Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A comprehensive computational pipeline for foramen magnum shape analysis integrating elliptic Fourier transform, polygon approximation, and Procrustes heatmaps from CT images: application to forensic sex estimation

Yasin Etli

yasinetli@yyu.edu.tr

Van Yuzuncu Yil University

Erhan Kartal

Van Yuzuncu Yil University

Mahmut Asirdizer

Bahcesehir University

Yavuz Hekimoglu

Ankara City Hospital, Health Sciences University

Sıddık Keskin

Van Yuzuncu Yil University

Uğur Demir

Harran University

Alparslan Yavuz

Antalya Training and Research Hospital, Health Sciences University

Osman Celbiş

Alanya Alaaddin Keykubat University

Research Article

Keywords: Foramen magnum, Sexual dimorphism, Geometric morphometrics, Elliptic Fourier transform, Polygon approximation, Computed tomography, Forensic anthropology, Sex estimation, Procrustes analysis, Machine learning

Posted Date: March 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8928788/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

The foramen magnum (FM) exhibits well-documented size-based sexual dimorphism, yet whether its shape independently differs between sexes remains unresolved. Previous attempts to characterize FM shape have relied on subjective visual classification into categorical types, producing inconsistent and irreproducible results. This study aimed to apply a comprehensive suite of computational geometric methods to CT-derived FM contours in order to objectively quantify shape variation and evaluate its discriminative capacity for sex estimation.

Methods

FM boundaries were automatically extracted from axial CT images of 473 adults (234 males, 239 females) from an Eastern Turkish population. Seven complementary geometric analyses—classical morphometry, Fitzgibbon and genetic-algorithm-optimized ellipse fitting, systematic polygon approximation (3–13 sides), normalized elliptic Fourier transform (20 harmonics), radial distance profiling, and Procrustes-based point-density heatmap construction—yielded 231 quantitative features per specimen. Sex differences were assessed using Mann–Whitney U tests with Bonferroni correction and Cohen's d effect sizes. Classification performance was evaluated with linear discriminant analysis, support vector machines, and random forest under stratified 10-fold cross-validation. Within-sex morphological subtypes were explored using Gaussian mixture model clustering.

Results

Of 234 morphometric features, 92 (39.3%) showed significant sex differences after Bonferroni correction. Dimorphism was overwhelmingly size-driven: all six size measurements showed large effect sizes (Cohen's d = 1.13–1.28, perimeter strongest at d = 1.284), while only 17 of 146 pure shape features (11.6%) reached significance and the FM index showed no sex difference (d = 0.088). The highest classification accuracy was 80.9% ± 4.9% (sensitivity 76.4%, specificity 85.3%), achieved by forward-selected shape features with LDA, outperforming size-only classification (75.6%). A novel Procrustes heatmap-based method yielded near-chance accuracy (54.1%), confirming size dominance. Gaussian mixture model analysis revealed greater male morphological heterogeneity (25.2% atypical variants) compared to more homogeneous female FM morphology (71.5% single dominant subtype).

Conclusions

FM sexual dimorphism is predominantly size-mediated, but computationally derived shape features provide incremental discriminative value that exceeds what size alone can achieve. This study

introduces a fully automated, reproducible geometric morphometric pipeline that replaces subjective shape categorization with objective, continuous measurements. The framework is openly available and readily transferable to other skeletal structures amenable to CT-based contour analysis.

Background

The foramen magnum (FM), the largest opening in the skull base formed by the occipital bone, serves as the primary conduit between the cranial cavity and the vertebral canal, transmitting the medulla oblongata, vertebral arteries, spinal accessory nerves, and associated meningeal membranes [1, 2]. Beyond its well-established clinical significance in cranio-cervical surgery and neuropathology, the FM has attracted considerable research interest due to its reported morphological variation across individuals, populations, and between sexes [3–6]. A central question that has persisted in the anatomical and forensic literature is whether sex-related differences in the FM are limited to overall size—with males exhibiting larger dimensions—or whether genuine shape differences also exist between the sexes [7–9].

The dimensional sexual dimorphism of the FM has been well documented across diverse populations. Numerous studies using dry skulls, radiographs, and computed tomography (CT) have consistently demonstrated that the anteroposterior diameter, transverse diameter, circumference, and area of the FM are significantly greater in males than in females [5, 10–14]. A recent systematic review and meta-analysis confirmed these dimensional differences across multiple populations with sex estimation accuracies ranging from approximately 62% to 90% depending on the population and analytical method employed [15]. In the Turkish population specifically, Meral et al. reported a maximum multivariate accuracy of 75% [16], while Kartal et al. achieved up to 86.7% with discriminant analysis and 88.2% with artificial neural networks using dimensional measurements from 720 individuals [17]. These findings collectively establish that the FM exhibits moderate but consistent size-based sexual dimorphism that is primarily driven by the overall larger cranial dimensions in males.

In contrast, the question of whether the FM also exhibits sex-related differences in shape—independent of size—remains substantially less resolved. Morphological descriptions in the forensic and anatomical literature have long suggested that the male FM tends to display a "rougher" or more irregular contour, while the female FM is described as "more oval" and "more regular" [4, 18]. However, these descriptions are based on qualitative, observer-dependent assessments and have not been quantitatively validated. The most common approach to characterizing FM shape has been subjective visual classification into categorical types such as oval, round, egg-shaped, tetragonal, pentagonal, hexagonal, and irregular forms [18–21]. This approach suffers from several well-documented limitations. First, the shape categories themselves are not standardized across studies, leading to marked inconsistencies in the reported prevalence of each type; for instance, the FM has been reported as oval in 87.3% of crania in one study but in only 7.39% in another that utilized the same visual categorization approach [22]. Second, studies that have investigated whether these categorical shape types differ between sexes have generally concluded that FM shape is not a reliable sex discriminator [7, 8], with the FM index

(anteroposterior/transverse diameter ratio) showing no significant sex difference in several populations [6, 7, 23]. Zdilla et al. critically observed that previous shape studies utilized "ambiguous and subjective descriptors" that are "difficult to reproduce" and called for more reproducible methods of characterizing FM morphology [22].

The reliance on subjective categorical classification or simple ratios to characterize a complex, irregularly bounded anatomical contour fundamentally limits the ability to detect subtle but potentially meaningful shape differences. A two-dimensional closed contour such as the FM boundary contains considerably more morphological information than can be captured by two perpendicular diameters or a single aspect ratio [24, 25]. Modern computational imaging and geometric analysis methods offer the potential to extract this richer shape information objectively. Elliptic Fourier analysis (EFA), introduced by Kuhl and Giardina [24], enables the mathematical decomposition of any closed contour into a series of harmonic ellipses, capturing both global shape properties and fine local boundary details in a compact numerical representation. This approach has been successfully applied to various anatomical contours, including the FM of the tropical raccoon, where Samuel and Bhagwat demonstrated that higher harmonics revealed fine morphological details relevant to sexual dimorphism that were invisible to conventional measurements [26]. Hayashizaki et al. applied Fourier analysis to pelvic contours from postmortem CT images and showed that frequency-domain shape descriptors could discriminate sex with greater objectivity than traditional morphometric methods [27]. Similarly, Gonzalez et al. utilized semi-landmark-based geometric morphometrics on the greater sciatic notch and ischiopubic complex to achieve over 90% sex estimation accuracy from shape variables alone [28]. In our own previous work, we applied advanced contour-based geometric analysis methods—including ellipse fitting—to pelvic CT images and demonstrated that computational shape features could objectively capture morphological variations that had previously been described only qualitatively [29]. These studies collectively suggest that computational geometric approaches, when applied to anatomical contour data extracted from CT images, can reveal shape information that is inaccessible to conventional linear measurements and subjective visual classification.

Despite the clear potential of these computational shape analysis methods, no study to date has applied a comprehensive suite of such techniques to the FM contour for the purpose of objectively quantifying and validating the morphological descriptions that have persisted in the anatomical literature. Specifically, it remains unknown whether computationally derived metrics—such as circularity, boundary roughness, contour regularity, left-right and antero-posterior asymmetry, and harmonic shape descriptors—differ significantly between male and female FM contours when size effects are controlled for. Furthermore, whether any such pure shape features, once objectively defined, can contribute to or improve upon size-based sex estimation has not been systematically evaluated. Resolving these questions is important not only for advancing evidence-based anatomical knowledge, but also for practical applications in forensic medicine and anthropology, where the ability to estimate sex from fragmentary cranial remains—particularly the skull base, which is resistant to taphonomic destruction—can be critical in identification cases [30, 31].

The present study aims to address these gaps through two integrated objectives. The first objective is to apply a comprehensive array of computational geometric methods—including Fitzgibbon ellipse fitting, genetic algorithm-optimized inner and outer bounding ellipses, systematic polygon approximation from 3-sided to 13-sided polygons, normalized elliptic Fourier transform with 20 harmonics, radial profile analysis, and point density heatmap construction—to FM boundary coordinates extracted from CT images of a large contemporary Turkish sample (n = 473), in order to objectively characterize and quantify morphological features that correspond to classical qualitative descriptions such as "oval," "round," "irregular," and "rough." The second objective is to evaluate the discriminative capacity of these objectively defined shape features for sex estimation, both independently and in combination with conventional size-based measurements, using multiple classification approaches including linear discriminant analysis, support vector machines, and random forest, with rigorous cross-validation; and to investigate whether unsupervised clustering reveals sex-specific morphological subtypes that may further elucidate the nature and extent of FM shape dimorphism.

Methods

Study design and ethical approval

This retrospective cross-sectional study utilized cranial computed tomography (CT) images originally acquired for a prior investigation of foramen magnum (FM) sexual dimorphism in an Eastern Turkish population [17]. The original study protocol received approval from the Non-Interventional Clinical Research Ethics Committee of Van Yuzuncu Yil University Faculty of Medicine (decision number: 02, dated: 19.02.2017) and was conducted in accordance with the Declaration of Helsinki and the principles of Good Clinical Practice. Since the present study represents a secondary analysis of the same anonymized imaging dataset using computational geometric methods, no additional ethical approval was required. Informed consent was waived due to the retrospective design and full anonymization of all data.

Study sample

The source dataset comprised 720 cranial CT examinations (360 males and 360 females) from adult patients who presented to Yüzüncü Yıl University Dursun Odabaş Medical Center between January and December 2016 for various clinical indications unrelated to craniocervical pathology [17]. The original sample was constructed using stratified sampling across six adult age groups (21–30, 31–40, 41–50, 51–60, 61–70, and ≥ 71 years) with equal sex distribution to control for potential age-related confounding effects. Exclusion criteria in the original study included skull trauma causing cranial deformities, congenital or acquired cranial bone diseases, any malignancy affecting the skull, and cases with missing or defective FM boundaries [17].

