Conference Paper/Konferans Bildirisi

GAZİ
JOURNAL OF ENGINEERING SCIENCES

Gender Detection via Voice Using Artificial Intelligence Algorithms

Serkan Gönen [a], Mehmet Ali Barışkan [b], Gökçe Karacayılmaz [c], Birkan Alhan [d], Ercan Nurcan Yılmaz [e], Harun Artuner [f]

Submitted: 01.09.2022 Revised: 09.12.2022 Accepted: 15.12.2022 doi:10.30855/gmbd.0705043

ABSTRACT

As a result of the developments in science and technology, all our living spaces, from health, education, and trade to our social life, have been moved to the digital environment. With this process, artificial intelligence, which is the ultimate goal of creating systems that think and act like human beings, has started to be used in all areas of our lives. This study focuses on gender determination by using artificial intelligence algorithms on voice data. Thanks to this determination, significant contributions will be made in various fields such as social engineering and cyber security such as fraud, person detection, and advertising investments. In the analysis of the study, R application, a completely open-source, for various artificial intelligence algorithms has been used. In this way, a solution has been provided to take the security as mentioned above measures with low cost instead of high-cost systems and increase the sales figures in areas such as marketing. In the study, supervised learning artificial intelligence algorithms have been examined. The artificial intelligence analysis results of the study have shown that the gender of the person could be determined above % 97 successful rates through the voice data.

Keywords: Artificial Intelligence, Gender Recognition, Cyber Security, Voice Analysis, Supervised Learning

[a] Istanbul Gelisim University, Dept. of Computer Engineering, 34310 - Istanbul, Turkey, Orcid: 0000-0002-1417-4461

[b] Istanbul Gelisim University, Dept. of Computer Engineering, 34310 - Istanbul, Turkey, Orcid: 000-0002-8039-2686

[c] Hacettepe University, Graduate School of Science and Engineering, Dept. of Computer Engineering 06532 - Ankara, Turkey, Orcid: 0000-0001-8529-1721

[d] Istanbul Gelisim University, Dept. of Computer Engineering, Istanbul, Turkey, Orcid: 0000-0003-1511-0109

[e,*] Gazi University, Technology Faculty, Dept. of Electrical and Electronic Engineering 06560 - Ankara, Türkiye Orcid: 0000-0001-9859-1600 e mail: enyilmaz@gazi.edu.tr

[f] Hacettepe University, Dept. of Computer Engineering, 06532 - Ankara, Turkey Orcid: 0000-0002-6044-379X

*Corresponding author: enyilmaz@gazi.edu.tr

Yapay Zekâ Algoritmaları Kullanılarak Sesle Cinsiyet Tespiti

ÖZ

Bilim ve teknolojideki gelişmeler sonucunda sağlıktan eğitime, ticaretten sosyal hayatımıza kadar tüm yaşam alanlarımız dijital ortama taşınmıştır. Bu süreçle birlikte insan gibi düşünen ve hareket eden sistemler oluşturmak amacıyla geliştirilmiş yapay zeka kavramı da hayatımızın her alanında kullanılmaya başlanmıştır. Bu çalışmada, yapay zeka algoritmaları kullanılarak ses verilerinin incelenmesiyle cinsiyet belirlemeyi hedefleyen bir algoritma geliştirilmiştir. Cinsiyet tespitine yönelik olarak yapılan Bu tespit sayesinde sosyal mühendislik gibi çeşitli alanlarda ve dolandırıcılık, kişi tespiti, reklam yatırımları gibi siber güvenlik alanlarında önemli katkılar sağlanması hedeflenmiştir. Uygulama geliştirilirken, çeşitli yapay zeka algoritmaları için tamamen açık kaynak kodlu R yazılımı kullanılmıştır. Bu sayede yukarıda bahsedilen güvenlik önlemlerinin yüksek maliyetli sistemler yerine düşük maliyetli önlemler alınmasına ve pazarlama gibi alanlarda satış rakamlarının artırılmasına çözüm aranmıştır. Ayrıca çalışmada yapay zeka algoritması olarak denetimli öğrenme kullanılmıştır. Çalışmanın yapay zeka analiz sonuçları, ses verileri aracılığıyla kişinin cinsiyetinin çok başarılı oranlarda belirlenebildiğini göstermiştir.

