Levenberg-marquardt backpropagation neural network with techebycheve moments for face detection

Ali Nadhim Razzaq[1], Rozaida Ghazali[2], Nidhal Khdhair El Abbadi[3], Hussein Ali Hussein Al Naffakh[4]
[1,2,4]Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Malaysia
[3]University of Kufa, Education College, Department of Computer Science, Iraq

Article Info

ABSTRACT

Face detection is an intelligent approach used in a variety of applications that identifies human faces in digital images. This work presents a new method which composes of a neural network and Techebycheve transforms for face detection. For feature extraction, Tchebychev transform was applied, in which a discrete Tchebychev transform is given for different sampling patterns and several samples here were performed on color images. A Levenberg-Marquardt backpropagation neural network was applied to the transformed image to find faces in the face detection dataset and FDDB benchmarked database. Model performance was measured based on its accuracy and the best result from the newly proposed method was 98.9%. Simulation results showed that the proposed method handles face detection efficiently.
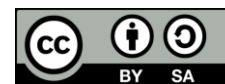
Corresponding Author:

Ali Nadhim Razzaq
Faculty of Computer Science & Information Systems
Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor, Malaysia
Email: gi180013@siswa.uthm.edu.my

1. INTRODUCTION

Face detection is one of the computer vision terms that covers a person's face and informs about his or her area. This technology can acquire facial features and does not account for anything else like bodies, houses, vehicles, and plants. Facial detection is considered to be the location of an object for a special case. In identifying the class of an object, the activity is to find and separate the total area of an element and the size of a particular class of scanned and digitalized images. The models include the supreme torso, walkers, and vehicles. The first step in image applications is to locate and detect the human face. It is important to filter and track facial recognition [1]-[3]. Face detection findings are difficult to obtain for three basic reasons: Firstly, though most faces are arranged with similar facial highlights arranged in a generally similar spatial configuration. There is a possibility that there can be very inflexible and textured contrasts between the faces. These components of volatility are usually due to the necessary contrasts in "facial appearance" between people [4]-[6]. Secondly, facial detection or recognition is difficult to identify on the specific premise but significant basic features, such as glasses or a mustache, may be available or completely absent from the face. Also, these reflections and features, when present can cause other basic features on the face (for example, glasses can reduce glare in eye-catching shades) and have a variable appearance on their own (for example, glasses are available in a wide range of designs) [7]. This adds even more to the scope of authorized facial designs that a comprehensive facial detection framework must address. Third, abnormal imaging conditions may interfere with facial detection in an unrestricted state. Because faces are essentially three-dimensional structures adjusting the shading of the light source can cast or eliminate huge shades from a particular face and is now much larger compared to 2D facial samples.

There are many types of moments for the extraction of highlighted features, Hu (1962) [8] presented the methods of the main invariants, which suggested a strategy for determining the second invariants using algorithmic techniques. He used mathematical techniques to create many invariants. In any case, geometric moments are not obtained from orthogonal functions or asymmetric ability groups and are difficult to noise especially for higher-order moments [9]. In this way, the second invariants of Hu have restricted applications. Many works in writing have introduced ways of dealing with second-construct moment invariants, for example, Zernike moment [10], pseudo-Zernike moment [11], and Legendre moment [12]. However, in these methodologies, the accuracy of the detection deteriorates due to the discrete estimation of the continuous integers [13]. To determine these problems, Mukundan proposed discrete orthogonal Tchebychev moments [14]. Once element extraction has to use a classification method, Neural Networks are one of the most popular machine-based artificial intelligence algorithmic calculations for characterization and classification at present. It was definitely proved after some time that neural network performs better and different algorithms with accuracy and speed. With different variants such as neural networks [15], convolutional neural networks (CNN), as such: [16]-[18], and deep learning and so on [19], [20]. Exploration techniques rely on layouts in the previous to extract specific features and classification, but there are many strategies for identifying faces by identifying the facial components of the neighborhood and command using factual and geometric models for the human face [21]-[23]. First, low-level review settings to share visual highlights using image properties, eg, edges, power, shading, motion, or general measurements. The paper is composed as; section 2 provides a summary of hypothetical and rational basic elements, section 3 presents the proposed technique, section 4 presents the results and discussion of the implementation of the proposed technique for facial detection, and section 5 concludes.

