❒   380

Predicting death and confirmed cases of coronavirus

Farqad Hamid Abdulraheem, Moatasem Yaseen Al-Ridha, Raid Rafi Omar Al-Nima
Engineering Technical Engineering College, Northern Technical University Mosul, Iraq

Article Info	ABSTRACT
	At the end of 2019, a new virus called coronavirus has globally spread causing severe effections. In this paper, an artificial intelligence (AI) method is proposed to predict numbers of death and confirmed coronavirus cases. Efficient machine learning (ML) network named the byesian regularization backpropagation (BRB) is employed. It can estimates numbers of death and confirmed cases from applied population density and date. So, the BRB uses the population density, month and day as inputs, and predicts the new cases per million and new deaths per million as outputs. The network was trained and assessed by using a daily coronavirus recorded dataset known as the our world in data (OWID). The considered dates here are from the 31st of December 2019 to the 13th of October 2020. Furthermore, recorded information from countries over all world are employed. The obtained results provided a good promising performance with a testing mean absolute error (MAE) equal to 0.0218.

Corresponding Author:

Moatasem Yaseen Al-Ridha
Engineering Technical Engineering College, Northern Technical University, Mosul, Iraq
Email: Moatasem@ntu.edu.iq

1. INTRODUCTION

Coronavirus has firstly started in China, more specifically in Wuhan. It is being a pandemic or global epidemic. The international committee on classifying the viruses named it as coronavirus disease 2019 (COVID-19) [1]-[3]. Its effectiveness may range between middle and severe and in some cases it can be fatal. The symptoms of coronavirus are usually fever, cough and breath difficulty. Some human infection is experiencing non-respiratory symptoms. Other people have almost no symptoms. Moreover, the new coronavirus is the seventh virus that characterized as being the main reason for respiratory infection [4], [5].

Coronavirus is also characterized for pulmonary infections among people. The efforts of international health and governments have focused on speeding the diagnosis. In addition of searching for treats, which is able to face the severe effects of the disease. However, efficient treats still do not exist [6]. Since coronavirus is appeared on the 12th of December 2019, it caused 2,794 laboratory-confirmed infections which contain 80 death cases by the 26th of January 2020. It killed more than 1800 and infected over 70,000 persons through 50 days of existing the disease [7].

Coronavirus protein works infection of targeted cells that hold the specific receptor. The protein of SARS-CoV type I trans membrane glycoprotein containing putative S1 and S2 ranges binds to angiotensin-converting.enzyme 2 (ACE2), a functional cell receptor, with high affinity [8]. It seems that it is necessary to establish a method for estimating the number of deaths and confirmed cases. In this work, an efficient network called the byesian regularization backpropagation (BRB) is utilized for this issue. The rest sections are organized as: section 2 reviews the previous studies, section 3 explains the theoretical concept of the BRB network, section 4 provides the implementation results and section 5 declares the conclusion.

2.   LITERATURE REVIEW

Many studies that consider various related subjects especially regarding the machine learning have been provided as in [9]-[17]. Other work that are related to COVID-19 can be reviewed as:

Mousa *et al.* [18] summarized the interfering characteristics of Hydroxychloroquine (HCQ) for strong human immunity against the growing of COVID-19. The authors supposed that the therapeutic efficiency of the HCQ can be enhanced by the concurrent management of zinc sulfate. Then, it can prevent spreading the virus inside biological cells. Riou and Althaus analyzed the 2019 novel coronavirus (2019-nCoV). It was suggested that the early pattern of person-to-person transmission of 2019-nCoV memorizes the virus of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) in 2002. It was also approved that the international coordination will be very important to prevent the spreading of 2019-nCoV [19].

Lu *et al.* [20] investigated the 2019-nCoV by analyzing its genome sequences. Samples from nine patients were collected, eight of them had visited markets in Wuhan. Sanger sequencing was used to get the full-length of genomes. The terminal regions were determined by quickly amplifying the ends of deoxyribonucleic acid (DNA). Chan *et al.* [21] worked on genetic sequences of the 2019-nCoV. Information were collected for the clinical, laboratory, epidemiological microbiological, radiological and results of five patients from a same family group. They infected by unexplained pneumonia after coming back to Shenzhen in China after visiting Wuhan. Furthermore, a member from additional family who did not travel to Wuhan was considered too.

