DEÜ FMD 25(74), 481-489, 2023

Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi
Dokuz Eylul University Faculty of Engineering Journal of Science and Engineering

Basılı/Printed ISSN: 1302-9304. Elektronik/Online ISSN: 2547-958X

Authentication with face recognition and sign language using ESP32-CAM

ESP32-CAM kullanarak yüz tanıma ve işaret dili ile kimlik doğrulama

Zafer Yalçın[1], Oktay Türkdağlı[1], Gökhan Dalkılıç[1], Ömer Aydın[2*]

[1] Dokuz Eylül Üniversitesi Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, İzmir, TÜRKİYE
[2] Manisa Celal Bayar Üniversitesi Mühendislik Fakültesi, Elektrik Elektronik Mühendisliği Bölümü, Manisa, TÜRKİYE

Sorumlu Yazar / Corresponding Author *: omer.aydin@cbu.edu.tr

Geliş Tarihi / Received: 21.09.2022 *Araştırma Makalesi/Research Article*
Kabul Tarihi / Accepted: 30.10.2022 *DOI:10.21205/deufmd.2023257417*
Atıf şekli/ How to cite: YALÇIN, Z., TÜRKDAĞLI, O., DALKILIÇ, G., AYDIN, Ö. (2023). Authentication with face recognition and sign language using ESP32-CAM. DEUFMD, 25(74), 481-489.

Abstract

Authentication is the confirmation of the accuracy of the data piece that any institution, person or system accepts as correct. Many methods are used for the authentication process. Some of these are the methods using biometric data such as face authentication, fingerprint authentication, and iris authentication. In this article, the way to create a secure system using face recognition and sign language as an authentication method is discussed and an application using ESP32-CAM is developed and tested. The results show that secure authentication cannot be achieved with facial recognition and sign language. The developed system is low cost and easy to implement. With this system, authentication can be done without requiring any physical contact, and it can be used for personal security and entrance and exit. Sign language, which is frequently used by hearing-impaired individuals, can play an active role as authentication. With low-cost modules such as ESP32, it will be an alternative to authentication in almost every environment.

Keywords: *Authentication, Face recognition, Sign language, ESP32-CAM*

Öz

ESP32-CAM Kullanılarak Yüz Tanıma ve İşaret Dili ile Kimlik Doğrulama Kimlik doğrulama, herhangi bir kurum, kişi veya sistemin doğru kabul ettiği veri parçasının doğruluğunun teyididir. Kimlik doğrulama işlemi için birçok yöntem kullanılmaktadır. Bunlardan bazıları yüzle kimlik doğrulama, parmak iziyle kimlik doğrulama ve iris kimlik doğrulama gibi biyometrik verileri kullanan yöntemlerdir. Bu makalede, kimlik doğrulama yöntemi olarak yüz tanıma ve işaret dilini kullanarak güvenli bir sistem oluşturmanın yolu tartışılmakta ve ESP32-CAM kullanan bir uygulama geliştirilip test edilmektedir. Sonuçlar, yüz tanıma ve işaret dili ile güvenli kimlik doğrulamanın sağlanamayacağını göstermektedir. Geliştirilen sistem düşük maliyetli ve uygulaması kolaydır. Bu sistem ile herhangi bir fiziksel temas gerektirmeden kimlik doğrulama yapılabilmekte, kişisel güvenlik ve giriş çıkış için kullanılabilmektedir. İşitme engelli bireylerin sıklıkla kullandığı işaret dili, kimlik doğrulama olarak aktif rol oynayabilir. ESP32 gibi düşük maliyetli modüller ile neredeyse her ortamda kimlik doğrulamaya alternatif olacaktır.

