A Review: Face Recognition Techniques using Deep Learning

Ghofran Khalid Hummady*, Asst. Prof. Mohand Lokman Ahmad**

*Department of Computer Technology Engineering, Technical College of Engineering, Northern Technical University / Mosul, Iraq, Email: ghofran.khalid@ntu.edu.iq
*Department of Computer Technology Engineering, Technical College of Engineering, Northern Technical University / Mosul, Iraq, Email: mohanadaldabag@ntu.edu.iq

Abstract

Face recognition (FR) is one of the most significant types of research that are widely used in various areas, such as finance, preventing crime, protecting the border, and for military purposes. Face recognition is a biometric identification technology based on human facial feature information. There are two main approaches first one is hand-crafted (HC) features which is the traditional method (geometry-based, holistic, feature based, and hybrid methods), and the recent one is based on deep learning (DL). The major purpose of this work to provide UpToDate literature review for face recognition (FR) Techniques. Furthermore, it summarizes the benchmark datasets and the most successful methods used on these datasets for face recognition.

Keywords- Face recognition, Deep Learning, CNN -Feature extraction.

I. INTRODUCTION

Computer vision technology has evolved and expanded steadily during the second part of the twentieth century.

Simultaneously, as software and hardware technologies connected to digital images become more widely used in people's lives, digital images have become an important component of information sources in modern civilization.

Face recognition is a biometric identification technique that uses information about a person's face features to identify them. Face image acquisition, preprocessing, face detection, face feature extraction, face recognition, and face body identification are all examples of automatic face recognition systems (Figure.1)[1]

In the recent era, the face recognition is important to make the world safer and used in many applications and is implemented in many systems such as checking the attendance access control for security reasons, preventing spoof attacks, education, information technology, banking & finance operation , management and many fields[2].

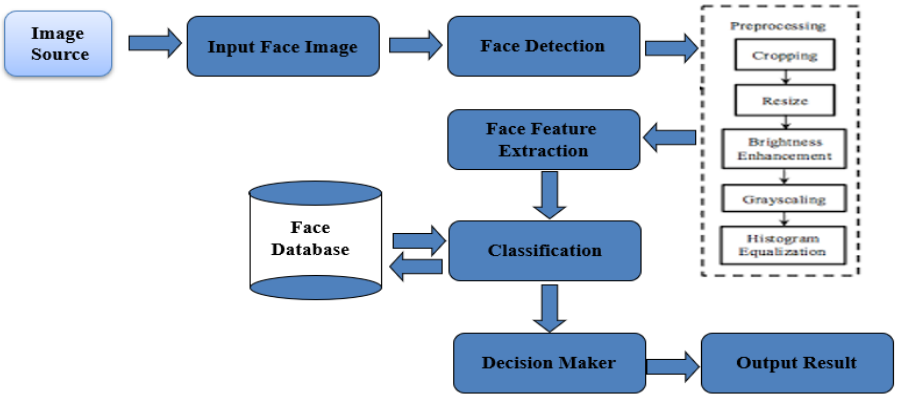

Figure 1: Steps of Facial Recognition system

AlphaGo is an artificial intelligence (AI) product that was published in 2016 by a team led by Deep-Demis Minda's Hassabis. Ke Jie, the top Come player in May 2017, was also defeated. In October 2017, the DeepMind team revealed AlphaGo Zero, the strongest version of AlphaGo.[3].

This work is categorized as following:

In the second section, we summarized an overview of a facial recognition steps with summarize each step, The Classification of Face Recognition Systems are reviewed in Section 3 with Section 4 discussed the concepts of Deep learning methods with three main types ((Convolutional neural network (CNN), Autoencoder, Generative adversarial network (GAN)), Eight papers are discussed in Section 5 with make comparison between them and summarized in a table including datasets, techniques, Conclusion are written in Section 6 .

2.FACIAL RECOGNITION STEPS

Face expression recognition (FER) algorithms have been used in a variety of ways. Both 2D and 3D approaches are taken into account and compared. The efficiency of FER is determined by the ease with which features for the descriptor can be retrieved and the descriptor's efficiency. Different expression descriptors should have a lot of variances, but the identical expressions should have minimal or no variance.