Anahtar Kelimeler: Yapay Zekâ, Cinsiyet Tanıma, Siber Güvenlik, Ses Analizi, Denetimli Öğrenme

To cite this article: S. Gönen, M. A. Barışkan, G. Karacayılmaz, B. Alhan, E. N. Yılmaz, and H. Artuner, "Gender Detection via Voice Using Artificial Intelligence Algorithms," *Gazi Journal of Engineering Sciences,* vol. 8, no. 3, pp. 567-575, 2022. doi:10.30855/gmbd.0705043

1. Introduction

Advances in science and technology have provided new solutions to the problems that existed in the past. The solution to all kinds of problems, from the field of health to the defense industry, is provided by technology. However, the most important factor in this process is the authentication systems, in which individuals are verified with systems such as fingerprint, voice print, and face recognition [1].

In ecosystems where the security of digitally stored data and the reliability of comparison systems are not given importance in authentication systems, the existence of cyber-attacks and identity theft becomes inevitable. This situation may damage individuals and cause material and moral damage and a loss of prestige and trust for system owners. As a result of the development of technology and the increase in cyberattack environments, the mentioned attack situations have been experienced. As a result, authentication systems have become the most important concept in recent times [2]. In this area, biometric, voice, and identity verification with a primitive question-answer method are performed [3]. Digitization of identity cards, voice signature, and authentication with an electronic signature can be given as examples of these methods. These systems have taken security measures in many countries and institutions using certain conditions and standards. However, considering the cost dimension of these measures, it is seen that their installation and maintenance are difficult and costly [4].

Gender recognition is based on the determination of gender information from people's speech. In other words, it focuses on determining the person from his voice characteristics [5]. Classification of speaker's gender information is one of the most important problems in speech processing. The most basic issue in determining the gender of the person is to produce stable features and to design a good classifier [6]. There are various studies in the literature focusing on gender recognition from voice data. Within the scope of the study, gender determination has been focused on by using the voice parameter, which is a sub-branch of biometric data, which is one of the security analyses. In this way, it will be possible to use the voice to detect gender with high accuracy and prevent spoofing at points where the sound factor is at the forefront, especially in call centers and security control points. In addition to the safety dimension, which is the first target, the gains obtained in the study will also contribute in many different areas such as disease detection via voice [7], use of gender-oriented information in the field of advertising, and e-commerce [8]. The gender of a customer contacting the call center for account transactions in a bank should be confirmed at first. If the customer is a woman but the person speaking is a man, they should be removed directly from the system without the need for identity verification. Considering that the voice signature used in banking systems cannot work properly from time to time, customer security will be ensured at the most basic level by at least determining the gender of the caller by voice analysis. Infiltration and action attempts in disguise in security areas are one of the benefits of working towards physical security. In this context, it will be possible to prevent security attacks in many regions at a low cost. In addition, by making gender determination with sound analysis in marketing activities, it will be possible to receive personalized advertisements, products, and campaigns.

In recent years, with the development of artificial intelligence applications and machine learning, studies in which voice is used in terms of banking, mobile applications, and security have also increased. Especially as fraud detection systems use these types of detection algorithms. The study conducted by Ertam focused on gender determination using the LSTM [9]. They used a small dataset with ten features. On other research, Büyükyılmaz et al., in their study, investigated gender determination using the Multilayer Perceptron (MLP) deep learning model [10]. Li et al. in their study, they stated that combining it with a Support Vector Machine can be successful in detection with a low computational cost [11]. Metze et al. examined different techniques for gender classification over phone applications. They compared the performance of the systems they designed with human listeners [12]. Another of the first studies in this field is the work of Acero and Huang. Researchers stated that they achieved a 30% reduction in error rate in their studies by using the Hidden Markov Model [13].

In addition to these studies, there are also studies that make sound analyzes and disease predictions. In this context, the discovery of tumors in the upper extremity postural and action tremor regions from vocal tremors were discussed in the study by Suppa et al. [14]. While Schultbraucks et al. combined sound analysis with image processing to examine post-traumatic learning [15], Robbotti et al. focused on the detection from sound analysis of those who have recovered from Covid 19 disease and those who have never been ill [16]. Shimon et al. used voice analyses with questionaries' to analyze sound quality to detect the effects of Covid19 on human speech [17]. In another area of voice analyses, Al-Dhief, et al. investigated the pathology of sounds using machine learning [18]. Other studies investigated the detection of depressed people through voice analysis [19, 20]. As an important example of sound-based detection, sound analysis has been used to detect Parkinson's disease [21,22, 23].

