences

Journal of Basic and Applied Sciences

Journal homepage: https://portal.bu.edu.sa/web/jbas/

Labeled Faces in the Wild Recognition Using Enhanced Convolution Neural Network Deep Learning

Mohammad E. Alzahrani [a,*]

[a] *Department of Computer Science, College of Computer and Information Technology, Albaha University, Albaha, Saudi Arabia*

ARTICLEINFO

Article history:
Received 26 August 2019
Received in revised form
22 October 2019
Accepted 18 November 2019

Keywords:
Face recognition
Deep learning
Convolutional neural network
Labeled faces in the wild (LFW)

ABSTRACT

This paper enhances the Convolutional Neural Network (CNN) to increase accuracy in recognizing faces in a crowd. The enhanced CNN is built upon the well-known VGG-19 which a very deep learning CNN technique. Prior to the recognition phase, the image data was preprocessed using standard image processing mechanism. Experiments on existing publicly available Labeled Faces in the Wild (LFW) dataset were conducted. The proposed enhanced-CNN outperforms the other existing methods and achieved up to 99.91% accuracy. The best accuracy was obtained from the experiment on LFW-100 sub-dataset, which has 5 persons with 100 images/person. In future, the proposed enhanced-CNN can be used for the development of Hajj safety alert system.

1. Introduction

Among the vision researchers, face recognition and verification have become one of deep focus research area [1, 2]. The ability of the face regonition/verification to identify and to process the emotional state, attention concentration, and the intent of others makes the face recognition/ verification tremendous societal importance [3-5]. Moreover, research works in [6] and [7] introduced Labeled Faces in the Wild (LFW) dataset to encourage further researches using images taken in common, everyday settings and Ethnics.

Labeled Faces in the Wild (LFW) is a database of face photographs intended for investigation the unconstrained face recognition problems. This database was created and maintained by researchers at the University of Massachusetts, Amherst [8]. The researchers in [8] collected 13,233 images of 5,749 people from the web, then used the Viola Jones face detector to detect and center them. 1,680 of the pictures have two or more distinct photos in the dataset. The original database contains four different sets of LFW images and three different types of "aligned" images. According to the researchers, deep-funneled images produced superior results for most face verification algorithms compared to the other image types.

On the other hand, the mass movement of people during hajj is the largest in the world and causing world's worst traffic jam. As the number of visitors rose each year, the problem got worse. Several crowd crush and stampede disasters have happened in the past. To avoid such disasters, it is imperative to develop a real-time alert and monitoring system that able to warn early on any movement anomalies and to recommend the best action. When the system recognizes numbers of faces of a group of ethnics at certain time and they are not supposed to be in the passage during that time, an analysis is performed and an alert may triggered.

Deep Learning is a part of Machine Learning in Artificial Intelligence (AI). Deep learning algorithms composed of Artificial Neural Networks with more than one hidden layer, called as Deep Neural Networks (DNNs). The DNNs are used to model complex non-linear relationships in both supervised and unsupervised settings [9]. Compared to traditional machine learning algorithms, deep learning models can provide significant improvement in areas such speech recognition and language translation as evidenced by the significant improvement in Google Translate after switching from Phrase Based Machine Translation (PBMT) to Neural Machine Translation (NMT). The different type of Deep Learning models includes Deep Auto encoders, Convolutional Neural Networks (CNNs), Restricted Boltzmann Machines (RBMs), Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) models [10].

This paper proposes an enhanced Convolutional Neural Network (CNN) deep learning architecture for a face recognition system. The enhanced CNN is built upon the well-known VGG-19, which is a very deep learning CNN technique. Thus, the proposed method contributes towards the core recognition engine for real time crowd accident alert system during hajj session and the main contribution of this paper (also the differences from the existing CNNs) are:

* Corresponding author: Department of Computer Science, College of Computer and Information Technology, Albaha University, Albaha Saudi Arabia.

Tel.: +966 53 509 9971.
E-mail address: meid@bu.edu.sa (M.E. Alzahrani).

1658-7537/

- A new architecture of VGG-19 CNN by implementing recurrent connections in neurons inside the convolutional layers.
- Introducing the use of :
  - Multi-Bias non-linear Activation (MBA) function instead of Batch normalization
  - Concatenated ReLU (CReLU) instead of ReLU function
  - Sequence of Enhanced Convolutional Layers (ECLs) for Enhancing Block
  - Data augmentation
  - Spatial Pyramid Pooling (SPP) layer before the fully connected layers to eliminate the fixed input image constraint.
- Replacing the use of the maxpooling layers by a new efficient Generalizing Pooling layers to adapt more to complex patterns.

The paper is structured as follows. Section 2 summarizes the previous works on LFW recognition and deep learning. Section 3 discusses the proposed recognition method. Section 4 discusses the experimental set up, results and discussion and lastly, conclusion is given in Section 5.

2. Related Works

Research on face recognition and verification started by recognizing face from one example view [1,2], and then it grows to some new issues [3]. In principle, face verification is to conclude the matching of two face images without any previous examples of those identities. Let *m* and *n* represent the two images, image *m* can be observed as a single training example as an identity of a specific person. Therefore, the face verification can be outlined as a binary classification problem, i.e. to decide whether image *n* is in the same class as image *m* or not.

The use of Convolutional Neural Network (CNN) to learn a metric between face images has been discussed by Chopra et al. [3]. A structure of the face recognition problem was specifically discussed as a large number of classes problem with small numbers of training examples per class. Observation on the difficulties in Augmented Reality database and indication of current positive application of CNNs to face recognition/verification was reported by Martinez and Benavente in [4].

