ORIGINAL RESEARCH ORİJİNAL ARAŞTIRMA

DOI: 10.5336/forensic.2023-96327

Investigation of the Effect of Ear Measurements on Sex Estimation in Forensic Sciences Using Machine Learning Techniques: Descriptive Research

Adli Bilimlerde Kulak Ölçümlerinin Cinsiyet Tahminine Etkisinin Makine Öğrenmesi Yöntemleri Kullanılarak İncelenmesi: Tanımlayıcı Araştırma

Serel ÖZMEN AKYOL[a], Nurdan SEZGİN[b]

[a]Department of Computer Engineering, Kütahya University of Health Sciences Faculty of Engineering and Natural Sciences, Kütahya, Türkiye
[b]Department of Forensic Sciences, Kütahya University of Health Sciences Faculty of Engineering and Natural Sciences, Kütahya, Türkiye

ABSTRACT Objective: The purpose of this study is to determine how ear measurements using machine learning approaches effect sex estimation. **Material and Methods:** Biometric features are used in forensics to detect or verify individuals. In this study, the effect of ear measurements on one of the biometric characteristics on sex estimation was investigated. Anthropometric landmarks on the faces of 345 persons were identified for this purpose, and a data set of 36 characteristics was created by measuring distances between these landmarks. Unlike many other studies in the field of image processing, measurements obtained on biometric front face, side face, and ear images, as well as a data set containing age information, were compared in the literature, as were different machine learning approaches and accuracy rates. **Results:** As a result of the findings, an artificial neural network-based sex estimate approach that appears to be consistent with the data set was developed. Two different artificial neural network models with the same structural features were built and trained. The first model uses all the data set's features as input parameters, whereas the second model does not. After ear measurements were included, the model's classification accuracy increased from 82.6% to 92.2%. **Conclusion:** When combined with other anthropometric parameters, ear measurements, one of the biometric characteristics, have been demonstrated to improve the success rate of sex estimation.

Keywords: Sex estimation; face biometrics; machine learning algorithms; artificial neural networks; artificial intelligence

ÖZET Amaç: Bu çalışmanın amacı, makine öğrenmesi yöntemleri kullanarak yapılan kulak ölçümlerinin cinsiyet tahminine ne kadar etki ettiğini incelemektir. **Gereç ve Yöntemler:** Adli alanda kişilerin tespit edilmesi ya da doğrulanması amacıyla biyometrik özelliklerden faydalanılmaktadır. Çalışmada, biyometrik ölçümlerden biri olan kulak ölçümlerinin cinsiyet tahminine etkisi incelenmiştir. Bu amaçla 345 kişiye ait görüntüden uzman tarafından belirlenen antropometrik noktalara göre noktalar arasındaki mesafe ölçümleri milimetrik olarak kaydedilerek 36 özniteliğe ait ölçümlerin yer aldığı bir veri seti oluşturulmuştur. Literatürde görüntü işleme alanında yapılan birçok çalışmadan farklı olarak biyometrik ölçümlerin yapılarak oluşturulduğu ön yüz, yan yüz ve kulak ölçümleri ile yaş bilgisinin yer aldığı veri seti farklı makine öğrenmesi yöntemleri uygulanarak doğruluk oranları karşılaştırılmıştır. **Bulgular:** Analizler sonucunda, veri setine uyumlu olduğu görülen yapay sinir ağları yöntemi ile cinsiyet sınıflandırması yapılmıştır. Aynı yapı özelliklerine sahip iki farklı yapay sinir ağı modeli oluşturularak eğitilmiştir. İlk modelde veri setinde yer alan tüm öznitelikler giriş parametresi olarak kullanılırken, ikinci modelde kulak ölçümlerine ait özniteliklere yer verilmemiştir. Kulak ölçümlerinin dâhil edilmesi ile modelin sınıflandırma doğruluk oranı %82,6'dan %92,2'ye yükselmiştir. **Sonuç:** Biyometrik özelliklerden biri olan kulak ölçümlerinin diğer antropometrik özellikler ile kullanıldığında kişinin cinsiyetinin tahmin edilmesinde başarı oranını artırdığı görülmüştür.

Anahtar Kelimeler: Cinsiyet tahmini; yüz biyometrisi; makine öğrenmesi algoritmaları; yapay sinir ağları; yapay zekâ

Face, which reveals identity, age and feelings is an important information in sex classification.[1] A large amount of visual data is produced as technology advances, majority of which is video, audio, and images.[2] Automatic sex classification is used in the virtual world, mostly in image processing applica-

Correspondence: Nurdan SEZGİN
Department of Forensic Sciences, Kütahya University of Health Sciences Faculty of Engineering and Natural Sciences, Kütahya, Türkiye
E-mail: nurdan.sezgin@ksbu.ed

Peer review under responsibility of Turkiye Klinikleri Journal of Forensic Medicine and Forensic Sciences.

Received: 02 Mar 2023 ***Received in revised form:*** 22 May 2023 ***Accepted:*** 01 Jun 2023 ***Available online:*** 06 Jun 2023

tions, and has a wide range of applications, including smart user interfaces, visual surveillance, and statistical dataset.[1,2]

Artificial Neural Networks (ANN) and Deep Learning (DL) are widely used in literature in fields such as smart and healthcare treatment, forensics, entertainment, interactive systems, and security.[2] In forensic research, Machine Learning (ML) methods are also used to identify and classify sex.[3,4] ML techniques, will play a significant role in development of decision-making procedures in forensic field.

In the forensic profession, biometric features, which are unique, indelible, and unpredictable structures, are used for identification or verification. Because of these advantages, it has become the most popular field in security systems.[5]

Fingerprint, palm print, retina, voice, gait, ear, handwriting, DNA, facial structure etc. are the main biometric characteristics.[6] A study explores face recognition and tracking in video images. Study uses color-based algorithm for skin-like face recognition. ANN utilized as neural-based filter with multilayer perceptron structure.[7] The effects of different activation functions on video images and forward-feedback ANN are investigated.

