International Journal on Advanced Science Engineering Information Technology

Vol.12 (2022) No. 6
ISSN: 2088-5334

Access Control System based on Voice and Facial Recognition Using Artificial Intelligence

José Capote-Leiva [a], Marco Villota-Rivillas [a], Julián Muñoz-Ordóñez [a,*]

[a] *Department of Engineering, Corporación Universitaria Comfacauca, Unicomfacauca, Popayán, 190003, Colombia*
Corresponding author: [*]jfmunoz@unicomfacauca.edu.co

***Abstract*—Computer security has become a matter of great concern at the global level. Whatever the economic sector, all companies handle confidential information related to clients and personnel. These latter therefore need to be seen as sensitive assets that require protection. Appropriate, consensual handling of personal information is thus a legal and delicate requirement that demands to be treated securely in all types of business. Many companies and clients have lost sums of money in the millions due to incorrect information protection, triggering complicated, expensive proceedings that are awkward and cumbersome to resolve. To implement an authentication system based on biometric parameters, which strengthens the security of sensitive assets in areas considered critical in various organizations by attaining the highest accuracy in user classification processes. By applying a convolutional neural network: *MobileNet*, viewing via computer and low-cost devices (Raspberry Pi 3). The system constitutes an authentication device for voice and face with 96% and 100% accuracy, respectively. Conclusion: The system shows that deep learning (Deep Convolutional Neural Networks), in combination with devices such as the Raspberry, generate a system capable of high performance in time and cost, as well as providing companies with a robust system featuring high accuracy in the correct recognition of the biometric patterns of users registered and trained by the system.**

***Keywords*— Artificial intelligence; computer vision; convolutional neural network; deep learning; spectrogram; biometric system.**

Manuscript received 14 Aug. 2021; revised 24 Sep. 2021; accepted 13 Dec. 2021. Date of publication 31 Dec. 2022.

I. INTRODUCTION

Computer security has become a global concern [1]–[4]. All companies-not only large or multinational companies, but medium and small companies from any economic sector - handle information, often confidential, about their clients, and at the same time, about their staff, making them sensitive assets that need to be protected [5], [6]. The appropriate, careful handling of personal information has long been a delicate legal matter requiring to be handled most securely in every type of company. Since, whenever the smallest of gaps has left this type of information vulnerable, many companies and clients have lost sums in the millions. They have likewise become embroiled in very complicated, expensive, and cumbersome legal proceedings [7]. For this reason, management of the information must ensure that it will not fall into fraudulent processes [8] and be used for beneficial purposes both for the companies and the client.

Within organizations, there are areas to which access is restricted, permitting entry only to certain users or collaborating personnel. However, the security required to authenticate the data of each permitted person must be stored in a very specialized system that manages to encrypt the information and ensure the security of the asset.

Today, leading countries in computing and electronics incorporate different objectives with their devices, including cost, efficiency, portability, security, and connectivity. The priority needs to be identified for all companies to implement and ensure the security of the information corresponding to their most important assets. As has been stated, generally in any company or institution, there is a private or specific area for maintaining assets of great importance, whether physical or virtual. This is kept in the custody of company personnel who have been granted access to it and what that area contains. A sufficiently secure system must avoid any vulnerability that puts the company's assets at risk, thus ensuring that only those with access can actually enter the site.

Various systems implemented today are controlled by biometric systems [9]–[12], but some tools can be compromised and remain susceptible to theft or fraud due to their lack of robustness and the fact that they fail to carry out a mix of biometric identifiers known to increase security. Deep learning based on the construction of neural network

architectures has shown its potential in face and voice recognition, reaching high values of accuracy, close to 97.25% [13]–[15]. These methods are novel and require a computational cost to their training, which is why they could not be embedded in low computational capacity devices.

Authentication has become the most reliable technique to protect information or valuables. For an authentication system to instill total confidence that a person actually is who they say they are, there are three ways, depending on what they know, what they have, or who it is [16]–[18]. For this project, "who it is" was used with the help of biometrics in the area of face and voice recognition, which are unique personal characteristics and parameters [19], [20]. In order to strengthen the authentication system, Deep Learning techniques were applied [21], [22] in the fields of artificial vision, natural language processing, and speech processing [19]. The research was carried out in the southern Colombian city of Popayan. According to the requirements, the following projects based on authentication with biometric characteristics, face, speech, speaker, and RFID technology were used as references.