Anahtar Kelimeler: *Kimlik doğrulama, Yüz tanıma, İşaret dili, ESP32-CAM*

1. Introduction

Authentication in online transactions has become more important than ever before. Authentication is the confirmation of the accuracy of the data piece that any institution, person or system accepts as correct. While most critical transactions require at least one authentication method, there are many that require two or more authentication methods. For example, you only need a student card to enter the school, this is a one-factor authentication method. Another example is when using your credit card, a password is requested with your card, which is a two-factor authentication method. Authentication methods can be done with more traditional methods such as signature, fingerprint image or identity card, as well as with more innovative methods such as face recognition, palm recognition or voice recognition. As a result of advances in artificial intelligence, the facial recognition system has reached a level that can distinguish even twin brothers from each other. However, it is still a structure that is open to deception with humanoid visuals. Therefore, when this verification method is used, most of the time people control the verification process instead of automated systems. Since the face recognition system is open to deception, another verification method, sign language visual authentication is used together with face recognition. Therefore, it is aimed to develop a system that uses both methods together and to investigate the results. In another word, this article presents the design of face recognition and sign language recognition as authentication techniques.

The rest of this article is organized as follows. Section 2 discusses previous and related studies in the literature. In Section 3, proposed work is given. Section 4 gives test results and discussions. The paper is concluded in Section 5. Finally, possible future works are given in Section 6.

2. Related Works

Today, many different studies are actively carried out in the name of identity verification. Biometric data is very reliable for this and deterrents for attackers. Nowadays, mobile devices can take high quality images, increased their processing power, etc. With technological advances such as facial recognition and authentication, it has become a part of daily life. Corporate companies and government agencies have already used fingerprints as a biometric authentication for many years. In addition, in some sectors, verification is done with iris. Today, some companies ask people to upload their photos, scan their faces, etc. for authentication. But users may request steps to ensure security by guaranteeing the protection of their personal data. As a positive alternative to such situations, the sign language and face recognition can also be used actively as a multi-factor authentication in order to verify the identity of the person.

Among the authentication methods, face recognition and sign language authentication provide a wide spectrum of possibilities that can be applied to many areas in daily life. For example, there are studies on authentication in computer-based exams using fingerprint as a biometric authentication method [1].

Considering the existing studies on face recognition and sign language authentication, it seems that some research has been done to make them unique and preferable. It has been suggested that hand gestures contain some individual nuances, that they are personal such as fingerprints and iris, and that it is a new biological authentication model that can be used for authentication even with low-cost cameras [2].

Visual cryptography is used for face identification. According to the research carried out by Ibrahim et al., it was revealed that a multifactor authentication system to be preferred instead of visual cryptography provides a faster verification opportunity and increases security with multiple verifications [3]. Using multiple authentication methods, especially using biometric data, is a good measure to deter attackers. In the developed study, the data is encrypted in three different templates, increasing the security, but this brings low accuracy and efficiency [3]. At this point, it is necessary to test the efficiency and accuracy rate when authentication is applied using both face and sign language at the same time.

Studies showing that the nearest-neighbor approach can be used when there is facial recognition similarity. A study has been developed considering the security of data by using secret sharing in public transportation [4]. In cases where face and sign language are considered as biometric data, data can be

confirmed for verification using visual sharing neural network (VSNN) [5].

There are many studies on the use of biometric data for authentication processes on Internet of things (IoT) devices. Mayut Badgujar et al. published a study in 2022 on automatic door opening with face and voice recognition using the convolutional neural network (CNN) algorithm. In this study, they succeeded in making it work using the Haar-cascade study method [6]. They used a three-step method with CNN to recognize a person's face. In this study, this method can be used as a basis for both face recognition and hand gesture recognition.

In a project developed for Malaysian sign language, sign language recognition operations were successfully performed using CNN algorithms with the ESP32 module and conversion from sign language to text was achieved [7].

In another study, it was tried to detect the movements with the Arduino NANO and the sensors placed on a glove instead of reading the frame from camera, and these signs were tried to be recognized with the random forest classifier. It was found that the method was not suitable in terms of both cost and sustainability [8].

