Available online at www.medicinescience.org

ORIGINAL ARTICLE

Medicine Science 2021;10(2):356-61

Medicine Science International Medical Journal

Gender prediction with parameters obtained from pelvis computed tomography images and decision tree algorithm

Yusuf Secgin[1], Zulal Oner[1], Muhammed Kamil Turan[2], Serkan Oner[3]

[1]*Karabuk University, Faculty of Medicine, Department of Anatomy, Karabuk, Turkey*
[2]*Karabuk University, Faculty of Medicine, Department of Medical Biology, Karabuk, Turkey*
[3]*Karabuk University, Faculty of Medicine, Department of Radiology, Karabuk, Turkey*

Received 05 November 2020; Accepted 29 December 2020
Available online 25.03.2021 with doi: 10.5455/medscience.2020.11.235

Abstract

Gender prediction is among the most critical topics in forensic medicine and anthropology since it is the basis of identity (height, weight, ancestry, age). Today, osteometry which is a low-cost, easily accessible method that requires no expertise is preferred when compared to DNA technology, which has several disadvantages such as high cost, accessibility, laboratory facilities, and expert personnel requirements. The Computed Tomography (CT) method, which is little affected by orientation and provides reconstruction opportunities, was selected instead of traditional methods for osteometry. This study aims to predict high and accurate gender with the Decision Tree (DT) algorithms used in the field of health recently. In the present study, CT images of 300 individuals (150 females, 150 males) without a pathology on the pelvic skeleton and between the ages of 25 and 50 were transformed into orthogonal form, landmarks were placed on promontorium, sacroiliac joint, iliac crest, terminal line, anterior superior iliac spine, anterior inferior iliac spine, greater trochanter, obturator foramen, lesser trochanter, femoral head, femoral neck, the body of femur, ischial tuberosity, acetabulum, and pubic symphysis, and the coordinates of these landmarks were determined. Then, parameters such as angle and length were obtained with various combinations. These parameters were analyzed with the DT algorithm.The analysis conducted with the DT algorithm revealed that accuracy (Acc) was 0.93, sensitivity was 0.95, specificity was 0.90, and the Matthews correlation coefficient was 0.86 for the pelvic skeleton. It was observed that the accuracy was quite high and more realistic when determined with the DT algorithm. In conclusion, the DT algorithm with multiple parameters and samples on pelvic CT images could improve the Acc of gender prediction.

Keywords: Computed Tomography, Decision Tree, gender prediction, osteometry, pelvis

Introduction

In high mortality cases where traumatic injuries are common and skeletal pieces are missing such as wars, accidents, terrorist attacks, and natural disasters (flood, earthquake, tsunami, landslide), the missing individuals should be identified quickly and accurately [1-3]. The first stage of identification entails gender prediction. Gender prediction is one of the important topics in forensic medicine and anthropology since it allows fast and accurate determination by reducing identity possibilities by almost half [4]. Since other biological factors (height, age, weight) depend on gender prediction, it greatly reduces identity possibilities [5].

The skeletal system may differ based on gender. The most useful skeletal system parts to reveal this difference are the pelvis and cranium [6]. However, the pelvis skeleton is considered the most dimorphic region due to the impact of natal morphology and sex hormones (estrogen, androgen) on the skeletal system. For natural birth, the bone structure should be at a certain width and the prenatal bone structure should meet certain biomechanical requirements in women. In men, the pelvis should carry a large body mass and bone volume. These properties made the pelvis a bone preferred by forensic experts and anthropologists [7-9].

Bone measurements conducted with instruments such as calipers, goniometers, electronic distance meters, etc. could lead to measurement difficulties due to the deformation of the bones after a certain period, while the Computed Tomography (CT) method offers both accurate and easy measurement with its ability to reform and image in 3 planes (coronal, sagittal, transversal) [1]. Furthermore, CT easily allows the superimposition of the 3-planar image, to examine a large number of images, to determine the age, and to identify the pathologies in the bone structure [10].