The Hybrid Biometric Person Authentication Project [23] published in 2001 presented a hybrid authentication prototype using two levels of identity validation, front face recognition and speaker voice recognition, depending on the word or phrase chosen as key. The extraction and classification algorithms for the processing of voice signals and image processing were evaluated with a real database, taking into account that the face samples were captured using a virtual template of the user's face [24]. However, the system in voice authentication recognizes a keyword as a password instead of the patterns of the speaker's voice for greater security in authentication, opening the possibility of working in the future by capturing the face without using the default templates and recognition of the speaker by the unique voice patterns.

Richardson *et al.* [25] developed approaches based on deep neural networks for recognizing the speaker and the language using deep neural networks. They improved the performance and response time of the classification and recognition of the speaker and language [25]. A system based on Raspberry Pi was published in 2019; the project developed a system based on face recognition for door unlocking. In this project, an automatic door opening system based on facial recognition was designed and implemented using computer vision and a single board (Raspberry Pi) computer responsible for controlling access to authorized persons in areas such as homes, bank lockers, and associated control operations [26]. The system showed high accuracy in face recognition and door opening, although it did not have a user-friendly graphical interface to show the authentication status and had only a single stage of identity validity.

Biometric authentication by fingerprint is one of the most widely used systems today. The incorporation of fingerprint authentication systems is regularly used as a digital signature in-state organizations to control access to critical areas, attendance control, and even the unlocking of electronic equipment such as laptops and Smartphones [27]. However, with it being so popular, techniques have been developed to spoof the fingerprint as pattern replicas and capture the fingerprint illegally. Moreover, these contact systems are not suitable for preventing the spread of infections due to the constant manipulation of establishment staff, which is why these systems require to be disabled in times of health crisis.

The above is framed in the "what I am" biometric security principle. However, in the field of the "what I have" principle, there are developments focused on access control, mainly in RFID-based technologies. Depending on its reading range, RFID technology is used in activities such as access control in organizations, public transport cards, and in specific applications such as livestock tracking, among others. It is characterized by a lower cost and allows information to be saved [28] despite the fact that the system for its optimal operation depends on the portability of the devices. In case of loss, theft, or damage, the system no longer functions.

The project exhibited at the International Conference on Quality in Research (International Conference on Quality in Research, 2017) presented the implementation of RFID and Raspberry Pi technology for the authentication and payment of public transport users using a card, allowing passengers to be registered automatically without the need for contact with access control devices. The system can identify the card data in a range of one to two meters [29]. The system works correctly in controlled environments. However, the RFID card can be easily forgotten or lost, and stolen. If it is not carried on entering public transport areas, the system will not recognize the user.

The Face++ company is currently an industry leader in real-world applications, engaged in face image processing research in the areas of face detection, facial points, face attributes, face comparison, and search. Face++ offers companies and organizations the authentication of people in everyday moments. Some of these are the daily registration of personnel, attendance control, pay payroll, and delivery of loans, even though the service it offers has a high cost in implementing the API and annual memberships and not having support for speaker authentication by voice.

In this research, recognition of biometric patterns such as voice and face were found to be the characteristics that offer the lowest vulnerability to spoofing and thus better security. In addition, staff's physical contact with the device is required, lending support to health processes in times of health crisis. The foregoing makes it possible to identify a gap in the research and thus implement a biometric system based on voice and face for control of access to restricted areas.

In Popayan, authentication processes are archaic and insecure, the most common being fingerprint identification and EMV chips [17]. The proposal detailed here combined two biometric parameters, each with a very low vulnerability index - the face of the user and voice of the user [19], [30], [31]. It furthermore aimed to generate a low-cost option for small and medium-sized companies. As such, the research put forward the following question: Is it possible to implement an authentication system, based on biometric parameters, which strengthens the security of sensitive assets in areas considered critical in various organizations, achieving high accuracy in Username classification processes?