2. BACKGROUND

Different techniques used in the proposed method have been presented in this section that will be explained in the subsequent:

2.1. Discrete tchebychev moments

There are many isolated symmetrical and orthogonal moments, for example, the Tchebychev moments. Using discrete symmetric Tchebychev polynomials as the main work for frame moments eliminates the discrete estimate associated with constant moments. In this work, use Tchebychev minutes to change the geodesic image rather than the image itself. The isolated normalized Tchebychev polynomials with the associated repeat condition are called [13]:

$$p_o(r) = \frac{1}{\sqrt{N}} \tag{1}$$

$$p_1(r) = (2r + 1 - N)\sqrt{\frac{3}{N(N^2-1)}} \tag{2}$$

$$p_k(r) = (a_k r + b_k)p_{k-1}(r) - c_r p_{k-2}(r) \; wit\square k, r = 0, \ldots, N-1 \tag{3}$$

where:

$$a_k = \frac{2}{k}\sqrt{\frac{4k^2-1}{N^2-k^2}} \tag{4}$$

$$b_k = \frac{1-N}{k}\sqrt{\frac{4k^2-1}{N^2-k^2}} \tag{5}$$

$$c_k = \frac{k-1}{k}\sqrt{\frac{2k+1}{2k-3} \cdot \frac{N^2-(k-1)^2}{N^2-k^2}} \tag{6}$$

Now the 2D normalized discrete Tchebychev moments T_{mn} for an N×N generalized geodesic transform $d_x = \tau(x) = [d_x(i,j)]$ of the image $x(i,j)$ are given by:

$$g_x(m,n,d_x) = T_{mn}\{t(x)\} = T_{mn}(d_x) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} p_m(i)\, p_n(j) \cdot d_x(i,j) \tag{7}$$

With $m, n = 0, \ldots, N-1$. $p_m(i)$: The mth normalized discrete Tchebychev polynomials for the variable i.

2.2. Levenberg-marquardt backpropagation neural network

The architecture of the backpropagation uses the design of the reprocessing algorithm [24]. It was re-propagation that was crafted by generalizing Widrow-Hoff's learning rule to multi-layered organizations and non-differential exchange work. Input vectors and associated target vectors are used to prepare an organization until it can approximate a resource. Associate input vectors with explicit target vectors or input vectors as defined in this study. Network with bias, a sigmoid level, and a direct linear level is equipped to approximate any resource with a limited number of gaps and discontinuities. There are two ways to calculate propagation; back and forth. The way forward is to put a network at the forefront, simulation, starting or initializing weight, and prepare the network. The biases and the weights of the network are updated on the contrary. See reference [24] for further subtleties. The Levenberg-Marquardt algorithm aimed to evolve towards the second query training speed without saving the Hesse framework matrix. By the time the presentation work contains an aggregate type of squares as is usual in the preparation of feed organizations, at this point the Hessian network can be approximated as:

$$H = J^T J \tag{8}$$

and the gradient can be computed as:

$$g = J^T e \tag{9}$$

Where J is the Jacobian matrix framework in which the first subordinates of the organization errors in relation to weights and trends, and a vector of network errors is e. The Jacobian matrix can be programmed through a standard reprocessing procedure (see [25]) which is much less complicated than the realization of the Hessian matrix. The Levenberg-Marquardt algorithm uses this approximation to the Hessian matrix in the Newton-like update:

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \tag{10}$$

At the point where the scalar μ is zero, this is just Newton's strategy, using the large Hessian matrix. When μ is huge, it becomes a small slope in advance. Newton's fastest and most accurate technique is at least almost a minimum error, so it's about switching to Newton's strategy as quickly as expected, given the circumstances. Consequently, μ decreases after each successful advancement (decrease in work) and grows just as speculative advancement would resume presentation work. Consequently, the presentation capacity will be systematically reduced with each algorithm cycle [24].