2.1 FACE DETECTION:

Face detection is the starting point for most face-related technologies, like face identification and verification. It is, can be very useful. The most successful application of face detection is likely to be image capturing. When you take a picture of your digital camera's face detection system determines where the faces are and changes the focus accordingly [4].

2.2 PREPROCESSING:

In the image of facial recognition, the first phase is pre-processing, which involves taking an input image and processing it. The image can be captured in a variety of ways that are not always in a conventional format. Various disturbances, such as noise, lighting variations, shadows, and size variations, might impact the input image. So, throughout this picture standardization process (preprocessing), which may include image enhancement, noise reduction, and scaling of the original image, these obstacles are removed from the image to prepare it for the next stage. the image is sometimes changed from color to gray scale to reduce the processing complicity. [5].

2.3 FEATURE EXTRACTION:

In the image compression process feature extraction is one of the stages that process will minimize the large dataset to the smaller one, in order to simplest the processing, this operation will extract the most significant feature. To process these variables, a large amount of computational system is needed. The common feature extraction techniques are Gabor filters , Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Local Binary Patterns (LBP) etc.[6].

2.4 CLASSIFIERS:

The features retrieved from the facial photos are assigned to the appropriate expression classes. The most widely used classifiers are the Haar Cascade and Fisher Face Classifier[7].

3. CLASSIFICATION OF FACE RECOGNITION SYSTEMS

There are three approaches that are divided into face recognition systems, based on (detection and recognition technology) [8] **:**

1) Local.
2) Holistic.
3) Hybrid.

A. Local approach:
In this approach is classified on bases of facial features like, mouth, nose and chin are extracted and fed to the classifier. Sometimes this approach is known as geometry feature-based approach. This technique is not used now a days.

B. Holistic approach:
The second method uses the full face as input data, which is then placed in a tiny subspace plane. This method was used by many researchers. This technique encompasses a number of methods are eigenfaces, fisher faces, support vector machine (SVM)[9].

C. Hybrid approach:
The third method improves facial recognition accuracy by combining local and global (holistic) data. Face recognition depends on the human facial features. According to research and studies, the eyes, mouth, and nose are among the most important features for recognition[10].

4.DEEP LEARNING METHODS

Scientist have developed various (DL) technologies to diverse tasks in recent years, and (FR) has benefited greatly from these techniques. Over the last few years, DL, which is part of a larger family of ML methods, has proven to be beneficial in multiple locations in the computer vision field. Working with data sets for large-scale training gives it a lot of advantages. Deep learning's basis is function learning. Its goal is to develop practical knowledge of hierarchical networks in order to tackle important challenges that necessitate artificial design[11].

4.1 CONVOLUTIONAL NEURAL NETWORK (CNN):

The name "CNN " refers to the network's use of the convolutional mathematical procedure, it is a (DL) system that can separate multiple aspects/objects in an image from an input image and is mostly used to classify photos and perform object detection in scenes by clustering them by similarity (photo search).

They are also known as shift invariant or space invariant artificial neural networks (SIANN), which are related to a shared-weight architecture of convolution kernels or filters that slide along the input features and form feature maps. The known translation provides the equivalent response. [12] .

It used in its work the convolution in place of the usual matrix multiplication in at least one of its layers. A convolutional neural network consists of an input layer, a hidden layer, and an output layer (Figure.2). Any layer in a feed-forward neural network is hidden.

since the activation function and the final convolution mask its inputs and outputs. ex., convolutional layers, and pooling layers. The common DL algorithm for image identification, pattern recognition, and other feature extraction operations from a picture is CNN.

CNN algorithms come in a variety of shapes and sizes. However, there are two types of explanations for the CNN algorithm, the extractor is one, while the classifier is the other. [13].

Figure.2: Basic CNN diagram

4.2 AUTOENCODER:

A moving autoencoder is a form of ANN that learns efficient information coding. An autoencoder's purpose is to train the network to ignore signal "noise" to learning a representation the encoder for a set of data, which is widely used to reduce dimensionality [14]. On the reduction front, the autoencoder learns a reconstruction side, it is trying to build a representation as like to the same input (Figure.3). Variants are used to require learned representations to take on useful features. Regular autoencoders are one example (sparse, denoising and convolutional). Autoencoders are used to solve a variety of problems, including facial recognition.[15].

