



OPEN Sex inference based on convolutional neural network analysis of fingerprint data

Yibo Zhang¹✉, Shengfeng Cui¹, Jiaqi Li¹, Xinlei Liu¹, Shumeng Li¹, Yaoru Jiang², Fang Wang³ & Weicheng Liu⁴

The analysis of fingerprint features for inferring biological sex is a growing area of research in forensic science. This study presents a lightweight and well-validated convolutional neural network (CNN) as an alternative approach for this task. A dedicated dataset of 1,000 fingerprint images was collected from 200 volunteers (100 males and 100 females). To ensure rigorous evaluation of generalisation ability, an independent test set of 100 images from an additional 20 volunteers (10 males and 10 females) was held out for final assessment. The proposed CNN, featuring a dual-convolutional-layer architecture, was optimised using a cross-entropy loss function and the Adam optimiser. It achieved a validation accuracy of 91.00% and a test accuracy of 95.00%, with AUC values of 0.974 and 0.983, respectively. Supplementary fivefold cross-validation on the development cohort yielded a mean accuracy of 90.60% (SD: 2.04%), confirming stable performance. Class activation mapping (CAM) was employed to visualise the model's focus regions, enhancing interpretability and providing insights into biometric relevance. These results demonstrate that the model compares favourably with traditional methods, suggesting its potential as an efficient and reliable complementary tool for forensic identification.

Keywords Fingerprint examination, Sex prediction, Convolutional neural networks, Biometric characterization

Personal identification has long been a cornerstone of forensic science, with various biometric modalities—including fingerprints, iris patterns, and facial features—serving as reliable means of identity verification^{1,2}. Among these, fingerprints are particularly valuable in forensic evidence analysis owing to their unique individual characteristics, lifelong stability, and tendency to leave latent impressions on contact surfaces³. As such, they remain one of the most fundamental and critical forms of evidence in criminal investigations. The distinctiveness of fingerprints arises primarily from the configuration of ridge patterns, which encompass level-one features (overall ridge flow), level-two minutiae (e.g., ridge endings, bifurcations), and level-three intraridge details, all of which demonstrate stable interindividual variability over a lifetime⁴. Previous studies have established correlations between fingerprint ridge density, pattern type, and sex, noting that males typically exhibit lower mean ridge density than females do, thereby providing a theoretical basis for sex inference^{5–7}. Specific regional ridge characteristics have been validated as potential markers for sex classification^{8,9}.

Conventional fingerprint identification predominantly relies on manual comparisons of minutia points (e.g., ridge endings and bifurcations), a process that is inherently subjective and labour-intensive, thus limiting both efficiency and reproducibility^{10,11}. Although automated fingerprint identification systems (AFISs) are extensively utilised in law enforcement, they remain ineffective for suspects absent from existing databases, underscoring the need for alternative approaches capable of inferring biometric attributes¹². In recent years, deep learning techniques—particularly convolutional neural networks (CNNs)—have shown promise in fingerprint analysis^{13–16}. These methods transcend traditional feature engineering approaches (e.g., ridge density and pattern classification) by autonomously learning discriminative representations, thereby enabling direct modelling of associations between fingerprint characteristics and sex¹⁷. However, many current studies are constrained by limited sample sizes (often fewer than 500 images) and challenges in model generalisability, which may hinder the effective learning of complex fingerprint features.

In light of these limitations, the present study developed a CNN-based framework for sex inference as a lightweight and well-validated alternative within existing approaches. A dataset comprising 1,000 high-

¹School of Forensic and Technology, Zhengzhou Police University, Zhengzhou City, Henan Province, China. ²Yuzhou Public Security Bureau, Yuzhou City, Henan Province, China. ³College of Animal Science, Shanxi Agricultural University, Taiyuan City, Shanxi Province, China. ⁴Guangzhou Railway Public Security Bureau, Huaihua City, Hunan Province, China. ✉email: 601887522@qq.com

resolution fingerprint samples (collected from 200 volunteers with balanced sex distributions) was constructed for this purpose. The proposed architecture integrates data augmentation strategies, a cross-entropy loss function optimised via the Adam algorithm, and class activation mapping to enhance the extraction of salient ridge features. By leveraging these techniques, this study aims to provide an efficient and interpretable solution for sex prediction in forensic contexts, offering potential practical value for criminal investigations and forensic casework.

Materials and methods

Data collection

A total of 220 volunteers aged 18–22 years were recruited through open channels after being fully informed of the study's objectives and procedures. Of these, 200 volunteers (100 males and 100 females) constituted the development cohort for model training and validation, whilst an additional 20 volunteers (10 males and 10 females) formed the independent test cohort for external validation assessment. Participation was entirely voluntary, and all individuals provided written informed consent prior to fingerprint collection, ensuring the protection of personal rights. Fingerprint acquisition was performed via a ZKTECO entropy-based fingerprint scanner in accordance with the standardized 500 DPI protocol to ensure procedural objectivity and consistency. This study collected right-hand fingerprints from all 220 volunteers (200 for the development cohort and 20 for independent testing), with all images consistently sized at 300 × 400 pixels. Postacquisition, images were systematically renamed: male samples were labelled with the initial "M" and female samples with "F," whereas the thumb, index, middle, ring, and little fingers were denoted by "A," "B," "C," "D," and "E," respectively (e.g., the right thumb fingerprint of male volunteer No. 001 was coded as "MA001").

In this study, we chose to collect only right-hand fingerprint samples for sex classification using a convolutional neural network, based on several key considerations. First, the decision was guided by practical applications. In forensic identification and biometric recognition, right-hand fingerprints are more frequently encountered, and impressions left at crime scenes are statistically more likely to originate from the right hand^{18,19}. Moreover, many fingerprint recognition systems, such as the Automated Fingerprint Identification System (AFIS), place greater emphasis on the processing of right-hand prints. Training models primarily on right-hand fingerprints therefore enhances both accuracy and efficiency in forensic contexts, making them more relevant for real-world applications. Second, from the perspective of standardisation, fingerprint ridge patterns vary markedly across different fingers²⁰. For deep learning models, the quality of the training data is often more critical than the sheer quantity^{21,22}. Restricting the dataset to right-hand fingerprints ensures greater consistency and reduces variability, allowing the model to focus on learning sex-related features. This improves training efficiency, reliability, and generalisability. Third, evidence from previous studies supports the superior performance of right-hand fingerprints in sex classification tasks. For example, Iloanusi and Ejiogu (ref. ²³) reported higher classification accuracy using right-hand prints compared with left-hand prints, whilst Qi et al. (ref. ²⁴) demonstrated that CNN models trained on right-hand data achieved better generalisation and higher accuracy. Finally, from the standpoint of feasibility and ethics, collecting fingerprints from one hand significantly improves efficiency, reduces the burden on participants, and ensures that both the quantity and quality of the data are maintained in line with ethical standards. Thus, models trained on right-hand fingerprint data are not only robust but also of greater practical and forensic value.

All collected data were stored on an encrypted server, with access restricted to authorized personnel. The research team strictly complied with data protection and privacy regulations, ensuring that no personal information or fingerprint data of the volunteers were disclosed to third parties.

Data preprocessing

Preprocessing is critical for enhancing CNN-based classification by standardizing inputs, improving convergence speed, stabilizing numerical computations, and improving model generalizability to real-world variations. Given that the original images were grayscale, min–max scaling was first applied to normalize pixel values within the range [0, 1], thereby reducing noise and unifying the data distribution.

