

Article

Machine Learning-Based Sex Classification Using Linear Dimensions of the Talocrural Joint

Songul Cuglan ^{1,*}, Muslume Kucukdemir ², Mustafa Durmaz ³, Kemal Altay ³ and Menduh Dursun ³¹ Department of Anatomy, Faculty of Medicine, Bezmiâlem Vakif University, 34093 Istanbul, Turkey² Department of Anatomy, Faculty of Medicine, Istanbul Beykent University, 34520 Istanbul, Turkey; muslume.kucukdemir@beykent.edu.tr³ Department of Radiology, Faculty of Medicine, Istanbul University, 34093 Istanbul, Turkey; mustafa.durmaz@istanbul.edu.tr (M.D.); kemalty16@gmail.com (K.A.); memduh.dursun@istanbul.edu.tr (M.D.)

* Correspondence: songul.cuglan@bezmialem.edu.tr; Tel.: +90-212-523-22-88

Abstract

Background: Pelvic and cranial markers are often fragmented in mass disasters. The resilient talocrural joint serves as a valuable secondary site. This study presents a novel 12-landmark configuration on conventional 2D AP ankle radiographs as a rapid, cost-effective auxiliary sex estimation tool, bypassing complex 3D reconstructions. **Methods:** Radiographs of 200 contemporary Turkish adults (100 males, 100 females) were calibrated via PACS DICOM metadata. Twelve landmarks mapping distal tibial (T1–T6) and fibular (F7–F12) cortical topography were tracked to derive linear metrics. Models were evaluated using 5-fold cross-validation. **Results:** Optimized logistic regression achieved the highest independent test accuracy of 80.0%, outperforming random forests (72.5%) and support vector machines (70.0%), with an AUC of 0.870 (95% CI: 0.765–0.975). Tibial T2–T3 and T6–T1 were the strongest predictors ($p < 0.001$, Cohen's $d > 1.5$). Although fibular F9–F10 lost significance after Bonferroni correction ($p = 0.552$), it maintained substantial multivariate weight, indicating spatial synergy. **Conclusions:** This interpretable, regularized logistic regression model trained on minimal linear ankle metrics shows that biological sex can be estimated with promising accuracy (80.0%) as an auxiliary tool capable of being seamlessly integrated into clinical PACS and active forensic workflows.

Keywords: radiography; ankle joint; sex determination analysis; forensic anthropology; machine learning; forensic radiology; ankle morphometry



Academic Editor: Thomas Lindner

Received: 18 June 2026

Revised: 17 July 2026

Accepted: 23 July 2026

Published: 3 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Forensic anthropology provides vital solutions for standard identification protocols when soft tissue is compromised. Its interventional value is particularly critical for analyzing skeletal remains yielded by mass disasters and active criminal investigation. Accurate characterization of an individual's biological profile, including age, sex, stature and ethnic background, constitutes a crucial step in this process [1,2]. The pelvic girdle and cranial structures traditionally serve as the definitive reference elements for skeletal sex estimation in anthropological profiling [3,4]. However, despite their high diagnostic accuracy, these primary markers are complex, thin, and hollow, making them highly susceptible to severe post-mortem fragmentation, taphonomic crushing, and incomplete recovery, often leaving them completely unavailable in forensic scenarios [1,5,6]. Long bones offer an essential

secondary diagnostic avenue in forensic contexts where the pelvis or skull remains incomplete or exhibits post-mortem damage [4,5,7]. Standard metric assessments of long bones are traditionally confined to single-dimensional linear features, primarily focusing on longitudinal extensions and diaphyseal shaft diameters [4]. Although researchers have utilized discriminant function analyses on other long bones like the femur and humerus, traditional osteometry remains restricted to simple shaft diameters, failing to capture plastic, load-bearing adaptations at joint surfaces [2,4,8,9].

Sex estimation in the literature has predominantly been performed either directly on dry skeletal material or indirectly through imaging modalities such as conventional radiography, computed tomography and magnetic resonance imaging [6,10–13]. Bone tissue is constantly subjected to dynamic processes, including biomechanical loading and the persistent action of gravitational forces [8,9]. The dynamics of structural bone remodeling can be more precisely elucidated through a rigorous morphometric characterization of skeletal segments immediately neighboring heavy force-transmission zones [8,9]. The ankle joint, notable for its anatomical and biomechanical complexity, coordinates the stabilization and transmission of kinetic energy during weight-bearing activities, with approximately 83% of the total compressive load of the lower limb transmitted directly through the tibio-talar interface [14,15]. Biomechanical loads exerted on this specific region of the locomotor system drive the geometric adaptation of articular surfaces, thereby fostering a pronounced sexual dimorphism in morphometric characteristics [12,16]. During puberty, the differentiation of the skeletal matrix by androgenic and estrogenic hormones results in larger macroscopic bone dimensions and greater bone mass in males compared to females [17]. However, patterns of sexual dimorphism in ankle morphology are not universal and exhibit population-specific variations. These variations stem from fundamental macro-structural dynamics—such as genetic drift, intra-population socio-economic stratification, and long-term adaptive selection—as well as epigenetic factors like dietary habits and microclimatic environmental conditions [18,19].

Modern imaging modalities encompass conventional radiographs, CT scans, and 3D reconstructions. These advanced techniques offer highly detailed, high-resolution representations that are critical for the evaluation of osseous macro- and microarchitecture. Digital methodologies have introduced a robust quantitative framework to forensic anthropology through cross-sectional, two-dimensional, and three-dimensional imaging. Consequently, this perspective enhances the accuracy of deriving core biological profiles—namely sex, age, height, and ethnicity—from postmortem skeletal remains [10–13,20]. In recent years, integrating large-scale morphometric datasets derived from these radiological modalities with machine learning algorithms has yielded a highly objective, reproducible, and accurate methodological framework for sex estimation, effectively transcending the limitations of traditional anthropometric approaches [2,12,13]. A review of the current literature reveals a significant scarcity of forensic sex estimation studies utilizing talocrural joint morphometry specifically within the Turkish population [12]. Although a recent study by Ray et al. [12] utilized machine learning algorithms for sex estimation from the talocrural region, their analysis was confined to conventional joint space distances and general linear dimensions without exploring the detailed peripheral cortical topography, which is highly susceptible to minor positioning errors and degenerative osteoarthritic alterations.

To address this gap, our study introduces a multi-point, interconnected linear configuration of both the distal tibial and fibular cortical frameworks. In doing so, rather than presenting this work as a novel machine-learning breakthrough, the primary methodological contribution of this study is the development of a population-specific radiographic morphometric framework combined with interpretable classification. By utilizing a highly structured, 12-landmark configuration across the distal tibia and fibula, we aim to provide

an accessible and reproducible osteometry reference tailored to a contemporary Turkish population. This approach prioritizes model interpretability and clinical ease-of-use over algorithmic complexity, offering a clear mathematical mapping of talocrural dimorphism on standard, widely available clinical radiographs, thereby demonstrating a robust and stable auxiliary capacity for sex classification.

2. Materials and Methods

This study was approved by the Ethics Committee for Non-Interventional Research at Istanbul University Faculty of Medicine via Decision No. 3552799 dated 1 September 2025. This retrospective study included imaging data from a total of 200 adults (100 women, 100 men) aged 19–75 years. Biological sex was ascertained from the official electronic medical records, which record sex based on national identity documentation at the time of registration. No cases with ambiguous or discordant records were encountered or included. The mean age was 40.01 ± 13.43 years for women and 38.11 ± 12.30 years for men. In our single-center study, anteroposterior (AP) ankle radiographs taken between January 2024 and May 2025 were retrieved via the institutional Picture Archiving and Communication System (PACS) of Istanbul University Faculty of Medicine. Inclusion criteria are images of individuals over 18 years of age, no history of trauma or surgery that would compromise the anatomical integrity of the area being examined, and sufficient image quality. Exclusion criteria include images of individuals under 18 years of age, a history of fracture, motion artifacts, and poor image quality such as inadequate positioning. To avoid artificial sample inflation and dependency clusters, only a single ankle radiograph per patient was included (104 right ankles, 96 left ankles). The selection of the side was based on image quality and strict adherence to the positioning criteria. Preliminary statistical analysis confirmed no significant morphometric differences between right and left configurations within genders ($p > 0.05$), consistent with prior lower-limb anthropological literature.