II. Materials and Method

The biometric system was built using a Raspberry Pi 3 in charge of executing the convolutional neural network models trained to recognize the voice and face of registered users.

These users simulated a real entry of people with different levels of access in the company. The general diagram for the hardware implementation is shown in Figure 1. The system was designed and implemented in a prototype consisting of a wooden door, a 5 Mega-pixel Raspberry Pi-type webcam module, a USB microphone, a 12 V push-pull solenoid, and an HC-SR04 ultrasound sensor and 5V relay module:

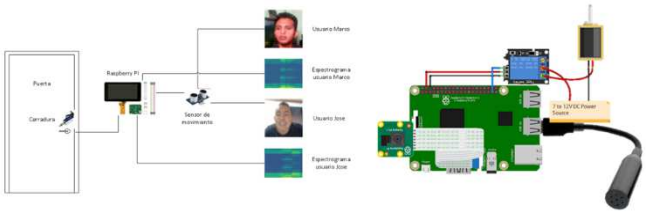

Fig.1 Modular representation of the constructed biometric system

Figure 1 indicates how samples corresponding to the user's face and voice are captured using a camera and microphone connected to the Raspberry. The audio samples are converted to spectrograms so that a convolutional neural network can process them. The system consists of a motion sensor to avoid executing the processes unnecessarily when the user is not in front of the biometric door.

The security levels initially consist of recognizing the user's face. If this method is effective, voice recognition is performed. Once the two stages have been completed, the Raspberry sends a signal to a solenoid responsible for opening the door. Samples of faces and voice (spectrograms) are presented in Figure 2.

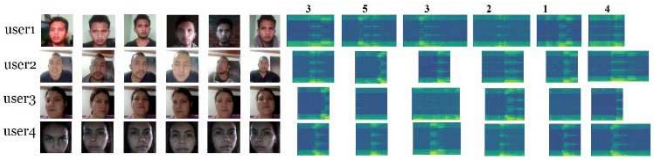

Fig. 2 Samples of faces and audios converted into a spectrogram of 4 different users. The audios correspond to pronouncing the numbers randomly: 1 to 5

To train a convolutional neural network for face recognition, a data set with 5 users with different access levels was built during the execution of the project. The data set consisted of 100 facial samples for each user. The age ranges varied to enrich the training samples. The biometric system captured audio from users and, through mathematical transformations, reconstructed the data into an image with the following characteristics: Y-axis encoded amplitude, X-axis encoded time, and the spectral content of the pixel corresponding to the intensity of the audio. The spectrograms were coded with a color palette, using the Matplotlib module, to show each user's different changes in intensity. The algorithm of the biometric access system software is shown in Table 1.

TABLE I
AUTHENTICATION SYSTEM PSEUDOCODE

Pseudocode: Authentication system
Input variables: facial and voice model
Output variables: valid access or access denied
Define interface elements (column 1 for text, column 2 for camera image)
Create interface (*column 1 of size 200*480 and column 2 with of size 200*480*)
Upload model containing facial information of users
Start camera capture. This is then displayed in column 2
Repeat 10 times
Capture image
Predict to which user the photo belongs
Save the prediction in a vector
Evaluate the predictions
if at least 7 photos belong to a specific user
Save the id of the corresponding user
Close the face authentication interface
Otherwise
Print access denied
Close the authentication system interface
Define interface elements (column 1 for text, column 2 for microphone image)
Create interface (*column 1 of size 200*480 and column 2 with of size 200*480*)
Upload model containing user voice information
Repeat 4 times
Print on the screen a random number from 1 to 5 that the user must say
Record user voice for 1.5 seconds
Save the user's voice in an audio file with a .WAV extension
Convert .WAV file to spectrum
Predict to which user the recording belongs
Save the prediction in a vector
Evaluate the predictions **if** at least 2 recordings belong to a specific user
Save the id of the corresponding user
Otherwise
Print access denied
Close the authentication system interface
If the user id saved in face authentication is the same as the user id saved in voice authentication, **Then**
Print on the screen the text welcome and the user's name
Otherwise
Print access denied

The pseudocode in Table 1 shows a functional biometric authentication system with a graphical interface. For operation, it requires input peripherals such as the webcam and microphone. In lines 1 and 2, the components that make up the graphical interface for facial authentication are defined and created. In line 3, the previously trained model is loaded with the facial information of the users with authorized access to the system. Line 4 starts the webcam and displays the image in column 2 of the interface. From lines 5 to 8, the process begins in which the webcam captures an image. Later it is predicted to which user it belongs.