2.3. Face image dataset

The dataset and the face detection benchmark (FDDB), a dataset of facial areas intended to examine the problem of unrestricted facial recognition. In this dataset, the 5171 annotations appear in many 2845 images taken from the Faces in the Wild data index [26], in this work using 5000 human faces for three-stage training, validation, and testing.

2.4. Performance measurement

The evaluation of the proposed technique will be measured by using the evaluation tools (precision, accuracy, recall, and F1_Score). These measurements depend on four initiatives TP, TN, FP, and FN.

3. THE PROPOSED METHOD

This section describes the design and implementation of the face detection system, the overall diagram of the proposed system is shown as Figure 1.

3.1. Tchebychev-neural network

Tchebychev moments and neural network (LMBP) are acceptable means of decoupling images and classification. In this work, a shift that was a collaboration between tchebychev and the neural network (LMBP) aimed to expose unusual exposure compared to using another shift or traditional (CNN). The use of Tchebychev moments changes and transformation give special highlights to the first image. Applying a neurological system (LMBP) will filter out the best of these specific features and provide a dominant performance in facial detection. The tchebychev -neural network measurement (called TNN) refers to the classification of measurements made on windows smaller than the first images, see Figure 2:
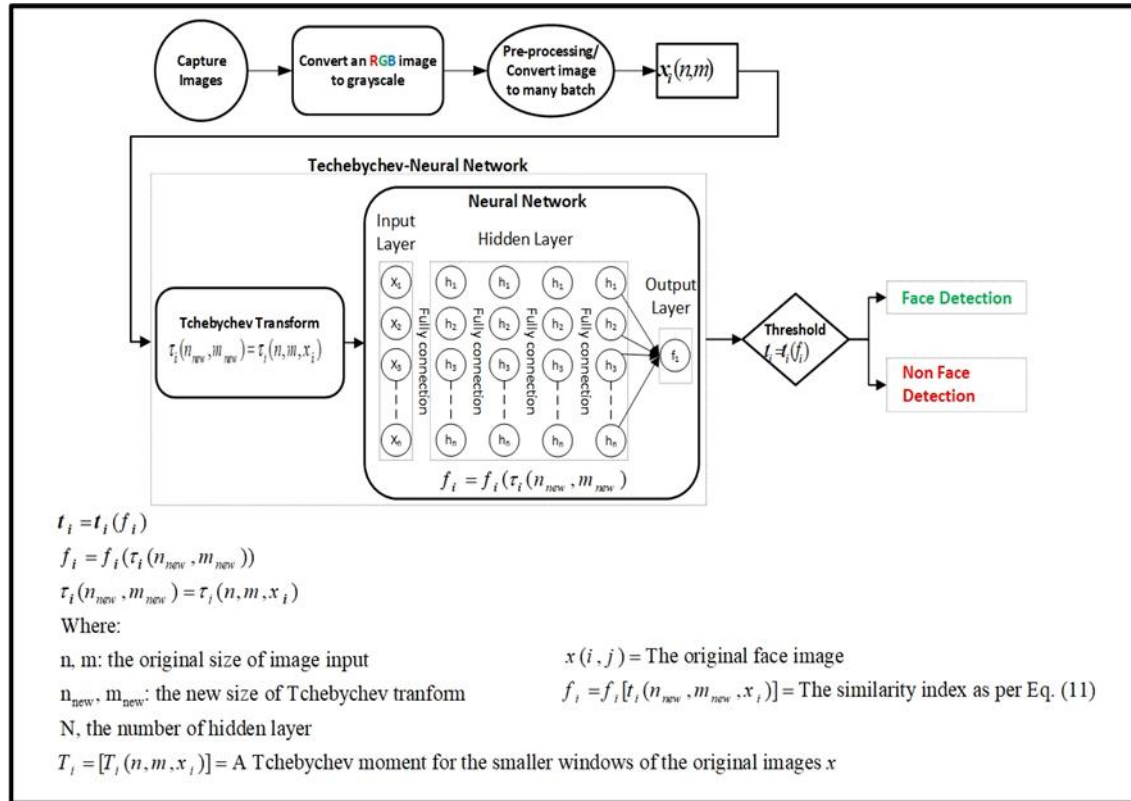
$$\tau_i = \tau_i(f_i) \tag{11}$$