For the present study, all 720 CT examinations were re-evaluated to determine their suitability for automated contour extraction. Cases were included if the axial CT slice displaying the widest FM cross-section yielded a boundary amenable to automated edge detection—specifically, sufficient contrast at the bone–soft tissue interface and a complete, uninterrupted FM margin. Cases with excessive image noise, partial volume artifacts obscuring the FM border, or incomplete boundary delineation after edge detection were excluded. After this additional selection, 473 individuals (234 males and 239 females) met the criteria and constituted the final study sample for geometric morphometric analysis.

CT image acquisition

All CT examinations had been performed on a 16-detector multislice CT system (Somatom Emotion 16-slice, CT2012E; Siemens AG, Berlin and Munich, Germany). Imaging parameters included a tube voltage of 120 kV with automatic tube current modulation ranging between 80 and 120 mA, a section thickness of 1.2 mm, a reconstruction interval of 1.0 mm, a pitch factor of 0.8, a field of view of 220 mm, and a matrix size of 256 × 256. Data acquisition was performed in the caudocranial direction without gantry tilt. Three-dimensional volume-rendered reformations were generated using Syngo VE.52A software (Siemens Healthcare) to minimize potential measurement errors arising from asymmetry or obliquity that may occur on conventional axial images [17]. The spatial calibration factor (mm per pixel) was individually extracted from the DICOM metadata of each examination and ranged from 0.271 to 0.513 mm/pixel (mean: 0.371 mm/pixel), ensuring accurate conversion of all pixel-based coordinate measurements to absolute millimeter values.

Foramen magnum boundary extraction

For each specimen, the axial CT slice displaying the widest cross-section of the FM was identified. The FM boundary was extracted using automated edge detection applied to the bone-windowed CT image at this level. The resulting contour coordinates (sequential X, Y pixel positions along the FM margin) were exported as individual comma-separated value (CSV) files. Each raw contour comprised between 36 and 187 boundary points (mean: 85.5).

A key methodological advantage of this approach is that the FM boundary was derived through computational image processing rather than manual tracing. As a consequence, the contour extraction procedure is fully deterministic: processing the same CT image repeatedly yields identical boundary coordinates. This eliminates inter-observer and intra-observer variability, which are inherent limitations of manual measurement approaches and have been documented as a source of error in conventional FM morphometric studies [5, 8]. To ensure comparability across specimens with different numbers of raw boundary points, all contours were resampled to 200 equally spaced points using cubic spline interpolation along the cumulative arc length prior to subsequent analysis [32].

Computational geometric analyses

All computational analyses were implemented in Python 3.8 + using NumPy, SciPy, scikit-learn, and Matplotlib [33–36]. The complete analysis pipeline is publicly available as a single, self-contained script (FM_Comprehensive_Analysis.py) in the Zenodo repository accompanying this paper (DOI: 10.5281/zenodo.18716445), along with all anonymized boundary coordinate data, enabling full reproducibility.

Classical morphometric measurements

Eleven standard morphometric parameters were computed from each resampled contour: anteroposterior (AP) diameter (maximum extent along the Y-axis), transverse diameter (maximum extent along the X-axis), maximum and minimum diameters (derived from the convex hull), perimeter (sum of consecutive point-to-point Euclidean distances along the closed contour), area (computed using the shoelace formula [37]), FM index (AP/transverse diameter ratio), circularity ($4\pi A/P^2$, where A is area and P is perimeter), eccentricity (maximum/minimum diameter ratio), area-to-perimeter ratio (A/P), and compactness ($P/\sqrt{A}$). These parameters correspond to those commonly reported in the FM literature [5, 6, 8, 16, 17], thereby facilitating direct comparison with previously published dimensional data.

Ellipse fitting

Three ellipse models were fitted to each FM contour using distinct optimization strategies, together with a best-fit circle.

The Fitzgibbon direct least-squares ellipse was obtained by minimizing the algebraic distance from contour points to a general conic equation, with an ellipticity constraint ensuring a valid ellipse solution [38]. Initial estimates were derived from the centroid and coordinate covariance, followed by geometric refinement via the Nelder-Mead simplex algorithm [39]. For each fitted ellipse, the following were recorded: center coordinates (cx, cy), semi-major axis (a), semi-minor axis (b), rotation angle (θ), four distance metrics (mean, standard deviation, and maximum Euclidean distance from each contour point to its nearest point on the ellipse boundary, and RMSE), the fitted ellipse area, the corresponding FM area, and the area ratio between the two. This yielded 12 features per ellipse model.

The genetic algorithm (GA)-optimized inner bounding ellipse (the largest ellipse fully inscribed within the FM contour) and the GA-optimized outer bounding ellipse (the smallest ellipse fully circumscribing the FM contour) were estimated using a stochastic optimization approach. An initial population of 60 candidate ellipses (parameterized by center coordinates, semi-axes, and rotation angle) was evaluated over 150 generations, followed by local refinement using the Nelder-Mead algorithm. These bounding ellipses capture the deviation of the FM contour from an ideal elliptical shape: a large gap between inner and outer ellipses indicates a more irregular boundary. Each bounding ellipse model produced 12 features identical in structure to the Fitzgibbon model.

The best-fit circle was obtained via algebraic least-squares fitting [40]. Ten features were extracted: center coordinates, radius, four distance metrics (mean, standard deviation, maximum, RMSE), circle area, FM area, and area ratio.

In total, the three ellipse models and the circle model yielded 46 features.

Polygon approximation

To capture FM boundary irregularity at multiple geometric resolution levels, systematic polygon approximation was performed for polygons with 3 to 13 sides (11 polygon orders). For each polygon order n, the optimal n-sided polygon was determined by angular sectoring from the FM centroid, with the vertex of each sector placed at the farthest contour point from the centroid within that sector. Twenty offset rotations were evaluated per polygon order and the configuration yielding the minimum RMSE was selected.

For each polygon order, six features were extracted: RMSE between the contour and the polygon edges (point-to-segment distance), mean distance, standard deviation of distance, maximum distance, polygon area, and the polygon-to-FM area ratio. This yielded 66 polygon features in total.

The rationale for this multi-scale approach is that low-sided polygons (3–5 sides) capture gross shape asymmetry and angularity, while high-sided polygons (9–13 sides) approximate the contour more closely and their residual errors primarily reflect fine boundary roughness and local irregularity. The progression of RMSE across polygon orders thus provides an objective, multi-resolution characterization of contour complexity.

Polygon elbow (optimal complexity) analysis

To identify the polygon order at which additional sides yield diminishing returns in contour approximation accuracy, an elbow-point analysis was performed on each specimen's RMSE curve using the Kneedle algorithm [53]. For each specimen, the RMSE values for polygon orders 3 through 13 were extracted and the curve was normalized to the unit interval on both axes. A reference diagonal line connecting the first and last normalized points was constructed, and the deviation of the normalized RMSE curve from this diagonal was computed at each polygon order. The elbow point was defined as the polygon order at which this deviation was maximal, representing the point of maximum curvature—i.e., the transition from rapid to marginal improvement in contour fit with each additional polygon side [53].

For each specimen, five derived metrics were computed at the identified elbow point: (1) the optimal polygon order (number of sides), (2) the RMSE at the optimal polygon, (3) the total improvement in RMSE from 3-sided to the optimal polygon as a percentage of the 3-sided RMSE, (4) the remaining improvement from the optimal polygon to 13-sided as a percentage, and (5) an efficiency ratio defined as the total improvement up to the elbow divided by the remaining improvement beyond it.

The elbow analysis was performed at both the individual specimen level and the group level. At the group level, RMSE values were averaged across all specimens within each sex and across the entire sample, and the Kneedle algorithm was applied to these mean curves. The distribution of individually determined optimal polygon orders was compared between males and females using the Mann–Whitney U test and Cohen's d to assess whether the two sexes differ in the geometric complexity required to adequately represent their FM boundaries. The step-by-step marginal improvement table—quantifying the percentage RMSE reduction at each polygon-order increment for males, females, and the pooled sample—was also computed and reported.

Elliptic Fourier transform

The normalized elliptic Fourier transform (EFT) was applied to each FM contour following the method of Kuhl and Giardina [24]. The EFT decomposes a closed two-dimensional contour into a series of harmonically related ellipses, where the first harmonic captures the gross elliptical shape and higher harmonics capture progressively finer morphological details. Twenty harmonics were computed, yielding 80 Fourier coefficients (four per harmonic: a□, b□, c□, d□). These coefficients were normalized for size, rotation, and starting-point invariance using the first-harmonic parameters, ensuring that the resulting descriptors reflect pure shape information independent of scale, orientation, and the arbitrary starting location of contour tracing [24, 25].

In addition to the 80 normalized coefficients, 14 summary descriptors were derived: individual power ratios for harmonics 1–10 (the proportion of total spectral power contributed by each harmonic), a higher-harmonic energy ratio (the proportion of total power from harmonics 4–20, reflecting fine morphological detail relative to gross shape), total spectral power, a first-harmonic power ratio (power of the first harmonic relative to total power), and a symmetry index (ratio of odd-harmonic to even-harmonic power, reflecting bilateral symmetry [41]). This yielded 94 EFT features per specimen.

Radial distance profile analysis

To provide intuitive, physically interpretable shape descriptors that correspond directly to classical qualitative morphological descriptions, a radial distance profile was computed for each FM contour. Radial distances from the centroid to each contour point were calculated and sorted by angular position. From this profile, six descriptors were extracted: radial mean and standard deviation (in mm), reflecting overall size and radial variability; radial coefficient of variation (CV), reflecting boundary regularity independent of size; roughness (mean absolute value of the second derivative of the normalized radial profile), quantifying high-frequency boundary undulation; left–right asymmetry (mean absolute difference in normalized radial distances between corresponding points in the left and right halves); and anterior–posterior (AP) asymmetry (the analogous measure for the anterior and posterior quarters).

Roughness, left–right asymmetry, and AP asymmetry were computed from size-normalized (unit-mean) radial profiles to ensure they reflect pure shape properties.

