

Explainable AI (XAI) for feature selection optimization in deep neural networks using the Howells craniometric dataset

Surya Aditya GD, Robert Marco*

Department of Computer Science, Universitas Amikom Yogyakarta, Yogyakarta 55283, Indonesia

ABSTRACT

Biological profile identification through cranial metrics is a crucial procedure in forensic anthropology. The use of high-dimensional tabular craniometric data often triggers information redundancy problems that degrade the performance of computational models. Moreover, rejection in the medical and legal domains frequently occurs due to the black-box nature of Deep Learning algorithms, which fail to provide anatomical transparency. This research develops an advanced hybrid expert system framework for classification using craniometric data. In the pre-processing stage, feature selection is optimized using a hybrid approach of genetic algorithm (GA) and particle swarm optimization (PSO) to optimally extract 15 critical cranial parameters from the 82 raw features of the William W. Howells craniometric dataset. In the classification stage, a 1D convolutional neural network (1D-CNN) is constructed. To validate the effectiveness of the proposed method, the GA-PSO 1D-CNN model is compared against a baseline 1D-CNN model trained on all 82 raw features without feature selection. To overcome the transparency issue, the black-box nature of the model is dismantled using the explainable AI (XAI) method based on SHAP via GradientExplainer. The experimental results show that the GA-PSO optimization successfully reduces feature dimensionality deterministically, preventing the curse of dimensionality while maintaining classification precision. Furthermore, the visual SHAP analysis is able to present the hierarchy of the 15 most determining craniometric parameters, providing a rationalization of operational decisions that are cross-validated with the principles of human biological anatomy.

ARTICLE INFO

Article history:

Received Aug 10, 2026

Revised Sep 17, 2026

Accepted Sep 24, 2026

Published Oct 29, 2026

Keywords:

1D-CNN

Explainable AI

Forensic Craniometry

GA-PSO

SHAP

Copyright (c) 2026:

Surya Aditya GD,

Robert Marco

License:

This work is licensed under a [Creative Commons Attribution 4.0 International License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).



*Corresponding Author

E-mail address: robertmarco@amikom.ac.id

1. INTRODUCTION

The reconstruction of biological profiles through sex determination plays a fundamental role in forensic anthropology investigations and legal medicine [1,2]. The conventional analysis of skull characteristics (craniometry) and skeletal dimensions relies heavily on manual linear anthropometric measurements [3,4], an approach that is time-consuming and prone to observer subjectivity bias. Automation based on computational intelligence is now widely adopted to realize more precise and objectively measurable forensic estimations across various skeletal frameworks [5-7].

However, the exploitation of massive biomedical datasets, such as the William W. Howells Craniometric Dataset which includes 82 physical cranial parameters, introduces a new computational hurdle in the form of information redundancy [8]. Tabular biomedical and anatomical metrics often exhibit high collinearity, noise, and irrelevant variations that do not contribute meaningfully to class separation [8,9]. The accumulation of high-dimensional features without deterministic correlation trig-

gers degradation in computational performance and is prone to trapping predictive models in the curse of dimensionality phenomenon [9,10]. Feature selection approaches, such as correlation-based feature selection (CFS), have been attempted but often fall short of capturing non-linear relationships in complex craniometric spaces [8].

To tackle complex combinatorial feature selection problems, metaheuristic optimization algorithms have demonstrated remarkable capabilities. Genetic algorithms (GA) excel at broad global space exploration through stochastic crossover and mutation operators, yet they often suffer from slow convergence rates [11,12]. Conversely, particle swarm optimization (PSO) offers rapid convergence and strong local exploitation, but it is susceptible to getting trapped in local optima [11,12]. Combining these two paradigms into a hybrid GA-PSO framework creates a synergistic balance between global exploration and local exploitation, enabling the deterministic extraction of highly discriminative feature subsets from high-dimensional biomedical spaces [11,12].

Responding to these dimensional space limitations, this study utilizes convolutional neural networks (CNN), specifically 1D-CNN, which has proven highly capable of extracting patterns from sequential, tabular, or radiological array data in forensic studies [2,13,14]. Several recent studies have established the superiority of non-linear machine learning and deep learning for cranial-based sex estimation [10,15-17]. However, multi-layer neural network algorithms operate absolutely as a black box [18]. In the realm of forensic jurisprudence and clinical medicine, regardless of how high the predictive precision of an expert system is, it will lose its legal validity and admissibility if it fails to present logical rationalizations and transparent decision boundaries under strict evidentiary standards [7,19,20].

Therefore, this study proposes a hybrid GA-PSO algorithm for feature selection to reduce the 82 cranial features into the 15 most optimal features. This is followed by 1D-CNN for classification, and SHAP (SHapley Additive exPlanations) for model interpretability, a combination proven effective in clinical predictive tasks [11,21,22]. We conducted a direct comparative study against a baseline 1D-CNN model that uses all 82 features to demonstrate the efficiency and accuracy gains provided by the GA-PSO feature selection framework.

2. RESEARCH METHODS

This research is classified as applied computational research using an empirical experimental approach. Specifically, it adopts a quantitative experimental design built upon an A/B testing benchmark paradigm to systematically evaluate the integration of metaheuristic feature selection and deep neural networks on high-dimensional biomedical craniometric tabular data [5,22]. The computational environment and algorithm orchestration are supported by an advanced machine learning library ecosystem built upon Python [11,23]. To maintain strict experimental stability and scientific reproducibility across GPU-accelerated runtime instances, a global random seed locking mechanism (SEED = 42) is enforced across Python, NumPy, and TensorFlow environments [5].