$t_i = t_i(f_i)$

$f_i = f_i(\tau_i(n_{new}, m_{new}))$

$\tau_i(n_{new}, m_{new}) = \tau_i(n, m, x_i)$

Where:

n, m: the original size of image input

n_{new}, m_{new}: the new size of Tchebychev tranform

N, the number of hidden layer

$T_i = [T_i(n, m, x_i)] = $ A Tchebychev moment for the smaller windows of the original images x

$x(i, j) = $ The original face image

$f_i = f_i[t_i(n_{new}, m_{new}, x_i)] = $ The similarity index as per Eq. (11)

Figure 1. Face detection based on tchebychev-neural network

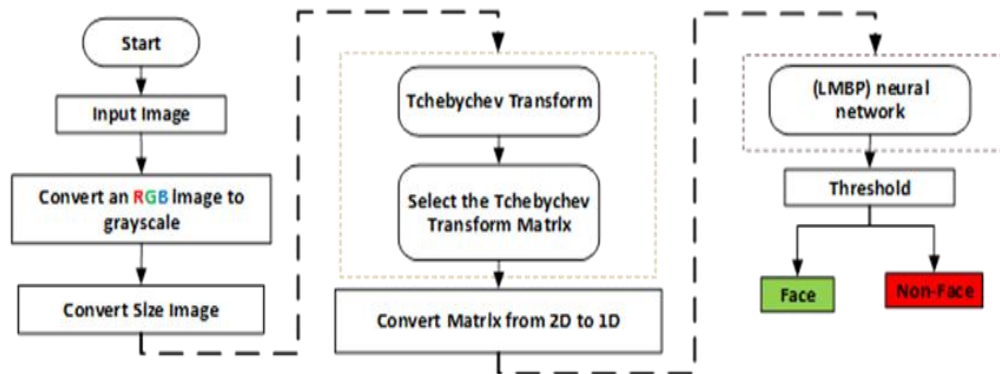
Figure 2. Flowchart (TNN)

3.2. Threshold

A threshold is the last step used to remove the ambiguity of the output from the neural network and to classify the input image as face or non-face. The threshold selected by (trial and error) process, in (12) shows the selected threshold.

$$\tau_i = \begin{cases} face, & if\ \tau_i > t \\ non - face, & else \end{cases} \quad where\quad t: \text{A threshold selection} \tag{12}$$

3.3. Face detection steps

This section describes all the steps for the proposed method:

a. The face area should be identified by using a bounding box in the original face image within the pre-processing phase. First, convert an RGB image to gray-scale and then using multiple sizes of the crop box, start with a sliding window from 16-by-16 pixels that scan for the entire image. After that, the sliding window (crop box) increased its size and did the same thing. The box depends on the original size of the captured image. It should be square of unified size N×N for all sub-images. In this work, we found that the crop box size has no significant effect on the detection except the increasing of execution time, for that the image size of 32×32 is selected.

b. The result from the first step xi (32, 32) passes through a discrete orthogonal Tchebychev transform to get Tchebychev moments Ti (32, 32). The low-pass area of size 12-by-12 was selected from the Tchebychev transform matrix.

c. The resulting dataset from the previous step was separated into three groups for training, validation, and testing. Common ratios used are:
   - 70% train, 15% validation, 15% test
   - 80% train, 10% validation, 10% test
   - 60% train, 20% validation, 20% test

   After the experiment, the best performance result with sizes 70.0%, 15.0%, and 15.0% is 3500, 750, and 750 human faces respectively.

d. The last step is to apply a neural network (LMBP) with many input features (144) get from Tchebychev moments with the best number of hidden layers (41), the image classified according to the threshold.

e. In this proposal the learning rate used is 0.6, activation function was (symmetric saturating linear transfer function), and the number of epochs is 17 with 5 number of runs.