Procrustes normalization and point density heatmaps

To enable shape comparison independent of size, position, and orientation, all resampled FM contours were subjected to Procrustes superimposition [42]. Each contour was: (1) translated to place its centroid at the origin, (2) scaled to unit centroid size, and (3) rotated to align its principal axis (determined by eigendecomposition of the coordinate covariance matrix) with the horizontal axis. A reflection correction ensured consistent orientation across all specimens.

Following Procrustes normalization, sex-specific point density heatmaps were generated by pooling all normalized contour points for each sex onto a 200 × 200 grid. Gaussian kernel smoothing ($\sigma = 2$ grid units) was applied to produce continuous density surfaces (Fig. 1). For each specimen, a male heatmap score and a female heatmap score were computed as the mean density value at the specimen's normalized contour point locations within the respective sex-specific heatmap. The heatmap difference score (male score minus female score) served as an additional descriptive feature, yielding 3 heatmap-derived features per specimen. These global heatmaps were used solely for descriptive visualization and feature extraction; heatmap-based classification employed a separate cross-validated procedure described below.

Statistical analysis

Sex comparison of individual features

All 234 numeric variables were compared between males and females using the Mann–Whitney U test, chosen for its robustness to non-normality [43]. Effect sizes were quantified using Cohen's d, computed as the difference in group means divided by the pooled standard deviation [44]. The Bonferroni correction was applied to all p-values to control the familywise error rate at $\alpha = 0.05$ across all 234 simultaneous comparisons. Features were considered statistically significant if the Bonferroni-corrected p-value was less than 0.05.

Sex classification

The discriminative capacity of the extracted features for sex estimation was evaluated using four classifiers: linear discriminant analysis (LDA) [45], support vector machine with linear kernel (SVM-linear), support vector machine with radial basis function kernel (SVM-RBF) [46], and random forest (RF, 100 trees) [47]. All classifiers were evaluated using 10-fold stratified cross-validation with a fixed random seed (42) to ensure reproducibility. Within each fold, features were standardized to zero mean and unit variance using the training set parameters only, and the same standardization was then applied to the test set to prevent data leakage.

Classification performance was assessed using overall accuracy (proportion of correctly classified specimens), sensitivity (proportion of males correctly identified), and specificity (proportion of females correctly identified).

Feature sets

To systematically evaluate the contribution of size-based versus shape-based features, classification was performed on the following feature sets:

S1 (Size): Six classical size measurements—AP diameter, transverse diameter, maximum and minimum diameters, perimeter, and area—that directly capture FM dimensions in millimeters.

S2 (Shape): All non-size shape features (174 features), including FM index, circularity, eccentricity, area-to-perimeter ratio, compactness, all EFT coefficients and summary descriptors, radial CV, roughness, asymmetry indices, polygon distance metrics, polygon area ratios, ellipse and circle distance metrics (RMSE, mean/std/max distances), and ellipse area ratios. Excluded from this set were absolute area measurements in mm^2 (polygon areas, ellipse areas, FM areas), ellipse geometric parameters (center coordinates, semi-axes, rotation angles), circle radius, radial mean and standard deviation in mm, and polygon elbow-derived metrics, as these either directly encode size information or represent pipeline metadata rather than morphometric descriptors.

S3 (Combined): The union of S1 and S2 (180 features).

S4 (Forward-selected Shape): The minimal subset of S2 features identified by sequential forward selection (see below).

S5 (Forward-selected Combined): The minimal subset of S3 features identified by forward selection.

S6 (Forward-selected All): The minimal subset from all available numeric features (including absolute area measurements and polygon elbow metrics) identified by forward selection.

In addition, two heatmap-based classification methods were evaluated as separate, feature-free approaches (see Heatmap-based classification below).

Forward feature selection

Sequential forward feature selection (SFS) was performed to identify minimal, non-redundant feature subsets that maximize classification accuracy [48]. Starting from an empty set, at each step the feature producing the greatest improvement in 10-fold stratified LDA cross-validation accuracy was added. The process terminated when no additional feature improved accuracy. Three separate forward selections were conducted corresponding to feature sets S4, S5, and S6, selecting from shape features exclusively (S2 pool), the combined size–shape pool (S3), and all available features, respectively.

Heatmap-based sex classification

A novel heatmap-based classification approach was developed as an alternative to conventional feature-based methods. Instead of extracting discrete features and applying a classifier, this method operates directly on the spatial distribution of Procrustes-normalized FM boundary points, comparing each test specimen's contour against sex-specific population density templates.

Two scoring strategies were evaluated under proper 10-fold stratified cross-validation with strict train/test separation—heatmaps were constructed exclusively from training-fold specimens, and test specimens were never included in the heatmaps against which they were scored, thereby preventing data leakage.

Method A — Point-sampling: For each fold, sex-specific density heatmaps were constructed from training specimens by accumulating their Procrustes-normalized contour points onto a shared 200 × 200 grid and applying Gaussian smoothing (σ = 2 grid units). Each test specimen's normalized contour was then overlaid on both the male and female training heatmaps, and the mean density value at the specimen's boundary point positions was computed. Classification was assigned to the sex whose heatmap yielded the higher score.

Method B — Kernel density estimation (KDE) with overlap integral scoring: This method addresses a limitation of point-sampling, which evaluates only the exact grid cells occupied by the test contour and is therefore sensitive to minor positional shifts and pixel-level gaps. In the KDE approach, each contour point—in both the training heatmaps and the test specimen—is treated as the center of a two-dimensional Gaussian kernel, producing continuous density fields rather than discrete point samples. The kernel has full intensity (1.0) at its center, diminishing to approximately half intensity at one pixel distance, one-quarter at two pixels, and so on—following a Gaussian decay profile. This ensures that nearby but not exactly coincident boundary points still contribute to the similarity measure.

The similarity between a test specimen's density field and each sex-specific reference field was quantified using the overlap integral—the element-wise product of the two fields summed over the entire grid. This metric is related to the Bhattacharyya coefficient [54] and provides a continuous, smooth similarity score that is robust to minor positional shifts. Classification was assigned to the sex whose reference field yielded the higher overlap score.

The kernel bandwidth (controlling the spatial extent of each Gaussian kernel in grid-pixel units) was optimized by evaluating values of 3, 5, 7, 9, and 11 pixels and selecting the value that maximized cross-validated accuracy within the same fold structure.

Within-sex morphological subtype discovery

To investigate whether the FM exhibits distinct morphological subtypes within each sex, unsupervised clustering was performed using Gaussian mixture models (GMM) [49]. For each sex independently: (1) Procrustes-normalized contour coordinates were flattened into feature vectors (200 points × 2 coordinates = 400 dimensions), (2) principal component analysis (PCA) was applied to reduce dimensionality to the first 20 components (retaining > 95% of variance), and (3) GMM with full covariance matrices was fitted for k = 2, 3, and 4 clusters (10 random initializations per k).

The optimal number of clusters was selected based on the silhouette score [50] and the Bayesian information criterion (BIC) [51]. Each cluster was characterized by computing its mean values for eight shape descriptors: circularity, radial CV, roughness, left–right asymmetry, AP asymmetry, FM index,

Fitzgibbon ellipse fitting RMSE, and area-to-perimeter ratio. For visualization, each subtype was represented by individual contour overlays (with adaptive transparency based on cluster size), mean contour shapes with ±1 SD and ±2 SD radial variation bands, density heatmaps, and sex-matched subtype comparisons showing the mean contour of the corresponding male and female cluster side by side.

Software and reproducibility

All analyses were implemented in Python 3 using the following libraries: NumPy 1.21+ for numerical computation [33], SciPy 1.7+ for statistical tests and optimization [34], scikit-learn 1.0+ for classification, clustering, and cross-validation [35], Matplotlib 3.4+ for visualization [36], and pandas 1.3+ for data management [52]. The complete analysis code, all 473 anonymized FM boundary coordinate files (500 equally-spaced points per specimen, resampled via periodic cubic spline interpolation), and associated metadata are publicly available in a Zenodo repository (DOI: 10.5281/zenodo.18716445) to ensure full reproducibility. The repository is currently under embargo and will be made publicly accessible upon publication; reviewer access can be arranged upon request to the corresponding author. Per-specimen morphometric measurements are provided in Table S1.

Results

Study sample

A total of 473 adult CT scans met the inclusion criteria. The sample comprised 234 males (49.5%) and 239 females (50.5%). Automated contour extraction yielded between 36 and 187 boundary points per foramen magnum (mean: 85.5), which were resampled to 200 equidistant points for all subsequent analyses. Six specimens were excluded from ellipse fitting due to convergence failure, leaving 467 specimens for the polygon elbow analysis and the full feature set.

Descriptive morphometry

Classical linear and area measurements revealed pronounced sexual dimorphism in all size-related parameters (Table 1). Male FM perimeters averaged 107.13 ± 11.16 mm compared to 95.02 ± 7.36 mm in females (Cohen's $d = 1.284$, $p < 10^{-36}$). FM area showed a comparable effect (males: 830.1 ± 159.9 mm², females: 665.3 ± 99.0 mm²; $d = 1.242$). Anteroposterior and transverse diameters were both significantly larger in males (AP: 36.53 ± 3.48 vs. 32.68 ± 2.77 mm, $d = 1.225$; transverse: 31.35 ± 3.16 vs. 28.23 ± 2.31 mm, $d = 1.132$).

In contrast, shape descriptors showed minimal dimorphism. The FM index did not differ significantly between sexes (males: 1.17 ± 0.08, females: 1.16 ± 0.09; $d = 0.088$, $p > 0.05$ after Bonferroni correction). Circularity showed a small, non-significant trend toward rounder foramina in females (females: 0.92 ± 0.04, males: 0.91 ± 0.07; $d = -0.264$, $p > 0.05$). Left–right and anteroposterior asymmetry indices likewise showed no significant sex differences ($d = 0.034$ and $d = 0.088$, respectively).

Feature significance

Of 234 morphometric features extracted per specimen, 92 (39.3%) showed statistically significant sex differences after Bonferroni correction ($p < 0.05$) (Table 1; Table S2). The top five features by effect size were all area-derived measures: polygon approximation areas ($d = 1.285$ for 10-sided polygon area), perimeter ($d = 1.284$), Fitzgibbon ellipse area ($d = 1.279$), and total FM area ($d = 1.242$). All six classical size measurements were significant.

Among pure shape features (dimensionless ratios and normalized coefficients), only 17 of 146 (11.6%) reached significance. All 11 polygon approximation areas in mm^2 were significant ($d = 1.237–1.285$), but none of the 14 polygon area ratios achieved significance, confirming that the discriminative information resided in absolute size rather than shape proportions. Elliptic Fourier transform (EFT) coefficients showed limited dimorphism: only the total power spectrum ($d = 0.855$, $p < 10^{-15}$) reached significance among the 94 EFT-derived features. The best-fitted ellipse semi-axes were significant (semi-major axis $d = 0.818$, semi-minor axis $d = 0.828$), again reflecting size.