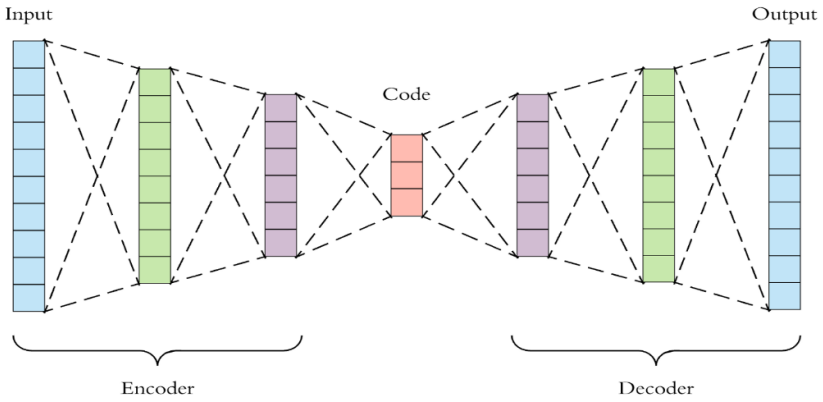

Figure.3: Basic Autoencoder diagram

4.3 GENERATIVE ADVERSARIAL NETWORK (GAN):

The GAN architecture is a type of artificial neural network. It is made up of two different network architectures that have been combined. GANs generate new data that has the same statistics as the dataset it was given.

GAN is made up of two distinct networks. The real data is received by one of these two independent networks, which then delivers it to a supervisory structure. The second network structure generates new data to reflect the original data and transmits it to the same controller structure as the first. These data are tested in the controller network or structure portion, and the resulting data is used to determine how similar the data from the copy structure is to the original data. If the representational network structure has not yet shown to be enough, (Figure.4). The basic idea of a GAN is based on "indirect" training through a discriminator[16].

