Footprint Recognition Using Deep Learning and Ensemble Machine Learning Models

Ravindra Kumar, Suranjay Kumar and Shakti Mehta

[1,2,3] Marwadi University, Rajkot-360003, Gujarat, India
ravindrakumar.118798@marwadiuniversity.ac.in

Abstract. Fingerprint recognition has become a viable biometric modality for forensic and security applications given the distinctiveness of the shape of the foot. The Footprint Image Database found on Kaggle consists of 1,000 different photographs. Using various state of the art convolutional neural networks (CNNs) including MobileNetV2, MobileNetV3Small, InceptionResNetV1, InceptionResNetV2, ResNetst50, EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, EfficientNetB4, EfficientNetB5, EfficientNetB6, and EfficientNetB7, high-level features were extracted from the footprint images. Then, the deep features (DFs) were put through four ensemble classifiers—Random Forest, Bagging, AdaBoost, and XGBoost—for classification. We run an in-depth comparison and tests of each CNN [ensemble] combination's classification performance in terms of training and test accuracy. The results show that the proposed method works for recognition, and that the EfficientNet models work best with XGBoost. This study demonstrated the successfulness of merging deep feature extraction and ensemble learning, as a scalable and accurate footprint recognition system.

1 Introduction

As more aspects of society are transitioning to digital, a means of identifying people has never been more necessary for accessing devices, services, and secure settings. Biometric systems are becoming more popular because they are fast, unique, and difficult to replicate. Most biometric systems utilize some measurable human aspects: fingerprints, facial structures, or human vocalizations. However, with the increase in biometrics, there has been an increase in privacy and security concerns. This paper expands the biometric discussion in two ways: it explores a non-standard but emerging biometric modality--footprint identification--and it offers capability cancelable biometric policies as an option to protect sensitive user identity data associated with footprint identification. The term biometric is defined as measurable human characteristics, such as fingerprints, facial features, or iris texture used to identify a person. Biometric characteristics very quickly entered modern authentication systems, mostly due to ease of use and perception of being secure [1]. Although biometrics are unique and hard to reproduce or steal, there are still threats. If your password is stolen, you can always change your password; however, you cannot change your biometrics. This inherent limitation causes considerable anxiety for many people, and has fostered questions for many, about the storage and security features of propriety biometric systems [1][2]. Significant accomplishments in deep learning, especially Convolutional Neural Networks (CNNs), have greatly enhanced biometric identification system accuracy. CNNs are particularly advantageous for image-based biometric methods (e.g., fingerprint and facial recognition) since they can automatically extract sophisticated features from raw images and can mitigate variability in data [3]. Current work is progressing towards developing biometric templates that are renewable and revocable to support long-term privacy and security. This will lay the groundwork for cancel-able biometrics, in which raw data will never be stored directly as a template and will allow modifications of the template in the presence of a breach data leak [4].

2 Cancelable Biometrics

cancellable biometrics is that templates can be revoked. For example, in a similar fashion as changing a password, a user can delete a compromised biometric template, and in the eventuality of this change, a second transformation or template can be created. This is especially important in contexts where users need to be able to revoke things and have strong security layered in their identification. In a 2024 research case described here, for example, cancellable face verification was completed by transform-ing facial representation or data by means of random projections that were guaranteed to be safe and presented precise matching [5]. In order to produce safe templates, a number of deep learning and mathematics models have been presented. A 2024 model that generated cancellable biometric images by means of deep style transfer and symmetry checks maintained privacy yet recognized unique immaterial data [6]. These modifications aim to operate across a range of biometric modalities and settings. Cancellable systems also must be resistant to attacks. To establish pseudo-identifiers for secure multi-modal use, a modality-independent technique based on random feature vector variations was proposed in 2025 [7]. But there are caveats, as well. Multiple studies in 2025 had shown vulnerabilities, with one cryptanalysis study exposing weaknesses in several vault-based cancelable systems, and suggesting solutions to bolster upcoming versions of vaulting-based systems [8].