A method for determining whether two images was the same object from a small number of examples was introduced by authors in [2, 5], using images available in [6, 7]. The images were taken from selected news articles, representing people in a wide variety of settings, lighting, poses, and expressions. Huang et al. [8] conducted further work by removing duplicate images, re-labelling some images by hand, preparing protocols to use the images. Then released the dataset as “Labeled Faces in the Wild” in the LFW technical report [8]. Many works have been carried out on face recognition using LFW dataset; Table 1 shows six standard methods on LFW recognition along with their accuracies.

Up to now, the highest reported accuracy on LFW is 99.63%, achieved by by Schroff et al. [17]. The authors reported that out of 6000 image pairs, the method can recognize correctly 5978 pairs. The real errors were only 17 pairs; due to five of the total 22 errors correspond to labeling errors in LFW.

Krizhevsky et al. [18] introduced AlexNet, as one of the most renowned CNN architectures for classification. The AlexNet has structure of feature maps of {96,256,384,384,256} kernels with pooling on the 1^{st}, 2^{nd}, and 5^{th} layers; kernel sizes were {11,5,3,3,3}, respectively. Two fully connected layers of 4096 units are added to the end of the network, which resulted in 60 million parameters. Researchers then tend to look for more deep models with more complex building blocks. Authors in [19, 20] showed that more deep networks have more ability to signify definite function classes more faster and effective. The authors also stated that in general, the more deep networks have a lesser memory footprint for the duration of inference. This benefit allows the deployment of the networks on mobile computing devices. The most popular deeper network VGG-19 is a deep network with 19 layers and uses fixed 3x3 sized kernels was developed by Simonyan and Zisserman [21]. The VGG-19 won the 2014 ImageNet challenge.

A 22-layer network named GoogLeNet was introduced by Szegedy et al. [22]. The GoogLeNet utilizes an inception blocks [23]. The inception blocks is a nested network, where the input is branched into several different convolutional sub-networks. At the end of the block, the nested network is concatenated. To reduce the dimensionality of the feature maps, the inception blocks use 1x1 kernel convolutions. When this paper is written, ResNet architecture [24] is the best deep architecture of CNN in term of performance.

Table 1 Selected six standard methods on LFW recognition.

Method	The use of LFW	Preprocessing	Descriptor	Training set up	Face alignment	Used Techniques	Accuracy
VisionLabs ver. 1.0 [11]	LFW-a (aligned images)		Metric learning & local image	Unrestricted (with lable-free outside data)	Use external data only		93.24%
betaface.com [12]	Original LFW	Converted to grayscale		Unrestricted (with labeled outside data)	Auto-aligned		98.08%
Colour & Imaging Technology (TCIT) [13]	Original LFW	Calculate the average position of the facial area		Unrestricted (with labeled outside data)		Face Feature Positioning	93.33%
Aurora-c-2014-1 [14]	Original LFW			Unrestricted (with labeled outside data)	aligned and funneled sets and some external data	Aurora’s proprietary algorithms, machine learning and computer vision tech.	93.24%
insky.so [15]	Original LFW			Unrestricted (with labeled outside data)			95.51%
Uni-Ubi [16]	Original LFW	Converted to grayscale		Unrestricted (with labeled outside data)	Auto-aligned		99.00%

Authors in [25] discuss a comprehensive review of historical and recent state-of-the-art approaches in deep learning applications such as audio, visual and text processing; social network analysis; and natural language processing. Issues faced in deep learning such as unsupervised learning, black-box models, and online learning are further discussed in the paper.

Several works on LFW recognition using CNN include: DeepFace [26], lightCNN [27], DeepID [28], DeepID2 [29], Baidu [30], DeepID3 [31], Face++ [32], FaceNet [33], and VGGFace [34]. Table 2 presents the characteristics of the Deep Learning methods. The results of experiment in this paper will be compared with the mentioned methods.

3. The Enhanced CNN for LFW Recognition

3.1 Data Collection

The experiments use 12339 images and 5653 people taken from Labeled Faces in the Wild Homepage [35], and each face has been labeled with the name of the person pictured. The number of images per person varies a lot; most of the people have just one image. The author uses pandas/scripting to group the images into four sub-datasets for training and validating classification as shown in Table 3.

Table 3 The dataset for experiments.

Group name	# of people	# of images /person	# of images for training	# of images for testing
LFW-100	5	100	90	10
LFW-20	74	25	20	5
LFW-10	152	12	9	3
LFW-5	423	8	6	2

3.2 The Enhanced CNN for LFW Recognition Method

The proposed method consists of two components. The first component is preprocessing. The preprocessing component reduces the image dimensional and keeps preserving the images' features. Downsizing the dimensionality (pixel size) of the processing images accelerates the CNN training process. Firstly, the video data is converted into image frames. Secondly, the Nearest-Neighbor algorithm is applied to shrink the image, followed by the grey scaling. Then, perform background color processing to remove noise using Frame differencing technique. Lastly, reduce the pixel sizes using Principle Component Analysis (PCA). Fig. 1 depicts the overall architecture of the proposed method.

Table 2 The characteristics of Deep Learning methods related to this work.

Method	Loss Function	Architecture	# of Networks	Features
DeepFace [26]	Softmax	Alexnet	3	Classification using 3D face frontalization
LightCNN [27]	Softmax	Light CNN	1	The use of Light model that is efficient in computational costs
DeepID [28]	Contrastive loss	Alexnet	10	Using supervised deep learning on large labeled data sets
DeepID2 [29]	Contrastive loss	Alexnet	25	Using supervised deep learning on large labeled data sets
Baidu [30]	Triplet loss	CNN-9	10	Using unsupervised deep learning on large labeled data sets
DeepID3 [31]	Contrastive loss	VGGNet-10	50	Using supervised deep learning on large labeled data sets
Face++ [32]	Triplet loss	GoogleNet-24	1	The use of unified embedding for face recognition and clustering and Inception modules for classification
FaceNet [33]	Triplet loss	GoogleNet-24	1	The use of unified embedding for face recognition and clustering and Inception modules from GoogLeNet
VGGFace [34]	Triplet loss	VGGNet-16	1	Improve on training time by reducing the number of parameters in the Convolution layers.

