Design and Enhancement of a CNN Model to Augment the Face Recognition Accuracy

Shakir Fattah Kak
Duhok Polytechnic University
Akre Technical Institutes
Information Technology Department
Duhok, Iraq
shakir.fattah@dpu.edu.krd
ORCID: 0000-0001-8183-9278

Firas Mahmood Mustafa
Chemical Engineering Dept., Technical College of Engineering, Duhok Polytechnic University
Department of Computer and Communication, College of Engineering, Nawroz University.
Duhok, Iraq.
firas.mustafa@ieee.org
ORCID: 0000-0002-8757-5303

Asaf Varol
The University of Tennessee
Dept. of Engineering Mgmt. & Tech.
College of Engineering and Comp. Sci.
Chattanooga, TN, United States
asaf-varol@utc.edu
asafvarol@maltepe.edu.tr
ORCID: 0000-0003-1606-4079

2022 3rd International Informatics and Software Engineering Conference (IISEC) | 978-1-6654-5995-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/IISEC56263.2022.9998236

***Abstract*— In the last decade, smart home security applications have relied more on human biometrics in their functions, due to the reliability and the high-precision results these technologies provide. Face recognition is one of the popular biometrics in the field of image processing technologies. Human face recognition processing is a complicated operation that involves different factors and circumstances such as the illumination degree and the position of the face that affects the final recognition rate. In this research, the Convolution Neural Network (CNN) architecture is used in the extracting phase of significant features of the face shape, and the SoftMax classification layer was used to identify faces in the fully connected CNN layer. This paper provides an update of CNN architecture by applying a three-batch normalization layer to the CNN design. By applying this modification, the system network speed increased with a better recognition rate. The recognition rate also increased by applying two DWT levels with a bio5.5 filter to the training group of the database images and the tested image before applying the PCA dimensional reduction algorithm instead of using the PCA algorithm alone. The obtained face recognition rates have been improved to 99.75% by applying the proposed CNN scheme. While applying the proposed hybrid approach (using the PCA next to applying DWT-2 levels with bior5.5 filter) has registered a 99.25% recognition rate compared to a 96.75% recognition rate when obtained by applying the PCA method alone. The research has adopted using a set of 360 training images and 40 test images set of the standard ORL Database in its work.**

Keywords*— CNN, DWT, PCA, Face Recognition, ORL Database.

I. INTRODUCTION

In public, a person's identification process involves the use of face recognition models. These models form the essential parts of most security systems, such as authentication control, video tracking, and the financial field in addition to their appearance in social media networks. Face recognition has become most attractive by the escalation of artificial intelligence applications in the last few decades. Face recognition acquired an essential role among other biometric technology models, which is a return to their neutral foundation, making it the most preferable members' identification technique [1].

The effortless inspection of a face recognition model in an uncontrolled environment with not previous attention by the person may be done easily [2, 3]. Many issues encounter older techniques used such as pose deviation, and face travesty, the illumination of the location, background diversity, and variations in facial expressions that may affect the performance of the algorithms negatively [4]. A face-scanning system, as in an iris scan, relies on taking specific facial reference features, keeping and then comparing them [5]. Feature extraction can be more effective using deep learning technology [6, 7]. Artificial Intelligence made an excellent advance by adopting Deep Learning technology in its progress course in solving many complex and hard to address problems. Multidimensional complex structures were analyzed and solved depending on the application of deep learning methods in their processing approaches. Hierarchical representations have been learned by implementing single or multiform on deep learning algorithms that achieved high results especially in many domains like image recognition, natural language processing, semantic segmentation, or other out frowned real-life problems [6, 8, 9].

Many deep learning applications have been presented in different areas, like Convolutional Neural Network (CNN), Stacked Autoencoder [10], and Deep Belief Network (DBN) [9]. In face recognition processing models, CNN has a powerful position as a type of artificial neural network that uses a feature extraction model by input data convolution approach to rise features amount. CNN algorithm had been formulated by LeCun in the implementation of a handwriting recognition project [6, 11]. In this work, an innovative CNN-based model has been constructed by employing three Normalization Batch function next to applying three layers of implementation. This paper consists of three steps for the general framework of the facial recognition process. The preprocessing operations are applied first on the groups of database images. Images resizing followed by feature extractions, finally the extracted features are passed into a classification process. SoftMax Classifier is used to obtain the results (Fig. 1).