2. Experimental Environment (Testbed)

Within the scope of the study, open-source technologies have been used to keep the cost at a minimum level. Within the scope of the study, the voice of the person whose gender determination has been requested has been recorded, and then, as a result of the analysis, the target voice recording has been divided into 22 acoustic property values. The trained artificial intelligence model has evaluated these values, and as a result, gender determination has been made. The stages of the experimental environment carried out in the study are depicted in Figure 1.

3. Voice Analysis

In order to make sound analysis in the study, firstly, the sound recording has been made in .WAV format. Secondly, sound analysis has been performed using the WarbleR package of the open-source R application for artificial intelligence model design [19]. Sound-specific acoustic properties have been extracted by using the specan function in order to measure and compare the recorded sound file. As a result of this process, the sound has been converted into 22 acoustic features.

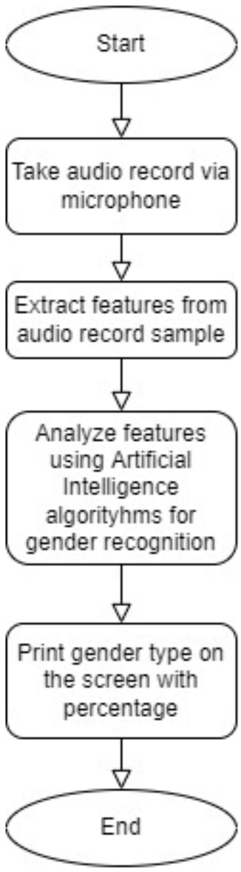

Figure 1. Flowchart of gender analysis over voice

These acoustic properties are given below:

- duration : length of signal
- meanfreq : mean frequency (in kHz)
- sd : standard deviation of frequency

- median : median frequency (in kHz)
- Q25 : first quantile (in kHz)
- Q75 : third quantile (in kHz)
- IQR : interquantile range (in kHz)
- skew : skewness
- kurt : kurtosis
- sp.ent : spectral entropy
- sfm : spectral flatness
- mode : mode frequency
- centroid : frequency centroid
- peakf : peak frequency (frequency with the highest energy)
- meanfun : average of fundamental frequency measured across an acoustic signal
- minfun : minimum fundamental frequency measured across an acoustic signal
- maxfun : maximum fundamental frequency measured across an acoustic signal
- meandom : average of dominant frequency measured across an acoustic signal
- mindom : minimum of dominant frequency measured across an acoustic signal
- maxdom : maximum of dominant frequency measured across an acoustic signal
- dfrange : range of dominant frequency measured across an acoustic signal
- modindx : modulation index

The sound file allocated to the features mentioned above has been instantly read, and gender determination has been carried out through the pre-trained artificial intelligence algorithm on the defining features of the voice.

There are 20 features of this data. The attributes of this dataset are listed in Figure 2. Also has the weight value and ranking of each features. As shown in the Figure 2, the best weight effect is mean fun.