Another study with high accuracy rates was carried out to provide convenience for the elderly and disabled people and to authenticate through sign language. In this study, Raspberry Pi 3, which is a more advanced board than used in this project, was used [9].

In addition, as observed in another study, DG5-V sensor gloves were used instead of images to recognize Arabic sign language. Wearing gloves and making hand signals are expected to activate these sensors. In addition, the aim of the study is to transform sign language into sound. CNN technique was used in the signal-to-sound conversion process [10]. Unlike our project, it requires continuous reading from the sensors and instantaneous translation. Since our study requires authentication, we can see that the cost will decrease when both the operation and the sensors are considered.

A 4-layer security system has been developed in a study on digital door lock systems to access personal belongings. Face authentication, one-time password generation and authentication, voice authentication and password authentication methods are used in this system. The ESP32-CAM card used for face authentication was considered sufficient in this research, which prioritizes security. It is understandable that this card, which performed quite well in this study, is sufficient for our project [11].

Another good example of the usage area of ESP32 is smart door lock systems. In a study conducted for this purpose, authorized and unauthorized users were detected by using ESP32 and ATMEGA328P together [12]. It is a useful study to compare the performance of the ESP32-CAM for face recognition at low dimensions.

3. Proposed Work

Under this title, the working principle, details and process steps of the proposed system are given. In addition, the test environment and test scenarios are explained.

3.1. Flow chart and design schemes

First of all, it is checked whether the camera module that is used in the study is turned on, and if the camera is working, a frame is taken. Secondly, the received frame is converted to gray image so that it can be used as an input to image recognition algorithms. It is checked whether this converted frame is a hand gesture or a face. If it is not a face or hand gesture that has been taught before, it is taught to the system and a frame is taken again so that the system is trained. Authentication is done with “n” hand movements that the person determines himself. The number of “n” to be used in authentication should be determined as a result of the tests. It is aimed to provide secure authentication with the person's own hand gestures and face identification. This flow is visualized on the flowchart in Figure 1.

The system to be used and required for a better understanding of the study is visualized in Figure 2 in a simplified manner. The system has an IoT system using a camera module, a visual recognition model to be used in face and sign language recognition, and an actor.

3.2. Defining system properties

Authentication was performed using the hand gestures listed in Table 1.

The hand gestures used in the system can be defined as follows;

Okay; the thumb and the forefinger are joined to form a circle and the remaining fingers are raised. Okay is a hand gesture that is used to express agreement or acceptance.