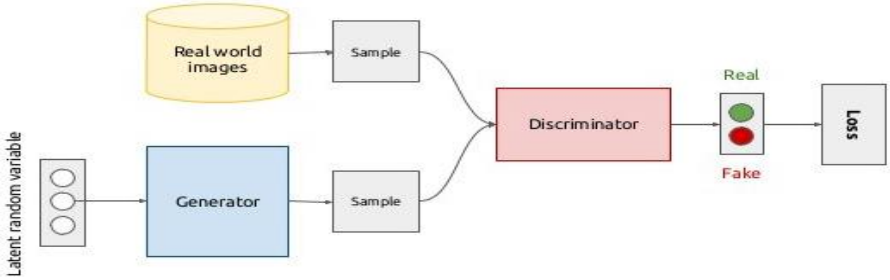

Figure.3: Basic Autoencoder diagram

5.OUTCOMES OF THE REVIEWED PAPERS

Pranav KB[13] Convolutional Neural Networks were used to construct and evaluate a Realtime facial recognition methods . Based on the AT&T dataset, to improve the recognition accuracy of the system constructed, several parameters of the CNN architecture are tuned. The suggested system achieves a recognition accuracy of 98.75%.Zied B.[17] use face recognition utilizing a combination of PCA, ICA, LDA with DWT, and SVM. It increases the rate of recognition. Strong lighting locations and face features affect the rate of recognition. Simulation using the AT&T dataset is used to assess the performance of different approaches. with recognition accuracy 96.00%. Zhang[18] presented a Posture-Weighted Generative Adversarial Network that worked in a generic way with large motion and photo-realistic frontal view synthesis changes (PW-GAN). With 98.38 % face verification precision on the LFW dataset, they frontalized the face picture through the 3D face, gave greater attention to big poses, and optimized the pose code in the loss function to overcome issues such as not being photo-realistic and losing 1D information. Nuclear Norm based Adapted Occlusion Dictionary Learning is a framework presented by Du and Hu[19] for dealing with illumination variations and occlusion in face recognition (NNAODL). In their framework, they used a two-dimensional structure and dictionary learning (DL). Using experiments on multiple public datasets LFW that achieve in Deep Learning (DL) technique this achieves 93.1% on the LFW . Wang[20] developed a pyramid diverse attention (PDA) method for learning multiscale distinct local representations automatically and adaptively. They claim that their model solves problems like pose changes, big expressions, and similarly local patches. They combined HBP and PDA to create the HPDA model. The stem CNN, local CNN, global CNN, and classification factors create this model. Face recognition was done using an artificial CNN network. They employed a pyramid diversified refer to create several attention-based local branches at different scales to stress distinct discriminating face regions automatically and adaptively at different scales. The accuracy was also tested on the popular LFW dataset, where the majority of the faces are frontal or near-frontal, and the result was 99.8%. Li[21] introduced a novel distance metric optimization technique that uses DCNN to integrate feature extraction, distance metric application, and interaction between them. It uses an end-to-end decision function to learn feature representation. They gathered photographs from people of all ages. Used CNN architecture is evaluating their method on the MORPH database with result accuracy was 93.6%. Lakshmi[22] used The ORL database will be used to train and test the models. Three different types of models are created, and their performance is evaluated. The ORL Database, which contains Convolutional Neural Networks (CNN), is used to train and test the models. Three different types of models are created, and their performance is evaluated and the suggested system achieves a recognition accuracy of 99.2%.ElBedwehy[23] introduced a new feature extraction method called Relative Gradient Magnitude Strength (RGMS). Deep Neural Networks are used in this procedure (DNNs). The experiments were performed out on popular datasets ORL, with the proposed method achieving a score of 98.75 %.The datasets, techniques, and accuracy results in eight studies are summarized in this table 1:

NO.	References	TECHNIQUE	DATASET	ACCURACY %
1.	[13]	CNN	AT&T	98.75
2.	[17]	Mixed combination of (PCA, ICA, LDA with DWT, and SVM)	AT&T	96
3.	[18]	GAN	LFW	98.38
4.	[19]	DL	LFW	93.1
5.	[20]	CNN	LFW	99.8
6.	[21]	CNN	MORPH	93.6
7.	[22]	CNN	ORL	99.2
8.	[23]	DNN	ORL	98.75

Table .1

6.DATASET

A. LFW （Labeled Faces in the Wild) It is the collection contains 13,000 facial photos gathered from the internet, each one labeled with the name of the individual who was photographed. There were 1,680 people who had two or more different photos. The main drawback to these data sets is that the original Viola Jones (Hamour) detector can detect them [24].

B. ORL (AT&T Dataset): The AT&T Laboratory at the University of Cambridge gathered the ORL dataset, which is a face dataset. Members of the laboratory from 1992.4 to 1994.4 are included. The photos in this data collection are grouped into 40 different subjects, with 10 photographs in each subject. The images for some of these subjects were shot at various periods. Care and facial emotions (eyes open, eyes closed, laughter, not smiling), face features (glasses), and so on differ. All photographs are shot from the front to the top and have a black consistent backdrop. The photos are in PGM format, with a size of 92 * 102 pixels and 256 gray channels[25].

C. MORPH: is a dataset well-known to estimate the facial age that contains a total number of 55,134 facial images from 13,617 participants old men 16 to 77. This dataset is a longitudinal face, formed for academics searching into all aspects of adult age development, like face modeling, image-realistic animation, and face recognition, amongst different things. It contributes to various active study fields, the utmost prominent of that is face recognition, through introducing the biggest samples of longitudinal images as publicly; longitudinal spans in the range between a (one month to twenty years); the inclusion of key physical parameters that affect old age appearance. This dataset has directly contributed to the face recognition algorithms by demonstrating the influence of age progression on recognition methods rate. [21].

7.CONCLUSION

In recent years, Due to the number of fake faces being created using artificial intelligence is increasing, giving a new dimension to disinformation and cyber-attacks, that cannot be detected by the naked eye the face recognition and detection system considerable attention from researchers.

Face recognition and detection technologies have increased strongly during the past years in many fields from security, and forensic applications requiring the use of face recognition technologies and become the most secure tool at the level of countries, institutions, or personal levels.

In this paper, we highlighted the recent researches by briefly discussing eight of the existing literature on face recognition and detection. Besides that, we were focused to compare CNN with various techniques that were used in this study with the same given dataset. The accuracy rate has been compared. We noticed that the CNN has greater accuracy in obtaining overall datasets, as compared with all other used methods.

