PROCEEDINGS OF
INTERNATIONAL CONFERENCE ON NEW TRENDS IN APPLIED SCIENCES
https://proceedings.icontas.org/
International Conference on New Trends in Applied Sciences (ICONTAS'23), Konya, December 1-3, 2023.

Development of Voice and Face Recognition Based Security Software for Biometric Systems

Emre AVUCLU
Aksaray University, Turkey, ORCID: 0000-0002-1622-9059
emreavuclu@aksaray.edu.tr

Murat KOKLU
Selcuk University, Turkey, ORCID: 0000-0002-2737-2360
mkoklu@selcuk.edu.tr

ABSTRACT: Nowadays, technology continues to make our life easier in every field. It is not possible to provide a high level of security access in security systems without the use of technology. Facial recognition and voice recognition technologies are encountered in many applications in every area of our lives. In this study, a sample application was developed by using a combination of face recognition and voice recognition technologies to create a security system. The Speech.dll library was used to control voice commands and the EmguCV libraries for face recognition. The application was developed in C # programming language. In this application, a security system with a very controlled structure was developed due to the importance of security systems. With the developed application, any high-security operation can be carried out anywhere with computer control. In the developed application, a method with two controls was proposed to provide security control of any system.

Keywords: Computer programming, Voice recognition, Security systems, Face recognition.

INTRODUCTION

Biometry is the general name given to the system which can make a definition as a result of the physical and behavioral behaviors of the individuals. Biometric systems are the methods that allow people to study biometrics. Face recognition technologies are one of the most important features of biometric systems. Biometrics, which includes all the techniques such as fingerprint recognition, retina and iris recognition, hand geometry recognition, voice recognition, face recognition, DNA recognition, signature recognition, and identity control were required in areas requiring a high level of security. It is an excellent solution in systems such as an automated personnel attendance control system, and it is a reliable technology that develops and is adopted very rapidly in social life applications [1-5]

There are some studies on face recognition, biometry, and voice recognition in literature.

Ozkaya et al. were tested in a database of 100 fingerprint images related to biometrics and gave successful results [6]. In Arslan and Sarıoğlu studies, biometric systems, methods, and technologies used in mobile devices were researched and their advantages and disadvantages were given [7]. They achieved a successful recognition rate for 10 people and also tried music, speech, and the application of different voice recognition algorithms [8-10]. Using different voice recognition algorithms," On TV", "Off TV", "Volume Up", "Volume Down" and "Channel One" command sets were tried [11-13]. In another study, a voice recognition-based security system was developed [14, 15]. Over 80% success was achieved in voice recognition on the letters "a", "e" and "i" [16]. Successful results were obtained in the study which carried out 40 commands. The numbers between 0-9 are perceived independently in time by means of the voice detection system [17-20]. Five different emotions of 30 people were tried to be detected and 70% success was achieved [21, 22]. The success rate was 67.5% in the system, which can perform simultaneous comparisons without registration [23-25]. Different algorithms have been tried and various studies have been done on face recognition. 3D face recognition [26, 27], development of face recognition system algorithms [28, 29], face recognition biometric security systems, real-time face recognition approaches in moving images, related to the presence of similar faces [30-32], and real-time facial recognition [33-36] have done different studies. In the United States, the National Security Agency (National Security Agency) uses a speech recognition type for keyword detection [37, 38]. A study was performed to improve the quality and accuracy of the speech recognition system using a multimodal audiovisual speech signal [39, 40]. They made a study called factored language models.

In this study, software for face detection, voice recognition, and password verification has been developed for security control. With this software, access to any process can be controlled with the highest level of security.