The result of the prediction is stored in a vector. This process is carried out 10 times, obtaining a vector of size 10 with the id of the users that was predicted in each interaction. Lines 9 to 14 evaluate if at least 7 id of the resulting vector belongs to the same user. If so, it saves the id and closes the facial authentication interface. Otherwise, it shows access denied on the screen and closes the authentication system interface. Lines 15 and 16 define and create the components that make up the graphical interface for voice authentication. In line 17, the previously trained model is loaded with information from users' voice with authorized access to the system. Lines 18 to 24 begin the process that is repeated 4 times: a random number from 1 to 5 is printed on the screen which the user must repeat. The system records the user saying that number. The recording is saved in a WAV file and is converted to a spectrogram obtaining an image in JPG format. The user is predicted to which the spectrogram belongs, and the user's id is saved in a vector. Lines 25 to 29 are evaluated if at least 2 of the IDs of the resulting vector belong to the same user and the user's id is saved. Otherwise, it shows access denied on the screen and closes the interface of the authentication system. In line 30, the system evaluates if the id of the user is saved in the authentication.

An interface is implemented for the collection of audios with a WAV extension. Using the set of audios collected with the different users, the audios are transformed into spectrograms. To develop the process of taking audio samples with an interface that allows the creation of audios in WAV extension, the code written in Python is executed. It is executed in a command line, writing Python and the name with which the program is realized. An interface is displayed where the number can be seen, chosen randomly from 1 to 5. In addition, a message is displayed on the screen that indicates that a recording process is being carried out. The user must pronounce the numbers as indicated.

The audio capture system is started, in which an interface with an illustration that shows the system startup is displayed. The system randomly chooses the number that the user must pronounce to capture the respective audio. The system has saved audio that guides the user in step-by-step the correct way to say the numbers. This process is carried out 100 times, achieving more information from the user's voice. The authentication system was developed with a series of consecutive interfaces, achieving greater user interaction with the authentication system, as can be seen in Figure 3.

Figure 3 shows the sequence performed by the authentication system when a user wishes to enter a restricted area. The video of the system operation can be found in the link shared in this document. For the project, a *MobileNet* architecture [32], [33] was used. This neural network model obtained competitive values in recognizing objects in the ImageNet competition, making it a powerful network architecture suited to the project's requirements. What is innovative about this architecture is that its configuration has a low computational cost. The architecture is made up of convolutional layers capable of processing image samples corresponding to the user's face and spectrograms. The result of training the convolutional neural network with 500 epochs, an initial learning rate of 0.0001, and a batch size of 16 generates a trained model that is stored in the following link: https://github.com/tesis-uni /System-de-authentication.git

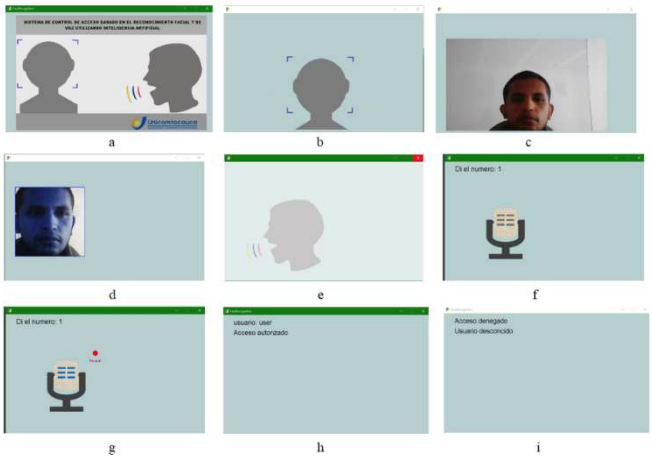

Fig. 3 Operation sequence of the biometric system: a) initial system screen, b) face detection started, c) company user who wishes to authenticate, d) face detection in order to eliminate outside noise, e) start voice recognition module, f) message on screen of the number that the user must pronounce, g) indication that the system is recording the user's voice, h) and i) possible results that can be obtained from the biometric evaluation of the face and the voice.