Data augmentation techniques, as detailed in Table 1, including random rotation, translation, and flipping, were subsequently employed to increase variability and improve the model's feature extraction capability. The preprocessed dataset (200 volunteers, 1,000 images) was partitioned into training and validation sets at an 8:2 ratio using stratified random sampling based on sex, thereby ensuring balanced representation across both groups. A

Parameter	Value
rotation_range	± 15°
width_shift_range	± 10%
height_shift_range	± 10%
zoom_range	5%
horizontal flipping	50% probability
shear_range	10%
brightness_range	[0.9, 1.1]
fill_mode	nearest

Table 1. Data augmentation parameters.

batch size of 32 was configured to meet the study's computational requirements. To objectively assess the model's generalisation and predictive performance on unseen data, an independent test cohort was established. This cohort comprised 20 volunteers (10 males and 10 females) recruited separately following identical acquisition protocols, yielding 100 fingerprint images. These 20 individuals were entirely distinct from the 200-volunteer development cohort and were reserved exclusively for final external validation. The test dataset was introduced only after model training and validation were completed, thereby providing an unbiased estimate of the model's discriminative ability on new individuals. All data splits (training, validation, and test) were subject-wise, ensuring that fingerprint images from the same individual appeared in only one split, thus preventing any overlap between subject groups. This subject-wise partitioning strategy was maintained consistently across all validation approaches, including the supplementary fivefold cross-validation described below.

Supplementary cross-validation protocol

To further assess model robustness and confirm that performance was not dependent on the specific 8:2 data partition, supplementary fivefold stratified cross-validation was performed on the entire development cohort (200 volunteers, 1,000 images). The dataset was partitioned into five subject-wise folds—ensuring all images from an individual resided within a single fold—with each fold serving sequentially as the validation set (200 images, 40 volunteers) whilst the remaining four constituted the training set (800 images, 160 volunteers). This process was repeated five times (once for each fold) with a fixed random seed (seed = 42) to ensure reproducibility. Model architecture and hyperparameters remained identical to those used in the primary training protocol (Table 1), with only the data partitioning strategy varying across folds. This cross-validation assessed internal consistency and stability across different training-validation configurations, whilst the independent test set (100 images from 20 additional volunteers collected separately) was retained exclusively for external validation, providing an unbiased estimate of generalisability to unseen individuals. The cross-validation results, including fold-wise performance metrics, are presented in the Results section to demonstrate convergent evidence with the primary hold-out validation approach.

Model architecture

The CNN architecture designed for this study (Fig. 1) consisted of two convolutional layers (conv1 and conv2), pooling layers, and fully connected layers optimized for binary classification. The convolution layers used sequential 3×3 kernels with 16 filters in the first layer and 32 in the second.

The rectified linear unit (ReLU) activation function, defined in Eq. (1), was applied after each convolution:

$$\text{ReLU}(x) = \max(0, x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \quad (1)$$

This nonlinear activation preserves positive signals ($x > 0$) whilst suppressing negative inputs ($x \leq 0$), thereby enhancing the network's representational capacity.

A 2×2 max pooling layer with a stride of 2 was employed to downsample the feature maps, reducing the spatial dimensions by half, lowering the computational cost, and mitigating overfitting by prioritizing feature presence over spatial precision. Following feature extraction, the flattened maps were connected to a fully connected layer of 128 units to integrate the learned features for classification.

Model optimization utilized the Adam optimizer with an initial learning rate of 0.001 over 100 epochs. An early stopping criterion was implemented, terminating training if validation loss failed to improve for 10 consecutive epochs, thereby reducing overfitting risk.

Evaluation metrics

Model performance was evaluated in terms of accuracy, loss, and the area under the receiver operating characteristic curve (AUC).

The accuracy, which is suitable for this study's balanced dataset (male-to-female ratio of 1:1), quantifies correct classifications as a proportion of total samples and is defined by Eq. (2):

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\% \quad (2)$$

where TP and TN denote true positives and true negatives, respectively, whilst FP and FN represent false positives and false negatives, respectively.

Loss, reflecting prediction-label divergence, was computed via the cross-entropy loss function (Eq. 3):

$$\text{Loss} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (3)$$

where y_i is the true label and p_i is the predicted probability.

The AUC was used to quantify the model's discriminative ability across thresholds and was calculated via Eq. (4):

$$\text{AUC} = \int_0^1 \text{TRU}(\text{FPR}^{-1}(x)) \, dx \quad (4)$$

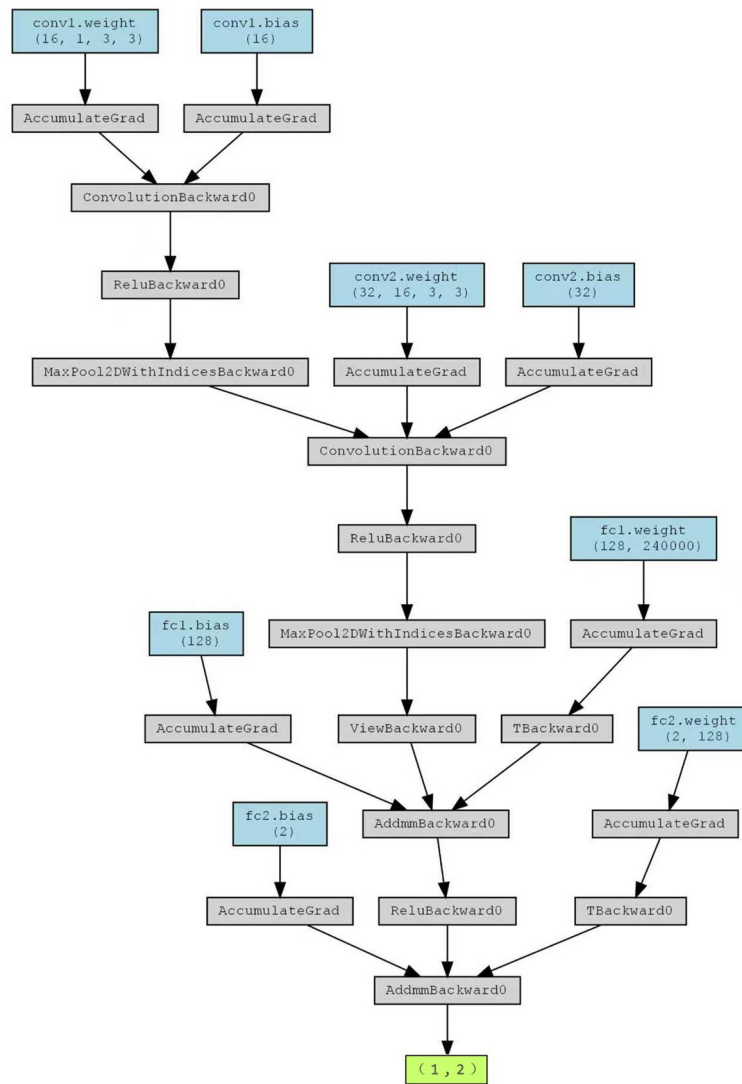


Fig. 1. Schematic architecture of the convolutional neural network (CNN) model used for fingerprint-based sex classification.

Higher AUC values indicate superior classification performance.

Ethical approval and consent to participate

All procedures involving human participants were conducted in accordance with relevant guidelines and regulations. The study protocol was reviewed and approved by the Institutional Review Board of Zhengzhou Police University. Written informed consent was obtained from all volunteers prior to data collection. The participants were fully informed about the study objectives and data handling protocols. No personal identifiers were recorded. Owing to limitations in consent coverage, the fingerprint dataset has not been made publicly available.

Experimental results

Fingerprint acquisition

A total of 1,100 high-resolution fingerprint samples were obtained via a ZKTECO entropy-based fingerprint acquisition device operating under a standardised protocol with a resolution of 500 DPI. The complete dataset comprised 1,000 images from 200 volunteers (100 males and 100 females) for model development, plus an additional 100 images from 20 volunteers (10 males and 10 females) reserved as an independent test set. All volunteers were aged between 18 and 22 years, and all images were standardised to a resolution of 300 × 400 pixels.