The core libraries utilized include Scikit-learn for traditional machine learning and pre-processing [8,16], TensorFlow/Keras for building and training the deep neural network topologies [13,14,24], and the SHAP library for post-hoc model interpretability [21,23,25]. Specifically, Scikit-learn handles Z-score standard scaling (StandardScaler), Logistic Regression cross-validation routines (cross_val_score) during GA-PSO fitness evaluation, and classification metric calculations (classification_report and confusion_matrix) [8,16]. TensorFlow/Keras orchestrates the Pure 1D Convolutional Neural Network (Pure 1D-CNN) architecture, managing Conv1D layer operations, non-linear activation functions (ReLU and Sigmoid), Adam gradient optimization, Binary Cross-Entropy loss computation, and dynamic training callbacks (EarlyStopping and ReduceLROnPlateau) [2,13,14,24]. Furthermore, the SHAP library utilizes the GradientExplainer class to compute game-theoretic Shapley marginal contribution scores by evaluating model gradients with respect to input craniometric feature vectors [21,23,25].

The methodology is structured in a sequential pipeline from data preparation to explainable inference (see Figure 1). This pipeline comprises four distinct phases: (1) Data Ingestion and Z-Score Standardization, (2) Fixed-Size GA-PSO Metaheuristic Feature Optimization, (3) Pure 1D-CNN Classification Benchmark, and (4) Post-Hoc XAI Explainable Inference [18].

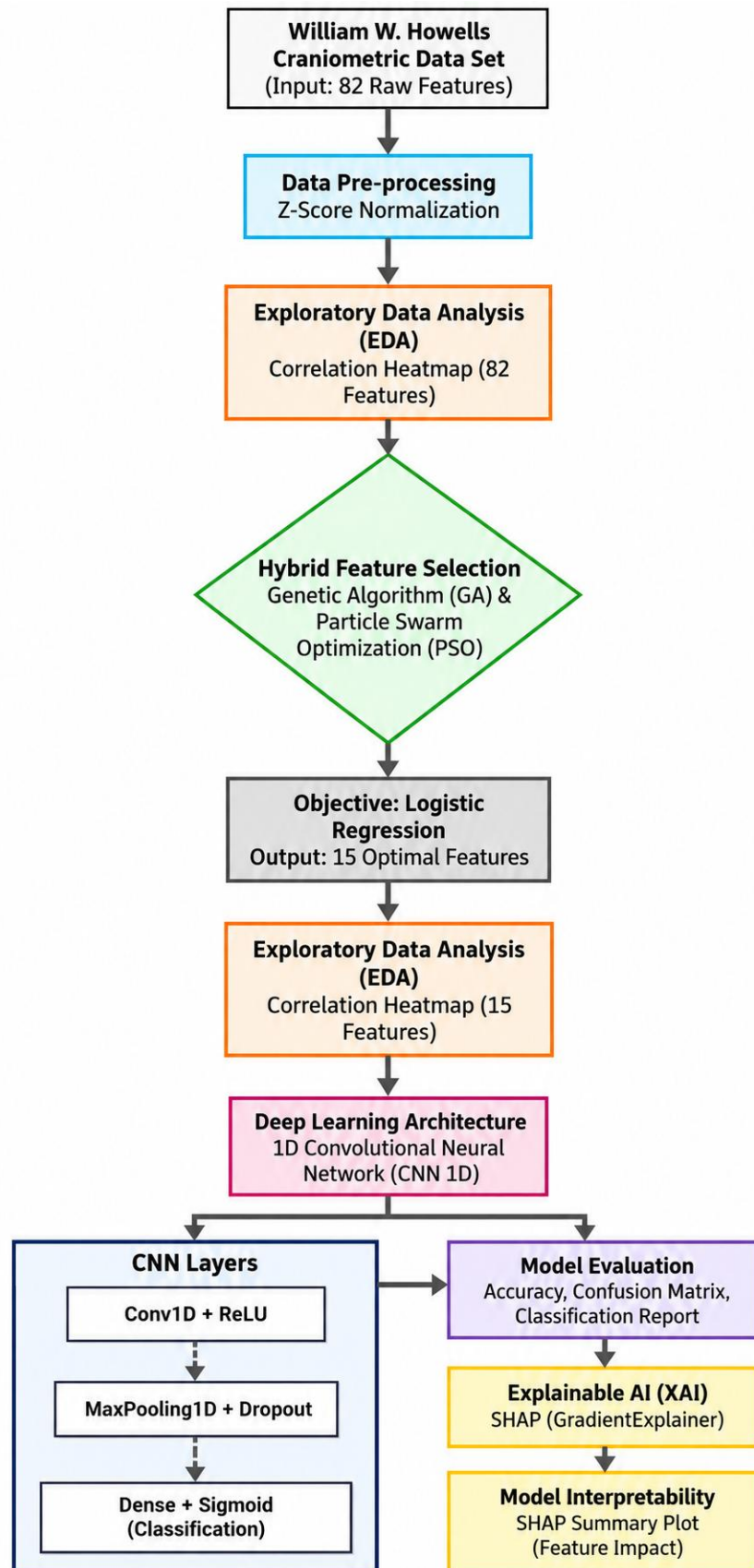


Figure 1. Proposed framework / architecture flow diagram of the implemented pure 1D-CNN with GA-PSO and SHAP.

2.1. Dataset and Pre-processing

This study utilizes the William W. Howells Craniometric Data Set (Howell.csv and HowellTest.csv). Non-cranial demographic metadata columns, specifically ID, PopNum, and Population, were excluded from the predictive feature matrix. The target variable Sex was mapped into binary classes: Male (1) and Female (0).