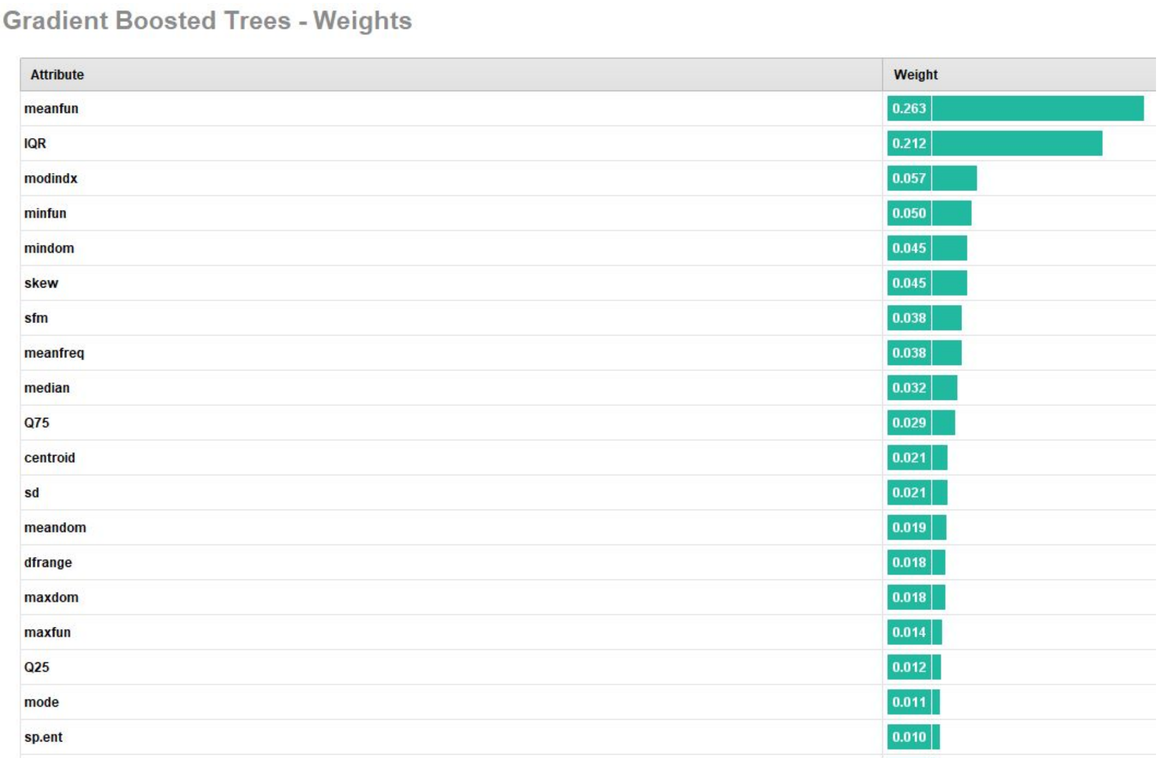

Figure 2. Weights of attributes of the used dataset

The training set has been run as 10 folds and the iteration values determined by using the 10-fold method for the GBM algorithm that has given the best results are shown in Figure 3.

```
Stochastic Gradient Boosting

2249 samples
  10 predictor
   2 classes: 'female', 'male'

No pre-processing
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 2024, 2024, 2025, 2024, 2023, 2025, ...
Resampling results across tuning parameters:

  interaction.depth  n.trees  Accuracy   Kappa
  1                   50      0.9221961  0.8443858
  1                  100      0.9466466  0.8932890
  1                  150      0.9573213  0.9146397
  2                   50      0.9533232  0.9066398
  2                  100      0.9657638  0.9315226
  2                  150      0.9697638  0.9395248
  3                   50      0.9564304  0.9128549
  3                  100      0.9670971  0.9341921
  3                  150      0.9706527  0.9413038

Tuning parameter 'shrinkage' was held constant at a value of 0.1
Tuning parameter 'n.minobsinnode' was held constant
 at a value of 10
Accuracy was used to select the optimal model using the largest value.
The final values used for the model were n.trees = 150, interaction.depth = 3, shrinkage = 0.1 and n.minobsinnode = 10.
```

Figure 3. 10 fold (Cross-validated) analysis

Density plots of men and women extracted from 20 of the sound features determined within the scope of the study are depicted in Figure 4.