978-1-6654-5995-2/22/$31.00 ©2022 IEEE

Authorized licensed use limited to: ULAKBIM UASL - Maltepe University. Downloaded on December 30,2022 at 11:16:21 UTC from IEEE Xplore. Restrictions apply.

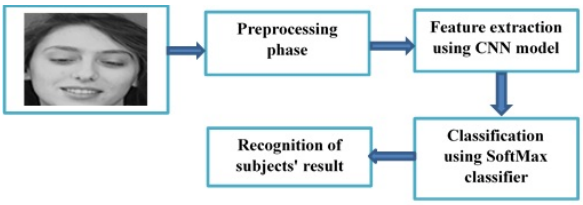

Fig. 1. The phases of face recognition with CNN algorithm

The rest of this research is organized as follows. In section 2, the related works are presented. In section 3, the proposed CNN structure model is introduced. In section 4, the methodology is discussed. In section 5, the face database used in this paper is presented. The experimental results are given in section 6. Finally, we discuss conclusions in section 7.

II. Related Works

The presented work's experiments have been implemented on the ORL database. Weihong Wang et al. [12] suggests using of SIAMESE neural network (SNN), which is a type of Convolutional Neural Network that is used by the authors to overcome the landmark detection and unrestricted recognition problems in face images. The used SIAMESE model of the Convolutional Neural Network provides a convenient solution for the multiple input data and classification type uncertainty.

An article has been published by Yajun Xu et al. [13] had first analyzed a Gabor function application on a one-dimensional dataset; while in the second stage, a two-dimensional Gabor filtering function has been applied in the image dataset by using stretching and rotation operations to obtain two-dimensional Gabor wavelet coefficients. A neural network model has been built by examining the back propagation (BP) Neural Network model on the Gabor wavelet resulted in variables of a human face image example in the image detection process. The authors reported a face detection accuracy rate above 0.93 percentages.

A new paper published by Fang Li et al. [14] applied a pre-training stage on labeled image inputs to obtain optimal factors, in order to design a model based on the supervised fine-tuning features on the ORL database samples with adding the dropout properties. The experiments show that deep autoencoder networks with dropout yield significantly lower test error. Through imposing of dropping items in the context of running the deep autoencoder networks, the model obtained better face recognition results with less time spent.

Another publication prepared by Kamencay P. et al. [15] where the authors proposed an image recognition method based on different image recognition methods, and it incorporates Principal Component Analysis (PCA) for dimensional reduction, Local Binary Patterns Histograms (LBPH), and the K–Nearest Neighbor measurement test method. The recognition rate of the proposed approach reached 98.3%. The experiments have shown result better recognition results can be obtained from applying the LBPH method rather than PCA and KNN approaches.

III. The Proposed CNN Structure Model

The architecture of the CNN module merges multiple types of layers, which are an Input Data Layer, Convolution Layers (Three), Batch Normalization Layers (Three), Activation Layers (Three), Pooling Layers (Two), a Fully Connected Layer, a SoftMax Layer in addition to a Classification (Output) Layer. The layer parameters of the CNN model have been set up and adjusted until the output of the structured model has the maximum accuracy rate. Fig. 1 shows the schematic flow phases of the proposed CNN recognition algorithm.

The CNN model has covered four paces of the overall improvement of the facial recognition technique. The paces start with the pre-processing stage, where the ROI is cropped and resized. Next, the facial feature sets are extracted. After that, at the Classification stage, the features' sets are classified with the SoftMax classification module. The performance of the CNN recognition system has been tested by applying it to the ORL dataset.

IV. Methodology

In this section, face recognition based on CNN with Bach normalization layer is evaluated. We try using multiple normalization layers and modifying the CNN layers parameters to observe the generalization performance of the network.

The three main steps of the proposed algorithm have been implemented are listed below:

- Resizing the input images as 32x32x1, and 64x64x1. The input images have been resized to 32x32 and 64x64.
- A multilayer CNN structure has been constructed. The structure is consisted of 14 layers; convolutional, max pooling, batch normalization, ReLU, convolutional, batch normalization, ReLU, max pooling, convolutional, batch normalization, ReLU, and max pooling layers respectively.
- At the final stage, the extracted features pass through a classification procedure, SoftMax classifier (Fig. 2).