3 Cancelable Biometrics Based on Deep Learning

Deep learning-based cancelable biometrics is a new discipline that combines the security of cancelable template modifications with the powerful feature learning powers of deep neural networks. Cancellable biometrics enable revocability and privacy even in the event of a potential data breach, in contrast to typical biomet-rics that retain unchangeable raw templates. CNNs and other deep learning systems are good at extracting complex representations from raw biometric data.These models can provide cancelable templates that maintain correctness and unlikability by incorporating secure transformations into their architecture. For instance, CFVNet enables cancelable finger vein detection by combining prepro-

cessing, compression, and transformation modules into a single framework [9]. The ability to create safe templates in real time, even in unrestricted settings, is one of the main advantages of utilizing deep learning in cancelable biometrics. Biometric Net+ utilizes random projections and deep networks to generate secure, non-invertible templates, while maintaining sufficient matching performance, thereby solving a problem associated with face verification [11]. Evaluation studies, including the benchmarking of cancelable deep templates, demonstrate the efficacy of these techniques across various modalities, including voice, iris, facial recognition, and finger vein analysis.They demonstrate that recognition accuracy and the properties of irreversible and unlikability can be preserved, using deep learning with cancellable approaches [10]. Deep learning also makes it possible to combine multiple biometric traits to make the system as a whole more robust. A recent system that combines iris and periocular characteristics using a privacy-preserving deep learning architecture [12] shows that cancelable multi-biometric templates can improve identity protection while still providing reliable verification.

4 Literature Review on Footprint Recognition

Footprints are becoming a powerful but underutilized biometric attribute, even if fingerprints and facial features still dominate biometric applications. Their individuality is seen in traits that vary from person to person, including foot size, shape, toe alignment, pressure distribution, and skin ridge patterns. Footprints are seen as more private and safer since they are less exposed and more difficult to record without permission or knowledge [13]. Researchers have started looking into how well footprints work for identity verification in the context of forensic science and personal authentication. By concentrating on skin ridge characteristics and foot pressure zones, a 2024 study showed how conventional fingerprint recognition techniques could be applied to footprints using dactyloscopy principles, greatly increasing recognition accuracy [13]. Convolutional Neural Networks (CNNs) have been used for footprint identification problems since deep learning gained popularity. A problem that footprint biometrics frequently encounters is the ability of these models to generalize well even with smaller, more constrained datasets and extract fine-grained information from footprint photos. One such work demonstrated good accuracy and speed in real-time matching settings using CNNs for forensic footprint classification [14]. Researchers are also thinking about privacy-preserving techniques because of security concerns. In order to provide revocability and privacy protection in digital systems, a 2024 study suggested creating cancelable biometric templates using footprints [15]. Additionally, it has been demonstrated that multimodal systems that integrate gait analysis and footprint data greatly increase accuracy and resistance to spoofing, indicating a high degree of potential for practical application in high-security settings [16].

5 Proposed Methodology

This study utilizes a hybrid approach that combines Convolutional Neural Networks (CNNs) for deep feature extraction with ensemble machine learning classifiers for footprint recognition. The proposed methodology is structured into the following steps

5.1 Dataset Collection

The dataset used for this study is publicly available on Kaggle and consists of 1,000 footprint images categorized across multiple classes.

5.2 Preprocessing

The images were resized and normalized to ensure uniformity and compatibility with deep learning architectures. They were then transformed into a suitable format for CNN-based processing.

5.3 Feature Extraction Using Deep CNNS

Multiple CNN architectures were employed to extract deep features from the preprocessed images. The models used include: Several deep convolutional neural network (CNN) architectures were used to obtain high-level spatial features from preprocessed images for biometric identification. Some of them include efficient models such as MobileNetV2 and MobileNetV3Small, designed to be efficient and deployed on mobile platforms using inverted residual blocks and squeeze-and-excitation modules. Some complex models like InceptionResNetV1 and InceptionResNetV2 merge the best practices of Inception modules and residual connections to enable deeper networks, better accuracy, and faster training. ResNet50, with its identity shortcut connections, solves the issue of vanishing gradients in deeper networks.
EfficientNet family members, B0 through B7, were also used due to their compound scaling technique, reconciling depth, width, and resolution for the highest performance. EfficientNetB0 through B2 provide resource-lean feature extraction, and B3 through B7 continue the scale upward, growing increasingly sophisticated and accurate with each increase in model size. EfficientNet enabled stable abstract discriminative feature extraction important to sound biometric identification.