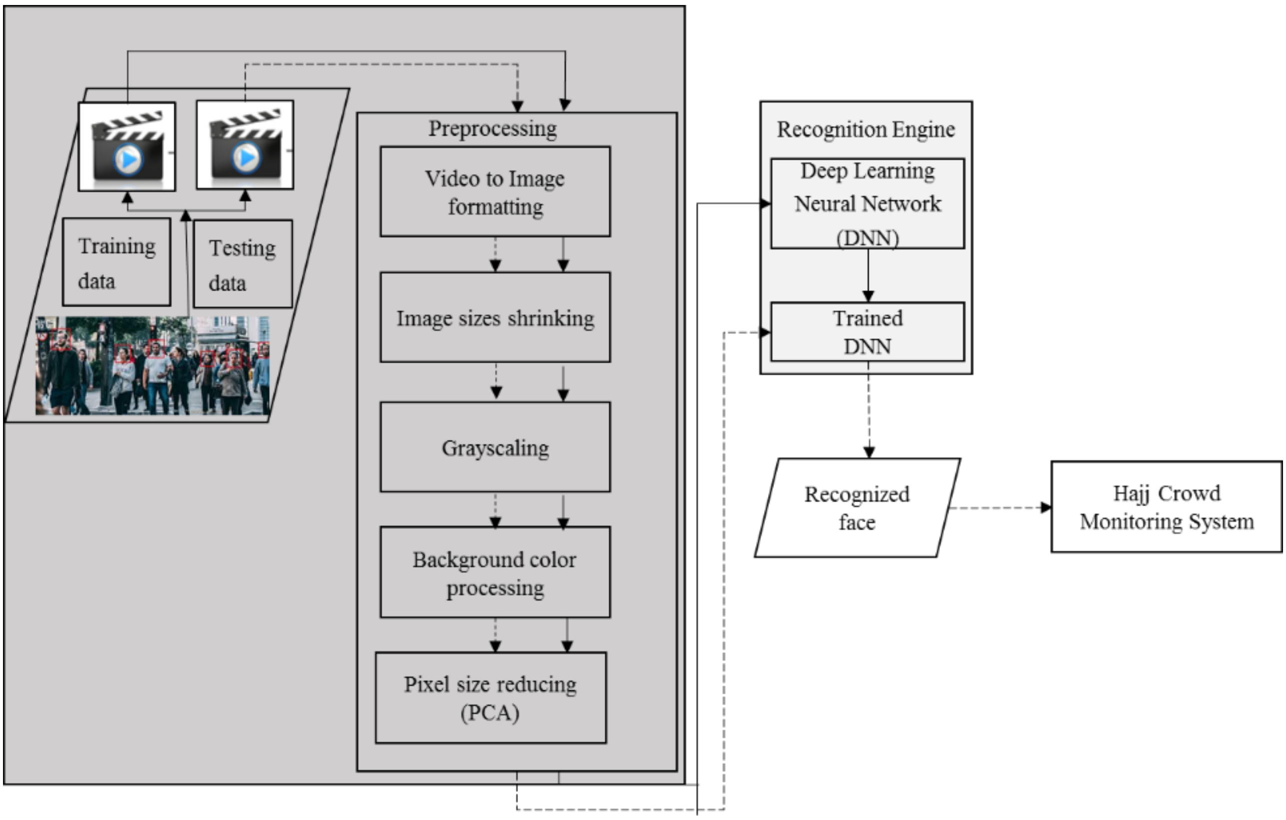


Fig. 1 Overall architecture of the proposed method.

The second component is the enhanced CNN. As dealing with huge image data, this work proposes a new CNN architecture for a face recognition system. The proposed (enhanced) CNN is built upon the well-known and a very deep learning CNN technique, VGG-19. The steps in the proposed architecture are as follow.

1. Integrate recurrent connections inside the convolutional layers.
2. Use Multi-Bias non-linear Activation (MBA) function instead of Batch normalization and use Concatenated ReLU (CReLU) function instead of ReLU function (Rectified Linear Unit). CReLu selects only the positive part of the activation whereas ReLU selects only the negative part of the activation function. It means that convolutional layer + recurrent connections inside + MBA + CReLU compose an Enhance Convolutional Layer (ECL). Then use sequence of ECLs as an Enhanced Block (EB).
3. Replace the max-pooling layers by a new efficient Generalizing pooling layers to adapt to complex patterns.
4. Use data augmentation
5. Eliminate fixed input image constraint using Spatial Pyramid Pooling (SPP) layer before the fully connected.

Figures 2 and 3 show the architecture of the proposed CNN and the enhanced convolutional layer, respectively.

4. Experimental Set up and Results Discussion

4.1 Experimental Set up

The proposed method is implemented using Python language programming Version 3.7.1, Keras and Tensorflow utilities/libraries [36] on a high-end PC with the following specifications. CPU: VPS 2 core, 16GB RAM, and 2TB SSD storage; GPU of 8 cores and 512MB RAM.

4.2 Experimental Results and Discussion

Figures 4(a) to 4(f) show the sequence of executing the CNN on the Tensorflow. The figure confirmed that the processes are working properly.

Figures 5 and 6 show the accuracy of training process and testing process, and loss of training process and testing process for the LFW-100 sub-dataset, respectively. As we can observe in Fig. 5(a), the accuracy during the training process converges fast, due to the good filtering in the Convolution Layer. The same trend is for the accuracy measurement during the testing as shown in Fig. 5(b). Accordingly, the value of the loss function during training converges quickly to zero value as shown in Fig. 6(a) Almost with the same trend, the loss during testing drop drastically after executing 5000 data, see Fig. 6(b).

