ICIC Express Letters
Volume **16**, Number **10**, October **2022**

ICIC International ©2022 ISSN 1881-803X
pp. **1027–1035**

DEVELOPMENT AND EVALUATION OF BIOMETRIC AUTHENTICATION SYSTEM BASED ON FACIAL AND VOICE RECOGNITION

Meennapa Rukhiran[1], Sorapak Pukdesree[2] and Paniti Netinant[2,*]

[1]Department of Social Technology
Rajamangala University of Technology Tawan-ok
131 Amphoe Khao Khitchakut, Chanthaburi 22210, Thailand
meennapa_ru@rmutto.ac.th

[2]College of Innovation Digital Technology
Rangsit University
52/347 Phaholyothin Road, Lak-Hok, Pathumthani 12000, Thailand
sorapak.p58@rsu.ac.th; *Corresponding author: paniti.n@rsu.ac.th

Received October 2021; accepted January 2022

Abstract. *End users are required to submit a variety of strong usernames and passwords to authenticate with various online applications. Biometric features can serve as a more effective substitute for a traditional restriction technique. Recently, customers reported feeling at ease and secured using their biometric to enter into the system. Biometrics has been utilized for security purposes in a variety of applications recently. However, multimodal biometrics that employs feature-level fusion integrates data from various sources, regardless of how good or bad the combination is. Voice biometric authentication is more comfortable for end-users since it does not require as many motions. It is fascinating to investigate the effect of feature fusions on the matching performance of a multimodal biometric system using face and voice biometrics. This research seeks to design and compare two biometrics that can be accurately integrated to determine authentication accuracy. We propose to analyze biometric authentication using the authentication accuracy matrix. The most advanced biometric authentication with a high degree of accuracy has been devised. On web-based applications, face authentication is the recommended biometric authentication method. Chi-Square statistics analysis findings help check and validate significant levels.*

Keywords: Biometrics, Authentication, Evaluation, Face recognition, Voice recognition

1. **Introduction.** Username and password verification is the most common method of user authentication. The username and password are standard and straightforward [1,2], but it has several drawbacks for both users and service providers. A user may have some online or offline services that require a password combination. As a result, even a complex password with uppercase and lowercase letters, special characters, and numbers may be challenging to remember. To crack a password, an attacker uses a massive dataset of leaked passwords [3].

Biometric characteristics such as the face, fingerprint, iris, ear, and hand can be used to authenticate services with more excellent reliability, simplicity, and integrity [4-6] than conventional username and password authentication. Unimodal biometrics has limitations for safe identification and verification [7,8]. Mason et al. [9] stated that biometric technology is rapidly evolving in healthcare. A new method for combining periocular biometrics and patient indexes in a healthcare information system is described. Environmental deterioration, sensor noise, and other factors limit each biometric modality. User information must be kept, accessed, retrieved, and managed securely, according to Shakil et al. [10].

DOI: 10.24507/icicel.16.10.1027

Signatures are used in a cloud-based behavioral biometric authentication platform. Cloud computing services have been researched for their biometric identification capabilities in a variety of fields. Face recognition using cloud services can result in much-reduced latency values compared to local servers [11].

Multimodal biometrics manage to avoid a single point of failure and make biometric systems more resilient [12-14]. Srivastva et al. [15] proposed an ECG-based biometric recognition system, representing the heart's electrical activity using electrode devices attached to the skin. PlexNet, a stacking model, was created by combining four fine-tuned models. PlexNet is evaluated using two public datasets. The intended biometric recognition system has a 99.66 percent accuracy score in testing. Face and palmprint biometrics identification is accomplished by the integration of non-stationary features [16]. With new goals, the algorithm can improve recognition rates. Kumari and Thangaraj [7] proposed a cloud-based system for multimodal biometric feature selection. Multimodal biometrics combines facial and fingerprint recognition. The recommended feature selection method used ant colony optimization. Particle swarm optimization is used to extract and classify features. Sarier [17] proposed a biometric identification method for Mobile Edge Computing (MEC). MEC is a secure cloud computing platform that a secure protocol defends against intruders. The new multimodal biometric authentication protocol allows for completely encrypted matching, with the final merged matching score hidden from other computers. Compared to traditional cloud computing, this method reduces latency by 75%.