4. RESULTS and DISCUSSION

This section shows the tests on how to choose the optimal variables used in the proposal and how they affect the detection accuracy, many tests were introduced to improve the accuracy of the proposed architecture. The neural network has three important parameters that highly affect the detection accuracy. These parameters are learning rate, number of hidden layers, and activation function. Various values of learning rate tested to find the best accuracy as in Figure 3. We found that the learning rate with a value of 0.6 is the best. Also, the effect of the number of hidden layers on the accuracy is tested. In this test many numbers of hidden layers are used and measuring the accuracy, the result of this test shows in Figure 4. It is concluded that the best number of hidden layers are forty-one layers.
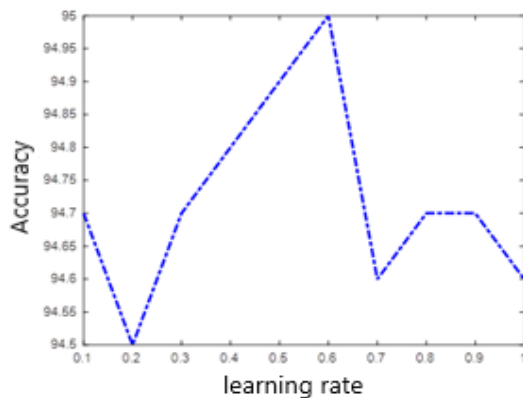

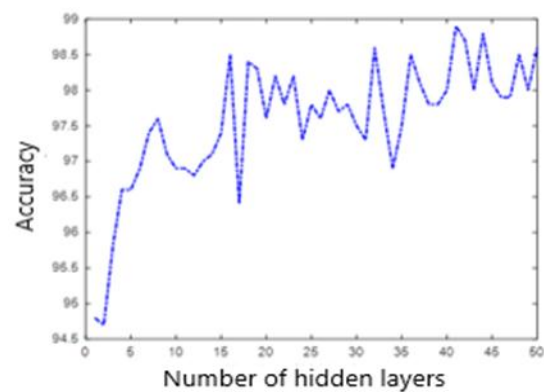Figure 3. Test the accuracy vs. learning rate

Figure 4. The effect of number of hidden layers on the accuracy

The third parameter is the activation function. Many activation functions can be used in this proposal and each one has a specific effect on the accuracy of the face detection. The activation functions can be divided into two groups according to the output value range, so each group is tested alone. The first group has an output value range from (-1 to 1), while the second group output value ranges from zero to one. So these activation function groups are tested to find which one gives the best accuracy as shown in Figure 5 and Figure 6. From these figures, it is found that the activation function symmetric saturating linear transfer function (Satlins) is the best.
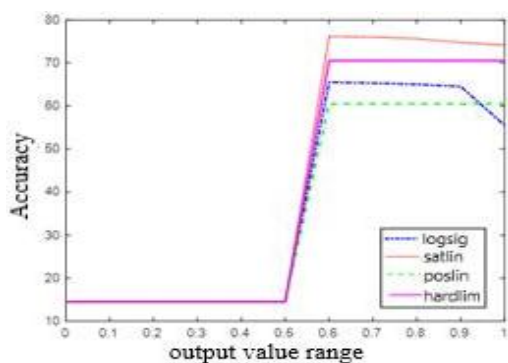
Figure 5. Effect of various activation functions (output range (0 to 1)) on the accuracy
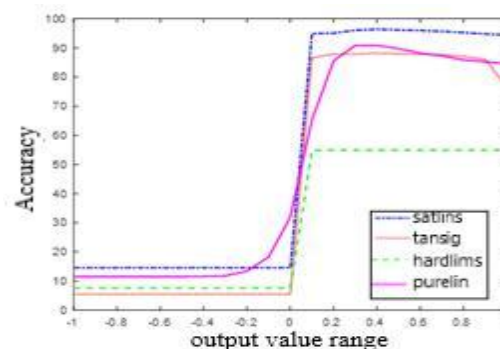