2.1. Imaging Protocol and Anatomical Measurements

To maintain rigorous measurement consistency and minimize geometric distortion, all digital radiographs were acquired using a digital radiography system in accordance with standardized departmental protocols. In terms of patient positioning, all radiographs were obtained in a weight-bearing (standing) position, with the lower limb placed in a neutral anatomical position and the foot directed perpendicular to the image receptor. Crucially, to clearly visualize the ankle joint space and the distal tibiofibular relationship (syndesmosis), the foot was placed in $15\text{--}20^\circ$ of internal rotation, achieving a standard mortise view. During exposure, technical parameters were automatically optimized between 55–65 kVp and 4–8 mAs, while the source-to-image receptor distance was strictly fixed at 100 cm (40 inches), with the central X-ray beam directed perpendicularly to the midpoint between the medial and lateral malleoli. Prior to any morphometric analysis, all acquired radiographic images were fully anonymized to protect patient confidentiality. These anonymized images were subsequently transferred to a Picture Archiving and Communication System (PACS) and reviewed on high-resolution diagnostic monitors. To guarantee metric precision and control for scale variations, all conventional radiographs were calibrated within the institutional ExtremePACS platform (Version 4.0; Extreme Pacs, Ankara, Turkey). Rather than using external physical calibration objects, the digital caliper tool automatically calibrated pixel dimensions into absolute millimeters (mm) by retrieving the ‘Imager Pixel Spacing’ (0018,1164) and ‘Detector Pixel Spacing’ attributes embedded within the DICOM file metadata. This automated calibration mathematically adjusts for the distance between the X-ray tube, the patient, and the digital detector. When combined with

these standardized acquisition and positioning parameters, this digital calibration protocol effectively eliminated projection magnification and ensured uniform scale reproducibility across all subjects. Finally, all morphometric measurements were performed independently by two distinct radiologists, each possessing a minimum of five years of professional experience, using this calibrated digital caliper tool.

Twelve reproducible bony landmarks were identified along the distal tibia (T1–T6) and distal fibula (F7–F12) on the digital radiographs, with their precise anatomical definitions and geometric criteria consolidated in Table 1. Based on these landmarks, twelve linear distance parameters were systematically measured to map the architectural dimensions of the talocrural joint: six intervals across the distal tibia (T1–T2, T2–T3, T3–T4, T4–T5, T5–T6, and T6–T1) and six parameters across the distal fibula (F7–F8, F8–F9, F9–F10, F10–F11, F11–F12, and F12–F7). These measurements are rigorously defined in Table 2 and schematically illustrated in Figure 1. To ensure absolute consistency and minimize interobserver variability across different reader panels, the landmarking and measurement process must be executed according to the following standardized, step-by-step protocol:

Step-by-Step Measurement Protocol

Step 1: Establish the horizontal and vertical reference axes of the distal tibia. Define the distolateral corner (T1) and proximolateral corner (T2) of the incisura fibularis.

Step 2: Draw an orthogonal horizontal line starting from T2 to the medial cortex to locate T3.

Step 3: Draw a horizontal line starting from T1 parallel to the T2–T3 axis. Measure the maximum perpendicular distance from this line to the medial tibial flare to define T4.

Step 4: Place T5 at the absolute distal tip of the medial malleolus.

Step 5: Locate T6 at the deepest, most superior inflection point of the subchondral tibial plafond.

Step 6: Move to the fibula. Identify F7 and F8 at the narrowest portion of the distal fibular shaft aligned with the superior edge of the incisura.

Step 7: Draw a vertical line along the medial fibular cortex from F8. Locate F9 at the medial-most projection point of the cortex.

Step 8: Locate F11 at the distal-most tip of the lateral malleolus.

Step 9: Establish a horizontal reference line from F9. Locate F12 at the maximum lateral extension of the lateral malleolus relative to this axis.

Step 10: Define F10 at the medial-most articular projection of the lateral malleolus. Measure all 12 consecutive linear paths as detailed in Table 2.

Table 1. Concise Anatomical Definition of Landmarks.

Landmark	Anatomical Landmark Name	Geometrical and Radiographic Definition
T1	Distolateral Tibial Corner	The distal-most lateral point of the tibia forming the inferior boundary of the incisura fibularis.
T2	Proximolateral Tibial Corner	The proximal-most lateral point of the tibia forming the superior boundary of the incisura fibularis.
T3	Medial Tibial Baseline Intersection	The point where a horizontal reference line drawn orthogonally from T2 intersects the medial cortical border of the tibia.
T4	Medial Tibial Inflection Point	The point on the medial margin of the tibia exhibiting the maximum perpendicular distance from the horizontal reference line extending from T1.
T5	Medial Malleolar Apex	The distal-most inferior tip of the medial malleolus.

Table 1. *Cont.*

Landmark	Anatomical Landmark Name	Geometrical and Radiographic Definition
T6	Tibial Plafond Deepest Point	The deepest point of maximum geometric curvature (inflection) along the subchondral cortical plafond of the distal tibia.
F7	Proximolateral Fibular Corner	The proximal-most lateral contour point of the fibula aligned at the level of the tibial notch.
F8	Proximomedial Fibular Corner	The proximal-most medial contour point of the fibula nearest the incisura fibularis.
F9	Fibular Shaft Width Anchor	The medial-most contour point at the maximum width of the fibular shaft, determined relative to a vertical axis drawn from F8.
F10	Medial Malleolar Inflection Point	The distal-medial corner of the lateral malleolus at its farthest projection from the lateral cortical border.
F11	Lateral Malleolar Apex	The distal-most inferior tip of the lateral malleolus.
F12	Lateralmost Fibular Contour	The lateral-most contour point of the fibula, determined at the maximum perpendicular distance from the horizontal reference axis drawn through F9.

Table 2. Definition of Derived Linear Distances.

Distance Parameter	Associated Anatomical Segment	Definition and Measurement Axis
T1–T2 (mm)	Incisura Fibularis Height	The vertical distance between the lower lateral corner (T1) and upper lateral corner (T2) of the tibial notch.
T2–T3 (mm)	Distal Tibial Metaphyseal Width	The linear distance between T2 and its orthogonal intersection with the medial tibial cortex (T3).
T3–T4 (mm)	Supramalleolar Medial Flare	The linear distance between the medial baseline intersection (T3) and the medial tibial inflection point (T4).
T4–T5 (mm)	Medial Malleolar Length	The linear distance between the medial inflection point (T4) and the medial malleolar apex (T5).
T5–T6 (mm)	Medial Plafond Width	The linear distance between the medial malleolar apex (T5) and the deepest subchondral inflection point of the plafond (T6).
T6–T1 (mm)	Lateral Plafond Width	The linear distance between the deepest subchondral inflection point of the plafond (T6) and the distolateral tibial corner (T1).
F7–F8 (mm)	Proximal Fibular Width	The linear distance between the proximolateral (F7) and proximomedial (F8) corners of the distal fibular shaft.
F8–F9 (mm)	Upper Fibular Expansion	The linear distance between the proximomedial corner (F8) and the medial shaft width anchor (F9).
F9–F10 (mm)	Lateral Malleolar Joint Interface	The linear distance between the medial shaft anchor (F9) and the medial malleolar inflection point (F10).
F10–F11 (mm)	Medial Lateral Malleolar Height	The linear distance between the medial malleolar inflection point (F10) and the lateral malleolar apex (F11).
F11–F12 (mm)	Malleolar Flare Width	The linear distance between the lateral malleolar apex (F11) and the lateralmost fibular contour point (F12).
F12–F7 (mm)	Lateral Fibular Height	The linear distance between the lateralmost fibular contour point (F12) and the proximolateral fibular corner (F7).

2.2. Statistical Analysis

Continuous variables were initially evaluated for normality using the Shapiro–Wilk test and visual inspection of Q–Q plots. Homogeneity of variances was assessed via Levene’s test. Accordingly, independent samples *t*-test was utilized for normally distributed parameters, while the Mann–Whitney U test was applied to non-normally distributed metrics. To control the family-wise error rate associated with conducting twelve simultaneous univariate comparisons, a Bonferroni correction was applied, adjusting the statistical significance threshold to $\alpha = 0.05/12 = 0.0041$.

The selection of 20 cases (10% of the total cohort) for interobserver reliability screening was based on well-established conventions in anthropometric and forensic imaging literature. We have now explicitly reported the 95% confidence intervals (CI) for all ICC values in the text. For eleven of the twelve evaluated parameters, the reliability coefficients remained highly robust, with the lower bounds of the 95% CIs comfortably exceeding 0.78, indicating good-to-excellent interobserver agreement. The single exception was the F12–F7 metric, which exhibited a wider confidence interval (95% CI: 0.40–0.91), reflecting localized geometric uncertainty in landmarking the continuous cortical borders of the distal fibular contour.