Figure 1 shows our previous system for biometric cloud service authentication using face and voice [18]. Our framework consists of three layers: end-user, application server, and database server layers. The end-user layer is the user layer that interacts with or utilizes the system. During registration or recognition, the system collects user information such as voice, face, and password. The application server layer contains system applications

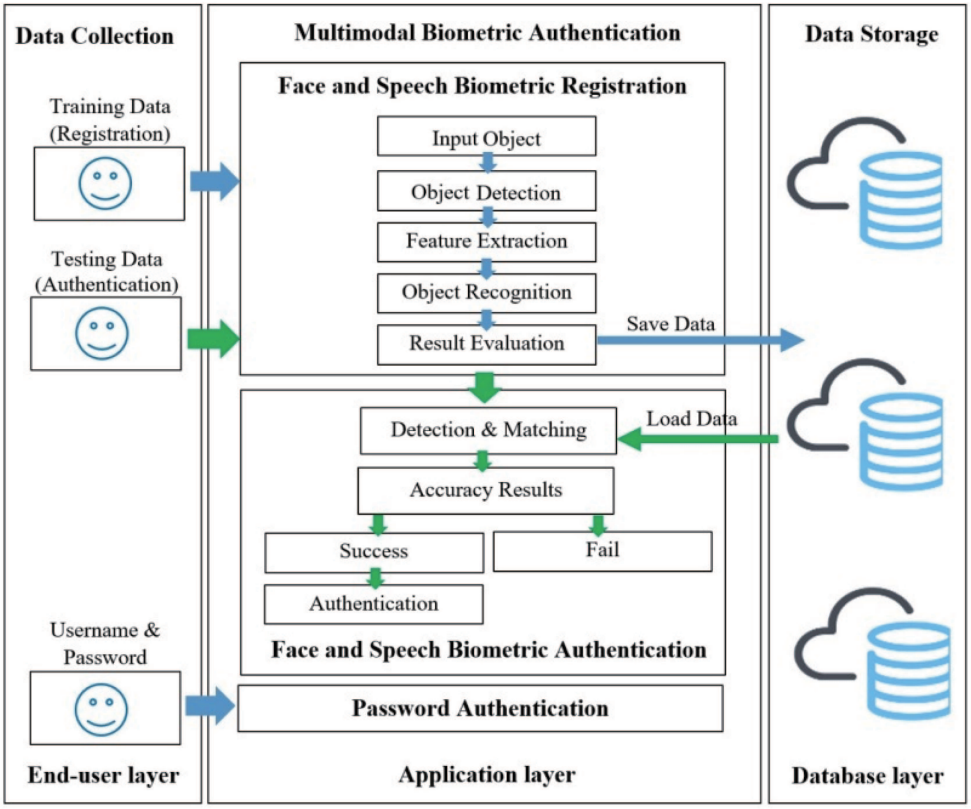

FIGURE 1. Framework of biometrics authentication cloud services

like speaker and facial recognition. The program runs on a virtual machine and a cloud VPS via a docker container. This application was developed using web technologies and is cross-platform. The database server layer manages data storage, inserting, updating, accessing, and deleting. The system's database is also hosted on a virtual private server in the cloud.

The proposed conceptual framework describes component layer concepts and procedures for a practical authentication system. However, multimodal biometrics combining face and voice authentication has been rarely found in most researches. Face and voice recognition may be more pleasant for end-users without requiring much movement than the well-known face and fingerprint recognition as well as face and ear recognition. The accuracy rates [19,20] of facial and speech recognition for test schedule authentication systems are essential concerns when performing multimodal biometrics authentication. By utilizing multimodal biometric authentication of facial and voice recognition, the primary contribution is to achieve the highest accuracy possible for genuine recognition. This research aims to develop and evaluate practically accurate factors of facial and voice authentication systems.

This article continues as follows. Section 2 describes related research. Section 3 describes the proposed system. Section 4 summarizes the results. Section 5 presents the conclusion.