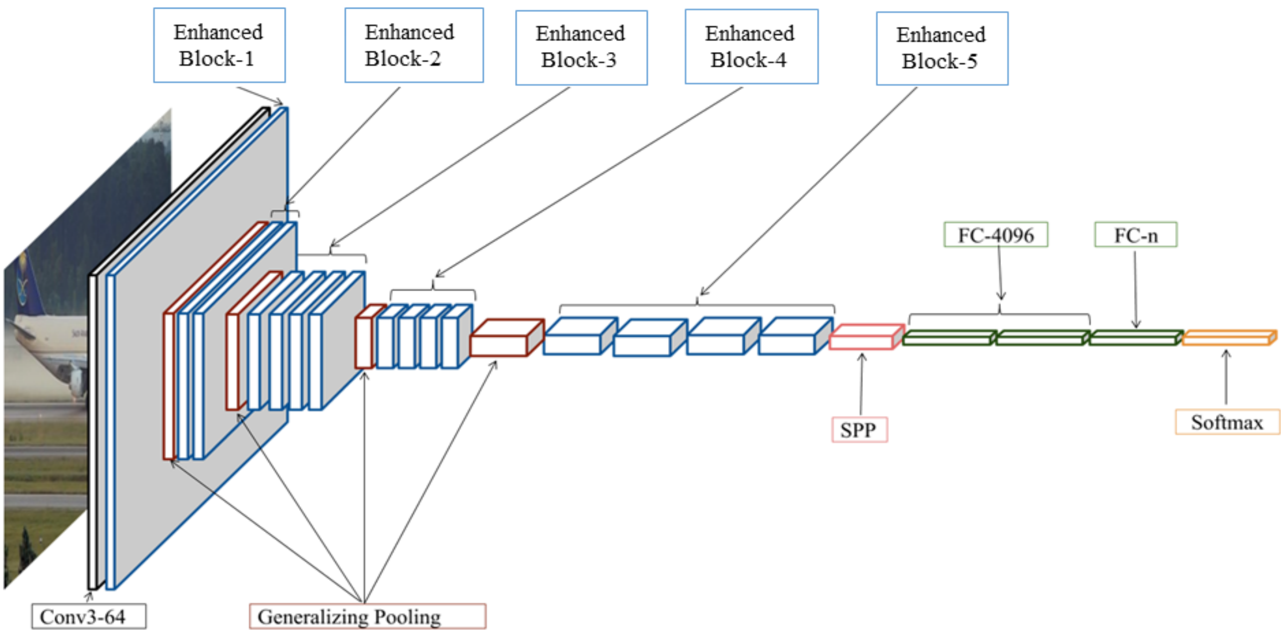


Fig. 2 The proposed enhanced CNN architecture.

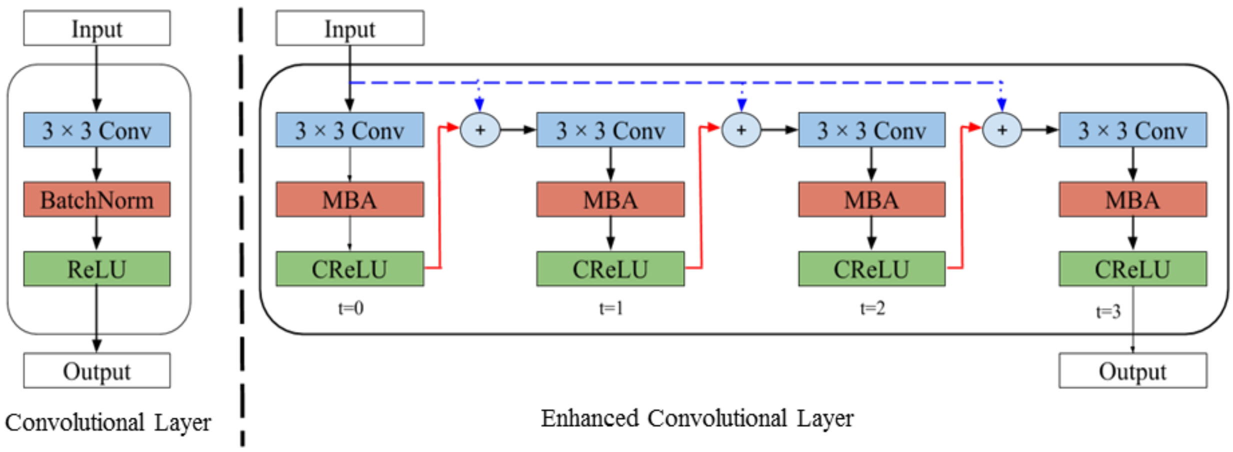
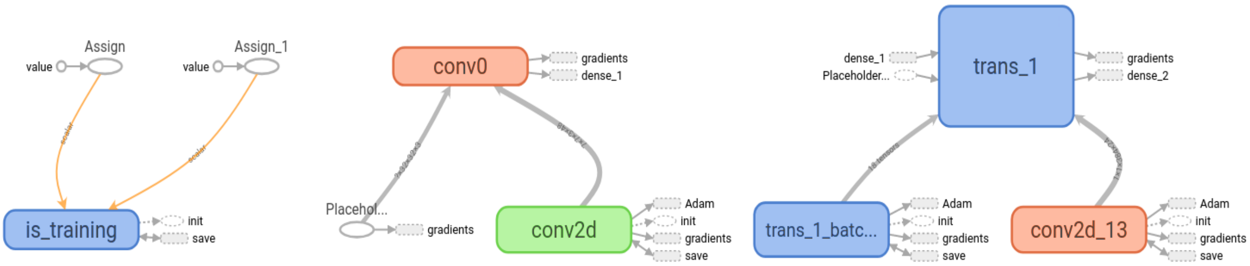
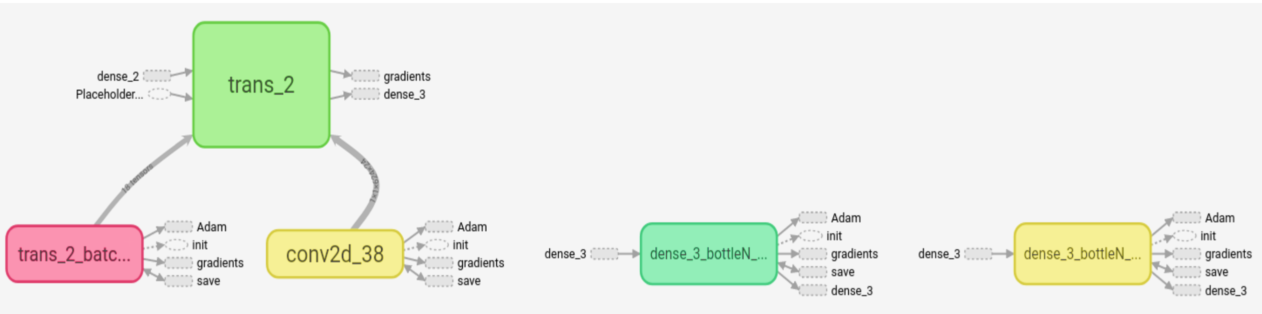


