Biometric Fingerprint Verification with Siamese Neural Network and Transfer Learning

[1]Suman M, [2]Shobha N, [3]Ashoka S B and [4]Job Prasanth Kumar Chinta Kunta
[1]Department of Artificial Intelligence and Data Science, BMS College of Engineering, Bangalore, Visvesveraya Technological Univerisity, Belagavi, Karnataka, India.
[2]Department of Computer Science and Design, Dayananda Sagar College of Engineering, Bangalore, Visvesveraya Technological Univerisity, Belagavi, Karnataka, India.
[3]Department of Computer Science, Maharani Cluster University Bangalore, Bengaluru (Bangalore), Karnataka, India.
[4]Lead Solution Data Architect, The Automobile Association, London, United Kingdom.
[1]sumanm@gmail.com, [2]shobha.venk20@gmail.com, [3]dr.ashoksbcs@gmail.com, [4]job.chintakunta@gmail.com,

Correspondence should be addressed to Suman M: sumanm@gmail.com

Article Info
Journal of Machine and Computing (https://anapub.co.ke/journals/jmc/jmc.html)
Doi : https://doi.org/10.53759/7669/jmc202606003
Received 23 April 2025; Revised from 18 August 2025; Accepted 01 October 2025.
Available online 13 October 2025.

Abstract – In this research, a fingerprint verification model, equipped with a Siamese Neural Network structure with MobileNetV2 and ConvNeXt models, was established in developing a robust fingerprint verification system using transfer learning. The algorithm is learned using a custom-made fingerprint database with image pairs dubbed as similar or different that facilitates the model to recognize the discriminative characteristics of biometric comparison. The Siamese structure enables the system to have generalization of unseen fingerprint classes without retraining again to be scalable and adaptable during applications in the real world. This MobileNetV2 performed very well with an accuracy of 96.35 and F1-scores of 0.96 on both classes and was very capable of differentiating between similar and dissimilar fingerprints. However, it was surpassed in accuracy (98.61% higher) and balanced F1-scores (0.99 higher) and on the generalization and classification error superiority by ConvNeXt. The overall scores of the two models were 1.00 with an area under the receiver operating characteristic curve, indicating that they were in a perfect state of class separability. ConvNeXt also demonstrated a cleaner convergence curve during the training process, a fact that further confirmed its superiority in cases of dynamic biometric verification. In general, the findings indicate the viability of deep learning based Siamese networks and elaborate convolutional networks in terms of scalable and precise fingerprint crack verification solutions.

Keywords – Fingerprint Verification, Siamese Neural Network, MobileNetV2, ConvNeXt, Transfer Learning, Biometric Authentication.

I. INTRODUCTION

Fingerprint verification is one of the most popular biometrics processes as it is quite reliable, unique and easy to obtain. Fingerprints being an important element in the identity verification system forms a broad aspect of applications including easy and secure access control as well as national identification systems [1][2]. The incorporation of machine learning and especially deep learning has made a significant breakthrough in the fingerprint recognition sector since it allows for automating feature analysis and their categorization. Nevertheless, older deep learning paradigms may depend on excessive retraining in situations when new fingerprint classes appear, which does not allow them to scale in dynamic settings.

A solution to this problem is presented by Siamese Neural Networks (SNNs). Whereas in standard classification models, a similarity function between two inputs is learned, SNNs do not require retraining on new classes and are allowed to generalize existing approaches [3]. This property ensures that SNNs have wide application in those applications that demand frequent state updates, including massive security systems. Moreover, the use of transfer learning in conjunction with SNNs also contributes to the efficiency of the latter, as the pre-trained models extract the robust and discriminative features of the fingerprint images [4].

Over the past few years, new types of architectures, including the MobileNetV2 and ConvNeXt ones, have become potentially strong tools in using deep-learning solutions. MobileNetV2 is characterized by its lightweight and computationally efficient model, so it is perfect to use in resource-constrained settings. Meanwhile, ConvNeXt has better feature extraction capabilities and generalization; this is highly regarded as a solution for fingerprint verification tasks that require a high overall rate. These architectures in combination with SNNs have shown impressive results in the way they improve the overall performance of the biometrics systems [5].