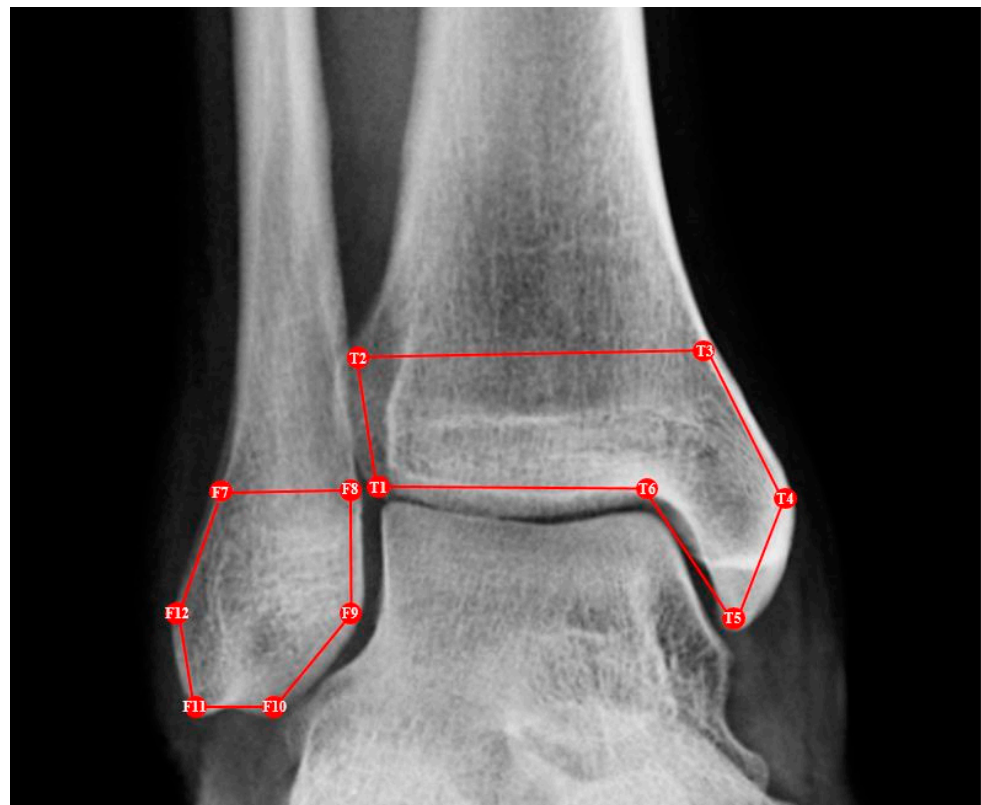


Figure 1. Consecutive anatomical landmarks utilized for the measurement of linear distances across the distal tibia and fibula. T1–T2 (mm): The linear distance between the lower lateral point closest to the incisura fibularis (T1) and the upper lateral point closest to the incisura fibularis (T2). T2–T3 (mm): The linear distance between T2 and the point where an orthogonal horizontal line drawn from T2 intersects the medial aspect of the tibia (T3). T3–T4 (mm): The linear distance between T3 and the point on the medial margin of the tibia at the maximum perpendicular distance from the horizontal reference line extending from T1 (T4). T4–T5 (mm): The linear distance between T4 and the lowermost tip of the medial malleolus (T5). T5–T6 (mm): The linear distance between the lowermost tip of the medial malleolus (T5) and the most concave point of the distal tibial articular surface (T6). T6–T1 (mm): The linear distance between the most concave point of the distal tibial articular surface (T6) and the lower

lateral point closest to the incisura fibularis (T1). F7–F8 (mm): The linear distance between the upper lateral point nearest the tibial notch (F7) and the upper medial point nearest the tibial notch (F8) on the distal fibula. F8–F9 (mm): The linear distance between F8 and the medial tip at the widest point of the fibular shaft relative to a vertical reference line drawn from F8 (F9). F9–F10 (mm): The linear distance between F9 and the lower medial tip of the lateral malleolus farthest from its lateralmost extension (F10). F10–F11 (mm): The linear distance between F10 and the lowermost tip of the lateral malleolus (F11). F11–F12 (mm): The linear distance between F11 and the lateralmost contour point of the fibula, determined at the maximum perpendicular distance from the horizontal reference axis drawn through F9 (F12). F12–F7 (mm): The linear distance between the lateralmost contour point of the fibula (F12) and the upper lateral point nearest the tibial notch (F7).

All data preprocessing, statistical analyses, and machine learning workflows were developed and executed using the Python software ecosystem (version 3.10.12). Prior to model training, all continuous features were standardized using StandardScaler (zero mean, unit variance) applied within each pipeline fold to prevent data leakage between training and evaluation stages. The dataset was divided into training and test sets using a stratified 80/20 split (`random_state = 42`), ensuring equal class proportions in both partitions. To provide stable performance estimates and prevent over-reliance on a single random split, a 5-fold stratified cross-validation protocol (`StratifiedKFold`, `n_splits = 5`, `shuffle = True`, `random_state = 42`) was implemented across the full training set ($n = 160$) prior to final test-set evaluation. Hyperparameter optimization for the logistic regression model was conducted via `GridSearchCV`, evaluating regularization strengths $C \in \{0.01, 0.1, 1, 10, 100\}$ with L2 penalty and `lbfgs` solver, scored by area under the ROC curve. The optimal configuration ($C = 0.1$, L2 regularization) was identified with a cross-validation AUC of 0.846 and was applied to all final evaluations. For the support vector machine, hyperparameters were tuned over $C \in \{0.1, 1, 10\}$ and $\gamma \in \{\text{scale}, 0.01, 0.1, 1\}$, identifying $C = 1$ and $\gamma = 0.1$ as optimal. To justify the study's cohort scale, a post hoc statistical power analysis was performed. Given the sample size of 200 subjects (100 males, 100 females) and a 12-predictor configuration, and taking into account the large to very large effect sizes exhibited by the primary talocrural metrics (ranging from Cohen's $d = 0.84$ to 1.65, as shown in Table 3), the statistical power of the logistic regression classifier was calculated to be well above the standard 80% threshold (achieved power > 0.95 at $\alpha = 0.05$). This confirms that the sample size is statistically robust and sufficient for detecting dimorphic skeletal variations without underpowering the classifier.

Table 3. Descriptive analytics and corresponding significance levels for all evaluated parameters.

Parameter	Males (n = 100) (Mean $\pm$ SD)	Females (n = 100) (Mean $\pm$ SD)	Uncorrected <i>p</i> -Value	Bonferroni-Adjusted <i>p</i> -Value	Cohen's d (Effect Size)
Age (years)	38.11 $\pm$ 12.30	40.01 $\pm$ 13.43	0.089	1.000 (Not Signif.)	−0.15 (Negligible)
T1–T2 (mm)	15.82 $\pm$ 2.11	12.72 $\pm$ 1.84	<0.001	<0.001 (Significant)	1.56 (Very Large)
T2–T3 (mm)	54.35 $\pm$ 4.12	47.69 $\pm$ 3.95	<0.001	<0.001 (Significant)	1.65 (Very Large)
T3–T4 (mm)	16.98 $\pm$ 2.25	13.70 $\pm$ 2.01	<0.001	<0.001 (Significant)	1.54 (Very Large)
T4–T5 (mm)	15.11 $\pm$ 2.80	13.21 $\pm$ 3.10	0.007	0.084 (Not Signif.)	0.64 (Medium)
T5–T6 (mm)	18.25 $\pm$ 2.05	15.61 $\pm$ 1.98	<0.001	<0.001 (Significant)	1.31 (Large)
T6–T1 (mm)	36.88 $\pm$ 3.42	31.60 $\pm$ 3.25	<0.001	<0.001 (Significant)	1.58 (Very Large)
F7–F8 (mm)	18.95 $\pm$ 2.10	15.85 $\pm$ 1.95	<0.001	<0.001 (Significant)	1.53 (Very Large)
F8–F9 (mm)	14.12 $\pm$ 2.30	12.10 $\pm$ 2.15	0.001	0.012 (Not Signif.)	0.91 (Large)
F9–F10 (mm)	12.58 $\pm$ 2.45	11.84 $\pm$ 2.38	0.046	0.552 (Not Signif.)	0.31 (Small)
F10–F11 (mm)	9.85 $\pm$ 1.75	8.43 $\pm$ 1.62	<0.001	<0.001 (Significant)	0.84 (Large)