Gorbalenya *et al.* [2] enhanced the understanding of virus-host interactions in a changing environment and improved the readiness for outbreaks in future. The authors presented an evaluation of the 2019-nCoV and explained the basis of renaming this virus to severe acute respiratory syndrome-associated coronavirus 2 (SARS-CoV-2). Chan *et al.* [22] provided base results for initializing additional coronavirus studies on the pathogenesis. In addition of improving the design of diagnostic, antiviral and vaccination approaches for this emerging infection.

Jin *et al.* [23] organized guidelines for the rapid advice of the diagnosis and treatment of 2019-nCoV. The first-hand management information were added for the Zhongnan hospital of Wuhan University. This work contains guidelines for the ways, epidemiological features, disease, checking and population prevention, control, treatment, and diagnosis. A whole operation for a good treatment was provided for the infection of 2019-nCoV. The guidelines can possibly be used by the first frontline people in hospitals such as managers, doctors and nurses. From the literature it can be observed that various studies were established for coronavirus. Furthermore, working on any version of this virus in any direction can be considered as a hot topic. As mentioned, this study is proposed for predicting death and confirmed cases of the COVID-19.

3.   THEORETICAL CONCEPT OF PROPOSED NETWORK

First of all, the proposed BRB is based on Backpropagation that is trained according to Byesian regularization. The key idea is that Bayesian regularization has the capability to minimize the linear combination between squared errors and weights [24]. The general block diagram of Byesian inference is given in Figure 1.

This figure demonstrates the scientific operation of Byesian inference with collected and modeled data. In particular, this operation can be employed for pattern interpolation, learning, classification. Bayes rule is basically applied for the only two double-framed blocks as they include the inference. Bayes does not tell you how to invent models, for example. The first block is for the goal of inferring where the model parameters may be given. Bayes can be utilized to explore the most appropriate parameter values [25]. The weights in a Bayesian network are randomly given at first. Then, the density function can be adjusted for the weights in the train based on the Bayes rule [26]:

$$P(\mathbf{w}|D, \alpha, \beta, M) = \frac{P(D|\mathbf{w}, \beta, M)\, P(\mathbf{w}|\alpha, M)}{P(D|\alpha, \beta, M)} \tag{1}$$

Where $P(\mathbf{w}|D, \alpha, \beta, M)$ is the density function, D is the input data set, α and β are objective function parameters, M is the employed neural network model, and $\mathbf{w}$ is the network weights vector. In (1) is explained in [25] as:

$$Posterior = \frac{Likelyhood \times Prior}{Evidence} \tag{2}$$

It seems that *Posterior* is equivalent to $P(\mathbf{w}|D, \alpha, \beta, M)$, *Likelyhood* is equivalent to $P(D|\mathbf{w}, \beta, M)$, *Prior* is equivalent to $P(\mathbf{w}|\alpha, M)$ and *Evidence* is equivalent to $P(D|\alpha, \beta, M)$. Moreover, information about the Backpropagation neural network are well known. It basically consists of three training stages: feedforward

inputs to outputs, backpropagation the calculated errors and updating the biases and weights. Whereas, the testing stage only involves the feedforward inputs to outputs after applying the obtained biases and weights from the train step. Such Backpropagation information are exploited in [27]-[33].

The specifications of our BRB network are as follows: it has three inputs (population density, month and day), one hidden layer with two hidden nodes, transfer functions of type binary sigmoid, two outputs (new cases per million and new deaths per million), recorded information from countries over all world, starting date of 31/12/2019 and ending date of 13/10/2020. The proposed architecture of the BRB neural network as shown in Figure 2. This network exploits in (1, 2) with the backpropagation algorithm, as mentioned. For more information, an example with a case study of similar network type can be discovered in [30].
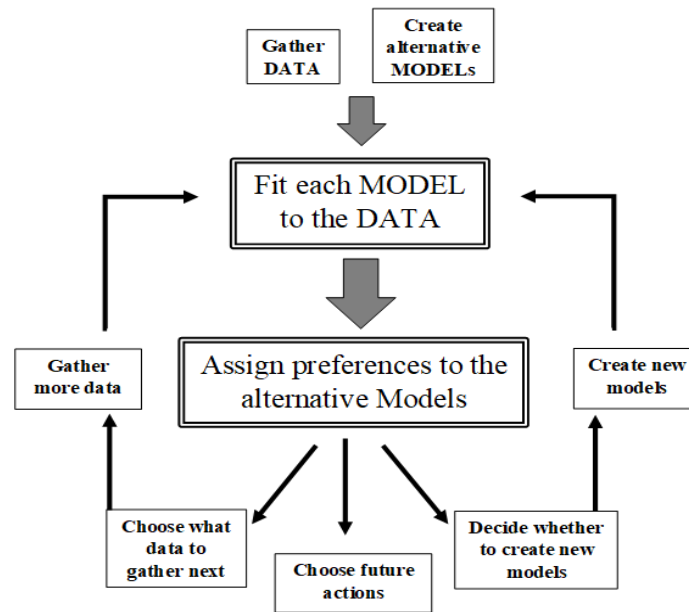