However, ear measurements outperform face measurements in terms of diagnosis and sex estimation (SE) because they require less spatial resolution, have a more regular color distribution, and are less affected by light changes. Because each individual's ear form and size differ, the ear is classified as a biometric feature with a fixed structure that changes little with age.[6]

Study uses ear biometrics for image processing-based identification.[8] Study trains software using analysis of images ear database images. Matching success in ear images from the same angle was 96%, and 72% when the angle was changed.

It was determined to improve image recognition success by combining Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Discriminant Common Vector Approach (DCVA) methods with a Support Vector Machine (SVM) in a study that compared success performances of different methods in ear biometry classification.[9] The MIDAS database was used for this, which contains left and right face images of 50 people. When feature vectors identified by PCA, LDA, and DCVA methods are performed with SVM instead of standard measures (Euclidean distance), it has been claimed that success rate increases.

Study uses ear shape for person identification.[6] The first part of the study, uses anthropometric canons to identify ear region. Second module utilized PCA for similarity analysis and identification. System achieves 80% accuracy in decision making with 35 profile pictures and 15 ear data. Combining ear, nose, lips, and forehead enhances person identification.[6]

SVM, neural network, and Adaboost methods were used to estimate sex using 3D ear images in University of Notre Dame (UND) database.[10] Researchers show 3D ear features outperform 2D methods for sex classification.

As stated before, there are successful studies in which sex analysis is done using different image analysis methods. Images form datasets for analysis. Limited studies on image-derived geometric measurements. Study uses ear images, anthropometric landmarks, distance, and area calculations for age and SE.[11] Convolutional neural networks were used to build the model. Model achieves 94% SE accuracy, 52% age estimation accuracy, superior appearance-based methods.

Study uses 509 mandible measurements to estimate sex using linear and angular parameters. Logistic Regression, Discriminant Analysis, and ANN methods were used to analyze measurement parameters, and based on shared accuracy rates, it was determined that ANN is a good SE tool that can be used in the field of forensics due to its high accuracy rate.[3]

Ear images aid in identity verification using biometric factors. In this paper, the effect of ear biometry on SE is examined. Expert measures evaluated ear biometric measurements using facial images. Results aid in estimating sex using face images. Larger sample size needed for effective sex identification models. ML algorithms improve forensic science identification using face images.

MATERIAL AND METHODS

Humans have a simpler classification of sex than machines. Machines face challenges in angles, positions, age, and facial expressions.[12] ML techniques are employed for classification/estimate methods to overcome challenges.

There are four stages to ML techniques used. The first step is to create database, and the second step is to preprocess measurement results obtained from images. Data preprocessing involves data conversion, grouping, missing values management, and noise cleaning.[13] The ML model is developed in the third stage based on data type and characteristics (Figure 1).[14] ML techniques are used to make estimations in the final stage.

ML techniques used in literature for feature selection, model development, evaluation, and estimation. Identify appropriate approach based on data structure, dependent/independent variables, and analysis goals. All of the study's independent variables are continuous variables. Variables are features determined from images' measurements. For feature selection, no ML techniques were utilized. ML techniques for SE classification based on categorical structure.

The landmarks and explanations used in the study are as follows:

Otobasion superior (Obs): It is the helix's attachment point in the temporal area, defining the top border where the ear conjugates the face.[15]

Otobasion inferior (Obi): It is the point at which the earlobe attaches to the cheek.[15]

Gonion (Go): They are the outermost of the angle formed by the lower jaw body and lower jaw arm.*

Nasion (N): It is the point on the median sagittal line where the nasal bone meets the frontal bone.*

Gnathion (Gn): On the median sagittal line, it is the lowest point on the lower jaw.*

Subnasal (Sbn): The location where the middle part of the nostrils meets the upper jaw.*

Pronasal (Prn): On the median sagittal line, it is the most projecting point at the tip of the nose.*

Stomion (Sto): It is where the lower and upper lips meet the median sagittal line, cutting the midline.*

Supramental (Sm): It is the deepest point of the concavity that runs from the lower lip to the chin.[16]

Alare (Al): It is the most projecting points that the nose wings make towards the sides.*

Glabella (G): On the median sagittal line, it is the point between the two eyebrows that project forward.*

Chellion (Ch): It refers to the locations on the sides of the mouth where the lower and upper lips meet.*

Tragus (T): It is the part of the auditory canal that projects in front of and above it.[16]

Superaurale (Sa): It is the highest point of the auricle.*

Subaurale (Sba): It is the lowest point of the auricle.*

Postaurale (Pa): It is the outermost point of the posterior arch of the auricle.*

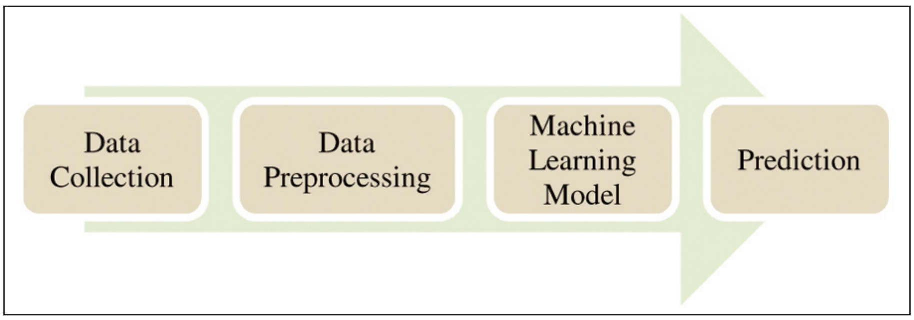

FIGURE 1: Prediction model with Machine Learning techniques.*

*Descriptions of gonion, nasion, gnathion, subnasal, pronasal, stomion, alare, glabella, chellion, superaurale, subaurale and postaurale points are quoted from Prof. Dr. İzzet Duyar's unpublished textbook (Duyar İ. Antropometri, Ankara Üniversitesi, DTCF Antropoloji Bölümü, 2000;s:11-15).

