Fingerprint-Based Gender Classification Using Convolutional Neural Networks (CNNs)

Maria Diaz Alba
Applied Artificial Intelligence
Florida International University, Miami, Florida, US
mdiaz683@fiu.edu

***Abstract*—In forensic sciences, the analysis of physical evidence plays a crucial role in identifying individuals or understanding the context of certain events. Fingerprint analysis is traditionally employed for person identification, but this study aims to address a different challenge: determining the gender of an individual based on fingerprint images using Artificial Intelligence (AI).**

This work explores whether gender classification, a task typically requiring significant expertise and domain knowledge, can be performed effectively through Convolutional Neural Networks (CNNs). The central question is whether AI models can uncover subtle patterns in fingerprint images that distinguish male and female characteristics, even when such patterns are not easily discernible by humans.

***Index Terms*—Artificial Intelligence, Biometrics, Convolutional Neural Networks (CNNs), Deep Learning, Feature Extraction, Fingerprint Analysis, Gender Classification.**

I. Motivation

In forensics, human experts analyze evidence such as fingerprints, bones, and wounds to draw conclusions about identity, age, gender, and cause of death. However, this process is subjective, time-consuming, and prone to error. AI has the potential to revolutionize this field by providing objective, efficient, and scalable solutions.

A recent study by *G. Guo, et al.* [1] from Columbia University has demonstrated that AI can uncover previously undetectable patterns in fingerprints, such as distinguishing fingerprints from different fingers of the same person. These advancements suggest that AI could also classify fingerprints by race, gender, or other demographic features. Although such classifications are less individual-specific, they may provide valuable insights for forensic and security applications.

Fingerprint-based gender classification could have applications in crime scene investigations, demographic studies, and personalized biometric systems. This research contributes to these fields by evaluating the performance of CNNs on a dataset of fingerprint images to classify gender.

II. Literature Review

Fingerprint analysis has traditionally relied heavily on human expertise and predefined rules, a method that, while effective, is subject to human error and lacks scalability. The National Institute of Standards and Technology (NIST) highlights these challenges in their study [2], emphasizing the critical role of human factors in latent print analysis. However, advancements in machine learning (ML) and deep learning (DL) have paved the way for more objective and automated approaches to fingerprint analysis.

Recent studies have explored the potential of ML and DL for various biometric tasks, including age and gender classification. The work by *M. Patel and U. Singh* [3] provides insights into how deep learning can enhance the accuracy of age and gender recognition, demonstrating its effectiveness in biometric applications. Similarly, *S. Hamdi and A. Moussaoui* [4] conducted a comparative study, showing the superior performance of DL over traditional machine learning methods in these tasks.

Additionally, the study by *G. Guo, et al.* [1] from Columbia University introduces an innovative approach to fingerprint analysis. Using deep contrastive learning, their work highlights how AI can uncover intra-personal fingerprint similarities, laying the groundwork for demographic classification tasks such as gender prediction.

Despite these advancements, prior research in gender classification using fingerprints faces significant challenges. For example, *S. Patil* [5] identified issues related to dataset robustness and feature extraction. Their integrated analysis of multimodal biometric traits underscores the need for high-quality datasets and advanced preprocessing techniques to enhance classification performance.

The present study builds upon these findings, leveraging deep learning techniques to tackle the challenge of gender classification in fingerprint analysis.

III. Dataset Information

The dataset used in this study is a biometric fingerprint database designed for academic research purposes. It is made up of 6,000 fingerprint images from 600 African subjects and contains unique attributes such as labels for gender and hand and finger name. Below, detailed information about the dataset and its preprocessing steps is provided:

A. Source

The dataset was obtained from Kaggle.com [6] and is titled "Fingerprint Dataset".

B. Class Distribution

This dataset comprises fingerprint images from both males and females

- **Male:** 4770 images.
- **Female:** 1230 images.