Preprocessing and construction of fingerprint datasets take a central position in the success of deep learning models. SNN-based systems are based on labelled pairs of similar and dissimilar fingerprints, whereby the network learns the similarity measure that has been underlying researchers. Preprocessing methods, such as feature fusion and pair generation have been used to enhance the data quality and representativity of the training data. Besides data preparation, one important aspect in biometric systems is scalability. Most of the traditional techniques necessitate retraining their models to adapt to new classes and due to this; more calculation expenses are incurred and time is consumed. This challenge is what SNNs overcome since they rely on comparing feature embeddings, which enables the system to classify new samples of fingerprints on the fly. This capability has been known to be a major strength of this capability in contemporary biometric usage. Transfer learning can be used to increase the power and flexibility of SNNs more. MobileNetV2 and ConvNeXt, as examples of pre-trained models, allow the network to use already known information, making the network more independent of large data and have a quicker convergence. This has been proven true with studies showing that the fingerprint verification effectiveness and scalability increase with the help of transfer learning, thus making it a crucial part of modern biometric systems [6].

Instead, this work assesses the results of SNNs that are combined with MobileNetV2 and ConvNeXt to use on fingerprint verification. A dataset consisting of pairs of labelled fingerprints was created and pre-processed in order to allow successful training and testing of the models. The main aim was to understand the scalability of these models and how accurate they were, and the remarkable evaluation was whether they could dynamically add new classes without retraining. The results showed that ConvNeXt was better performing than MobileNetV2, having better generalization, precisions, and recalls and convergence-free training. The results in this paper provide evidence of the SNN-based systems as a scalable robust system to perform fingerprint verification even in dynamic applications. More so, the combination of SNNs with transfer learning to create flexible and efficient biometric systems is an active stride forward with regard to meeting the changing needs of real time deployment.

The organization of this paper is as follows: Section II covers related studies. Section III details the proposed methodology. Section IV discusses results and comparisons. The paper ends with essential results alongside possible future research in Section V.

II. LITERATURE SURVEY

The paper provides a review of the latest discoveries of deep learning models, especially MobileNetV2, ConvNeXt and Siamese networks, used to solve various tasks within the framework of image classification, biometric recognition, and medical counterproposal. It focuses on the betterment of precision, speed and feasible implementation and puts an emphasis on any mode of transfer learning, data augmentation and lightweight structures in the context of real-world applications.

MobileNetV2-Based Image Classification Models

MobileNetV2 has become a widely used lightweight convolutional neural network structure; it has become commonly used in the domain of image classification because of its efficiency and accuracy in resource-constrained environments. A model of fruit images classification by past researcher, which is built on the framework of a modified MobileNetV2 architecture augmented with deep transfer learning, or TL-MobileNetV2. With a massive data set containing 26,149 images and 40 classes of fruit, the model swapped out the classification head of MobileNetV2 in favor of five unique layers, as well as delivering an impressive 99 percent accuracy and beating the mean of the previously existing traditional architectures such as AlexNet, VGG16, InceptionV3, and ResNet. This paper describes the efficiency of transfer learning and dropout layers in diminishing overfitting and also increasing generalization. In the same way, researchers used MobileNetV2 to classify household waste into four categories with an accuracy of 82.92%, 15.15% better than a baseline CNN. Through their work, they showed that MobileNetV2 could implement lightweight, real time classification systems on mobile devices, like a WeChat applet, and yet some areas of concern were identified in situations where the waste is highly deformed or contaminated [7]. Researchers MobileNetV2 [8] further into the applicability range with the background of the classification of agricultural products, supplementing it with a Res-Inception module and an Efficient Multi-scale cross-space Attention (EMA) mechanism. The specified hybrid model performed better in comparison with the original MobileNetV2 on the Fruit-360 dataset by 1.86 percent especially in correct recognition of similar subcategories [9]. Future developments represented in the study also included augmenting images in the network and parameter optimizing the network to increase scalability and efficiency. In addition to agriculture, researchers applied MobileNetV2-based model to classify melanoma images into benign and malignancy. The model used transfer learning and a user-defined classification head to extract features and identified up to 85 percent accuracy on the ISIC-Archive dataset, superior to other heftier networks like ResNet50V2 and InceptionV3 [10]. The low weight of MobileNetV2 facilitated a possible development of early melanoma detection on the mobile, but there were still issues with the imbalance of data. Researches in face mask detection during the COVID-19 pandemic applied MobileNetV2 transfer learning and Deep Convolutional Neural Networks (DCNN) to record high accuracies of 98-99 percent in two different datasets. MobileNetV2 performed better as compared to DCNN, which further highlights that the model can be used to deploy real-life solutions in mitigating the pandemic. In totality, these results confirm the versatility of MobileNetV2 in a wide range of fields starting with agriculture, healthcare, and ending with the public security realms, primarily, when transferred and embedded in with a transfer learning framework and a tailored classifier head [11]. The stable improvement in accuracy in various domains confirms the flexibility and effectiveness of the MobileNetV2 to work in tandem with its transfer learning. **Table 1** summarizes some important literature applicable to the study with attention to relevant datasets and accuracy values to show the overall effectiveness of this method and such a base.