***Coresponding Author:** Zulal Oner, Karabuk University, Faculty of Medicine, Department of Anatomy, Karabuk, Turkey
E-mail: zulaloner@karabuk.edu.tr

A Decision Tree (DT) is a simple and strong Machine Learning (ML) algorithm that continuously divides and classifies the inputs based on adequate criteria. The DT algorithm is a controlled algorithm that allows us to obtain fast and clear results due to its features such as data prediction and easy interpretation without pre-processing [11].

The present study aimed to predict gender with DT algorithm analysis instead of descriptive statistics of the data obtained by adding new length and angle parameters to the previously determined standard pelvis length and angle parameters on CT images.

Materials and Methods

Population Sample and Image Scanning Protocol

The present study was conducted after the approval of the ethics committee (Act no: 6/23 on 07/10/2019). The present study was conducted on pelvic CT images of 150 female and 150 male patients (n = 300), who were between the ages of 25 and 50 (mean female age = 40, mean male age = 37) and admitted to the hospital with various health problems and without any pelvic skeleton pathology between 2018 and 2019.

The CT images were obtained with a Toshiba Medical Systems (Otawara, Japan) Aquilion 16 model 16-row Multi-Detector CT device in the supine position. Images were scanned with 5 mm slide thickness, 0.75-sec gantry rotation, 1.0 mm pitch, and 120 kV tube voltage.

Image Analysis Protocol

The pelvic CT images were recorded on Horos Medical Image Viewer 3.0 (United States) software in Digital Imaging and Communications in Medicine (DICOM) format. 3D Curved Multiplanar Reconstruction (MPR) was conducted on the recorded images to obtain 3-planar images (coronal, sagittal, transversal). To minimize the image bias, all images were transformed into orthogonal form based on the clearest promontorium and superposed (Figure 1).

Landmarks were placed on 35 anatomical points of the orthogonal images (promontorium, terminal line, superior obturator foramen (for.), inferior obturator for., greater trochanter, lesser trochanter, femoral head, femoral neck, the body of femur, ischial tuberosity, acetabulum superior, acetabulum inferior, superior pubic symphysis, inferior pubic symphysis, anterior superior iliac spine, anterior inferior iliac spine, iliac crest, sacroiliac joint) and the coordinates of these landmarks were obtained (Figure 2). This was transferred to a graphic-based program developed by the faculty members of the Department of Medical Biology, Anatomy and Radiology of Karabük University Medical Faculty, Sekazu (Version 3.0), which can transform the obtained coordinates into length and angle units in various variations and perform ML algorithm analysis according to these transformed units (Tables 1, 2).