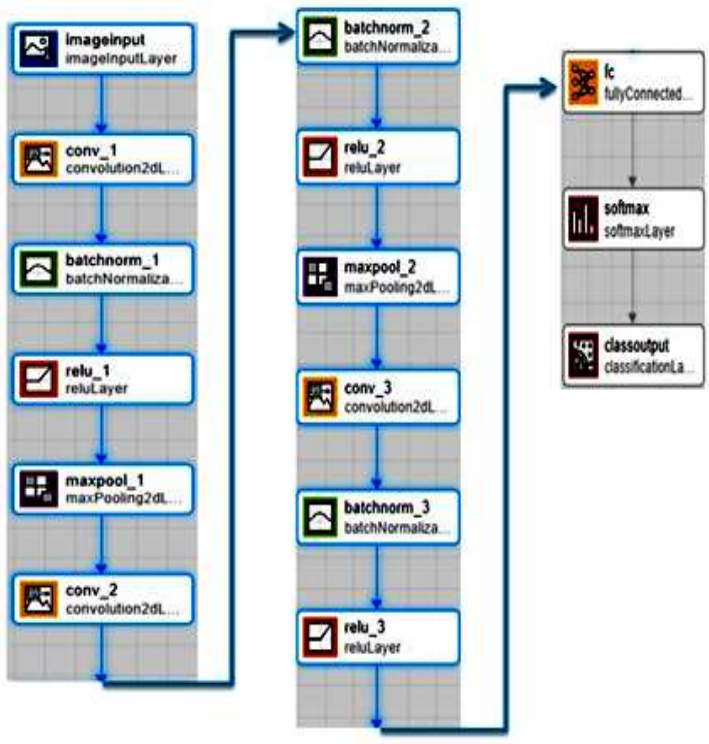

Fig.2. shows the proposed CNN layers structure steps.

Authorized licensed use limited to: ULAKBIM UASL - Maltepe University. Downloaded on December 30,2022 at 11:16:21 UTC from IEEE Xplore. Restrictions apply.

The ORL database and the local-facial datasets have been trained and validated by applying the CNN model methodology on them, analyzing the system network scheme. MATLAB platform has been used in the system structure design and the processes of recognition performance are applied stepwise. Next to the preprocessing, the stage the image is resized to 32x32x1, and 64x64x1. Based on the subjected images, the CNN model performance is examined by applying it on two groups of the subject images, training set, and test set. The recognition rate that has registered for the examined images were 50%, 80%, and 90% for the training images dataset and 50%, 20%, and 10% for the test dataset. Fig. 3 shows the architecture of the convolutional neural network.

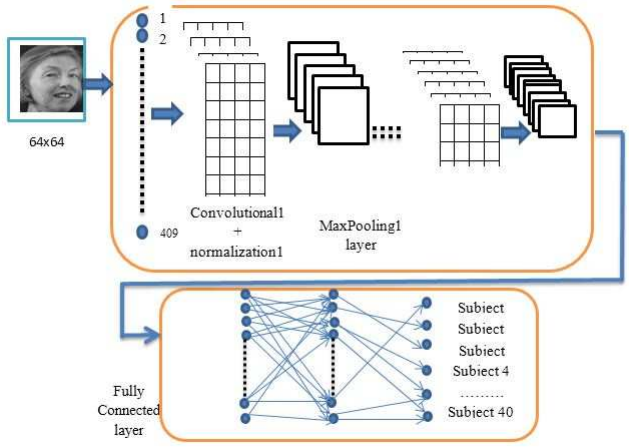

Fig. 3. CNN model architecture.

In this stage, many changes have been made on image features such as image sizes, learning rate, the size of the batch, the epoch number, kernel convolutions set Maxpooling parameters...etc. The CNN model has been examined according to the prediction correction of the test against the training datasets.

V. DATABASE CHARACTERISTICS

The standardized ORL database is the most used facial database in examining face recognition models. In this project, the ORL database has used to test the constructed facial recognition system. The ORL database includes 400 face images of 40 persons, 10 for each, that have had been taken in different poses and illumination situation conditions. Each facial image in the ORL dataset has 10,304-pixel counts. The properties of the ORL database are illustrated in Table I.

TABLE I THE OVERALL RECOGNITION SYSTEM PERCENTAGE.