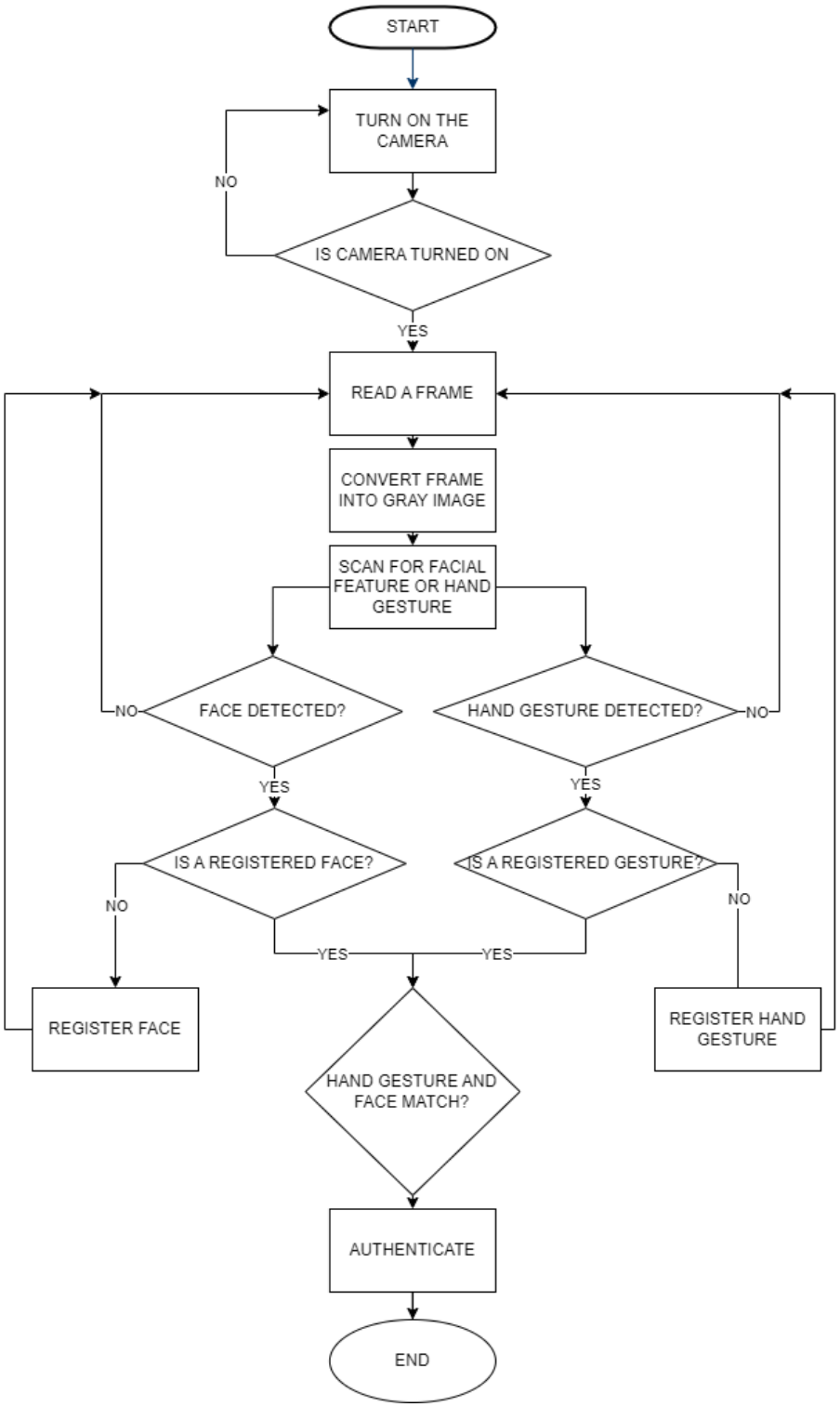

Figure 1. Flow chart of the system
Şekil 1. Sistemin akış şeması

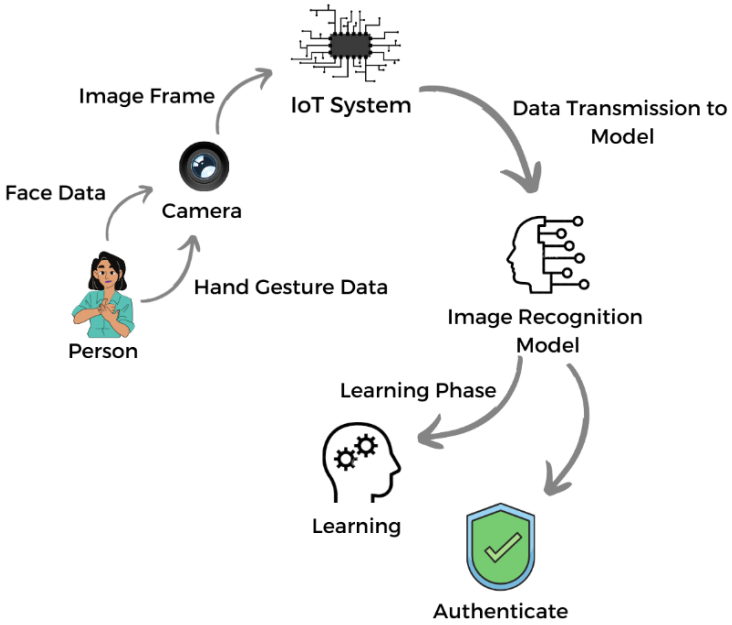

Figure 2. Design scheme
Şekil 2. Tasarım şeması

Table 1. Hand gestures

ID	Hand Gesture	Impression
1	okay	
2	peace	
3	thumbs up	
4	thumbs down	
5	call me	
6	stop	
7	rock	
8	live long	
9	fist	
10	smile	