2. Material and Method.

2.1. **Type of research.** This research is exploratory in nature and takes a systematic approach to research, exploring, and inventing in a trustworthy manner. This research aims to develop a web-based biometric authentication system for undergraduate students' online examinations that makes use of cloud-based services. Students may now learn from anywhere or at any time, thanks to advancements in network technology, notably Internet technology. However, a previous process for authenticating students taking online tests will be in place. This study aims to provide multimodal biometric authentication for web-based biometric cloud services that are compatible with mobile devices.

2.2. **Research sample.** The demographic for this study is students enrolled in the Computer Science and Technology Department at Bangkok University's School of Information Technology and Innovation during the first semester of 2021. The population is around four-hundred undergraduate students. Purposive sampling was used in this study to choose students from the computer organization course, which is appropriate for the current circumstances in our environment. This course had around eighty students enrolled per class. Seventy-two genuine students and eight phony students are included in the sampling.

2.3. **Research instrument and procedure.** The study's instrument is a web application written in Python 3.9 and based on the Django framework. This web application makes use of the sqlite3 database management system. This web application may be accessed and used on iOS, Android devices, and a personal computer equipped with a web browser. Before students take the final examinations via online assessments, researchers organize meetings with them.

2.4. **Data collection.** This research collects data through experimental testing. The sampling will enable anyone to use the multimodal biometric authentication system on their devices. The sampling will be provided with the information necessary to log into the system on a specific day and time. To ensure the accuracy of biometric information, the outcome variables of each student's training and testing processes must achieve a 65 percent accuracy rate for face and fingerprint recognition findings. Eighty students are engaged in this study project, which is divided into two groups. Each sample group has

forty students, thirty-six authentic and four fictitious. We create fictitious students with similar looks or voices to those on the biometric database's student identification.

2.5. **Data analysis.** The gathered data is an enumeration count of multimodal biometric authentication matches and mismatches. Qualitative data is gathered. Because of this, the Chi-Square test is used to analyze and compare one sample group's statistical data. The model's prediction accuracy is compared to the actual values in the table. This matrix is used in predictive analytics, a branch of statistics that uses data to forecast trends and behavior patterns. Predictive analytics uses data modeling, machine learning, artificial intelligence, deep learning algorithms, and data mining. The framework was evaluated using the metrics below.

1) A statistical metric's authentication accuracy score is the ratio of accurate prediction results to comprehensive testing. Equation (1) can be expressed of the accuracy.

$$\begin{aligned} &Authentication\ Accuracy\ (ACC) \\ &= \frac{Genuine\ pass\ authentication + Fake\ fail\ authentication}{\sum\ (Total\ Authentications)} \end{aligned} \tag{1}$$

2) Authentication precision denotes the precision of the prediction's outcome. Authentication precision can be seen in two ways: positively or negatively. Authentication precision positive refers to the precision with which the member aspect is predicted, whereas negative refers to the precision with the nonmember aspect is predicted. Equations (2) and (3) can be used to express the authentication accuracy score.

$$\begin{aligned} &Authentication\ Precision\ Positive \\ &= \frac{Genuine\ pass\ authentication}{(Genuine\ pass\ authentication + Fake\ fail\ authentication)} \end{aligned} \tag{2}$$

$$\begin{aligned} &Authentication\ Precision\ Negative \\ &= \frac{Fake\ fail\ authentication}{(Fake\ fail\ authentication + Genuine\ fail\ authentication)} \end{aligned} \tag{3}$$

3) An authentication recall measures the prediction results' accuracy compared to the actual authentication values; it may be thought of as the percentage of correcting predictions made by members named authentication recall positive and the percentage of correcting predictions made by nonmembers named authentication recall negative. Equations (4) and (5) can be used to express the recall formula.

$$\begin{aligned} &Authentication\ Recall\ Positive \\ &= \frac{Genuine\ pass\ authentication}{(Genuine\ pass\ authentication + Fake\ pass\ authentication)} \end{aligned} \tag{4}$$

$$\begin{aligned} &Authentication\ Recall\ Negative \\ &= \frac{Fake\ fail\ authentication}{(Fake\ fail\ authentication + Genuine\ fail\ authentication)} \end{aligned} \tag{5}$$