The captured fingerprints clearly demonstrated essential ridge features and sweat pore distributions. These high-quality samples provided comprehensive morphological fingerprint details, forming a robust and reliable



Fig. 2. | Representative raw fingerprint images (a–d) showing ridge features and sweat pore distributions, captured at 500 DPI and standardised to 300×400 pixels.

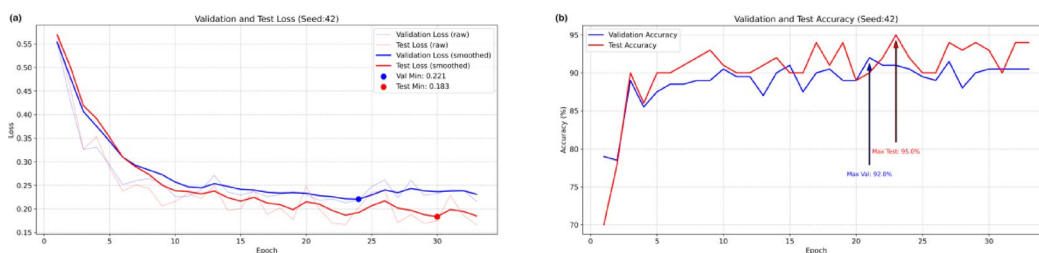


Fig. 3. Model training curves: (a) Loss curves for the training and validation datasets; (b) accuracy curves for the training, validation, and test sets.

foundation for model construction and subsequent biometric analyses. Representative raw fingerprint images are presented in Fig. 2.

Model training outcomes

The original dataset was partitioned into training and validation sets at an 8:2 ratio. The training set was used to train and learn the model, whilst the validation set was employed to assess the model's training performance and detect potential overfitting. Additionally, the test set, which was collected separately, was used to evaluate the model's generalisation capability. The full training workflow is illustrated in Fig. 3. An early stopping mechanism was applied, which terminated training if the validation loss did not decrease over 10 consecutive epochs. Accordingly, the model ceased training at epoch 33, effectively preventing overfitting.

As shown in the loss curve (Fig. 3a), the loss values decreased rapidly during the initial epochs, indicating efficient feature capture. With continued training, the loss gradually stabilized, ultimately converging to a low and stable value. The synchronous downwards trend of validation and test losses, along with a consistent gap between the two curves, suggests balanced model performance with no evident overfitting.

The accuracy curves (Fig. 3b) demonstrated that the model achieved a validation accuracy of 91.00% and a test accuracy of 95.00%, indicating high predictive reliability. Both the validation accuracy and the test accuracy increased steadily across epochs, with rapid initial improvements followed by convergence. The consistency between the two curves confirms the stable training behavior and effective utilization of the fingerprint data features.

Collectively, these findings demonstrate the reliability of the CNN-based sex inference method using fingerprint images. The training results validate the robustness of the dataset, the appropriateness of the network architecture, and the effectiveness of the training strategy, confirming the feasibility of applying this model to fingerprint-based classification tasks.

Model performance evaluation

A confusion matrix was constructed to visualise the model's classification performance. Each row corresponds to the actual class, whilst each column represents the predicted class. As shown in Fig. 4a, the validation confusion matrix indicates that the recognition rate for female fingerprints (92.00%) is slightly higher than for male fingerprints (90.00%). Similarly, in Fig. 4b, the test confusion matrix shows that the recognition rate for female fingerprints (96.00%) is also slightly higher than for male fingerprints (94.00%). This finding is consistent with previous studies, suggesting that female fingerprints may exhibit more distinguishable features. Furthermore, the differences between the validation and test confusion matrices are minimal, ruling out the risks of overfitting or data distribution inconsistencies^{6,7}.

To evaluate the binary classification performance at different thresholds, Receiver Operating Characteristic (ROC) curves were plotted, with the x-axis representing the false positive rate (FPR) and the y-axis representing the true positive rate (TPR). As shown in Figs. 4c,d, the model achieved an AUC value of 0.974 (95% CI: 0.955–0.988) on the validation dataset and 0.983 (95% CI: 0.961–0.998) on the test dataset, indicating strong discriminative power and the model's ability to effectively extract sex classification features from fingerprint data. The narrow confidence intervals and high AUC values demonstrate robust performance with minimal uncertainty. The small difference between validation and test performance further supports the absence of overfitting or data distribution inconsistencies.

To assess the reliability of the model's predicted probabilities and support its potential deployment readiness, calibration curves were generated alongside Brier scores (Figs. 4e,f). A perfectly calibrated model would align with the diagonal reference line, where predicted probabilities match the actual fraction of positive outcomes. The validation set achieved a Brier score of 0.0708, whilst the test set yielded a Brier score of 0.0485, both indicating excellent calibration. These low Brier scores suggest that the model's predicted probabilities are well-calibrated and reliable, with the test set demonstrating particularly strong alignment between predicted and observed probabilities. The calibration curves reveal that the model's predictions are neither systematically overconfident nor underconfident across different probability ranges, further supporting its suitability for practical forensic applications where reliable probability estimates are essential for decision-making. Detailed model evaluation metrics are provided in Table 2.

Supplementary cross-validation analysis

To confirm that the hold-out validation results were not dependent on a particular data partition, fivefold cross-validation was performed on the entire development cohort (200 volunteers, 1000 images) following the protocol described in Materials and Methods. Table 3 presents the detailed performance metrics for each fold.

The cross-validation results demonstrated high internal consistency, with mean accuracy of 90.60% (SD: 2.04%) across the five folds. The low standard deviation (SD < 2.1% for accuracy, < 1.0% for AUC) confirms stable performance regardless of data partition. Notably, the cross-validation mean (90.60%) closely aligns with the hold-out validation performance (91.00%), providing convergent evidence that the primary validation results were representative and not due to a fortuitous split. The mean AUC of 0.9687 (SD: 0.0100) and mean Brier score of 0.0706 (SD: 0.0094) further corroborate the robustness of model performance across different training-validation configurations. The minimal performance variation across folds (coefficient of variation < 2.3%) demonstrates that the model has learned stable, generalisable sex-discriminative features rather than partition-specific patterns.

These supplementary findings, combined with the independent test set results (95% accuracy, 0.983 AUC; Table 2), provide evidence of model reliability encompassing both internal consistency and external validity. This dual validation approach demonstrates that the lightweight CNN architecture offers a practical alternative for fingerprint-based sex inference, with performance characteristics suitable for consideration in forensic applications alongside existing methodologies.

Heatmap analysis

The heatmap visualization (Fig. 5) revealed that the CNN model exhibited distinct spatial attention patterns during fingerprint processing. The model focused primarily on the central region of the fingerprint, which accounted for approximately 60% of the attention weights. This region typically exhibits structural stability and rich discriminative features.

In addition, the delta region attracted approximately 30% of the model's attention. This region's characteristic ridge bifurcations contributed auxiliary information to the model's understanding of the overall fingerprint structure. Together, these two regions form the core of model attention, providing visual insight into the areas contributing most significantly to classification decisions. This offers interpretability for model inference and guidance for optimizing future fingerprint feature extraction methods.