Although these techniques have achieved a lot of success and can beat the CNN But need more development to improve the image quality to solve the challenges such as lighting conditions and facial expressions and keep pace with the development of counterfeiting techniques and solve the problems of image filtering, image reconstruction, rotation, and occlusion.

REFERENCES

[1] M. T. H. Fuad *et al.*, "Recent advances in deep learning techniques for face recognition," *IEEE Access*, vol. 9, pp. 99112–99142, 2021, doi: 10.1109/ACCESS.2021.3096136.

[2] A. Mutrak, "Intelligent Virtual Assistant - VISION," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 9, no. 5, pp. 2057–2060, 2021, doi: 10.22214/ijraset.2021.34757.

[3] D. Silver *et al.*, "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm," pp. 1–19, 2017, [Online]. Available: http://arxiv.org/abs/1712.01815

[4] C. Vision, D. Algorithm, F. Expression, and H. Liu, "Face Technologies on Mobile Devices Face Detection and Face Direction Esti- mation Using Color and Shape Features," 2015.

[5] Institute of Electrical and Electronics Engineers. and IEEE Photonics Society., "2011 Symposium on Photonics and Optoelectronics (SOPO) : May 16-18, 2011, Wuhan, China," no. 2, pp. 3–6, 2011.

[6] N. Samadiani *et al.*, "A review on automatic facial expression recognition systems assisted by multimodal sensor data," *Sensors (Switzerland)*, vol. 19, no. 8, pp. 1–27, 2019, doi: 10.3390/s19081863.

[7] S. O. Adeshina, H. Ibrahim, S. S. Teoh, and S. C. Hoo, "Custom face classification model for classroom using haar-like and lbp features with their performance comparisons," *Electron.*, vol. 10, no. 2, pp. 1–15, 2021, doi: 10.3390/electronics10020102.

[8] I. Technology, "A REVIEW ON FACIAL RECOGNITION INCLUDING LOCAL, HOLISTIC AND HYBRID APPROACHES Prince Goyal, Heena wadhwa," vol. 21, no. 2, pp. 210–216, 2020.

[1] M. T. H. Fuad *et al.*, "Recent advances in deep learning techniques for face recognition," *IEEE Access*, vol. 9, pp. 99112–99142, 2021, doi: 10.1109/ACCESS.2021.3096136.

[2] A. Mutrak, "Intelligent Virtual Assistant - VISION," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 9, no. 5, pp. 2057–2060, 2021, doi: 10.22214/ijraset.2021.34757.

[3] D. Silver *et al.*, "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm," pp. 1–19, 2017, [Online]. Available: http://arxiv.org/abs/1712.01815

[4] C. Vision, D. Algorithm, F. Expression, and H. Liu, "Face Technologies on Mobile Devices Face Detection and Face Direction Esti- mation Using Color and Shape Features," 2015.

[5] Institute of Electrical and Electronics Engineers. and IEEE Photonics Society., "2011 Symposium on Photonics and Optoelectronics (SOPO) : May 16-18, 2011, Wuhan, China," no. 2, pp. 3–6, 2011.

[6] N. Samadiani *et al.*, "A review on automatic facial expression recognition systems assisted by multimodal sensor data," *Sensors (Switzerland)*, vol. 19, no. 8, pp. 1–27, 2019, doi: 10.3390/s19081863.

[7] S. O. Adeshina, H. Ibrahim, S. S. Teoh, and S. C. Hoo, "Custom face classification model for classroom using haar-like and lbp features with their performance comparisons," *Electron.*, vol. 10, no. 2, pp. 1–15, 2021, doi: 10.3390/electronics10020102.

[8] I. Technology, "A REVIEW ON FACIAL RECOGNITION INCLUDING LOCAL, HOLISTIC AND HYBRID APPROACHES Prince Goyal, Heena wadhwa," vol. 21, no. 2, pp. 210–216, 2020.

[9] S. Dhawan and N. Khurana, "Volume 2 , Issue 2 (February 2012) A REVIEW OF FACE RECOGNITION ISSN : 2249-3905 IJREAS Volume 2 , Issue 2 (February 2012) ISSN : 2249-3905," vol. 2, no. 2, pp. 835–846, 2012.