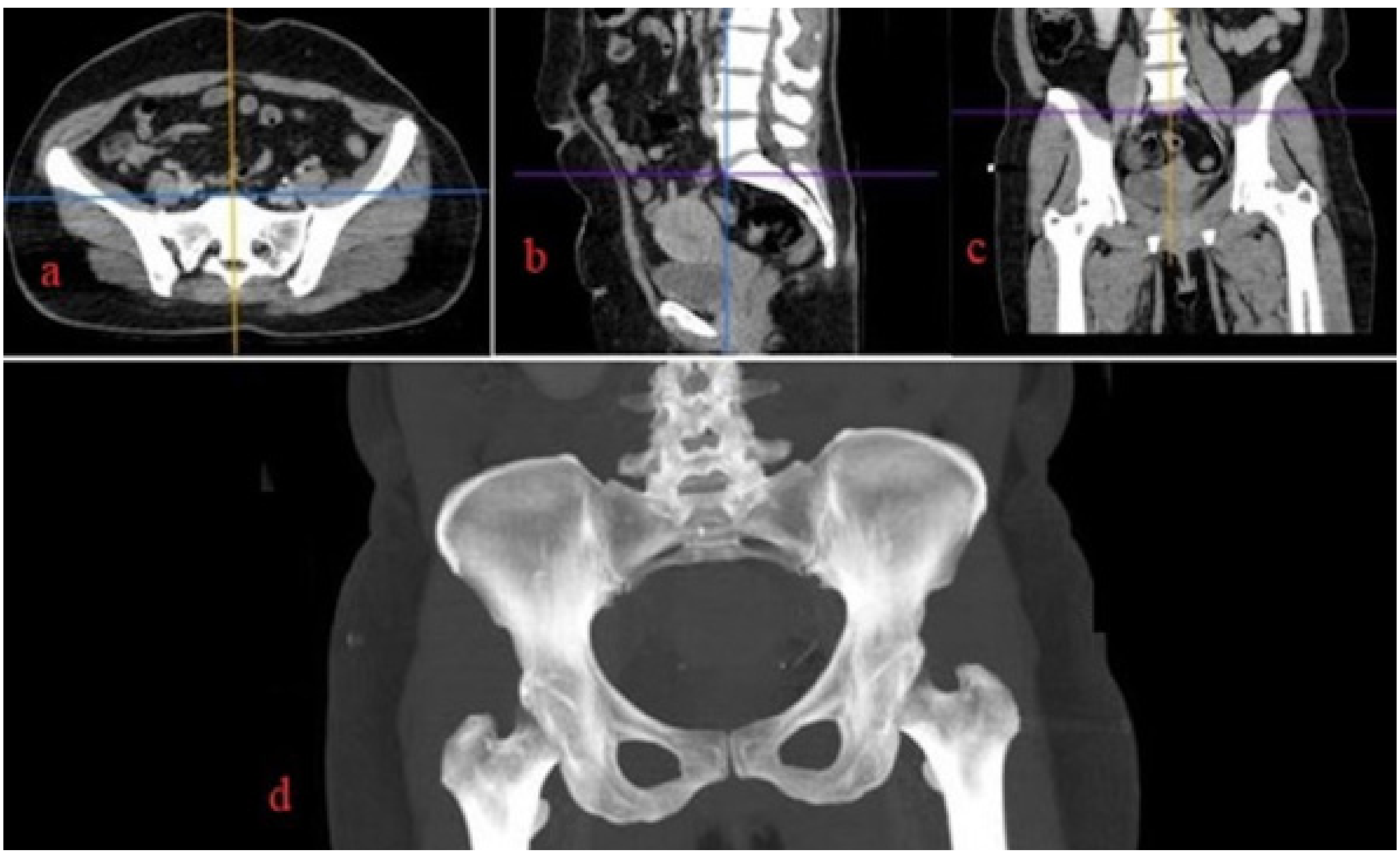

Figure 1. (a) the transversal image in orthogonal form, (b) the sagittal image in orthogonal form, (c) coronal image in the orthogonal form (d) superposed image

Various DT algorithms were applied to these parameters and the models with an Acc ratio of 0.90 or above were selected. The models with an Acc of 0.80 - 0.90 were excluded from the study (Table 3).

Table 1. Determined length parameters

Units	Defined landmarks	
Length 1	Left iliac crest	Right iliac crest
Length 2	Right iliac crest	Promontorium
Length 3	Left iliac crest	Promontorium
Length 4	Superior pubic symphysis	Promontorium
Length 5	Inferior pubic symphysis	Promontorium
Length 6	Right terminal line	Promontorium
Length 7	Left terminal line	Promontorium
Length 8	Left terminal line	Right sacroiliac joint
Length 9	Right terminal line	Left sacroiliac joint
Length 10	Left anterior superior iliac spine	Right anterior superior iliac spine
Length 11	Left anterior inferior iliac spine	Right anterior inferior iliac spine
Length 12	Left terminal line	Right terminal line
Length 13	Left greater trochanter	Right greater trochanter
Length 14	Left lesser trochanter	Right lesser trochanter
Length 15	Left ıschial tuberosity	Right ıschial tuberosity
Length 16	Acetabulum inferior dextra	Acetabulum superior dextra
Length 17	Acetabulum inferior sinistra	Acetabulum superior sinistra
Length 18	Left sacroiliac joint	Right sacroiliac joint
Length 19	Left femoral head	Right femoral head
Length 20	Left superior obturator for.	Right superior obturator for.
Length 21	Right inferior obturator for.	Left inferior obturator for.
Length 22	Right inferior obturator for.	Right superior obturator for.
Length 23	Left inferior obturator for.	Left superior obturator for.
Length 24	Inferior pubic symphysis	Superior pubic symphysis
Length 25	Acetabulum superior sinistra	Acetabulum superior dextra
Length 26	Acetabulum inferior sinistra	Acetabulum inferior dextra
Length 27	Superior pubic symphysis	Right sacroiliac joint
Length 28	Right sacroiliac joint	Inferior pubic symphysis
Length 29	Superior pubic symphysis	Left sacroiliac joint
Length 30	Inferior pubic symphysis	Left sacroiliac joint
Length 31	Left femoral neck	Right femoral neck
Length 32	Left body of femur	Right body of femur
Length 33	Right femoral head	Promontorium
Length 34	Left femoral head	Promontorium