MATERIALS AND METHODS

Speech.dll and Software development kit libraries were used for voice recognition in the system. EmguCv library was used for facial recognition. Face recognition applications are software applications that can identify the person in a unique way by analyzing the face lines of the person. These systems are the fastest biometric methods with minimal errors. Face recognition [24, 41] is often used for security purposes. Eigenfaces are eigenvectors of the covariance matrix of the distribution of images in the training set. The face space for m facial images in the data set to be used for educational purposes is expressed as s=[3I_m]. Where I represent a gray-scale image vector in n × n dimensions in the face space. Eigenfaces technology is used in this application. The resulting view of the eigenfaces is shown in Figure 1 [42].

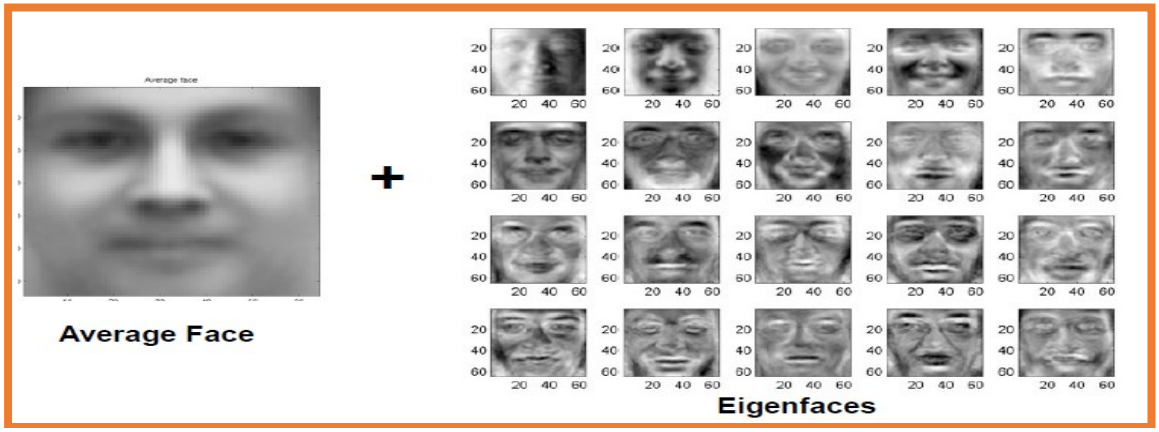

Figure 1. The examples of eigenface

Eigenfaces take the average of all the images in memory and perform the operation by taking the difference of the resulting image from the average. As a result, each image is evaluated as a differential image. The general interface of the software developed is shown in Figure 2.