Preaurale (Pra): It is the front of the ear.[15]

Intertragic notch (Intno): It is the deep notch between the tragus and the antitragus.[16]

DATASET

To determine measurements used in this study, images in the data set created by Sezgin's doctoral thesis prepared in accordance with Helsinki Declaration 2008 principles were used.[17] The data obtained in that study were approved by Non Interventional Scientific Research Ethics Committee of İstanbul University Cerrahpaşa Faculty of Medicine with the approval number 114,364 dated 01.07.2014. The images in data set were taken with a flat lighting setup with 800 ASA at 50 mm and 1 m distance. First, using GIMP, anthropometric landmarks were placed on face in each image (Figure 2). The distances between the landmarks determined following this method were measured in millimeters (Figure 3). Photos are published with consent of the participant.

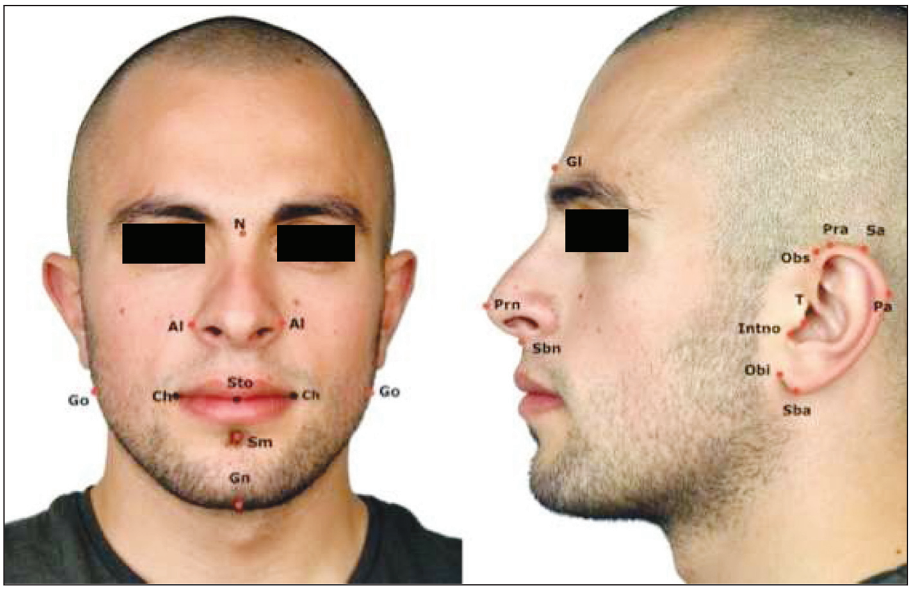

FIGURE 2: Landmarks (N: Nasion; Al: Alare; Ch: Chellion; Sto: Stomion; Go: Gonion; Sm: Supramental; Gn: Gnathion; Gl: Glabella; Prn: Pronasal; Sbn: Subnasal; Obs: Otobasion superior; Obi: Otobasion inferior; Pra: Preaurale; Sa: Superaurale; Pa: Postaurale; Sba: Subaurale; Intno: Intertragic notch; T: Tragion).

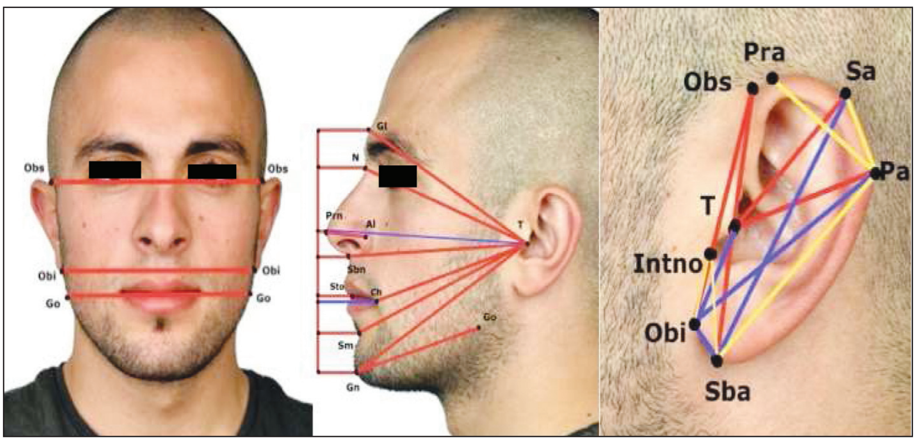

FIGURE 3: Measurements between landmarks.

N: Nasion; Al: Alare; Ch: Chellion; Sto: Stomion; Go: Gonion; Sm: Supramental; Gn: Gnathion; Gl: Glabella; Prn: Pronasal; Sbn: Subnasal; Obs: Otobasion superior; Obi: Otobasion inferior; Pra: Preaurale; Sa: Superaurale; Pa: Postaurale; Sba: Subaurale; Intno: Intertragic notch; T: Tragion.

By examining the images in the database, measurements are made according to the measurement points/landmarks determined by the expert. As a result, the data set created includes measurements of 36 features for 345 samples.

Data preprocessing applied after collection. Data preprocessing is crucial for ML technique development. Improve ML model accuracy, efficiency, consistency by identifying missing values, cleaning raw data. Missing data, repeated measurements, data set division, independent variables scaled.

The data set was normalized, and all analyses were performed on it. Normalization is the process of rescaling data from 0 to 1. In the study, the Min-Max method, which is one of the most commonly utilized normalizing methods, was used. Calculates normal distribution values using formula given in Equation 1, comparing mean to standard deviation:[18]

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (1)$$

COMPUTATIONAL TECHNIQUES

Analyses was implemented on MATLAB (MATLAB is a registered trademark of The MathWorks, Inc.). The dataset is divided into training and test datasets at random. dataset: 230 samples trained, 115 samples analyzed for system testing.

Different ML methods were applied, and results were examined while deciding on the best classification method for the dataset.

Decision trees classify data using induction and tree-shape structure. Decision trees are preferred for easy creation and interpretation.[19] All decision tree sub-models created in the study used Gini's diversity index as the Split criterion. As a result, in terms of maximum number of divisions, fine decision tree model is 10, coarse decision tree model is 4, and medium decision tree model is 30.

Discriminant analysis is a ML method for categorical data identification.[20] Proposed diagonal used in discriminant analysis for limited data. Low training and test accuracy rates required full covariance analysis.