Polygon approximation analysis

The Kneedle-based elbow detection identified the optimal polygon as 6-sided for both sexes (mode = 6, mean = 5.60 ± 0.94 sides) (Table S3). The distribution was: 6-sided (n = 241, 51.6%), 5-sided (n = 84, 18.0%), 4-sided (n = 83, 17.8%), 7-sided (n = 54, 11.6%), and 8-sided (n = 5, 1.1%) (Fig. 4).

The transition from 3 to 4 sides provided the largest marginal RMSE improvement (50.1% for males, 50.7% for females), followed by 5→6 sides (37.6% and 35.5%, respectively) (Table 2; Fig. 3). Beyond 6 sides, each additional vertex contributed diminishing returns (< 25%). Mean optimal polygon RMSE was 1.80 ± 0.47 mm, representing 74.0% improvement over the initial 3-sided approximation. Males and females showed no significant difference in optimal polygon selection (male mean = 5.68, female mean = 5.53; $p > 0.05$), indicating that the geometric complexity of the FM boundary is sex-invariant.

Sex classification

Classification performance was evaluated using seven feature sets across four classifiers with stratified 10-fold cross-validation (Table 3; Table S4; Fig. 4).

The highest accuracy was achieved by forward-selected shape features (S4) with LDA: 80.9% ± 4.9% (sensitivity = 76.4%, specificity = 85.3%). The forward selection procedure identified seven optimal shape features: area-to-perimeter ratio, circularity, Fitzgibbon ellipse RMSE, and four EFT coefficients (harmonics 48, 21, 13, and 51) (Table 4). Notably, the first feature alone (area-to-perimeter ratio) achieved 74.9%, and the addition of circularity raised accuracy to 77.7%.

Forward-selected combined features (S5, LDA) and the full feature set (S6, LDA) both reached 80.5% ± 6.8%, with higher specificity (85.7%) but lower sensitivity (75.1%) than S4 (Table S5). The combined

forward selection started with perimeter (78.1% alone) and added five shape refinements to reach the final accuracy.

Size features alone (S1) reached 75.6% ± 5.7% with LDA, and 77.3% ± 7.0% with SVM-rbf. The full shape feature set without selection (S2) achieved 67.9% with LDA but 77.1% with Random Forest, suggesting non-linear shape relationships captured better by ensemble methods (Fig. 4).

Among classifiers, LDA consistently outperformed others on forward-selected features, while Random Forest showed superior performance on unselected high-dimensional feature sets (S2: RF 77.1% vs. LDA 67.9%).

Heatmap-based density classification

A novel classification approach was implemented using Procrustes-normalized point-density heatmaps. Two scoring strategies were compared under proper 10-fold cross-validation with strict train/test separation.

Point-sampling — where each test contour point was scored against the reference heatmap at its exact grid position — achieved only 50.3% ± 3.2%, effectively at chance level. The method exhibited extreme sensitivity bias (96.2% for males) with near-zero specificity (5.4% for females), indicating systematic classification toward one sex.

KDE overlap scoring — where both the reference and test contours were rendered as Gaussian density fields (optimal bandwidth = 9 pixels) and their element-wise product integrated — improved accuracy to 54.1% ± 7.1% with more balanced sensitivity (70.5%) and specificity (38.1%).

The near-chance performance of both heatmap methods is attributable to Procrustes normalization, which removes all size information. Since the dominant sexual dimorphism in FM morphology is size-based (top features: d = 1.28), the normalized contours retain only subtle shape differences insufficient for density-based discrimination. This negative finding is consistent with the observation that pure shape features without feature selection achieved only 67.9% accuracy even with LDA, a method with far greater discriminative capacity than density overlap (Fig. 1).

Within-sex morphological subtypes

GMM clustering on Procrustes-normalized contour coordinates, evaluated for k = 2, 3, and 4 clusters, revealed distinct patterns of within-sex variation (Table 5; Fig. 5; Table S6; Table S7, Figure S1).

For males, the optimal cluster number was k = 2 based on silhouette score (0.72) and BIC (Fig. 6). The dominant subtype (Subtype-1, n = 175, 74.8%) showed higher circularity (0.914) and a more elliptical FM index (1.177), representing the typical male FM morphology. The minor subtype (Subtype-0, n = 59, 25.2%) exhibited lower circularity (0.884) and higher radial variability (CV = 7.7%), indicating a distinct, more irregular morphological pattern (Fig. 5; Figure S2).

For females, the silhouette score peaked at k = 3 (0.78), though k = 2 also performed well (0.74) (Fig. 6). At k = 2, the dominant subtype (Subtype-0, n = 171, 71.5%) showed high circularity (0.923) with an FM index of 1.164. The secondary subtype (Subtype-1, n = 68, 28.5%) had comparable circularity (0.920) but slightly higher asymmetry (0.058 vs. 0.050), suggesting size-related rather than shape-related differentiation (Fig. 5; Figure S3).

The subtype analysis revealed that male FM morphology exhibits greater heterogeneity than female. At k = 4, male subtypes included a clearly atypical group (Subtype-1, n = 16, 6.8%) with the highest radial CV (10.5%) and lowest circularity, while female subtypes remained more morphologically coherent. These findings suggest that the male foramen magnum encompasses a wider range of morphological variants (Figure S4-S17).

Summary of key findings

The foramen magnum shows clear sexual dimorphism, but this dimorphism is overwhelmingly driven by size rather than shape. Males have FM perimeters approximately 12.7% longer and areas roughly 24.8% larger than females, yielding large effect sizes (d ≈ 1.28). However, when size is removed through normalization, male and female FM boundaries are remarkably similar — shape-only features achieve only modest classification gains over chance.

The most effective sex estimation (80.9% accuracy) required a combination of shape features selected through forward selection, including geometric ratios, ellipse fitting residuals, and specific Fourier harmonics. This suggests that while no single shape feature strongly discriminates sex, certain combinations of subtle boundary irregularities can collectively approach the discriminative power of size-based measurements. The finding that shape features outperform size features at optimal selection (80.9% vs. 75.6%) indicates that shape encodes complementary — though individually weak — dimorphic information.

In practical forensic terms, these results mean that: (1) the foramen magnum can correctly predict biological sex in approximately 4 out of 5 individuals; (2) female FMs are better classified (specificity 85.3%) than male FMs (sensitivity 76.4%); and (3) when only fragmentary remains are available where size cannot be assessed, automated shape analysis can still provide meaningful, though less precise, sex estimation.

Discussion

Principal findings and novelty

This study presents the first comprehensive computational geometric morphometric analysis of the foramen magnum that integrates multiple complementary outline-based approaches — elliptic Fourier transform, systematic polygon approximation, ellipse fitting, radial distance profiling, Procrustes-based

density heatmaps, and Gaussian kernel density estimation — to jointly characterize FM shape and its relationship to biological sex. While conventional FM studies have relied on a small number of linear measurements and subjective shape classification [4, 8, 17], the present work extracted 231 quantitative features per specimen from automatically detected contour coordinates, enabling an objective, reproducible, and high-dimensional evaluation of FM morphology.

The best classification accuracy of 80.9% was achieved using forward-selected shape features with LDA — a result that compares favorably with the reported range of 62.5%–86.7% for conventional FM morphometry across various populations [5, 6, 10, 17, 55]. Our most consequential finding, however, is not the classification accuracy itself but rather the quantitative demonstration that FM sexual dimorphism is overwhelmingly driven by size rather than shape. When size information was removed through Procrustes normalization, classification accuracy dropped sharply, and the novel density-heatmap approach — which operates exclusively on size-normalized contours — achieved only 54.1%. This provides the first direct, quantitative evidence that residual FM shape differences between sexes are minimal once allometric scaling is accounted for.

Size dominates FM sexual dimorphism

All six classical size measurements differed significantly between sexes with large effect sizes (Cohen's $d = 1.13–1.28$). Perimeter showed the strongest discriminative capacity ($d = 1.284$, $p < 10^{-36}$), followed by polygon area ($d = 1.285$) and total FM area ($d = 1.242$). Males had FM perimeters approximately 12.7% longer and areas roughly 24.8% larger than females. These findings are consistent with the broader FM literature. Gapert et al. [5] reported that FM width and area were among the best univariate discriminating parameters in a British sample, achieving a maximum multivariate accuracy of 70.3%. Madadin et al. [6] similarly found that all FM dimensions were significantly larger in males in a Saudi population, with width yielding the highest univariate accuracy of 65.0%. A recent systematic review and meta-analysis by Fernandes et al. [15] confirmed that FM dimensions are consistently larger in males across populations, though classification accuracy varies substantially by population.

In contrast, shape indices showed negligible dimorphism in our sample. The FM index — the most widely used shape descriptor in the literature — did not differ significantly between sexes ($d = 0.088$, $p > 0.05$). Circularity showed only a weak, non-significant trend ($d = -0.264$), and left–right and AP asymmetry were virtually identical between sexes. These results align with Toneva et al. [8], who found that FM shape did not provide substantial sex discrimination in a Bulgarian population, and with our previous study [17], which reported that the H/W index and FMI had no significant sex difference ($p = 0.622$ and $p = 0.440$, respectively) and yielded univariate accuracies of only 51.9% and 52.8%.

From subjective categories to objective shape quantification

A particularly notable contribution of this work concerns the long-standing problem of FM shape classification. Numerous studies have attempted to characterize FM morphology using categorical

descriptors such as "oval," "round," "tetragonal," "pentagonal," "egg-shaped," "hexagonal," and "irregular" [4, 18, 56]. As Zdilla et al. [22] noted in their comprehensive review, these shape categories rely on ambiguous terminology that is inconsistent between studies, observer-dependent, and difficult to reproduce. Indeed, visual shape classification is inherently subjective: the same foramen may be classified as "oval" by one observer and "egg-shaped" by another, and studies vary widely in how many and which shape categories they employ [22].