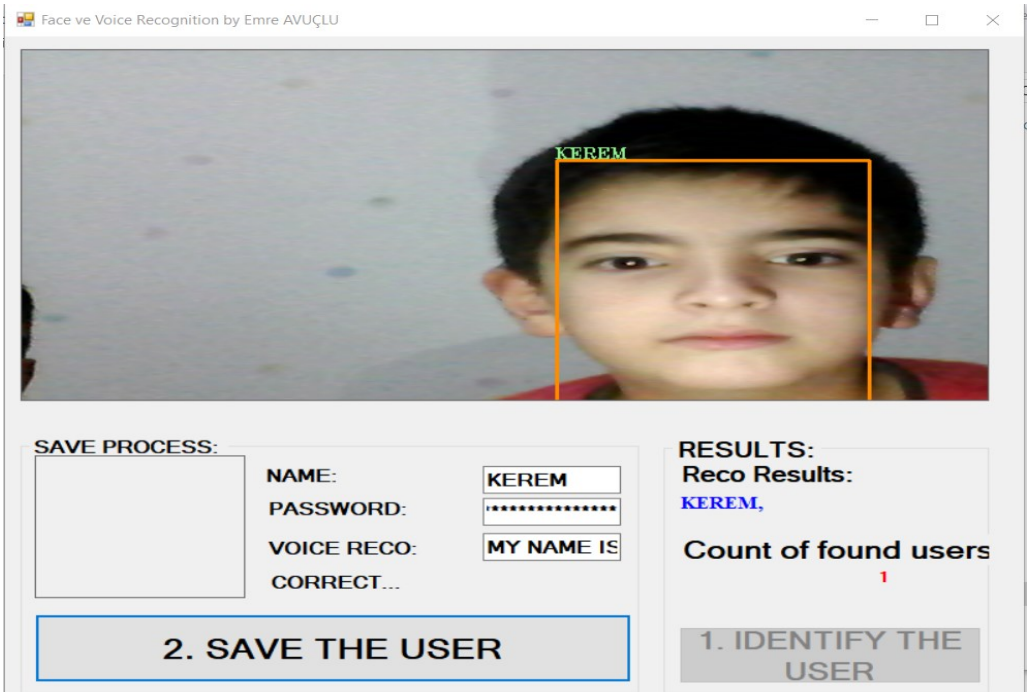

Figure 2. Interface of the developed system

The application developed in this study is also carried out a series of operations. The sequence and general description of these processes are shown in Figure 3.

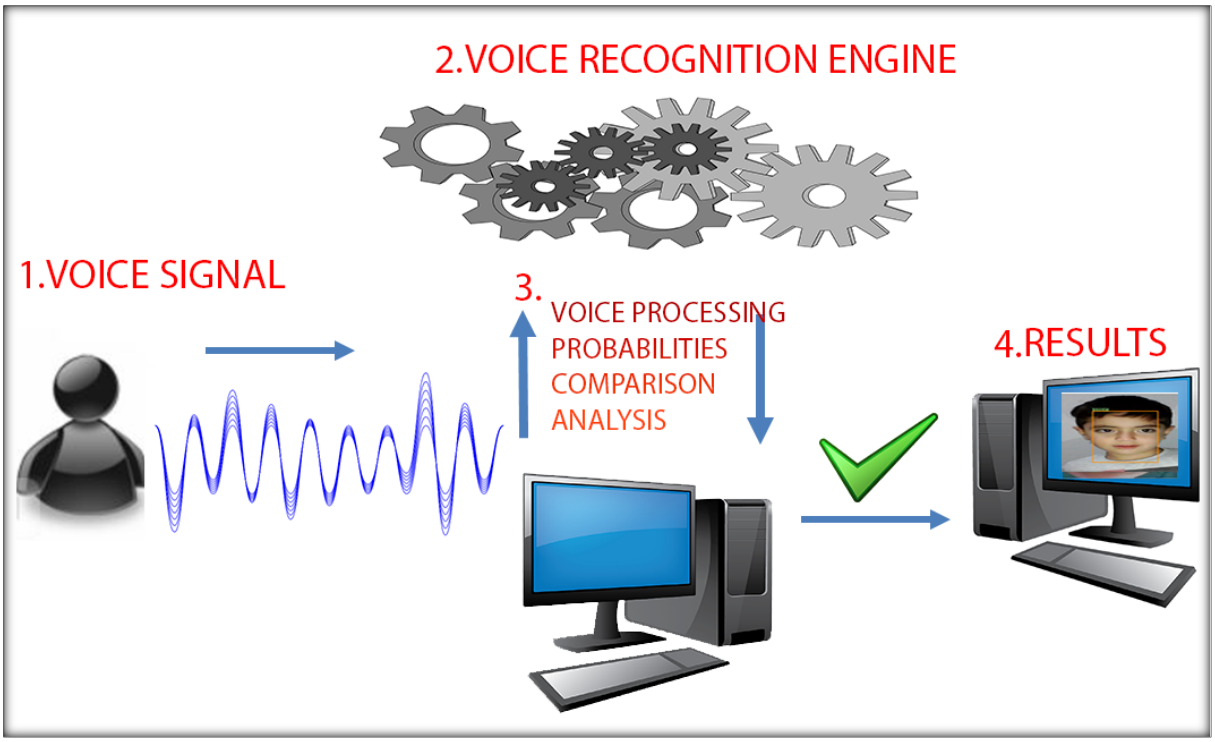

Figure 3. The general structure of the system

Voice detection [43] can be performed with the methods and classes in the System.Speech library in the .Net Framework. The following code blocks are used first for voice recognition and the use of the necessary libraries in the system.

```
SpeechSynthesizer Speech = new SpeechSynthesizer();
PromptBuilder founder = new PromptBuilder();
SpeechRecognitionEngine recog= new SpeechRecognitionEngine();
```