Bayes Classifier uses existing classifier data to classify data.[21] Gaussian and Kernel methods represent numeric predictors in NB analysis.[22]

Data transferred to high dimension, separated using SVM linearly. Then it verifies that they are separated by the highest boundary achievable.[23] Models are developed in the study using various SVM kernel functions (Linear, Quadratic, Cubic, Gaussian). The box restriction level in all SVM models is set at 1.

The K-Nearest Neighbor (KNN) is one of the most well-known and used algorithms among ML methods. Classification uses closeness between selected and closest features.[24] The nearest neighbor numbers (1-10-100) differ while using the Euclidean Distance metric in the Fine, Medium, and Coarse KNN models. The number of nearest neighbors in the Cosine and Cubic KNN models is set to 10. The distance metrics of two methods, however, differ. Cosine KNN model uses cosine metric distance, Cubic KNN uses minkowski.

Boosted Trees are ensemble classifiers for classification and regression.[25] AdaBoost ensemble method and decision tree learner type in study.

Bagged Trees is another ensemble classifier that aims to increase accuracy and reliability of ML methods.[26] Combining random trees in training sets. The Bag method and decision tree learning developed Bagged Trees model.

When all features are used in the study, the Subspace Discriminant ensemble classifier method is used to create different training sets. Each basic learning algorithm is trained on smaller training sets using the sub-space method, which divides the existing training set into small training sets.[27] Ensemble method Subspace, Discriminant learner type, 30 learners, and Subspace dimension 18 are selected in the model.

ANNs are hardware and software realizations of nerve cells in human brain in the digital world.[28] Because of its high performance, ANN is commonly used in a variety field of the classification problems.

Different neural network models were created in this study depending on the number of layers and activation function. The activation function in single-layer Narrow, Medium, and Wide Neural Network models was Rectified Linear Unit (ReLU), and number of neurons were 10, 25, and 100, respectively. The ReLU activation function was used with 10 neurons in each layer in 2-layer Bilayered and 3-layer Trilayered Neural Network models.

Different ML methods were used to compare classification performance, and the classifier method that would be used in the study was chosen accordingly.

To compare success of classification algorithms, various metrics are used. One of them is confusion matrix, which is used to determine the classifier's accurate classification value. Based on actual and expected classifications generated by a classifier, the confusion matrix evaluates the accuracy of the classifier.[23] Using confusion matrix for train and test results, correctly classified positive sample (True Positive, TP), correctly classified negative sample (True Negative, TN), misclassified positive sample (False Positive, FP) and misclassified negative sample (False Negative, (FN) values were calculate.[29]

Accuracy (ACC), Precision, Recall and F-score values of the model are determined according to these values. These metrics are calculated with formulas given in Equations (2), (3), (4), and (5):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{5}$$

In this study, different ML methods were evaluated according to accuracy value, which is widely used in classification for train and test datasets.[22] Table 1 shows percentage values for training and test classification accuracy obtained from models during the method selection stage.

According to the model classification accuracy values in Table 1, among the applied ML methods, ANN and Subspace Discriminant methods provide the most accurate predictions using dataset.[4]

Subspace Discriminant is a preferred ensemble learning method due to its fast classification and optimum memory usage. It's a method for improving a single model's estimation performance by training multiple models and combining their predictions. It has been demonstrated that when compared to a single classifier model, a higher accuracy rate can be obtained.[30] However, depending on data, the method's classification accuracy differs. As a result, analyses will be conducted using ANN method, which provides the best results in this study.

ANN

ANN are a form of computing for predicting a set of outputs from a set of input variables.[31] ANN is made up of neurons connected by numerically weighted links. Each neuron calculates weighted sum of input connections and compares it to threshold value. Depending on activation function used, each neuron's output differs.[3]

An ANN model is consisting of three layers: an input layer, a hidden layer, and an output layer. A basic ANN model structure is shown in Figure 4. The number of hidden layers and activation function used depend on data structure and ANN model that was constructed.

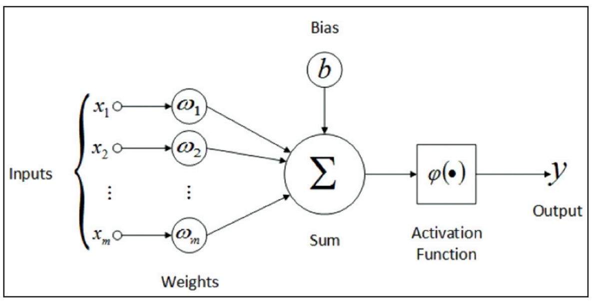

FIGURE 4: Structure of Artificial Neural Networks models.

TABLE 1: Accuracy values of ML methods.

		Accuracy rates	
ML techniques		**Results of train**	**Results of test**
Decision tree	Fine tree	79.2%	82.3%
	Medium tree	79.2%	83.4%
	Coarse tree	81.1%	85.8%
Discriminant analysis	Linear discriminant	70.8%	74.6%
	Quadratic discriminant	62.5%	64.2%
Logistic Regression		68.8%	70.3%
Naive bayes network	Gaussian naive bayes	75.4%	74.3%
	Kernel naive bayes	70.2%	69.5%
Support vector machines	Linear SVM	77.2%	76.5%
	Quadratic SVM	72.5%	75.1%
	Cubic SVM	70.8%	70.0%
	Gaussian SVM	77.9%	76.3%
K-Nearest neighbors	Fine KNN	69.5%	70.9%
	Medium KNN	73.5%	71.8%
	Coarse KNN	67.3%	66.2%
	Cosine KNN	68.2%	67.3%
	Cubic KNN	74.1%	74.6%
Ensemble classifiers	Boosted trees	70.9%	72.3%
	Bagged trees	80.2%	82.4%
	Subspace discriminant	89.1%	90.4%
Neural network	Narrow neural network	88.4%	91.2%
	Medium neural network	81.9%	84.6%
	Wide neural network	82.7%	85.2%
	Bilayered neural network	86.3%	82.5%
	Trilayered neural network	82.1%	83.0%

ML: Machine Learning; SVM: Support Vector Machine; KNN: K-Nearest Neighbor.

RESULTS

Measurements are performed based on expert-defined measurement points/features obtained by examining images in the database. As a result, the dataset obtained contains measurements of 36 features for 345 samples (151 female, 194 male). Table 2 shows the dataset's features, which include front face, side face, and ear measurements, as well as age information.