Our polygon approximation analysis offers a direct, objective alternative. By fitting optimal polygons of 3 through 13 sides to each FM contour and identifying the elbow point where additional sides yield diminishing accuracy improvements, we determined that the typical FM is best approximated by a 6-sided polygon (mode = 6 in both sexes, median = 6, mean = 5.60 ± 0.94). This finding objectively confirms and quantifies what the traditional literature has described qualitatively: the FM boundary is broadly hexagonal with substantial individual variation spanning 4 to 8 sides (Fig. 2). Furthermore, the absence of a significant sex difference in optimal polygon order (male mean = 5.68, female mean = 5.53; $p > 0.05$) provides the first quantitative evidence that the geometric complexity of the FM boundary is sex-invariant — a finding that could not have been established through subjective visual categorization.

Similarly, the elliptic Fourier transform decomposed each FM contour into 20 harmonics, enabling objective quantification of boundary features at multiple spatial scales. Of the 94 EFT-derived features, only the total spectral power ($d = 0.855$) reached significance, reflecting the global size component captured by the power spectrum. Individual harmonic power ratios and normalized coefficients showed minimal dimorphism, confirming that fine boundary details — the "roughness," undulations, and local irregularities that distinguish one FM shape from another — do not differ systematically between sexes.

The value of shape features despite modest individual effects

Although no single shape feature showed strong dimorphism, the forward selection procedure demonstrated that particular combinations of individually weak shape features can collectively approach and even exceed the discriminative capacity of size-based measurements. The optimal shape feature subset (area-to-perimeter ratio, circularity, Fitzgibbon ellipse RMSE, and four EFT coefficients) achieved 80.9% accuracy — surpassing size features alone (75.6%) (Table 3). This pattern suggests that shape encodes complementary information: each shape feature captures a subtle, partially independent aspect of FM morphology, and their combination reveals dimorphic patterns that individual features or simple size measurements cannot.

This finding has practical implications for forensic anthropology. In cases involving fragmentary cranial remains — such as fire-damaged or blast-affected specimens — size measurements may be compromised by incomplete margins, while shape features derived from even a partially preserved FM contour could still contribute to sex estimation [5, 30]. Furthermore, the forward selection identified specific features — area-to-perimeter ratio, circularity, and ellipse fitting residuals — that are computationally simple and could be readily integrated into automated forensic analysis tools.

Heatmap-based classification: an informative negative result

The density heatmap approach, particularly the novel KDE overlap scoring method, represents a conceptually distinct strategy that bypasses discrete feature extraction entirely, instead comparing continuous spatial distributions of boundary points. The near-chance performance of this method (54.1%) is itself an informative finding. Because heatmap classification operates on Procrustes-normalized contours — from which all size information has been removed — its failure provides direct confirmation that the spatial distribution of FM boundary points does not differ meaningfully between sexes after size normalization. This result is consistent with the statistical finding that pure shape features collectively explain only modest variance in sex (LDA on the full shape feature set yielded 67.9%), and offers a population-level visualization of this principle.

We note that this approach may prove more effective for anatomical structures where shape differences between groups are more pronounced. For example, EFA applied to the greater sciatic notch — a structure with well-documented qualitative shape dimorphism — has achieved classification rates exceeding 80% [57]. The heatmap methodology developed here could be readily extended to such structures, where it may provide superior discrimination precisely because the shape differences are larger.

Within-sex morphological heterogeneity

The GMM-based subtype analysis revealed an asymmetry in within-sex morphological variation. Males showed greater morphological heterogeneity, with k = 2 as the optimal cluster number but a clearly identifiable atypical minority subtype (25.2%) characterized by lower circularity and higher radial variability (Figure S5). Females, by contrast, exhibited more morphological homogeneity, with the majority (71.5%) belonging to a single dominant subtype. This asymmetry in within-sex variability is also reflected in the descriptive statistics reported in our previous study [17], where males showed higher standard deviations than females in most FM measurements (e.g., FM area SD: 121.5 mm² vs 101.6 mm²; perimeter SD: 8.15 mm vs 6.8 mm). The differential classification performance observed in the present study — females more reliably classified (specificity 85.3%) than males (sensitivity 76.4%) — is likely a consequence of this narrower distribution of female FM morphology providing a more distinct separation from the male distribution.

Methodological contributions and broader applicability

Beyond the specific findings regarding the FM, this study demonstrates the feasibility and value of applying a comprehensive, automated geometric morphometric pipeline to skeletal structures that lack well-defined anatomical landmarks. The FM is particularly challenging for traditional landmark-based morphometrics because its margin is a smooth, continuous curve without discrete, reproducible points that would qualify as Type I or Type II landmarks [58]. Our approach — treating the entire contour as the

unit of analysis through resampling, Fourier decomposition, polygon approximation, and density mapping — circumvents this limitation entirely.

Caple et al. [59] highlighted the potential of EFA in forensic anthropology, noting that outline-based methods can capture a large amount of shape information in contrast to landmark approaches where information falling between landmarks fails to be acquired. Similarly, Kilmer and Garvin [57] demonstrated that EFA of skeletal outlines can support qualitative descriptions of sex differences with objective, quantitative data. Our work extends this paradigm by combining EFA with multiple additional geometric analysis methods, providing a more complete morphometric characterization than any single technique alone.

The complete analysis pipeline — from automated contour extraction through feature computation, statistical testing, classification, and subtype discovery — is implemented as a single, self-contained, publicly available script. This design facilitates direct application to other skeletal structures where shape variation has traditionally been described only qualitatively, such as the obturator foramen [57, 60] or piriform aperture. The objective quantification of shapes that are currently assessed visually (e.g., "oval vs. triangular obturator foramen") could enable more reproducible forensic assessments and contribute to a more precise anatomical vocabulary for morphological description.

Comparison with the source study

Our dataset derives from the same CT image collection used in our previous study [17], which reported an accuracy of 86.7% using linear discriminant function analysis with leave-one-out cross-validation on the full 720-case sample, and 88.2% using artificial neural networks. The apparent difference from our best accuracy of 80.9% in the present study likely reflects sampling effects rather than methodological inferiority. First, 247 cases were excluded in the present study due to insufficient contour quality for automated edge detection, reducing the sample to 473. This exclusion was not random with respect to morphology: cases with lower image quality, partial volume artifacts, or less distinct FM margins were preferentially removed, potentially yielding a subsample that — by chance — captured a less dimorphic subset of the original population. Second, the present study employed 10-fold rather than leave-one-out cross-validation, which can yield slightly different accuracy estimates depending on fold composition. Despite this reduction in sample size, the present study achieved a comparable cross-validated accuracy (80.9%) while simultaneously demonstrating that shape features, when optimally selected, can match and even marginally exceed the performance of classical size-based measurements — a finding that was not addressed in the original study.

Limitations

Several limitations should be acknowledged. First, the study sample was drawn from a single Eastern Turkish population, and the population-specificity of FM morphometry is well established [6, 15]. As noted in our previous study [17], even within Turkey, discriminant formulas from western Turkish samples showed variable performance when applied to the eastern Turkish sample, reflecting potential ancestral

differences between regions. The classification models and feature importance rankings derived here may therefore not generalize directly to other populations. Second, the contour extraction relied on 2D axial CT slices rather than 3D surface reconstructions, potentially missing three-dimensional shape information such as FM depth, curvature, and condylar morphology. Third, the automated edge detection required sufficient CT image quality, which led to the exclusion of 247 cases from the original dataset; this introduces a potential selection bias toward cases with clearer FM margins. Fourth, age-related variation was not explicitly modeled. Although our previous study found no significant correlation between FM measurements and age apart from a weak trend in perimeter [17], age effects on shape features remain unexplored. Finally, the heatmap and KDE methods employed fixed grid resolutions and kernel bandwidths that may not be optimal for all contour morphologies; adaptive or multi-scale approaches could potentially improve performance.

Future directions

The computational framework developed here opens several avenues for future research. Application to skeletal structures with more pronounced qualitative shape dimorphism — such as the greater sciatic notch, obturator foramen, or frontal sinus — could test whether the polygon approximation and heatmap methods achieve higher discriminative accuracy when shape differences are more substantial. Multi-population studies using the same pipeline would clarify which FM shape features, if any, transcend population boundaries. Integration of 3D surface data, enabled by modern CT reconstruction capabilities, could capture additional morphological information inaccessible from 2D contours. Finally, the subtype discovery approach could be extended to investigate whether FM morphological variants correlate with other craniometric traits, functional anatomy, or developmental influences.

Conclusions

This study demonstrates that the sexual dimorphism of the foramen magnum is predominantly mediated by overall size rather than by outline shape. Among 234 CT-derived morphometric features, all six conventional size measurements produced large, highly significant sex differences, whereas only 11.6% of pure shape features reached significance after correction for multiple testing. Perimeter emerged as the single strongest discriminator (Cohen's d = 1.284), and size features alone achieved 75.6% classification accuracy. Importantly, however, a forward-selected combination of seven shape features—spanning ellipse fitting residuals, polygon approximation metrics, Fourier harmonic descriptors, and radial asymmetry indices—surpassed size-based classification, reaching 80.9% under rigorous 10-fold cross-validation. This finding establishes that individually weak shape signals, when optimally combined, carry discriminative information that is not redundant with overall size.

From a methodological standpoint, this study replaces the subjective and non-reproducible categorical shape classification that has prevailed in the FM literature with a fully automated, deterministic pipeline that extracts objective and continuous shape descriptors from standard axial CT images. The seven

complementary geometric analyses—classical morphometry, Fitzgibbon and genetic-algorithm-optimized ellipse fitting, systematic polygon approximation, normalized elliptic Fourier transform, radial distance profiling, and Procrustes-based density heatmap construction—collectively provide a multiscale characterization of the FM boundary that is reproducible across observers and institutions. The polygon approximation analysis, in particular, revealed that both sexes are best described by a hexagonal template, offering a data-driven, objective alternative to the inconsistent subjective shape taxonomies reported in the literature. The pipeline is distributed as a single open-source Python script with no proprietary software dependencies, facilitating independent replication and extension.

The unsupervised clustering analysis further revealed a previously unreported asymmetry in within-sex morphological variation: male foramina magna display greater heterogeneity, with a distinct atypical subtype comprising approximately one quarter of male specimens, whereas female FM morphology is substantially more homogeneous. This observation suggests that the developmental and biomechanical constraints shaping the FM may act more uniformly in females, and it warrants further investigation across geographically diverse samples.

In practical forensic and clinical terms, the foramen magnum can predict biological sex correctly in approximately four out of five individuals using the computational approach presented here. As 80.9% accuracy approaches the upper boundary reported in the conventional FM literature, these results suggest that further gains from two-dimensional FM morphometry alone may be limited. Future studies should validate these findings across multiple populations, extend the pipeline to three-dimensional surface data enabled by modern CT reconstruction capabilities, and apply the framework to anatomical structures where shape dimorphism is expected to be more pronounced, such as the greater sciatic notch and the pelvic inlet.