Properties	Explanation
Purpose	It is used primarily for face recognition
NO. of Subjects	40
NO. of Person	10
NO. of Images	400
Video / Static	Static
Gray / Color	Gray
Image Format	Portable Gray Map (*.pgm)
Image Resolution	92x112
Facial expression	Three facial expression: natural, smiling and closed eye
Face Pose	Moderate pose variation (up and down, quarter profile to frontal view)
Image Background	Dark homogeneous
Face Accessories	Glasses

A set of facial images from the ORL Database have selected at random as showed in Fig. 4. The face samples have special image effects such as lighting accessories and positioning.

Fig. 4. Random samples selected form ORL database.

VI. EXPERIMENTAL RESULTS

According to the Table II, the results of implementing three methods PCA algorithm, DWT with PCA algorithm, and the proposed CNN algorithm are distributed among three testing stages.

The first experiment set has been applied on 200 (50%) training images and 200 (50%) test images of the targeted database. Whilst the second set of experiments has used 320 (80%) training images versus 80 (20%) test images. Eventually, the last set of experiments has been applied to 360 (90%) training images and 40 (10%) test images. The obtained experimental results are presented in Table II. The last row shows the results obtained by applying the proposed CNN algorithm on the targeted database.

TABLE II THE OVERALL RECOGNITION SYSTEM PERCENTAGE.

Technique used	Number of training (Tr) and test (Ts)images		
	I (Tr=50%, TS=50%)	II (Tr=80%, TS=20%)	III (Tr=90%, TS=10%)
PCA (%)	94.35	96.37	96.75
DWT (2 Levels+bior5.5) + PCA (%)	96.15	98.25	99.25
Proposed CNN model (%)	96.5	98.75	99.75

Authorized licensed use limited to: ULAKBIM UASL - Maltepe University. Downloaded on December 30,2022 at 11:16:21 UTC from IEEE Xplore. Restrictions apply.

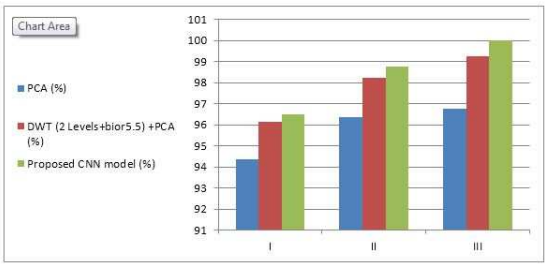

Fig.5. A recognition accuracy comparison of three numerous facial recognition techniques

Fig.5 presents the overall comparison recognition accuracy results for the three testing progresses.

The best experimental precision results appear in applying CNN algorithm on 360 training images of the targeted database that has 100% accuracy rate, whereas the worst precision results have obtained from applying the PCA algorithm on 200 training images of the targeted database (94.35%). Moreover, the efficiency of the new approach has been examined to previous approaches pieces of literature that used CNN algorithm in their models such as CNN [12], GW and the neural network [13], Deep Autoencoder Networks with Dropout [14], and CNN [15] (Table III).

TABLE III COMPARATIVE ANALYSIS WITH OTHER RELATED WORK.

Related Works	Technique	Recognition Accuracy
Wang et al., 2015, [12]	CNN	91%
Xu et al. 2016, [13]	G.W. and Neural Network	93%
Li. et al., 2017, [14]	Deep Autoencoder Networks with Dropout	97.5%
Patrik AMENCAY et al. ,2017[15]	CNN	98.3%
Proposed technique	CNN	98.75%

The performance of multiple recognition methods has been compared by taking 80% of the training samples to 20% of the testing samples from the ORL dataset. The experiments have shown that the accuracy records obtained from the new proposed CNN structure are overweight the recorded accuracy from the other test methods, as shown in Fig. 6.

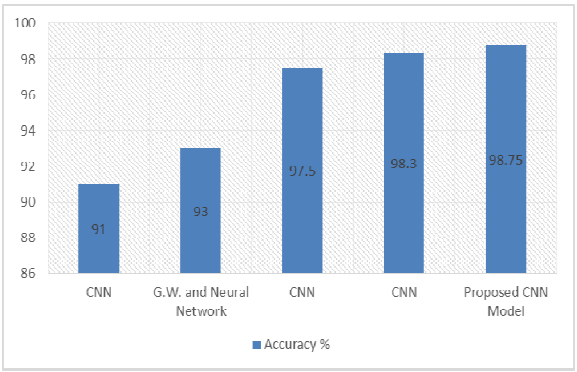

Fig. 6. Comparability of the performance exactness for several methods of face recognition

The comparability Table III above shows that, compared to other face recognition approaches that used different CNN models which had been recorded in 2017 and achieved 99.3% accuracy results [14], the new proposed CNN model got exactly 98.75% accuracy rate. The input parameters of the CNN approach have been chosen depending on the previously applied experiments. The obtained results by applying the proposed method were taken by examining the method on the ORL test database. The minimized imposed features vector in the experiments speeds the classifier up at the same time. To get a high recognition result, it is important to balance between the number of examined features and the targeted accuracy rate.