Table 1. Transfer-Learning & Advanced CNN Classification

Architecture / Approach	**Dataset**	**Accuracy (%)**
TL-MobileNetV2 [9]	26 149 fruit images (40 classes)	99
DCNN + MobileNetV2 [10]	2 500 + mask images	98–99
MobileNetV2 [11]	Fruit-360 (4 waste category)	82.92
Melanoma classification (MobileNetV2) [12]	ISIC-Archive (benign vs. malignant)	85
LSMSC (1D spectral + HiASPP) [13]	PaviaU / Salinas / Indian Pines	99.22 (PaviaU)

ConvNeXt and Advanced CNN Architectures for Specialized Classification

Although MobileNetV2 has entirely and successfully made use of lightweight application performance, newer current developments use deeper and more intricate designs such as ConvNeXt to be applied to special tasks in classification. Researchers developed a semi-supervised learning method of weeds recognition in the form of a ConvNeXt-based encoder and a custom decoder. Their model, through the use of consistency regularization, was able to exploit plentiful unlabeled data, as well as a few labelled ones, to outdo state of the art supervised and semi-supervised baselines on agricultural weed datasets [12][13][14]. The scheme tries to address the practical issues of high cost of data labelling in precision farming. Researchers proposes a high-accuracy method of classifying cervical precancerous lesions on a basis of ConvNeXt. They used data augmentation (self-supervised) and ensemble learning method (random forests) that improved the accuracy of the previous models by 8.85 percent on the DCCL cytology dataset. The authors confirmed effectiveness of augmentation and ensemble techniques with the help of many ablation experiments which are indicative of clinical diagnostic assistance. According to these studies, ConvNeXt and its updated convolutional architecture based on transformer structures are impressive in medical imaging and farming situations, especially when combined with complex data augmentation and learning planning [15]. These illustrations show that ConvNeXt and other sophisticated backbones provide high performance and generalization in medical or agricultural classification. **Table 2** gives a contrasting overview of these designs and the documented performance of the related architectures in recent studies.

Table 2. Multimodal & Behavioral Biometric Systems

Modality / Task	Dataset	Accuracy (%)
FP + face fusion (SVM) [16]	500 fingerprint+face pairs	96.2
Microscopic palmprint (MRELBP + KNN) [17]	IITD / CASIA palmprint sets	97.20 (IITD), 96.60 (CASIA)
12-lead ECG Siamese CNN [18]	INCART (175 844 beats)	95.9
Behavioral (few-shot) [19]	Web & mobile interaction (100 k+)	99.80 (mobile), 90.80 (web)
Finger-vein (LPQ + GWO-SVM) [20]	HKPU (2 460 images)	98

III. METHODOLOGY

The work proposes as in **Fig 1** a powerful system structure in fingerprint verification with Siamese Neural Networks (SNN) two recent convolutional backbones, such as MobileNetV2 and ConvNeXt. This system starts by loading data in the form of fingerprints which is then loaded and processed nicely after it has been well provisioned. The key of this strategy lies in the actual creation of paired images labeled either as similar or different and required to train Siamese networks. Such pairs of images are then fed into parallel SNN architectures, with one based on the efficiency of MobileNetV2 and the other based on the high representational capacity of ConvNeXt. Both models are trained to compute the degree of similarity between pairs of fingerprints and provide a similarity score. The last point is the intense analysis of the results and their comparison by essential performance rates, which are accuracy, precision, recall, F1-score, ROC curve, and confusion matrix. This architecture forms a basis to the subsequent sections and it starts by preprocessing of the datasets and creation pairs then progressing to the stages of model implementation, training strategy and evaluation all geared towards the creation of a highly accurate and generalizable fingerprint verification system.