Table 2. Determined angle parameters

Units	Defined landmarks		
Angle 35	Left iliac crest	Promontorium	Right iliac crest
Angle 36	Left terminal line	Promontorium	Right terminal line
Angle 37	Left ıschial tuberosity	Inferior pubic symphysis	Right ıschial tuberosity
Angle 38	Right body of femur	Right femoral neck	Right femoral head
Angle 39	Left body of femur	Left femoral neck	Left femoral head
Angle 40	Left obturator for.	Promontorium	Right superior obturator for.
Angle 41	Left terminal line	Superior pubic symphysis	Right terminal line
Angle 42	Left femoral head	Promontorium	Right femoral head

Table 3. DT algorithm models

Model No	Parameters
Model 1	Length 2 – 7, Length 33 – 34, Angle 35 – 36, Angle 40
Model 2	Angle 35 – 42
Model 3	Angle 35 – 42, Length 2 – 7, Length 33 – 34
Model 4	Angle 35 – 42, Length 2 – 7, Length 10 – 14, Length 33 – 34

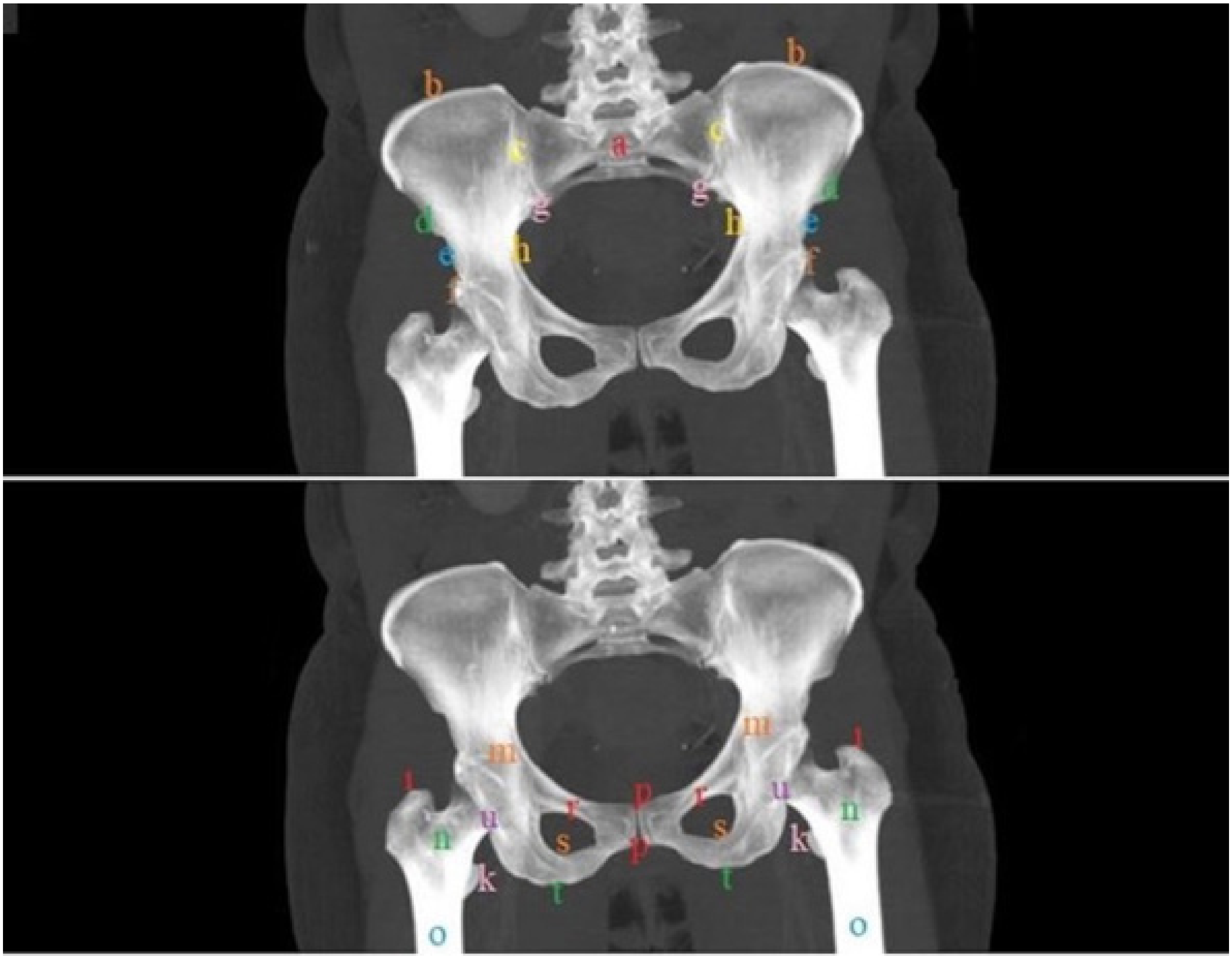

Figure 2. (a) promontorium, (b) iliac crest, (c) sacroiliac joint, (d) anterior superior iliac spine, (e) anterior inferior iliac spine, (f) acetabulum superior, (g) sacroiliac joint lower end, (h) terminal line, (ı) greater trochanter, (k) lesser trochanter, (m) femoral head, (n) femoral neck, (o) body of femur, (u) acetabulum inferior, (p) pubic symphysis, (r) superior obturator foramen, (s) inferior obturator foramen, (t) ischial tuberosity

Decision Tree Algorithm and Performance Criteria

DT is an ML algorithm that models multiple variables in the form of the root, node, branch, and leaf in an inverted tree structure, where the root is at the top and the leaves are at the bottom. It categorizes the target variable by continuously dividing multiple variable inputs. Nodes are divided into three classes of the root node, inner node, and final (leaf) node. The root node divides the input data into subunits, the inner node represents one of