BMC Medical Imaging

RESEARCH

Open Access

Sex estimation with parameters of the facial canal by computed tomography using machine learning algorithms and artificial neural networks

Yusuf Secgin[1*], Seren Kaya[2], Oğuzhan Harmandaoğlu[3], Oğuzhan Öztürk[3], Deniz Senol[2], Ömer Önbaş[4] and Nihat Yılmaz[5]

Abstract

Background The skull is highly durable and plays a significant role in sex determination as one of the most dimorphic bones. The facial canal (FC), a clinically significant canal within the temporal bone, houses the facial nerve. This study aims to estimate sex using morphometric measurements from the FC through machine learning (ML) and artificial neural networks (ANNs).

Materials and methods The study utilized Computed Tomography (CT) images of 200 individuals (100 females, 100 males) aged 19–65 years. These images were retrospectively retrieved from the Picture Archiving and Communication Systems (PACS) at Düzce University Faculty of Medicine, Department of Radiology, covering 2021–2024. Bilateral measurements of nine temporal bone parameters were performed in axial, coronal, and sagittal planes. ML algorithms including Quadratic Discriminant Analysis (QDA), Linear Discriminant Analysis (LDA), Decision Tree (DT), Extra Tree Classifier (ETC), Random Forest (RF), Logistic Regression (LR), Gaussian Naive Bayes (GaussianNB), and k-Nearest Neighbors (k-NN) were used, alongside a multilayer perceptron classifier (MLPC) from ANN algorithms.

Results Except for QDA (Acc 0.93), all algorithms achieved an accuracy rate of 0.97. SHapley Additive exPlanations (SHAP) analysis revealed the five most impactful parameters: right SGAs, left SGAs, right TSWs, left TSWs and, the inner mouth width of the left FN, respectively.

Conclusions FN-centered morphometric measurements show high accuracy in sex determination and may aid in understanding FN positioning across sexes and populations. These findings may support rapid and reliable sex estimation in forensic investigations—especially in cases with fragmented craniofacial remains—and provide auxiliary diagnostic data for preoperative planning in otologic and skull base surgeries. They are thus relevant for surgeons, anthropologists, and forensic experts.

Clinical trial number Not applicable.

Keywords Sex estimation, Machine learning, Artificial neural network, Facial canal, Fallopian canal

*Correspondence:
Yusuf Secgin
yusufsecgin@karabuk.edu.tr

Full list of author information is available at the end of the article

Introduction

Identification is crucial for forensic anthropology and forensic medicine, and accounts for a significant proportion of the workload in these fields. In adults, factors such as sex, age, height, and weight form the first step in identifying an individual [1]. Sex determination is the key biological characteristic that plays the most important and pivotal role in forensic and medical identification. This is because, in situations where rapid action is required, sex reduces the potential identification possibilities by half and facilitates correct decisions regarding other identification characteristics. Sex determination is therefore crucial for the rapid and accurate resolution of forensic cases [2, 3]. DNA analysis is reliable for estimating sex. However, due to the high cost of DNA analysis, its limited availability, and the need for expertise to perform the analysis, more applicable methods such as osteometry and odontometry are preferred. Morphometric analyses are used to determine sex based on bones [4].

The facial nerve is the seventh cranial nerve and runs within the facial canal (FC-Fallopian canal). The facial nerve, which contains somatomotor, parasympathetic and sensory fibers, follows a complex course along the base of the skull. Understanding the anatomy of the FC and the facial nerve within it is critical during surgeries such as resection of skull base tumors that are closely associated with the nerve, particularly tumors in the cerebellopontine angle [5]. The FC, which extends from the internal acoustic meatus to the stylomastoid foramen, is divided into three segments: the labyrinthine segment, the tympanic segment, and the mastoid segment. The angle formed between the labyrinthine and tympanic segments is referred to as the first genu angle (FGA), while the angle between the tympanic and mastoid segments is referred to as the second genu angle (SGA) [6]. The study of the FC and the identification of its divisions contribute to the differential diagnosis of middle ear pathology. Furthermore, the study of the diameter, course and opening of the canal facilitates the identification of lesions in this area and helps to reduce facial nerve damage during surgical procedures [5].

Data mining is the process of discovering new data from previously known data sets [7]. The algorithms and methodologies used in this field facilitate the application of machine learning (ML) in healthcare fields such as medicine, forensic anthropology, and forensic medicine [8]. ML algorithms are a subset of artificial intelligence. By using ML techniques, sex estimation and identification can be done in a short time in the healthcare field through data processing [9]. Python, the most popular and general-purpose programming language in ML, is often used to write ML algorithms. The availability of various modules and libraries makes Python even more popular among data scientists [10]. One of the most common applications of ML in medicine is sex estimation [11–13]. Studies have shown that different ML algorithms are used for sex estimation and that these algorithms produce different results [11]. It is important to determine which algorithm has a higher estimation rate for sex estimation in medicine and forensic anthropology. Diagnostic techniques that use human-machine collaboration allow for more accurate results by leveraging the strengths of both humans and machines [14].

A review of the available literature revealed that sex estimation has not been previously performed using measurements taken from the FC. The aim of this study is to estimate sex using ML by utilizing clinically significant morphometric features of the FC, such as internal and external opening widths, segment widths, and genu angles. This study aims to address a gap in the current literature by exploring the potential of FC-based morphometry for sex estimation through ML methods.

Method

Study design, ethical clearance, and image processing

The present study was conducted retrospectively on the basis of a series of Computed Tomography (CT) scans obtained from 100 male and 100 female subjects aged between 19 and 65 years. Before the study started, ethical approval was obtained from Düzce University Non-interventional Health Research Ethics Committee (DU-HR-IRB) with protocol number 2024/237. Informed consent was waived due to lack of direct patient contact during data collection. The study was also conducted in accordance with the relevant guidelines and regulations as per the Declaration of Helsinki. CT images were obtained from Picture Archiving and Communication Systems (PACS) at Düzce University Faculty of Medicine, Department of Radiology, by retrospectively scanning between 2021 and 2024. The study was a retrospective study based on the hospital image archive. Therefore, informed consent is not required. The exclusion criteria encompassed the presence of pathology in the temporal bone and invasive intervention. Images that met the exclusion criteria in the archive system were selected using simple random sampling between 2021 and 2024. The minimum sample size was also determined by Power power analysis. Since the studies used as a guide for the analysis did not include algorithms, the number was partially increased and divided into equal sex ratios for the reliability of the study.

The radiologic images utilized in the study were acquired using a Siemens (model Somatom Definition AS) 128-slice CT device. The following values were scanned: kV (kilovoltage): 100, mAs (milliamperseconds): 49, scan time: 2 s, FOV (field of view): 163 millimeters, and slice thickness: 0.6 millimeters. A total of 18 parameters were measured in the axial, coronal, and sagittal planes on the right and left sides of the image. The measurement parameters are as follows;

1. Inner opening width of the FC (IFC): The width of the initial level of the FC was measured bilaterally on axial sections (Fig. 1).
2. Length of the labyrinthine segment (LSL): The length from the internal acoustic meatus to the fossa of the geniculate ganglion was measured bilaterally on axial sections (Fig. 1).
3. Width of the labyrinthine segment (LSW): The width at the midpoint of the labyrinthine segment was measured bilaterally on axial sections (Fig. 1).
4. Length of the tympanic segment (TSL): The distance between the fossa of the geniculate ganglion and the pyramidal eminence was measured bilaterally on axial sections (Fig. 2).
5. Width of the tympanic segment (TSW): The width at the midpoint of the TSL on the right side was measured bilaterally on axial sections (Fig. 2).
6. Length of the mastoid segment (MSL): The distance between the pyramidal eminence and the stylomastoid foramen was measured bilaterally on coronal sections (Fig. 2).
7. Width of the mastoid segment (MSW): The width at the midpoint of the MSL was measured bilaterally on coronal sections (Fig. 2).
8. First genu angle (FGA): The angle between the labyrinthine segment and the tympanic segment was measured bilaterally on axial sections (Fig. 3).
9. Second genu angle (SGA): The angle between the tympanic segment and the mastoid segment was measured bilaterally on sagittal sections (Fig. 3).