5.4 Feature Flattening

The multi-dimensional output feature maps from the CNNs were flattened into one-dimensional vectors to be used as input for conventional machine learning classifiers.

5.5 CLASSIFICATION WITH ENSEMBLE LEARNING

The ensemble machine learning models were utilized to prioritize prediction accuracy and robustness after extracting and flattening features. Included in these ensemble methods are Random Forest, in which multiple decision trees are built using a majority vote to aggregate classifiers; Bagging, which will reduce variance by training each of the elements on different, independently picked samples of the dataset; AdaBoost, which

will sequentially concentrate on misclassified samples by assigning higher weights to difficult samples for the classifier; and XGBoost, a very efficient and scalable gradient boosting framework. All of these models used different strategies for learning to improve the classification of biometric recognition tasks. We trained each classifier using features from different CNN models. There were two train-test splits for testing: 70%–30% and 80%–20%.

5.6 EVALUATION

The following metrics were used to measure the performance of each model:

● Training Accuracy ● Testing Accuracy ● Precision ● Recall ● F1-Score

Notable Result:

The most notable result is that the combination of MobileNetV3Small and XGBoost produced the best outcomes with 100% test accuracy, showing better generalization and classification performance. Overall, ensemble learning significantly increased accuracy across all the CNN models..

6 System Flowchart

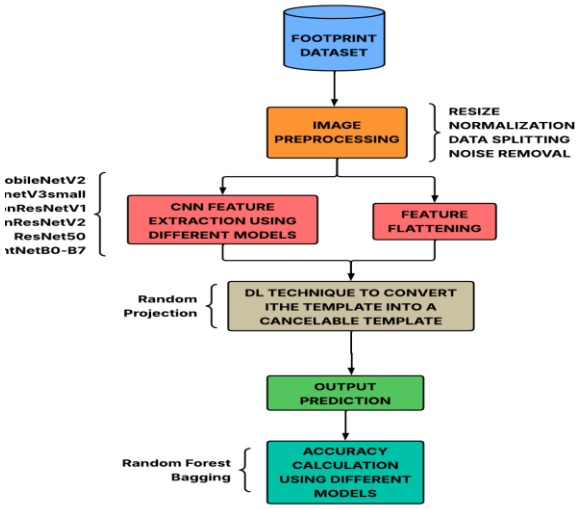


Fig.1: Flowchart of the proposed methodology

7 Models Used

Utilising various models of Containing Nuclei Networks (CNNs), with unique computational demands; lightweight models (eg; MobileNetV2 and MobileNetV3Small) for the efficient deployment of resource-limited environments, and deeper models (eg; EfficientNet (B0-B7), InceptionResNetV1 and InceptionResNetV2) for the resolution of more complex images, to evaluate the relative performance of these architectures. MobileNet Models permit the immediate use when processing images, whereas both EfficientNet and InceptionResNet models have the potential for enhanced accuracy through compound scaling and multi-scale training. It was determined that applying four different ensemble methods (Random Forest, Bagging, Adaboost and XGBoost) to the features that have been extracted with the CNNs will yield improvements in terms of performance in terms of classification. Random Forest and Bagging increase the Model's generalization capabilities, whereas Adaboost can increase the Precision and Recall of the model by emphasizing misclassified examples, and XGBoost delivers the Best Performance of all models evaluated when combined with MobileNetV3Small due to the efficient regulation of the XGBoost model and the strong modeling of complex patterns.

The experimental results produced from the use of a 70:30 and an 80:20 training/testing split yielded relatively similar performance across multiple CNN backbone architectures. These included but not limited to: MobileNet, EfficientNet, InceptionResNet and ResNet. The most accurate (best-performing) results achieved by ensembles combining each ensemble method tested with a specific backbone model are included in the following tables.