IV. Data Preprocessing

A. Ensuring Uniformity and Size Consistency

All images in the dataset were resized to 128x128 pixels to ensure uniformity across the dataset. This step was taken to standardize the input dimensions, which is essential for consistent processing in the CNN model. The resizing introduced minimal distortion that does not compromise the integrity of the fingerprint patterns, making the data suitable for analysis. Additionally, the size distribution of the images was analyzed to confirm that all images now have the same dimensions, ensuring a consistent and balanced dataset for training and evaluation.

B. Balancing the Dataset

The dataset originally consisted of $1,230$ images labeled as *female* and $4,770$ images labeled as *male*, introducing a significant class imbalance. Two strategies were evaluated to balance the dataset:

- **Undersampling:** Reducing the number of *male* samples to match the *female* samples.
- **Oversampling:** Duplicating and augmenting *female* samples to match the number of *male* samples.

The latter approach was adopted to maximize the dataset size. Each *female* sample was duplicated randomly until the total count reached $3,000$ images per class, ensuring an equal representation of both categories in the dataset.

C. Final Dataset Distribution

After preprocessing and balancing, the dataset consisted of $6,000$ images, equally distributed between the two classes (*female* and *male*). This final dataset was split into training, validation, and testing sets to facilitate model training and evaluation.

- **Training set:** 72% of the images (4320 images).
- **Validation set:** 18% of the images (1080 images).
- **Testing set:** 10% of the images (600 images).

D. Data Augmentation

To address potential overfitting and enhance the model's generalization ability, data augmentation was applied to the training dataset. This approach artificially expanded the dataset by introducing variations, making the model more robust to unseen data. The following augmentation techniques were implemented:

- **Random Rotations:** Images were rotated randomly within a specified range to simulate various fingerprint orientations.
- **Zoom Adjustments:** Minor zooms were applied to simulate different scales of fingerprint images.
- **Width and Height Shifts:** Horizontal and vertical translations were performed to simulate small shifts in fingerprint placements.

These augmentations were only applied to the training dataset to avoid artificially altering the validation and test datasets, which are intended to measure the model's performance on untouched data. In figure 1 we can see a sample of the appearance of the images that will be used for training after applying data augmentation.

Additionally, pixel values for all images were normalized to the range $[0, 1]$. This scaling process was essential for improving model training stability by ensuring numerical consistency across the dataset.

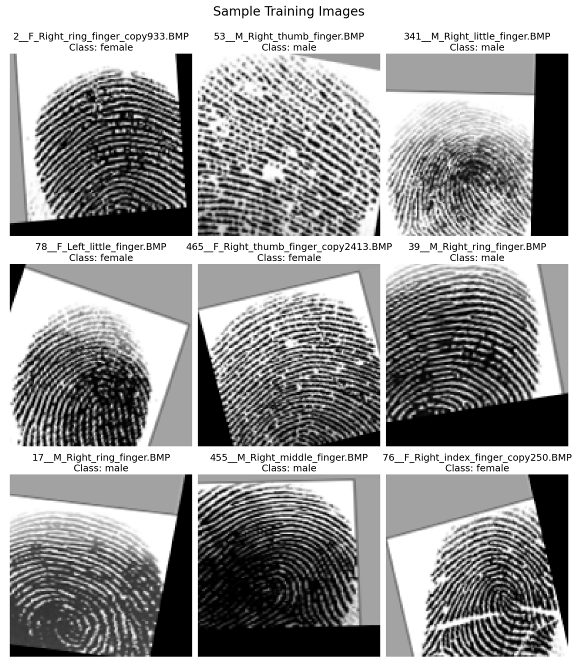

Fig. 1. Sample training images after data augmentation process

E. The Effect of Filters on Fingerprint Images

CNNs apply filters to images during convolutional operations to extract meaningful features as part of their learning process. To provide a clearer understanding of how these filters interact with fingerprint images, four examples are plotted in the figure 2 showcasing how different filters can enhance specific characteristics. For instance, edge-enhancing filters and pattern-emphasizing transformations were applied to highlight key features such as ridges and edges in the fingerprint. These visualizations offer valuable information about how the model might interpret and process features during training, helping to understand the role of convolutional operations in extracting relevant details from raw input data.