The code block required to activate or deactivate the application is as follows.

```
Recocon.Text = Result.PhraseInfo.GetText(0, -1, true); // activate
objRecoContext = null; // deactivate
```

The flow diagram of the developed application is shown in Figure 4.

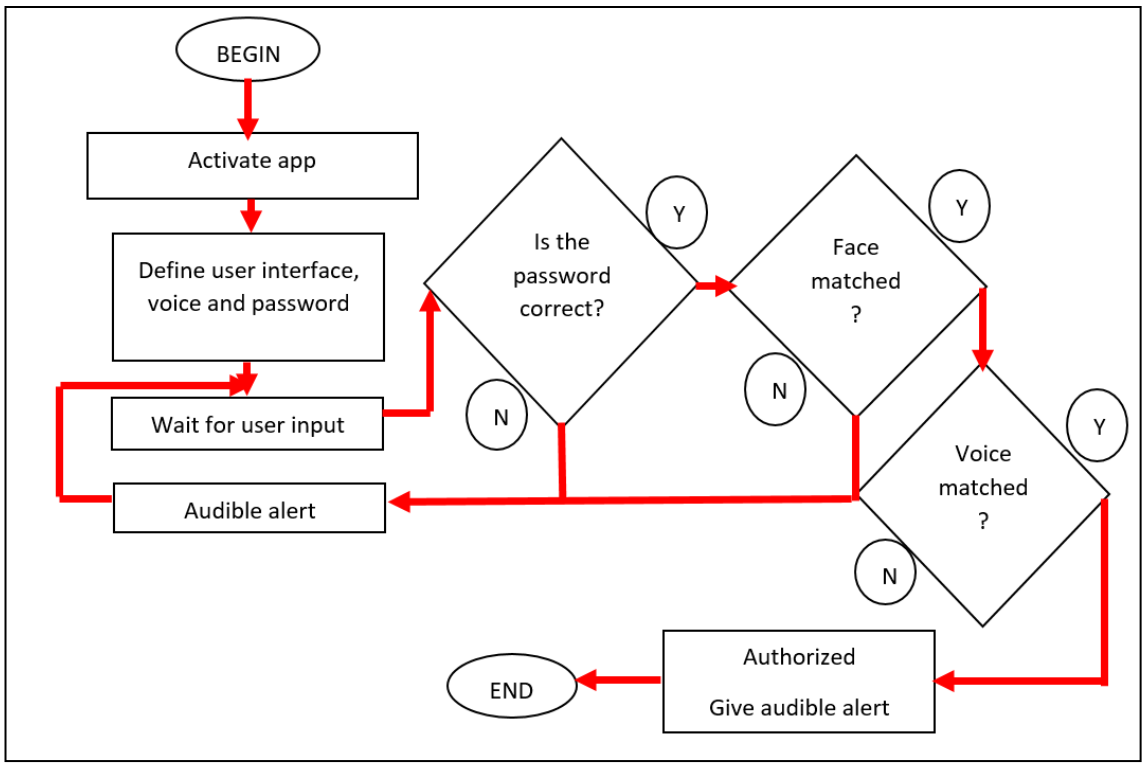

Figure 4. Flow diagram of the application

The system first creates a profile for the user. The information is recorded for this process as shown in Figure 5.

