MCBU SBED
MANİSA CELAL BAYAR ÜNİVERSİTESİ SAĞLIK BİLİMLERİ ENSTİTÜSÜ DERGİSİ
MANISA CELAL BAYAR UNIVERSITY JOURNAL OF INSTITUTE OF HEALTH SCIENCE
ISSN: 2147-9607

ARAŞTIRMA MAKALESİ
RESEARCH ARTICLE
CBU-SBED, 2024, 11 (3): 447-451

Gender Prediction From Foramen Magnum Using Machine Learning Algorithms

Makine Öğrenimi Algoritmaları Kullanılarak Foramen Magnum'dan Cinsiyet Tahmini

Nesibe Yılmaz[1], Yusuf Secgin[1], Rabia Demirci[2], Nevin Köremezli Keskin[3]

[1]Karabük University, Faculty of Medicine, Department of Anatomy, Karabük, Turkey
[2]Karabük University, Graduate School of Education, Department of Anatomy, Karabük, Turkey
[3]Karabük University, Faculty of Medicine, Department of Radiology, Karabük, Turkey

e- mail: nesibeyilmaz@karabuk.edu.tr, yusufsecgin@karabuk.edu.tr, demircirabia@gmail.com, nevinkoremezlikeskin@karabuk.edu.tr
ORCID: 0000-0002-5527-8507
ORCID: 0000-0002-0118-6711
ORCID: 0009-0002-0159-1909
ORCID: 0000-0002-3169-9083

*Sorumlu Yazar / Corresponding Author: Nesibe Yılmaz
Gönderim Tarihi / Received: 05.06.2024
Kabul Tarihi / Accepted: 11.07.2024
DOI: 10.34087/cbusbed.1496208

Öz

Giriş ve Amaç: Bireyin ölüm sonrası doğru ve yüksek doğrulukta kimliklendirilmesi adli bilimler için önemlidir. Kimlik tespitinin ana unsuru ise cinsiyetin belirlenmesidir. Bu çalışmanın amacı, Bilgisayarlı Tomografi (BT) görüntülerinden elde edilen foramen magnum parametreleri ile Makine öğrenimi (ML) algoritmalarını kullanarak cinsiyeti yüksek doğrulukla tahmin etmektir.

Gereç ve Yöntemler: Çalışma 18-65 yaş arası 214 bireyin BT görüntüleri üzerinde gerçekleştirilmiştir. Görüntüler üzerinde foramen (for.) magnum uzunluğu, genişliği, çevresi ve alanı ölçülmüştür. Ölçümler cinsiyet tahmini için ML algoritmalarında kullanıldı ve performans değerleri kaydedildi.

Bulgular: Çalışma sonucunda ML algoritmaları ile 0,84'e varan yüksek doğrulukta cinsiyet tahmin sonuçları elde edilmiştir. Ayrıca SHapley Additive exPlanations analizörü ile bu yüksek orana en çok for. magnum yükseklik parametresinin katkı sağladığı tespit edilmiştir.

Sonuç: Çalışmamız sonucunda for. magnum'dan elde edilen parametrelerin ML algoritması ile analiz edilerek Türk popülasyonunda cinsiyet tahmini için kullanılabileceği tespit edilmiştir. Bu açıdan Türk popülasyonunda yapılacak diğer metrik çalışmalarına katkı sağlayacağını düşünüyoruz.

Anahtar kelimeler: Foramen Magnum, Bilgisayarlı Tomografi, Makine Öğrenmesi Algoritmaları, Cinsiyet Tahmini.

Abstract

Aim; Accurate and highly accurate postmortem identification of the individual is important for forensic sciences. The main element of identification is the determination of gender. The aim of this study is to predict gender with high accuracy using Machine learning (ML) algorithms with parameters of the foramen magnum obtained from Computed Tomography (CT) images.

Method; The study was performed on CT images of 214 individuals aged 18-65 years. Foramen (for.) magnum length, width, circumference and area were measured on the images. The measurements were used in ML algorithms for gender prediction and performance values were recorded.

Results; As a result of the study, gender prediction results with high accuracy up to 0.84 were obtained with ML algorithms. In addition, it was found that the for. magnum height parameter contributed the most to this high rate with the SHapley Additive exPlanations analyzer.

Conclusion; As a result of our study, it was found that the parameters obtained from for. magnum can be used for sex prediction in the Turkish population by analyzing them with ML algorithm. In this respect, we think that it will contribute to other metric studies in the Turkish population.