The purpose of this study was to see how ear measurements affected SE. Two ANN models (ANN Model-I and ANN Model-II) with different features were created and compared for this purpose. All measurement values, including face and ear measurements, were included in the first model, however ear measurements, were not included in the second model, and their contribution to classification accuracy was evaluated.

Age, side face, front face, and ear measurements were used as input features in the ANN Model-I. All the features are used as predictors for SE.

Ear measurements were not included as input parameters in the ANN Model-II; instead, only age, side, and front face measurements were used. The effect of ear measurements on SE was studied by comparing the estimation accuracy performances of the two models.

First, different activation functions, layers, and neuron numbers were tested for both models, and models were optimized and developed accordingly. As a result, both models employ a three-layer optimal network structure. The ANN model's first layer

TABLE 2: Facial features measurements.

Features		Minimum	Maximum
1	Age	18	92
Front face measurements			
2	Otobasion superior-Otobasion superior	150.65	201.62
3	Otobasion inferior-Otobasion inferior	126.67	193.23
4	Gonion-Gonion	115.81	1148.83
Side face measurements			
5	Nasion-Gnathion	120.60	174.60
6	Nasion-Subnasal	51.01	82.48
7	Nasion-Pronasal	42.36	110.84
8	Nasion-Stomion	75.03	146.10
9	Nasion-Supramental	88.49	1214.89
10	Pronasal-Subnasal	15.15	34.57
11	Pronasal-Alare	24.00	51.72
12	Glabella-Nasion	18.00	43.26
13	Glabella-Pronasal	67.58	109.22
14	Gnathion-Gonion	79.33	154.99
15	Tragus-Glabella	127.19	198.65
16	Tragus-Nasion	110.30	178.62
17	Tragus-Pronasal	50.31	204.10
18	Tragus-Subnasal	112.92	181.32
19	Tragus-Chellion	99.99	161.09
20	Tragus-Supramental	115.65	193.25
21	Tragus-Gnathion	125.86	1321.87
22	Chellion-Gnathion	35.72	68.75
Ear measurements			
23	Otobasion superior-Otobasion inferior	33.60	99.72
24	Tragus-Otobasion superior	25.20	52.87
25	Tragus-Superaurale	38.68	74.63
26	Tragus-Postaurale	29.03	64.76
27	Tragus-Subaurale	24.12	58.80
28	Tragus-Otobasion inferior	16.32	47.95
29	Superaurale-Subaurale	65.47	119.59
30	Otobasion inferior-Subaurale	3.79	30.71
31	Otobasion inferior-Postaurale	41.91	97.87
32	Preaurale-Postaurale	26.29	74.21
33	Postaurale-Subaurale	34.11	94.32
34	Superaurale-Postaurale	19.47	74.88
35	Intertragic notch-Otobasion inferior	9.98	44.65
36	Intertragic notch-Subaurale	18.68	50.43

size is 35, the second layer size is 13, and the third layer size is 102. As the activation function, ReLU was used. Because of its low complexity and rapid operation, ReLU, which is used as the activation function in the model, is widely used in computer vision, visual feature extraction, and DL methods.[32] All negative states are converted to zero with this function, and the network's nonlinear properties are increased.[33] The ReLU activation function used is given in Equation (6).[34]

$$x_{out} = \max(0, wx_{in} + b) \tag{6}$$

The input and output values of the equation are indicated by the x_{in}, and x_{out} expressions, while the weight and bias values are expressed by the w and b expressions.

The first model, ANN Model-I, includes all the measurements and features used in the study. Figure 5 shows the ANN Model-I, which is composed of 36 features such as age, front face, side face, and ear measurements that are used in the estimation of sex using the specified network model.

The training and test confusion matrices used in calculating the accuracy values of Model-I are given in Figure 6 and Figure 7. The Training Accuracy value of the model using Bayesian optimization was found to be 83.5% and the Test Accuracy value to be 92.2%.

The ANN Model-II structure, which was created to examine effect of ear measurements on SE, is given in Figure 8. In the model, only 22 features with age, front face and side face measurements are included as input parameters.

Training and test confusion matrices of ANN Model-II are given in Figure 9 and Figure 10. While

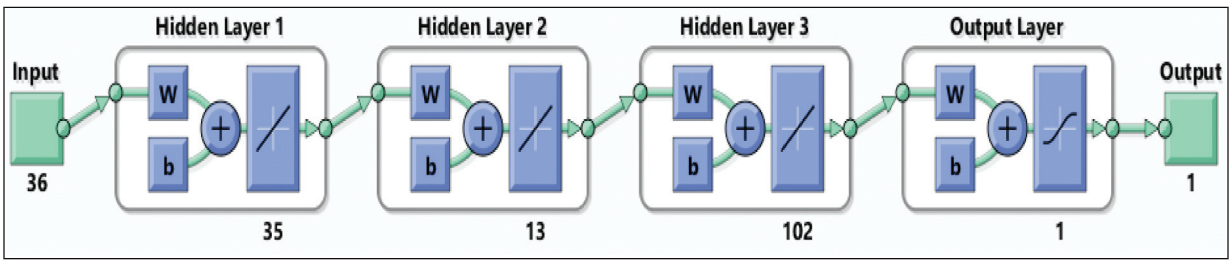

FIGURE 5: Artificial Neural Networks Model-I (all features).

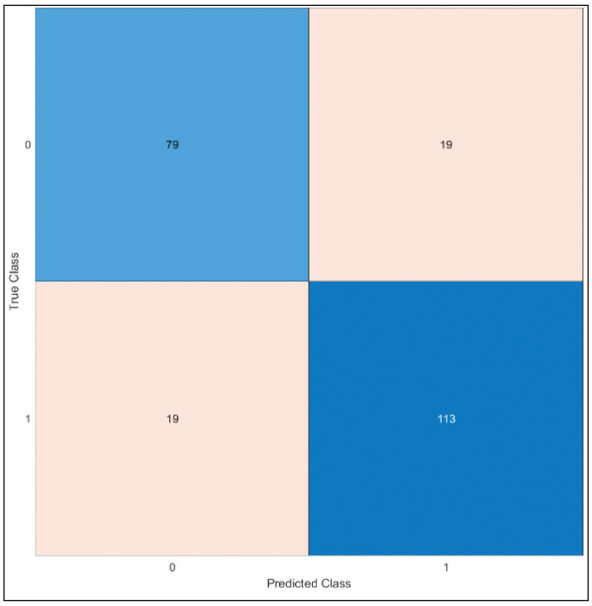

FIGURE 6: Model-I Confusion matrix (train).