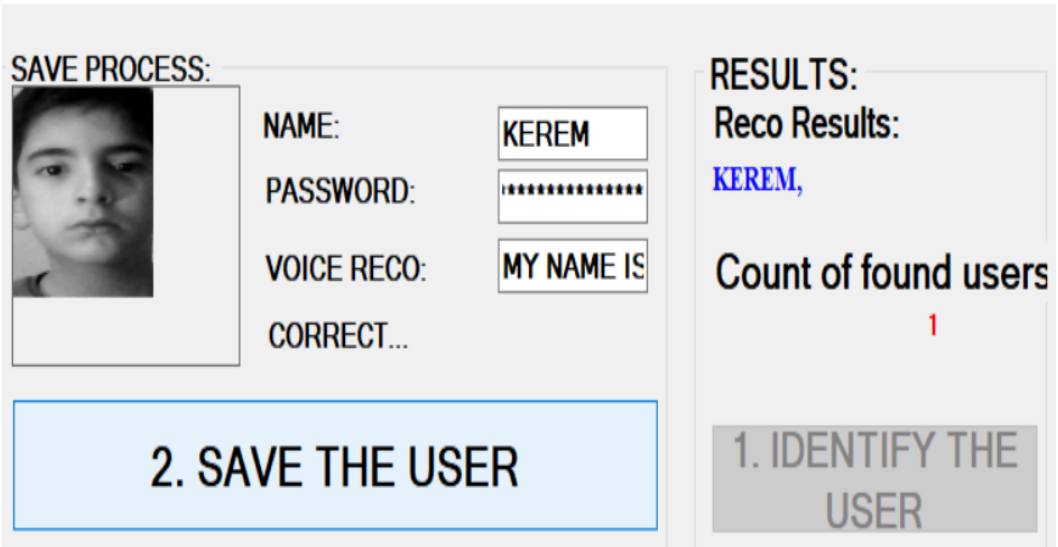

Figure 5. User identification process

In Figure 5. the name and surname of the person are entered in the "NAME" section, the password of the person is entered in the "PASSWORD" section and the voice description required for the approval of the system in the "VOICE DESCRIPTION" section. After these operations, the system is prepared according to the user. The following code blocks are used for "AUTHORIZED AUDIO DEFINITION", which is one of the parameters used to customize the application.

```
Choices recolist = new Choices();
```

```
recolist.Culture = new System.Globalization.CultureInfo("en-US");
recolist.Add(new string[] {textBox2.Text });
Grammar gr = new Grammar(new GrammarBuilder(recolist));
   try
   {
   Recog.RequestRecognizerUpdate();
   Recog.LoadGrammar(gr);
   Recog.SpeechRecognized += Recognize_SpeechRecognized;
   Recog.SetInputToDefaultAudioDevice();
   Recog.RecognizeAsync(RecognizeMode.Multiple);
    }
  catch
    {
    MessageBox.Show("ERROR..", e.ToString());
     return;
     }
```

In the developed application, the process of giving voice feedback indicating whether it is the right or wrong user is carried out with the following code blocks.

```
Founder.ClearContent();
Founder.AppendText(textBox1.Text);
SpeechSynth.Speak(Founder);
```

In voice recognition, the acknowledgment process is executed with the following code blocks. Both the username and the previously defined voice must match.

```
void Recognize_SpeechRecognized(object sender,SpeechRecognizedEventArgs e)
      {
         if (textBox2.Text == "MY NAME IS HOME")
         {
            if (e.Result.Text == "MY NAME IS HOME")
            {
               label8.Text = "CORRECT VOICE... ";
               SpeechSynth.Speak("HI, WELCOME TO SYSTEM..");
            }
         }
       }
```

CONCLUSION

In this study, the results obtained from experimental studies on voice and facial recognition are explained in this section. The face detection results are shown in Table 1. In this study, the results of the tests carried out 10 times over 20 people are as follows.