4) The F-test of authentication accuracy score is calculated by averaging the accuracy and recall values. The metric score is sometimes referred to as the harmonic mean. The F-test of authentication accuracy metric score has two values: F-test score negative and F-test positive. Assume the score is high, indicating the model is doing well. Equation (6) can be used to express the F-test of authentication accuracy scoring.

$$\begin{aligned} &F\text{-}test\ Authentication\ Accuracy \\ &= \frac{2 * Genuine\ pass\ authentication}{(Fake\ fail\ authentication + Fake\ pass\ authentication)} \end{aligned} \tag{6}$$

3. **Proposed System Implementation.** The system was created by researchers utilizing a variety of design techniques. On the other hand, this research includes design tools such as a use case diagram and a flow diagram. According to the use case model depicted in Figure 2, the system interacts with three actors: students, teachers, and administrators. Students can register, retrieve, edit, display, and learn courses as well as tests using this system. Teachers can create an account in the system, update their information, show courses and tests, and log into examinations. Finally, administrators can create, retrieve, edit, and remove courses and tests, as well as assign teachers to the course.

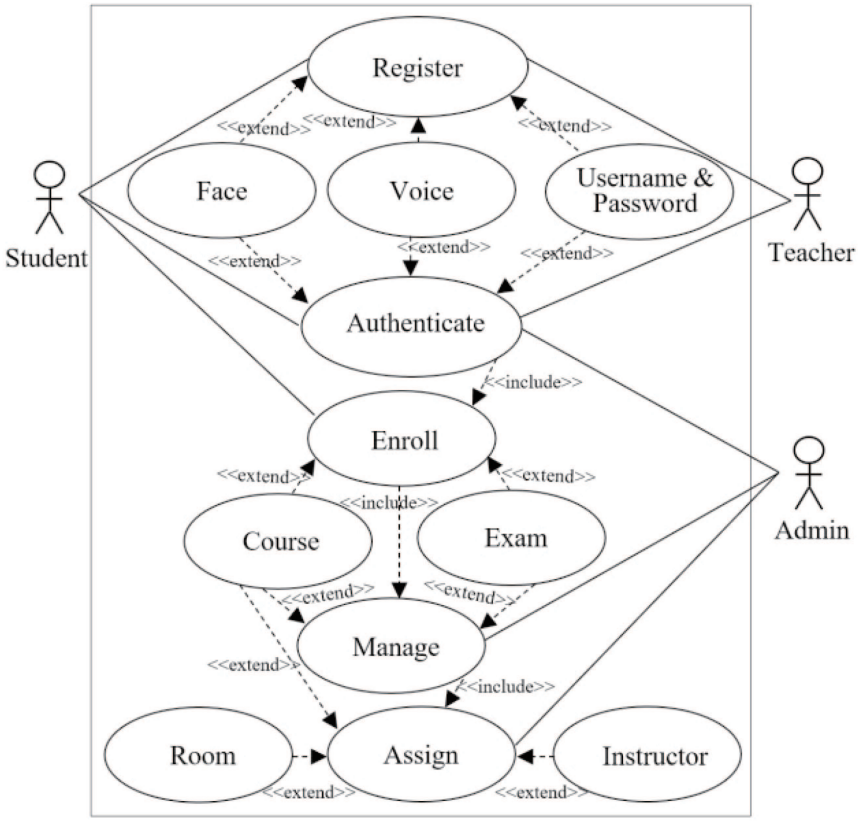

FIGURE 2. Use case biometrics authentication system

To begin, students can register biometric into the system using their devices, such as personnel and laptop computers and smartphones, as seen in Figure 3. Students must enter their facial biometric information into the system authentication through the camera; the facial biometric information is stored in a temporary folder on the server. Then, students must enter their voice into the system using a microphone; the speech biometric is stored in a temporary folder on the server. When students click the login button, the system extracts facial biometric characteristics and compares