Abbreviations

FM
Foramen magnum
CT
Computed tomography
AP
Anteroposterior
EFT
Elliptic Fourier transform
GA
Genetic algorithm
LDA
Linear discriminant analysis
SVM
Support vector machine

RF
Random forest
KDE
Kernel density estimation
GMM
Gaussian mixture model
BIC
Bayesian information criterion
RMSE
Root mean square error
CV
Coefficient of variation
SD
Standard deviation
SFS
Sequential forward selection

Declarations

Ethics approval and consent to participate

The original study protocol received approval from the Non-Interventional Clinical Research Ethics Committee of Van Yuzuncu Yil University Faculty of Medicine (decision number: 02, dated: 19.02.2017) and was conducted in accordance with the Declaration of Helsinki and the principles of Good Clinical Practice. Since the present study represents a secondary analysis of the same anonymized imaging dataset using computational geometric methods, no additional ethical approval was required. Informed consent was waived due to the retrospective design and full anonymization of all data.

Consent for publication

Not applicable.

Availability of data and materials

The complete analysis pipeline is publicly available as a single, self-contained Python script (FM_Comprehensive_Analysis.py). Per-specimen morphometric measurements are provided in the supplementary materials (Table S1 in Additional file 1). The source CT images cannot be shared publicly due to institutional data protection regulations but are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions

YE conceived the study, developed the computational pipeline and software, performed programming and data analyses, conducted the literature review, and wrote the original draft of the manuscript. EK contributed to image acquisition and selection, literature review, ethics committee procedures, and radiological image analysis. MA contributed to project organization, supervision, conceptualization, and critical revision of the final manuscript. YH contributed to supervision and conceptualization. SK contributed to data curation and statistical analysis. UD contributed to the literature review, preparation of radiological images for the study, and critical revision of the final manuscript. AY contributed to radiological image acquisition, development of analysis methodology, and preparation of images for final analysis. OC contributed to project organization, supervision, data analysis, conceptualization, and critical revision of the final manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Standring S. Gray's Anatomy: The Anatomical Basis of Clinical Practice. 42nd ed. Amsterdam: Elsevier; 2020.
2. Tubbs RS, Griessenauer CJ, Loukas M, Shoja MM, Cohen-Gadol AA. Morphometric analysis of the foramen magnum: an anatomic study. Neurosurgery. 2010;67(5):1479–84.
3. Catalina-Herrera CJ. Study of the anatomic metric values of the foramen magnum and its relation to sex. Acta Anat (Basel). 1987;130(4):344–7.
4. Murshed KA, Çiçekcibaşi AE, Tuncer I. Morphometric evaluation of the foramen magnum and variations in its shape: a study on computerized tomographic images of normal adults. Turk J Med Sci. 2003;33(5):301–6.
5. Gapert R, Black S, Last J. Sex determination from the foramen magnum: discriminant function analysis in an eighteenth and nineteenth century British sample. Int J Legal Med. 2009;123(1):25–33.
6. Madadin M, Menezes RG, Al Saif HS, Abu Alola H, Al Muhanna A, Gullenpet AH, et al. Morphometric evaluation of the foramen magnum for sex determination: a study from Saudi Arabia. J Forensic Leg Med. 2017;46:66–71.
7. Aljarrah K, Packirisamy V, Al Anazi N, Nayak SB. Morphometric analysis of foramen magnum and occipital condyle using CT images for sex determination in a Saudi Arabian population.

Morphologie. 2022;106(355):260–8.
8. 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. Leg Med (Tokyo). 2018;35:50–60.
9. Ajanovic Z, Dervisevic L, Dervisevic E, Lujinovic A, Ajanovic U, Biscevic-Tokic J, et al. Sex estimation based on foramen magnum: a three-dimensional geometric morphometrics approach. Int J Morphol. 2023;41(2):410–6.
10. Uthman AT, Al-Rawi NH, Al-Timimi JF. Evaluation of foramen magnum in gender determination using helical CT scanning. Dentomaxillofac Radiol. 2012;41(3):197–202.
11. Jaitley M, Phulambrikar T, Kode M, Gupta A, Singh SK. Foramen magnum as a tool for sexual dimorphism: a cone beam computed tomography study. Indian J Dent Res. 2016;27(5):458–62.
12. Singh PK, Tamrakar D, Paudel S, Timsinha S, Pokhrel RP, Shrestha R, et al. Morphometric analysis of the foramen magnum in sex estimation: an additional 3DCT study from Nepal on a larger sample. Anat Cell Biol. 2022;55(4):401–11.
13. Mehta M, Saini V, Patel MN, Menon SK. Applicability and reliability of foramen magnum for sex determination in contemporary Western Indian population: a computed tomographic study. J Forensic Radiol Imaging. 2019;17:13–9.
14. Akay G, Güngör K, Peker İ. Morphometric analysis of the foramen magnum using cone beam computed tomography. Turk J Med Sci. 2017;47(6):1715–22.
15. 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 Sci Med Pathol. 2024;20(1):268–79.
16. 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. Leg Med (Tokyo). 2020;47:101775.
17. Kartal E, Etli Y, Asirdizer M, Hekimoglu Y, Keskin S, Demir U, Yavuz A, Celbis O. Sex estimation using foramen magnum measurements, discriminant analyses and artificial neural networks on an Eastern Turkish population sample. Leg Med (Tokyo). 2022;59:102143.
18. Chethan P, Prakash KG, Murlimanju BV, Prashanth KU, Prabhu LV, Saralaya VV, et al. Morphological analysis and morphometry of the foramen magnum: an anatomical investigation. Turk Neurosurg. 2012;22(4):416–9.
19. Aragão JA, de Oliveira Pereira R, da, Cruz de Moraes RZ, Reis FP. Morphological types and the dimorphism of the foramen magnum of dry skulls from northeastern Brazil. Int J Morphol. 2014;32(4):1372–8.
20. Gargi V, Yadav PK, Kumar Y. Foramen magnum as a tool for sexual dimorphism. Natl J Med Res. 2018;8:28–31.
21. Singh G, Gargi V. Discriminant function analysis of foramen magnum variables in South Indian population: a study of computerised tomographic images. J Clin Diagn Res. 2018;12(5):HC01–5.

22. Zdilla MJ, Russell ML, Bliss KN, Mangus KR, Koons AW. The size and shape of the foramen magnum in man. J Craniovertebr Junction Spine. 2017;8(3):205–21.
23. Singh G, Talwar I. Morphometric analysis of foramen magnum for sex determination using CT scan. Indian J Forensic Med Toxicol. 2013;7(2):220–3.
24. Kuhl FP, Giardina CR. Elliptic Fourier features of a closed contour. Comput Graph Image Process. 1982;18(3):236–58.
25. Haines AJ, Crampton JS. Improvements to the method of Fourier shape analysis as applied in morphometric studies. Palaeontology. 2000;43(4):765–83.
26. Samuel OM, Bhagwat AA. Elliptical Fourier descriptors of outline and morphological analysis in caudal view of foramen magnum of the tropical raccoon (Procyon cancrivorus) (Linnaeus, 1758). J Morphol Sci. 2018;35(1):24–32.
27. Hayashizaki Y, Usui A, Hosokai Y, Sakai J, Funayama M. Sex determination of the pelvis using Fourier analysis of postmortem CT images. Forensic Sci Int. 2015;246:e1221–9.
28. Gonzalez PN, Bernal V, Perez SI. Geometric morphometric approach to sex estimation of human pelvis. Forensic Sci Int. 2009;189(1–3):68–74.
29. Etli Y, Kartal E, Demir U, Etli E, Hekimoğlu Y, Aşirdizer M. Sex determination using image processing reveals elliptical features of the female pelvic inlet. Sci Rep. 2025;15:33800.
30. Krishan K, Chatterjee PM, Kanchan T, Kaur S, Baryah N, Singh RK. A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. Forensic Sci Int. 2016;261:e1651–8.
31. Tsiouris C, Triantafyllou G, Karangeli N, Botis GG, Papadopoulos-Manolarakis P, Kalamatianos T, et al. Morphometric assessment of occipital condyles and foramen magnum reveals enhanced sexual dimorphism detection via 3D imaging: a systematic review and meta-analysis utilizing classification and regression trees. Diagnostics (Basel). 2025;15(11):1359.
32. De Boor C. A Practical Guide to Splines. Rev. ed. New York: Springer; 2001.
33. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature. 2020;585(7825):357–62.
34. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17(3):261–72.
35. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res. 2011;12:2825–30.
36. Hunter JD. Matplotlib: a 2D graphics environment. Comput Sci Eng. 2007;9(3):90–5.
37. Braden B. The surveyor's area formula. Coll Math J. 1986;17(4):326–37.
38. Fitzgibbon A, Pilu M, Fisher RB. Direct least square fitting of ellipses. IEEE Trans Pattern Anal Mach Intell. 1999;21(5):476–80.
39. Nelder JA, Mead R. A simplex method for function minimization. Comput J. 1965;7(4):308–13.
40. Kasa I. A circle fitting procedure and its error analysis. IEEE Trans Instrum Meas. 1976;25(1):8–14.

41. Lestrel PE. Fourier Descriptors and Their Applications in Biology. Cambridge: Cambridge University Press; 1997.
42. Rohlf FJ, Slice D. Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst Biol. 1990;39(1):40–59.
43. Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. Ann Math Stat. 1947;18(1):50–60.
44. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale: Lawrence Erlbaum Associates; 1988.
45. Fisher RA. The use of multiple measurements in taxonomic problems. Ann Eugen. 1936;7(2):179–88.
46. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20(3):273–97.
47. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32.
48. Whitney AW. A direct method of nonparametric measurement selection. IEEE Trans Comput. 1971;C–20(9):1100–3.
49. McLachlan GJ, Peel D. Finite Mixture Models. New York: Wiley; 2000.
50. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20:53–65.
51. Schwarz G. Estimating the dimension of a model. Ann Stat. 1978;6(2):461–4.
52. McKinney W. Data structures for statistical computing in Python. Proc 9th Python Sci Conf. 2010;445:51–6.
53. Satopaa V, Albrecht J, Irwin D, Raghavan B. Finding a kneedle in a haystack: detecting knee points in system behavior. In: 2011 31st International Conference on Distributed Computing Systems Workshops. IEEE; 2011. pp. 166–71.
54. Bhattacharyya A. On a measure of divergence between two statistical populations defined by their probability distributions. Bull Calcutta Math Soc. 1943;35:99–109.
55. 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 Sci Rep. 2023;6(1):e999.
56. Vinutha SP, Suresh V, Shubha R. Discriminant Function Analysis of Foramen Magnum Variables in South Indian Population: A Study of Computerised Tomographic Images. Anat Res Int. 2018;2018:2056291.
57. Kilmer K, Garvin HM. Outline analysis of sex and population variation in greater sciatic notch and obturator foramen morphology with implications for sex estimation. Forensic Sci Int. 2020;314:110346.
58. Bookstein FL. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Med Image Anal. 1997;1(3):225–43.