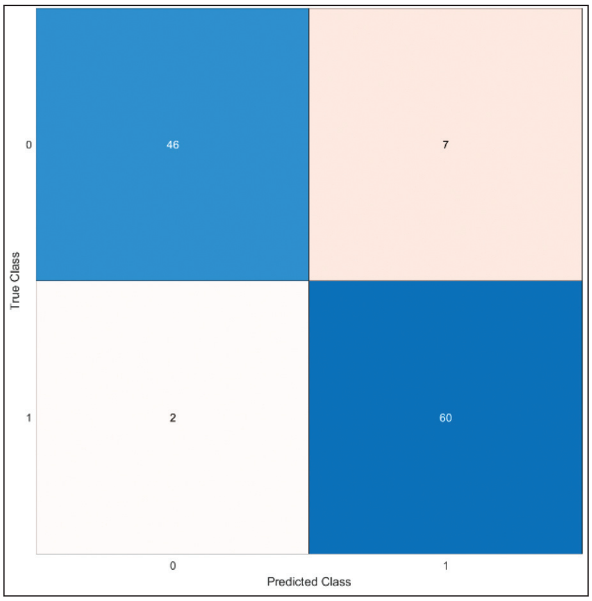

FIGURE 7: Model-I Confusion matrix (test).

Training Accuracy value of ANN Model-II is 80.9%, Test Accuracy value is calculated as 82.6%.

DISCUSSION

Local binary patterns (LBP)-Histogram bins were learned as compact face representation for sex classification in a study that investigated sex classification on real-life faces obtained under unconstrained conditions. By adopting SVM with selected LBP-Histogram bins, a classification rate of 94.81% was achieved in the Labeled Faces in the Wild database.[1]

A study developed a new filter structure based on the MatConvNet-1.0-beta15 architecture, which is based on regional convolutional neural network model in DL algorithm. To improve SE accuracy, filter sizes and contents were changed, and features that could be used in sex classification were extracted from face images. According to the study, DL algorithm's features were classified by SVM, and sex classification was performed with an average of 94% success based on complexity matrix calculation.[2]

In a study where different distance criteria were intuitively limited according to the Borda voting method and success rate of the system was evaluated, 80% success was achieved in tests performed on 15 data in the database.[6] With the experiments conducted on the public UND ear biometric database, sex classification was achieved with an average accuracy of 92% in the F dataset and approximately 92% in the J2 dataset.[10] In a study in which it was stated that the methods based on appearance were superior to the methods based on geometric features, an accuracy of 94% was obtained in sex classification and 52% in age classification.[11]

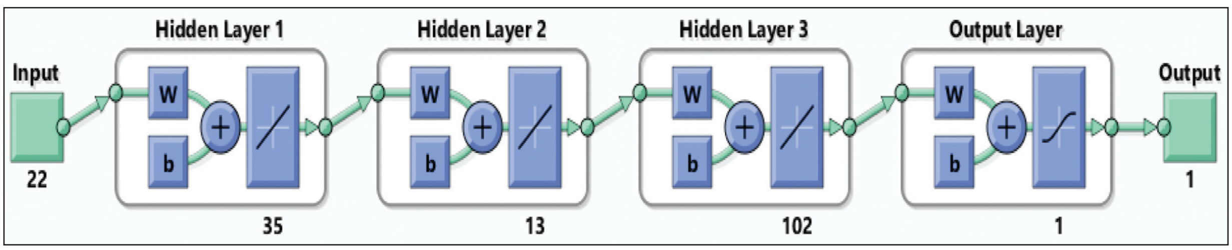

FIGURE 8: Artificial Neural Networks Model-II (excluding ear measurements).

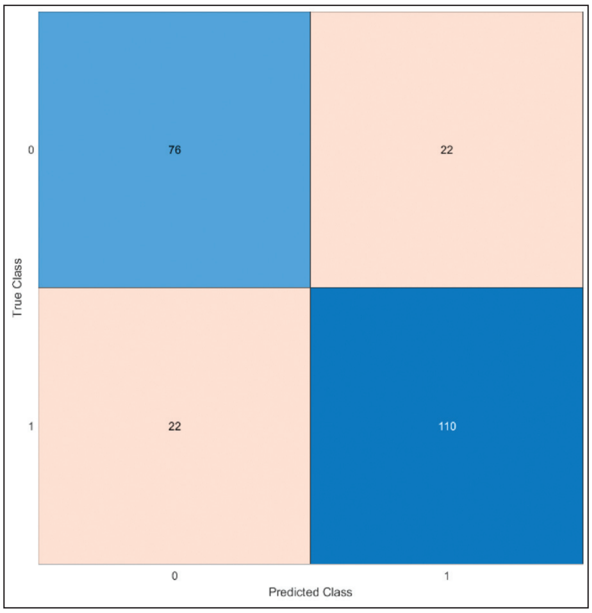

FIGURE 9: Model-II Confusion matrix (train).

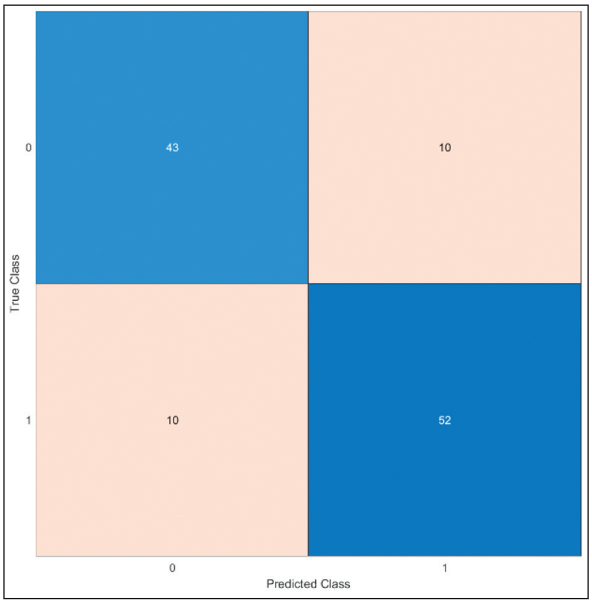

FIGURE 10: Model-II Confusion matrix (test).