VII. CONCLUSION

This research focuses on the process of human identification by using an innovative face recognition approach through exploring the performance and the accuracy of the approach at different proceeding levels. At the first level, the PCA algorithm is impended. The second level, DWT (2level, bior5.5) with PCA technique is applied, where the face image samples go through successive processing stages include of DWT against ORL Database, followed by feature extraction and dimension reduction stages within PCA model. At the final level, the CNN technique is applied. The eminent characteristics of the designed model exploit batch normalization of the values resulted from the initial and the last convolution layers of the CNN model.

At the face classification stage, the SoftMax Classifier was used. A performance test for the model was made on the ORL Face Database. The recorded recognition rates of the used studies were outstanding. The optimum recorded recognition rate from manipulating the face recognition model for the three techniques are as follow: 96.75% recognition accuracy achieved using PCA algorithm, and 99.25% is the recognition accuracy achieved by using (DWT-2Levels with bior5.5 filter) followed by PCA algorithm, and the recognition accuracy of using CNN structure is 99.75%.

In this paper, an experimental estimation of CNN-based face recognition architecture was manipulated. The method gave a robust recognition improvement in a lower time. The new approach offers more an enriched accurate capability for an individual face's recognition at different positions and brightness grades.

References

[1] Coşkun, M., et al. (2017). Face recognition based on convolutional neural network. 2017 International Conference on Modern Electrical and Energy Systems (MEES), IEEE.

[2] Kak, S. F., et al. (2018). "A review of person recognition based on face model." Eurasian Journal of Science & Engineering 4(1): 157-168.

[3] Jain, A. K., et al. (2011). Introduction to biometrics, Springer Science & Business Media.

[4] Kak, S. F., et al. (2018). "Discrete wavelet transform with eigenface to enhance face recognition rate." Academic Journal of Nawroz University (AJNU) 7(4): 9-17.

[5] Varol, A. and B. Cebe (2011). Yüz Tanıma Algoritmaları Algorithms of Face Recognition. 5th International Computer & Instructional Technologies Symposium.

[6] Patterson, J. and A. Gibson (2017). Deep learning: A practitioner's approach, "O'Reilly Media, Inc.".

[7] Obaid, K. B., et al. (2020). "Deep Learning Models Based on Image Classification: A Review." International Journal of Science Business(ijsab), 4(11): 75-81.

Authorized licensed use limited to: ULAKBIM UASL - Maltepe University. Downloaded on December 30,2022 at 11:16:21 UTC from IEEE Xplore. Restrictions apply.

[8] Déniz, O., et al. (2011). "Face recognition using histograms of oriented gradients." Pattern Recognition Letters 32(12): 1598-1603.
[9] Wright, J., et al. (2008). "Robust face recognition via sparse representation." IEEE transactions on pattern analysis machine intelligence 31(2): 210-227.
[10] Xia, R., et al. (2014). “Modeling gender information for emotion recognition using denoising autoencoder”. 2014 IEEE International conference on acoustics, speech and signal processing (ICASSP), IEEE.
[11] Krizhevsky, A., et al. (2012). "Imagenet classification with deep convolutional neural networks." Communications of the ACM 25: 1097-1105.
[12] Wang, W., et al. (2014). Face recognition based on deep learning. International Conference on Human Centered Computing, Springer.
[13] Xu, Y., et al. (2016). "Image intelligent detection based on the Gabor wavelet and the neural network." Symmetry 8(11): 130.
[14] Li, F., et al. (2017). Face Recognition Based on Deep Autoencoder Networks with Dropout. 2017 2nd International Conference on Modelling, Simulation and Applied Mathematics Advances in Intelligent Systems Research.
[15] Kamencay, P., et al. (2017). "A new method for face recognition using convolutional neural network.”, Digital Image Processing and Computer Graphics 15: 663-672.

Authorized licensed use limited to: ULAKBIM UASL - Maltepe University. Downloaded on December 30,2022 at 11:16:21 UTC from IEEE Xplore. Restrictions apply.