the probable possibilities, and the final node represents the result based on the given data. Structures are called branches to provide the connection between the nodes. The number of leaves is associated with the initial DT node; when a correct node is not selected, the leaf density increases. DT is commonly used in health and forensic medicine due to several advantages such as accessibility, cost, ease of application, and usability even in the absence or lack parameters [12].

For the Sekazu (Version 3.0) employed in the present study, the selected DT parameters included unlimited max depth of the tree, unlimited max features, unlimited max number of leaf nodes, and Gini analytical criteria.

Performance criteria included Acc, specificity (Spe), sensitivity (Sen), Matthews correlation coefficient (Mcc), and F1 score obtained with a confusion matrix. 80% of the data were used in the training set and 20% were used in the test set.

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

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

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

$$Mcc = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)\times(TP+FN)\times(TN+FP)\times(TN+FN)}}$$

$$F1 = 2\,\frac{Specificity \times Sensitivity}{Specificity + Sensitivity}$$

Equation 1. (TP: True positive, TN: True negative, FP: False positive, FN; False negative)

Results

In the present study, the DT algorithm was applied to models with 34 length and 8 angle variations, and it was determined that the accof 4 models was 0.90 or above. Model 1 revealed 0.91 Acc, 0.93 Sen, 0.88 Spe, 0.82 Mcc. Model 2 revealed 0.92 Acc, 0.93 Sen, 0.90 Spe, 0.84 Mcc. Model 3 revealed 0.92 Acc, 0.93 Sen, 0.90 Spe, 0.84 Mcc. Model 4 revealed 0.93 Acc, 0.95 Sen, 0.90 Spe, 0.86 Mcc. The most successful model was model 4 with 0.86 Mcc (Table 4).