Table 1. Face detection levels in different situations

Light status	The viewing angle of the face	The distance of the face	Accurate recognition
Excessively	Front Against +/- 20-30 Degree	70-80 cm +/-15 cm	19
Middle	Front Against +/- 20-30 Degree	70-80 cm +/-15 cm	20
Little	Front Against +/- 20-30 Degree	70-80 cm +/-15 cm	17

Looking at the table, it can be seen that high light, low light, and distance affected negatively the number of recognitions. In addition, the number of recognitions has decreased in images that are not from the front.

As the pronunciation of the voice command becomes more difficult and the number of letters in it increases, the level of accurate voice recognition decreases. 100% accurate recognition can be achieved by using short words and words with full pronunciation when making voice definitions. In this study, test procedures were performed with 20 people. In some word tests, more than one test was performed over the same person's voice. The following in Table 2 shows the results of the experimental studies.

Table 2. Experimental results.

Words	Number of trials	Accurate recognition	False recognition	Error rate
My	15	15	0	%0
Name	20	20	0	%0
Home	20	19	1	%6,66
My name is home	20	17	3	%15

As can be seen from the results, it was more difficult to identify words with a high number of words and difficult to pronounce.

Face recognition and voice recognition technologies are increasing day by day in our lives. In this study, a new system with multiple controls is proposed for any security situation. This system can be developed for any application. Thanks to the application, control operations will be more reliable in areas requiring special security. EmguCv, and Speech.dll libraries were used in the application which was developed by using Eigenfaces technology in C # programming language. The developed application provides access to the system if the information matches the password, face recognition, and voice recognition control of a user registered to the system [44]. It is possible to use the developed application on any system that requires high security.

REFERENCES

[1] L. C. Jain, U. Halici, I. Hayashi, S. Lee, and S. Tsutsui, *Intelligent biometric techniques in fingerprint and face recognition*. Routledge, 2022.

[2] A. Jain, L. Hong, and R. Bolle, "On-line fingerprint verification," *IEEE transactions on pattern analysis and machine intelligence,* vol. 19, no. 4, pp. 302-314, 1997.

[3] A. K. Jain, L. Hong, S. Pankanti, and R. Bolle, "An identity-authentication system using fingerprints," *Proceedings of the IEEE,* vol. 85, no. 9, pp. 1365-1388, 1997.

[4] A. K. Jain, S. Prabhakar, L. Hong, and S. Pankanti, "FingerCode: a filterbank for fingerprint representation and matching," in *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, 1999, vol. 2: IEEE, pp. 187-193.

[5] I. Cinar, "Disaster Detection Using Machine Learning Methods with Deep Features," *disasters,* vol. 19, p. 21.

[6] N. Ozkaya, S. Sagiroglu, and E. Besdok, "Genel Amaçlı Otomatik Parmakizi Tanıma Sistemi Tasarımı ve Gerçekleştirilmesi," *Politeknik Dergisi,* vol. 8, no. 3, pp. 239-247, 2005.

[7] B. Arslan and S. Sagiroglu, "Mobil Cihazlarda Biyometrik Sistemler Üzerine Bir İnceleme," *Politeknik Dergisi,* vol. 19, no. 2, pp. 101-114, 2016.

[8] G. Dede and M. H. Sazlı, "Biyometrik sistemlerin örüntü tanıma perspektifinden incelenmesi ve ses tanıma modülü simülasyonu," *EEBM Ulusal Kongresi,* pp. 1-5, 2010.

[9] B. Bolat, U. Kucuk, and T. Yildirim, "Aktif Öğrenen Pnn İle Konuşma/Müzik Siniflandirma."