59. Caple J, Byrd J, Stephan CN. Elliptical Fourier analysis: fundamentals, applications, and value for forensic anthropology. Int J Legal Med. 2017;131(6):1675–90.
60. Bierry G, Le Minor JM, Schmittbuhl M. Oval in males and triangular in females? A quantitative evaluation of sexual dimorphism in the human obturator foramen. Am J Phys Anthropol. 2010;141(4):626–31.

Tables

Tables are available in the Supplementary Files section.

Figures

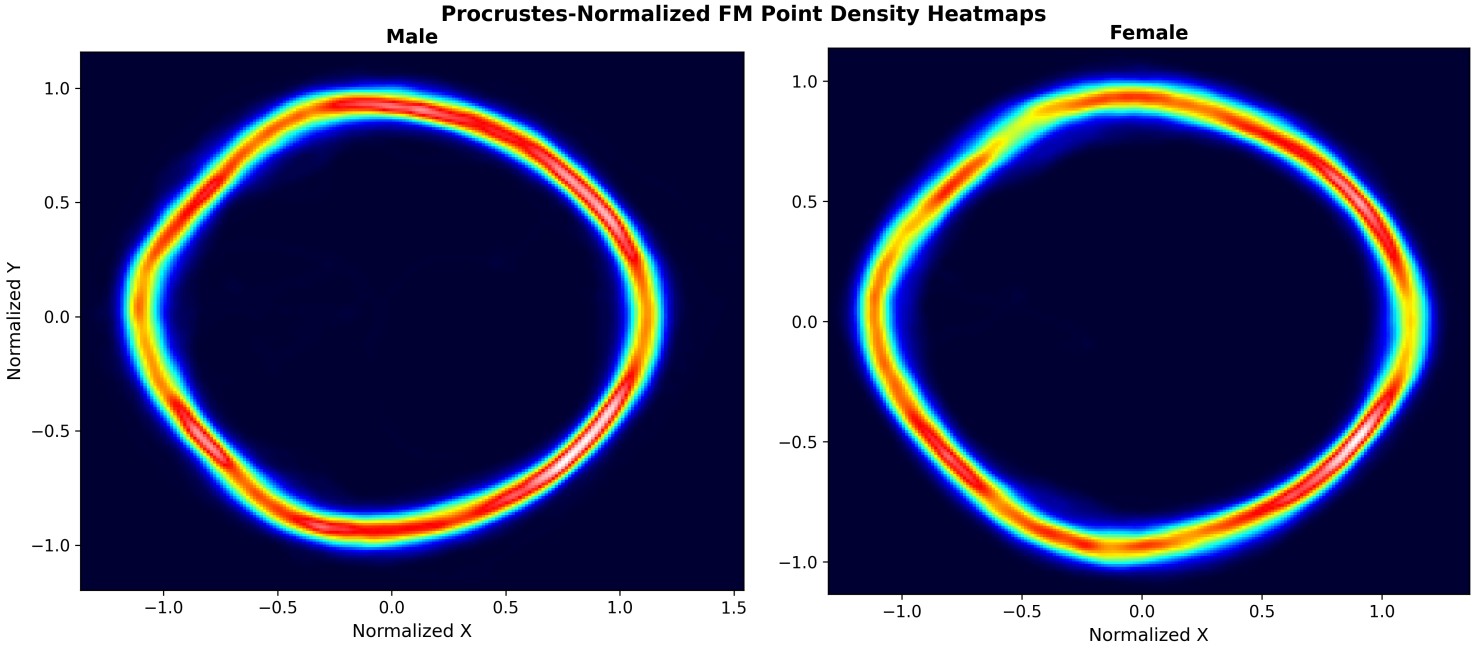


Figure 1

Procrustes-normalized foramen magnum point density heatmaps for males (left) and females (right). Gaussian kernel smoothing (σ = 2 grid units) was applied to pooled contour points on a 200 × 200 grid. Color intensity represents point density, with warmer colors indicating higher concentrations.

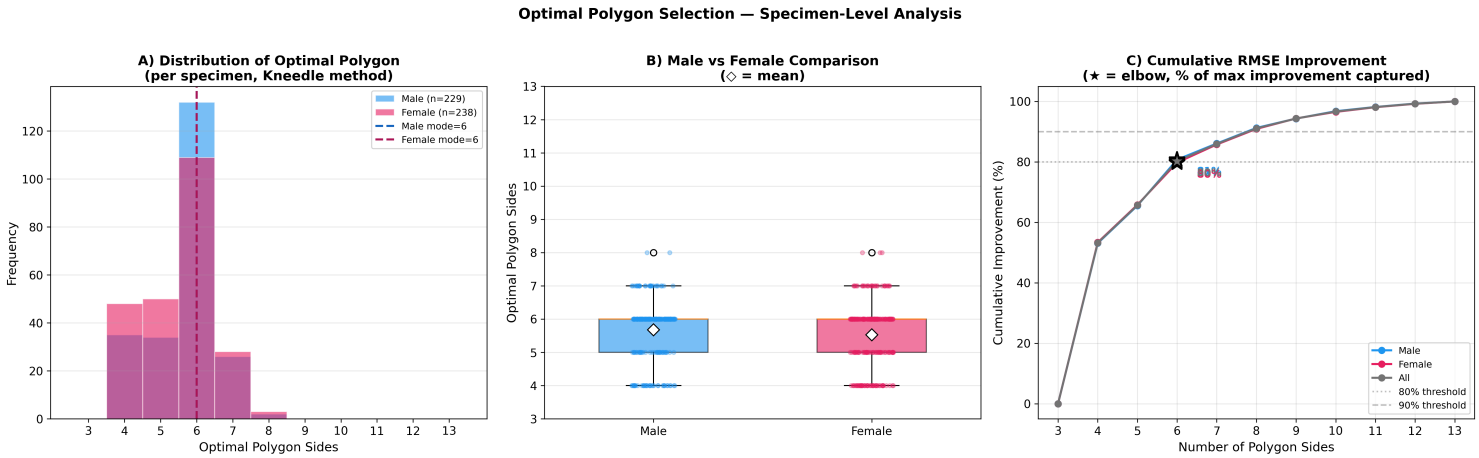


Figure 2

Distribution of individually determined optimal polygon sides using the Kneedle algorithm. (A) Histogram of optimal polygon order by sex (mode = 6 for both). (B) Box plot comparison between males and females. (C) Cumulative RMSE improvement as a function of polygon sides, with star indicating the elbow point.

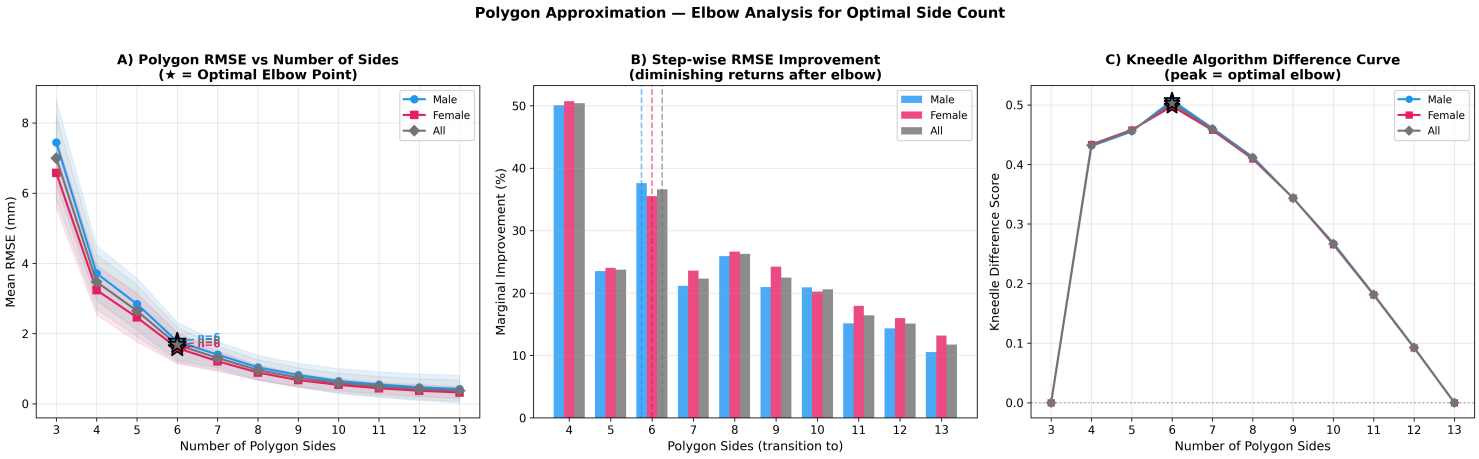


Figure 3

Polygon approximation elbow analysis for optimal side count. (A) Mean RMSE versus number of polygon sides by sex, with stars indicating optimal elbow points. (B) Step-wise marginal RMSE improvement showing diminishing returns beyond 6 sides. (C) Kneedle algorithm difference curve with peak at 6 sides.