Table 4. Performance criteria results for models

No	Acc	Sen	Spe	F1	Mcc
Model 1	0.91	0.93	0.88	0.90	0.82
Model 2	0.92	0.93	0.90	0.92	0.84
Model 3	0.92	0.93	0.90	0.92	0.84
Model 4	0.93	0.95	0.90	0.93	0.86

Acc: Accuracy, Sen: Sensitivity, Spe: Spesificity, F1 score, and Mcc: Matthews correlation coefficient for the models

Discussion

Gender prediction based on the skeletal system could be conducted with osteometry or morphological methods. The morphological methods are primitive, do not reflect the facts accurately, and could lead to different results based on the expertise of the observer, while osteometry is a more realistic and advanced method that does not depend on observer expertise that much [10, 13]. In this study, it has been tried to increase the Acc rate that can be obtained from osteometric measurement by using today's modern imaging technology CT.For this purpose, osteometry measurements were made from pelvic CT images of 300 people and analyzed with the DT algorithm. The analysis revealed 0.93 Acc, 0.95 Sen, 0.90 Spe, 0.86 Mcc for model 4.

The advantages of the DT algorithm include its ability to easily classify large and complex inputs, categorize missing data and allow these to contribute to the results, its non-parametric algorithmic structure, facility to predict subsequent tree models based on the pre-created tree models, its ability to determine the specific contribution of input parameters to the results, its ability to analyze outliers and extreme values and easy interpretation. The disadvantages include the problems that could be experienced with small inputs and the inability to establish a causality hypothesis [12]. In the present study, the DT algorithm, commonly used in health and forensic medicine of late, was utilized due to several advantages of the DT algorithm and to contribute to the literature.

Commonly, discriminant analysis is preferred in gender prediction; however, it was observed that the correlation between gender and bone measurements is mostly non-linear [14]. This emphasizes the significance of current methods instead of conventional methods in gender prediction. The preferred bone for testing with current methods is the pelvic skeleton that has been studied for several years and is considered as the most dimorphic structure due to the sex hormones [15].

Savall et al. placed landmarks on 17 anatomical points in os coxae CT images of 59 females and 54 males and determined the Acc of 0.92 with DT algorithm analysis [16]. In the present study, unlike previous studies, the whole pelvic skeleton was analyzed. After the CT images were transformed into orthogonal form based on the promontory point, the landmarks were placed on 35 anatomical points in the pelvis skeleton by superimposition, and 0.93 Acc, 0.95 Sen, 0.90 Spe, 0.86 Mcc were obtained with model 4.

Coelho and Curate conducted 38 metric measurements (33 lengths, 5 angles) on 131 female and 125 male pelvic skeletons in the Coimbra skeleton collection, and reported 0.85 Acc with DT algorithm analysis [15]. In the present study, CT, a modern imaging technique, was employed for a current study and to prevent the impact of orientation on Acc and to obtain more realistic results.

Pretorius et al. placed landmarks on the incisura ischiadica section of the pelvis skeletons of 31 females and 29 males in Pretoria skeleton collection and reported 0.87 Acc in females and 0.93 in males with the Geometric Morphometric method [17]. McBride et al. analyzed 31 morphological features in 115 os coxae bones in Terry skeleton collection with the DT algorithm using 70% for training and 30% for the tests and reported 0.90 Acc [18]. Bonczarowska et al. placed landmarks on the anatomical points in the sciatic notch, auricular surface, angular apex, and straight inferior portion on the os ilium of 88 females and 106 males, and reported 0.831 Acc with logistic regression [5]. In the present study, the sample size (150 females, 150 males) and the number of parameters (34 lengths, 8 angles) were higher and a higher Acc was obtained with the DT algorithm.

Steyn and İşcan analyzed 17 anatomical points on pelvic skeletons of 95 females and 97 males in Crete bone collection and reported 0.935 Acc with the differential function and 0.594 Acc by applying the differential function on sacrum width and length [19]. The results demonstrated that the sacrum was not dimorphic. In the study, 0.90 Acc were obtained with length and angle measurements based on the promontorium in all four models and it was observed that the pelvic skeleton was extremely dimorphic.

In the above-mentioned studies, it was observed that Acc was selected as the performance criterion. Since Acc does not provide direct information on Spe and Sen, the Mcc value, which reflects

Acc, Spe, and Sen, was employed in the present study. Mcc varies between +1 and -1. A value of +1 indicates that there are no errors in the prediction, a value of 0 indicates that the results are random, and a value of -1 indicates that the classes are inversely predicted [20].