them to biometric template stored in the system, therefore confirming facial biometric authentication. The OpenCV based on LBPH algorithm [5,21,22] produces a matching score for the face biometric that represents the degree of similarity between the comparisons. The system then obtains speech biometric characteristics and compares them to the voice biometric authentication performed by when students click the login button. The Mel-Frequency Cepstral Coefficients (MFCC) algorithm [23] generates a matching score for the compared speech biometric as a secret score indicating the similarity of the comparisons [24-26]. The technology then calculates the fusion scores by combining the secret face and voice biometrics scores. Finally, the system makes an automatic judgment based on the fusion scores obtained from the multimodal biometric authentications. If the fusion scores fall between the range of 60 percent, the system permits the learner to access a home page. On the

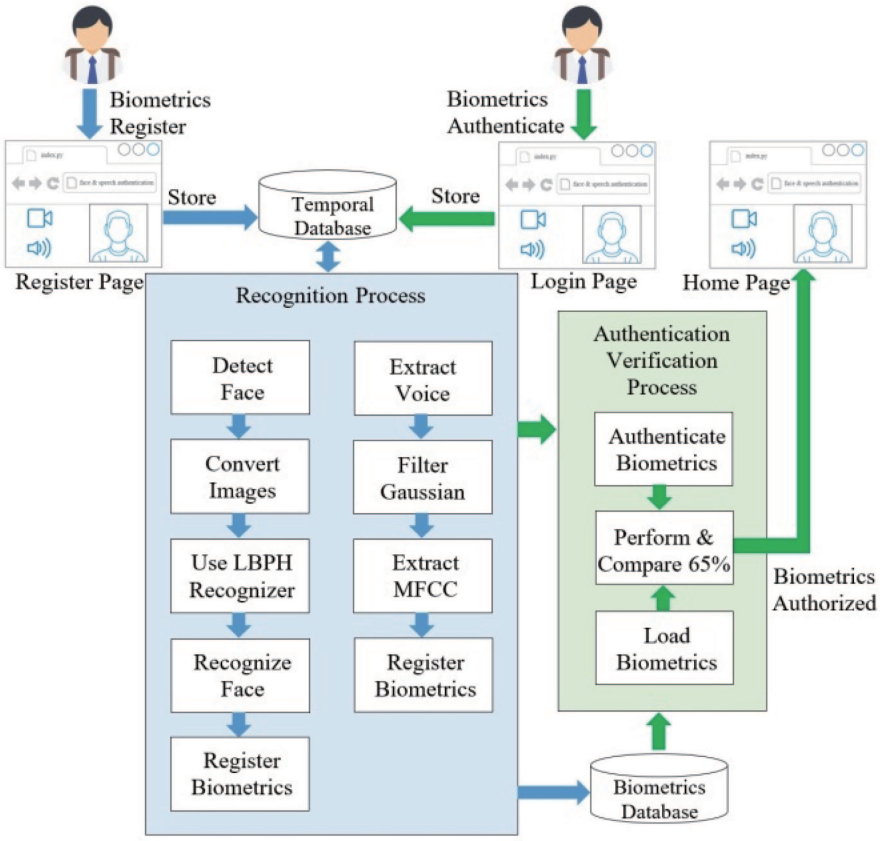

FIGURE 3. Biometrics authentication components

other hand, if the fusion scores fall beyond the system's threshold values range, the system decides to reject the student authentication. The student has the option of utilizing their password for login in a conventional manner.

4. **Performance and Result Analysis.** The suggested system's recognition accuracy and methodological validity have been evaluated. The sample group evaluates several techniques for biometric authentication. A sample group of eighty students is composed of seventy-two genuine students and eight phony students. Three steps comprise the examination of three techniques. The first phase is to train all legitimate students' facial and voice biometrics. The following phase secures the system using a variety of biometric authentication methods. For the face and voice authentication of a sample group provided in Table 1, the accuracy metrics of several identity authentications are computed.