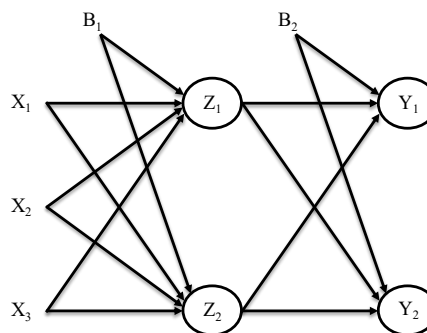Figure 1. The general block diagram of Byesian inference as shown in [25]

Figure 2. The proposed architecture of the BRB neural network

4. IMPLEMENTATION RESULTS

Coronavirus database is acquired from our world in data (OWID). It has different information from all world. Examples of these information are: positive rate, tests per case, new tests smoothed per thousand, new tests smoothed, new tests per thousand, total tests per thousand, total tests, new tests, new deaths smoothed per million, new deaths per million, total deaths per million, new cases smoothed per million, new cases per million, total cases per million, new deaths smoothed, new deaths, total deaths, new cases smoothed, new cases, total cases and date [31].

As mentioned, recorded data for countries from all world are considered in this study. The concentrated dates are for the period from 31/12/2019 to 13/10/2020. Five types of information are employed, these are the population density, month, day, new cases per million and new deaths per million. Moreover, the

proposed BRB has been trained for the odd-ordered information and tested for the even-ordered information. For avoiding the BRB over-loads, the values of inputs and outputs are decreased/normalized. Total of 5945 pairs of inputs and targets are used in the training stage. Similarly, total of 5945 pairs of inputs and outputs are used in the testing stage. Figures 3 and 4 shows training curves of the suggested BRB network
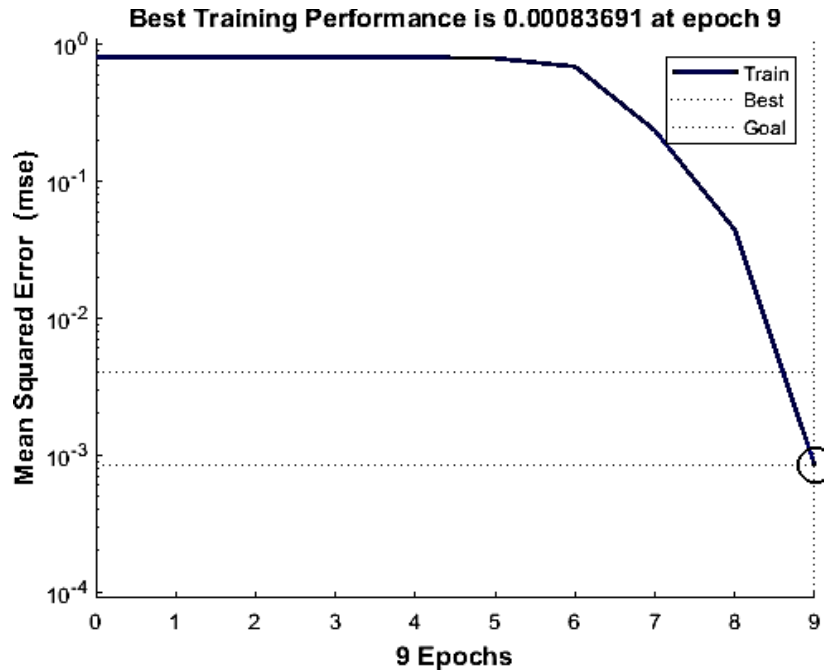