In conclusion, the present study demonstrated that the pelvic skeleton could be used in gender prediction with high Acc (0.91-0.93) with DT analysis performed on parameters based on the promontorium in CT images. Furthermore, the literature review suggested that the present study will contribute to the literature by determining angle and length values that were not previously determined for the pelvic skeleton.

Conflict of interests
The authors declare that they have no competing interests.

Financial Disclosure
All authors declare no financial support.

Ethical approval
The present study was conducted after the approval of the ethics committee (Act no: 6/23 on 07/10/2019).

References

1. Colman KL, van der Merwe AE, Stull KE, 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. Giurazza F, Schena E, Del Vescovo R, et al. Sex determination from scapular length measurements by CT scans images in a Caucasian population. Conf Proc IEEE Eng Med Biol Soc. 2013;2013:1632-5.
3. Sinhorini PA, Costa IAP, Lopez-Capp TT, et al. Comparative analysis of four morphometric methods for sex estimation: A study conducted on human skulls. Leg Med (Tokyo). 2019;39:29-34.
4. Ahmed AA. Estimation of sex from the upper limb measurements of Sudanese adults. J Forensic Leg Med. 2013;20:1041-7.
5. Bonczarowska JH, Bonicelli A, Papadomanolakis A, et al. The posterior portion of the ilium as a sex indicator: A validation study. Forensic Sci Int. 2019;294:1- 6.
6. Kim DI, Kwak DS, Han SH. Sex determination using discriminant analysis of the medial and lateral condyles of the femur in Koreans. Forensic Sci Int. 2013;233:121-5.
7. Best KC, Garvin HM, Cabo LL. An investigation into the relationship between human cranial and pelvic sexual dimorphism. J Forensic Sci. 2018;63:990-1000.
8. Blake KAS, Hartnett-McCann K. Metric assessment of the pubic bone using known and novel data points for sex estimation. J Forensic Sci. 2018;63:1472-8.
9. Karakas HM, Harma A, Alicioglu B. The subpubic angle in sex determination: anthropometric measurements and analyses on Anatolian Caucasians using multidetector computed tomography datasets. J Forensic Leg Med. 2013;20:1004-9.
10. Hayashizaki Y, Usui A, Hosokai Y, et al. Sex determination of the pelvis using Fourier analysis of postmortem CT images. Forensic Sci Int. 2015;246:1-9.
11. Climent MT, Pardo J, Munoz-Almaraz FJ, et al. Decision tree for early detection of cognitive impairment by community pharmacists. Front Pharmacol. 2018;9:1232.
12. Song YY, Lu Y. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry. 2015;27:130-5.
13. Wankhede KP, Bardale RV, Chaudhari GR, et al. Determination of sex by discriminant function analysis of mandibles from a Central Indian population. J Forensic Dent Sci. 2015;7:37-43.
14. Oner Z, Turan MK, Oner S, et al. Sex estimation using sternum part lenghts by means of artificial neural networks. Forensic Sci Int. 2019;301:6-11.
15. d'Oliveira Coelho J, Curate F. CADOES: An interactive machine-learning approach for sex estimation with the pelvis. Forensic Sci Int. 2019;302:109873.
16. Savall F, Faruch-Bilfeld M, Dedouit F, et al. Metric sex determination of the human coxal bone on a virtual sample using decision trees. J Forensic Sci. 2015;60:1395-400.
17. Pretorius E, Steyn M, Scholtz Y. Investigation into the usability of geometric morphometric analysis in assessment of sexual dimorphism. Am J Phys Anthropol. 2006;129:64-70.
18. McBride DG, Dietz MJ, Vennemeyer MT, et al. Bootstrap methods for sex determination from the os coxae using the ID3 algorithm. J Forensic Sci. 2001;46:427-31.
19. Steyn M, Iscan MY. Metric sex determination from the pelvis in modern Greeks. Forensic Sci Int. 2008;179:1-6.
20. Turan MK, Oner Z, Secgin Y, et al. A trial on artificial neural networks in predicting sex through bone length measurements on the first and fifth phalanges and metatarsals. Comput Biol Med. 2019;115:103490.