Peace; the forefinger and the middle finger are straight to form the first letter of Victory and tighten the remaining fingers. Peace is a hand gesture that means peace, peaceful agreement.

Thumbs up; clench all the fingers except the thumb (raise the thumb). Thumbs up is a hand gesture that means approval.

Thumbs down; clench all the fingers except the thumb (thumb points down). Thumbs down is a hand gesture that means approval.

Call me; is performed by fingers drawn into a fist while the thumb and pinkie finger are sticking out. Call me means phone or contact me by phone.

Stop; is performed by extending any hand, palm upward. Stop is used to end or stop any situation.

Rock; forefinger and the pinky finger are raised and the remaining fingers are tightened. Rock is used to mean horned hand, which has become the symbol of the rock and roll music genre.

Live long; show inside of your hand by raising it, and also raising all your fingers and separate the long finger and the fourth finger by gathering pinky and fourth fingers as one group, and the forefinger and the long finger as the other group. Live long is used to salute the Vulcans in the popular television series Star Trek.

First; is performed by bending the fingers towards the palm and held there tightly. Fist is used to express "I care" in sign language.

Smile; is performed by raising the thumb and towards the chin, with the index finger pointing forward at mouth level and the other fingers facing the palm and the palm facing the face. Smile is used to express happy moments, and a smile.

3.3. Test environment and test scenarios

The test environment meets the minimum requirements for the project to run. Therefore, the test environment consists of the OV2640 development board containing the camera module, the ESP32 Future Technology Devices International (FTDI) programmer board that provides the connection between the computer and the module, the Arduino IDE to run the development boards, the Python IDE since artificial intelligence applications are written in Python, and a personal computer. Using a specially defined sign language for users creates a two-step authentication process.

The test scenarios determined for testing the system are given in Table 2. In the T1 scenario, the user registers to the system and then tries to log into the system with his/her face and correct signs. In the T2 scenario, the user registers in the system and tries to trick the system with wrong sign language letters. In the T3 scenario, the user registers in the system, then he/she tries to trick the system by showing another face instead of his/her face. T4, an unregistered user tries to access the system by showing his/her face and random signs. In the T5 scenario, the twin sister of a user who has already registered in the system tries to authenticate. It is aimed to test the system with these scenarios.

Table 2. Test Scenarios

ID	Task	Scenario
T1	Registered User Authentication	The user registers in the system, then the system reboots and the user tries to access the system by showing his/her face and correct signs.
T2	Registered User Sign Language Authentication	The user registers in the system, then the system reboots and the registered user tries to trick the system by showing the wrong sign language letters.
T3	Registered User Face Recognition Authentication	The user registers in the system, then the system restarts, and the user tries to trick the system by showing another face instead his/her face.
T4	Unregistered User Authentication	An unregistered user tries to access the system by showing his/her face and random signs.
T5	Twin Sisters Authentication	The twin sister of the previously registered user tries to authenticate.

During the tests, the user was informed as shown in Figure 3. For example, let's say there is a hand gesture "peace" defined for user X. After the user passes the face recognition stage, the user is expected to make the hand gesture (peace) that is pre-registered in the system. When the user

makes the “rock” sign as seen in Figure 3, an incorrect authentication entry is reported. If the registered hand gesture is performed correctly, a message will be notified that the authentication process has been successful.