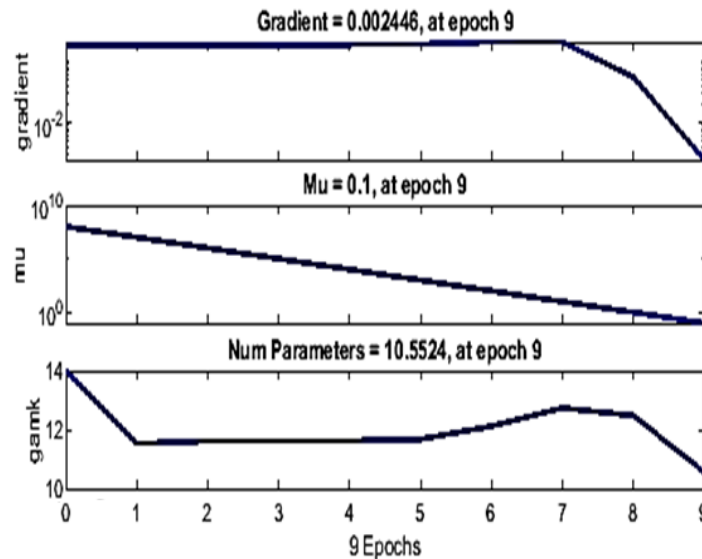Figure 3. Main training curve for the suggested BRB network

Figure 4. Training curves for different BRB network parameters

It can be noted from Figure 3 that the training curve is quickly degraded to such small mean square error (MSE) value of approximately equal to 8×10^{-4}. Furthermore, training curves for different BRB network factors in Figure 4 demonstrates how they could attained acceptable values at the end of learning. That is, the factors of gradient, mu and number of parameters are observed during the training stage. They almost show successful progressing. They also almost achieve acceptable values of gradient=2.446×10^{-3}, mu=0.1 and number of parameters=10.55 at the end of learning. As a result of training, it can be yield that it is quick and successful for the employed data values of the strong pandemic of coronavirus. It is worth mentioning that the

suggested BRB network is very fast as it required 9 epochs to achieve its goal, this can be justified by the training facilities of this network as well as the form of training information. As far as the training information are appropriately distributed, the BRB has the ability to reach its goal with a significant low number of epochs.

The implementation can be explained and justified according to the testing stage. A testing error of type the mean absolute error (MAE) is considered between the outputs and targets, it reports a promising MAE value equal to 2×10^{-2}. This error value justifies how the obtained outputs are close to the desired targets, which explains how the suggested BRB network has the ability to successfully be implemented and employed. Especially, recorded data from countries over all world are exploited in this study and total of 5945 pairs of inputs and outputs are evaluated, this can add more significant performance to the suggested approach. So, it can be illustrated that the obtained MAE in the testing stage is such interested and promised.

5. CONCLUSION

This paper provided an approach for predicting death and confirmed cases of a current pandemic named coronavirus. The approach was based on the BRB technique. It is a type of neural network that has the ability to train and test groups of data according to Backpropagation algorithm and Byesian regularization. It has three inputs of population density, month, and day and two outputs of new cases per million and new deaths per million. In addition to one hidden layer of nodes. A big number of data was used for recorded values over all world (5945×2 pairs of inputs and targets). The results yielded promising and interesting performance ($MAE = 2 \times 10^{-2}$) which can be exploited for making preparations to save vulnerable incidents.

REFERENCES

[1] J. Cui, F. Li, and Z. L. Shi, "Origin and evolution of pathogenic coronaviruses," *Nature Reviews Microbiology*, vol. 17, no. 3. pp. 181–192, 2019. doi: 10.1038/s41579-018-0118-9.

[2] A. E. Gorbalenya *et al*., "Severe acute respiratory syndrome-related coronavirus: The species and its viruses – a statement of the Coronavirus Study Group," *BioRxiv*, 2020. doi: 10.1101/2020.02.07.937862.

[3] WHO Guidance Note, "Laboratory testing for coronavirus disease (COVID-19) in suspected human cases," *World Health Organization WHO*, 2020, [Online]. Available: https://apps.who.int/iris/handle/10665/334254.

[4] C. Huang *et al*., "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *The Lancet*, vol. 395, no. 10223, pp. 497-506, 2020, doi: 10.1016/S0140-6736(20)30183-5.

[5] N. Zhu *et al*., "A Novel Coronavirus from Patients with Pneumonia in China, 2019," *New England Journal of Medicine*, vol. 382, no. 8, pp. 727-733, 2020, doi: 10.1056/nejmoa2001017.

[6] C. A. Devaux, J. M. Rolain, P. Colson, and D. Raoult, "New insights on the antiviral effects of chloroquine against coronavirus: what to expect for COVID-19?," *International Journal of Antimicrobial Agents*, vol. 55, no. 5, 2020, doi: 10.1016/j.ijantimicag.2020.105938.