III. Results and Discussion

The training of neural networks using the *Keras* deep learning library, the performance of the machine running the models, and the evaluation metrics: confusion matrix, precision, recall, and error are presented in this section. The results begin with the test of the model on its face recognition module. The system was tested with 5 users. Thirty (30) different scenes of simulated arrivals to the company were generated (assuming that the user would enter the restricted area throughout one full month). For a total of 150 samples in the test set, the system was able to perform the task correctly in all cases.

TABLE II
CONFUSION MATRIX FOR FACE RECOGNITION

		Real data				
		User1	**User2**	**User3**	**User4**	**User5**
Classifier results	**User1**	30	0	0	0	0
	User2	0	30	0	0	0
	User3	0	0	30	0	0
	User4	0	0	0	30	0
	User5	0	0	0	0	30
General accuracy	**100%**					
Kappa	**1.0**					

The training results can be visualized in confusion matrices (Table 2 and Table 3). The vertical axis corresponds to the prediction result (Classifier Results), and the horizontal axis

represents the real sample class (Real Data). Table 2 shows the evaluation metrics for the face recognition model. The accuracy, precision, recall, and F1Score metrics have a score of 1, equivalent to 100% general accuracy with a Kappa coefficient of one.

Table 2 shows the confusion matrix for speech recognition. The experiment consists of 30 examples of speech spectrum samples for each class/user. For a total of 150 samples, the system was able to get it right in class 3. With respect to classes 2 and 5, 29 of 30 examples were classified correctly. Finally, for class 1, corresponding to user 1, 93.33% of the total samples were classified correctly.

The evaluation metrics for the voice model obtained an accuracy between 96.67% and 99.33%, precision between 0.86 and 1, recall between 0.93 and 0.97, and F1Score a score of 0.96, equivalent to 96% confidence in the prediction of the model (See Table 4). In Table 3, a Kappa coefficient of 0.95 and general accuracy of 96% are observed.

TABLE III
CONFUSION MATRIX FOR VOICE RECOGNITION

		Real data				
		User1	**User2**	**User3**	**User4**	**User5**
Classifier results	**User1**	28	0	0	0	1
	User2	0	29	0	0	0
	User3	2	1	30	2	0
	User4	0	0	0	28	0
	User5	0	0	0	0	29
General accuracy	**96%**					
Kappa	**0.95**					

TABLE IV
EVALUATION METRICS FOR VOICE RECOGNITION

	Accuracy	**Precision**	**Recall**	**F1 Score**
User1	98%	0.97	0.93	0.95
User2	99.33%	1.0	0.97	0.98
User3	96.67%	0.86	1.0	0.92
User4	98.67%	1.0	0.93	0.97
User5	99.33%	1.0	0.97	0.98

The results obtained by a remote connection to the *Raspberry Pi 3* with the VNC Server to test the performance of the prototype built in this project are presented below. After installing the necessary libraries for the execution of the project, the measurement tests of the capacity of the *Raspberry* Pi 3 to execute *Tensorflow, Keras*, *OpenCV*, *Numpy, Pyaudio*, *Pillow* was carried out. The tests were implemented in a virtual environment to avoid problems with future versions or updates. We worked with Python 3.7, in which the import tests of the libraries were carried out with a success factor of 100%.

After verifying the performance and consumption of resources, the project was executed. In Figure 4, the execution of the biometric authentication system with the home interface is shown. The consumption of 11% of the CPU was observed. This performance varied in the project according to the process of loading and reading the training models, and subsequently, their classification.