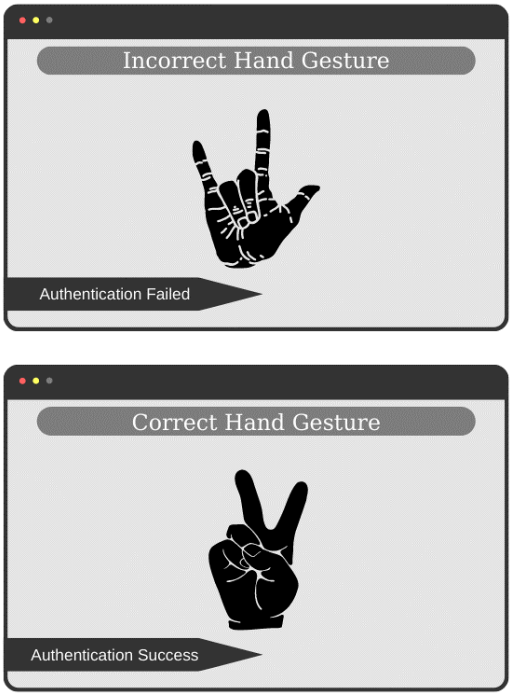

Figure 3. Authentication (Hand gesture step)
Şekil 3. Doğrulama (El hareketi adımı)

4. Test Results and Discussion

At this stage, depending on the test results, face recognition, sign language and test stages are described in the first order.

4.1. Face recognition

The face recognition part of our application works in the form of live verification of the person whose photo is given. The photo of the person to be authenticated is saved into the folder used as the database before the authentication can be done. Thus, the user's data to be verified is obtained. When the application is run, if the person is verified, the person's face is framed in green, and his name is written at the bottom of the frame. If the person could not be verified, the person's face is framed in red and the expression not recognized is written under it.

4.2. Sign language

The sign language recognition part of our application works in the form of live verification of the person whose photo is given with sign language. The photo of the gestures of the person to be authenticated with sign language is put in the folder used as the database before the authentication can be proceeded. Thus, the user's data to be verified is obtained. When the application is run, if the person is verified, the person's face is framed in green and "sign language verification successful" is written at the bottom of the frame. If the person cannot be verified, the person's face is framed in red and the phrase "sign language not recognized" is written under it.

4.3. Testing

Photos of the user's face and sign language were uploaded for testing as explained with the examples below. Then the application created for the user is run. During the tests, there is a hand gesture specially defined for the users. Authentication is not completed until this sign is made by the user. If the hand gesture defined for the X user is 'peace', the user receives false feedback from all movements other than this hand gesture, if it is true, the authentication is confirmed.

In the first scenario, codenamed T1, first the user's face photo was uploaded. Then, in order to verify this person who was introduced to the system, the live face of the person was shown with the OV2640 integrated ESP32 card and he was asked to make defined hand movements. The person's face was recognized, and the hand gesture was correct, the test was successful.

In the second scenario, codenamed T2, the user correctly showed his face and misrepresented his hand gestures. The user's face was verified, but the hand gestures could not be verified, resulting in the message "X hand gesture is incorrect".

In the third scenario, codenamed T3, a user whose face and hand gestures are not registered in the system has been selected. The user showed his face and hand gestures as in the T1 scenario. However, because the user was not registered in the system, the "Unrecognized Person" message was displayed on the screen.

In the fourth scenario, codenamed T4, a user whose face and hand gestures are not registered in the system has been selected. The user showed his face and hand gestures as in the T1 scenario. However, because the user was not

registered in the system, the "Unrecognized Person" message was displayed on the screen.

In the fifth scenario, codenamed T5, identical twins were selected to be introduced into the system. One of the brothers showed his face and hand gestures as in the T1 scenario. Facial recognition was done, as hand gestures in other scenarios already gave correct results. The system successfully performed facial recognition on identical twins, except when they were next to each other.