TABLE 1. Chi-Square different statistics of face and voice recognition comparison

Model	Authentication metric		Chi-Square values		χ^2	*df*	*p*
	Negative	Positive	Negative	Positive			
Genuine student	13	59	0.3457	0.0933	4.3905	1	0.0361
Fake student	4	4	3.1117	0.8397			

Figure 4 illustrates the authentication accuracy rate for several biometrics. Equation (1)'s accuracy metric is used to calculate authentication and F-test accuracies. Face authentication has an accuracy of 0.9650, whereas the F-test has an accuracy of 0.9375. The Chi-Square χ^2, *df*, and *p* values are, respectively, 5.3333, 1, and 0.0209. The authentication accuracy of voice authentication is 0.8437, whereas the F-test authentication accuracy is 0.75. The Chi-Square χ^2, *df*, and *p* values are, respectively, 8.5714, 1, and 0.0034. Face

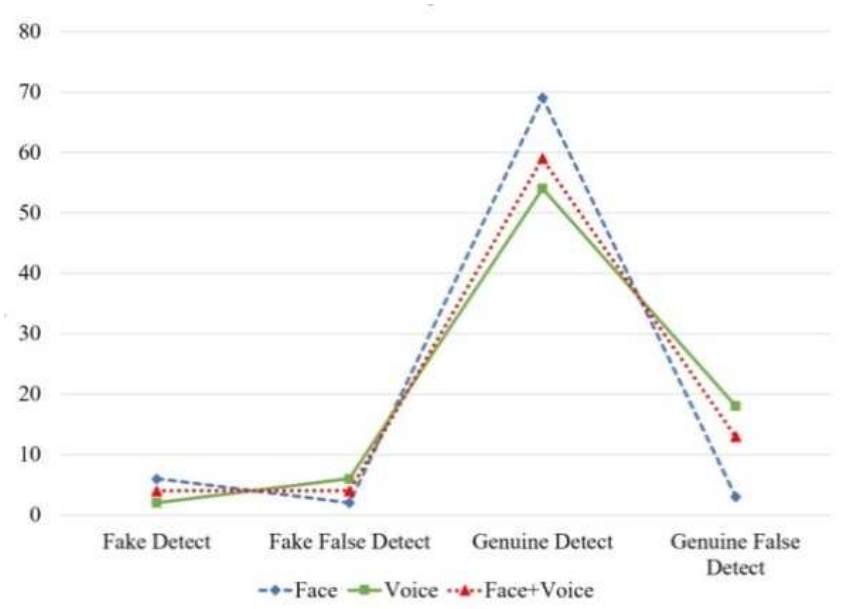

FIGURE 4. Factors accuracy rates of biometric authentications

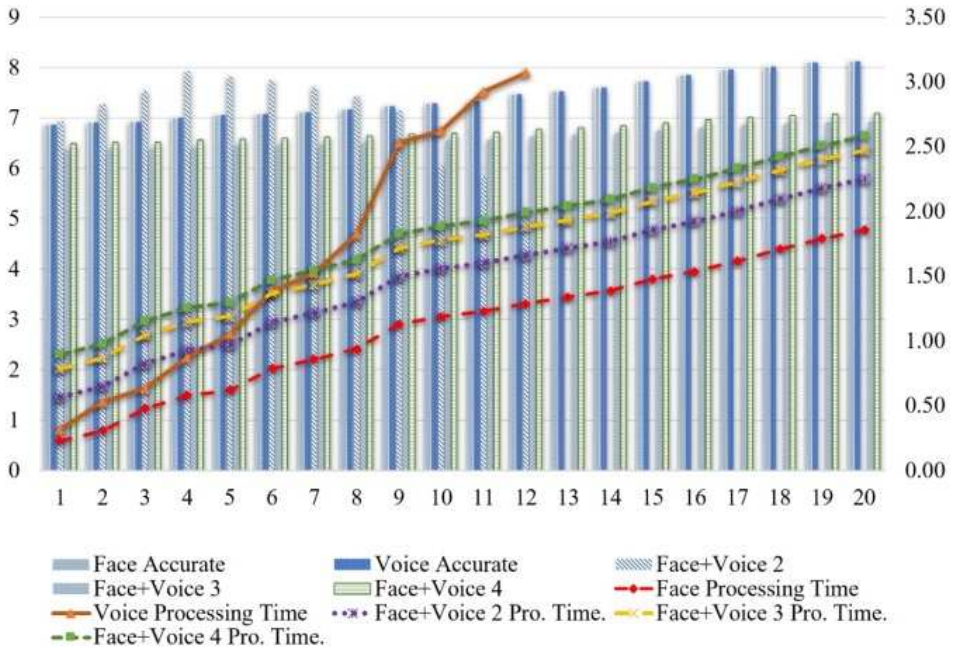

FIGURE 5. Accuracy rate and processing time per images and voice durations

and voice authentications have an accuracy of 0.8740, while the F-test has an accuracy of 0.7875. The Chi-Square χ^2, *df*, and p values are 4.3905, 1, and 0.0361, respectively, as shown in Table 1. As indicated in Figure 5, three essential assessing variables that impact biometric authentication are processing time, voice lengths, and authentication accuracy. Seconds are the unit of processing time, and the number of students is the unit of accuracy rate. The brief processing time appears to be preferable on the surface, particularly for face technology. The more face photos provided, the higher the authentication accuracy rate. The voice dramatically increases processing time as the duration of the speech increases. As a result, the longer the duration of the voice, the longer the processing time, and the lower the authentication accuracy rate. The optimal duration of a voice is around four seconds.

5. **Conclusions.** We propose a framework to reflect the compatibility of three system layers in this article and the creation of facial and voice authentication systems for cloud services that operate within the exam system. The web-based program detects and matches the students' faces using real-time optimized OpenCV libraries based on the LBPH algorithm. To identify the students' speech, the