Table 3. *Cont.*

Parameter	Males (n = 100) (Mean $\pm$ SD)	Females (n = 100) (Mean $\pm$ SD)	Uncorrected <i>p</i> -Value	Bonferroni-Adjusted <i>p</i> -Value	Cohen's d (Effect Size)
F11–F12 (mm)	12.18 $\pm$ 2.60	11.12 $\pm$ 2.52	0.038	0.456 (Not Signif.)	0.41 (Small)
F12–F7 (mm)	15.08 $\pm$ 2.95	13.42 $\pm$ 2.80	0.001	0.012 (Not Signif.)	0.58 (Medium)

Note: T1–T2 (mm): The linear distance between the lower lateral point closest to the incisura fibularis (T1) and the upper lateral point closest to the incisura fibularis (T2). T2–T3 (mm): The linear distance between T2 and the point where an orthogonal horizontal line drawn from T2 intersects the medial aspect of the tibia (T3). T3–T4 (mm): The linear distance between T3 and the point on the medial margin of the tibia at the maximum perpendicular distance from the horizontal reference line extending from T1 (T4). T4–T5 (mm): The linear distance between T4 and the lowermost tip of the medial malleolus (T5). T5–T6 (mm): The linear distance between the lowermost tip of the medial malleolus (T5) and the most concave point of the distal tibial articular surface (T6). T6–T1 (mm): The linear distance between the most concave point of the distal tibial articular surface (T6) and the lower lateral point closest to the incisura fibularis (T1). F7–F8 (mm): The linear distance between the upper lateral point nearest the tibial notch (F7) and the upper medial point nearest the tibial notch (F8) on the distal fibula. F8–F9 (mm): The linear distance between F8 and the medial tip at the widest point of the fibular shaft relative to a vertical reference line drawn from F8 (F9). F9–F10 (mm): The linear distance between F9 and the lower medial tip of the lateral malleolus farthest from its lateralmost extension (F10). F10–F11 (mm): The linear distance between F10 and the lowermost tip of the lateral malleolus (F11). F11–F12 (mm): The linear distance between F11 and the lateralmost contour point of the fibula, determined at the maximum perpendicular distance from the horizontal reference axis drawn through F9 (F12). F12–F7 (mm): The linear distance between the lateralmost contour point of the fibula (F12) and the upper lateral point nearest the tibial notch (F7).

3. Results

Descriptive statistics and *p*-values for all variables in the study group are presented in Table 3. The measurement values for T1–T2, T2–T3, T3–T4, T5–T6, T6–T1, F7–F8, and F10–F11 emerged as strong predictors in the gender distinction ($p < 0.001$).

3.1. Interobserver Agreement

All 12 measurements demonstrated good to excellent agreement based on their point estimates. Six parameters showed excellent agreement (ICC > 0.90), with the highest being T2–T3 (ICC = 0.95; 95% CI: 0.88–0.98). The remaining six parameters showed good agreement, with the lowest being F12–F7 (ICC = 0.77; 95% CI: 0.40–0.91) (Figure 2.). The wide 95% confidence interval and the lower confidence bound (0.40) observed for the F12–F7 metric indicate a degree of interobserver uncertainty. This variation is primarily attributed to the anatomical nature of the F7 (proximolateral fibular corner) and F12 (lateral-most fibular contour point) landmarks, which are situated along smooth, continuous cortical curves rather than prominent, sharp osseous projections. Consequently, placing the digital caliper along these curving margins introduces minor spatial variations. Additionally, because the mean distance of F12–F7 is relatively small (13.42 mm), these minor discrepancies represent a larger relative percentage of the measurement, depressing the ICC value despite a low mean absolute difference (1.03 mm).

To complement the ICC results, mean absolute differences (MAD) were calculated for each parameter. These values ranged from 0.53 mm (T4–T5) to 1.25 mm (T6–T1), all falling below the 2 mm clinical acceptability threshold. Relative percentage differences also remained low (mean absolute difference < 8% across all parameters), supporting the reproducibility of the measurements (Table 4). Based on these findings, all 12 parameters were retained for model development.

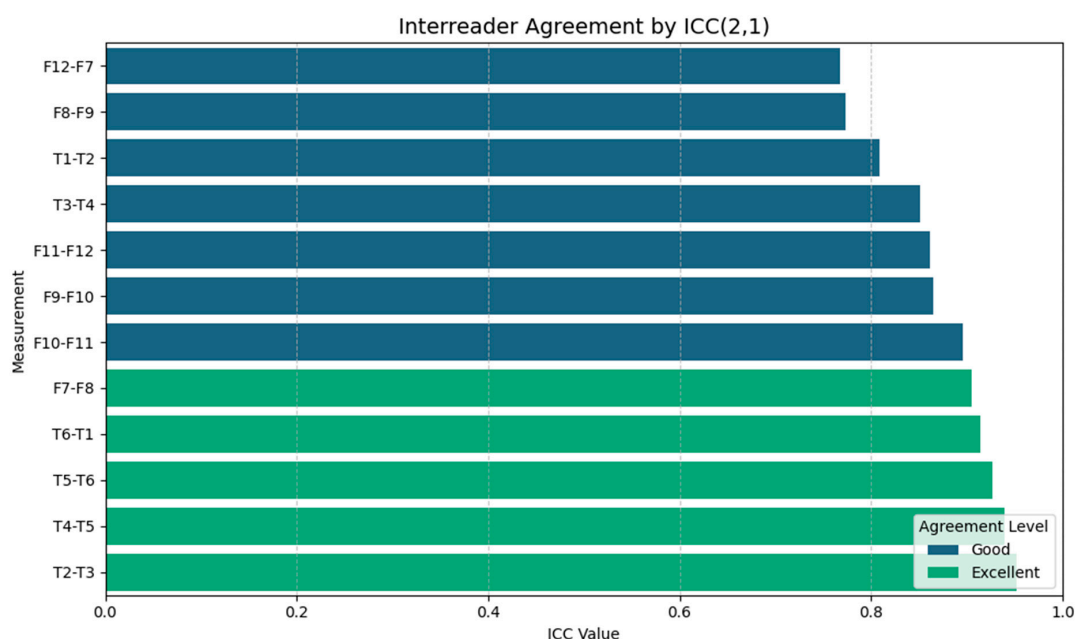


Figure 2. Intraclass correlation coefficients (ICCs) for interobserver agreement across all 12 measurements. Six parameters showed excellent agreement (ICC > 0.90); all others demonstrated good agreement.

Table 4. Interobserver Reliability and Agreement Assessment.

Measurement	Mean Absolute Difference (mm)	Mean Value of the Reliability Subsample (n = 20)	Relative Difference (%)
T1–T2	0.58	11.93	4.83
T2–T3	0.99	50.01	1.98
T3–T4	0.93	13.26	6.98
T4–T5	0.53	12.15	4.38
T5–T6	0.77	15.22	5.05
T6–T1	1.25	32.53	3.84
F7–F8	0.67	16.53	4.07
F8–F9	0.81	12.74	6.32
F9–F10	0.63	11.76	5.38
F10–F11	0.56	8.58	6.49
F11–F12	0.90	11.48	7.85
F12–F7	1.03	13.42	7.65

Note: Interobserver agreement assessment was supplemented by mean absolute differences (MAD) and relative differences for each anatomical measurement. “mean_abs_diff” represents the average absolute difference (in mm) between the two readers; “mean_val” denotes the average of both readers’ measurements; and “abs_diff_” reflects the relative difference as a percentage of the mean. All absolute differences were below 1.5 mm, and relative differences remained under 8%, supporting the reproducibility of the manual measurements.

3.2. Data Preprocessing and Model Development

To ensure robust validation and prevent performance estimation instability arising from a single random split, a 5-fold stratified cross-validation protocol was implemented across the entire training set (n = 160). Hyperparameter optimization confirmed L2 regularization with C = 0.1 as optimal for logistic regression (CV AUC = 0.846), and this configuration was used for all final test-set evaluations reported below.

3.3. Model Evaluation and Performance

Three machine learning classifiers—logistic regression, support vector machine (SVM), and random forest—were initially trained and validated. While the optimized logistic regression model achieved the highest classification accuracy of 80.0% on the independent test set, outperforming the random forest (72.5%) and support vector machine (70.0%)

classifiers (Table 5), it was selected as the final model due to its superior combination of discriminative depth (AUC = 0.870), perfectly balanced class performance (F1-score = 0.80), and direct clinical interpretability through its regression coefficients. On the independent evaluation pipeline, the optimized logistic regression model achieved a mean classification accuracy of 80.0% and an area under the ROC curve (AUC) of 0.870 (95% CI: 0.765–0.975), which provides a highly satisfactory binary discriminatory capacity for forensic screening (Figure 3). The classification performance appeared perfectly balanced between the sexes, demonstrating identical female-specific and male-specific F1-scores of 0.80 (Figure 4). The confusion matrix (Figure 5) illustrates these classification outcomes, showing that the model correctly classified 16 out of 20 females (80.0% sensitivity) and 16 out of 20 males (80.0% specificity), resulting in a reduced rate of only 8 misclassified cases out of 40 overall. To further elucidate the anatomical drivers of this decision boundary, the relative predictive contribution of each talocrural measurement was extracted and analyzed using the absolute values of the final model's regression coefficients (Figure 6).