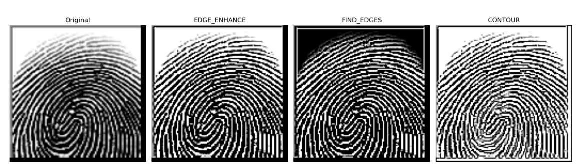

Fig. 2. Filter effects for feature extraction

V. AI Model Building

A. Model Architecture

The Convolutional Neural Network (CNN) designed for this task consisted of the following components:

- **Input Layer:** The model takes images maintaining the original dimensions as input.
- **Convolutional Layers:** Three convolutional blocks were used, each comprising:
  - **Conv2D Layers:** Extract features using 32, 64, and 128 filters, respectively, with kernel sizes of (3×3) and padding set to 'same'.
  - **Activation:** ReLU was used as the activation function.
  - **Pooling Layers:** MaxPooling was applied with a pool size of (2×2) to reduce spatial dimensions.
  - **Batch Normalization:** Added after each pooling layer to stabilize and accelerate training.
- **Fully Connected Layers:**
  - A Dense layer with 256 neurons, followed by a Dropout layer (rate = 0.5) to prevent overfitting.
  - The final Dense layer served as the output layer.
- **Output Layer:** Initially, the output layer used a sigmoid activation function for binary classification. However, this was later replaced with a softmax activation function to handle categorical labels more effectively, yielding better results.

B. Training Configuration

The model was trained with the following configurations:

- **Loss Function:**
  - Binary Cross-Entropy: Used during initial experiments for binary classification.
  - Categorical Cross-entropy: Adopted in the final implementation with a softmax activation function, as it yielded better classification results.
- **Optimizer:** Adam optimizer with an adaptive learning rate.
- **Learning Rate Scheduler:**
  - **ReduceLROnPlateau:** Monitored validation loss and reduced the learning rate by a factor of 0.3 after 3 epochs without improvement, with a minimum learning rate of 1×10^{-6}.
- **Metrics:** Accuracy was used as the primary evaluation metric during training.
- **Callbacks:** Early stopping was implemented to terminate training if validation performance did not improve after 8 epochs, preventing overfitting and saving resources.

C. Model Observations

Switching to categorical cross-entropy with a softmax output layer significantly improved the model's classification performance compared to the binary cross-entropy approach. The softmax activation allowed the model to better handle the two-class problem in this specific context, ensuring a more balanced learning process.

Furthermore, the decision to gradually increase the number of convolutional filters across layers allowed the network to capture lower-level features initially and progressively learn more abstract and complex patterns deeper in the network.

The model was trained over 30 epochs; however, training often stopped earlier due to the early stopping criterion, which monitored validation performance. This ensured that the model avoided overfitting and generalized well to unseen data.

These architectural and training adjustments underscore the importance of iterative experimentation and optimization when designing convolutional neural networks for domain-specific challenges such as fingerprint classification.

VI. Evaluating the Results

A. Model Training Dynamics

Figure 3 shows the evolution of the training and validation accuracy, as well as their loss over the epochs. The learning rate adjustments, managed by the ReduceLROnPlateau callback, are also depicted in the third plot. These metrics provide a clear overview of the model's convergence behavior and its generalization ability during training.

The training accuracy steadily improves, reaching 64.56%, while validation accuracy fluctuates around 0.64. Training loss consistently decreases, indicating effective learning, whereas validation loss shows some irregularities but remains relatively stable. The learning rate plot demonstrates the model's adaptation during training, with reductions at specific epochs to fine-tune performance.