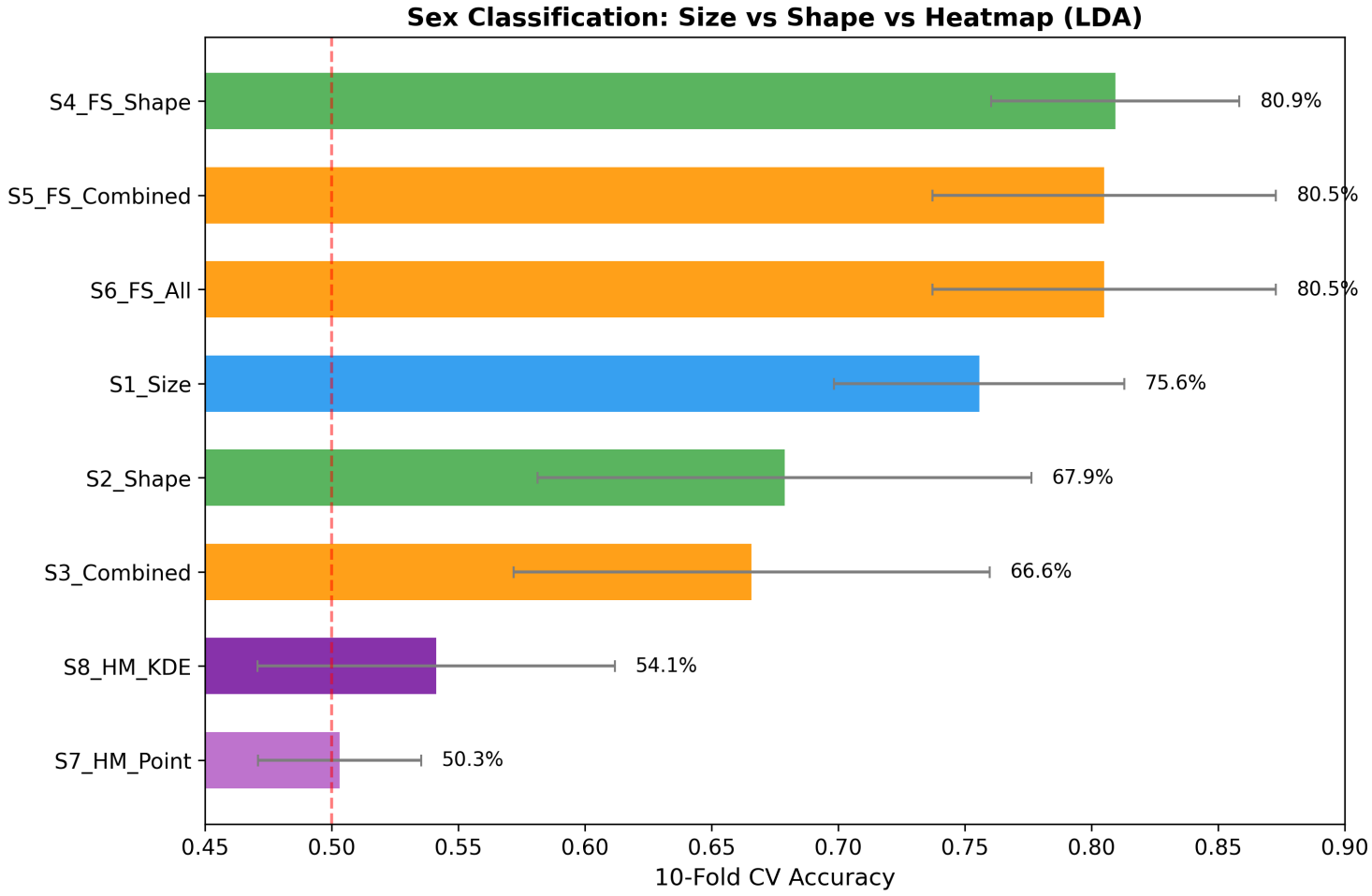


Figure 4

Sex classification performance comparison across feature sets using LDA. Horizontal bars show 10-fold cross-validation accuracy with error bars indicating standard deviation. The dashed red line indicates chance level (50%). Feature sets are ordered by accuracy.

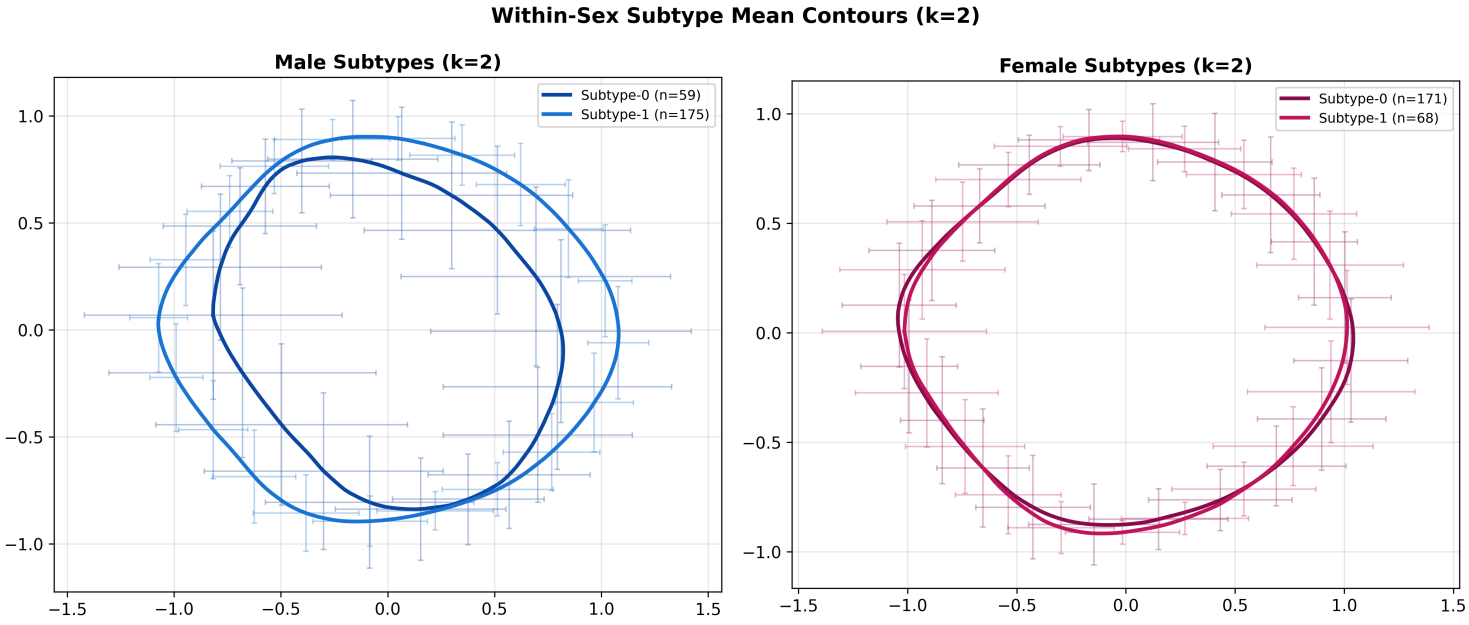


Figure 5

Within-sex subtype mean contours at k = 2. Male subtypes (left, blue) and female subtypes (right, pink) with error bars showing ±1 SD at sampled contour points. Male Subtype-0 (n = 59) shows a distinctly irregular, less circular shape compared to Subtype-1 (n = 175). Female subtypes show minimal shape differentiation.

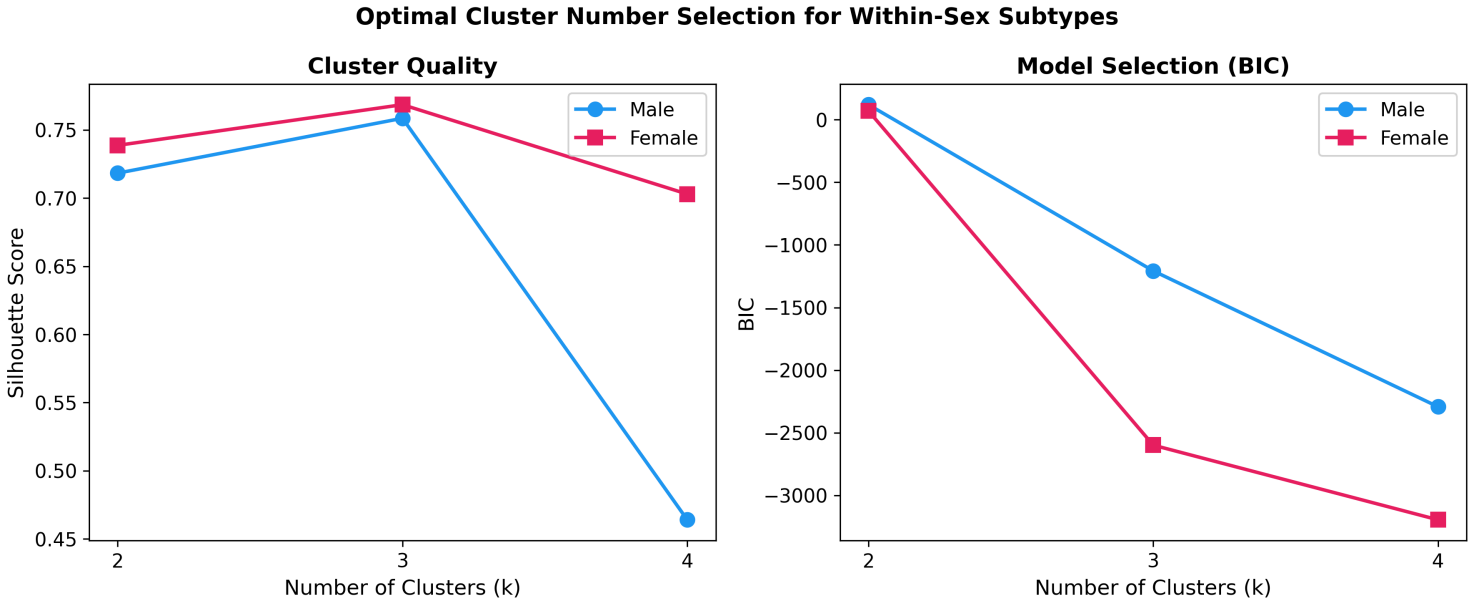


Figure 6

Optimal cluster number selection for within-sex morphological subtypes. Left: Silhouette score versus number of clusters (k). Right: Bayesian information criterion (BIC) versus k. For males, k = 2 is optimal based on silhouette score. For females, silhouette score peaks at k = 3 though k = 2 also performs well.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigureCaptions.docx
- FIGURES1.png
- FIGURES4.png
- FIGURES5.png
- FIGURES6.png
- FIGURES7.png
- FIGURES8.png
- FIGURES16.png
- FIGURES17.png

- TABLES3.xlsx
- TABLES4.xlsx
- TABLES2.xlsx
- TABLES6.xlsx
- TABLES7.xlsx
- TABLES5.xlsx
- TABLES1.xlsx
- FIGURES2.png
- FIGURES3.png
- FIGURES9.png
- FIGURES10.png
- FIGURES11.png
- FIGURES12.png
- FIGURES13.png
- FIGURES14.png
- FIGURES15.png
- Tables.docx