[7] P. Zhou *et al*., "A pneumonia outbreak associated with a new coronavirus of probable bat origin," *Nature*, vol. 579, no. 7798, pp. 270-273, 2020, doi: 10.1038/s41586-020-2012-7.

[8] Y. He, J. Li, W. Li, S. Lustigman, M. Farzan, and S. Jiang, "Cross-Neutralization of Human and Palm Civet Severe Acute Respiratory Syndrome Coronaviruses by Antibodies Targeting the Receptor-Binding Domain of Spike Protein," *The Journal of Immunology*, vol. 176, no. 10, pp. 6085-6092, 2006, doi: 10.4049/jimmunol.176.10.6085.

[9] A. Andreas, C. X. Mavromoustakis, G. Mastorakis, S. Mumtaz, J. M. Batalla and E. Pallis, "Modified Machine Learning Techique for Curve Fitting on Regression Models for COVID-19 projections," *IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, 2020, pp. 1-6, doi: 10.1109/CAMAD50429.2020.9209264.

[10] S. Kamley, S. Jaloree, and R. S. Thakur, "Performance forecasting of share market using machine learning techniques: A review," *International Journal of Electrical and Computer Engineering*, vol. 6, no. 6. pp. 3196-3204, 2016. doi: 10.11591/ijece.v6i6.13323.

[11] J. A. Jupin, T. Sutikno, M. A. Ismail, M. S. Mohamad, S. Kasim, and D. Stiawan, "Review of the machine learning methods in the classification of phishing attack," *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 4, pp. 1545-1555, 2019, doi: 10.11591/eei.v8i4.1344.

[12] N. M. Nawi, M. Makhtar, Z. A. Afip, M. Z. Salikon, and K. A. M. Sukri, "Parkinson disease classification: A comparative analysis on classification techniques," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, no. 3, pp. 1351-1358, 2020, doi: 10.11591/ijeecs.v18.i3.pp1351-1358.

[13] S. Verma, G. T. Thampi, and M. Rao, "ANN based method for improving gold price forecasting accuracy through modified gradient descent methods," *IAES International Journal of Artificial Intelligence*, vol. 9, no. 1, pp. 46-57, 2020, doi: 10.11591/ijai.v9.i1.pp46-57.

[14] A. J. Mohammed, M. H. Arif, and A. A. Ali, "A multilayer perceptron artificial neural network approach for improving the accuracy of intrusion detection systems," *IAES International Journal of Artificial Intelligence*, vol. 9, no. 4, pp. 609-615, 2020, doi: 10.11591/ijai.v9.i4.pp609-615.

[15] Z. Zainuddin, E. A. P. Akhir, and M. H. Hasan, "Predicting machine failure using recurrent neural network-gated recurrent unit (RNN-GRU) through time series data," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 2, pp. 870-878, 2021, doi: 10.11591/eei.v10i2.2036.

[16] Z. Hamadouche, M. Khiat, and M. A. Iqbal, "Intelligent voltage regulator for distributed generation-based network," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 1, pp. 98-109, 2021, doi: 10.11591/ijeecs.v23.i1.pp98-109.

[17] R. M. A. Ujjan, Z. Pervez, K. Dahal, A. K. Bashir, R. Mumtaz, and J. González, "Towards sFlow and adaptive polling sampling for deep learning based DDoS detection in SDN," *Future Generation Computer Systems*, vol. 111, pp. 763-779, 2020, doi: 10.1016/j.future.2019.10.015.

[18] Y. J. Mousa, M. B. Mahmood, F. A. Isihak, and A. A. Mohammed, "Are promising mechanisms of hydroxychloroquine abolish COVID-19 activity? A review study," *Iraqi Journal of Veterinary Sciences*, vol. 34, no. 2, pp. 345-349, 2020, doi:

10.33899/ijvs.2020.127049.1449.

[19] J. Riou and C. L. Althaus, "Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020," *Eurosurveillance*, vol. 25, no. 4, pp. 1-5, 2020, doi: 10.2807/1560-7917.ES.2020.25.4.2000058.

[20] R. Lu *et al.*, "Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding," *The Lancet*, vol. 395, no. 10224, pp. 565-574, 2020, doi: 10.1016/S0140-6736(20)30251-8.

[21] J. F. W. Chan *et al.*, "A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster," *The Lancet*, vol. 395, no. 10223, pp. 514-523, 2020, doi: 10.1016/S0140-6736(20)30154-9.

[22] J. F. W. Chan *et al.*, "Correction to: Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan," *Emerging Microbes and Infections*, vol. 9, no. 1, pp. 221-236, vol. 9, no. 1. Pp. 221-236, 2020. doi: 10.1080/22221751.2020.1737364.