Table 5. Cross-validation and test-set classification performance across machine learning models.

Model	CV Accuracy	CV F1	CV AUC	Test Accuracy	Test F1	Test AUC
Logistic Regression	0.794	0.794	0.858	0.800	0.800	0.870
SVM	0.806	0.798	0.859	0.700	0.684	0.818
Random Forest	0.800	0.793	0.866	0.725	0.703	0.846

CV = mean across 5 stratified folds on training set (n = 160). Test = independent hold-out set (n = 40). All models evaluated with the same random seed (42) and stratified split.

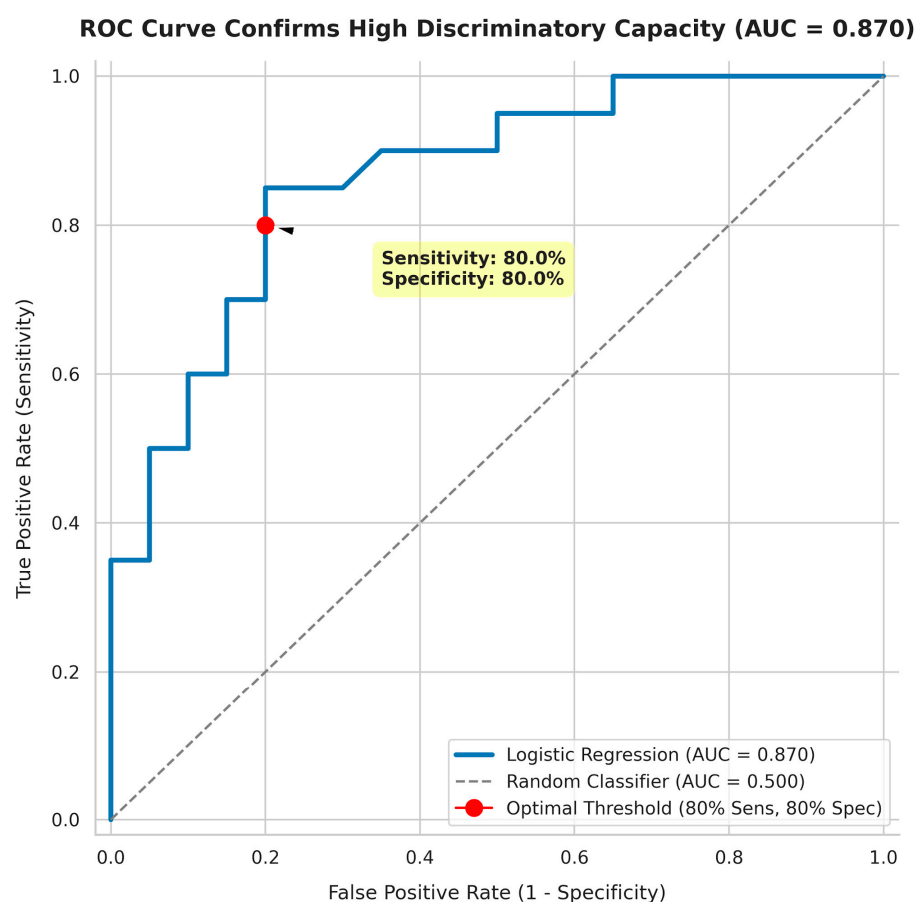


Figure 3. The solid line depicts sensitivity as a function of 1 – specificity on the independent test set (n = 40). The red dot and arrow indicate the optimal threshold (80.0% sensitivity and 80.0% specificity). The area under the curve (AUC) is 0.870.

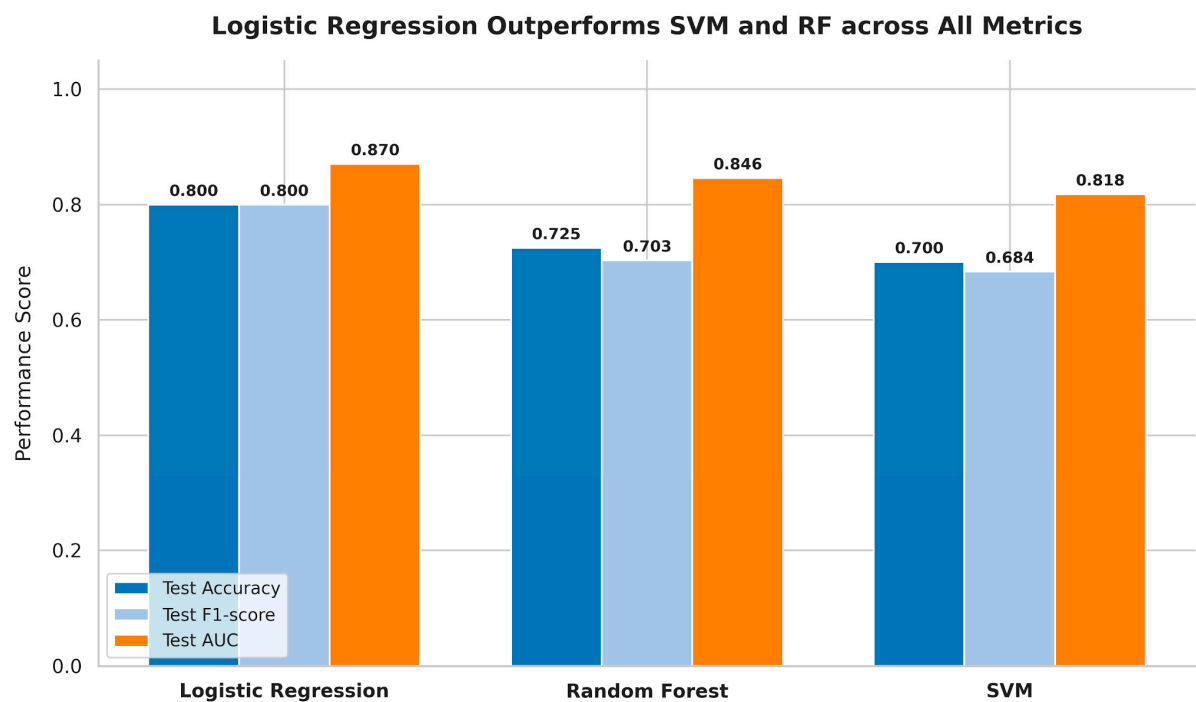


Figure 4. Comparison of classification performance across three machine learning models. Logistic regression achieved the highest area under the curve (AUC = 0.870) and F1-score, while support vector machine (SVM) had the lowest performance. All models were evaluated on the same independent test set ($n = 40$).

Optimized Logistic Regression Achieved Balanced 80.0% Accuracy (8 Misclassifications Out of 40 Test Cases)