Keywords: Foramen Magnum, Computed Tomography, Machine Learning Algorithms, Gender Prediction.

1. Introduction

Identification of the deceased individual or individuals is important for forensic processes. In this identification phase, skeletal tissue comes to the fore due to its ability to be preserved in nature for many years. Within the skeletal tissue, pelvis and skull bones stand out due to their high accuracy in identification. The foramen magnum, one of the structures of the occipital bone, can be preserved in nature for many years because it is located at the base of the skull and surrounded by muscle and connective tissue. Because of this feature, it stands out among the skull bones in terms of identification [1-3]. The first biological characteristic to look at for identification from the skeletal tissue obtained is gender [1]. This is followed by age, height, ancestry, time of death and cause of death [3]. Sex estimation is seen as a simple categorization of an individual as male or female, but it is complex in the case of unidentified or mutilated bodies in situations of mass disaster or crime [4].

Morphological and metric methods for sex estimation include [1]. Among these methods, the morphological ones have serious disadvantages as they depend on the experience of the evaluator. However, they are preferable to metric methods as they do not require special equipment [1, 5]. Radiodiagnostic techniques come to the forefront for metric methods. In the 19th century, radiodiagnostic methods were used for diseases, but they have also been used in forensic medicine, albeit infrequently. Among radiodiagnostic methods, Computed Tomography (CT) stands out because it does not require tissue maceration, has high contrast resolution, and offers three-dimensional measurement and correction. However, it has the disadvantage of being expensive compared to basic metric methods [4-6].

Classification (prediction, categorization), pattern recognition, image processing, data analytics are important topics. Machine learning algorithms (ML) divide the inputs for classification into training and testing based on certain principles. It then applies the algorithm suitable for the model used to the training set and trains it. The trained data is then tested on the test set to demonstrate the performance of the model. ML is an engineering-based methodology that is more reliable and more accurate than basic statistical analysis because it tests the data [7-9]. The Extra Tree Classifier (ETC) is similar in structure to the Random Forest (RF) algorithm, but the ETC algorithm grows trees using all the learning data and splits the nodes at random points. RF is an ensemble algorithm that classifies by combining the prediction results of multiple Decision Tree (DT) predictions. Gaussian Naive Bayes (GaussianNB) classifies inputs according to a Gaussian distribution. Linear Discriminant Analysis (LDA) is an extension of Fisher's linear discriminant analysis, which performs well in classifying non-linear data, to machine learning (ML) algorithms. K-Nearest Neighbors (k-NN) is a simple non-parametric algorithm that performs classification based on the nearest neighbor. DT is a tree-like classification algorithm. Quadratic discriminant analysis (QDA) is an algorithm that, unlike LDA, works well on nonlinear data and has no hyperparameters. Logistic regression (LR) is a classification method that can analyze multivariate data [10-12].

The aim of this study is to predict gender using machine learning algorithms with parameters of the foramen (for.) magnum obtained from CT.

2. Material and Methods

Retrospective study Karabük approved by the decision of the local non-interventional ethics committee of the University of 2023/1483 numbered and dated 08.11.2024. The CT images in the study were obtained randomly from the archive of the Department of Radiology, Karabük University Training and Research Hospital. The study was performed on CT images of 89 males and 125 females aged 18-65 years. Exclusion criteria were pathology or surgical intervention in the cranium.

Images were obtained from 5-mm slice thickness CT images acquired in the supine position using a 16-row MDCT scanner (Aquilion 16; Toshiba Medical Systems, Otawara, Japan). Scan protocol values were tube voltage: 120 kV, gantry rotation: 0.75 s and inclination: 1.0 mm.

Images in Digital Imaging and Communications in Medicine (DICOM) format were imported into Radiant Dicom Wiewer (64-bit version) and all images were brought into axial, sagittal and coronal planes using the 3D MPR console of the program. Sagittal images were superimposed and contrast was adjusted to measure the length, width, curvature length (circumference) and area of the foramen magnum. Foramen magnum circumference: The curvature length of the foramen magnum, foramen magnum length: the distance from the top to the bottom of the foramen magnum, foramen magnum width: the widest distance of the foramen magnum, foramen magnum area: the area of the foramen magnum in cm^2. (Figure 1).