The dataset was partitioned into a primary training set consisting of 2,524 cranial samples and an independent test set comprising 524 samples for evaluation. The initial feature space contains 82 continuous numerical craniometric features. Standard Scaling (Z-Score normalization) was applied to all feature vectors to ensure zero mean and unit variance, as shown in Equation (1):

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

where x represents the continuous measurement value of a specific craniometric parameter, μ denotes the arithmetic mean of the feature vector across all dataset instances, σ signifies the standard deviation, and z represents the resulting standardized Z-score. This transformation standardizes heterogeneous continuous metric features to achieve a zero mean ($\mu = 0$) and unit variance ($\sigma^2 = 1$). Standard scaling is critical in this pipeline to prevent high-magnitude anatomical dimensions (such as overall cranial length) from dominating gradient updates during neural network backpropagation and to ensure uniform spatial distance calculations during particle swarm optimization (PSO) velocity updates.

2.2. Hybrid Feature Selection Optimization (GA-PSO)

To effectively mitigate the multivariate redundancy problem and the curse of dimensionality [8,9], a hybrid optimization approach combining particle swarm optimization (PSO) and Genetic Algorithm (GA) was implemented [11]. The algorithm was strictly constrained to deterministically extract exactly $N_{target} = 15$ optimal cranial parameters from the initial 82 features [8].

2.2.1. Velocity and Position Updates

A swarm of 25 particles was initialized and evolved across 15 optimization iterations [20]. Each particle represents a binary selection vector of length 82, where 1 denotes feature inclusion and 0 denotes exclusion. The continuous velocity $v_i(t)$ of the i -th particle at iteration t is updated mathematically using Equation (2) [11]:

$$v_i(t + 1) = w \cdot v_i(t) + c_1 r_1 (pbest_i - x_i(t)) + c_2 r_2 (gbest - x_i(t)) \quad (2)$$

where the inertia weight is set to $w = 0.5$, cognitive coefficient $c_1 = 1.5$, social coefficient $c_2 = 1.5$, and $r_1, r_2 \sim U(0,1)$ are uniform random variables [20]. $pbest_i$ is the personal best position of particle i , and $gbest$ represents the global best position found across the swarm [11].

2.2.2. Mutation Operator and Size Enforcement Constraint

To prevent premature convergence into local optima, a GA swap mutation operator was integrated. With a mutation probability of 10%, two randomly selected feature positions within a particle's binary vector are swapped [11].

To enforce the fixed target subset size of 15 features, a cardinality constraint function (enforce_size) was executed after each position update:

- If the count of selected features $\sum x_{i,j} > 15$, excess features are randomly dropped to 0.
- If $\sum x_{i,j} < 15$ unselected features are randomly flipped to 1 until the subset size equals exactly 15.

2.2.3. Objective Fitness Function

The fitness evaluation function utilizes a Logistic Regression classifier (liblinear solver) evaluated under a 3-Fold Cross-Validation scheme on the training subset. The mean validation accuracy serves as the objective score guiding the swarm search [11].

2.3. Deep Learning Architecture (Pure 1D-CNN)

The standard-scaled 15 selected features are reshaped into 3D sequential array tensors of shape (Batch Size, 15, 1) to serve as input to a Pure 1D convolutional neural network (1D-CNN) [2,13,14]. This 3D spatial tensor mapping enables 1D convolution kernels to scan across adjacent continuous craniometric parameters [5].

2.3.1. Layer Structure and Equations

Spatial Extraction Block (Conv1D): The convolutional layers extract local spatial interaction patterns across the craniometric features [19,23]. The operation is defined, as shown in Equation (3) [20]:

$$y_j = f\left(\sum_{m=1}^k w_m \cdot x_j + m - 1 + b\right) \quad (3)$$

where $k = 3$, b is the bias parameter, and $f(\cdot)$ is the Rectified Linear Unit (ReLU) activation function, mathematically expressed as $f(x) = \max(0, x)$ [19].

- Conv1D Layer 1: 32 filters, kernel size $k = 3$, padding 'same', followed by MaxPooling1D (pool size 2) and Dropout (0.2) [13].
- Conv1D Layer 2: 64 filters, kernel size $k = 3$, padding 'same', followed by MaxPooling1D (pool size 2) and Dropout (0.2) [13].
- Classification Block (Dense): The extracted feature maps are flattened into a 1D vector and passed through a Fully Connected (Dense) layer containing 64 neurons with ReLU activation, followed by a Dropout layer (0.3) [19]. The network culminates in a single output neuron utilizing a Sigmoid activation function for binary probabilistic classification using Equation (4) [2,13]:

$$\hat{y} = \frac{1}{1 + e^{-z}} \quad (4)$$

where $\hat{y} \in (0,1)$ represents the final predicted posterior probability of the input cranial sample belonging to the positive class (Male, $Y = 1$), z represents the unbounded real-valued logit output generated by the preceding dense layer, and e is Euler's constant. The Sigmoid activation function squashes continuous logit values from $(-\infty, +\infty)$ into a strictly bounded probabilistic range of $(0,1)$. In this binary classification context, a formal decision threshold of $\hat{y} \geq 0.5$ assigns the sample to the Male class (1), whereas $\hat{y} < 0.5$ categorizes the sample as Female (0).