Discussion

Fingerprints, regarded as the "gold standard" of biometric evidence, are indispensable in criminal investigations¹⁷. The application of Artificial Intelligence (AI), particularly Convolutional Neural Networks (CNNs), has expanded their forensic utility. CNN models enable automatic feature extraction from fingerprint images, achieving high-accuracy sex predictions even from partial or low-quality prints, which are common at crime

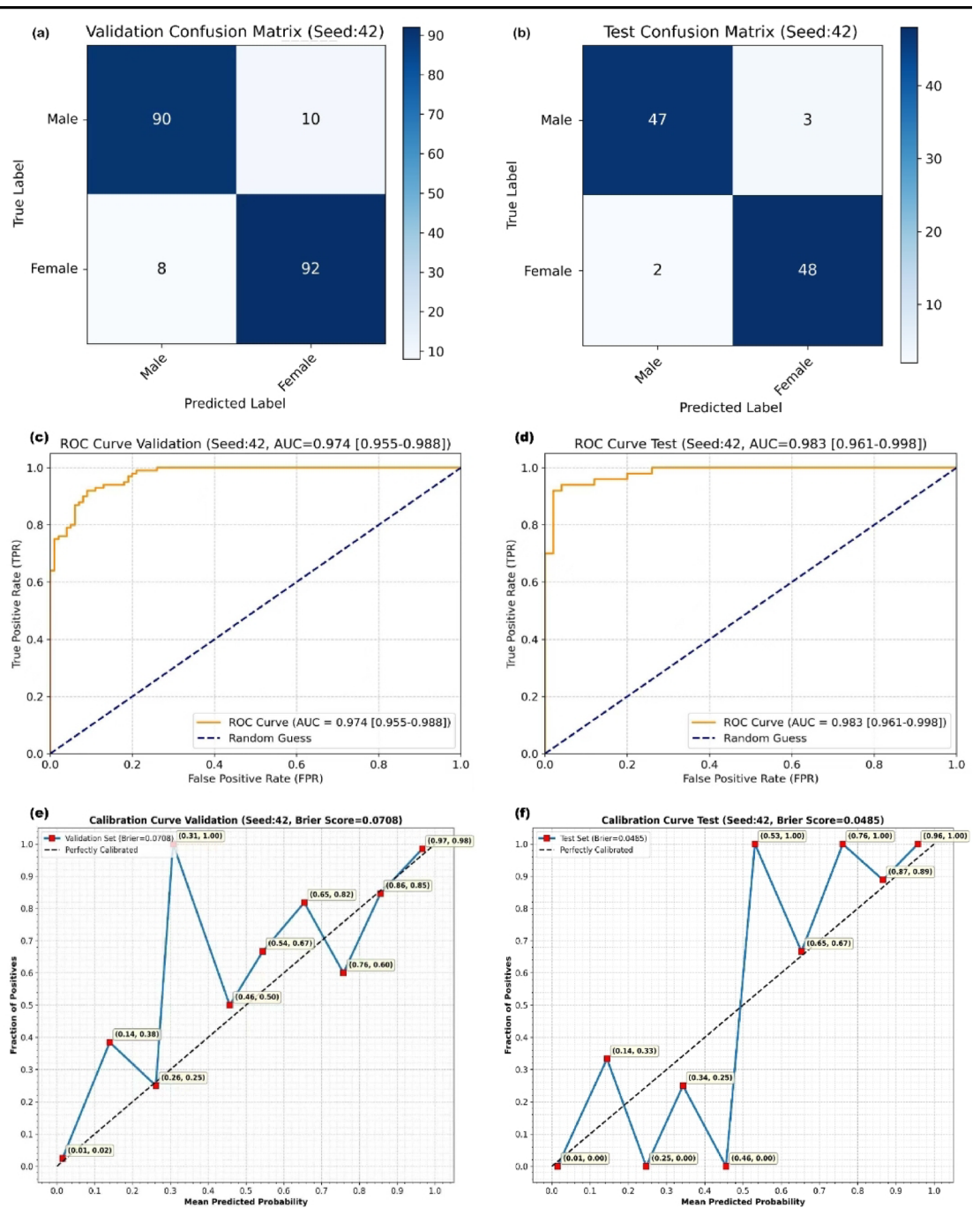


Fig. 4. Comprehensive model performance evaluation. (a) Validation confusion matrix showing classification results with 92% accuracy for female and 90% for male fingerprints. (b) Test confusion matrix demonstrating improved performance with 96% accuracy for female and 94% for male fingerprints. (c) ROC curve for the validation dataset with AUC=0.974 (95% CI 0.955–0.988), indicating strong discriminative ability. (d) ROC curve for the test dataset with AUC=0.983 (95% CI 0.961–0.998), confirming robust generalisation. (e) Calibration curve for the validation dataset with Brier score=0.0708, demonstrating good alignment between predicted probabilities and actual outcomes. (f) Calibration curve for the test dataset with Brier score=0.0485, showing excellent calibration and reliability of predicted probabilities. The dashed diagonal line in calibration plots represents perfect calibration.

scenes^{25–27}. This automation transforms subjective, expert-driven analysis into objective, quantifiable processes, enhancing classification accuracy and efficiency^{28–31}.

This capability further demonstrates fingerprints’ potential as a forensic tool: extending beyond identity verification to serve as a source of investigative information, complementing existing approaches to biometric profiling. When fingerprint matches are absent in databases, AI models can provide biometric characteristics

Evaluation metric	Validation set	Test set
Accuracy	0.91	0.95
Loss	0.22	0.17
Precision	0.91	0.95
Recall	0.91	0.95
F1-Score	0.91	0.95
AUC	0.974	0.983
AUC [95% CI]	[0.955–0.988]	[0.961–0.998]
Brier Score	0.0708	0.0485

Table 2. Model evaluation metrics on validation set and test set.

Fold	Accuracy	Loss	Precision	Recall	F1-score	AUC [95% CI]	Brier score
1	0.920	0.1885	0.921	0.920	0.920	0.9815[0.967–0.993]	0.0574
2	0.915	0.2319	0.915	0.915	0.915	0.9605[0.930–0.985]	0.0712
3	0.910	0.2403	0.910	0.910	0.910	0.9698[0.946–0.988]	0.0690
4	0.915	0.2501	0.916	0.915	0.915	0.9745[0.956–0.989]	0.0684
5	0.870	0.3168	0.870	0.870	0.870	0.9572[0.930–0.977]	0.0868
Mean	0.9060	0.2455	0.9064	0.9060	0.9060	0.9687	0.0706
SD	0.0204	0.0414	0.0207	0.0204	0.0204	0.0100	0.0094

Table 3. Model Performance Metrics from 5-Fold Cross-Validation.

like sex and handedness, offering key leads that may refine suspect lists and direct investigative resources³². Additionally, AI can support tasks such as preliminary classification, potentially improving the objectivity and efficiency of forensic workflows. This expanded role suggests that fingerprints could serve as useful tools for constructing biometric profiles and advancing investigations.

This study developed a convolutional neural network (CNN)-based model for fingerprint-based sex inference, demonstrating its capacity to extract sex-related features from fingerprint images. Throughout the training phase, the model exhibited stable convergence characteristics, with the validation loss decreasing from an initial value of 0.55 to 0.21 and the test loss decreasing from 0.56 to 0.19. The consistent trends observed in both loss curves indicate the model's ability to capture discriminative features associated with sex in fingerprint patterns. Based on a dataset comprising 1,000 fingerprint samples, the model achieved a validation accuracy of 91.00% and a test accuracy of 95.00%. Comprehensive performance evaluation revealed strong discriminative power, with AUC values of 0.974 (95% CI: 0.955–0.988) on the validation set and 0.983 (95% CI: 0.961–0.998) on the test set. The narrow confidence intervals demonstrate consistent performance with minimal uncertainty. Furthermore, precision, recall, and F1-scores all exceeded 0.91, indicating balanced classification performance across both sexes. Calibration analysis yielded Brier scores of 0.0708 for the validation set and 0.0485 for the test set, both considerably below 0.10, suggesting that the model's predicted probabilities are well-calibrated and reliable. These results collectively suggest robust generalisation capabilities, effective avoidance of overfitting, and suitability for forensic decision-making contexts where reliable probability estimates are essential. The model represents a lightweight, computationally efficient alternative to more complex deep learning architectures, maintaining competitive performance whilst offering practical advantages for resource-constrained forensic applications.