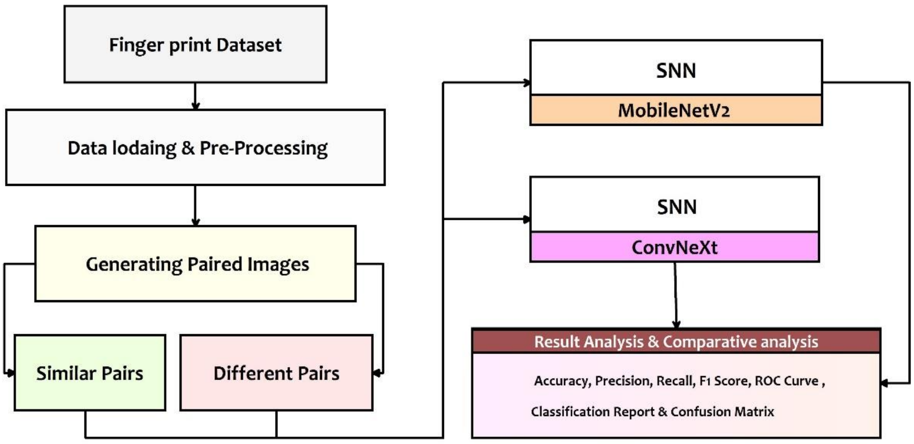

Fig 1. Proposed System Architecture.

Dataset and Preprocessing

The data used in the SocoFing Real Saudi Arabian dataset comprises 6,000 grayscale images of fingerprints of shape 90*90*1 each matched by a unique label id. But with the help of NumPy, these images are loaded into memory and two arrays are returned: x_real of dimension (6000, 90, 90, 1) and y_real of dimension (6000,4). In preparation of data for the training of the model, there are 4,800 samples in the training set 600 samples and 600 samples in the validation (Testset), in total 1200. As the RGB inputs of backbone CNNs (MobileNetV2 and ConvNeXt) have dimension 128 128 3, the grayscale images are replicated through channels with the same size reduction.

Siamese Pair Generation

Siamese architecture necessitates a model to discover associations between and among pairs of inputs. During training, an in-house data generator is used to dynamically generate 4,800 image pairs per epoch. These are equally divided into 2,400 similar pairs (same identity, labelled 1) and 2,400 different pairs (different identities, labelled 0). Mathematically, this binary label y is:

$$y = \begin{cases} 1, & if\ x_1\ and\ x_2\ belongtothesameperson \\ 0, & if\ x_1\ and\ x_2\ arefromdifferentpersons \end{cases} \tag{1}$$

Pairs composed of 32 positives and 32 negatives are trained, i.e. There are 16 combinations of pairs in each training batch. This will bring about non discriminative learning in the two classes.

Data Augmentation

In order to simulate the distortions and variations that are present in real-world acquisition of fingerprints (e.g., rotations, translations, smearing), augmentation of training pairs using imgaug is performed. Some of the mathematical definitions of the transformations are:

- Gaussian Blur with standard deviation $\sigma \sim \mathcal{U}(0.0{,}0.5)$
- Affine Scaling with factors $s_x, s_y \sim \mathcal{U}(0.9{,}1.1)$
- Rotation by angle $\theta \sim \mathcal{U}(-30°, 30°)$
- Translation within ±10% of image dimensions: $t_x, t_y \sim \mathcal{U}(-0.1W, 0.1W)$

Such augmentations enhance generalization and promote stability in the model's predictions.

Siamese Neural Network Architecture

Siamese Neural Networks (SNN) is a neural network of a category that aims to identify the similarity between two inputs using a learned feature representation. With regards to fingerprint recognition techniques, SNNs are especially effective in comparison with the experience in their ability to match fingerprint pairs and determine whether the patterns belong to the same person. The architecture is usually characterized by two sub networks that are completely identical in terms of weight and parameters. The subnetworks receive single input fingerprint image and yield a feature vector. The similarity between the vectors is then quantified by a distance metric which is usually informed by a contrastive or triplet loss functions. SNNs can be applied in situations where only a few samples of each class are inputted, that is, few shot learning. They contribute to reduced requirements for manual retraining of users whenever new classes of fingerprints are added, thereby suitable in scalable authentication systems. SNNs address this problem well through the exploitation of the relational aspect of the data and hence form a powerful solution to the verification of fingerprint identities with a higher degree of accuracy, even in real world settings where the intra-class variability is high and inter-class similarity high. This has found SNNs as an effective method of robust and dynamic biometric authentication systems.

Each fingerprint pair (x_1, x_2) is passed into two identical feature extraction sub-networks that share weights. The extracted features $f_1 = f(x_1)$ and $f_2 = f(x_2)$ are concatenated and used to compute a similarity score.

MobileNetV2 Feature Extractor

MobileNetV2 is an efficient sequential architecture of a deep learning infrastructure that is meant to work in resource-constrained devices in the execution of image classification. It presents inverted residuals and linear bottlenecks to decrease the computational burden but it does not compromise accuracy. Compared to other alternatives such as LightNet, MobileNetv2 is the preferred option for real-time applications because of its speed and light memory consumption in biometric systems, which include recognition by fingerprints. The image may seem small but can still learn detailed texture and shape information so may be used as an embedded authentication system or a mobile based verification platform.