Data processing using machine learning algorithms and artificial neural networks

The implementation of machine learning (ML) algorithms and artificial neural networks (ANNs) was conducted using a Monster Abra A7 V12.5 model computer with an i5 operating system. The analysis was performed using the Python 3.9 programming language and the scikit-learn 1.1.1 (for ML algorithms) and TensorFlow (for ANN algorithms) framework. The 18 measurement parameters obtained from CT images were utilized in the input layer of the ML and ANN algorithms, while sex was designated as the output layer. The following machine learning algorithms were employed:

Linear Discriminant Analysis (LDA) performs classification by projecting high-dimensional data onto a lower-dimensional subspace to achieve maximum separation between classes [15]. Quadratic Discriminant Analysis (QDA) works similarly to LDA but assumes that different classes may have different covariance structures. It allows for more accurate drawing of nonlinear boundaries [15].

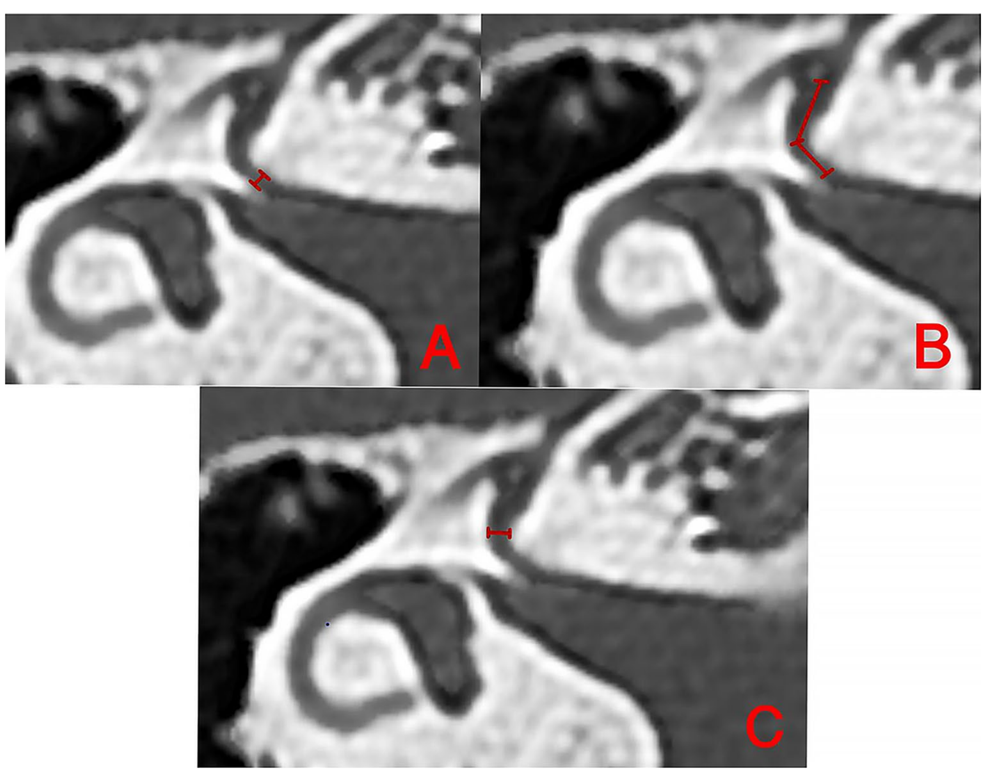


Fig. 1 Measurements related to the labyrinthine segment of the facial canal. **A**: Inner opening width of the facial canal (IFC), **B**: Length of the labyrinthine segment (LSL), **C**: Width of the labyrinthine segment (LSW)

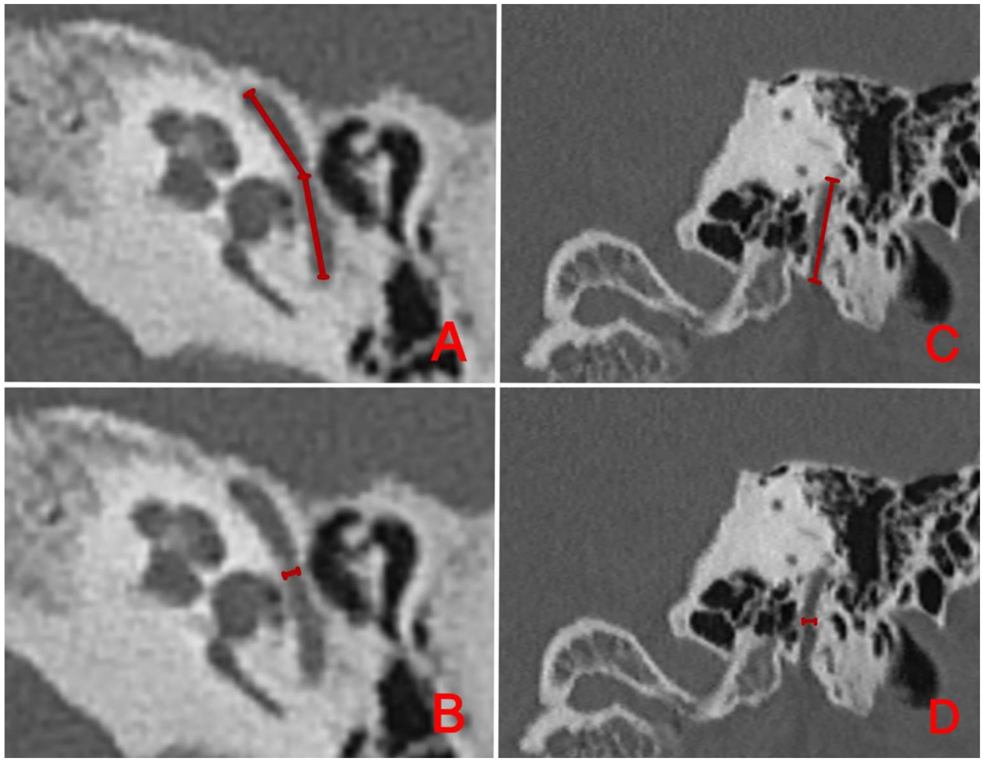


Fig. 2 Measurements related to the tympanic and mastoid segments of the facial canal. **A**: Length of the tympanic segment (TSL), **B**: Width of the tympanic segment (TSW), **C**: Length of the mastoid segment (MSL), **D**: Width of the mastoid segment (MSW)