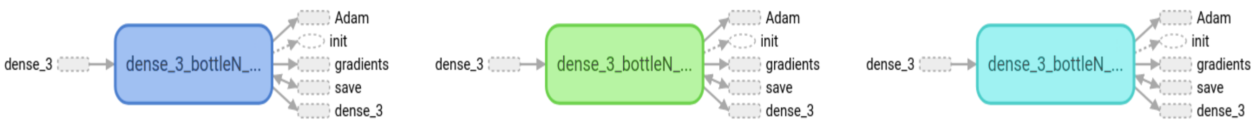
Fig. 3 The enhanced convolution layer.



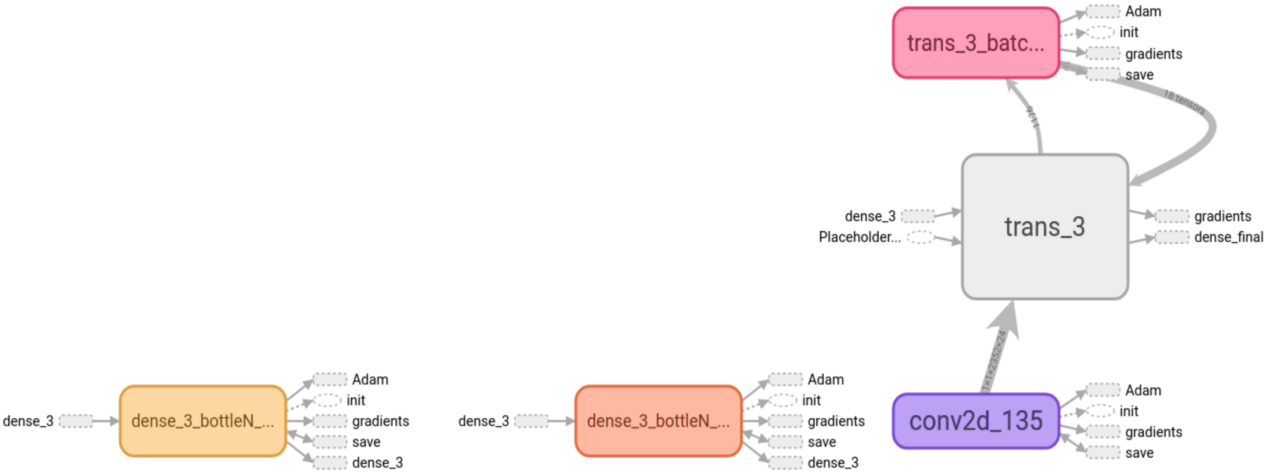
(a)



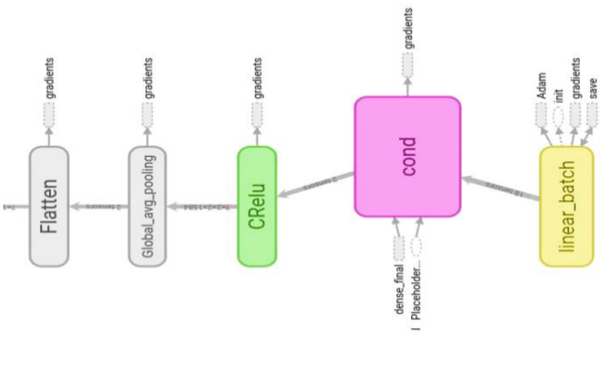
(b)



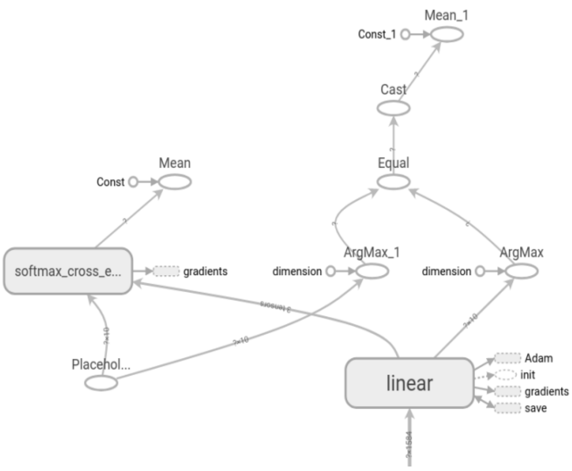
(c)



(d)