When the system was subjected to processes of face capture, a consumption of 26% of the CPU was observed. In this process, the CPU does not require a high consumption because face detection is carried out with a fairly light Haar classifier that ensures that the face is found when the user is 40cm from the capture device. The convolution operations in the face recognition process using the *MobileNet* architecture have an approximate consumption of 42% of the CPU; while consumption shoots up to almost half of the capacity, the system can execute the module without altering or disrupting the consumption and performance of the Raspberry. Once the face recognition stage was finished, the system proceeded to capture the audio, randomly generating digits that the user was required to pronounce. The CPU consumption reached 45% because, after the audio capture, an audio-to-image (spectrogram) signal transformation was required to be performed. When the face and voice recognition processes are successful, the Raspberry sends an electrical pulse to the lock authorizing the opening of the biometric door. The system frees the resources ensuring to maintain a suitable temperature on the device to avoid failures due to overheating.

Fig. 4 Execution of biometric system on a Raspberry Pi 3.

IV. CONCLUSION

According to the study conducted, it can be concluded that deep learning artificial neural networks in the field of biometric security allow greater control of authentication in the various areas in which the biometric system is installed. The above can be defined based on the results of the evaluation metrics carried out in this project.

To design a deep learning neural network, it is important to determine the area in which it will be implemented. The type of information it will receive and the hardware device in which it will be executed, thus identifying fundamental parameters of that neural network - such as the number of hidden layers and their associated neurons, batch sizes in the input samples, learning rates, probability of neuron deactivation to avoid overfitting and definition of loss functions. In the case of the project, the source of information was images. As such, convolutional neural networks were chosen because these are specialists in this type of information.

The *MobileNet* neural network was appropriate for this research project since it is a very light deep neural learning network, which can ensure its optimal operation on the Raspberry Pi 3 development board. This architecture does not have a very high processing level due to its particular characteristics, making it possible to adjust parameters such as training times, sample batch size, and learning rate.

The implemented system managed to recognize the faces of the users participating in the project with precision and accuracy of 100%. Meanwhile, the voice authentication system obtained a general accuracy of 96%, meaning that voice recognition needs to be improved in the training

processes of the neural network. The execution of libraries such as *Keras-Tuner* could help by adjusting the hyperparameters of the network. In this project, no search strategy for optimal values was used, only the authors' experience in training neural networks.

In the research carried out, it was determined that the system's accuracy in face authentication is improved by adjusting regions of interest, making the system work independently of the background in which it is installed, maintaining adequate accuracy. Adjusting the regions of interest eliminates the peculiarity of identity theft, using a photo that is shown to the system, both physical and virtual, for example, through a telephone device.

The evaluation of the biometric system, both in the field and in-situ tests, showed high values in its accuracy metrics in its face and voice recognition processes, achieving 100% and 96% accuracy, respectively. The system is a low-cost proposal with a minimum level of contamination or disease transmission due to the absence of manipulation by users of the devices that it comprises.

As future work, the authors of this research proposal to develop a learning transfer module in convolutional neural networks to expand the number of users belonging to a specific organization and to implement a richer data set in terms of users, thereby facilitating the strengthening of the final model of the neural network. It ought to be clarified that this would depend on the company where the system operates. Finally, the authors propose to implement the prototype within a real environment.

ACKNOWLEDGMENT

The authors are grateful to the Computing and Applied Informatics Group (MIND) of the Corporación Universitaria Comfacauca–Unicomfacauca. We are especially grateful to Colin McLachlan for suggestions related to the English text.

REFERENCES

[1] R. Blanco-Gonzalo, C. Lunerti, R. Sanchez-Reillo, and R. M. Guest, "Biometrics: Accessibility challenge or opportunity?," *PLoS One*, vol. 13, no. 3, p. e0194111, Mar. 2018, doi: 10.1371/journal.pone.0194111.

[2] E. Bertino, M. Kantarcioglu, C. G. Akcora, S. Samtani, S. Mittal, and M. Gupta, "AI for Security and Security for AI," in *CODASPY 2021 - Proceedings of the 11th ACM Conference on Data and Application Security and Privacy*, Apr. 2021, pp. 333–334, doi: 10.1145/3422337.3450357.