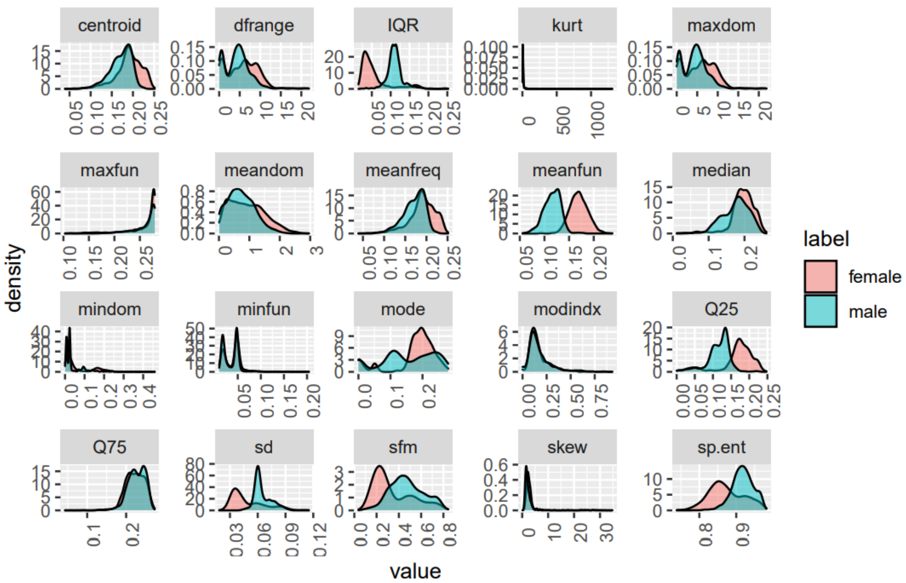

Figure 4. Density plots for male-female values

The confusion matrix of study is given in Figure 5.

```
   Cell Contents
|-------------------------|
|                       N |
|         N / Table Total |
|-------------------------|


Total Observations in Table:  962


              | predicted gender
actual gender |    female |      male | Row Total |
--------------|-----------|-----------|-----------|
       female |       470 |        10 |       480 |
              |     0.489 |     0.010 |           |
--------------|-----------|-----------|-----------|
         male |        10 |       472 |       482 |
              |     0.010 |     0.491 |           |
--------------|-----------|-----------|-----------|
 Column Total |       480 |       482 |       962 |
--------------|-----------|-----------|-----------|
```

Figure 5. Confusion Matrix

4. Artificial Intelligence Model

The analyzed audio file has been sent to the artificial intelligence integration as a dataset divided into 22 acoustic property values. Within the scope of this study, comparisons have been made with eight different algorithms to determine the most performance and most efficient artificial intelligence algorithm. As a test dataset, the universally accepted voice gender dataset on Kaggle [24], which has voice frequencies and acoustic property values, has been used by adding new voice data.

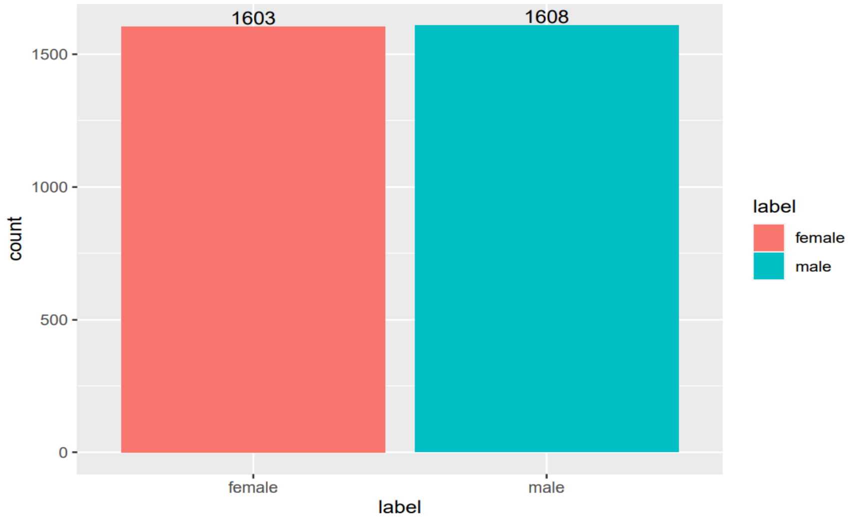

Figure 6. Male-female distribution of the dataset with voice analysis

As depicted in Figure-6, the new dataset created in the study consists of acoustic features of 3211 (1603 female, 1608 male) balanced and gender-labeled voice data. In artificial intelligence algorithms, 70% of values for the train and 30% for the control dataset have been used to reach the most optimum values and result values. Table 1. shows the accuracy values obtained as a result of the study of artificial intelligence algorithms and the metrics of the performance of the algorithms. When these performance metrics are examined together with the accuracy value, it is seen that the Gradient Boosted Model (GBM) gives the best results and this algorithm is used as a model in the study.

Table 1. Artificial intelligence algorithms and success rates used for gender analysis over voice

Algorithm	Accuracy	Classification Error	Standard Deviation	AUC	Standard Deviation	F Measure	Standard Deviation	Precision	Standard Deviation	Recall	Standard Deviation
Naive Bayes	87.6%	12.4%	±1.0%	0.945	±0.013	87.7%	±0.9%	87.3%	±2.1%	88.0%	±1.1%
Generalized Linear Model	97.1%	2.9%	±0.6%	0.994	±0.007	97.1%	±0.6%	95.8%	±1.6%	98.5%	±0.6%
Logistic Regression	94.9%	5.1%	±0.7%	0.994	±0.007	95.1%	±0.7%	91.7%	±1.4%	98.7%	±0.9%
Deep Learning	96.1%	3.9%	±0.3%	0.997	±0.004	96.2%	±0.2%	93.4%	±1.0%	99.1%	±0.9%
Decision Tree	96.5%	3.5%	±0.3%	0.987	±0.005	96.5%	±0.3%	97.0%	±1.1%	96.1%	±1.6%
Random Forest	97.2%	2.8%	±0.4%	0.996	±0.004	97.2%	±0.5%	95.4%	±0.9%	98.0%	±1.4%
Gradient Boosted Model	97.9%	2.1%	±0.5%	0.995	±0.003	97.8%	±0.5%	97.4%	±1.2%	98.3%	±1.5%
Support Vector Machine	96.2%	3.8%	±0.5%	0.88	±0.014	96.1%	±0.5%	95.1%	±1.3%	97.8%	±1.4%