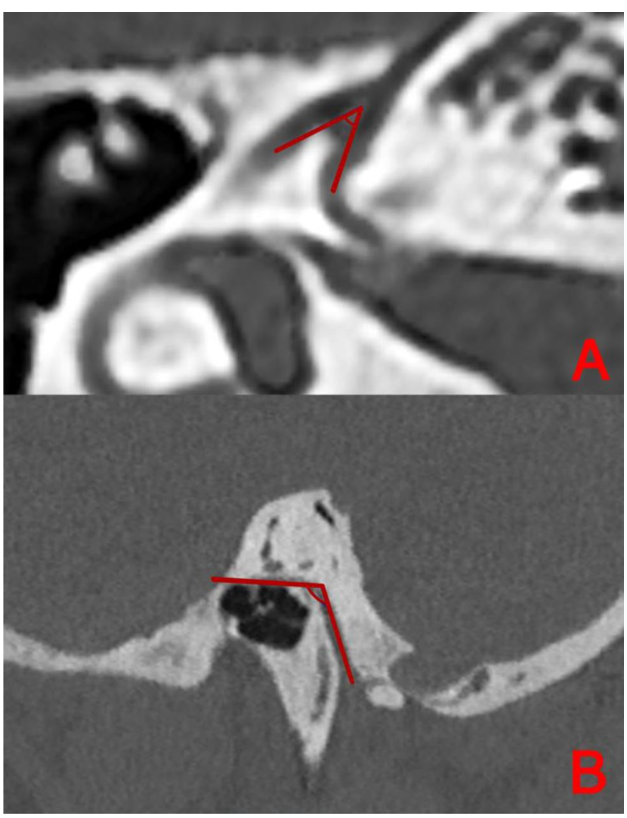


Fig. 3 The first (**A**) and second (**B**) genu angles of the facial canal

Decision Tree (DT) classifies data by reaching a result through a branching structure in the form of a tree, based on decision rules [16]. Random Forest (RF) is based on the principle of creating many decision trees and combining their predictions through majority voting. It is resistant to overfitting [17]. Extra Tree Classifier (ETC) applies more randomization by using all data in multiple decision trees [17]. Logistic Regression (LR) is a linear model that calculates the probability of an instance belonging to a particular class [18]. Gaussian Naive Bayes (GaussianNB) tries to determine the class membership of a data point by referencing the probability distributions of the variables belonging to that class [19]. K-Nearest Neighbors (k-NN) classifies new data based on the class of its k nearest neighbors. There is no training phase [19]. Multilayer Perceptron (MLCP) is a type of artificial neural network (ANN) consisting of input, hidden, and output layers. It can learn complex nonlinear relationships in the data using the backpropagation algorithm during training [15]. Deep learning (DL), due to its layered architecture and automatic feature extraction capability, requires tens of thousands (or even hundreds of thousands to millions) of samples. Traditional ML and ANNs, on the other hand, can yield satisfactory results with smaller datasets (ranging from hundreds to thousands of samples) [20]. Overall, the data requirements for the ML algorithms and ANNs used in this study are lower compared to those of DL.

In DT algorithm, max-depth was set as 5, criterion gini, min_samples_leaf was set as 2, in ETC algorithm n_estimators was set as 100, min_samples_split was set as 2, criterion entropy was set as 2, k was set as 5 in k-NN algorithm. In addition, the SHapley Additive exPlanations (SHAP) analyzer was employed to assess the efficacy of the parameters.

The architecture of the MLCP comprised 18 neurons in the input layer, 2 neurons in the output layer, and 2 hidden layers, with 5 neurons in the first layer and 10 neurons in the second layer. To ensure that the topology reflects reality, the data was retrained 100, 500, and 1,000 times. In the study, the dropout method was used to prevent over-learning of the MLCP model and to ensure adaptation. The dropout value was determined as 0.05.

The dataset was divided into two, a training set that accounts for 80% of the total data and a test set that accounts for the remaining 20%. Additionally, 10-fold cross-validation results were also included in the ML algorithms for the reliability of our study. The performance of the models was measured by various performance metrics such as accuracy (Acc), specificity (Spe), sensitivity (Sen) and F1 score.

$$Acc = \frac{TP}{TP + FN + FP + TN}$$

$$\mathrm{Sen} = \frac{TP}{TP + FN}$$

$$\mathrm{Spe} = \frac{TN}{TN + FP}$$

$$\mathrm{F1} = 2\frac{Precision \times \ Recall}{Precision + Recall}$$

Equation 1. (TP; True positive, TN; True negative, FP; Falsepositive, FN; Falsenegative).

Table 1 Descriptive statistics and group comparisons for normally distributed parameters

Parameters	Sex	Mean ± standard deviation	*p**
Right MSL	Male	12.491 ± 2.522	**0.001**
	Female	10.627 ± 1.917	
Left MSL	Male	12.739 ± 2.571	**0.001**
	Female	10.584 ± 1.866	
Right SGA	Male	143.636 ± 8.511	**0.001**
	Female	121.347 ± 5.605	
Left SGA	Male	141.509 ± 8.697	**0.001**
	Female	121.533 ± 5.504	

*Two sample t test

Basic statistical analysis

Basic statistical analyses were performed using Minitab 17 and SPSS 21 package programs. A p-value less than 0.05 was considered to be significant. In instances of sex comparison, the Two Sample T test was employed for data that conformed to a normal distribution, while the Mann-Whitney U test was utilized for data that did not adhere to a normal distribution. ROC analysis was performed to evaluate the effect of the parameters on the overall outcome. The measurements in our study were taken by an expert radiologist with at least 10 years of experience in the field. In addition, the measurements of 30 individuals for each parameter were measured by another investigator in the study to assess inter-rater reliability and the Intraclass Correlation Coefficient (ICC) test was performed. Since the measurements were performed by an expert radiologist and tested by applying the ICC test, no intra-observer assessment was deemed necessary. This does not constitute a limitation. The minimum sample size was also performed using the G*Power program (Version 3.1.9.4). The minimum sample size was found to be 188 with an alpha value of 0.05 and 95% confidence interval.

Results

The right/left MSL and the right/left SGA, which exhibited normal distribution, were evaluated in terms of sex, and all parameters were greater in males compared to females, with statistically significant differences ($p < 0.05$) (Table 1).

Among the parameters that did not show normal distribution, except for the right/left TSL, all other parameters were more common in males than females and all parameters had a significant difference in terms of sex ($p < 0.05$), (Table 2).

ROC analysis was performed to see the contribution of each parameter to the overall outcome, and it was found that the right TSL parameter had the highest contribution ($p < 0.05$), (Table 3), (Fig. 4).

In our study, 30 measurements of each parameter were measured by another investigator and ICC test was performed and all parameter measurements were found to be reliable ($p < 0.05$).

In our study, in addition to the 20% test set, 10-fold cross validation results were also included in terms of the reliability of ML algorithms. Accordingly, the highest Acc rate was found to be 0.98 ± 0.026 with the LDA algorithm (Table 5).

Confusion matrix table of the algorithms with the highest Acc values is given in Table 5. In the test set, 20 out of 21 male individuals and all 19 female individuals were correctly estimated (Table 6).

The impact of the parameters on the overall result was also evaluated using the SHAP analyzer, and it was found

Table 2 Descriptive statistics and group comparisons for non-normally distributed parameters

Parameters	Sex	Median (Minimum-Maximum)	*p***
Right IFC	Male	1.300 (0.800–2.700)	**0.001**
	Female	1.100 (0.800–1.900)	
Right LSL	Male	4.900 (2.400-8.000)	**0.001**
	Female	3.500 (2.200-7.000)	
Right LSW	Male	1.150 (0.800–2.300)	**0.001**
	Female	0.950 (0.600–2.100)	
Right TSL	Male	9.500 (2.800–15.800)	**0.009**
	Female	10.700 (6.600–13.800)	
RightTSW	Male	1.300 (0.700–3.100)	**0.001**
	Female	0.800 (0.500–1.700)	
Right FGA	Male	71.150 (36.000–90.000)	**0.001**
	Female	65.600 (48.300–88.400)	
Right MSW	Male	1.800 (0.700–3.800)	**0.004**
	Female	1.600 (0.900–2.400)	
Left IFC	Male	1.600 (0.900–3.100)	**0.001**
	Female	1.100 (0.800–2.100	
Left LSL	Male	4.900 (3.000-7.200)	**0.001**
	Female	3.400 (2.300–5.900)	
Left LSW	Male	1.200 (0.700–2.900)	**0.001**
	Female	0.900 (0.600–1.700)	
Left TSL	Male	10.050 (4.900–17.100)	**0.044**
	Female	10.500 (6.500–24.000)	
LeftTSW	Male	1.200 (0.600–2.400)	**0.001**
	Female	0.800 (0.600–2.200)	
Left FGA	Male	72.400 (25.000–90.000)	**0.001**
	Female	63.200 (49.600–87.600)	
Left MSW	Male	2.000 (0.700–5.300)	**0.001**
	Female	1.600 (1.100–2.300)	

**Mann Whitney U test

that the top 5 influential parameters were the right SGA, left SGA, right TSW, left TSW, and left IFC (Fig. 5).