B. Model Predictions

The evaluation of the trained CNN model was performed using the reserved test set comprising 600 samples evenly distributed between the "male" and "female" classes. The model achieved a test accuracy of 67.33% and a test loss of 0.611, showing a notable improvement over earlier iterations.

To provide information on the performance of the model, a sample of predictions from the test data set is visualized in Figure 4. Each image is displayed alongside its corresponding true class and predicted class. Correct predictions align the true and predicted labels, while mismatches highlight errors made by the model.

C. Confusion Matrix

Figure 5 displays the confusion matrix, which provides insights into the model's ability to correctly classify each class. While the model performs better in identifying "male" fingerprints, it still demonstrates some challenges in distinguishing "female" fingerprints, as evident by the misclassifications.

D. Classification Report

Table I presents the precision, recall, and F1-score for each class. The metrics indicate that while the model shows a relatively balanced performance across classes, there is room for improvement in recall for the "female" class.

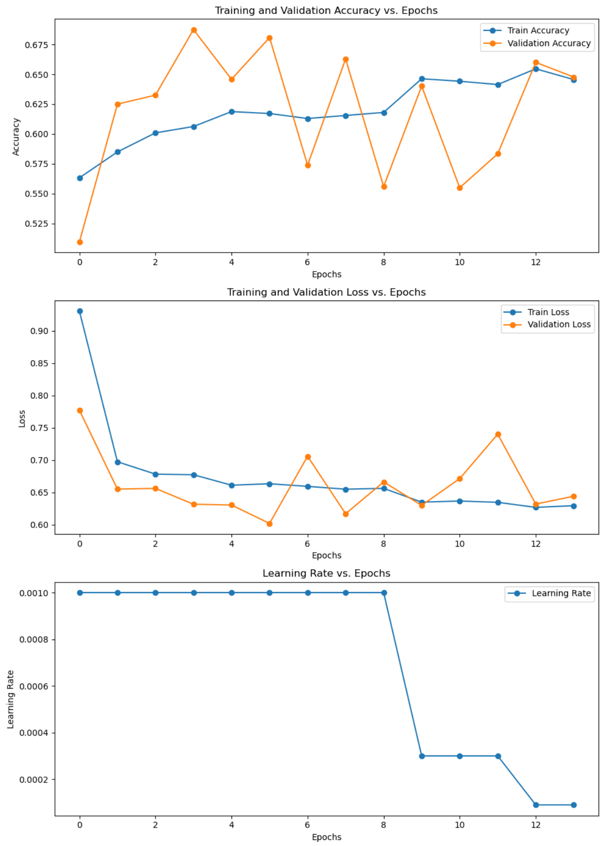

Fig. 3. Evolution of training and validation metrics over epochs

TABLE I
CLASSIFICATION REPORT FOR TEST DATASET

Class	**Precision**	**Recall**	**F1-Score**	**Support**
Female	0.70	0.60	0.65	300
Male	0.65	0.75	0.70	300
Accuracy	0.67 (600 samples)			
Macro Average	0.68	0.67	0.67	600
Weighted Average	0.68	0.67	0.67	600

E. Key Findings

Key observations derived from the evaluation include:

- The classification of "male" fingerprints exhibited higher recall (0.75) compared to "female" fingerprints (0.60), suggesting that certain features are more distinctive in "male" fingerprints.
- Adjustments to the model's architecture, particularly in the number of filters in convolutional layers, significantly improved performance.
- The use of learning rate adjustment (ReduceLROnPlateau) allowed the model to stabilize during training, resulting in better convergence.

Despite these improvements, challenges in distinguishing between classes persist, highlighting the inherent complexity of the task.