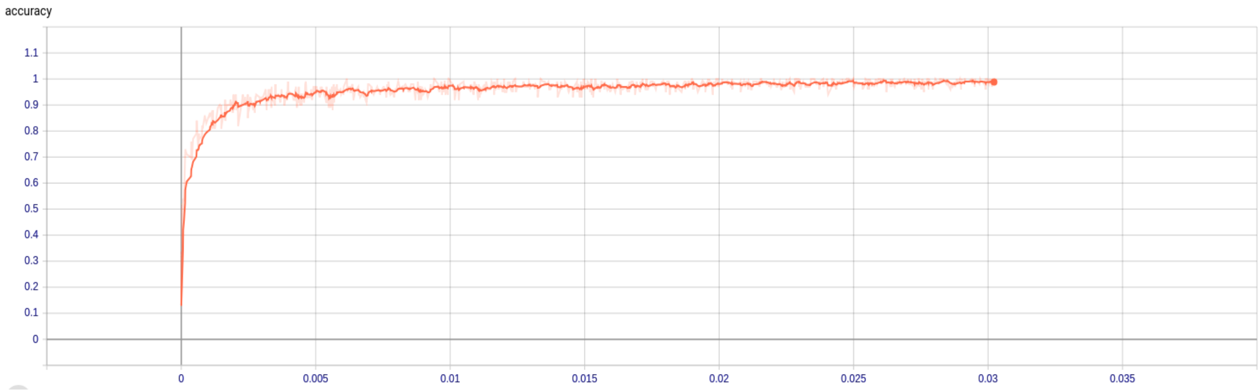


(e)

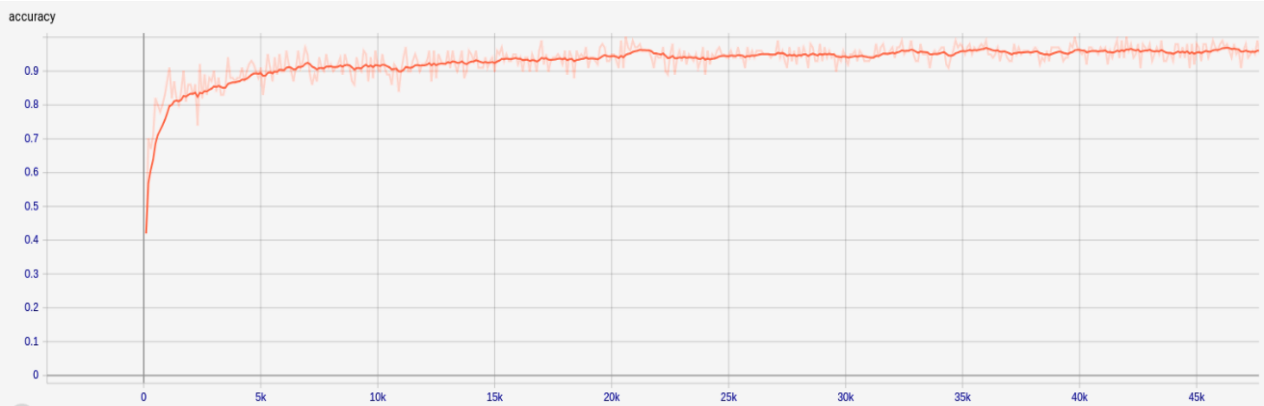


(f)

Fig. 4 Sequence of process execution.

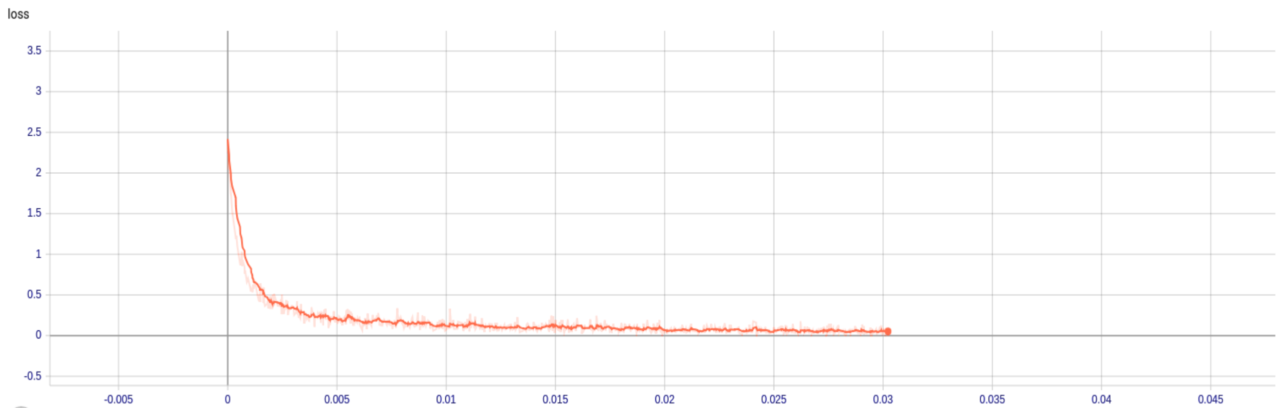


(a) Training Accuracy results.