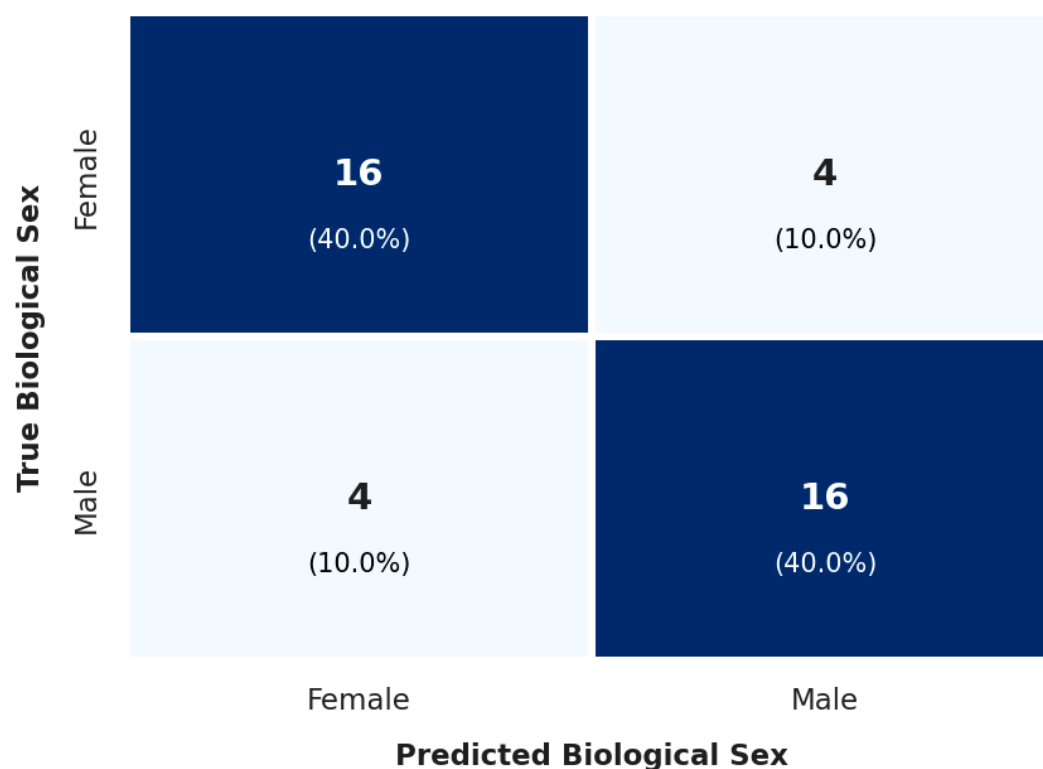


Figure 5. Confusion matrix of the logistic regression model on the test set. The model correctly classified 16/20 females and 16/20 males (80.0% accuracy).

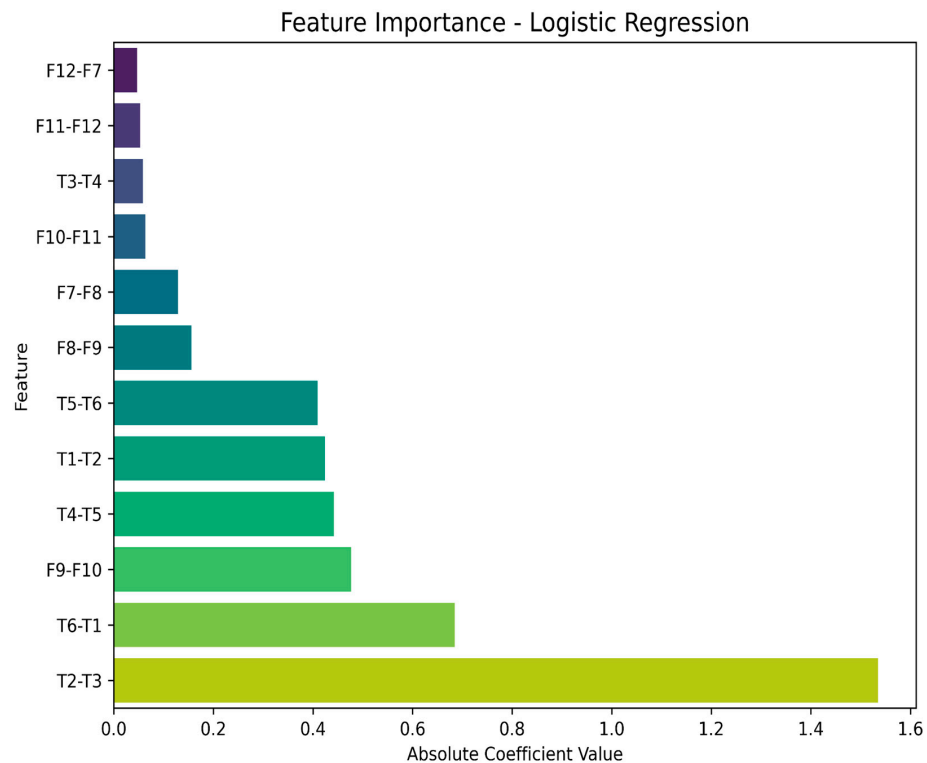


Figure 6. Feature importance plot based on the absolute value of logistic regression coefficients. After the Bonferroni correction, T2–T3 and T6–T1 stood out as the main predictors. Even though F9–F10 lost its individual significance ($p = 0.552$), its high coefficient shows that it plays a vital role in the model when combined with the other variables.

3.4. Error and Misclassification Analysis

A detailed post hoc analysis of the misclassified test cases ($n = 8$) revealed that errors predominantly occurred in female subjects exhibiting larger-than-average skeletal frameworks and male subjects with gracile morphometric patterns. The majority of false positives and false negatives were clustered at the operational boundaries of features T2–T3 and T6–T1, where overlapping phenotypic dimensions blurred binary boundaries. These cases likely represent individuals at the tails of the sex-specific size distributions, where the linear decision boundary of logistic regression is unable to further discriminate without additional predictor variables. Patient-level retrospective data lacked longitudinal parameters (e.g., body mass index, exact occupational loading, or sports history) that may have contributed to these atypical morphometric profiles, precluding definitive biomechanical attribution of the misclassification errors.

4. Discussion

This study demonstrates the feasibility of utilizing a targeted, interconnected linear landmark network across the distal tibia and fibula to perform biological sex estimation. Specifically, we evaluated twelve sequential linear parameters—comprising T1–T2, T2–T3, T3–T4, T4–T5, T5–T6, and T6–T1 on the tibia, and F7–F8, F8–F9, F9–F10, F10–F11, F11–F12, and F12–F7 on the fibula—derived from standard anteroposterior (AP) ankle radiographs. By modeling these metrics through an optimized logistic regression classifier, we successfully captured sex-specific morphological variations without requiring expensive three-dimensional reconstructions or resource-intensive shape modeling. This foundational framework bridges the gap between traditional manual osteology and modern computational workflows, demonstrating how routine clinical imaging can be successfully

repurposed for auxiliary forensic screening when primary diagnostic indicators, such as the pelvis or skull, are compromised.

Feature importance analysis based on the absolute values of logistic regression coefficients identified T2–T3 (distal tibial metaphyseal width, coefficient = 0.784) and T6–T1 (lateral plafond width, coefficient = 0.531) as the strongest contributors, followed by T5–T6 (medial plafond width, coefficient = 0.299). T2–T3 reflects the transverse breadth of the distal tibial metaphysis, a dimension directly shaped by androgenic periosteal expansion during puberty—males in our cohort exhibited a mean value approximately 6.66 mm wider than females (54.35 vs. 47.69 mm, Cohen's $d = 1.65$, Table 3). T6–T1 captures the articular load-bearing surface breadth of the plafond, which is biomechanically determined by the approximately 83% axial load transmitted through the tibiotalar joint during weight-bearing activities. Under Wolff's Law of bone functional adaptation, the increased body mass and dynamic muscular forces in males drive appositional bone deposition at these primary joint contact areas to distribute mechanical stress [15]. Notably, F9–F10 (lateral malleolar joint interface) and T4–T5 (medial malleolar length) carried negative coefficients, indicating that larger values in these dimensions are associated with female classification—a pattern consistent with the relatively wider lateral malleolar articular surfaces reported in females in geometric morphometric analyses of fibular dry bone. This anatomical configuration likely relates to differences in joint kinematics, as females exhibit greater subtalar joint laxity and different rotational kinematics during stance phases compared to males [18,21]. These anatomical interpretations are offered as mechanobiological hypotheses grounded in the observed coefficient directions and should be regarded as preliminary; they require prospective validation with dynamic loading and microarchitectural data before definitive conclusions can be drawn.

Biomechanical studies have shown that approximately 83% of the total load on the ankle is transmitted through the tibiotalar joint, while 17% is transmitted via the fibula, with this load increasing particularly during the dorsiflexion phase [15]. These variations in mechanical loading are key determinants in the formation of sexual dimorphism in distal fibular morphology. Indeed, the literature has shown that in three-dimensional (3D) geometric morphometric analyses of dry bones, women have narrower fibular articular surfaces compared to men, and that fibular morphology is a strong predictor of sex determination with high accuracy rates ranging from 80% to 93% [18]. In addition, fluoroscopy-based kinematic analyses confirm a smooth dynamic differentiation between the sexes in the ranges of motion for dorsiflexion, plantar flexion, and subtalar rotation [21]. On the other hand, it is known that morphometric variations in anteroposterior radiographic measurements are closely related to variations in the ligament complex that provides stability to the talocrural joint and to potential pathologies [20]. In this context, the high-accuracy quantitative datasets generated in our study and augmented by machine learning algorithms clearly demonstrate the robust predictive capacity of both the static morphometric architecture and functional biomechanics of the talocrural joint in sexual dimorphism analyses.