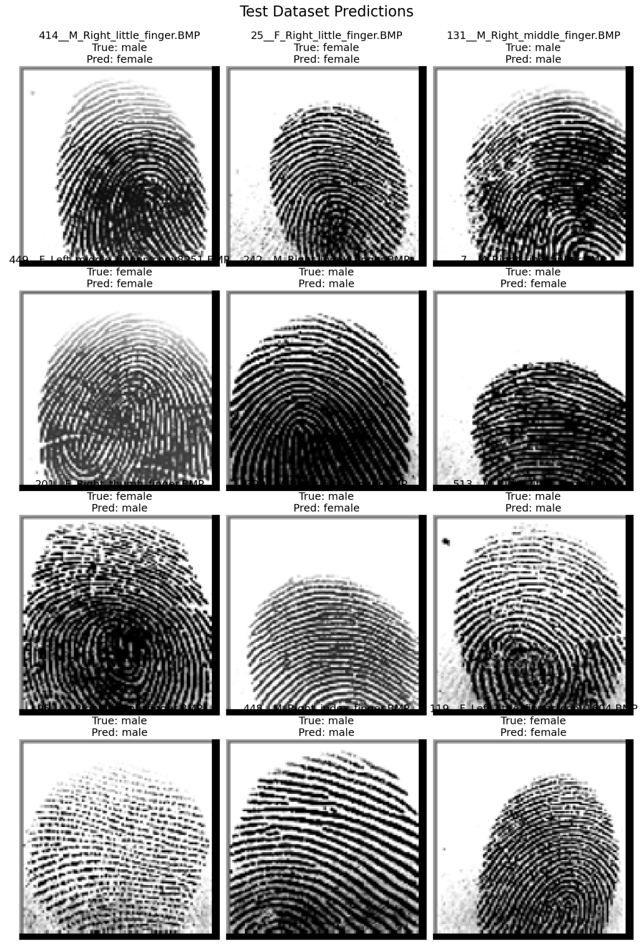

Fig. 4. Sample predictions from the test dataset

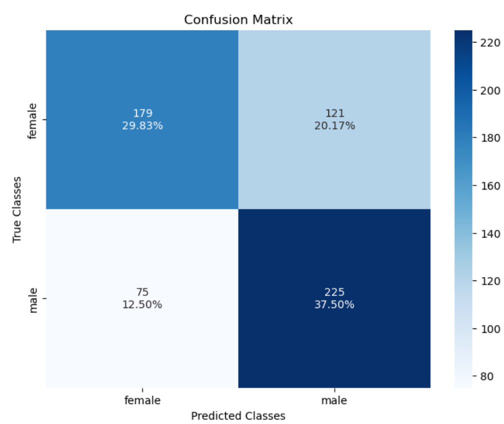

Fig. 5. Confusion Matrix for Test Dataset Predictions

VII. CONCLUSION

The results of this study demonstrate that CNN-based architectures can offer moderate success in gender classification from fingerprint images, but also underscore the complexity of gender classification based on fingerprint images. While the final model achieved a moderate test accuracy of 67.33%, with relatively balanced performance across both classes, the results suggest significant challenges remain.

A. Limitations and Challenges

Key limitations affecting performance include:

- **Dataset limitations:** Publicly available datasets for fingerprints are limited in quantity and often lack the diversity required for robust generalization.
- **Inherent complexity of the task:** Fingerprint patterns may not exhibit sufficient distinctiveness for gender classification, especially in the absence of domain-specific preprocessing techniques.
- **Technical constraints:** While the CNN architecture used in this study showed moderate success, deeper architectures or transfer learning methods might be needed to achieve significant improvements.

B. Implications and Applications

The ability to classify gender from fingerprint images has several potential applications and implications, particularly in forensic science, as well as broader scientific and judicial advancements:

- **Forensic investigations:** Gender classification can aid investigators in narrowing down suspects in criminal investigations, particularly when fingerprints are one of the few pieces of evidence available.
- **Identity reconstruction:** In scenarios such as natural disasters or accidents, gender classification from incomplete fingerprint evidence can support identity reconstruction efforts.
- **Judicial applications:** Automating gender classification could provide additional tools for verifying identity claims in judicial processes, improving objectivity and reducing human error.
- **Advancements in biometric research:** Understanding gender-based fingerprint patterns could drive further innovations in biometric technologies, expanding their use cases beyond traditional identification tasks.