[3] A. De Keyser, Y. Bart, X. Gu, S. Q. Liu, S. G. Robinson, and P. K. Kannan, "Opportunities and challenges of using biometrics for business: Developing a research agenda," *J. Bus. Res.*, vol. 136, pp. 52–62, Nov. 2021, doi: 10.1016/j.jbusres.2021.07.028.

[4] R. Blanco-Gonzalo *et al.*, "Biometric Systems Interaction Assessment: The State of the Art," *IEEE Trans. Human-Machine Syst.*, vol. 49, no. 5, pp. 397–410, Oct. 2019, doi: 10.1109/THMS.2019.2913672.

[5] G. C. Yang and H. Oh, "Implementation of a personal authentication system T-time," in *Proceedings - 2019 IEEE 12th Conference on Service-Oriented Computing and Applications, SOCA 2019*, Nov. 2019, pp. 79–81, doi: 10.1109/SOCA.2019.00019.

[6] A. Barros *et al.*, "Data improvement model based on ecg biometric for user authentication and identification," *Sensors (Switzerland)*, vol. 20, no. 10, p. 2920, May 2020, doi: 10.3390/s20102920.

[7] M. Gautier and D. Jaafar, "Legal issues surrounding data protection," *Soins*, vol. 64, no. 838, pp. 36–39, Sep. 2019, doi: 10.1016/j.soin.2019.06.007.

[8] R. De La Rocha Ladeira and R. Rodrigues Obelheiro, "Automatic challenge generation for teaching computer security," in *Proceedings - 2018 44th Latin American Computing Conference, CLEI 2018*, Oct. 2018, pp. 774–783, doi: 10.1109/CLEI.2018.00098.

[9] A. Das, S. K. Mohapatra, and L. P. Mishra, "Biometric detection using stroke dynamics," in *Lecture Notes in Networks and Systems*, vol. 109, Springer, Singapore, 2020, pp. 458–466.

[10] P. Kumar, R. Saini, B. Kaur, P. P. Roy, and E. Scheme, "Fusion of neuro-signals and dynamic signatures for person authentication," *Sensors (Switzerland)*, vol. 19, no. 21, Nov. 2019, doi: 10.3390/s19214641.

[11] P. Panasiuk, M. Dąbrowski, and K. Saeed, "Keystroke dynamics and face image fusion as a method of identification accuracy improvement," in *Advances in Intelligent Systems and Computing*, 2017, vol. 567, pp. 187–196, doi: 10.1007/978-981-10-3409-1_13.

[12] V. Sujitha and D. Chitra, "A Novel Technique for Multi Biometric Cryptosystem Using Fuzzy Vault," *J. Med. Syst.*, vol. 43, no. 5, pp. 1–9, Mar. 2019, doi: 10.1007/s10916-019-1220-x.

[13] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, pp. 1–74, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[14] Y. Hu, M. Lu, C. Xie, and X. Lu, "FIN-GAN: Face illumination normalization via retinex-based self-supervised learning and conditional generative adversarial network," *Neurocomputing*, vol. 456, pp. 109–125, Oct. 2021, doi: 10.1016/j.neucom.2021.05.063.

[15] J. Zeng, X. Qiu, and S. Shi, "Image processing effects on the deep face recognition system," *Math. Biosci. Eng.*, vol. 18, no. 2, pp. 1187–1200, 2021, doi: 10.3934/MBE.2021064.

[16] C. S. Hsiao, C. P. Fan, and Y. T. Hwang, "Iris location and recognition by deep-learning networks-based design for biometric authorization," in *LifeTech 2021 - 2021 IEEE 3rd Global Conference on Life Sciences and Technologies*, Mar. 2021, pp. 144–145, doi: 10.1109/LifeTech52111.2021.9391787.

[17] A. C. Weaver, "Biometric authentication," *Computer (Long. Beach. Calif).*, vol. 39, no. 2, pp. 96–97, Feb. 2006, doi: 10.1109/MC.2006.47.

[18] K. M. Sudar, P. Deepalakshmi, K. Ponmozhi, and P. Nagaraj, "Analysis of Security Threats and Countermeasures for various Biometric Techniques," in *2019 International Conference on Clean Energy and Energy Efficient Electronics Circuit for Sustainable Development, INCCES 2019*, Dec. 2019, doi: 10.1109/INCCES47820.2019.9167745.