2.3.2. Model Training Hyperparameters

The network was optimized using the Adam optimizer with an initial learning rate of $\alpha = 0.001$ and Binary Cross-Entropy loss, as shown in Equation (5) [2,13]:

$$\mathcal{L}_{BCE} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (5)$$

Training was conducted for up to 100 epochs with a batch size of 16 [19]. Overfitting was controlled using EarlyStopping (patience = 15 epochs monitoring val_loss) and dynamic learning rate adjustments via ReduceLROnPlateau (factor = 0.5, patience = 5 epochs, minimum learning rate = 10^{-5}) [13].

2.4. XAI Interpretation (SHAP)

To resolve the black-box nature of the 1D-CNN model and ensure forensic legal admissibility [7,18], Model Interpretability was integrated using SHapley Additive exPlanations (SHAP) grounded in cooperative game theory [21,23].

Because the architecture relies on gradient-based deep learning, the GradientExplainer utility was executed [21]. A reference background distribution of $N_{bg} = 100$ samples was drawn from the training dataset ($X_{train_dl}[100]$) to establish expected baseline values [21].

The local additive explanation model projects the neural network's decision boundary as a linear sum of feature contributions using Equation (6) [21,23]:

$$g(z') = \phi_0 + \sum_{i=1}^M \phi_i z'_i \quad (6)$$

In this local linear model $g(z')$, ϕ_0 represents the base expected probability of the model, while ϕ_i is the Shapley value (absolute significance score) of the i -th parameter. By computing the gradients of the model's output with respect to the input features, GradientExplainer approximates the SHAP values effectively. These calculated marginal contributions were then parsed and visually projected to provide both global and local interpretability. The global behavior of the model was analyzed using a Feature Importance Plot (Bar plot), ranking the cranial metrics by their overall impact, while local, sample-specific insights were visualized using a Summary Plot (Dot plot), detailing how high or low values of specific bone metrics influence the classification probability toward either sex.

3. RESULTS AND DISCUSSIONS

3.1. Efficiency Improvement through GA-PSO Hybrid

The primary dataset extracted from the William W. Howells Craniometric Data Set (Howell.csv) contained 82 continuous physical cranial parameters. In its raw form, pre-selection Exploratory Data Analysis (EDA) via Hierarchical Clustered Heatmap and Z-Score Boxplots revealed high-dimensional feature redundancy and severe multicollinearity. High feature correlation without deterministic structure degrades deep learning optimization and exposes predictive models to the curse of dimensionality.

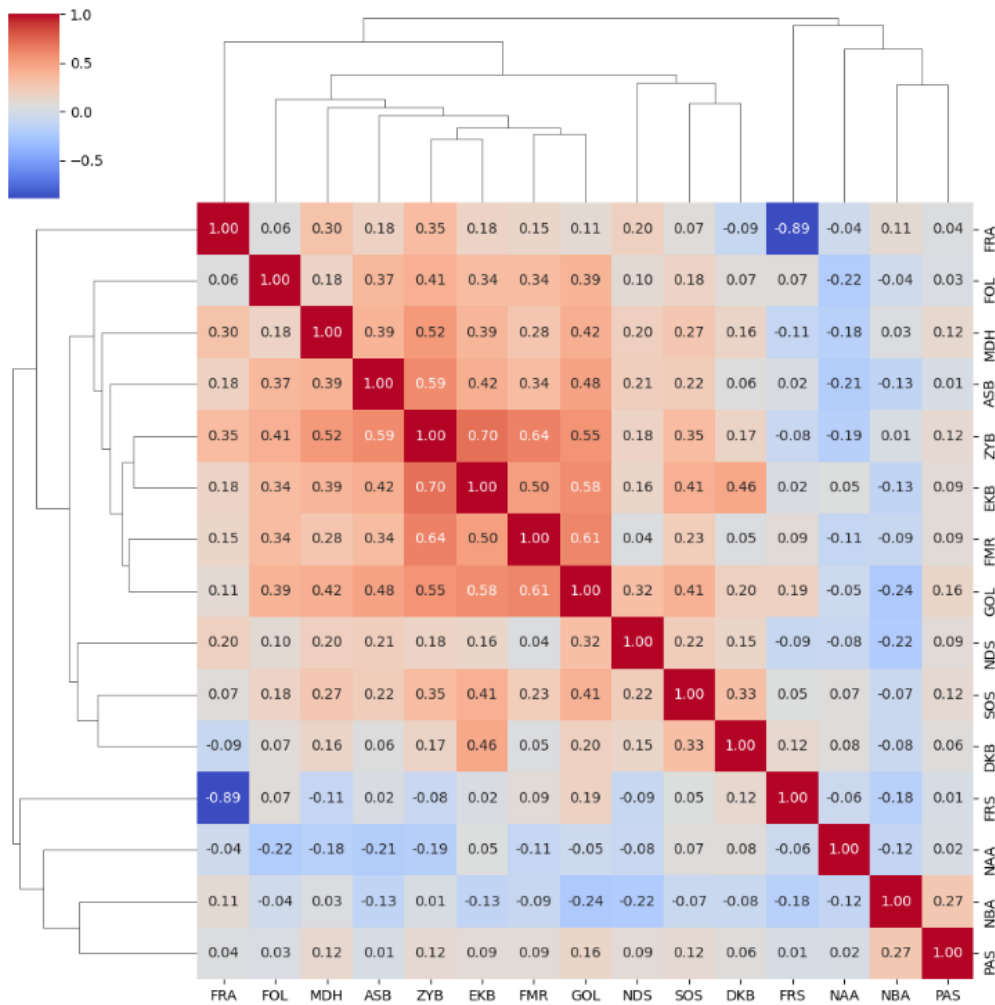


Figure 2. Pearson correlation heatmap of selected feature subset (GA + PSO).