The tests performed were mostly successful in authentication with face recognition and personalized hand gestures. All existing hand gestures in the system were tried in order and it was observed that the correct hand gesture did not conflict with any other hand gesture.

5. Conclusion

The goal of this research is to provide two-factor low-cost biometric authentication method utilizing facial recognition and sign language via the ESP32-CAM and OV2640 camera. Authentication was accomplished by receiving common intermediate format (CIF) resolution (352 x 288) broadcasts via Wi-Fi using the camera modules employed. Users were registered to the system using their pictures (1 picture for each user) in this study, and a personalized hand gesture for each of them was determined. In the authentication phase, each frame received via the ESP32-CAM was simultaneously analyzed by the system written in Python, first detecting the face, and then determining whether the discovered face belongs to a registered user. If the detected face belongs to a registered user, it was intended that the hand gesture (without specifying which sign) would be performed by the user that must match with the already recorded hand gesture in the system. The user's login to the system is canceled if the hand gesture recognized using the joints on the hand is not the hand gesture registered for that user in the system. If the authentication process was completed successfully, a notification was displayed.

Test scenarios, including identical twins, were run, and the two-step authentication process was effectively performed. In the identical twins' case, in 10 different tests, the system gave wrong alarm only once. For the regular tests, the system has 87% success rate, and the failures were because of low light. It has been discovered that this low cost system, which was created with a relatively small camera module having a resolution of 352 x 288, can be utilized as a basic authentication mechanism.

It has been established that sign language, which hearing-impaired people are familiar with, can be utilized as an authentication mechanism. This very low cost system may be utilized for security with an identity verification system in dormitory lockers, personal spaces, access and exit to particular places, and similar areas.

5. Sonuç

Bu araştırmanın amacı, ESP32-CAM ve OV2640 kamera aracılığıyla yüz tanıma ve işaret dili kullanan iki faktörlü düşük maliyetli biyometrik kimlik doğrulama yöntemi ortaya koymaktır. Çalışmada kimlik doğrulama işlemi, kamera modülleri kullanılarak Wi-Fi aracılığıyla ortak ara format (CIF) resimler (352 x 288) alınarak gerçekleştirilmiştir. Bu çalışmada kullanıcılar kendilerine ait resimler (Her kullanıcı için 1 resim) ile sisteme kayıtlanmıştır ve her bir kullanıcı için kişiselleştirilmiş bir el hareketi belirlenmiştir. Kimlik doğrulama aşamasında, ESP32-CAM aracılığıyla alınan her çerçeve, Python'da yazılmış sistem tarafından eş zamanlı olarak analiz edildi, önce yüz algılandı ve ardından keşfedilen yüzün kayıtlı bir kullanıcıya ait olup olmadığı belirlendi. Tespit edilen yüz kayıtlı bir kullanıcıya ait ise, sistemde kayıtlı olan el hareketi ile eşleşmesi gereken el hareketinin (hangi işaret belirtilmeden) kullanıcı tarafından yapılması amaçlandı. Eldeki eklemler kullanılarak tanınan el hareketi, o kullanıcı için sistemde kayıtlı olan el hareketi değilse, kullanıcının sisteme girişine izin verilmez. Kimlik doğrulama işlemi başarıyla tamamlandıysa, bir bildirim görüntülenmektedir.

Tek yumurta ikizlerini içeren test senaryoları çalıştırılmış ve iki aşamalı kimlik doğrulama süreci etkin bir şekilde uygulanmıştır. Tek yumurta ikizlerine ait 10 farklı testte sadece bir kez sistem yanlış uyarı vermiştir. Normal testler için sistem %87 başarı oranına sahip ve başarısızlıkların sebebi düşük ışık olarak tespit edilmiştir. 352 x 288 çözünürlüğe sahip nispeten küçük bir kamera modülü ile oluşturulan bu düşük maliyetli sistemin temel bir kimlik doğrulama mekanizması olarak kullanılabileceği belirlenmiştir.