The integration of AI in forensic sciences opens the door to unprecedented levels of objectivity and efficiency. For instance, AI has already been shown to detect patterns in fingerprints that human experts often overlook, such as intra-personal similarities across fingerprints of different fingers, as highlighted in recent studies. These advancements suggest the potential for AI to refine existing forensic methodologies and even uncover new patterns previously undetectable by human analysis.

VIII. Future Work

To address the limitations encountered in this study and further enhance the performance of gender classification from fingerprint images, several directions for future research are proposed:

- **Dataset Expansion:** Acquiring larger, high-quality fingerprint datasets that encompass diverse populations and fingerprint patterns is essential. This will improve the model's generalization capabilities across various demographic groups.
- **Advanced Data Augmentation:** Exploring more sophisticated data augmentation techniques, such as GAN-generated fingerprints or geometric and photometric transformations, could help increase the variability and size of the dataset.
- **Deeper Architectures and Transfer Learning:** Implementing deeper CNN architectures or leveraging transfer learning with pre-trained models on biometric or image classification tasks might significantly improve performance by capturing more complex patterns in the data.
- **Domain-Specific Feature Extraction:** Collaborating with forensic science experts to integrate domain knowledge into the feature extraction process. This could include identifying gender-specific minutiae or ridge patterns in fingerprints that are not easily detectable by current AI methods.
- **Exploring Hybrid Models:** Investigating hybrid approaches that combine CNNs with other machine learning models, such as ensemble techniques or classical methods, to improve robustness and accuracy.
- **Multi-task Learning:** Expanding the scope of the model to perform multiple tasks simultaneously, such as gender and age prediction or even identifying unique characteristics within fingerprint groups.
- **Real-world Applications:** Testing the model in real-world forensic scenarios to evaluate its utility, reliability, and performance in practical applications.

This future work aims to overcome the current study's challenges and limitations, paving the way for more accurate, robust, and impactful AI-based solutions in biometric and forensic sciences.

References

[1] G. Guo, et al., "Unveiling intra-person fingerprint similarity via deep contrastive learning," *Science Advances*, vol. 10, no. 2, pp. 1-12, Jan. 2024. [Online]. Available: https://doi.org/10.1126/sciadv.adi0329

[2] M. Taylor, et al., "Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach," *NIST Interagency/Internal Report (NISTIR) 7842, National Institute of Standards and Technology*, Gaithersburg, MD, 2012. [Online]. Available: https://doi.org/10.6028/NIST.IR.7842

[3] M. Patel and U. Singh, "Age and Gender Recognition using Deep Learning Technique," *2023 3rd International Conference on Smart Data Intelligence (ICSMDI)*, Trichy, India, 2023, pp. 238-245, doi: 10.1109/ICSMDI57622.2023.00052.

[4] S. Hamdi and A. Moussaoui, "Comparative study between machine and deep learning methods for age, gender and ethnicity identification," *2020 4th International Symposium on Informatics and its Applications (ISIA)*, M'sila, Algeria, 2020, pp. 1–6, doi: 10.1109/ISIA51297.2020.9416549.

[5] S. Patil, "Gender Classification: An Integrated Analysis of Multimodal Biometric Data Using Deep Learning Techniques," *International Journal of Research Publication and Reviews (IJRPR)*, vol. 4, no. 10, pp. 2855-2865, Oct. 2023. [Online]. Available: https://ijrpr.com/uploads/V4ISSUE10/IJRPR18480.pdf

[6] Kaggle, "Fingerprint Dataset." [Online]. Available: https://www.kaggle.com/datasets/ruizgara/socofing/data, [Accessed: Dec. 8, 2024]