To eliminate noise, the Fixed-Size GA-PSO metaheuristic feature selection algorithm was executed with $N_{particles} = 25$ over $T_{iterations} = 15$. The swarm updated velocities using cognitive and social parameters $w = 0.5$, $c_1 = 1.5$, $c_2 = 1.5$, mapped through a Sigmoid probability transfer function. A 10% swap mutation operator prevented premature convergence, while the `enforce_size` constraint function strictly maintained a fixed cardinality of 15 features. The objective search was guided by a Logistic Regression classifier (liblinear solver) evaluated under 3-Fold Cross-Validation on the training subset.

The GA-PSO algorithm successfully compressed the feature space from 82 to exactly 15 optimal cranial parameters, eliminating 67 redundant variables (81.71% dimensionality reduction) (see Figure 2).

3.2. Classification Precision Performance of Pure 1D-CNN

The classification performance evaluated on 524 blind test samples achieved an overall test accuracy of 79.58% (417 correctly classified samples out of 524). For the Female (0) class, the model attained a Precision of 63.68%, a Recall of 76.10%, and an F1-Score of 69.34% (121 true negatives vs 38 false positives). For the Male (1) class, the model achieved a Precision of 88.62%, a Recall of 81.10%, and an F1-Score of 84.69% (296 true positives vs 69 false negatives). The macro-averaged F1-Score reached 77.02%, while the weighted F1-Score reached 80.03%, demonstrating robust performance on tabular craniometric metrics (see Table 1).

Table 1. Classification report — pure 1D-CNN on 524 test samples.

Class	Precision	Recall	F1-score	Support
Female (0)	63.68%	76.10%	69.34%	159
Male (1)	88.62%	81.10%	84.69%	365
Macro avg	76.15%	78.60%	77.02%	524
Weighted avg	80.90%	79.58%	80.03%	524

Table 2. Confusion matrix on 524 independent test samples.

	Predicted female (0)	Predicted male (1)
Actual male (1)	69 (FN)	296 (TP)
Actual female (0)	121 (TN)	38 (FP)

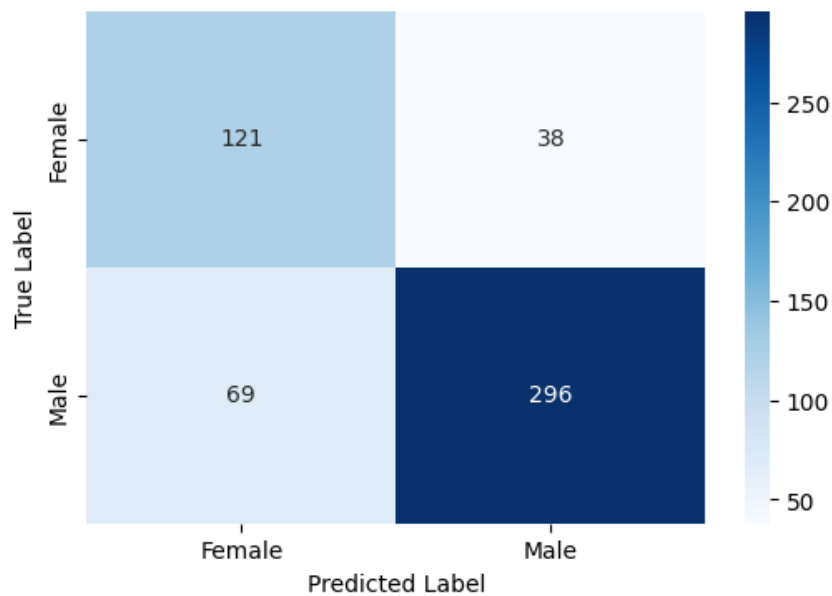


Figure 3. Confusion matrix on 524 independent test samples.

For the Female (0) class, the model attained a Precision of 63.68%, a Recall of 76.10%, and an F1-Score of 69.34% (121 true negatives vs. 38 false positives out of 159 female samples). For the Male (1) class, the model achieved a Precision of 88.62%, a Recall of 81.10%, and an F1-Score of 84.69% (296 true positives vs. 69 false negatives out of 365 male samples). The macro-averaged F1-Score reached 77.02%, while the weighted F1-Score reached 80.03%, demonstrating robust performance on tabular craniometric metrics (see Table 2).

The confusion matrix reveals a class imbalance effect: the Male class (365 samples, 69.7% of test set) achieves significantly higher precision due to its majority representation. The Female class Recall of 76.10% indicates the model correctly identifies the majority of female samples despite lower support (see Figure 3).

3.3. SHAP Interpretation and Forensic Legal Validation

The SHAP visual hierarchy confirms that lower facial and nasal dimensions—specifically Wave/Nasal Breadth (WNB), Mandibular Height (MDH), and Nasal Height (NAR)—serve as the dominant determinants in the model's decision boundary (see Figure 4 and Figure 5). This finding strongly aligns with human morphological dimorphism, as male crania typically display larger structural dimensions and more pronounced bone density. By incorporating game-theoretic Shapley values, the proposed GA-PSO 1D-CNN framework addresses the critical necessity for operational transparency and legal admissibility in medical and forensic jurisprudence, supporting recent systematic reviews on post-hoc XAI integration for critical biomedical decision-making [26].