In this study, the Training Accuracy value of the ANN Model-I using Bayesian optimization was found to be 83.5% and the Test Accuracy value to be 92.2%. While Training Accuracy value of ANN Model-II is 80.9%, Test Accuracy value is calculated as 82.6%.

The effect of ear measurements made using ML approaches on SE was investigated in this study. Analyzes were made on a dataset containing the measurement results. While creating dataset for this study, all images were standardized in a biometric template. This is important in that metric measurements give results close to real-time measurements. It also differs in terms of providing weight to side and ear images in terms of measurement when compared to research in which front images are usually analyzed.

The data set, which includes the actual measurement results of 345 samples, is divided into training and test data and ML methods commonly used for classification in the literature (Decision Tree, Discriminant Analysis, Naive Bayes Network, Logistic Regression, SVM, KNNs, Boosted Trees, Bagged Trees was trained with Subspace Discriminant and Neural Network) to determine the highest performance classifier. As a result, the ANN and Subspace Discriminant methods achieved the highest accuracy rate among the models constructed using the specified methods. Because the Subspace Discriminant method's accuracy in classification varies based on the data and has a relatively low accuracy rate, advanced analyses were performed using the ANN method. The aim of advanced analyses is to investigate the effect of ear measurements on SE, which has never been studied before in the literature.

CONCLUSION

The dataset of this study was trained using widely known ML methods and the best performing classifier was determined. The effect of ear measurements on SE was evaluated by ANN method, and high-accuracy SE was performed. For this purpose, two neural network models with same structure are trained for different features. While all measurement results were included in one model, ear measurements were subtracted from the other model and its effect on classification performance was evaluated. When accuracy values of two models were compared, classification accuracy of model increased from 82.6% to 92.2% with the addition of ear measurements.

In the subject of identification, forensic sciences, particularly sex and age estimation research, play a vital role. ML with various models is widely used to test features extracted from facial images. Given that facial images cannot always be obtained from the front, side face and ear images are undeniably important subjects for this field. It is investigated to what extent the examination of many factors and facial features affects identification. This study investigated the effect of ear measurements on SE when paired with other anthropometric parameters and concluded that it raised the success rate. In this regard, it is stated that the data generated by using ML algorithms will contribute to forensic studies of SE from facial images.

In the future, research to increase classification performance can be aided by the development of simplified models and estimation features.

Source of Finance

During this study, no financial or spiritual support was received neither from any pharmaceutical company that has a direct connection with the research subject, nor from a company that provides or produces medical instruments and materials which may negatively affect the evaluation process of this study.

Conflict of Interest

No conflicts of interest between the authors and / or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Authorship Contributions

Idea/Concept: *Serel Özmen Akyol, Nurdan Sezgin;* ***Design:*** *Serel Özmen Akyol;* ***Control/Supervision:*** *Serel Özmen Akyol, Nurdan Sezgin;* ***Analysis and/or Interpretation:*** *Serel Özmen Akyol;* ***Literature Review:*** *Nurdan Sezgin, Serel Özmen Akyol;* ***Writing the Article:*** *Serel Özmen Akyol, Nurdan Sezgin;* ***Critical Review:*** *Serel Özmen Akyol, Nurdan Sezgin.*

REFERENCES

1. Shan C. Learning local binary patterns for gender classification on real-world face images. Pattern Recognit Lett. 2012;33(4):431-7. [Crossref]
2. Metlek S, Kayaalp K. Derin öğrenme ve destek vektör makineleri ile görüntüden cinsiyet tahmini [Image gender prediction with deep learning and support vector machines]. Düzce Üniversitesi Bilim ve Teknol Derg. 2020;8:2208-28. [Crossref]
3. Patil V, Vineetha R, Vatsa S, Shetty DK, Raju A, Naik N, et al. Artificial neural network for gender determination using mandibular morphometric parameters: a comparative retrospective study. Cogent Eng. 2020;7(1):1-12. [Crossref]
4. Alunni V, Jardin Pd, Nogueira L, Buchet L, Quatrehomme G. Comparing discriminant analysis and neural network for the determination of sex using femur head measurements. Forensic Sci Int. 2015;253:81-7. [Crossref] [PubMed]
5. Jain AK. Biometric recognition: how do I know who you are? Image Analysis and Processing-ICIAP. 2005:19-26. [Crossref]
6. Nabiyev V. Kulak biyometrisine göre kimlik tespiti. 2. Mühendislik ve Teknoloji Sempozyumu Bildiriler Kitabı. 2009:277-87. [Link]
7. Eser S. Yapay Sinir Ağları ile Yüz Sezimi ve Takibi, Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, 2006. [Cited: Jan 20, 2023]. Available from: [Link]
8. Yenisarı E. Kulak biyometrisi kullanarak görüntü işlem tabanlı kimlik tespiti [Yüksek lisans tezi]. Denizli: Pamukkale Üniversitesi; 2015.
9. Kaçar Ü, Kırcı M, Güneş EO, İnan T. DVM ile kulak biyometrisi sınıflandırmasında TBA, DAA ve AOVY'nın performansının karşılaştırılması. 2015 23rd Signal Processing and Communications Applications Conference, SIU 2015-Proceedings IEEE. 2015:15-8. [Crossref]
10. Lei J, Zhou J, Abdel-Mottaleb M. Gender classification using automatically detected and aligned 3D ear range data. Proceedings-2013 International Conference on Biometrics, ICB. 2013:1-7. [Crossref] [PMC]
11. Yaman D, Eyiokur Fİ, Sezgin N, Ekenel HK. Age and gender classification from ear images. IWBF 2018-Proc. 2018 6th Int. Work. Biometrics Forensics. 2018:1-7. [Crossref]
12. Verma R, Bhardwaj N, Singh PD, Bhavsar A, Sharma V. Estimation of sex through morphometric landmark indices in facial images with strength of evidence in logistic regression analysis. Forensic Sci Int Reports. 2021;4:100226. [Crossref]
13. Özçakır FC, Çamurcu AY. Birliktelik kuralı yöntemi için bir veri madenciliği yazılımı tasarımı ve uygulaması [A data mining software design and application for association rule technique]. Istanbul Commer Univ J Sci. 2007;12(2):21-37. [Link]
14. Şenol A, Canbay Y, Kaya M. Trends in outbreak detection in early stage by using machine learning approaches. Bilişim Teknol Derg. 2021;14(4):355-66. [Link]
15. Auricular Acupuncture [Internet]. [Cited: May 19, 2023]. Anthropometry of The Outer Ear. Available from: [Link]
16. Tıp Terimleri Sözlüğü [İnternet]. [Erişim tarihi: 19 Mayıs 2023]. Erişim linki: [Link]
17. Sezgin N. Dijital Ortamda İnsan Yüzlerindeki Yaşa Bağlı Değişimlerin Metrik Olarak İncelenmesi, Doktra Tezi, İstanbul Üniversitesi, Adli Tıp Enstitüsü, 2017. [Cited: Jan 18, 2023]. Available from: [Link]
18. Tunç A, Ülger İ. Veri madenciliği uygulamalarında özellik seçimi için finansal değerlere binning ve five number summary metotları ile normalizasyon işleminin uygulanması. XVIII. Akademik Bilişim Konferansı. Adnan Menderes Üniversitesi. 2016:1-8. [Link]
19. Albayrak AS, Yılmaz ŞK. Veri madenciliği: karar ağacı algoritmaları ve İMKB verileri üzerine bir uygulama [Data mining: decision tree algorithms and an application on ise data]. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilim Fakültesi Derg. 2009;14(1):31-52. [Link]
20. Filiz Z, Yaprak B. A Study on classifying parenting styles through discriminant analysis. J Theory Pract Educ. 2009;5(2):195-209. [Link]