According to the RF algorithm SHAP analyzer, it was found that the following parameters had minimal impact on the algorithm's accuracy: left FGA, left LSL, left LSW, left MSL, right MSL, right FGA, right IFC, and right LSL (< 0.05). After these parameters were removed, the sex evaluation was re-conducted, and it was observed that the algorithms' accuracy rate was affected by 0.02 (QDA: 0.91, other algorithms: 0.95).

As a result of the ML algorithms, the accuracy rate of all other algorithms was 0.97 except for the QDA algorithm (Acc 0.93) (Table 7).

As a result of the neural network analysis, a 0.97 sex estimation accuracy was obtained across all training sets. In the test set of all trainings, 20 out of 21 male individuals and all 19 female individuals were correctly estimated (Table 8).

Table 3 ROC curve performance results

Parameters	AUC (95%)	Cutt off	*p*	Sensitivity%	Speci-ficity%
Right IFC	0.198 (0.137–0.259)	1.150	**0.000**	31.0	20.0
Right LSL	0.167 (0.109–0.225)	4.050	**0.000**	23.0	24.0
Right LSW	0.279 (0.209–0.350)	1.050	**0.000**	34.0	39.0
Right TSL	0.607 (0.524–0.690)	10.450	**0.009**	58.0	57.0
RightTSW	0.150 (0.097–0.202)	1.050	**0.000**	21.0	23.0
Right FGA	0.366 (0.288–0.444)	68.950	**0.001**	40.0	40.0
Right MSL	0.275 (0.205–0.345)	11.450	**0.000**	35.0	35.0
Right MSW	0.384 (0.305–0.463)	1.650	**0.005**	43.0	36.0
Right SGA	0.022 (0.003–0.041)	129.650	**0.000**	0.07	0.07
Left IFC	0.128 (0.078–0.177)	1.250	**0.000**	20.0	17.0
Left LSL	0.142 (0.091–0.194)	4.150	**0.000**	22.0	20.0
Left LSW	0.219 (0.155–0.283)	1.050	**0.000**	26.0	30.0
Left TSL	0.583 (0.500-0.665)	10.350	**0.044**	55.0	57.0
LeftTSW	0.167 (0.110–0.225)	1.050	**0.000**	21.0	29.0
Left FGA	0.350 (0.272–0.428)	67.950	**0.000**	38.0	38.0
Left MSL	0.237 (0.170–0.303)	11.550	**0.000**	31.0	32.0
Left MSW	0.306 (0.231–0.381)	1.750	**0.000**	35.0	34.0
Left SGA	0.040 (0.010–0.069)	129.350	**0.000**	0.08	0.09

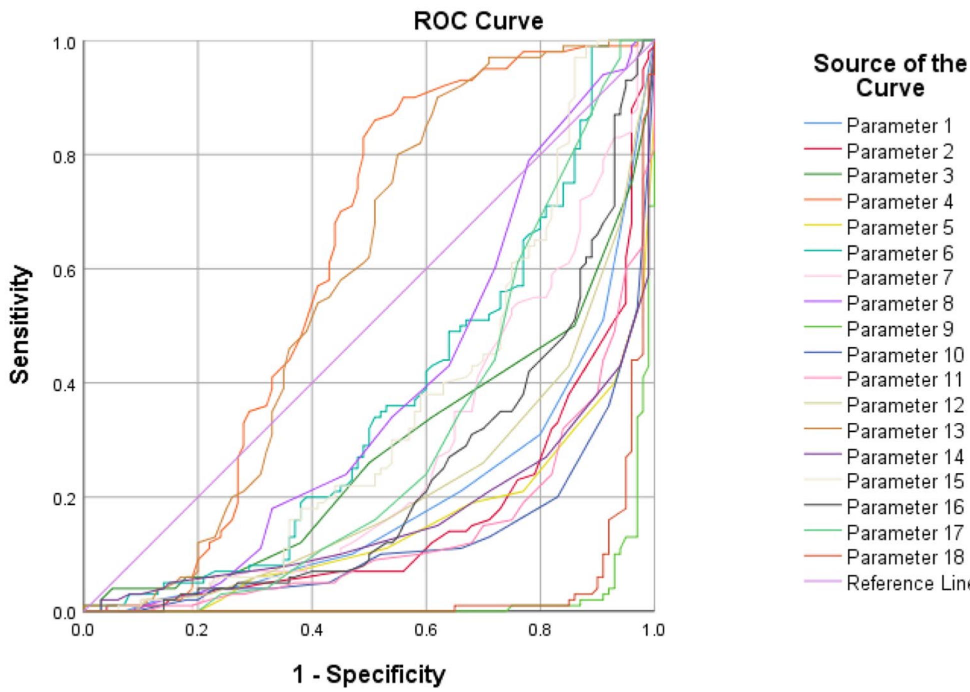


Fig. 4 ROC curve graph (Parameter 1: Right inner opening width of the facial canal (right IFC), 2: Right labyrinthine segment length (right LSL), 3: Right labyrinthine segment width (right LSW), 4: Right tympanic segment length (right TSL), 5: Right tympanic segment width (right TSW), 6: Right first genu angle (right FGA), 7: Right mastoid segment length (right MSL), 8: Right mastoid segment width (right MSW), 9: Right second genu angle (right SGA), 10: Left inner opening width of the facial canal (left IFC), 11: Left labyrinthine segment length (left LSL), 12: Left labyrinthine segment width (left LSW), 13: Left tympanic segment length (left TSL), 14: Left tympanic segment width (left TSW), 15: Left first genu angle (left FGA), 16: Left mastoid segment length (left MSL), 17: Left mastoid segment width (left MSW), 18: Left second genu angle (left SGA))

Table 4 Intraclass correlation coefficient test results

Parameters	Intraclass Correlation	*p*
Right IFC	0.827	**0.001**
Right LSL	0.883	**0.001**
Right LSW	0.811	**0.001**
Right TSL	0.879	**0.001**
RightTSW	0.801	**0.003**
Right FGA	0.833	**0.001**
Right MSL	0.823	**0.001**
Right MSW	0.812	**0.001**
Right SGA	0.888	**0.001**
Left IFC	0.805	**0.003**
Left LSL	0.886	**0.001**
Left LSW	0.875	**0.001**
Left TSL	0.801	**0.003**
LeftTSW	0.866	**0.001**
Left FGA	0.846	**0.001**
Left MSL	0.812	**0.001**
Left MSW	0.844	**0.001**
Left SGA	0.826	**0.001**

Discussion

In this study, which evaluated the morphometric properties of the FC in the temporal bone, it was found that sex could be estimated with 97% accuracy using ML algorithms and ANN from FC measurements It was found that all of the evaluated parameters had significant differences between the sexes; in ML algorithms, the most effective parameter for sex estimation was the right SGA. Furthermore, both right and left SGAs and TSWs were found to exhibit more dimorphic characteristics compared to other parameters. In the algorithms with the highest accuracy, the sex of all female individuals in the test set was estimated correctly, while the sex of 20 out of 21 male individuals was estimated correctly.