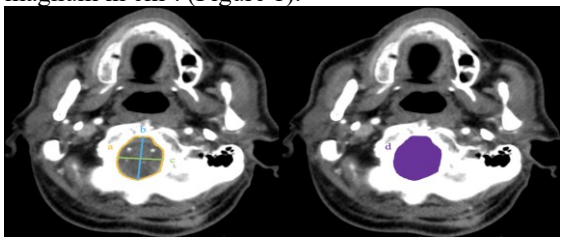

Figure 1. Demonstration of parameters (a: foramen magnum circumference, b: foramen magnum length, c: foramen magnum width, d: foramen magnum area)

2.1. Analyzing Machine Learning Algorithms

For ML algorithms analysis, measurement parameters were set as input and gender as output. The distribution of the data was set as 80% training and 20% testing. The analysis was performed on a personal computer with Monster Abra i5 operating system. The analysis was performed with Python 3.9 programming language and scikit-learn 1.1.1 framework. LR, ETC, RF, LDA, QDA, DT, GaussianNB, k-NN were used as machine learning models. Accuracy (Acc), Specificity (Spe), Sensitivity (Sen), F1 score (F1) values were used for model performance. The contribution of the parameters was evaluated using the SHapley Additive exPlanations (SHAP) solver of the RF algorithm.

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

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

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

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

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

2.2. Statistical analysis

The conformity of the data to normal distribution was tested using the Anderson Darling test. Mean±standard deviation was used for descriptive statistics of normally distributed parameters and median (minimum and maximum) values were used for non-normally distributed parameters. Two simple T test was used to compare normally distributed data in terms of gender, and Mann Whitney U test was used for non-normally distributed data. Minitab 17 package program was used in statistical analyses and $p<0.05$ was considered significant.

3. Results

In this study of 89 males and 125 females, it was found that all parameters except foramen magnum length were not normally distributed. According to the two simple T test, the foramen magnum length parameter was found to be 3.387±0.313 cm in females and 3.856±0.293 cm in males and had a significant difference in terms of gender ($p=0.000$). Descriptive statistics and Mann-Whitney-U test results of the parameters that were not normally distributed are given in Table 1.

Parameters	Gender	Median	Minimum	Maximum	p
For. magnum width (cm)	M	3.18	2.86	3.89	**0.000**
	F	2.78	2.22	3.49	
For. magnum circumference (cm)	M	11.2	10.2	12.73	**0**
	F	9.88	7.89	11.63	
For. magnum field (cm^2)	M	9.17	7.47	12.21	**0.000**
	F	7.09	4.80	9.97	

*M: Male, F: Female

Table 1. Descriptive statistics of non-normally distributed parameters and gender comparison

As a result of ML analysis, the highest Acc ratio was found to be 0.84 with ETC, LDA and LR algorithm. The Acc ratio of the other algorithms ranged between 0.79 and 0.81 (Table 2).

Algorithms	Acc	Spe	Sen	F1
ETC	0.84	0.84	0.88	0.83
LDA	0.84	0.83	0.86	0.83
QDA	0.81	0.81	0.84	0.81
k-NN	0.81	0.83	0.86	0.81
LR	0.84	0.83	0.86	0.83
GaussianNB	0.79	0.78	0.75	0.76
RF	0.81	0.83	0.86	0.81
DT	0.81	0.83	0.86	0.81

Table 2. Analysis of machine learning algorithms

The confusion matrix table of the ETC algorithm with the highest Acc and Spe values is given in Table 3.

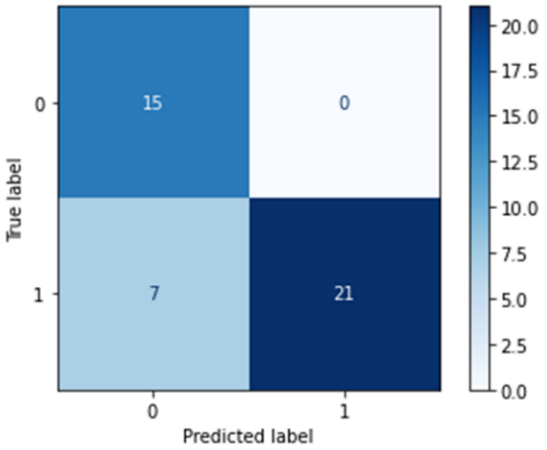

Table 3. Confusion matrix tables

The contribution of the parameters to the overall result was tested with the SHAP analyzer and it was found that for. magnum height contributed the most to sex (Figure 2).