[10] O. Meral, "Doğrusal öngörülü kodlama ve adaptif algoritma tabanlı konuşmacı tanıma," Master Thesis, 2008.
[11] L. Muda, M. Begam, and I. Elamvazuthi, "Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques," *arXiv preprint arXiv:1003.4083,* 2010.
[12] M. K. Baygun and A. K. Yaldir, "Linear Predictive Coding Ve Dynamic Time Warping Teknikleri Kullanilarak Ses Tanima Sistemi Geliştirilmesi."
[13] A. Bala, A. Kumar, and N. Birla, "Voice command recognition system based on MFCC and DTW," *International Journal of Engineering Science and Technology,* vol. 2, no. 12, pp. 7335-7342, 2010.
[14] I. A. Ozkan and M. Koklu, "Skin lesion classification using machine learning algorithms," *International Journal of Intelligent Systems and Applications in Engineering,* vol. 5, no. 4, pp. 285-289, 2017.
[15] B. Ozturk and T. Cakar, "Gerçek Zamanli Ses Tanima," Engineering Faculty, Electrical-Electronics Engineering, Istnabul University, 2007. [Online]. Available: https://www.emo.org.tr/ekler/6a70544c7953a16_ek.pdf
[16] M. H. Asyali, M. Yilmaz, M. Tokmakci, K. Sedef, B. H. Aksebzeci, and R. Mittal, "Design and implementation of a voice-controlled prosthetic hand," *Turkish Journal of Electrical Engineering and Computer Sciences,* vol. 19, no. 1, pp. 33-46, 2011.
[17] C. G. Babu, R. H. Kumar, and P. Vanathi, "Performance analysis of hybrid robust automatic speech recognition system," in *2012 IEEE International Conference on Signal Processing, Computing and Control*, 2012: IEEE, pp. 1-4.
[18] M. A. Abushariah, R. N. Ainon, R. Zainuddin, M. Elshafei, and O. O. Khalifa, "Natural speaker-independent Arabic speech recognition system based on Hidden Markov Models using Sphinx tools," in *International Conference on Computer and Communication Engineering (ICCCE'10)*, 2010: IEEE, pp. 1-6.
[19] V. A. Petrushin, "Emotion recognition in speech signal: experimental study, development, and application," in *Sixth international conference on spoken language processing*, 2000.
[20] Y. S. Taspinar, I. Cinar, and M. Koklu, "Classification by a stacking model using CNN features for COVID-19 infection diagnosis," *Journal of X-ray science and technology,* vol. 30, no. 1, pp. 73-88, 2022.
[21] N. Ozturk and U. Unozkan, "Microprocessor based voice recognition system realization," in *2010 4th International Conference on Application of Information and Communication Technologies*, 2010: IEEE, pp. 1-3.
[22] I. A. Ozkan, M. Koklu, and I. U. Sert, "Diagnosis of urinary tract infection based on artificial intelligence methods," *Computer methods and programs in biomedicine,* vol. 166, pp. 51-59, 2018.
[23] V. Blanz and T. Vetter, "A Morphable model fort he synthesis of 3D faces," *Max-Planck-Institut für biologische Kybernetik, Germany,* 2011.
[24] M. Koklu and K. Tutuncu, "Performance of Classification Techniques on Medical Datasets," *Int. J. Biomed. Sci. Bioinforma,* vol. 3, no. 1, pp. 5-9, 2016.
[25] O. Ozaltın, M. Koklu, A. Yonar, and O. Yeniay, "Automatically Image Classification Based on a New CNN Architecture," in *III International Applied Statistics Conference (UYIK-2022), Skopje, N Macedonia*, 2022, pp. 22-24.
[26] K. Kıymacı, "Yüz tanıma sistemi algoritmalarının geliştirilmesi," *Bitirme Tezi, Yıldız Teknik Üniversitesi, Bilgisayar Mühendisliği Bölümü, İstanbul,* 2010.
[27] E. Avuclu, "A Novel Method Using Covid-19 Dataset and Machine Learning Algorithms For The Most Accurate Diagnosis That Can Be Obtained In Medical Diagnosis," *Biomedical Signal Processing and Control,* vol. 77, p. 103836, 2022.
[28] R. Samli and M. E. Yuksel, "Biyometrik güvenlik sistemleri," *XI. Akademik Bilişim Konferansı Bildirileri 11-13 Şubat 2009 Harran Üniversitesi, Şanlıurfa,* 2009.
[29] M. Koklu and K. Sabanci, "The classification of eye state by using kNN and MLP classification models according to the EEG signals," *International Journal of Intelligent Systems and Applications in Engineering,* vol. 3, no. 4, pp. 127-130, 2015, doi: https://doi.org/10.18201/ijisae.75836.
[30] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision,* vol. 57, pp. 137-154, 2004.
[31] M. Koklu, I. Cinar, and Y. S. Taspinar, "CNN-based bi-directional and directional long-short term memory network for determination of face mask," *Biomedical signal processing and control,* vol. 71, p. 103216, 2022.
[32] A. Tunc, S. Tasdemir, and M. Koklu, "Fuzzy Based Noise Removal, Age Group and Gender Prediction with CNN," in *Intelligent and Fuzzy Techniques: Smart and Innovative Solutions: Proceedings of the INFUS 2020 Conference, Istanbul, Turkey, July 21-23, 2020*, 2021: Springer, pp. 204-212.
[33] E. Hjelmas and B. K. Low, "Face detection: A survey," *Computer vision and image understanding,* vol. 83, no. 3, pp. 236-274, 2001.