In the talocrural joint—one of the body's primary load-bearing and weight-distribution centers—kinetic forces are transmitted to the talus via the tibia and fibula, and from there to the anterior and posterior segments of the foot [20,22]. Asymmetric mechanical stresses acting on this region of the locomotor system induce population-based patterns of sexual dimorphism in the morphometric parameters of the bones forming the joint. Indeed, discriminant analyses conducted on the architecture of the talus and calcaneus in the Korean population reveal that all morphometric variables show statistically significant differentiation between the sexes and achieve high predictive accuracy rates ranging from 65% to 89.4% [23,24]. The classification performance achieved herein (80.0% accuracy, 0.870 AUC) is not a gold standard to replace primary forensic identification markers

like DNA or pelvic/cranial morphology [3,4,16]. Instead, the proposed talocrural metric framework is designed as a rapid and cost-effective screening or exclusion tool for scenarios where skeletal remains are severely compromised due to mass disasters, or where the pelvis and cranium are unavailable [1,14]. Compared to single-bone morphometric studies (e.g., with accuracy rates ranging from 65% to 89.4% for tarsal bones), our model provides a reproducible and realistic baseline using standard anteroposterior radiographs, without the need for complex 3D reconstructions [23,24].

Sex-specific dimorphism significantly influences the structural configuration and anatomical layout of the talocrural assembly [10,16]. In a study conducted by Uzuner et al. [16], anteroposterior (A-P) talocrural joint radiographs from 274 anatomically normal individuals aged between 10 and 85 years were evaluated. Following the assessment of linear metrics and angular parameters, the authors noted that males consistently exhibited higher values than females [16]. Consistent with this trend, our findings demonstrated significantly higher values in male subjects compared to their female counterparts. Moreover, there is a notable paucity of published sex classification studies that specifically leverage anteroposterior ankle radiographs to evaluate the talocrural joint [12]. Although Ray et al. [12] achieved 92% accuracy, their study may have been at risk of overfitting, or they may have used a different set of algorithms. The superior test-set performance of logistic regression relative to SVM and random forest (Table 5) can be explained through the lens of the bias-variance trade-off. With $n = 200$ subjects and 12 predictors, more complex models such as SVM with an RBF kernel and random forest face an elevated variance relative to the available sample size, making them susceptible to overfitting. The large to very large univariate effect sizes observed for the primary talocrural metrics (Cohen's $d = 0.84$ – 1.65 , Table 3) suggest an approximately linear separability in the feature space, which is inherently well-suited to logistic regression's linear decision boundary. L2 regularization further stabilizes coefficient estimation under mild multicollinearity among correlated anatomical landmark distances, explaining why logistic regression achieves superior generalization on the independent test set despite comparable or slightly lower cross-validation accuracy. However, since skeletal dimorphism exhibits profound population-specific variations driven by genetic, environmental, and socioeconomic differences, these metrics must be interpreted as specific to the evaluated regional Turkish cohort until validated on external, heterogeneous population. Consistently, in the present study, two parameters pertaining to the inferior surface of the tibia (the T1–T6 and T2–T3 distances) were found to differ significantly between male and female subjects. Both the distal portions and the joint-forming articular facets of the tibiofibular complex underwent a rigorous morphometric characterization in this investigation. Additionally, our findings indicate that integrating distal fibular metrics—specifically the F9–F10 parameter (Lateral Malleolar Joint Interface)—contributes valuable morphological data to the overall model, though its interpretation must be carefully balanced. In our univariate analysis (Table 3), the F9–F10 distance demonstrated a small independent effect size (Cohen's $d = 0.31$) and failed to maintain significance after the conservative family-wise error controls of the Bonferroni correction (adjusted $p = 0.552$). Rather than analyzing features in isolation as univariate tests do, regularized multivariate approaches, such as our Logistic Regression framework, are designed to leverage joint geometric covariance and capture complementary, non-redundant spatial relationships among landmarks. Through this mechanism, F9–F10 appears to operate alongside distal tibial variables to subtly refine the classifier's overall decision boundary. From a biomechanical perspective, these observations can only serve to formulate tentative hypotheses regarding the sexual dimorphism of the talocrural joint surfaces, rather than definitive conclusions. We hypothesize that potential variations in internal trabecular architectures, theoretically driven by different gravitational and compressive loads, might influence force distribution

across the ankle complex. However, because our study design did not directly measure dynamic loading, gait kinematics, or internal bone microarchitecture, these prospective functional adaptations and their role in sex-specific load transmission remain entirely speculative [25]. They represent potential mechanobiological explanations that must be rigorously tested in future prospective investigations rather than established outcomes of the current model.

The assessment of skeletal sexual dimorphism continues to be a cornerstone of anthropological research, though traditional osteometric methodologies are progressively giving way to sophisticated artificial intelligence frameworks and machine learning algorithms. Moreover, recent studies increasingly move away from single-bone and simple linear assessments towards multiple linear models and analyses incorporating two or three bones in combination. Accordingly, the detailed morphometric model used in our study offers morphometric evidence that can be incorporated into these advancing technological approaches.

Given its minimal imaging requisites and the high reproducibility demonstrated by inter-observer agreement, this approach holds substantial promise for applications in forensic anthropology, population-based studies, and potentially automated clinical workflows. Future research could extend the anatomical scope or integrate automated landmark detection to further optimize scalability and precision.

Limitations

Several limitations must be acknowledged. First, this retrospective, single-center study analyzed a modest cohort of 200 subjects; while statistically sufficient for initial classifier development, this scale limits broader machine learning stability and generalizability beyond the studied Turkish population. Second, although the age distribution was homogeneous between sexes, potential non-linear interactions of age-related cortical thinning were not modeled. Third, the dataset lacked patient-specific metadata—including height, weight (body mass), BMI, and occupational or athletic loading profiles—restricting definitive biomechanical conclusions regarding active bone remodeling. Fourth, despite standard weight-bearing mortise positioning, deriving metrics from 2D planar radiographs introduces minor risks of projection, magnification, and rotational errors compared to 3D tomographic imaging. Fifth, while our final model achieved a promising test AUC of 0.870, its relatively wide 95% confidence interval (0.765–0.975) reflects the statistical variance inherent to a 40-case test fold, highlighting the necessity of larger, multi-center cohorts to stabilize model performance metrics. Regarding observer dynamics, although interobserver reliability was high and confirmed the robustness of our landmarks, intraobserver reliability was not formally evaluated in this investigation. Furthermore, the manual landmarking protocol was restricted to two experienced radiologists, representing expert dependency and a scalability bottleneck for high-volume forensic triage. Notably, the F12–F7 metric (Lateral Fibular Height) exhibited a wide confidence interval due to localized landmarking uncertainty. Because F7 and F12 lie along smooth, continuous cortical curves lacking sharp osseous projections, digital caliper placement is highly sensitive to minor inter-rater drift. While this absolute discrepancy was minimal (mean absolute difference = 1.03 mm), its relative impact was magnified by the small physical scale of the F12–F7 distance. Additionally, regarding practical deployment, the measurement protocol described herein was calibrated to the DICOM pixel spacing metadata of a single institutional PACS platform. Inter-institutional magnification differences across PACS platforms with different calibration workflows, as well as positioning variability in non-standardized clinical settings (such as minor deviations from the 15–20° internal rotation mortise standard), may introduce systematic measurement drift relative to our acquisition protocol. We therefore recommend that any future clinical or forensic implementation include a site-specific calibration vali-

dation step—applying the measurement protocol to a local reference cohort with known sex—before applying the logistic regression coefficients reported here to an independent setting. Consequently, future multi-center studies must focus on transitioning to automated or semi-automated landmark detection algorithms to eliminate observer subjectivity and minimize human error.

5. Conclusions

This study presents a validated, population-specific 12-landmark cortical morphometric framework for the Turkish talocrural joint, showing that an optimized logistic regression classifier trained on standard anteroposterior ankle radiographs can serve as a rapid, cost-effective forensic sex screening tool when primary skeletal indicators such as the pelvis and skull are unavailable or compromised, thereby demonstrating a robust and stable auxiliary capacity for sex classification. By achieving 80.0% test-set accuracy and an AUC of 0.870 using only routine clinical imaging and a reproducible set of linear measurements, this framework bridges the gap between traditional manual osteometry and modern computational forensic workflows. Ultimately, this research establishes that conventional ankle radiographs carry clear dimorphic signatures when analyzed using a simplified network of linear landmarks. Instead of chasing complex, hard-to-integrate classification algorithms, we focused on proving the practical viability of an optimized logistic regression classifier trained on a minimal and highly reproducible set of distal tibial and fibular measurements. This provides forensic experts with a swift, cost-effective backup screening tool when primary skeletal indicators like the pelvis and skull are damaged or absent.