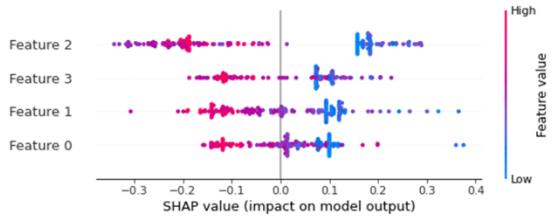

Figure 2. SHAP analyzer (Feature 0: foramen magnum area, 1: foramen magnum circumference, 2: foramen magnum height, 3: foramen magnum width)

4. Discussion

In this study, 8 different models were tested and an accuracy rate between 0.79-0.84 was obtained with the parameters of the foramen magnum obtained from CT using ML algorithms. With the SHAP analyzer, it was found that foramen magnum height contributed the most to this accuracy.

In forensic cases where the pelvis cannot be preserved, the structures of the skull come to the fore. In cases where the skull is also fragmented, the for. magnum comes to the fore because it is located at the base of the skull and is surrounded by large soft tissue compared to other cranial structures [13].

ML algorithms are a method that combines information theory, artificial intelligence and detailed computational statistics to train and classify data. In addition, ML algorithms are slowly gaining ground in forensic studies due to their ability to quickly and accurately reveal complex relationships between unknown data [14, 15].

Gapert et al. used linear discriminant analysis and regression analysis in their study of 158 human skulls from the 18th and 19th centuries from St. Bride's Church and obtained an accuracy rate of 68% and reported that for. magnum can be used for sex estimation [16].

Gender estimation may give different results for individuals of different ethnicities, so it is important to estimate gender on different ethnicities [15]. Edwards et al. used discriminant analysis and binary logistic regression analysis in their study on CT images of 144 male and 106 female individuals and reported a gender prediction rate of up to 66.4% from for. magnum [17]. Mehta et al. In their study on CT images of 291 male and 262 female individuals aged 18-60 years from the West Indian population, they obtained a sex prediction rate of up to 69.1% from for. magnum and concluded that it cannot be used for sex prediction in the West Indian population [18]. Atreya et al. In their study on CT images of 261 Nepalese individuals, they obtained an accuracy rate between 70.5-71% with discriminant function analysis and reported that for. magnum was less accurate in determining the sex of Nepalese individuals [19]. Kalbouneh et al. In their study on CT images of 500 individuals in the Jordanian population, they took measurements of the for. magnum and occipital condyle and obtained an accuracy rate of up to 78.8% and reported higher prediction rates in male individuals compared to female individuals [20]. Madadin et al. In their study on CT images of 200 adult individuals from the Saudi Arabian population, they obtained a sex prediction rate of up to 71% [21]. In this study, we used a recent methodology, engineering-based ML algorithms, which offer higher accuracy and precision than classical statistical methods, and obtained gender prediction results with up to 84% accuracy in the Turkish population.

Meral et al. In their study on CT images of 300 female and 300 male individuals in the Turkish population, they found for. magnum length as 37.54±2.86 mm in male individuals and 34.76±2.64 mm in female individuals, for. magnum width as 32.75±2.46 mm in male individuals and 29.98±2.43 mm in female individuals, for. magnum area as 967.66±125.40 mm^2 in male individuals and 820.49±109.30 mm^2 in female individuals [5]. In this study, we obtained similar results for. magnum length as 3.856±0.293 cm in males and 3.387±0.313 cm in females, for. magnum width as 3.180 (2.860-3.890) cm in males and 2.780 (2.220-3.490) cm in females, for. magnum area as 9.168 cm^2 (7.468-12.210) in males and 7.092 cm^2 (4.799-9.970) in females.

Kartal et al. [3] In their study on 360 female and 360 male individuals belonging to the Eastern Turkish population, they obtained a sex prediction rate of 88.2% using artificial neural networks. The difference of our study from this study is the use of ML algorithms, the second algorithm of artificial intelligence. As far as we know, there is no study in the literature using for. magnum morphometry and ML algorithms. In this respect, this study will make an important contribution to the literature.

5. Conclusion

As a result of the study, it was found that sex can be estimated using for. magnum on CT images of individuals in the Turkish population. In this respect, we believe that it will support for. magnum studies in the Turkish population.

6. Acknowledgements and Disclosures Conflict of Interest: The authors declare that they have no conflict of interests regarding content of this article.