To address potential concerns regarding validation robustness with limited sample sizes, this study employed a dual validation strategy combining hold-out validation (8:2 split) and supplementary fivefold cross-validation. The hold-out validation set (40 volunteers, 200 images) served as the primary assessment during model training, achieving 91.00% accuracy with an AUC of 0.974. To confirm that these results were not dependent on a fortuitous data partition, supplementary fivefold cross-validation was conducted on the entire development cohort (200 volunteers, 1000 images), yielding a mean accuracy of 90.60% (SD: 2.04%) and mean AUC of 0.9687 (SD: 0.0100) across folds. The low standard deviations (SD < 2.1% for accuracy, < 1.0% for AUC) and close alignment between hold-out validation (91.00%) and cross-validation mean (90.60%) provide convergent evidence of internal consistency, confirming that model performance is stable across different data partitions.

Importantly, whilst both hold-out validation and cross-validation assess internal consistency within the development cohort, the independent test set—comprising 100 images from 20 additional volunteers recruited separately—provides the critical assessment of external validity. The test set achieved 95% accuracy and 0.983 AUC, demonstrating performance that compares favourably with internal validation results (90.60–91.00%). This consistency between internal and external validation suggests that the model has learned robust sex-discriminative features, offering a computationally efficient approach that may serve as a practical complement to existing forensic methodologies. Such external validation is recognised as the gold standard for demonstrating deployment readiness, where operational performance must be estimated on entirely new cases not present during model development^{33,34}.

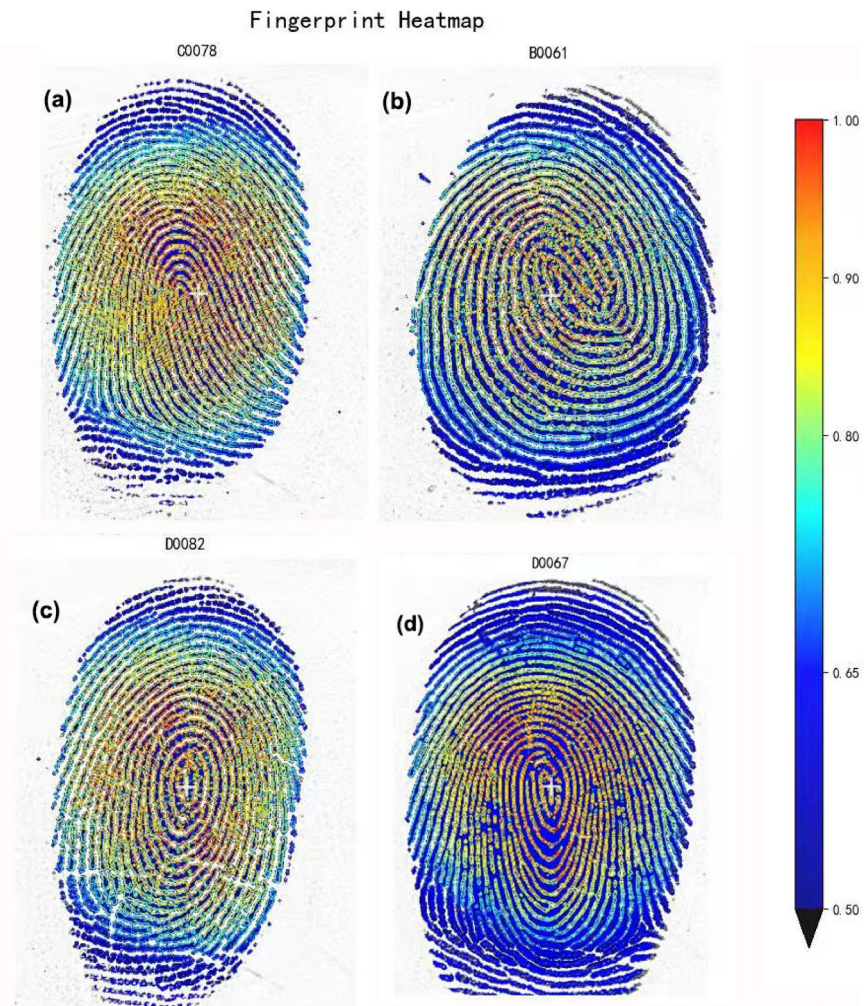


Fig. 5. Class activation mapping (CAM) visualisations (a–d) showing spatial attention patterns, with primary focus on central whorl regions (~60%) and delta regions (~30%). Warmer colours indicate higher importance.

This comprehensive validation framework—combining internal consistency assessment (hold-out + cross-validation) with external generalisability evaluation (independent test)—aligns with best practices in machine learning model validation³⁵. The consistent internal performance and favourable external validation collectively provide evidence supporting the model's potential reliability as a lightweight alternative for fingerprint-based sex inference in forensic contexts, where both stability across data samples and generalisation to new individuals are valued characteristics.

Table 4 summarises the performance of the proposed model alongside previous studies. Gnanasivam and Muttan³⁶ combined a six-level discrete wavelet transform (DWT) and singular value decomposition (SVD) with a K-nearest neighbour (KNN) classifier, achieving an accuracy of 91.67% for males, 84.89% for females, and an overall classification rate of 88.28%. Abdullah et al.³⁷ achieved a success rate of approximately 74.5% using a ridge density-based method. Beanbonyka et al.³⁸ directly employed various advanced deep learning architectures (including VGG-19, ResNet-50, and EfficientNet-B3) for end-to-end learning on raw fingerprint images, with their best model achieving a classification accuracy of 63.05% on the test set. Furthermore, multi-finger fusion deep learning methods²³ achieved an overall classification accuracy of 91.3%.

The model demonstrates performance that compares favourably with recent studies whilst maintaining both precision and recall consistently above 0.91, which may help address the commonly encountered issue of sex identification bias. Importantly, beyond standard classification metrics, the model's well-calibrated probability estimates (Brier scores < 0.071) ensure that predicted probabilities appropriately reflect classification confidence—a characteristic valued in forensic applications where probabilistic evidence must be weighed alongside other investigative information. When compared to recent studies utilising alternative biological traits—such as orbital measurements and mandibular CBCT morphology—for sex inference, the model shows comparable or favourable performance. For example, studies employing random forest models for sex prediction reported precision values of 0.65, recall values of 0.70, and an F1-score of 0.675³⁹. Whilst another study achieved an accuracy of 97.95% for specific dental categories, recall rates fluctuated significantly between classes (ranging from 0.33 to 1.0), indicating potential imbalance⁴⁰. A study by Baban et al.⁴¹ based on mandibular CBCT

Author(s)	Dataset used	Model used	Accuracy	AUC	Precision	Recall	F1-score
Present Study (Validation Set)	Fingerprint	Customized CNN	0.91	0.974	0.91	0.91	0.91
Present Study (Test Set)	Fingerprint	Customized CNN	0.95	0.983	0.95	0.95	0.95
Iloanusi and Ejiogu ²³	Multi-fingerprint Fusion	Customized CNN	0.913	–	–	–	–
Rim et al. ³⁸	Fingerprint	CNN (EfficientNet-B3)	0.6986 (Val) / 0.6305 (Test)	–	–	–	–
Gnanasivam and Muttan ³⁶	Fingerprint (DWT + SVD)	K-Nearest Neighbor	0.8828	–	–	–	–
Abdullah et al. ³⁷	Fingerprint (Ridge Density)	Custom Algorithm	0.745	–	–	–	–
Triantafyllou et al. ³⁹	Orbital Measurements	Random Forest	0.68	0.72	0.65	0.70	0.675
Natarajan et al. ⁴⁰	Dental Morphology	Random Forest	0.88	–	0.82 (F) 0.94 (M)	0.91 (F) 0.87 (M)	–
Baban et al. ⁴¹	Mandibular CBCT	Gaussian Naive Bayes	0.90	–	0.86 (F) 0.95 (M)	0.95 (F) 0.86 (M)	0.90