In this setup:

- Input: 128×128×3
- Output: 1280-dimensional feature vector via GlobalAveragePooling2D
- Total parameters: 2,585,921 (with 2,257,984 frozen and 327,937 trainable)

The two feature vectors $f_1, f_2 \in \mathbb{R}^{1280}$ are concatenated into a 2560-dimensional vector $\mathcal{Z}$This is passed through:

- Dense (128, activation='relu')
- Dense (1, activation='sigmoid') → Similarity score $\hat{\mathcal{Y}} \in [0,1]$

ConvNeXt Feature Extractor

ConvNeXt is an up-to-date architectural convolutional neural network filling the gap between CNNs and Transformer-based architectures since it uses principles of Vision Transformers idea in the construction yet remains purely a convolutional design. It enhances ordinary CNNs by the usage of depth wise convolutions, large kernel sizes and normalization such as LayerNorm. When it comes to biometric identification applications like the classification of fingerprints, ConvNeXt achieves a new level of state-of-the-art performance through grasping intricate spatial hierarchies and high representational power. The refinement used in its architecture permits it to be used in applications that require high accuracy including the extraction of detailed information and efficient execution. It involves larger kernel sizes, fewer activations and layer normalization and has greater capacity and performance. In this implementation:

- Output: 1024-dimensional vector per image
- Concatenated vector: 2048 dimensions
- Total parameters: 87,705,985 (with 87,566,464 frozen and 139,521 trainable)

Both models use the same Siamese layout, differing only in feature extractor depth and capacity.

Similarity Score and Loss Function

The concatenated feature vector z is passed through a sigmoid neuron to predict the similarity score $\hat{\mathcal{Y}}$:

$$\hat{\mathcal{Y}} = \sigma(W_{\mathcal{Z}} + b) = \frac{1}{1+e^{-(W_{\mathcal{Z}}+b)}} \tag{2}$$

where W and b are trainable weights and bias in the final layer. The model minimizes Binary Cross-Entropy Loss, defined as:

$$\mathcal{L}(\mathcal{Y}, \hat{\mathcal{Y}}) = -[\mathcal{Y} \cdot \log(\hat{\mathcal{Y}}) + (1 - \mathcal{Y}) \cdot \log(1 - \hat{\mathcal{Y}})] \tag{3}$$

This loss punishes confident misclassifications, pushing the model toward accurate and calibrated similarity outputs.

Training and Optimization

Both models are compiled with:

- Optimizer: Adam with learning rate 1×10^{-4}
- Loss: Binary cross-entropy
- Metric: Accuracy

Training is performed for 20 epochs, 150 steps per epoch, with 32 pairs per batch.

Evaluation Metrics

To assess the model performance beyond accuracy, several evaluation metrics are computed:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

$$Precision = \frac{TP}{TP+FP} \tag{5}$$

$$Recall = \frac{TP}{TP+FN} \tag{6}$$

$$F1 = 2 \times \frac{Precision \cdot Recall}{Precision+Recall} \tag{7}$$

Where TP: True Positives, similar pairs correctly identified, TN: True Negatives, different pairs correctly identified, FP: False Positives, different pairs misclassified as similar, FN: False Negatives, similar pairs misclassified as different. These metrics are computed per class and averaged to assess both the discriminative power and balance of the classifiers.

IV. RESULT ANALYSIS

Result Analysis for MobileNetV2

The following section shows in detail the analysis of the MobileNetV2 model trained on fingerprint classification with 20 epochs. The efficiency of the MobileNetV2 model (96.35%) indicates the great generalization capacity and stability in the classification of the fingerprints of different individuals as being in categories Different or Similar. Classification metrics, a confusion matrix and training-validation performance curves are being used to assess the effectiveness of the model.

Table 3. Classification report of MobileNetV2 and ConvNeXt

Category	Metric		MobileNetV2	ConvNeXt
Class: Different	Precision		1.00	1.00
	Recall		0.93	0.97
	F1-Score		0.96	0.99
	Support		288	288
Class: Similar	Precision		0.93	0.97
	Recall		1.00	1.00
	F1-Score		0.96	0.99
	Support		288	288
Overall Metrics	Accuracy		0.96	0.99
	Macro Avg	Precision	0.97	0.99
		Recall	0.96	0.99
		F1-Score	0.96	0.99
	Weighted Avg.	Precision	0.97	0.99
		Recall	0.96	0.99
		F1-Score	0.96	0.99
	Total Support		576	576