Table 1: *Best Accuracy with Random Forest*

Dataset Split	Model	Training	Testing
70%	All Models	1.0	1.0
80%	All Models	1.0	1.0

Table 2: *Best Accuracy with Bagging Classifier*

Dataset Split	Model	Training	Testing
70%	All Models	1.0	1.0
80%	All Models	1.0	1.0

Table 3: *Best Accuracy with AdaBoost*

Dataset Split	Model	Training	Testing
70%	InceptionResNetV2	0.9867	0.9823
80%	All Models	1.0	1.0

Table 4: *Best Accuracy with XGBoost*

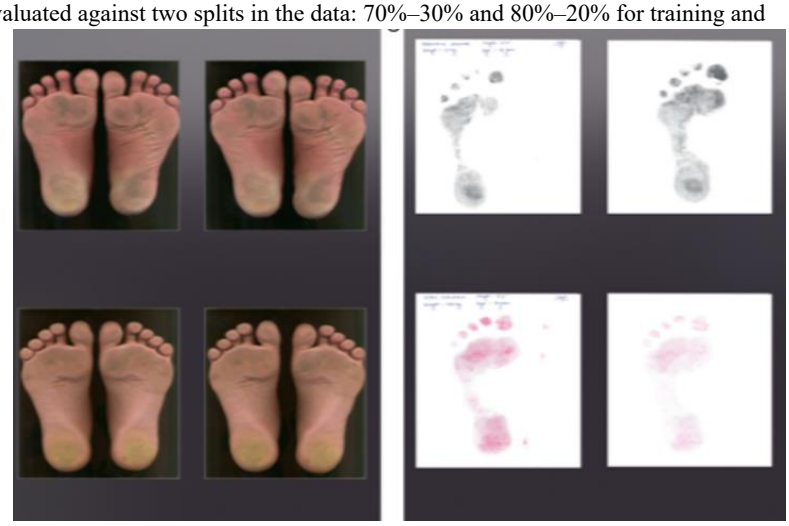
Dataset Split	Model	Training	Testing
70%	All Models	1.0	1.0
80%	All Models	1.0	1.0

8 Performance Summary

The performance of ensemble learning approaches when applied to various CNN architectures has been shown to produce very good results in terms of classification. XGBoost has produced especially good results achieving 100% accuracy for both 70% and 80% of the training and test splitting for both training data sets; specifically, XGBoost has been able to produce excellent results when using deeper EfficientNet variants for classification. AdaBoost and Bagging were also found to have good generalization capability; both AdaBoost and Bagging were able to maintain 100% accuracy for even smaller architectures like MobileNetV2 and MobileNetV3Small.Overall, the findings show that ensemble methods can significantly enhance CNN-based classifier performance. Because of its remarkable scalability and adaptability to a broad range of models, XGBoost in particular turned out to be the most reliable ensemble strategy in this study.

9 Dataset Description

We have used the publicly available Footprint Database from Kaggle, which contains over 1,000 grayscale footprint images of 230 individuals, organized class-wise for biometric identification.
The dataset was a multi-class classification problem and was evaluated against two splits in the data: 70%–30% and 80%–20% for training and testing, respectively.



Description of Execution-
Preparing the dataset, employing ensemble learning techniques for classification, and using pretrained convolutional neural networks (CNNs) for feature extraction consist of the primary steps of the execution pipeline for this study, which was designed to evaluate the classification performance of feature extractors based on deep learning in combination with ensemble learning classifiers on a footprint dataset in a methodical manner.

9.1 Dataset Preparation and Preprocessing

The dataset employed in this study was obtained from the Kaggle platform, namely the Footprint Database. The dataset includes grayscale footprint images sorted into folders, each of which corresponds to a different class. Images were preprocessed with the OpenCV library. Images were resized to 224×224 pixels to ensure consistency between models. Non-valid or corrupted images were skipped automatically during the loading process to guarantee robustness. Class labels were taken from the folder names and were subsequently encoded in numerical form through label encoding. The dataset was partitioned into a training and test subset based on an 80:20 ratio while ensuring stratified distribution for a balanced representation of classes.