21. Bilgin M. Gerçek veri setlerinde klasik makine öğrenmesi yöntemlerinin performans analizi [Classical machine learning methods on real datasets performance analysis]. Breast. 2017;2(9):683-8. [Link]

22. Tantuğ AC. Metin sınıflandırma [Text classification]. Türkiye Bilişim Vakfı Bilgi Bilim ve Mühendisliği Derg. 2012;5(2):1-12. [Link]

23. Khehra BS, Pharwaha APS. Classification of clustered microcalcifications using MLFFBP-ANN and SVM. Egypt Informatics J. 2016;17(1):11-20. [Crossref]

24. Kılınç D, Borandağ E, Yücalar F, Tunali V, Şimşek M, Özçift A. KNN algoritması ve R dili ile metin madenciliği kullanılarak bilimsel makale tasnifi [Classification of scientific articles using text mining with kNN algorithm and R language]. Marmara Fen Bilim Derg. 2016;28(3):89-94. [Crossref]

25. Barkana BD, Saricicek I, Yildirim B. Performance analysis of descriptive statistical features in retinal vessel segmentation via fuzzy logic, ANN, SVM, and classifier fusion. Knowledge-Based Syst. 2017;118:165-76. [Crossref]

26. Bin Alam MS, Patwary MJA, Hassan M. Birth Mode Prediction Using Bagging Ensemble Classifier: A Case Study of Bangladesh. 2021 International Conference on Information and Communication Technology for Sustainable Development, ICICT4SD. 2021:95-9. [Crossref]

27. Ashour AS, Guo Y, Hawas AR, Xu G. Ensemble of subspace discriminant classifiers for schistosomal liver fibrosis staging in mice microscopic images. Health Inf Sci Syst. 2018;6(1):21. [Crossref] [PubMed] [PMC]

28. Burke HB, Goodman PH, Rosen DB, Henson DE, Weinstein JN, Harrell FE Jr, et al. Artificial neural networks improve the accuracy of cancer survival prediction. Cancer. 1997;79(4):857-62. [Crossref] [PubMed]

29. Acı Çİ, Çırak A. Türkçe haber metinlerinin konvolüsyonel sinir ağları ve Word2Vec kullanılarak sınıflandırılması [Turkish news articles categorization using convolutional neural networks and Word2Vec]. Bilişim Teknol Derg. 2019;12(3):219-28. [Crossref]

30. Karaca BK, Akşahin MF, Öcal R. EEG tutarlılık analizi ile multipl skleroz hastalığının belirlenmesi. Tıp Teknolojileri Kongresi. 2019:235-8. [Link]

31. Özmen-Akyol S. Estimating breast cancer class using artificial neural network and logistic regression methods. Eskişehir Türk Dünyası Uygul ve Araştırma Merk Bilişim Derg. 2022;3(1):26-31. [Crossref]

32. Çılgın C, Ünal C, Alıcı S, Akkol E, Gökşen Y. Metin sınıflandırmada yapay sinir ağları ile bitcoin fiyatları ve sosyal medyadaki beklentilerin analizi [In text classification, bitcoin prices and analysis of expectations in social media with artificial neural networks]. Mehmet Akif Ersoy Üniversitesi Uygulamalı Bilim Derg. 2020;4(1):106-26. [Crossref]

33. Özyurt F, Avcı E. İmge sınıflandırması için yeni öznitelik çıkarım yöntemi: ADD-TDA Algısal özet fonksiyonu tabanlı evrişimsel sinir ağ (ADD-TDA-ESA) [A new feature extraction method for ımage classification: DWT-SVD perceptual hash -based convolutional neural network (DWT-SVD-CNN)]. Türkiye Bilişim Vakfı Bilgi Bilim ve Mühendisliği Derg. 2019;12(1):30-8. [Link]

34. Tokgöz GK, Tekbıyık K, Kurt GK, Yarkan S. A Comparison on Energy Detector and CNN-based Detector. 2020 28th Signal Processing and Communications Applications Conference, SIU. 2020. [Crossref]