Figure 6. Effect of various activation functions (output range (-1 to 1)) on the accuracy

The image in this proposal is divided into many blocks. Various block sizes were checked, the accuracy for each block size is measured. The best accuracy is achieved when the block size is (12x12) as shown in Figure 7. The performance of the suggested method is evaluated by measuring the root mean square error when using the Tchebychev moments, with block size (12×12), activation function (satlins), and threshold (0.4). Two types of datasets were used for this test (random and specific dataset). The minimum RMSE for the specific dataset is (0.1104) and for the random dataset is (0.1326) when using the number of the hidden layer equal to forty-one as in Figure 8. It is concluded that the use of the specific dataset gives a better result than using the random dataset. The face detection accuracy of the proposed method was up to 98%. Figure 9 shows samples of detected face images. While Figure 10 shows samples of face images where the system fails to detect. Figure 11 evaluated the proposed method via specificity, accuracy, precision, recall, and F1_Score.
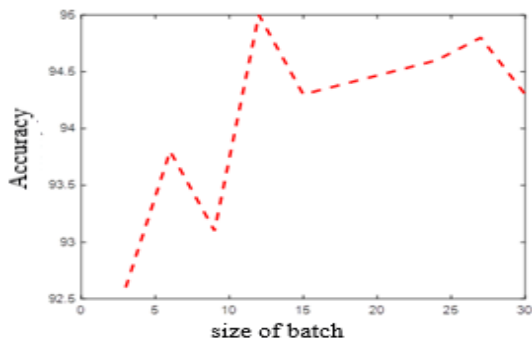

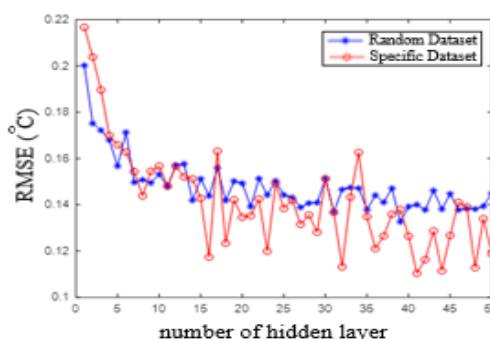Figure 7. The variation of accuracy according to the block size

Figure 8. Determine the RMSE for the specific and random dataset

Figure 9. Sample detected faces by proposal method

Figure 10. Sample of faces images where the proposed method fails to detect
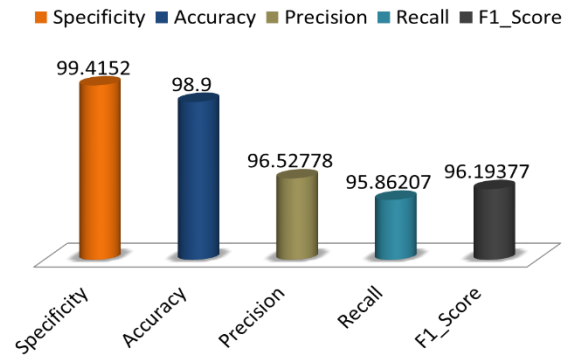
Figure 11. Proposed method evaluated based on five measures

Finally, we summarized the training process such that the input features were (144) get from Tchebychev moments, the best number of hidden layers (41), the image classified according to the threshold equal to (0.4), The learning rate used is 0.6, the number of epochs was seventeen. At the end of the training.