[23] Y. H. Jin and J. Cai, "A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version)," *Military Medical Research*, vol. 7, no. 1. 2020. doi: 10.1186/s40779-020-0233-6.

[24] H. Demuth and M. Beale, "Neural Network Toolbox - For Use with MATLAB," pp. 1-847, 2002.

[25] D. J. C. MacKay, "Bayesian Interpolation," *Neural Computation*, vol. 4, no. 3, pp. 315-447, 1992, doi: 10.1162/neco.1992.4.3.415.

[26] F. Dan Foresee and M. T. Hagan, "Gauss-Newton approximation to Bayesian learning," *Proceedings of International Conference on Neural Networks (ICNN'97)*, 1997, pp. 1930-1935 vol. 3, doi: 10.1109/ICNN.1997.614194.

[27] A. S. Anaz, R. R. O. Al-nima and M. Y. Al-Ridha "Multi-Encryptions System Based on Autoencoder Deep Learning Network," *Solid State Technology*, vol. 63, no. 6, pp. 3632-3645, 2020.

[28] R. R. O. Al-Nima, M. Y. Al-Ridha, and F. H. Abdulraheem, "Regenerating face images from multi-spectral palm images using multiple fusion methods," *Telecommunication Computing Electronics and Control TELKOMNIKA* vol. 17, no. 6, 2019, doi: 10.12928/TELKOMNIKA.v17i6.12857.

[29] M. A. M. Abdullah, R. R. Al-Nima, S. S. Dlay, W. L. Woo, and J. A. Chambers, "Cross-Spectral Iris Matching for Surveillance Applications," *Advanced Sciences and Technologies for Security Applications*, pp. 105-125, 2018. doi: 10.1007/978-3-319-68533-5_5.

[30] R. R. Al-nima, "Design a Biometric Identification System Based on the Fusion of Hand Geometry and Backhand Patterns," *IRAQI Journal Of Statistical Sciences,* vol. 10, no. 17, pp. 169-180, 2010.

[31] R. R. Al-nima, L. Muhanad, and S. Q. Hassan, "Data encryption Using Backpropagation Neural Network," *IRAQI Academic Scientific Journals*, vol. 15, no. 2, pp. 112-117, 2009.

[32] M. Khalil, M. Majeed, and R. Omar, "Personal Identification with Iris Patterns," *AL-Rafidain Journal of Computer Sciences and Mathematics*, vol. 6, no. 1, pp. 13-26, 2009, doi: 10.33899/csmj.2009.163762.

[33] M. T. Al-Kaltakchi, R. R. Omar, H. N. Abdullah, T. Han, and J. A. Chambers, "Finger Texture Verification Systems Based on multiple spectrum Lighting Sensors with Four Fusion Levels," *Iraqi Journal of Information & Communications Technology*, vol. 1, no. 3, 2019, doi: 10.31987/ijict.1.3.28.

BIOGRAPHIES OF AUTHORS

Farqad Hamid Abdulraheem 🆔 🔗 SC Ⓟ received the BSc degrees in Computer science in 1994, he accomplished his master's degree in the college of computer and arithmetic science in university of mosul at 2004. In 2017, he accomplished his PhD in Babylon University. His research interests are in the fields of information security, image processing and soft comuting technique. He can be contacted at email: frqhm.70@ntu.edu.iq.

Moatasem Yaseen Al-Ridha 🆔 🔗 SC Ⓟ received the BSc degrees in Technical Computer Engineering in 2006. In 2013, he accomplished his master's degree in the School of Electrical and Computer Engineering at Southern Illinois University, USA. His research interests are in the fields of image processing, signal processing, pattern recognition, security, and artificial intelligence. He can be contacted at email: Moatasem@ntu.edu.iq.

Raid Rafi Omar Al-Nima 🆔 🔗 SC Ⓟ received the BSc and MSc degrees in Technical Computer Engineering in 2000 and 2006, respectively. During 2006, he worked as an Assistant Lecturer in the Technical College of Mosul, Iraq. In 2011, he obtained the Lecturer scientific title in the same college. In 2017, he accomplished his PhD in the School of Electrical and Electronic Engineering at Newcastle University, UK. In 2020, he achieved the title of Assistant Professor in the Northern Technical University. His research interests are in the fields of pattern recognition, security, artificial intelligence and image processing. He can be contacted at email: raidrafi3@ntu.edu.iq.