Sex estimation based on the morphometric features of bone structures is widely practiced. In forensic medicine, it may be crucial to estimate sex from structures located in deeper positions due to the lower likelihood of damage [21]. Studies investigating the effectiveness and accuracy of sex estimation from structures in deeper regions have been conducted to contribute to forensic medicine and forensic anthropology [22–25]. However, in the available literature, no study has evaluated the accuracy and effectiveness of morphometric analyses of the FC in sex estimation.

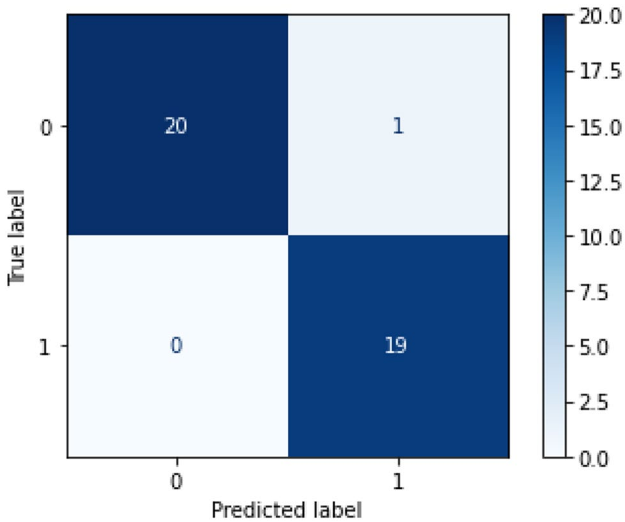
The use of machine learning algorithms in health domains, such as disease prediction and sex estimation, is becoming increasingly widespread. The achievement of fast, reliable, and accurate results, as well as the usability of digitized health data, has led to an increase in studies evaluating the application areas and effectiveness of

Table 5 10-fold cross validation performance of machine learning algorithms (Training fold results)

Training folds	QDA	LDA	DT	ETC	RF	LR	GaussianNB	k-NN
1	0.90	1.00	0.90	0.95	1.00	0.89	0.95	0.90
2	1.00	1.00	0.95	0.95	1.00	0.89	1.00	0.95
3	0.90	0.95	1.00	1.00	0.88	0.89	0.90	0.85
4	0.90	0.95	0.95	0.95	0.94	0.96	0.85	0.90
5	0.90	0.95	0.95	0.95	1.00	0.81	0.90	0.90
6	0.95	1.00	1.00	1.00	1.00	0.92	0.95	1.00
7	0.95	0.95	0.95	1.00	0.94	0.92	0.95	1.00
8	0.90	1.00	0.90	0.95	0.88	0.92	0.90	0.95
9	0.95	1.00	1.00	0.95	1.00	0.85	1.00	1.00
10	0.90	1.00	0.70	0.95	1.00	0.81	0.95	1.00
Mean	0.92	0.98	0.93	0.97	0.96	0.89	0.94	0.95
Std	0.035	0.026	0.089	0.024	0.051	0.049	0.047	0.055

Std: Standard deviation

Table 6 Confusion matrix table for the highest accuracies



ML algorithms [26]. In this study, the accuracy rates of sex estimation from measurements obtained from the segments and angles of the FC were evaluated using ML algorithms. Due to both the use of ML algorithms and the sex estimation from the FC, no similar study has been found for this unique research. There are studies in the literature where sex estimation is performed using data from different bone structures with ML algorithms. The studies show that different accuracy rates are obtained depending on the bone evaluated, the morphometric feature and the algorithm used. In a study where sex estimation was made using morphometric measurements obtained from the mandible with machine learning (ML) algorithms, the lowest accuracy rate was achieved with linear regression (LR) (69%), while the highest accuracy rates were obtained with k-nearest neighbor (k-NN) and random forest (RF) (76%) algorithms [27]. In another study, where sex estimation was made with ML algorithms using measurements obtained from the foramen incisivum, the lowest accuracy rates were attained with QDA and k-NN (76%) algorithms, while the highest accuracy rates were achieved with LR, LDA, and RF (82%) algorithms [23]. A further study utilized machine learning algorithms with measurements obtained from the first and fifth metatarsals and phalanges to estimate sex, with the DT algorithm demonstrating the highest accuracy rate of 85%. The ETC algorithm yielded the lowest accuracy rate of 74% [28]. In the present study, the lowest accuracy value of 93% was achieved with QDA, while an accuracy of 97% was obtained with all other algorithms. It was found that the morphometric properties of the FC exhibit sexual dimorphic characteristics, and high accuracy and reliability were obtained with all algorithms.

In sex estimation studies conducted in the craniofacial region, Parlak et al. reported an accuracy rate of 84.2% using ML algorithms based on the morphometric characteristics of the piriform aperture [29]. In another study that estimated sex based on measurements of the mastoid triangle and the volume of the mastoid air cells, an accuracy rate of 88.5% was achieved using ML algorithms