Table 4. Comparison of the performance of the proposed model with recent sex inference models reported in the literature. F, Female; M, Male; Val, Validation Set; Test, Test Set.

morphology showed that the best-performing Gaussian Naive Bayes model achieved an overall test accuracy of 0.90, with a precision of 0.86 and recall of 0.95 for the female category, and a precision of 0.95 and recall of 0.86 for the male category, achieving a macro-averaged F1-score of 0.90. However, these studies did not report calibration metrics, making it difficult to assess whether their probability estimates would be reliable for real-world forensic decision-making. In comparison, our model maintains consistently high and balanced values across classification metrics whilst demonstrating strong calibration, suggesting robust practical applicability for forensic contexts where both accurate classifications and reliable probability estimates are essential.

It is important to emphasise that the model's performance reflects both its architectural design and the selection of appropriate biological features. The information inherent in fingerprint traits related to sexual dimorphism, combined with the feature extraction capabilities of CNNs, collectively contribute to the performance demonstrated in this study. These findings offer insights for future research on fingerprint-based sex inference and highlight the potential role that fingerprint features can play in forensic biometrics.

Rather than pursuing maximal architectural complexity, the proposed model employs a dual-layer CNN architecture (3×3 convolution kernels with ReLU activation functions), enabling automatic extraction of both global ridge structures and local minutia features⁴². This design choice prioritises computational efficiency and practical deployability, offering a more accessible alternative to resource-intensive deep networks. Compared to studies that rely on deep pre-trained models (such as VGG16) or complex multi-task architectures, the lightweight dual-layer structure employed here represents a more targeted approach. The 3×3 convolution kernels are designed to match the micro-scale ridge patterns of fingerprints, whilst the ReLU activation function helps filter noise from low-quality samples. This approach aims to avoid parameter redundancy and overfitting whilst enabling efficient extraction of sex-discriminative features, such as sweat pore distribution and local ridge density.

Furthermore, by incorporating two levels of 2×2 max pooling layers, the model performs dimensionality reduction and key information aggregation. Experimental results suggest that this hierarchical pooling design enhances feature representation capability⁴³. Compared to the study by Khazaei et al.⁴⁴, which applied deep architectures such as DenseNet121 for sex classification, our approach focuses on a more lightweight, dedicated network rather than relying on complex models dependent on large-scale pre-training. This design choice improves computational efficiency whilst maintaining feature extraction capability. Notably, despite the use of partial fingerprint samples from a single hand, the CNN achieves high test accuracy, confirming its potential suitability for real-world forensic scenarios where crime scene fingerprints are often incomplete or blurred^{25–27,45}. The model achieved a per-sample inference time of only 15 ms whilst maintaining a 95.00% prediction rate—an improvement over traditional manual analysis workflows.

In addition to the modelling framework, extensive data augmentation techniques, inspired by successful U-Net image enhancement strategies^{46,47}, have been adopted to address the common limitation of small sample sizes in forensic research. This enhancement methodology has contributed to generalisation performance on a limited yet demographically balanced dataset ($n=200$ individuals). The volunteers' ages were controlled between 18 and 22 years, a period during which fingerprint features remain stable and well-defined, reducing variations in fingerprint patterns that may arise from factors other than sex-related variables. Furthermore, we employed a ZKTECO entropy-based fingerprint acquisition device adhering to a 500 DPI high-resolution standard for standardised data collection. In contrast to studies relying on pre-existing image libraries or datasets potentially suffering from quality loss⁴³, our dataset provides clarity, consistency, and completeness of information, offering a foundation for feature extraction and classification. Additionally, during data collection, all samples were uniformly named and securely encrypted, enhancing the standardisation and security of data management. When compared to the similar CNN-based approach reported by Hsiao et al.¹⁷, our model showed an 11.4% improvement in AUC, which may be attributed to the combination of standardised high-resolution image acquisition (500 DPI), systematic data augmentation strategies, and controlled demographic sampling. This comparison suggests that lightweight CNN architectures, when paired with high-quality data collection and appropriate preprocessing pipelines, can provide a viable and efficient alternative to more complex deep learning frameworks for forensic sex classification.

Importantly, the forensic applicability of machine learning models depends not only on performance metrics but also on interpretability^{31,48–51}. To this end, we implemented class activation mapping (CAM) to generate

high-resolution heatmaps, which visually localise regions contributing most to the model's decisions^{52,53}. These attention maps revealed a predominant focus on central whorl regions and triradial areas—structural zones previously reported as sex-discriminative in classical studies (e.g., differential ridge density in specific regions, as confirmed by Sharma et al.⁸). This alignment with prior empirical knowledge provides supporting evidence for the model's biological plausibility, whilst acknowledging that such interpretability techniques offer approximate rather than definitive explanations of model behaviour. The correspondence between data-driven focus regions and prior empirical knowledge lends biological plausibility to the model and may enhance expert interpretability. These maps could potentially guide forensic analysts in prioritising key areas during manual evaluations, supporting scientific rigour and evidentiary credibility.

Despite these promising results, several limitations remain. The current dataset is limited to young adults aged 18–22 years, predominantly students, with all fingerprint samples obtained from the right hand. This narrow demographic focus restricts the model's applicability in real-world forensic contexts, where individuals from a wider range of age groups, body statures, and occupational backgrounds must be considered. As such, the model may not generalise effectively to more diverse and complex populations, which poses limitations on its practical utility. Additionally, during the model training process, finger identity (e.g., individual fingers) was not treated as a parameter. Consequently, the model was trained on fingerprints from all five fingers without distinguishing between them. This lack of differentiation may limit the model's ability to focus specifically on sex-related features, potentially affecting both its performance and its ability to extract relevant discriminatory features. Whilst 1,000 fingerprint images were collected using image augmentation techniques to ensure uniformity and quality, the dataset remains restricted to right-hand fingerprints from just 200 volunteers. This limitation in sample diversity may further restrict the model's generalisability. Future work should therefore incorporate a broader demographic spectrum, including age, height, and occupation-related fingerprint variability.

Key limitations

The principal limitations of this study can be summarised as follows:

1. **Narrow demographic scope:** The dataset is restricted to young adults aged 18–22 years, predominantly university students, limiting generalisability to broader age groups, body statures, and occupational backgrounds encountered in real-world forensic contexts.
2. **Non-public dataset:** Due to ethical constraints and the absence of consent for public release, the raw fingerprint images cannot be shared publicly, which may limit independent validation efforts, although trained model weights are available upon request.
3. **Single-hand sampling without finger-specific parameterisation:** All fingerprint samples were obtained exclusively from the right hand, and finger identity was not treated as a distinct parameter during model training. This may limit the model's ability to focus on finger-specific sex-related features.
4. **Limited sample diversity:** Despite employing image augmentation techniques, the dataset comprises 220 volunteers in total (200 for development, 20 for testing), which may restrict the model's generalisability to more diverse and complex populations.

These constraints should be carefully considered when interpreting the study's findings and planning future validation studies.