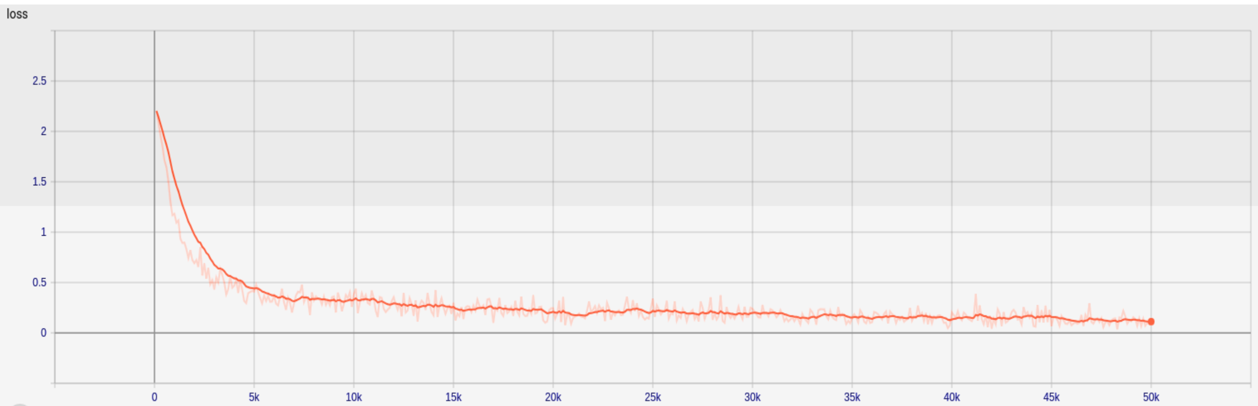


(b) Testing Accuracy results.

Fig. 5 Training and Testing Accuracy results.



(a) Training loss.



(b) Testing loss.

Fig. 6 Training and testing loss results.

Figure 7 depicts a snapshot of running method during the testing process on LFW-100 data.

```
Epoch [17/20], Step [200/600], Loss: 0.0523,Acc: 0.99
Epoch [17/20], Step [300/600], Loss: 0.0214,Acc: 1.00
Epoch [17/20], Step [400/600], Loss: 0.0836,Acc: 0.99
Epoch [17/20], Step [500/600], Loss: 0.0407,Acc: 0.99
Epoch [17/20], Step [600/600], Loss: 0.1038,Acc: 0.98
Epoch [18/20], Step [100/600], Loss: 0.0583,Acc: 0.98
Epoch [18/20], Step [200/600], Loss: 0.0203,Acc: 1.00
Epoch [18/20], Step [300/600], Loss: 0.0461,Acc: 0.98
Epoch [18/20], Step [400/600], Loss: 0.0492,Acc: 0.98
Epoch [18/20], Step [500/600], Loss: 0.0385,Acc: 1.00
Epoch [18/20], Step [600/600], Loss: 0.0460,Acc: 0.99
Epoch [19/20], Step [100/600], Loss: 0.0546,Acc: 0.99
Epoch [19/20], Step [200/600], Loss: 0.0825,Acc: 0.96
Epoch [19/20], Step [300/600], Loss: 0.1295,Acc: 0.97
Epoch [19/20], Step [400/600], Loss: 0.0563,Acc: 0.98
Epoch [19/20], Step [500/600], Loss: 0.0734,Acc: 0.99
Epoch [19/20], Step [600/600], Loss: 0.0262,Acc: 1.00
Epoch [20/20], Step [100/600], Loss: 0.0135,Acc: 1.00
Epoch [20/20], Step [200/600], Loss: 0.0236,Acc: 1.00
Epoch [20/20], Step [300/600], Loss: 0.1148,Acc: 0.96
Epoch [20/20], Step [400/600], Loss: 0.0628,Acc: 0.98
Epoch [20/20], Step [500/600], Loss: 0.0250,Acc: 1.00
Epoch [20/20], Step [600/600], Loss: 0.0466,Acc: 0.99
Test Accuracy of the model on the 10000 test images: 93.12 %
```

Fig. 7 Snapshot of accuracy during the testing process.

Table 4 shows the best results of accuracy for each sub-dataset. As we can observe the use of LFW-100 dataset provides the highest accuracy of the recognition/verification. The experiments do not apply any learning rate strategies that is why, the sub-datasets with smaller size do not perform well.

Table 4. The accuracy for each sub-dataset.

Group name	# of people	# of images/person	# of images (training)	# of images (testing)	The best accuracy (%)
LFW-100	5	100	90	10	99.91
LFW-25	74	25	20	5	94.88
LFW-12	152	12	9	3	86.90
LFW-8	423	8	6	2	76.84

Table 5 shows the comparison of the accuracy of the proposed method with the existing methods presented in [37]. The best accuracy of the proposed method during the testing phase is 99.91% for the LFW-100 dataset. Six methods (DeepID2, Baidu, Face++, FaceNet, VGGFace) achieved accuracy above 99%. DeepID2 and DeepID3 achieved good accuracy because it uses a greater number of networks compare the DeepID that uses same architecture (AlexNet). Baidu achieved a good performance because it implements CNN-9 which is better that AlexNet, and it also apply unsupervised learning. Face++ and FaceNet also performed well because they use Inception modules from GoogLeNet and unified embedding for face recognition and clustering. They are more efficient due to the decrease of parameters to 20 times lesser. Not only give a good accuracy, these two methods provide fast training and testing processing time. Finally, VGGNet provided a good result because it uses efficient mechanism in Convolution Layer that decreases the parameters.

Table 5 Accuracy Comparison.

Method	# Model	Outside data	# Layers	Accuracy (%)
DeepFace [26]	4	4M	8	97.35
LightCNN [27]	60	203K	7	96.45
DeepID [28]	60	203K	7	97.45
DeepID2 [29]	25	203K	7	99.15
Baidu [30]	25	290K	7	99.47
DeepID3 [31]	25	290K	10-15	99.53
Face++ [32]	1	5M	10	99.50
FaceNet[33]	1	260M	22	99.60
VGGFace[34]	20	1M	12	99.65
Proposed Method (on LFW-100 dataset)	1	5M	128	99.91

5. Conclusion

This paper proposed an enhanced CNN for LFW recognition. The best experimental result using benchmark dataset from LFW portal achieved 99.91% accuracy for sub-dataset with 100 images/person. As carried out by Parkhi et al. [34], this paper compared the enhanced CNN with other similar methods that used the same LFW dataset. The comparison results showed that the enhanced CNN outperformed existing CNNs.