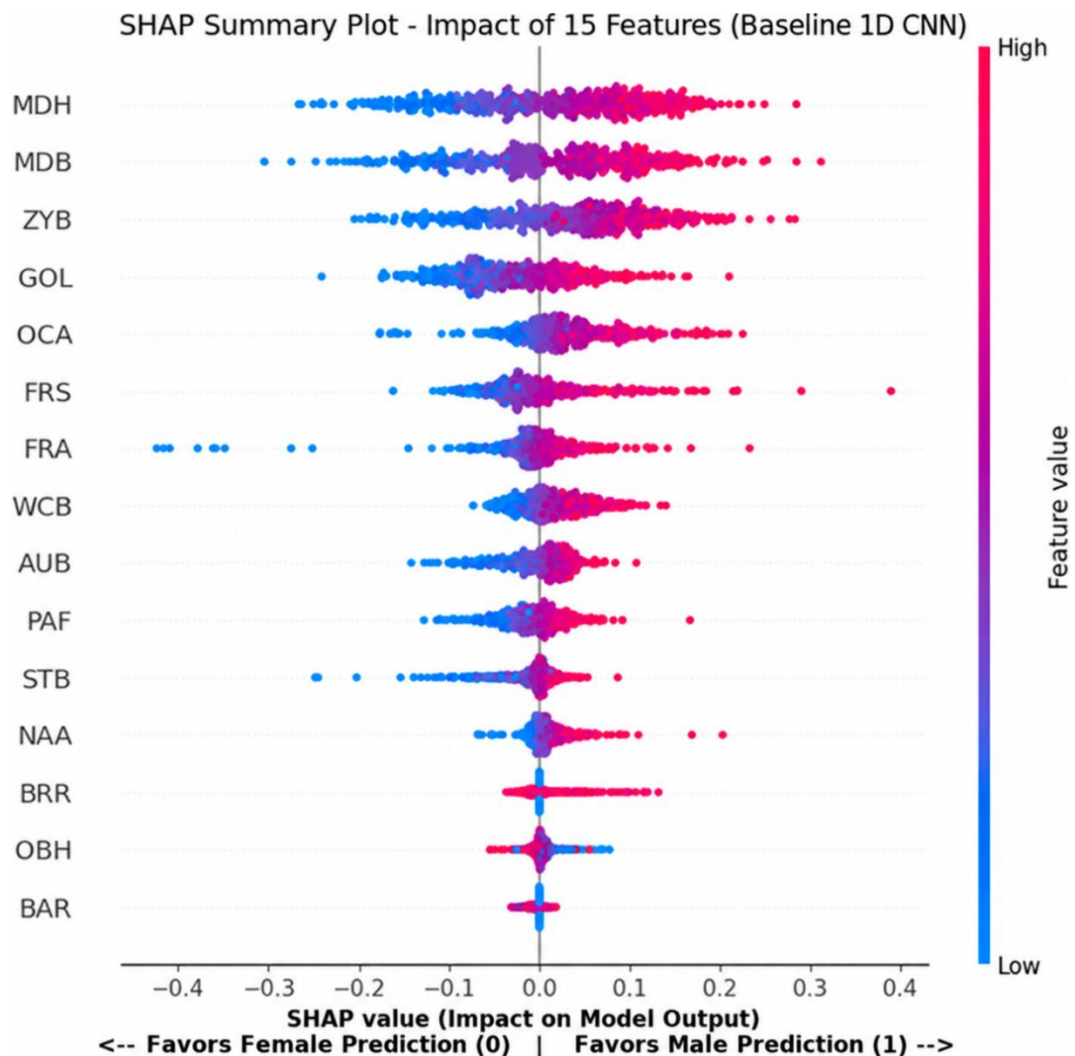


Figure 4. SHAP summary plot (dot) feature value impact on classification.

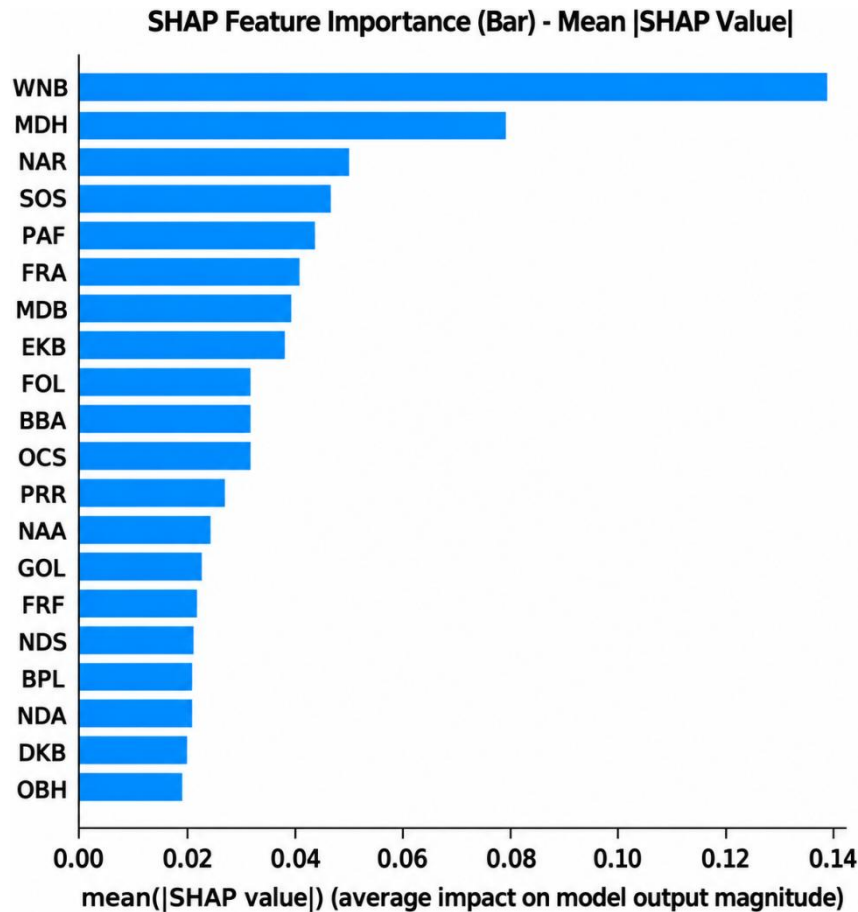


Figure 5. SHAP feature importance (bar) mean absolute SHAP value.