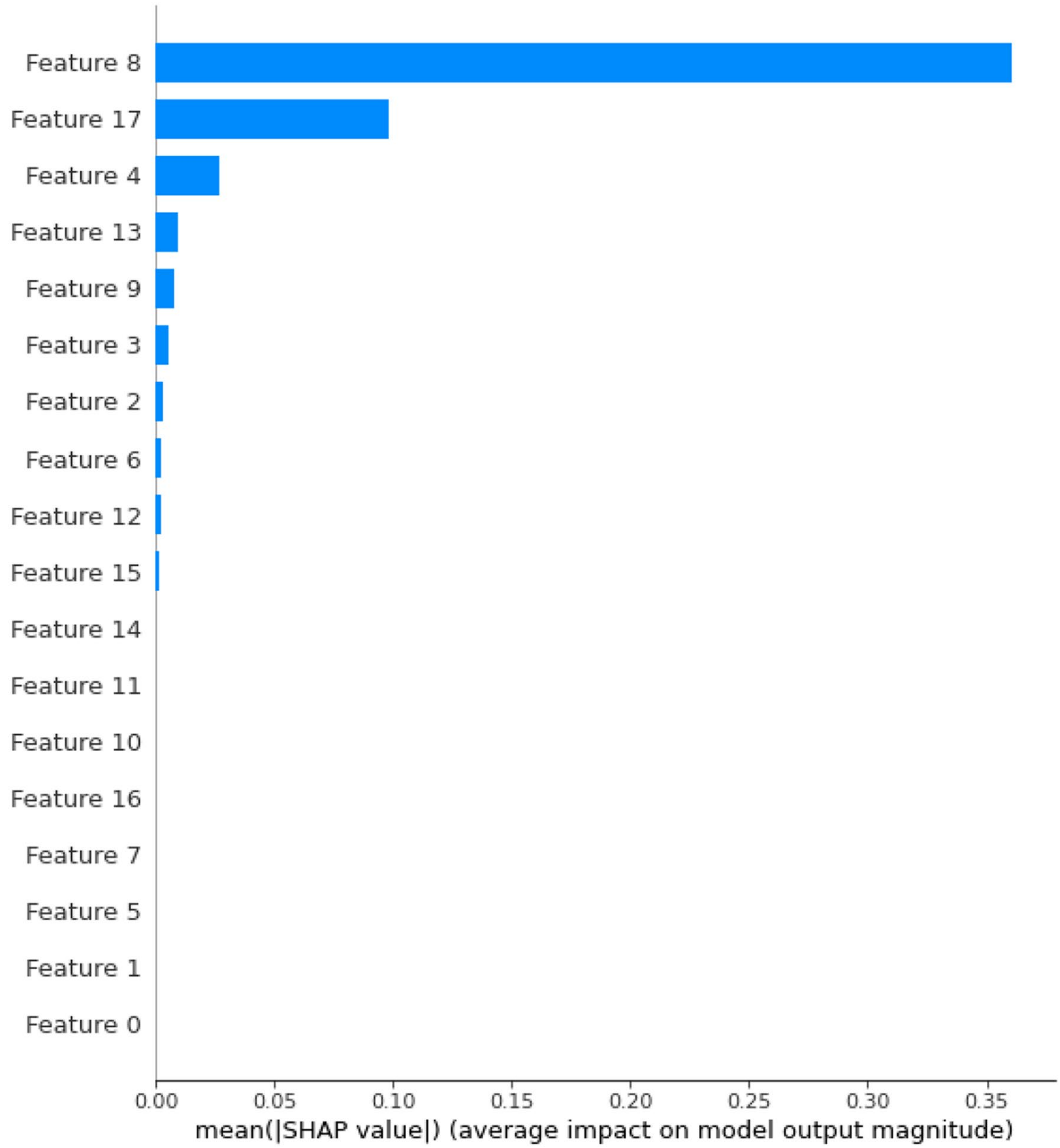


Fig. 5 SHapley Additive exPlanations (Feature 0: Right inner opening width of the facial canal (right IFC), 1: Right labyrinthine segment length (right LSL), 2: Right labyrinthine segment width (right LSW), 3: Right tympanic segment length (right TSL), 4: Right tympanic segment width (right TSW), 5: Right first genu angle (right FGA), 6: Right mastoid segment length (right MSL), 7: Right mastoid segment width (right MSW), 8: Right second genu angle (right SGA), 9: Left inner opening width of the facial canal (left IFC), 10: Left labyrinthine segment length (left LSL), 11: Left labyrinthine segment width (left LSW), 12: Left tympanic segment length (left TSL), 13: Left tympanic segment width (left TSW), 14: Left first genu angle (left FGA), 15: Left mastoid segment length (left MSL), 16: Left mastoid segment width (left MSW), 17: Left second genu angle (left SGA))

[30]. In a study utilizing various morphometric variables such as orbital width, height, intraorbital, and extraorbital distances, an overall accuracy of approximately 68% was reported [31]. When these sex estimation studies in the craniofacial region using ML algorithms are considered, it is seen that lower accuracy rates were obtained compared to our study. In this context, morphometric data obtained from the FC may exhibit stronger anatomical sexual dimorphism.

In the sex estimation study by Baban and Mohammed, the results of the ROC analysis indicated that the average mandibular canine width exhibited the most significant effect on sex estimation. However, the SHAP analysis revealed that the average mandibular canine width emerged as the second most significant parameter. According to the SHAP analysis, the mandibular right canine height had the most significant effect [32]. The results of our study, employing traditional ROC

Table 7 Machine learning algorithms analysis results (Test set results)

Algorithms	Acc	Spe	Sen	F1
QDA	0.93	0.93	0.92	0.92
LDA	0.97	0.97	0.98	0.97
DT	0.97	0.97	0.98	0.97
ETC	0.97	0.97	0.98	0.97
RF	0.97	0.97	0.98	0.97
LR	0.97	0.97	0.98	0.97
GaussianNB	0.97	0.97	0.98	0.97
k-NN (k=5)	0.97	0.97	0.98	0.97

QDA: Quadratic Discriminant Analysis; LDA: Linear Discriminant Analysis; DT: Decision Tree; ETC: Extra Tree Classifier; RF: Random Forest; LR: Logistic Regression; GaussianNB: Gaussian Naive Bayes; k-NN: k-Nearest Neighbors

Table 8 Performance results of artificial neural networks

Number of Trainings	Acc	Spe	Sen	F1
100	0.97	0.97	0.98	0.97
500	0.97	0.97	0.98	0.97
1000	0.97	0.97	0.98	0.97

analysis, indicated that the most effective parameter for sex estimation was the right TSL. However, the SHAP analysis identified the right SGA, left SGA, right TSW, left TSW, and the internal opening width of the left FC as the five most effective parameters. It is hypothesized that this discrepancy stems from the utilization of the entire dataset as a training set in ROC analysis, in contrast to the 80%/20% training/test set configuration employed in machine learning algorithms, which may indicate a higher reliability of the SHAP-based findings due to the separation of training and testing phases inherent to machine learning validation. Further studies are needed to test these differences and investigate their causes.

Celik et al. reported an average diameter of 1.22 mm for the labyrinthine segment of the FC, 1.39 mm for the tympanic segment, and 1.61 mm for the mastoid segment [33]. In a separate study by Yurdabakan et al., the mastoid segment diameter was measured as an average of 1.57 mm in females and 1.68 mm in males. According to the findings of Yurdabakan et al., a statistically significant variation in diameter was observed according to sex. Specifically, males presented with higher measurements [34]. Shin et al. evaluated the diameters of FC segments by measuring them from different points. The values obtained from the measurement of the labyrinthine segment along the longest axis were found to be 1.5 mm on average, with the measurement taken close to the start and end points of the segment. The width measurement corresponding to the center of the tympanic segment was 1.6 mm, and the diameter measurement corresponding to the center of the mastoid segment was 2.3 mm [35]. In another study that employed CT images, the labyrinthine segment diameter in axial images was determined to be 1.14 mm in females and 1.17 mm in males on the left side, and 1.13 mm in females and 1.14 mm in males on the right side. The TSL was measured at 1.11 mm in females and 1.14 mm in males on the left side, while it was recorded as 1.06 mm for both sexes on the right side. The diameter of the mastoid segment, as determined from coronal images, was recorded as 1.49 mm in female subjects and 1.58 mm in male subjects on the left side. On the right side, the diameter was 1.51 mm in females and 1.56 mm in males. In the aforementioned study, no statistically significant sex disparities were observed in the diameters of the tympanic and labyrinthine segments measured from axial images. However, significant sex differences were identified in the coronal images. While the mastoid segment diameter exhibited a significant variation between sexes in axial sections, a significant difference between sexes was only observed on the left side in coronal sections [5]. In our study, a statistically significant difference was found between the labyrinthine segment and tympanic segment diameters measured on axial images, as well as the mastoid segment diameters measured on coronal images, between females and males. The measured values are consistent with literature.

In their study, Shin et al. found the LSL to be an average of 4.5 millimeters in both females and males. The TSL was measured at 12.2 millimeters in females and 12.7 millimeters in males. The MSL was found to be 14 mm in females and 15 mm in males [35]. In a cadaver study, the average LSL was found to be 4.27 mm; TSL was 10.4 mm; and MSL was 12.34 mm [36]. In a separate study, Li et al. utilized statistical analysis to ascertain the MSL to be 14 mm in females and 14.92 mm in males, thereby identifying a statistically significant difference between the two groups [37]. The findings of the present study are consistent with the extant literature on the subject.

In their study, Ishita et al. measured the FGA as 71.7 degrees and the SGA as 103.5 degrees on 30 cadavers [36]. In a subsequent study, the SGA was found to be 104.66 degrees on 30 adult cadavers [38]. Shin et al. measured the FGA at 60 degrees and the SGA at 112 degrees [35]. A further study by Li et al. revealed a statistically significant disparity in the SGA between sexes, with the angle measuring 116.14 degrees in female subjects and 114.85 degrees in male subjects [37]. Concurrently, Zhao et al. documented an average FGA of 70.12 degrees in a sample of 100 healthy individuals [39]. In our study, the FGA showed similarity with the literature, while the SGA was found to have higher angle values, particularly in males.