Currently, the author is working on feature extraction of ethnical faces using the available LFW dataset. Once the ethnical face features available and since the proposed CNN achieved good results in accuracy, the author plans to use the proposed CNN as the core recognition engine for real time crowd accident alert system during hajj session, in the near future.

Other future work, the author considers also working on Cross-Pose LFW (CPLFW) as a challenge in face recognition with the aim to increase the accuracy due to big data driven machine learning methods, the performance on the database approaches nearly 100%.

References

[1] Beymer D, Poggio T. Face recognition from one example view. The Fifth IEEE Conference on Computer Vision. 1995:500–507.

[2] Ferencz A, Learned-Miller EG, Malik J. Building a classification cascade for visual identification from one example. The Tenth IEEE Conference on Computer Vision. 2005;1:286–293.

[3] Chopra S, Hadsell R, LeCun Y. Learning a similarity metric discriminatively, with application to face verification. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2015:539–546.

[4] Martinez A, Benavente R. The AR face database. CVC Technical Report 24, Ohio State University. 1998.

[5] Learned-Mille EG, Ferencz A, Malik J. Learning hyper-features for visual identification. Advances in Neural Information Processing Systems 17: Proceedings of the 2004 Conference. MIT Press, 2005;17:425.

[6] Berg TL, Berg AC, Edwards J, Maire M, White R, Teh Y-W. Learned-Miller E, Forsyth DA. Names and faces in the news. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR. 2004;2(II):848.

[7] Berg TL, Berg AC, Edwards J, Forsyth DA. Whos in the picture? In NIPS 2005.

[8] Huang GB, Ramesh M, Berg T, Learned-Miller E. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst October 2007.

[9] Waehner K. Open source deep learning frameworks and visual analytics. [Online]. Available: http://www.datasciencecentral.com/profiles/blogs/open-source-deep-learning-frameworks-and-visual-analytics. [Accessed 15 July 2019].

[10] Deng L, Yu D. Deep Learning: Methods and applications, Microsoft Publisher. 2014.

[11] Visionlabs ver. 1.0. http://www.visionlabs.ru/face-recognition.

[12] betaface.com. http://betaface.com.

[13] Taiwan colour & imaging technology (tcit). http://www.tcit-us.com/.

[14] Aurora-c-2014-1. http://www.facerec.com/, 2014.

[15] inksy.so. http://www.insky.so/.

[16] Uni-ubi. http://uni-ubi.com

[17] Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering. IEEE

Conference on Computer Vision and Pattern Recognition (CVPR). 2015:815–823.
[18] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. The NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems. 2012;(1):1097-1105.
[19] Bengio Y. Practical recommendations for gradient-based training of deep architectures. In: Neural Networks: Tricks of the Trade. Springer Berlin Heidelberg. 2012: 437–478.
[20] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research. 2014;15(1):1929–1958.
[21] Simonyan K, Zisserman A. A very deep convolutional networks for large-scale image recognition. arXiv:1409.15, 2014.
[22] Szegedy C, Ioffe S, Vanhoucke, A. Alemi, Inception-v4, Inception-ResNet and the impact of residual connections on learning. arXiv:1602.07261, 2016.
[23] Lin M, Chen Q, Yan S. Network in network. arXiv:1312.4400, 2013.
[24] He K, Zhang X, Ren S, Sun J, Identity mappings in deep residual networks. arXiv:1603.05027, 2016.
[25] Pouyanfar S, Sadiq S, Yan Y, Tian H, Tao Y, Reyes MP, Shyu M-L, Chen S-C, Iyengar S.S. A survey on deep learning: Algorithms, techniques, and applications. September 2018 ACM Computing Surveys (CSUR). October 2018;51(5).
[26] Taigman,Y Yang M, Ranzato M, Wolf L. Deepface: Closing the gap to human-level performance in face verification. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2014;1701–1708.
[27] Wu X, He R, Sun Z, Tan T. A light CNN for deep face representation with noisy labels. 2015, arXiv preprint arXiv:1511.02683.
[28] Sun Y, Wang X, Tang X. Deep learning face representation from predicting 10,000 classes. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2014:1891–1898.
[29] Sun Y, Chen Y, Wang X, Tang X. Deep learning face representation by joint identification-verification. In Advances in Neural Information Processing Systems. 2014;1988–1996.
[30] Liu J, Deng Y, Bai T, Wei Z, Huang C. Targeting ultimate accuracy: Face recognition via deep embedding. 2015, arXiv preprint arXiv:1506.07310, 2015.
[31] Sun Y, Liang D, Wang X, Tang X. DeepID3: Face recognition with very deep neural networks. arXiv preprint arXiv:1502.00873, 2015.
[32] Zhou E, Cao Z, Yin Q. Naive-deep face recognition: Touching the limit of LFW benchmark or not? arXiv preprint arXiv:1501.04690.
[33] Schroff F, Kalenichenko D, Philbin S. FaceNet: A unified embedding for face recognition and clustering. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015:815–823.
[34] Parkhi O, Vedaldi A, Zisserman A. Deep face recognition. British Machine Vision Conference 2015. January 2015. DOI: 10.5244/C.29.41.
[35] LFW Dataset, http://vis-www.cs.umass.edu/lfw/lfw.tgz
[36] Tensorflow - https://www.tensorflow.org. 2017.
[37] Learned-Miller E, Huang GB, RoyChowdhury A, Li H, Hua G. Labeled faces in the wild: A survey, in Advances in face detection and facial image analysis. Springer, 2016:189–248.

ALBAHA UNIVERSITY JOURNAL OF

BASIC AND APPLIED

SCIENCES

Dar Al Manar for Printing +966 17 7223212