[34] V. Bruce and A. Young, *In the eye of the beholder: The science of face perception*. Oxford university press, 1998.
[35] K. Singh, "Speech recognition: A review of literature," *International Journal of Engineering Trends and Technology (IJETT),* vol. 37, no. 6, pp. 302-310, 2016.
[36] M. Koklu and Y. Unal, "Analysis of a population of diabetic patients databases with classifiers," *International Journal of Biomedical and Biological Engineering,* vol. 7, no. 8, pp. 481-483, 2013.
[37] E. E. El Maghraby, A. M. Gody, and M. H. Farouk, "Enhancing quality and accuracy of speech recognition system by using multimodal audio-visual speech signal," in *2016 12th International Computer Engineering Conference (ICENCO)*, 2016: IEEE, pp. 219-229.
[38] E. Avuçlu and A. Elen, "Evaluation of train and test performance of machine learning algorithms and Parkinson diagnosis with statistical measurements," *Medical & Biological Engineering & Computing,* vol. 58, pp. 2775-2788, 2020.
[39] G. Donaj and Z. Kacic, "Context-dependent factored language models," *EURASIP Journal on Audio, Speech, and Music Processing,* vol. 2017, no. 1, pp. 1-16, 2017.
[40] M. M. Sarıtas, M. Koklu, and İ. A. OZKAN, "A Novel Embedded System on Cold Box Design for The Cold Chain," *International Journal of Applied Mathematics Electronics and Computers,* vol. 5, no. 4, pp. 67-70, 2017.
[41] A. Tunc, S. Tasdemir, M. Koklu, and A. C. Cinar, "Age group and gender classification using convolutional neural networks with a fuzzy logic-based filter method for noise reduction," *Journal of Intelligent & Fuzzy Systems,* vol. 42, no. 1, pp. 491-501, 2022, doi: 10.3233/JIFS-219206.
[42] G. Biricik, "Metin sınıflama için yeni bir özellik çıkarım yöntemi," 2011.
[43] Y. S. Taspinar, M. M. Saritas, İ. Cinar, and M. Koklu, "Gender Determination Using Voice Data," *International Journal of Applied Mathematics Electronics and Computers,* vol. 8, no. 4, pp. 232-235, 2020, doi: https://doi.org/10.18100/ijamec.809476.
[44] I. Cinar, Y. S. Taspinar, R. Kursun, and M. Koklu, "Identification of Corneal Ulcers with Pre-Trained AlexNet Based on Transfer Learning," in *2022 11th Mediterranean Conference on Embedded Computing (MECO)*, 2022: IEEE, pp. 1-4, doi: 10.1109/MECO55406.2022.9797218.