5.   CONCLUSION

The proposed face detection method based on Tchebychev transform using the LMBP neural network has been introduced and evaluated. The use of the LMBP neural network improved the face detection performance, as shown from comparing with other methods. The failure in the detection of some of the faces may be due to facial expression, rotation of faces with an angle more than 45 degrees, or maybe affected by glasses, and occlusion. The performance in general was promised and may be enhanced when adding some other features to the current suggested algorithm.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Z. Wei, J. Lidong, L. Yueyuan, and Y. Sen, "Design of Multi-task Cascade Convolutional Face Tracking system," *Microcontrollers Embed. Syst.*, p. 7, 2020.
[2]   A. Kumar, A. Kaur, and M. Kumar, "Face detection techniques: a review," *Artificial Intelligence Review*, vol. 52, no. 2, pp. 927-948, 2019, doi: 10.1007/s10462-018-9650-2
[3]   B. Rehman, W. H. Ong, A. C. H. Tan, and T. D. Ngo, "Face detection and tracking using hybrid margin-based ROI techniques," *Visual Computer*, vol. 36, no. 3, pp. 633-647, 2020, doi: 10.1007/s00371-019-01649-y
[4]   O. N. A. Al-Allaf, "Review of face detection systems based artificial neural networks algorithms," *The International Journal of Multimedia & Its Applications*, arXiv Prepr. arXiv1404.1292, pp. 1-16, 2014, doi: 10.5121/ijma.2013.6101.
[5]   B. Chhetri, I. Sharma, S. Thapa, C. Pradhan, and S. Kumari, "A Review on Face Detection and Recognization Techniques," *International Journal of Computer Science And Technology*, vol. 9, no. 2, pp. 33-36, 2017.
[6]   Y. Wu, and Q. Ji, "Facial landmark detection: A literature survey," *International Journal of Computer Vision*, vol. 127, no. 2, pp. 115-142, 2019.
[7]   Z. Zhao, P. Zheng, S. Xu, and X. Wu, "Object Detection With Deep Learning: A Review," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3212-3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.
[8]   Ming-Kuei Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179-187, February 1962, doi: 10.1109/TIT.1962.1057692.
[9]   Bingcheng Li, "High-order moment computation of gray-level images," *IEEE Transactions on Image Processing*, vol. 4, no. 4, pp. 502-505, April 1995, doi: 10.1109/83.370680.
[10]   C.-W. Chong, P. Raveendran, and R. Mukundan, "Translation invariants of Zernike moments," *Pattern Recognition*, vol. 36, no. 8, pp. 1765-1773, 2003, doi: 10.1016/S0031-3203(02)00353-9.
[11]   C.-W. Chong, P. Raveendran, and R. Mukundan, "The scale invariants of pseudo-Zernike moments," *Pattern Analysis & Applications*, vol. 6, no. 3, pp. 176-184, 2003.

[12] C.-W. Chong, P. Raveendran, and R. Mukundan, "Translation and scale invariants of Legendre moments," *Pattern Recognition*, vol. 37, no. 1, pp. 119-129, 2004, doi: 10.1016/j.patcog.2003.06.003.

[13] R. Mukundan, "Some computational aspects of discrete orthonormal moments," *IEEE Transactions on Image Processing*, vol. 13, no. 8, pp. 1055-1059, Aug. 2004, doi: 10.1109/TIP.2004.828430.

[14] R. Mukundan, S. H. Ong, P. A. Lee, "Image analysis by Tchebichef moments," *IEEE Transactions on Image Processing*, vol. 10, no. 9, pp. 1357-1364, Sept. 2001, doi: 10.1109/83.941859.

[15] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A Comprehensive Survey on Graph Neural Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4-24, Jan. 2021, doi: 10.1109/TNNLS.2020.2978386.

[16] G. Guo, H. Wang, Y. Yan, J. Zheng, and B. Li, "A fast face detection method via convolutional neural network," *Neurocomputing*, vol. 395, pp. 128-137, 2020, doi: 10.1016/j.neucom.2018.02.110.

[17] R. Qi, R. -S. Jia, Q. -C. Mao, H. -M. Sun, and L. -Q. Zuo, "Face Detection Method Based on Cascaded Convolutional Networks," *IEEE Access*, vol. 7, pp. 110740-110748, 2019, doi: 10.1109/ACCESS.2019.2934563.

[18] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artificial Intelligence Review*, vol. 53, pp. 5455-5516, 2020.

[19] Y. S. El-Din, M. N. Moustafa, and H. Mahdi, "Deep convolutional neural networks for face and iris presentation attack detection: survey and case study," *IET Biometrics*, vol. 9, no. 5, pp. 179-193, 2020.