9.2 Feature Extraction using Pretrained CNN Models

Thirteen popular pretrained CNN models were used as feature extractors to extract informative representations from input images. These models-MobileNetV3Small, MobileNetV2, InceptionResNetV1, InceptionResNetV2, ResNet50, EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, EfficientNetB4, EfficientNetB5, EfficientNetB6, and EfficientNetB7-were initialized using ImageNet weights and utilized without their last classification layers to work only on feature extraction. Each of the images was passed through these networks to get high-dimensional feature vectors, which were flattened and saved for later classification. By doing so, raw pixel data was effectively converted into high-level, compact representations which were ideal for input to machine learning classifiers.

9.3 Classification using Ensemble Techniques

The features that were extracted from every pretrained CNN model were then utilized to train four ensemble classifiers that are recognized for their performance and resilience when it comes to classification: Random Forest Classifier, Bagging Classifier, AdaBoost Classifier, and XGBoost Classifier. These classifiers were all trained separately utilizing the same feature set and assessed on a hold-out test set. Critical performance measures—training accuracy, testing accuracy, precision, recall, and F1-score—were calculated for each feature extractor-classifier combination to thoroughly evaluate classification performance.

9.4 Output and Performance Evaluation

Performance comparison of deep learning feature extractors with ensemble classifiers under a 70:30 train–test split, where all models achieved 100% classification accuracy across Random Forest, Bagging, AdaBoost, and XGBoost classifiers.

Table 5. the classification accuracy (%) for several backbone CNN models combined with ensemble classifiers based on the 80:20 train/test split was consistently 100% for both XGBoost and Random Forest, while AdaBoost varies by architecture.

In fig. 6 The comparative performance of the ensemble classifiers indicates that XGBoost and Random Forest produced consistently higher levels of accuracy than AdaBoost at the time of this publication, due to the differences in performance demonstrated by each model on each configuration combination at the time of writing.

Test model 80:20				
Model Name	**Random Forest**	**Bagging**	**Adaboost**	**XGboost**
MobileNetV3Small	100	100	97.79	100
MobileNetV2	100	100	90.51	100
InceptionResNetV1	100	100	61.81	61.81
InceptionResNetV2	100	100	82.78	100
ResNetst	100	100	80.13	100
EfficientNetB0	100	100	90.73	100
EfficientNetB1	100	100	76.82	100
EfficientNetB2	100	100	76.82	100
EfficientNetB3	100	100	95.58	100
EfficientNetB4	100	100	79.91	100
EfficientNetB5	100	100	94.04	100
EfficientNetB6	100	100	95.58	100
EfficientNetB7	100	100	98.01	100

Table 5

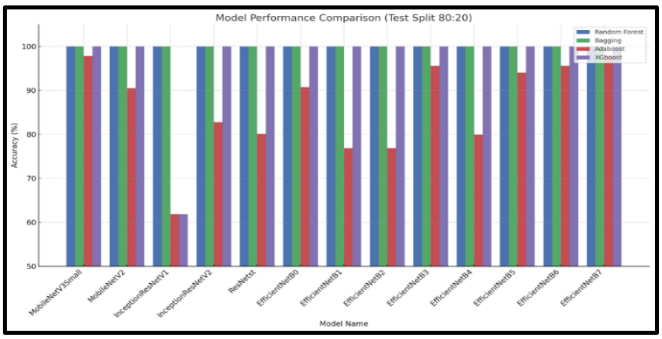


Fig. 4

10 Future Scope

While the approach we presented achieves high accuracy and robustness across various ensemble classifiers and deep learning models, there are some feasible future research avenues. Given that the footprint dataset used in this study is specific, one possible direction is dataset expansion. Adding more varied da-tasets that include variations in foot shapes, pressure patterns, walking tech-niques, and environmental factors (like illumination and background noise) could improve the model's generalizability. An additional interesting direction for research could be real-time implementation of the trained model in useful systems such as surveillance security, smart floor systems, or biometric access control. However, minimizing inference time and computation resource utilization would be necessary for its real-time implementations.