The Gradient Boosting Machine (GBM) has given the highest accuracy, F-measure and precision rate among the specified artificial intelligences. So considering these results in the continuation of the study, the GBM algorithm has been used. As a result of running the GBM algorithm, the returned value (female or male) has been transmitted to the relevant user.

5. Discussion

Success statistics are given in studies on sound. Such as Ramadhan et al. examined the random forest algorithm in his study. In the study, classification of the data was carried out using parameter optimization. As a result of the study, it was reported that a success rate of 96.7% was determined [25]. DARPA, according to Ertem, has achieved up to 84% accuracy in resource management documentation. This dataset is from 160 American English-speaking interviews. [26]. Huestegge et al. on the other hand, proposed a new gender-based approach based on a generic voice classifier. The method with 92% success was used to classify the language independently [27]. Moreover, Zorumand et al. In their study, tried to determine the gender of Malay children's voices. They examined the use of formal and fundamental frequencies to discriminate the sex of 7-12-year old children. In their study, they achieved 99.8% performance with MLP [28]. Büyükyilmaz et al. also used a multilayer perceptron deep learning model with 96.8% accuracy to identify gender from voice features [10]. Zvarevashe et al. [29], on the other hand, managed to determine the gender with 97.58% accuracy by using random forest classifier and gradient boosting machine algorithm Livieris, et al, on the other hand used iCST Votting a deviant of CST Votting algorithm with success rate of 98.42%. The ResNet 50 fine-tuned gender data study, by Alnuaim et al., achieved an accuracy of 98.57% [31]. In this study, multiple methods were tried and the most successful one that was the Gradient Boosted Model, with a rate of 97.8%.

6. Conclusion

The design and implementation of voice authentication and gender detection applications is still a developing technology. Still, it has started to take an active role, especially in banking and marketing applications. However, their installation costs are quite high. The study aims to eliminate this high cost by using open source technologies and minimize human-induced errors and/or spoofing attacks by using artificial intelligence algorithms. In this way, an effective solution with low cost has been brought to the cyber security applications and marketing applications that use voice data. The main purpose of this study is to analyze the identity verification and voice signature approach, which is included in specific sectors such as the banking sector, especially for cyber security, with more stable and artificial intelligence analysis and a humanistic approach. In addition, it is aimed to evaluate people as men or women with voice analysis, regardless of disguise, by preventing the abuse of women's / men's rights used in security attacks. In the study, gender determination has been made on the voice data using supervised artificial intelligence algorithms, and an accuracy rate of 97% has been reached with the Gradient Boosting Machine. Thanks to this approach of the study, realizing identity and gender determination with artificial intelligence has been provided in sectors where technology and identity requirements are used.

In the future study, it aims to develop the existing system by combining it with IoT technologies. To make studies to use sound more effectively in physical and cyber security areas such as speaker recognition and emotion analysis.

Conflict of Interest

The authors declare that there is no conflict of interest

References

[1] A. Ross, S. Banerjee, and A. Chowdhury, "Security in smart cities: A brief review of digital forensic schemes for biometric data," *Pattern Recognition Letters*, vol. 138, pp. 346-354, 2020. doi:10.1016/j.patrec.2020.07.009

[2] M. A. Ferrag, L. Maglaras, A. Derhab, and H. Janicke, "Authentication schemes for smart mobile devices: Threat models, countermeasures, and open research issues," *Telecommunication Systems*, vol. 73, no. 2, pp. 317-348, 2020. doi:10.1007/s11235-019-00612-5

[3] A. Abozaid, A. Haggag, H. Kasban, and M. Eltokhy, "Multimodal biometric scheme for human authentication technique based on voice and face recognition fusion." *Multimedia tools and applications*, vol. 78, no. 12, pp. 16345-16361, 2019. doi:10.1007/s11042-018-7012-3