In the present study, a comparative analysis of previously reported facial canal (FC) morphometric measurements was conducted and summarized in Table 9. This table emphasizes how the reported values vary according to several factors, including the country of origin, sample size, and the measurement technique employed (e.g.,

Table 9 Comparison of studies based on sample size, geographic origin, and measurement technique

Article	Number	Country	Method	IFC	LSL	LSW	TSL	TSW	MSL	MSW	FGA	SGA
[33]	34	Turkey	CT			1.23 ± 0.22		1.39 ± 0.28		1.61 ± 0.31		
[38]	30 cadaver	India	Measuring probe				10.44 ± 1.43		13.33 ± 2.2			104.66 ± 7.76
[5]	169 (85 Male; 84 Female)	Turkey	CT			Male: (right:1.14 ± 1.23; left:1.17 ± 0.21) Female: (right:1.13 ± 0.27; left:1.14 ± 0.27)		Male: (right:1.06 ± 0.25; left:1.14 ± 0.24) Female: (right:1.06 ± 0.25; left:1.11 ± 0.25)		Male: (r:1.89 ± 0.26; l:1.92 ± 0.28) Female: (r:1.79 ± 0.23; l:1.82 ± 0.22)		
[36]	30 cadaver	India	Caliper				10.40 ± 1.42		12.34 ± 0.92		71.7 ± 9.53	103.50 ± 8.32
[37]	234 (118 Male; 116 Female)	China	CT									Male: 114.85 ± 6.30 Female: 116.14 ± 6.51
[41]	35 cadaver	Romania	Photograph-based software measurement				10.25 ± 0.75		13.78 ± 1.12		48 to 86	92 to 125
[40]	34	China	CT		2.88 ± 0.45		12.50 ± 1.04		16.51 ± 1.44		80.43 ± 6.43	114.15 ± 8.51
[34]	479	Turkey	CBCT							1.62 ± 0.35		
[35]	32 cadaver	Korea	CT	1.5 ± 0.2	4.5 ± 0.5	1.5 ± 0.2	12.5 ± 0.9	1.6 ± 0.2	14.7 ± 1.6	2.3 ± 0.3	60.2 ± 14.7	112.0 ± 10.2

direct dissection vs. radiological imaging with software-based measurement). In our CT-based study, the second genu angle was found to be higher than those reported in earlier literature. Moreover, the TSL and MSL parameters were measured at lower values compared to data from Chinese and Korean populations [35, 40]. Despite methodological and demographic differences, the remaining morphometric parameters generally showed concordance with values previously reported in the literature.

Given that an increased sample size in machine learning algorithms will result in a greater quantity of training and test data, which in turn provides enhanced data diversity, it stands to reason that studies employing larger sample sizes may yield divergent results. In this context, it is our hope that our study will serve as a foundation for future research by increasing the number of samples, a deficiency that our study addresses. Although this study achieved high classification accuracy within a well-defined population, the use of data from a single center and the relatively broad age range (19–65 years) without stratified analysis may limit the generalizability of the findings to other populations or age groups. Anatomical variations in FC morphology may exist across different racial, ethnic, and age groups due to genetic and environmental factors. Previous craniofacial morphometric studies have demonstrated population-specific patterns that could influence the predictive performance [42, 43]. Future studies incorporating multicenter data and stratified age analyses are recommended to enhance the robustness and applicability of the models developed in this study, and to validate and refine their performance across diverse demographic groups.

Conclusion

In this study, we found that all morphometric features of the FC exhibit sexual dimorphism, and sex estimation using ML algorithms achieved a high accuracy rate of 97%. Among the measured parameters, the genu angles and transverse segment widths on both sides were identified as the most sexually dimorphic features. These morphometric data, which showed statistically significant differences between males and females, may play a valuable role in clinical settings—particularly in otologic and skull base surgeries—by helping surgeons anticipate anatomical variations and reduce the risk of complications. They may also serve as auxiliary diagnostic data for more precise preoperative planning. In forensic practice, the application of ML-based sex estimation methods using FC measurements may facilitate rapid and reliable identification, especially in cases involving severely fragmented or decomposed remains where traditional methods are limited. These findings therefore hold promise for supporting both surgical planning and postmortem identification workflows.

Acknowledgements
We thank all scientists who have contributed to science.

Author contributions
Concept – Y.S., D.S., S.K., O.H., O. Ö.; Design – Y. S., D.S., S.K., O.H., O.Ö.; Supervision – Y.S., D.Ş., Ö.Ö., N.Y.; Resources – Y.S., D.Ş., Ö.Ö., N.Y.; Materials – Y.S., D.Ş., Ö.Ö., N.Y.; Data Collection and/or Processing – Y.S., S.K., O.H., O.Ö.; Analysis and/or Interpretation – Y.S.; Literature Search – S.K., O.H., O. Ö.; Writing – Y.S., D.S., S.K., O.H., O. Ö.; Critical Review – Y.S., D.Ş., Ö.Ö., N.Y.

Funding
The authors declare that this study received no financial support.

Data availability
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics declarations and consent to participate
Before the study started, ethical approval was obtained from Düzce University Non-interventional Health Research Ethics Committee (DU-HR-IRB) with protocol number 2024/237. Informed consent was waived due to lack of direct patient contact during data collection. The study was also conducted in accordance with the relevant guidelines and regulations as per the Declaration of Helsinki.

Consent for publication
Not applicable.

Main points
1. Examination of fascial canal morphologies obtained by computed tomography. 2. Determination of the variation of the fascial canal according to gender. 3. To reveal the variation of the fascial canal according to gender with machine learning algorithms and artificial neural networks, which is a current methodology. 4. To explain the clinical importance of fascial canal morphology.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Anatomy, Faculty of Medicine, Karabük University, Karabük, Türkiye
[2]Department of Anatomy, Faculty of Medicine, Düzce University, Düzce, Türkiye
[3]Department of Therapy and Rehabilitation, Çatalzeytin Vocational School, Kastamonu University, Kastamonu, Türkiye
[4]Department of Radiology, Faculty of Medicine, Düzce University, Düzce, Türkiye
[5]Department of Otorhinolaryngology, Faculty of Medicine, Karabük University, Karabük, Türkiye