Future research directions should include:

- **Automated Landmark Detection:** The development of deep learning-based automatic landmark detection algorithms to eliminate manual measurement subjectivity and overcome the scalability bottleneck identified in the F12–F7 metric.
- **Patient-Specific Metadata Integration:** Prospective incorporation of patient-specific metadata—including BMI, occupational loading profiles, and age-stratified cohorts—to model non-linear biomechanical confounders currently excluded from this framework.
- **Multimodal Fusion:** Multimodal fusion combining these 2D linear metrics with CT-derived trabecular density indices to determine whether internal bone microarchitecture contributes independent predictive value to talocrural sex classification.

Author Contributions: Conceptualization, S.C. and M.K.; methodology, S.C., M.K. and M.D. (Mustafa Durmaz); formal analysis, M.D. (Mustafa Durmaz) and K.A.; writing—original draft preparation, S.C. and M.K.; writing—review and editing, S.C. and M.K.; supervision, S.C. and M.D. (Menduh Dursun). All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee for Non-Interventional Research at Istanbul University Faculty of Medicine (Decision No. 3552799, dated 1 September 2025).

Informed Consent Statement: Patient consent was waived due to the retrospective design of the study and the use of anonymized radiographic data, as approved by the Ethics Committee for Non-Interventional Research at Istanbul University Faculty of Medicine (Decision No. 3552799, dated 1 September 2025).

Data Availability Statement: The datasets generated and analyzed during the current study are not publicly available due to institutional regulations and ethical restrictions governing patient confidentiality. However, de-identified morphometric data may be obtained from the corresponding

author upon reasonable request, subject to prior approval from the Ethics Committee for Non-Interventional Research at Istanbul University Faculty of Medicine.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Blau, S.; Briggs, C.A. The role of forensic anthropology in Disaster Victim Identification (DVI). *Forensic Sci. Int.* **2011**, *205*, 29–35. [[CrossRef](#)] [[PubMed](#)]
2. Duangchit, S.; Poodendaen, C.; Phetnui, P.; Tasu, P.; Boonthai, W.; Tangsrisakda, N.; Iamsaard, S. Sex determination and stature estimation using logistic and linear regression models: A population-specific study of tibia and fibula in Northeastern Thais. *Transl. Res. Anat.* **2025**, *40*, 100420. [[CrossRef](#)]
3. Krogman, W.M. The human skeleton in forensic medicine. I. *Postgrad. Med.* **1955**, *17*, A-48. [[CrossRef](#)] [[PubMed](#)]
4. Buikstra, J.E. *Standards for Data Collection from Human Skeletal Remains*; Arkansas Archaeological Survey: Fayetteville, AR, USA, 1994; Volume 44, p. 44.
5. Christensen, A.M.; Passalacqua, N.V.; Bartelink, E.J. *Forensic Anthropology: Current Methods and Practice*; Academic Press: San Diego, CA, USA, 2019.
6. Krishan, K.; Chatterjee, P.M.; Kanchan, T.; Kaur, S.; Baryah, N.; Singh, R. A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. *Forensic Sci. Int.* **2016**, *261*, 165.e1–165.e8. [[CrossRef](#)] [[PubMed](#)]
7. Spradley, M.K.; Jantz, R.L. Sex estimation in forensic anthropology: Skull versus postcranial elements. *J. Forensic Sci.* **2011**, *56*, 289–296. [[CrossRef](#)] [[PubMed](#)]
8. Ruff, C.; Holt, B.; Trinkaus, E. Who's afraid of the big bad Wolff?: "Wolff's law" and bone functional adaptation. *Am. J. Phys. Anthropol.* **2006**, *129*, 484–498. [[CrossRef](#)] [[PubMed](#)]
9. Pearson, O.M.; Lieberman, D.E. The aging of Wolff's "law": Ontogeny and responses to mechanical loading in cortical bone. *Am. J. Phys. Anthropol.* **2004**, *125*, 63–99. [[CrossRef](#)]
10. Sholts, S.B.; Wärmländer, S.K.; Flores, L.M.; Miller, K.W.; Walker, P.L. Variation in the measurement of cranial volume and surface area using 3D laser scanning technology. *J. Forensic Sci.* **2010**, *55*, 871–876. [[CrossRef](#)] [[PubMed](#)]
11. Ebert, L.C.; Thali, M.J.; Ross, S. Getting in touch—3D printing in forensic imaging. *Forensic Sci. Int.* **2011**, *211*, e1–e6. [[CrossRef](#)] [[PubMed](#)]
12. Ray, A.; Ray, G.; Kürtül, İ.; Şenol, G.T. Sex estimation from the variables of talocrural joint by using machine learning algorithms. *J. Forensic Leg. Med.* **2025**, *114*, 102912. [[CrossRef](#)] [[PubMed](#)]
13. El Morsi, D.A.; Gaballah, G.; Mahmoud, W.; Tawfik, A.I. Sex determination in Egyptian population from scapula by computed tomography. *J. Forensic Res.* **2017**, *8*, 383.
14. Yurek, J.W.; Gianakos, A.L.; Mulcahey, M.K. Ankle anatomy and biomechanics. In *The Female Athlete*; Springer: Cham, Switzerland, 2022; pp. 161–167.
15. Brockett, C.L.; Chapman, G.J. Biomechanics of the ankle. *Orthop. Trauma* **2016**, *30*, 232–238. [[CrossRef](#)] [[PubMed](#)]
16. Uzuner, M.B.; Sargon, M.F.; Sancak, T.; Demiryürek, D.; Geneci, F.; Ocak, M.; Bilecenoğlu, B. Radiographic Examination of Talocrural Joint Morphometry in Individuals with Normal Joint Anatomy. *Kafkas J. Med. Sci.* **2018**, *8*, 141–147. [[CrossRef](#)]
17. Callewaert, F.; Sinnesael, M.; Gielen, E.; Boonen, S.; Vanderschueren, D. Skeletal sexual dimorphism: Relative contribution of sex steroids, growth hormone-insulin-like growth factor-I (GH-IGF-I) and mechanical loading. *J. Endocrinol.* **2010**, *207*, 127–134. [[CrossRef](#)] [[PubMed](#)]
18. Pietrobelli, A.; Sorrentino, R.; Durante, S.; Marchi, D.; Benazzi, S.; Belcastro, M.G. Sexual dimorphism in the fibular extremities of Italians and South Africans of identified modern human skeletal collections: A geometric morphometric approach. *Biology* **2022**, *11*, 1079. [[CrossRef](#)] [[PubMed](#)]
19. Lazenby, R.A. Population variation in second metacarpal sexual size dimorphism. *Am. J. Phys. Anthropol.* **2002**, *118*, 378–384. [[CrossRef](#)] [[PubMed](#)]
20. Martis, R.S. A study of usefulness of X-Ray in diagnosis of ligamentary tear in talocrural joint. *Int. J. Radiol. Diagn. Imaging* **2020**, *3*, 43–45. [[CrossRef](#)]
21. Fukano, M.; Fukubayashi, T.; Banks, S.A. Sex differences in three-dimensional talocrural and subtalar joint kinematics during stance phase in healthy young adults. *Hum. Mov. Sci.* **2018**, *61*, 117–125. [[CrossRef](#)] [[PubMed](#)]
22. Patil, M.; Raza, S.; Ali, M.N. Anthropometric measurements of ankle mortise for evaluating mortise fracture reductions with an aim to develop contoured implants. *Al Ameen J. Med. Sci.* **2012**, *5*, 381–387.
23. Lee, U.Y.; Han, S.H.; Park, D.K.; Kim, Y.S.; Kim, D.I.; Chung, I.H.; Chun, M.H. Sex determination from the talus of Koreans by discriminant function analysis. *J. Forensic Sci.* **2012**, *57*, 166–171. [[PubMed](#)]

24. Kim, D.-I.; Kim, Y.-S.; Lee, U.-Y.; Han, S.-H. Sex determination from calcaneus in Korean using discriminant analysis. *Forensic Sci. Int.* **2013**, *228*, 177.e1–177.e7. [[CrossRef](#)] [[PubMed](#)]
25. Kivell, T.L. A review of trabecular bone functional adaptation: What have we learned from trabecular analyses in extant hominoids and what can we apply to fossils? *J. Anat.* **2016**, *228*, 569–594. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.