[20] A. Shrestha, and A. Mahmood, "Review of Deep Learning Algorithms and Architectures," *IEEE Access*, vol. 7, pp. 53040-53065, 2019, doi: 10.1109/ACCESS.2019.2912200.

[21] S.-H. Jeng, H. Y. M. Liao, C. C. Han, M. Y. Chern, and Y. T. Liu, "Facial feature detection using geometrical face model: an efficient approach," *Pattern Recognition*, vol. 31, no. 3, pp. 273-282, 1998, doi: 10.1016/S0031-3203(97)00048-4.

[22] Z. Arsalane, M. Aicha, and M. El-Rhazi, "Survey on the approaches based geometric information for 3D face landmarks detection," *IET Image Process*, vol. 13, no. 8, pp. 1225-1231, 2019, doi: 10.1049/iet-ipr.2018.6117.

[23] A. N. Razzaq, Z. Hussain, and H. R. Mohammed, "Structural Geodesic-Tchebychev Transform: An image similarity measure for face recognition," *Journal of Computer Sciences*, vol. 12, no. 9, pp. 464-470, 2016, doi: 10.3844/jcssp.2016.464.470.

[24] E. H. Aria, J. Amini, and M. R. Saradjian, "Back propagation neural network for classification of IRS-1D satellite images," *Joint Workshop of High Resolution Mapping from Space*, Tehran University, Iran, 2003.

[25] M. T. Hagan, and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989-993, Nov. 1994, doi: 10.1109/72.329697.

[26] V. Jain, and E. Learned-Miller, "Fddb: A benchmark for face detection in unconstrained settings," 2010.

BIOGRAPHIES OF AUTHORS

Ali Nadhim Razzq received the B.Sc. in computer science from University of Kufa/ Faculty of Education in (2010)- Iraq; He's doing H.Dip. Degree in Higher Diploma of Mathematics, Faculty of Mathematics and Computer Science, University of Kufa, Iraq in (2013). He's doing M.Sc. Degree in Computer Science Department, Faculty of Mathematics and Computer Science, University of Kufa, Iraq in (2017). Ph.D. candidate in Information Technology at Universiti Tun Hussein Onn Malaysia (UTHM). He's interested in the following field (Information security, image processing, and artificial intelligence).

Rozaida Ghazali is currently a Professor at the Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM). She graduated with a Ph.D. degree in Higher Order Neural Networks from the School of Computing and Mathematical Sciences at Liverpool John Moores University, United Kingdom in 2007. Earlier, in 2003 she completed her M.Sc. degree in Computer Science from Universiti Teknologi Malaysia (UTM). She received her B.Sc. (Hons) degree in Computer Science from Universiti Sains Malaysia (USM) in 1997. In 2001, Rozaida joined the academic staff in UTHM. Her research area includes neural networks, swarm intelligence, optimization, data mining, and time series prediction. She has successfully supervised a number of PhD and master students and published more than 150 articles in various international journals and conference proceedings. She acts as a reviewer for various journals and conferences. She has also served as a conference chair, and as a technical committee for numerous international conferences.

Prof. Dr. Nidhal El Abbadi, received B.Sc. in Chemical Engineering, B.Sc. in computer science, M.Sc. and Ph.D. in computer science, worked in industry and many universities, reviewer for a number of international journals, has many published papers and six published books (Programming with Pascal, C++ from beginning to OOP, Data structures in C++, Principles of Computer Graphics, effect of modern technology on talent, and chapter in Handbook of Research on Threat Detection and Countermeasures in Network Security book), his research interests are in image processing, security, and steganography, He's Professor in Computer Science in the University of Kufa – Najaf, IRAQ.

Hussein Ali Hussein received the B.Sc. in computer science from University of Kufa/ Faculty of Education in (2014)-Iraq; He is doing M.Sc. (Computer Science) in Software Division from University of Babylon/College of Information Technology in 2017. Ph.D. candidate in Information Technology at Universiti Tun Hussein Onn Malaysia (UTHM). He is interested in the following field (Steganography, Cryptology, Information security and Information system, image processing, artificial intelligence).