MFCC algorithm is used. Additionally, we assess the authentication accuracy rates of unimodal and multimodal biometrics, emphasizing the face and voice. The assessment findings are based on a random sample of physical and bogus detections. The Chi-Square test is used to verify statistical connections. The statistical metric of face authentication accuracy score delves into the most precise kind of biometric authentication. Biometric authentication elements demonstrate a potentially adaptable and compatible system for optimizing additional face images and voice durations. Further study and refinement of authentication accuracy rates and processing methodologies are required.

REFERENCES

[1] A. Czeskis, M. Dietz, T. Kohno, D. Wallach and D. Balfanz, Strengthening user authentication through opportunistic cryptographic identity assertions, *Proc. of 2012 ACM Conference on Computer and Communications Security*, Raleigh, USA, pp.404-414, 2012.

[2] B. R. Naidu and M. S. P. Babu, Biometric authentication data with three traits using compression technique, HOG, GMM and fusion technique, *Data in Brief*, vol.18, pp.1976-1986, 2018.

[3] J. Blocki and A. Datta, CASH: A cost asymmetric secure Hash algorithm for optimal password protection, *Proc. of IEEE 29th Computer Security Foundations Symposium (CSF)*, Lisbon, Portugal, pp.371-386, 2016.

[4] J. Blasco, T. M. Chen, J. Tapiador and P. Peris-Lopez, A survey of wearable biometric recognition systems, *ACM Computing Surveys*, vol.49, no.3, pp.1-35, 2016.

[5] M. Rukhiran, P. Netinant and T. Elrad, Effecting of environmental conditions to accuracy rates of face recognition based on IoT solution, *Journal of Current Science and Technology*, vol.10, no.1, pp.21-33, 2020.

[6] N. M. G. Al-Saidi, A. J. Mohammed, R. J. Al-Azawi and A. H. Ali, features via fractal functions for authentication protocols, *International Journal of Innovative Computing, Information and Control*, vol.15, no.4, pp.1441-1453, 2019.

[7] P. Kumari and P. Thangaraj, A fast feature selection technique in multi modal biometrics using cloud framework, *Microprocessors and Microsystems*, vol.79, no.2, DOI: 10.1016/j.micpro.2020.103277, 2020.

[8] F. Kausar, Iris based cancelable biometric cryptosystem for secure healthcare smart card, *Egyptian Informatics Journal*, vol.22, no.4, pp.447-453, 2021.