Referanslar

1. Babu YR, Kanchan T, Attiku Y, Dixit PN, Kotian M. Sex estimation from foramen magnum dimensions in an Indian population. Journal of forensic and legal medicine. 2012;19(3):162-7.
2. Toneva D, Nikolova S, Harizanov S, Georgiev I, Zlatareva D, Hadjidekov V, et al. Sex estimation by size and shape of foramen magnum based on CT imaging. Legal Medicine. 2018;35:50-60.
3. 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. Legal Medicine. 2022;59:102143.
4. Fernandes ALVC, Shetti A, Lagali-Jirge V, Keluskar V. Accuracy of sex estimation by morphometric evaluation of foramen magnum using computed tomography–a systematic review and meta-analysis. Forensic Science, Medicine and Pathology. 2024;20(1):268-79.
5. Meral O, Toklu BB, Meydan R, Kaya A, Karadayı B, Acar T. Sex estimation from foramen magnum parameters in

adult Turkish population: A computed tomography study. Legal Medicine. 2020;47:101775.
6. Secgin Y, Oner Z, Turan MK, Oner S. Gender prediction with the parameters obtained from pelvis computed tomography images and machine learning algorithms. Journal of the Anatomical Society of India. 2022;71(3):204-9.
7. Singh A, Thakur N, Sharma A, editors. A review of supervised machine learning algorithms. 2016 3rd international conference on computing for sustainable global development (INDIACom); 2016: Ieee.
8. Senol D, Secgin Y, Duman BS, Toy S, Oner Z. Sex and age estimation with machine learning algorithms with parameters obtained from cone beam computed tomography images of maxillary first molar and canine teeth. Egyptian Journal of Forensic Sciences. 2023;13(1):27.
9. Erkartal HŞ, Tatlı M, Secgin Y, Toy S, Duman BS. Gender Estimation with Parameters Obtained From the Upper Dental Arcade by Using Machine Learning Algorithms and Artificial Neural Networks. European Journal of Therapeutics. 2023;29(3):352-8.
10. Saez Y, Baldominos A, Isasi P. A comparison study of classifier algorithms for cross-person physical activity recognition. Sensors. 2016;17(1):66.
11. Attia MH, Attia MH, Farghaly YT, Abulnoor BAE-S, Curate F. Performance of the supervised learning algorithms in sex estimation of the proximal femur: A comparative study in contemporary Egyptian and Turkish samples. Science & Justice. 2022;62(3):288-309.
12. Kuan Y-C, Hong C-T, Chen P-C, Liu W-T, Chung C-C. Logistic regression and artificial neural network-based simple predicting models for obstructive sleep apnea by age, sex, and body mass index. Math Biosci Eng. 2022;19(11):11409-21.
13. Seifert Z, Friedl L, Chaumoitre K, Brůžek J. Applicability and limitations of sex assessment based on foramen magnum. Forensic science international. 2017;271:126. e1-. e9.
14. Navega D, Vicente R, Vieira DN, Ross AH, Cunha E. Sex estimation from the tarsal bones in a Portuguese sample: a machine learning approach. International journal of legal medicine. 2015;129:651-9.
15. Toy S, Secgin Y, Oner Z, Turan MK, Oner S, Senol D. A study on sex estimation by using machine learning algorithms with parameters obtained from computerized tomography images of the cranium. Scientific Reports. 2022;12(1):4278.
16. Gapert R, Black S, Last J. Sex determination from the foramen magnum: discriminant function analysis in an eighteenth and nineteenth century British sample. International journal of legal medicine. 2009;123:25-33.
17. Edwards K, Viner M, Schweitzer W, Thali MJ. Sex determination from the foramen magnum. Journal of forensic radiology and imaging. 2013;1(4):186-92.
18. Mehta M, Saini V, Patel M, Menon S. Applicability and reliability of foramen magnum for sex determination in contemporary Western Indian population: A computed tomographic study. Journal of Forensic Radiology and Imaging. 2019;17:31-5.
19. Atreya A, Shrestha R, Bhandari K, Malla SK, Acharya S, Menezes RG. Morphometric analysis of the foramen magnum in sex estimation: An additional 3DCT study from Nepal on a larger sample. Health Science Reports. 2023;6(1):e999.
20. Kalbouneh H, Mubarak N, Shwaqfeh J, Nawaiseh M, Ashour O, Alkhatib AM, et al. Estimation of sex from measurements of foramen magnum region in a contemporary Jordanian population: A computed tomographic study. Anthropologischer Anzeiger. 2021;78(4).
21. Madadin M, Menezes RG, Al Saif HS, Alola HA, Al Muhanna A, Gullenpet AH, et al. Morphometric evaluation of the foramen magnum for sex determination: A study from Saudi Arabia. Journal of Forensic and Legal Medicine. 2017;46:66-71.