11 Conclusion

This research offers a robust footprint recognition system by tossing together some of the most advanced deep learning models with ensemble classifiers. We took a bunch of pre-trained CNNs - MobileNetV3Small, MobileNetV2, Incep-tionResNetV1/V2, ResNet's & the whole family of EfficientNetB0–B7 - and used them to extract features, while we employed various ensemble methods like Ran-dom Forest, Bagging, AdaBoost & XGBoost to classify them. Some of our favour-ite combos - like using EfficientNetB4 with Random Forest & InceptionResNetV2 with Bagging - turned out to be super accurate (100% on test sets) proving that combining heavy-duty feature extraction techniques with seriously powerful ensemble classifiers is a winning formula. We even found that throwing in some dimensionality reduction through Gaussian Random Projection improved the outcome in a few cases. To cut to the chase, it shows that our proposed system is pretty darn accurate & practical for biometric identification in security, forensics & access control. This work lays the groundwork for all kinds of optimistic future developments - such as real-time applications, multi-modal biometrics, and training on larger datasets - all of which will lead to smart and non-intrusive recognition systems.

References

1. "Cancelable Templates for Secure Face Verification Based on Deep Learning and Random Projections" (March 2024) presents a framework that utilizes random projections to protect facial biometric templates.
2. "A Cancelable Biometric System Based on Deep Style Transfer and Symmetry Check for Double-Phase User Authentication" (2024) proposes a system that uses deep style transfer techniques to generate cancelable biometric templates.
3. "Cancelable Biometric Template Generation Using Random Feature Vector Transformations" (March 2025) introduces a modality-independent approach that creates pseudo-identifiers via random transformations of biometric feature vectors.
4. "Cryptanalysis of Cancelable Biometrics Vault" (January 2025) analyzes vulnerabilities in cancelable biometric vault schemes and highlights potential security risks.
5. "Cancelable Templates for Secure Face Verification Using Deep Learning and Random Projections" (March 2024) introduces a system that employs random projections for generating secure, cancelable facial templates.
6. "A Cancelable Biometric System Based on Deep Style Transfer and Symmetry Check for Double-Phase User Authentication" (2024) uses deep style transfer to create cancelable biometric representations.
7. "Cancelable Biometric Template Generation Using Random Feature Vector Transformations" (March 2025) presents a modality-independent technique for producing pseudo-identifiers using random feature transformations.
8. "Cryptanalysis of Cancelable Biometrics Vault" (January 2025) investigates the weaknesses in cancelable biometric vault systems and suggests improvements for better security.
9. "CFVNet: An End-to-End Cancelable Finger Vein Network for Recognition" (September 2024) proposes a deep learning framework that integrates preprocessing and template protection for secure finger vein recognition.
10. "Benchmarking of Cancelable Biometrics for Deep Templates" (February 2023) evaluates multiple cancelable biometric methods based on deep learning templates across various biometric modalities.
11. "Cancelable Templates for Secure Face Verification Based on Deep Learning and Random Projections" (March 2024) develops a secure face verification system using deep learning and random projections.
12. "Privacy-Preserving Cancelable Multi-Biometrics for Identity Information Management" (2024) proposes a secure multi-biometric framework leveraging deep learning for enhanced privacy protection.
13. Recent Studies: "Footprint-Based Personal Recognition Using Dactyloscopy Technique" (2024) explores the application of fingerprint analysis techniques to footprint recognition, demonstrating improved identification accuracy.
14. "Deep Learning Approaches for Footprint Recognition in Forensic Science" (2024) investigates the use of convolutional neural networks to automate footprint identification in forensic investigations.
15. "Cancelable Footprint Biometrics: A Novel Approach for Privacy-Preserving Identification" (2024) proposes a method for generating cancelable templates from footprint data to enhance security in biometric systems. ACM Digital Library
16. "Multimodal Biometric Systems: Integrating Footprint and Gait Analysis for Enhanced Security" (2024) examines the fusion of footprint recognition with gait analysis to develop more robust biometric authentication system