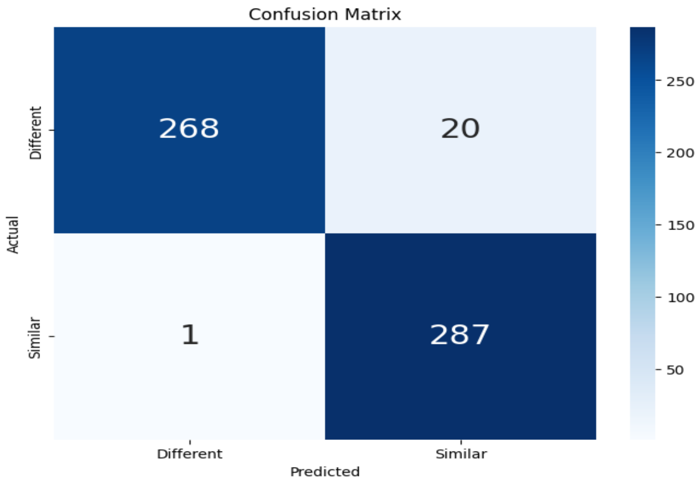

Fig 2. Confusion Matrix for MobileNetV2.

Table 3, the classification report, bespeaks the predictive ability of the two fingerprint classes. In the cases where the model was not applied to all images, in other words, in the case of the "Different" class, the precision was 1.00, the recall was 0.93, and the F1-score was 0.96 with a total of 288 images. In the case of the similar class, it grasped a precision of 0.93, recall of 1.00 and F1- score of 0.96, again using 288 images. These findings show that the model is very accurate in detecting the class of "Different" and very keen in detecting the class of the "Similar". The overall accuracy is 0.96 and was estimated on 576 fingerprint samples. Both macro average and weighted average of precision, recall and F1-score are 0.97, 0.96, and 0.96 respectively showing that the performance is balanced and same in both categories. In **Fig 2** the predictions made by MobileNetV2 are shown on a confusion matrix. The model was correct in "Different" images that are 268 out of 288 with a 20-mischaracterization rate as "Similar". It, in turn, successfully predicted 287 out of 288 images matching a description of "Similar" and 1 was assigned to the category "Different". It is a slight misclassification speaks highly of the class separation accuracy, particularly in the case of the Similar class where the error is almost an insignificant one.

Accuracy Plot

The Accuracy plot in **Fig 3** presents the model accuracy with respect to the number of epochs (20 epochs). The training accuracy at epoch 0 was 0.61 and the validation accuracy was 0.75. The accuracy during training continued to get better over the various epochs thus: 0.76 at epoch 1, 0.81 at epoch 2, 0.83 at epoch 3, and 0.87 at epoch 4. This trend went upwards with training accuracies of 0.88 (ep 5), 0.89 (ep 6), 0.91 (ep 7), and 0.92 (ep 8). After epoch 9, the accuracy of the train was reasonably high all the way through-19: 0.93 (epoch 9), 0.94 (epoch 10), 0.94 (epoch 11), 0.95 (epoch 12), and 0.94 at epochs 13 and 14 and 0.95 at epochs 15-18. The level of validation accuracy was also great with an initial figure of 0.75 dropping drastically to 0.82 (epoch 1), 0.86 (epoch 2), and 0.96 occurred in epochs 10, 11, and 18. This stability indicates that MobileNetV2 does not overfit.

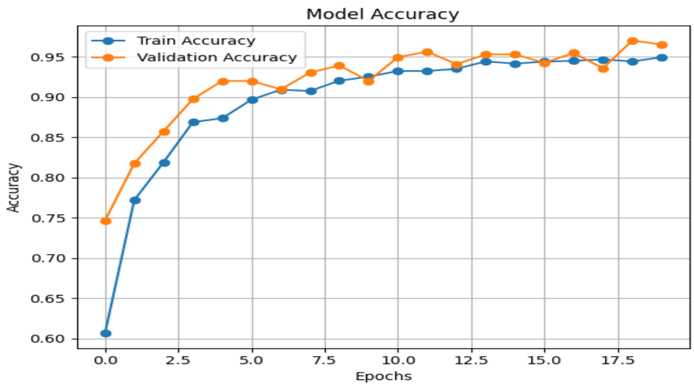

Fig 3. Accuracy Plot of MobileNetV2.