[19] K. Sundararajan and D. L. Woodard, "Deep learning for biometrics: A survey," *ACM Comput. Surv.*, vol. 51, no. 3, May 2018, doi: 10.1145/3190618.

[20] H. Mehraj and A. H. Mir, "A Survey of Biometric Recognition Using Deep Learning," *EAI Endorsed Trans. Energy Web*, vol. 8, no. 33, pp. 1–16, 2021, doi: 10.4108/eai.27-10-2020.166775.

[21] T. Gwyn, K. Roy, and M. Atay, "Face recognition using popular deep net architectures: A brief comparative study," *Futur. Internet*, vol. 13, no. 7, p. 164, Jun. 2021, doi: 10.3390/fi13070164.

[22] A. N. Razzaq, R. Ghazali, and N. K. El Abbadi, "Face Recognition – Extensive Survey and Recommendations," Jul. 2021, pp. 1–10, doi: 10.1109/icoten52080.2021.9493444.

[23] N. Poh and J. Korczak, "Hybrid biometric person authentication using face and voice features," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2001, vol. 2091 LNCS, pp. 348–353, doi: 10.1007/3-540-45344-x_51.

[24] K. K. A. Ghany and H. M. Zawbaa, "Hybrid biometrics and watermarking authentication," in *Securing Government Information and Data in Developing Countries*, IGI Global, 2017, pp. 37–61.

[25] F. Richardson, D. Reynolds, and N. Dehak, "Deep neural network approaches to speaker and language recognition," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1671–1675, Oct. 2015, doi: 10.1109/LSP.2015.2420092.

[26] T. H. Iwan, H. M. Fahrezy, and R. P. Merliasari, "The Design and the Implementation of Security System Office Door Using Raspberry Pi Face Detection," Mar. 2020, pp. 307–311, doi: 10.2991/assehr.k.200303.074.

[27] K. KumarNagwanshi and S. Dubey, "Biometric Authentication using Human Footprint," *Int. J. Appl. Inf. Syst.*, vol. 3, no. 7, pp. 1–6, Aug. 2012, doi: 10.5120/ijais12-450568.

[28] R. Want, "An introduction to RFID technology," *IEEE Pervasive Computing*, vol. 5, no. 1. pp. 25–33, Jan. 2006, doi: 10.1109/MPRV.2006.2.

[29] A. Noer, Z. B. Hasanuddin, and D. Djamaluddin, "Implementation of RFID based raspberry Pi for user authentication and offline intelligent payment system," in *QiR 2017 - 2017 15th International Conference on Quality in Research (QiR): International Symposium on Electrical and Computer Engineering*, Dec. 2017, vol. 2017-Decem, pp. 251–255, doi: 10.1109/QIR.2017.8168491.

[30] X. Zhang, D. Cheng, P. Jia, Y. Dai, and X. Xu, "An Efficient Android-Based Multimodal Biometric Authentication System with Face and Voice," *IEEE Access*, vol. 8, pp. 102757–102772, 2020, doi: 10.1109/ACCESS.2020.2999115.

[31] M. A. M. El-Bendary, H. Kasban, A. Haggag, and M. A. R. El-Tokhy, "Investigating of nodes and personal authentications utilizing multimodal biometrics for medical application of WBANs security," *Multimed. Tools Appl.*, vol. 79, no. 33–34, pp. 24507–24535, Jun. 2020, doi: 10.1007/s11042-020-08926-2.

[32] I. B. Venkateswarlu, J. Kakarla, and S. Prakash, "Face mask detection using MobileNet and global pooling block," in *4th IEEE Conference on Information and Communication Technology, CICT 2020*, Dec. 2020, doi: 10.1109/CICT51604.2020.9312083.

[33] Y. Zhou, Y. Liu, G. Han, and Y. Fu, "Face Recognition Based on the Improved MobileNet," in *2019 IEEE Symposium Series on Computational Intelligence, SSCI 2019*, Dec. 2019, pp. 2776–2781, doi: 10.1109/SSCI44817.2019.9003100.