Received: 12 May 2025 / Accepted: 16 July 2025
Published online: 18 July 2025

References

1. Colman KL, van der Merwe AE, Stull KE, Dobbe JGG, Streekstra GJ, van Rijn RR, et al. The accuracy of 3D virtual bone models of the pelvis for morphological sex Estimation. Int J Legal Med. 2019;133:1853–60.
2. Decker SJ, Foley R, Hazelton JM, Ford JM. 3D analysis of computed tomography (CT)-derived lumbar spine models for the Estimation of sex. Int J Legal Med. 2019;133:1497–506.
3. Saukko P, Knight B. Knight's forensic pathology. CRC; 2015.
4. Scheuer L. Application of osteology to forensic medicine. Clin Anat. 2002;15:297–312.
5. Inal M, Bayar Muluk N, Asal N, Şahan MH, Şimşek G, Arikan OK. Is there a relationship between mastoid pneumatisation and facial Canal dimensions? J Laryngol Otol. 2019;133:546–53.
6. Beger O, Erdoğan O, Kara E, Vayisoğlu Y, Görür K, İsmi O, et al. Morphometric properties of the facial Canal in children: A retrospective computed tomography study. Int J Pediatr Otorhinolaryngol. 2019;124:59–67.
7. George T, Rufus E, Alex Z. Simulation of microwave induced thermo-acoustical imaging technique for cancer detection. 2015;10:9424–8.
8. Islam MMF, Ferdousi R, Rahman S, Bushra H. Likelihood prediction of diabetes at early stage using data mining techniques. Computer vision and machine intelligence in medical image analysis. Singapore: Springer. 2020:113–25.
9. Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF diabetes atlas: global estimates for the prevalence of diabetes for 2015 and 2040. Diabetes Res Clin Pract. 2017;128:40–50.
10. Karakuş C. Makine Öğrenmesi Temelleri Ders Notu Çalışma Notları. 2021. https://ckk.com.tr/ders/ML.html
11. Bidmos MA, Olateju OI, Latiff S, Rahman T, Chowdhury MEH. Machine learning and discriminant function analysis in the formulation of generic models for sex prediction using patella measurements. Int J Legal Med. 2023;137:471–85.
12. Kartal E, Etli Y, Asirdizer M, Hekimoglu Y, Keskin S, Demir U, et al. Sex Estimation using foramen magnum measurements, discriminant analyses and artificial neural networks on an Eastern Turkish population sample. Leg Med. 2022;59:102143.
13. Ortega RF, Irurita J, Campo EJE, Mesejo P. Analysis of the performance of machine learning and deep learning methods for sex Estimation of infant individuals from the analysis of 2D images of the ilium. Int J Legal Med. 2021;135:2659–66.
14. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health. 2019;1:e271–97.
15. Sarker IH. Machine learning: algorithms, Real-World applications and research directions. SN Comput Sci. 2021;2:160.
16. Sun J, Zhao H. The application of sparse Estimation of covariance matrix to quadratic discriminant analysis. BMC Bioinformatics. 2015;16:48.
17. Mavrogiorgou A, Kiourtis A, Kleftakis S, Mavrogiorgos K, Zafeiropoulos N, Kyriazis D. A catalogue of machine learning algorithms for healthcare risk predictions. Sensors. 2022;22:8615.
18. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Learn. 2006;63:3–42.
19. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inf Decis Mak. 2019;19:281.
20. Cheng Y, Petrides KV, Li J. Estimating the minimum sample size for neural network model Fitting—A Monte Carlo simulation study. Behav Sci. 2025;15:211.
21. Christensen AM, Passalacqua NV, Bartelink EJ. Forensic Anthropology: Current Methods and Practice. 2nd edition. Oxford; San Diego, California: Academic Press; 2019.
22. Murali S, Bagewadi A, Lagali-Jirge V, Panwar SLK, Keluskar A. Reliability of gender determination from paranasal sinuses and its application in forensic identification-a systematic review and meta-analysis. Forensic Sci Med Pathol. 2023;19:409–39.
23. Senol D, Secgin Y, Harmandaoglu O, Kaya S, Duman SB, Oner Z. Gender prediction using Cone-Beam computed tomography measurements from foramen incisivum: application of machine learning algorithms and artificial neural networks. J Anat Soc India. 2024;73:152.
24. Şimşek HA, Gültekin T, Bilecenoğlu B, Kolsuz ME, Atakan C. Konik Işınlı Bilgisayarlı tomografi Ile frontal sinüsten Morfolojik cinsiyet Tayini. Turkiye Klinikleri J Foren Sci Leg Med. 2020;17:117–21.
25. Şimşek HA, Gültekin T, Bilecenoğlu B, Kolsuz ME, Atakan C. Konik Işınlı Bilgisayarlı tomografi Ile Maksiller sinüsten cinsiyet Tayini. Turkiye Klinikleri J Dent Sci. 2021;27:364–71.
26. Chen P-HC, Liu Y, Peng L. How to develop machine learning models for healthcare. Nat Mater. 2019;18:410–4.
27. Basso IB, de Jesus Freitas PF, Ferraz AX, Borkovski AJ, Borkovski AL, Santos RS, et al. Sex prediction through machine learning utilizing mandibular condyles, coronoid processes, and sigmoid notches features. PLoS ONE. 2024;19:e0312824.

28. Senol D, Bodur F, Secgın Y, Bakıcı RS, Sahin NE, Toy S, et al. Sex prediction with morphometric measurements of first and fifth metatarsal and phalanx obtained from X-ray images by using machine learning algorithms. Folia Morphol. 2023;82:704–11.
29. Parlak ME, Etli Y, Beyhan M, Kanat K, Kızıloğlu HA. Sex Estimation with ensemble learning: an analysis using anthropometric measurements of piriform aperture. Egypt J Forensic Sci. 2025;15:10.
30. Sasani H, Etli Y, Tastekin B, Hekimoglu Y, Keskin S, Asirdizer M. Sex Estimation from measurements of the mastoid triangle and volume of the mastoid air cell system using classical and machine learning methods: A comparative analysis. Am J Forensic Med Pathol. 2024;45:51–62.
31. Triantafyllou G, Botis GG, Piagkou M, Papanastasiou K, Tsakotos G, Paschopoulos I, et al. Sex Estimation through orbital measurements: A machine learning approach for forensic science. Diagnostics (Basel). 2024;14:2773.
32. Baban MTA, Mohammad DN. A new approach for sex prediction by evaluating mandibular arch and canine dimensions with machine-learning classifiers and intraoral scanners (a retrospective study). Sci Rep. 2024;14:27974.
33. Celik O, Eskiizmir G, Pabuscu Y, Ulkumen B, Toker GT. The role of facial Canal diameter in the pathogenesis and grade of bell's palsy: a study by high resolution computed tomography. Braz J Otorhinolaryngol. 2016;83:261–8.
34. Yurdabakan ZZ, Okumuş Ö, Orhan K. Evaluation of the morphometry of the stylomastoid foramen in the Turkish population: A retrospective study using Cone-Beam computed tomography (CBCT). J Neurol Surg B Skull Base. 2023;84:521–9.
35. Shin K-J, Gil Y-C, Lee J-Y, Kim J-N, Song W-C, Koh K-S. Three-Dimensional study of the facial Canal using microcomputed tomography for improved anatomical comprehension. Anat Rec. 2014;297:1808–16.
36. Ishita BS, Singla RK, Sharma DK, Yadav V. Intratemporal facial nerve anatomy and its variations in 30 cases of cadaveric Temporal bones. Indian J Otolaryngol Head Neck Surg. 2022;74(Suppl 3):4183–8.
37. Li T, Lai ZC, Wang XD, Feng Y, Li YQ, Fang XD. Measurement and analysis of facial nerve on fully displayed multislice computed tomographic multiplanar reconstruction image. J Craniofac Surg. 2013;24:1411–3.
38. Gupta N, Gupta R, Singh IP, Kotwal S, Suri A, Raina S. To study the anatomy of tympanomastoid segment of facial nerve and its variations in human cadaveric Temporal bone. Indian J Otology. 2017;23:108.
39. Zhao S, Han D, Wang Z, Li J, Qian Y, Ren Y, et al. An imaging study of the facial nerve Canal in congenital aural Atresia. Ear Nose Throat J. 2015;94:E6–13.
40. Ni Y, Sha Y, Dai P, Li H. Quantitative morphology of facial nerve based on three-dimensional reconstruction of Temporal bone. Otolaryngol Head Neck Surg. 2008;138:23–9.
41. Măru N, Cheiţă AC, Mogoantă CA, Prejoianu B. Intratemporal course of the facial nerve: morphological, topographic and morphometric features. Rom J Morphol Embryol. 2010;51:243–8.
42. Franklin D, Cardini A, Flavel A, Kuliukas A. Estimation of sex from cranial measurements in a Western Australian population. Forensic Sci Int. 2013;229:e1581–8.
43. Belaldavar C, Kotrashetti VS, Hallikerimath SR, Kale AD. Assessment of frontal sinus dimensions to determine sexual dimorphism among Indian adults. J Forensic Dent Sci. 2014;6:25–30.

Publisher's note