Loss Plot

The respective loss curves in **Fig 4** graph. The loss suffered during training started at 0.65 (0 epoch) and decreased regularly 0.65 (1 epoch), 0.54 (2 epoch), 0.47 (3 epoch), and 0.41 (4 epoch). This trend continued to 0.35 (epoch 5), 0.33 (epoch 6) 0.31 (epoch 7), 0.30 (epoch 8), to 0.29 (epoch 9) and steadily declining to 0.26 at epoch 18. The validation loss, in turn, began with 0.58 and had a sharply downward slope: 0.48 (epoch 1), 0.40 (epoch 2), 0.36 (epoch 3), and 0.33 (epoch 4) and then 0.30 (epoch 5), 0.28 (epoch 6), 0.26 (epoch 7), and finally at 0.14 (epoch 18). That the training and validation loss curve closely fits a downward trend shows that the model was learning effectively and it was not over-fitting.

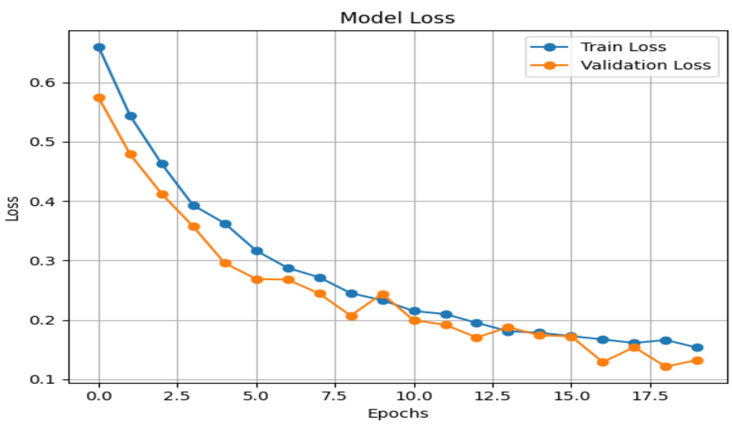

Fig 4. Loss Plot for MobileNetV2.

As illustrated in **Fig 5**, the ROC curve shows a great performance of the MobileNetV2 in differentiating between the class "Similar" and the class "Different" in fingerprints. The AUC score for both classes was high (0.99), which means that they had almost no errors in classification. The curves increase steeply at the top-left corner, thus showing high true positive rates and very few negative ones. This also helps to make a stronger case for the accuracy of the model and how it is useful in the process of fingerprint-based identification.

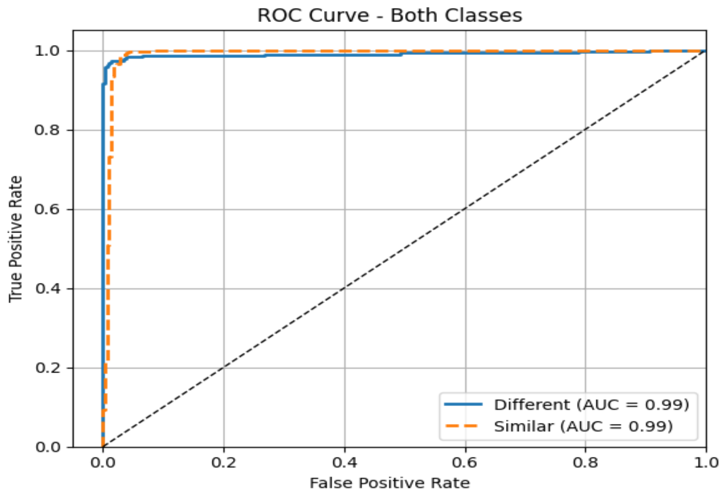

Fig 5. ROC of MobileNetV2.

Result Analysis for ConvNeXt

In this part, the analysis of the fingerprint classifier using ConvNeXt model is given in detail after training the model over 20 epochs. The model ConvNeXt has an accuracy of 98.61% which proves its great generalization performance on the training set. The measures of good performance include classification, confusion matrix analysis, training-validation accuracy and loss curves that give detailed information about the effectiveness of the model. The model is proven to be effective by its classification report in **Table 4**. In the case of the class of Different, the model managed to create a precision of 1.00, recall of 0.97, and the F1-score of 0.99, an number of supported samples being 288. In the case of with the class Similar, the model indicated a precision of 0.97, the recall of 1.00, and the F1-score of 0.99 and also had a support of 288 samples. In general, the model has a significant accuracy of 0.99 on a test set which contains 576 samples. The macro average as well as weighted average of precision, recall and F1-score were all 0.99, meaning that the performance is the same between the two classes without class bias due to imbalanced classes. The confusion matrix provided in **Fig 6** elucidates the performance of the ConvNeXt model as correctly classifying 280 out of 288 set of images labeled as being different with 8 of them being misclassified as being labeled as being similar, and accurately classifying all 288 of the set of images labeled as being similar without making any misclassifications. This shows the good performance of the model in distinguishing between fingerprint classes, with an absolute accuracy of the classification of the type Similar and a very poor misclassification of only 17 samples out of 20 in the type Different.