[4] S. Tanwar, M. S. Obaidat, S. Tyagi, and N. Kumar, "Online signature-based biometric recognition," *in Biometric-based physical and cybersecurity systems: Springer*, pp. 255-285. 2019. doi:10.1007/978-3-319-98734-7_10

[5] C. Chen, W. Wang, Y. He, & J. Han, “A bilevel framework for joint optimization of session compensation and classification for speaker identification.” *Digital Signal Processing*, 89, 104-115, 2019. doi:10.1016/j.dsp.2019.03.008

[6] Z. Qawaqneh, A. A. Mallouh, & B. D. Barkana, “Deep neural network framework and transformed MFCCs for speaker's age and gender classification,” *Knowledge-Based Systems*, vol. 115, pp. 5-14, 2017. doi:10.1016/j.knosys.2016.10.008

[7] J. Chu, “Artificial intelligence model detects asymptomatic Covid-19 infections through cellphone-recorded coughs,” 2020. [Online]. Available: https://news. mit. edu/2020/covid-19-cough-cellphone-detection [Access Date: 08.04. 2022].

[8] P. Gupta, S. Goel, and A. Purwar, "A stacked technique for gender recognition through voice," *in 2018 Eleventh International Conference on Contemporary Computing (IC3), Noida, India,* 2018. pp. 1-3. doi:10.1109/IC3.2018.8530520

[9] F. Ertam, "An effective gender recognition approach using voice data via deeper LSTM networks," *Applied Acoustics*, vol. 156, pp. 351-358, 2019. doi:10.1016/j.apacoust.2019.07.033

[10] M. Buyukyilmaz and A. O. Cibikdiken, "Voice gender recognition using deep learning," in *2016 International Conference on Modeling, Simulation and Optimization Technologies and Applications (MSOTA2016), Xiamen, China, Dec. 18-19, 2016,* Atlantis Press, 2016. pp. 409-411, 2016,. doi: 10.2991/msota-16.2016.90

[11] M. Li, C. S. Jung, & K. J. Han, “Combining five acoustic level modeling methods for automatic speaker age and gender *recognition.” In Eleventh annual conference of the international speech communication association (INTERSPEECH 2010), Makuhari, Chiba, Japan, Sept. 26-30, 2010*, International Speech Communication Association (ISCA), 2010. pp. 2826-2829. doi:10.21437/Interspeech.2010-74

[12] F. Metze, J. Ajmera, R. Englert, U. Bub, F. Burkhardt, J. Stegmann, & B. Littel, “Comparison of four approaches to age and gender recognition for telephone applications.” *In 2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07, Honolulu, HI, USA, Ap. 15-20,* 2007*.* vol. 4, pp. IV-1089. doi:10.1109/ICASSP.2007.367263

[13] A. Acero, & X. Huang, “Speaker and gender normalization for continuous-density hidden Markov models.” *In 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings, Atlanta, GA, USA, May, 1996*, IEEE Publication, 1996. vol. 1, pp. 342-345. doi:10.1109/icassp.1996.541102

[14] A. Suppa et al., "Voice analysis with machine learning: one step closer to an objective diagnosis of essential tremor," *Movement Disorders,* vol. 36, no. 6, pp. 1401-1410, 2021. doi:10.1002/mds.28508

[15] K. Schultebraucks, V. Yadav, and I. R. Galatzer-Levy, "Utilization of machine learning-based computer vision and voice analysis to derive digital biomarkers of cognitive functioning in trauma survivors," *Digital biomarkers*, vol. 5, no. 1, pp. 16-23, 2021. doi:10.1159/000512394

[16] C. Robotti et al., "Machine learning-based voice assessment for the detection of positive and recovered COVID-19 patients," *Journal of Voice*, vol. In Press., 2021. doi:10.1016/j.jvoice.2021.11.004

[17] C. Shimon, G. Shafat, I. Dangoor, and A. Ben-Shitrit, "Artificial intelligence enabled preliminary diagnosis for COVID-19 from voice cues and questionnaires," *The Journal of the Acoustical Society of America*, vol. 149, no. 2, pp. 1120-1124, 2021. doi:10.1121/10.0003434

[18] F. T. Al-Dhief et al., "Voice pathology detection and classification by adopting online sequential extreme learning machine," *IEEE Access,* vol. 9, pp. 77293-77306, 2021. doi:10.1109/ACCESS.2021.3082565