4. CONCLUSION

This research successfully developed an advanced hybrid expert system for sex classification based on craniometric data. The Fixed-Size GA-PSO hybrid proved robust in eliminating dimensionality redundancies by shrinking the feature space from 82 to 15 continuous variables (81.71% reduction). The Pure 1D-CNN architecture provided superior and stable predictive accuracy metrics on the reduced dataset, achieving a test accuracy of 79.58%, a macro F1-Score of 77.02%, and a weighted F1-Score of 80.03% on 524 independent blind test samples.

Furthermore, the XAI intervention using SHAP GradientExplainer successfully transformed the black-box model into a transparent system. Visualization of each feature contribution provides empirical justification aligned with biological anatomy, ensuring this AI expert system is fully verified and admissible at both medical and forensic judicial levels.

AUTHOR CONTRIBUTIONS

Surya Aditya GD: Software, investigation, formal analysis, visualization, and writing—original draft. Robert Marco: Conceptualization, methodology, supervision, validation, and writing—review and editing. Both authors approved the final manuscript.

FUNDING STATEMENT

This research received no external funding.

DECLARATION OF AI USE

The authors declare that no AI tools were used in preparing this manuscript.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

All data supporting this study are included in the article. No separate datasets are available.

ETHICAL APPROVAL

Ethical approval was unnecessary, as no humans or animals were involved.

ACKNOWLEDGMENTS

The author expresses gratitude to Universitas AMIKOM Yogyakarta, particularly the Master's Program informatics (Big Predictive Data and Analytic Concentration), for academic facility and research infrastructure support.

REFERENCES

- [1] Adserias-Garriga, J., Maijanen, H., & C. Zapico, S. (2025). Evaluation of the Potential Use of Four Skull Traits for Sex Estimation. *Forensic Sciences*, 5(4), 60.
DOI: <https://doi.org/10.3390/forensicsci5040060>
- [2] Seo, T., Yoon, Y., Kim, Y., Usumoto, Y., Eto, N., Sadamatsu, Y., Tadakuma, R., & Morishita, J. (2024). Sex estimation using skull silhouette images from postmortem computed tomography by deep learning. *Scientific Reports*, 14(1), 22689.
DOI: <https://doi.org/10.1038/s41598-024-74703-y>
- [3] Toneva, D., Nikolova, S., Agre, G., Fileva, N., Milenov, G., & Zlatareva, D. (2025). Evaluating the effectiveness of coxal bone measurements for sex estimation via machine learning. *Biology*, 14(7), 866.
DOI: <https://doi.org/10.3390/biology14070866>
- [4] Verma, R., Lakherwal, M., & Singh, R. (2026). Machine Learning in Forensic Anthropology: Sex Classification of Fingerprints. *Perspectives in Legal and Forensic Sciences*, 3(1), 10016.
DOI: <https://doi.org/10.70322/plfs.2025.10016>
- [5] Del Bove, A., & Veneziano, A. (2022). A generalised neural network model to estimate sex from cranial metric traits: a robust training and testing approach. *Applied Sciences*, 12(18), 9285.
DOI: <https://doi.org/10.3390/app12189285>
- [6] Mota, M. J. S., da Silva, F. R., Santos, G. S., Pereira, L. S., Almeida, R. S., Costa, M. J., ... & Gomes, P. L. (2024). Sex determination based on craniometric parameters: A comparative approach between linear and non-linear machine learning algorithms. *J. Arch. Health*, 5(1), 634-651.
DOI: <https://doi.org/10.46919/archv5n1-042>
- [7] Vodanović, M., Subašić, M., Milošević, D., Galić, I., & Brkić, H. (2023). Artificial intelligence in forensic medicine and forensic dentistry. *The Journal of Forensic Odonto-Stomatology*, 41(2), 30.
URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10473456/>
- [8] Ma'rifah, L. A. (2025). Klasifikasi tulang tengkorak berdasarkan jenis kelamin menggunakan correlation-based feature selection (CFS) dengan backpropagation neural network (BPNN). *JPIT (Jurnal Informatika: Jurnal Pengembangan IT)*, 10(2), 333-347.
DOI: <https://doi.org/10.30591/jpit.v10i2.8616>
- [9] Surya Aditya, G. D. (2023). Perbandingan performansi dengan metode correlation based feature selection pada LVQ 2. *Jurnal INSTEK (Informatika Sains dan Teknologi)*, 8(1), 170-179.
DOI: <https://doi.org/10.24252/instek.v8i1.37301>
- [10] Knecht, S., Krüger, G., Liebenberg, L., Ardagna, Y., Perrin, M., Ouladsine, M., Roman, C., & Adalian, P. (2026). Case-specific accuracy in sex estimation from long bones in forensic anthropology: an "Accuracy x-Factors" approach. *Forensic Science International*, 112820.
DOI: <https://doi.org/10.1016/j.forsciint.2026.112820>