To address these limitations, future research could focus on extending this framework by developing more advanced multidimensional fingerprint recognition models. These models could potentially be integrated into portable devices designed for use in crime scene investigations and forensic evidence analysis, thereby enhancing their real-world applicability. Additionally, exploring correlations between fingerprint features and individual traits such as age, stature, and occupation could enable more precise multidimensional profiling of suspects, thereby potentially narrowing investigative scopes and improving case resolution efficiency. The integration of dynamic thresholding mechanisms and the development of multi-feature intelligent systems capable of inferring such traits from fingerprints represent promising avenues for future work. Ultimately, these advancements aim to establish more efficient and user-friendly fingerprint recognition systems within forensic sciences, supporting technological innovation in crime prevention and the administration of justice.

Conclusion

This study presents a CNN-based approach for sex inference from fingerprints, contributing to biological profiling methodologies in forensic science. The convolutional neural network model developed in this work demonstrated robust performance, offering a lightweight and well-validated alternative within existing approaches. A standardised fingerprint dataset comprising 1,100 samples from 220 volunteers was established, with 1,000 images from 200 volunteers allocated for model development and 100 images from an additional 20 volunteers reserved for independent external validation. A dual-layer CNN architecture was designed to balance computational efficiency with predictive accuracy. The model achieved a test accuracy of 95.00% with AUC values of 0.974–0.983, whilst maintaining balanced performance across precision, recall, and F1-scores (all exceeding 0.91).

The incorporation of class activation mapping (CAM) enhanced the interpretability of classification outcomes by highlighting biologically relevant regions, such as central whorl and triradial areas previously identified as sex-discriminative features. This visualisation approach provides potential support for forensic analysts in manual evaluations, thereby strengthening the evidential foundation of the method. Compared to previous approaches, the lightweight architecture enables efficient feature extraction whilst avoiding parameter redundancy, making it potentially suitable for practical forensic applications where computational resources may be limited.

This work contributes to the ongoing development of biometric data analysis methodologies integrated with deep learning approaches to support forensic practice. The lightweight CNN framework presented here offers a well-validated alternative within existing approaches, demonstrating competitive performance whilst maintaining computational efficiency. The model may serve as one among several complementary tools for suspect profiling in cases where database matches are unavailable, potentially aiding investigative processes alongside established methodologies. Ultimately, these efforts aim to contribute to the development of more diverse and accessible biometric recognition options within forensic sciences.

Data availability

The datasets generated during the current study are not publicly available due to the lack of consent for public release from all participants, as the dataset contains identifiable biometric information that cannot be anonymised. However, to facilitate reproducibility and transparency, the trained model weights and inference scripts are available from the corresponding author upon reasonable request and with appropriate ethical oversight. Researchers interested in validating or extending this work may contact the corresponding author with a detailed description of their intended use. All experimental procedures, model architecture details, and hyperparameters are fully described in the Methods section to ensure reproducibility of the analytical framework.

Received: 14 August 2025; Accepted: 31 October 2025

Published online: 01 December 2025

References

1. Trokielewicz, M., Maciejewicz, P. & Czajka, A. Post-mortem iris biometrics - field, applications and methods. *Forensic Sci. Int.* **365**, 112293 (2024).
2. Tome, P., Vera-Rodriguez, R., Fierrez, J. & Ortega-Garcia, J. Facial soft biometric features for forensic face recognition. *Forensic Sci. Int.* **257**, 271–284 (2015).
3. Cao, K. & Jain, A. K. Automated latent fingerprint recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **41**(4), 788–800 (2019).
4. Yoon, S. & Jain, A. K. Longitudinal study of fingerprint recognition. *Proc. Natl. Acad. Sci. U. S. A.* **112**(28), 8555–8560 (2015).
5. Jalali, S., Boostani, R. & Mohammadi, M. Efficient fingerprint features for gender recognition. *Multidim. Syst. Sign. Process.* **33**, 81–97 (2022).
6. Das, D. et al. Sexual dimorphism and topological variability in fingerprint ridge density in a north-west Indian population. *Sci. Nat.* **111**, 23. <https://doi.org/10.1007/s00114-024-01911-x> (2024).
7. Thakar, M. K., Kaur, P. & Sharma, T. Validation studies on gender determination from fingerprints with special emphasis on ridge characteristics. *Egypt. J. Forensic Sci.* **8**, 20 (2018).
8. Sharma, S., Shrestha, R., Krishan, K. & Kanchan, T. Sex estimation from fingerprint ridge density. *Acta Biomed.* **92**(5), e2021366 (2021).
9. Huynh, C., Brunelle, E., Halámková, L., Agudelo, J. & Halámek, J. Forensic identification of gender from fingerprints. *Anal. Chem.* **87**(22), 11531–11536 (2015).
10. Levanon, L. & Tully, G. Fingerprint analysis and reporting in legal trials: A critical re-evaluation. *Crim. Law Forum* **36**, 1–32. <https://doi.org/10.1007/s10609-025-09497-3> (2025).
11. Martins, N., Silva, J. S. & Bernardino, A. fingerprint recognition in forensic scenarios. *Sensors* **24**(2), 664 (2024).
12. Abraham, J., Champod, C., Lennard, C. & Roux, C. Modern statistical models for forensic fingerprint examinations: A critical review. *Forensic Sci. Int.* **232**(1–3), 131–150 (2013).
13. Guo, G. et al. Unveiling intra-person fingerprint similarity via deep contrastive learning. *Sci. Adv.* **10**(2), eadi0329 (2024).
14. Zhang, Z., Liu, S. & Liu, M. A Multi-task fully deep convolutional neural network for contactless fingerprint minutiae extraction. *Pattern Recognit.* **120**, 108189 (2021).
15. Singh, R., Singh, R., Tripathi, R. K. & Agarwal, P. Fingerprint recognition using artificial neural networks. *Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci.* **95**, 127–135 (2025).
16. Kaplesh, P., Gupta, A., Bansal, D., Sofat, S. & Mittal, A. Vision transformer for contactless fingerprint classification. *Multimed. Tools Appl.* **84**, 31239–31259 (2025).
17. Hsiao, C. T., Lin, C. Y., Wang, P. S. & Wu, Y. T. Application of convolutional neural network for fingerprint-based prediction of gender, finger position, and height. *Entropy.* **24**(4), 475 (2022).
18. Turner, D. A., Pichtel, J., Rodenas, Y., McKillip, J. & Goodpaster, J. V. Microbial degradation of gasoline in soil: Effect of season of sampling. *Forensic Sci. Int.* **251**, 69–76 (2015).
19. Marietta, P. P. et al. Human handedness: a meta-analysis. *Psychol. Bull.* **146**(6), 481–524 (2020).
20. Kapoor, N., Badiye, A., & Mishra S. D. Fingerprint analysis for the determination of hand origin (right/left) using the axis slant in whorl patterns. *Forensic Sci. Res.* **7**(2), 285–289 (2020).
21. Stylianou, N. & Vlahavas, I. TransforMED: end-to-end transformers for evidence-based medicine and argument mining in medical literature. *J Biomed. Inform.* **117**, 103767 (2021).
22. Houy, N. & Le Grand, F. Personalized oncology with artificial intelligence: the case of temozolomide. *Artif. Intell. Med.* **99**, 101693 (2019).
23. Iloanusi, O. N. & Ejiogu, U. C. Gender classification from fused multi-fingerprint types. *Inf. Secur. J. Glob. Perspect.* **29**(5), 209–219. <https://doi.org/10.1080/19393555.2020.1741742> (2020).
24. Tarare, S., Anjekar, A. & Turkar, H. Fingerprint based gender classification using DWT transform. In: *Proceedings of the 1st International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, Pune, India, 689–693 (2015).
25. Muthusamy, D. & Muniyappan, S. Enhancement comparison of laplace kernelized piecewise regression-based progressive generative adversarial network for latent fingerprint. *Pattern Anal. Appl.* **28**, 118 (2025).
26. Chhabra, M., Ravulakollu, K. K., Kumar, M., Sharma, A. & Nayyar, A. Improving automated latent fingerprint detection and segmentation using deep convolutional neural network. *Neural Comput. Appl.* **35**, 6471–6497 <https://doi.org/10.1007/s00521-022-07894-y> (2023).
27. Zhu, Y., Yin, X. & Hu, J. Robust fingerprint matching based on convolutional neural networks. In *Mobile Networks and Management* (eds Hu, J., Khalil, I., Tari, Z. & Wen, S.). MONAMI 2017. Lecture Notes of the Inst. for Comput. Sci., Soc. Informatics Telecommun Eng. **235**, (Springer, 2018).
28. Yager, N. & Amin, A. Fingerprint classification: A review. *Pattern Anal. Appl.* **7**, 77–93 (2004).
29. Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Proc. Adv. Neural Inf. Process. Syst.* **2012**, 1097–1105 (2012).
30. TensorFlow. TensorFlow tutorials, convolutional neural network (CNN). <https://www.tensorflow.org/tutorials/images/cnn> (2024).