[9] J. Mason, R. Dave, P. Chatterjee, I. G. Allen, A. Esterline and K. Roy, An investigation of biometric authentication in the healthcare environment, *Array*, vol.8, DOI: 10.1016/j.array.2020.100042, 2020.

[10] K. A. Shakil, F. J. Zareen, M. Alam and S. Jabin, BAMHealthCloud: A biometric authentication and data management system for healthcare data in cloud, *Journal of King Saud University – Computer and Information Sciences*, vol.32, pp.57-64, 2020.

[11] M. D. M. Soares, *MEC vs MCC: Performance Analysis of Real-Time Applications*, Master Thesis, Polytechnic Institute of Porto, Portugal, 2019.

[12] S. Wong-In and P. Netinant, Revised software model design for biometric examiner personal verification system, *Proc. of 2017 International Conference on Information Technology*, New York, USA, pp.237-242, 2017.

[13] S. Sarhan, S. Alhassan and S. Elmougy, Multimodal biometric systems: A comparative study, *Arabian Journal for Science and Engineering*, vol.42, no.2, pp.443-457, 2017.

[14] I. Stylios, S. Kokolakis, O. Thanou and S. Chatzis, Behavioral biometrics & continuous user authentication on mobile devices: A survey, *Information Fusion*, vol.66, pp.76-99, 2021.

[15] R. Srivastva, A. Singh and Y. N. Singh, PlexNet: A fast and robust ECG biometric system for human recognition, *Information Sciences*, vol.558, pp.208-228, 2021.

[16] M. I. Ahmad, W. L. Woo and S. Dlay, Non-stationary feature fusion of face and palmprint multimodal biometrics, *Neurocomputing*, vol.177, pp.59-61, 2016.

[17] N. D. Sarier, Multimodal biometric authentication for mobile edge computing, *Information Sciences*, vol.573, pp.82-99, 2021.

[18] S. Pukdesree and P. Netinant, Biometric authentication for accessing cloud computing services using iPhone, *Advances in Intelligent Systems and Computing*, vol.463, pp.209-216, 2016.

[19] W. Yang, S. Wang, J. Hu, G. Zheng and C. Valli, Security and accuracy of fingerprint-based biometrics: A review, *Symmetry*, vol.11, no.2, DOI: 10.3390/sym11020141, 2019.

[20] A. Ometov, V. Petrov, S. Bezzateev, S. Andreev, Y. Koucheryavy and M. Gerla, Challenges of multi-factor authentication for securing advanced IoT applications, *IEEE Network*, vol.33, no.2, pp.82-88, 2019.

[21] M. Ahsan, Y. Li, J. Zhang, T. Ahad and S. Yazdan, Face recognition in an unconstrained and real-time environment using novel BMC-LBPH methods incorporates with DJI vision sensor, *Journal of Sensor and Actuator Networks*, vol.9, no.4, pp.1-12, 2020.

[22] K. Deepti, Implementation of safety survilliance system with face recognition using IoT, *International Journal of Engineering Research & Technology*, vol.9, no.5, pp.434-437, 2021.

[23] Q. Li, Y. Yang, T. Lan, H. Zhu, Q. Wei, F. Qiao, X. Liu and H. Yang, MSP-MFCC: Energy-efficient MFCC feature extraction method with mixed-signal processing architecture for wearable speech recognition applications, *IEEE Access*, vol.8, pp.48720-48730, 2020.

[24] R. D. Kumar, A. B. Ganesh and S. S. Kala, Speaker identification system using Gaussian Mixture Model and Support Vector Machines (GMM-SVM) under noisy conditions, *Indian Journal of Science and Technology*, vol.9, no.19, DOI: 10.17485/ijst/2016/v9i19/93870, 2016.

[25] E. S. Wahyuni, Arabic speech recognition using MFCC feature extraction and ANN classification, *Proc. of the 2nd International conferences on Information Technology, Information Systems and Electrical Engineering*, Yogyakarta, Indonesia, pp.22-25, 2017.

[26] N. P. Patel and A. Kale, Optimize approach to voice recognition using IoT, *Proc. of International Conference on Advances in Communication and Computing Technology*, pp.251-256, 2018.