[19] D. Shin et al., "Detection of minor and major depression through voice as a biomarker using machine learning," *Journal of Clinical Medicine*, vol. 10, no. 14, p. 3046, 2021. doi:10.3390/jcm10143046

[20] C. W. Espinola, J. C. Gomes, J. M. S. Pereira, and W. P. dos Santos, "Detection of major depressive disorder using vocal acoustic analysis and machine learning - an exploratory study," *Research on Biomedical Engineering*, vol. 37, no. 1, pp. 53-64, 2021. doi:10.1007/s42600-020-00100-9

[21] J. Carrón, Y. Campos-Roca, M. Madruga, and C. J. Pérez, "A mobile-assisted voice condition analysis system for Parkinson’s disease: assessment of usability conditions," *BioMedical Engineering OnLine*, vol. 20, no. 1, pp. 1-24, 2021. doi:10.1186/s12938-021-00951-y

[22] L. Moro-Velazquez, J. A. Gomez-Garcia, J. D. Arias-Londoño, N. Dehak, and J. I. Godino-Llorente, "Advances in Parkinson's disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects," *Biomedical Signal Processing and Control,* vol. 66, pp. 102418, 2021. doi:10.1016/j.bspc.2021.102418

[23] E. Majda-Zdancewicz, A. Potulska-Chromik, J. Jakubowski, M. Nojszewska, and A. Kostera-Pruszczyk, "Deep learning vs feature engineering in the assessment of voice signals for diagnosis in Parkinson’s disease," *Bulletin of the Polish Academy of Sciences. Technical Sciences,* vol. 69, no. 3, 2021. doi: 10.24425/bpasts.2021.137347.

[24] K. Becker, “Gender recognition by voice,” *Kaggle.com* 2016. [Online] Available: https://www.kaggle.com/datasets/primaryobjects/voicegender. [Accessed: 08.04. 2022].

[25] M. M. Ramadhan, I. S. Sitanggang, F. R. Nasution, & A. Ghifari, “Parameter tuning in random forest based on grid search method for gender classification based on voice frequency,” *DEStech Transactions on Computer Science and Engineering*, vol. 10, pp. 625-629, 2017. doi:10.12783/dtcse/cece2017/14611

[26] F. Ertam, “An effective gender recognition approach using voice data via deeper LSTM networks,” *Applied Acoustics,* vol.156, pp. 351–358. 2019. doi:10.1016/j.apacoust.2019.07.033

[27] S. M. Huestegge and T. Raettig, “Crossing Gender Borders: Bidirectional Dynamic Interaction Between Face-Based and Voice-Based Gender Categorization,” *JVoice,* vol. 34, no. 3 pp. 487 -487, 2018. doi:10.1016/j.jvoice.2018.09.020

[28] A. Zourmand, H. N. Ting andS. M. Mirhassani, “Gender classification in children based on speech characteristics: using fundamental and formant frequencies of Malay vowels,” *J Voice*, vol. 27 no. 2 pp. 201-209, 2013. doi:10.1016/j.jvoice.2012.12.006.

[29] K. Zvarevashe and O. O. Olugbara, “Gender Voice Recognition Using Random Forest Recursive Feature Eliminationwith Gradient Boosting Machines,” *In Proceedings of the 2018 International Conference on Advances in Big Data,Computing and Data Communication Systems (icABCD),* pp. 1–6, 6-7 Augest 2018, Durban, South Africa. doi: 10.1109/ICABCD.2018.8465466

[30] I. E. Livieris, E. Pintelas, P. Pintelas, “Gender Recognition by Voice Using an Improved Self-Labeled Algorithm,” *Mach. Learn. Knowl.* Extr., vol. 1, no. 1, pp. 492–503, 2019. doi;10.3390/make1010030

[31] A. A. Alnuaim, M. Zakariah, C. Shashidhar, W A. Hatamleh, H. Tarazi, P. K. Shukla, and R. Ratna, "Speaker Gender Recognition Based on Deep Neural Networks and ResNet50," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1-13, 2022. doi:10.1155/2022/4444388

** Bu makale Uluslararası Mühendislikte Yapay Zeka ve Uygulamalı Matematik Konferansı - ICAIAME 2022 isimli konferansta sunulmuş bildirinin genişletilmiş halidir.*