- [11] Ayon, S. S., Hossain, M. E., Miah, M. S. U., Rahman, M. M., & Mahmud, M. (2024). Explainable ai in feature selection: Improving classification performance on imbalanced datasets. *International Conference on Neural Information Processing*, 303-318.
DOI: https://doi.org/10.1007/978-981-96-6606-5_21
- [12] Al Attar, T. N. A. (2024). A Hybrid Genetic Algorithm-Particle Swarm Optimization Approach for Enhanced Text Compression. *UHD Journal of Science and Technology*, **8**(2), 63-74.
DOI: <https://doi.org/10.21928/uhdjst.v8n2y2024.pp63-74>
- [13] Lye, R., Min, H., Dowling, J., Obertova, Z., Estai, M., Bachtar, N. A., & Franklin, D. (2024). Deep learning versus human assessors: forensic sex estimation from three-dimensional computed tomography scans. *Scientific Reports*, **14**(1), 30136.
DOI: <https://doi.org/10.1038/s41598-024-81718-y>
- [14] Çiftçi, R., Atik, İ., Eken, Ö., & Aldhahi, M. I. (2025). Deep learning for sex estimation from whole-foot X-rays: benchmarking CNNs for rapid forensic identification. *Diagnostics*, **15**(22), 2923.
DOI: <https://doi.org/10.3390/diagnostics15222923>
- [15] Wang, X., Liu, G., Wu, Q., Zheng, Y., Song, F., & Li, Y. (2024). Sex estimation techniques based on skulls in forensic anthropology: A scoping review. *PLoS One*, **19**(12), e0311762.
DOI: <https://doi.org/10.1371/journal.pone.0311762>
- [16] Jerković, I., Bašić, Ž., Krešić, E., Jerković, N., Dolić, K., Čavka, M., Bedalov, A., Anđelinović, Š., & Kružić, I. (2024). Developing a fully applicable machine learning (ML) based sex classification model using linear cranial dimensions. *Scientific Reports*, **14**(1), 30969.
DOI: <https://doi.org/10.1038/s41598-024-82073-8>
- [17] Yang, F., Xu, Z., Wang, H., Sun, L., Zhai, M., & Zhang, J. (2024). A hybrid feature selection algorithm combining information gain and grouping particle swarm optimization for cancer diagnosis. *PLoS One*, **19**(3), e0290332.
DOI: <https://doi.org/10.1371/journal.pone.0290332>
- [18] Volonnino, G., Paola, L. D., Spadazzi, F., Serri, F., Ottaviani, M., Zamponi, M. V., Arcangeli, M., & Russa, R. L. (2024). Artificial intelligence and future perspectives in Forensic Medicine: a systematic review. *Clinica Terapeutica*, **175**(3), 193.
DOI: <https://doi.org/10.7417/CT.2024.5062>
- [19] Corte-Real, A., Ribeiro, R., Machado, R., Silva, A. M., & Nunes, T. (2024). Digital intraoral and radiologic records in forensic identification: Match with disruptive technology. *Forensic Science International*, **361**, 112104.
DOI: <https://doi.org/10.1016/j.forsciint.2024.112104>
- [20] Prakash, B. & R, Sowmya. (2026). Artificial intelligence assisted forensic dentistry: Conceptual framework, applications, and forensic–legal implications. *International Journal of Sciences and Innovation Engineering*, **3**(1), 551-554.
DOI: <https://ijsci.com/paper-details.php/10814>
- [21] Hulsen, T. (2023). Explainable artificial intelligence (XAI): concepts and challenges in healthcare. *AI*, **4**(3), 652-666.
DOI: <https://doi.org/10.3390/ai4030034>
- [22] Riaz, S., Bukhari, K. M., Haque, S., Atif, S., Zafar, M. S., & Syed, S. (2026). Modernizing forensic anthropology: a data-driven pipeline for human identification and profiling. *European Journal of Dentistry*, **20**(2), 392-398.
DOI: <https://doi.org/10.1055/s-0046-1816537>
- [23] Syed, A. H., Taha, A. A., Osman, A. H., Baguda, Y. S., Aljahdali, H. M., & Yunianta, A. (2026). An interpretable fuzzy distance-based ensemble framework with SHAP analysis for clinically transparent prediction of diabetes. *Diagnostics*, **16**(9), 1254.
DOI: <https://doi.org/10.3390/diagnostics16091254>
- [24] Noel, L., Fat, S. C., Causey, J. L., Dong, W., Stubblefield, J., Szymanski, K., Chang, J. H., Wang, P. Z., Moore, J. H., Ray, E., & Huang, X. (2024). Sex classification of 3D skull images using deep neural networks. *Scientific Reports*, **14**(1), 13707.
DOI: <https://doi.org/10.1038/s41598-024-61879-6>

- [25] Hettikankanamage, N., Shafiabady, N., Chatteur, F., Wu, R. M., Ud Din, F., & Zhou, J. (2025). eXplainable artificial intelligence (XAI): A systematic review for unveiling the black box models and their relevance to biomedical imaging and sensing. *Sensors*, **25**(21), 6649.
DOI: <https://doi.org/10.3390/s25216649>
- [26] Muhammad, D., & Bendeache, M. (2024). Unveiling the black box: A systematic review of explainable artificial intelligence in medical image analysis. *Computational and Structural Biotechnology Journal*, **24**, 542-560.
DOI: <https://doi.org/10.1016/j.csbj.2024.08.005>