[10] A. A. Fathima, S. Ajitha, V. Vaidehi, M. Hemalatha, R. Karthigaiveni, and R. Kumar, "Hybrid approach for face recognition combining Gabor Wavelet and Linear Discriminant Analysis," *2015 IEEE Int. Conf. Comput. Graph. Vis. Inf. Secur. CGVIS 2015*, pp. 220–225, 2016, doi: 10.1109/CGVIS.2015.7449925.

[11] D. Graupe, "Deep Learning Convolutional Neural Network," *Deep Learn. Neural Networks*, pp. 41–55, 2016, doi: 10.1142/9789813146464_0005.

[12] S. Almabdy and L. Elrefaei, "Deep convolutional neural network-based approaches for face recognition," *Appl. Sci.*, vol. 9, no. 20, 2019, doi: 10.3390/app9204397.

[13] K. B. Pranav and J. Manikandan, "Design and Evaluation of a Real-Time Face Recognition System using Convolutional Neural Networks," *Procedia Comput. Sci.*, vol. 171, no. 2019, pp. 1651–1659, 2020, doi: 10.1016/j.procs.2020.04.177.

[14] W. H. Lopez Pinaya, S. Vieira, R. Garcia-Dias, and A. Mechelli, "Autoencoders," *Mach. Learn. Methods Appl. to Brain Disord.*, pp. 193–208, 2019, doi: 10.1016/B978-0-12-815739-8.00011-0.

[15] H. Sewani and R. Kashef, "An autoencoder-based deep learning classifier for efficient diagnosis of autism," *Children*, vol. 7, no. 10, 2020, doi: 10.3390/children7100182.

[16] A. Aggarwal, M. Mittal, and G. Battineni, "Generative adversarial network: An overview of theory and applications," *Int. J. Inf. Manag. Data Insights*, vol. 1, no. 1, p. 100004, 2021, doi: 10.1016/j.jjimei.2020.100004.

[17] Z. B. Lahaw, D. Essaidani, and H. Seddik, "Robust Face Recognition Approaches Using PCA , ICA , LDA Based on DWT , and SVM algorithms," *2018 41st Int. Conf. Telecommun. Signal Process.*, no. January 2022, pp. 1–5, 2018, doi: 10.1109/TSP.2018.8441452.

[18] S. Zhang, Q. Miao, X. Zhu, Y. Chen, Z. Lei, and J. Wang, "POSE-WEIGHTED GAN FOR PHOTOREALISTIC FACE FRONTALIZATION University of Chinese Academy of Sciences 2 National Laboratory of Pattern Recognition , Institute of Automation Chinese Academy of Sciences , Beijing , China , 100190," *2019 IEEE Int. Conf. Image Process.*, pp. 2384–2388, 2019.

[19] L. Du and H. Hu, "Neurocomputing Nuclear norm based adapted occlusion dictionary learning for face recognition with occlusion and illumination changes," *Neurocomputing*, vol. 340, pp. 133–144, 2019, doi: 10.1016/j.neucom.2019.02.053.

[20] Q. Wang, T. Wu, H. Zheng, and G. Guo, "Hierarchical pyramid diverse attention networks for face recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2020, pp. 8323–8332. doi: 10.1109/CVPR42600.2020.00835.

[21] D. Metric *et al.*, "PT," 2017, doi: 10.1016/j.patcog.2017.10.015.

[22] L. Patil and V. D. Mytri, "Face recognition with CNN and inception deep learning models," *Int. J. Recent Technol. Eng.*, vol. 8, no. 3, pp. 1932–1938, 2019, doi: 10.35940/ijrte.C4476.098319.

[23] M. N. Elbedwehy and G. M. Behery, "Face Recognition Based on Relative Gradient Magnitude Strength," *Arab. J. Sci. Eng.*, 2020, doi: 10.1007/s13369-020-04538-y.

[24] N. Zhang and W. Deng, "Fine-grained LFW database," *2016 Int. Conf. Biometrics, ICB 2016*, pp. 1–11, 2016, doi: 10.1109/ICB.2016.7550057.

[25] C. Engineering and S. Domain, "Deep Learning based Human Recognition using Integration of GAN and Spatial Domain Techniques," vol. 21, no. 8, 2021.