31. Finzel, B. Current methods in explainable artificial intelligence and future prospects for integrative physiology. *Pflugers Arch. Eur. J. Physiol.* **477**, 513–529. <https://doi.org/10.1007/s00424-025-03067-7> (2025).
32. Maltoni, D., Maio, D., Jain, A. K. & Feng, J. Fingerprint matching. In: *Handbook of Fingerprint Recognition* (Springer, 2022). https://doi.org/10.1007/978-3-030-83624-5_4.
33. Steyerberg, E. W. & Harrell, F. E. Prediction models need appropriate internal, internal–external, and external validation. *J. Clin. Epidemiol.* **69**, 245–247 (2016).
34. Vabalas, A., Gowen, E., Poliakoff, E. & Casson, A. J. Machine learning algorithm validation with a limited sample size. *PLoS One.* **14**(11), e0224365 (2019).
35. Riley R. D. et al. Calculating the sample size required for developing a clinical prediction model. *BMJ.* **368**, m441 (2020).
36. Gnanasivam, P. & Muttan, S. Fingerprint gender classification using wavelet transform and singular value decomposition <https://doi.org/10.48550/arXiv.1205.6745> (2012).
37. Abdullah, S. F., Rahman, A. F. N. A., Abas, Z. A. & Saad, W. H. M. Development of a fingerprint gender classification algorithm using fingerprint global features. *Int. J. Adv. Comput. Sci. Appl.* **7**(6), 275–279. [https://doi.org/10.1016/S1002-0721\(14\)60405-1](https://doi.org/10.1016/S1002-0721(14)60405-1) (2016).
38. Rim, B., Kim, J. & Hong, M. Gender classification from fingerprint-images using deep learning approach. In *Proceedings of the International Conference on Research in Adaptive and Convergent Systems* (RACS '20). Association for Computing Machinery, New York, NY, USA, 7–12. <https://doi.org/10.1145/3400286.3418237> (2020).
39. Triantafyllou, G. et al. Sex estimation through orbital measurements: A machine learning approach for forensic science. *Diagnostics (Basel)*. Dec 10, **14**(24), 2773 (2024).
40. Natarajan, S. et al. Tooth shape and sex estimation: a 3D geometric morphometric landmark-based comparative analysis of artificial neural networks, support vector machines, and Random Forest models. *3 Biotech.* **15**(8), 273 (2025).
41. Baban, M. T. A. & Mohammad, D. N. The accuracy of sex identification using CBCT morphometric, easurements of the mandible, with different machine-learning algorithms—A retrospective study. *Diagnostics* **13**(14), 2342. <https://doi.org/10.3390/diagnostics13142342> (2023).
42. Ningthoujam, C., Singh, T. C., Brahma, B. & Bhoi, A. K. Hybrid cnn-knn model for image annotation: Combining deep learning and instance-based learning. *J. Inst. Eng. India Ser. B* <https://doi.org/10.1007/s40031-025-01239-8> (2025).
43. Mahajan, A. et al. Deep learning-based CNN model for multiclass classification of fingerprint patterns. *Med. Sci. Law.* (2025).
44. Khazaei, M., Mollabashi, V., Khotanlou, H. & Farhadian, M. Sex determination from lateral cephalometric radiographs using an automated deep learning convolutional neural network. *Imag. Sci. Dent.* **52**(3), 239–244 (2022).
45. Parvathy, J. & Patil, P. G. Fingerprint recognition model using improved firebug swarm optimization and tanh-based fuzzy activated neural network. *SN Comput. Sci.* **5**, 575. <https://doi.org/10.1007/s42979-024-02885-3> (2024).
46. Askarin, M. M., Wang, M., Yin, X., Jia, X. & Hu, J. U-net-based fingerprint enhancement for 3D fingerprint recognition. *Sensors (Basel)*. **25**(5), 1384 (2025).
47. Cheng, Y. H., Su, S. G., Lin, Y. L. & Hsu, H. C. Fingerprint image enhancement method based on u-net model. In: Zhao, F. & Miao, D. (eds) *AI-generated Content. AIGC 2023. Commun. in Comput. Inf. Sci.* **1946**, (Springer, Singapore, 2024).
48. Şahin, E., Arslan, N. N. & Özdemir, D. Unlocking the black box: an in-depth review on interpretability, explainability, and reliability in deep learning. *Neural Comput. Appl.* **37**, 859–965 (2025).
49. Hassija, et al. Interpreting black-box models: A review on explainable artificial intelligence. *Cogn. Comput.* **16**, 45–74 <https://doi.org/10.1007/s12559-023-10179-8> (2024).
50. Elbeialy, H. et al. Visual steering for deep neural networks using explainable artificial intelligence. In: Hassanien A, Rizk R.Y., Pamucar D, Darwish A. & Chang, K.C. (eds) *Proceedings of the 9th International Conference on Advanced Intelligent Systems and Informatics 2023 (AISI 2023)*. Lecture Notes on Data Eng. Commun. Technol., **184**, (Springer, 2023).
51. Garrett, B. L. & Rudin, C. Interpretable algorithmic forensics. *Proc. Natl. Acad. Sci. U. S. A.* **120**(41), e2301842120 (2023).
52. Truong, N., Pesenti, D. & Hasson, U. Explaining human comparisons using alignment-importance heatmaps. *Comput. Brain Behav* <https://doi.org/10.1007/s42113-025-00235-x> (2025).
53. Du, S. & Ikenaga, T. Bidirectionally learning heatmaps for 2D human pose estimation. In: *Human Pose Analysis* (Springer, 2025). https://doi.org/10.1007/978-981-97-9334-1_2.

Author contributions

Y.Z. conceived the study, proposed the core idea, designed the experimental framework, and determined the research direction. Y.Z. also drafted the initial manuscript and wrote the main technical sections. Y.J. and W.L. conducted data collection, including subject recruitment, fingerprint acquisition, and sample preparation under standardized protocols. J.L. and X.L. performed data preprocessing, statistical analysis, and model performance evaluation. F.W. developed the core methodology, built the convolutional neural network model, and designed the overall technical pipeline. S.C. and S.L. provided essential resources, including funding, equipment, materials, and infrastructure support for fingerprint acquisition and computational analysis. All authors reviewed and approved the final manuscript.

Funding

This research was partially supported by the National Natural Science Foundation of China (No. 21805208), the Applied Innovation Program of the Ministry of Public Security (2024YY50), the Henan Provincial Science and Technology Research Project (212102310487 and 252102310375), and the Central University Basic Research Fund of China (2024TJJBKY041).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025