Bu çalışma ile işitme engellilerin aşina olduğu işaret dilinin kimlik doğrulama mekanizması olarak kullanılabileceği tespit edilmiştir.

Maliyeti oldukça düşük olan bu sistem, yurt dolaplarında, kişisel alanlarda, belirli yerlere giriş çıkışlarda ve benzeri alanlarda kimlik doğrulama sistemi olarak güvenlik amacıyla kullanılabilir.

6. Future Works

In the future studies, an integration that enables this system to be controlled by structures such as a smart home, cabinet, door, safe etc. or by applications on a smart phone can be developed. In this developed authentication system, a new authentication method as the third factor can be added for high security. And also, as the IoT devices evolve each year, the success rates will get much better with low cost cameras even in low light.

7. Ethics committee approval and conflict of interest statement

There is no need to obtain permission from the ethics committee for the article prepared.

There is no conflict of interest with any person / institution in the article prepared.

Acknowledgment

We would like to thank Dokuz Eylül University Computer Engineering Department.

References

[1] Evwiekpaefe, A.E., Eyinla, V.O. 2021. Implementing fingerprint authentication in computer-based tests. Nigerian Journal of Technology, Vol. 40(2), pp. 284-291.

[2] Fong, S., Zhuang, Y., Fister, I., Fister, Jr. I. 2013. A biometric authentication model using hand gesture images. BioMed Engineering OnLine, Vol. 12(111), pp. 1-18.

[3] Ibrahim, D.R., The, J.S., Abdullah, R. 2021. Multifactor authentication system based on color visual cryptography, facial recognition, and dragonfly optimization. Information Security Journal: A Global Perspective, Vol. 30(3), pp. 149-159.

[4] Lin, W.H., Wu, B.H., Huang, Q.H. 2018. A face-recognition approach based on secret sharing for user authentication in public-transportation security. 2018 IEEE International Conference on Applied System Invention (ICASI), Chiba &Tokyo, Japan, 13-17 April 2018.

[5] Gayathri, M., Malathy, C. 2021. Novel framework for multimodal biometric image authentication using visual share neural network. Pattern Recognition Letters, Vol. 152(December 2021), pp. 1-9.

[6] Badgujar, M., Wagh, A., Chavan, S., Chumbhale, P., Sonawane, R.C. 2022. IoT Based Automatic Door Lock System by Face and Voice Recognition. International Research Journal of Modernization in Engineering Technology and Science, Vol. 4(3), pp. 542-545.

[7] Van Murugiah, K., Subhashini, G., Abdulla, R. 2021. Wearable IOT based Malaysian sign language recognition and text translation system. Journal of Applied Technology and Innovation, Vol. 5(4), pp. 51-58.

[8] Ajay, S., Potluri, A., George, S.M., Gaurav, R., Anusri, S., 2021. Indian Sign Language Recognition Using Random Forest Classifier. 2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 9-11 July 2021.

[9] Palarimath, S., Blessing, N. R., Sujatha, T., Pyingkodi, M., Ugalde, B. H., & Palarimath, R. D. 2022. A Robust Authentication and Authorization System Powered by Deep Learning and Incorporating Hand Signals. In Intelligent Data Communication Technologies and Internet of Things (pp. 1061-1071). Springer, Singapore.

[10] Rwelli, R. E., Shahin, O. R., & Taloba, A. I. 2022. Gesture based Arabic Sign Language Recognition for Impaired People based on Convolution Neural Network. arXiv preprint arXiv:2203.05602.

[11] Harini, N., Lavanya, P., & Sravya, G. V. N. S. K. 2022. High-Security Locking System Using Arduino.

[12] Reddy, K. Y., Reddy, A. J., Reddy, K. B. P., & Rao, M. B. S. 2022. IoT Based Smart Door Lock System.