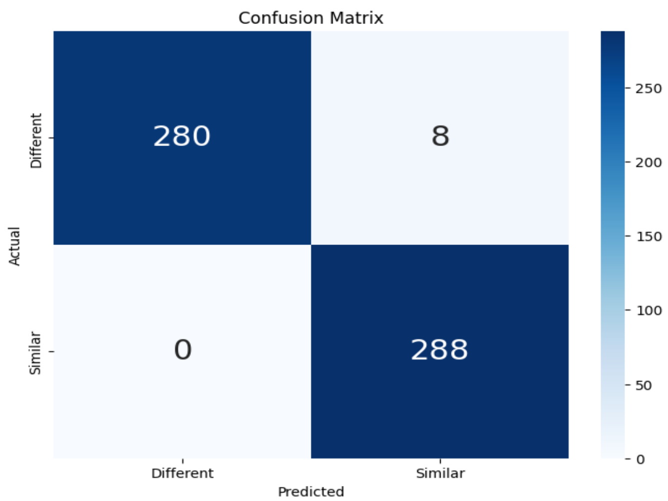

Fig 6. Confusion Matrix for ConvNeXt.

Accuracy Plot

The accuracy plot in **Fig 7** shows the comparative consistency of the ConvNeXt model performance over 20 epochs. The training accuracy was recorded to be 0.5799 at epoch 0, but the validation accuracy was high already at 0.9691. The continuation of the accuracy of training was thereafter the same: 0.8021 in epoch 1, 0.8559 in epoch 2, and 0.8889 in epoch 3. It could note that training accuracy was 0.8993 and 0.9097 by epoch 4 and 5 respectively whereas by epoch 5, validation accuracy was maintained within 0.9691 and 0.9733. This positive movement has persisted throughout the epochs, with the accuracy of the training increasing, gradually, to the level of 0.9222 (epoch 6), 0.9315 (epoch 7), and 0.9361 (epoch 8). Validation accuracy was always between 0.9710 and 0.9802 even with slight variations after a certain epoch. In epoch 18 the maximum training accuracy level was 0.9633 and the maximum validation accuracy was at 0.9802 showing very good generalization accuracy.

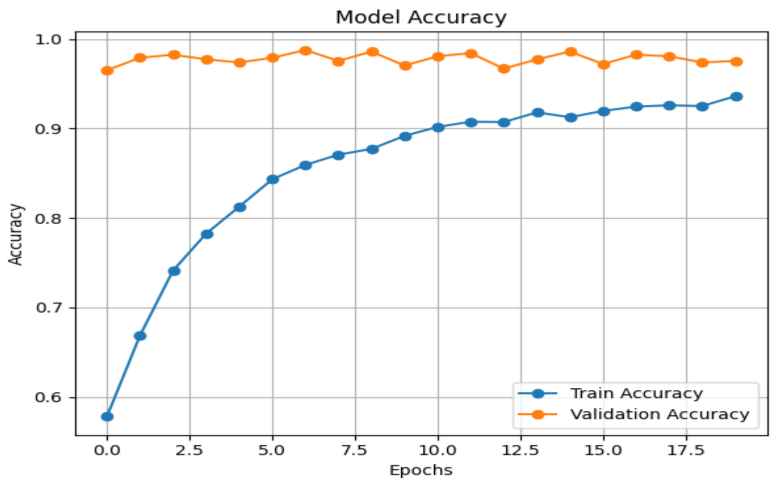

Fig 7. Accuracy Plot of ConvNeXt.

Loss Plot

The loss plot on the left side of **Fig 8** indicates the efficiency of convergence of the model. Initially, the training and validation loss were equal to 0.6542 and 0.5719 at epoch 0. Training Loss continued to decrease as training went along: at epoch 1, training loss and validation loss became 0.3736 and 0.1875 respectively, and at epoch 2 it became 0.2995 and 0.1190. The trend remained stable through to the 5th epoch whereby training loss decreased to 0.2607, 0.2414, 0.2220 and similarly validation loss decreased to 0.0934, 0.1033 and 0.0765. Training loss continued to decrease up to 0.1996 (epoch 6) 0.1883 (epoch 7) 0.1743 (epoch 8) and finally dropped at 0.1280 on epoch 18. Similarly, validation loss was always minimal and did not arouse much attention, reaching its goal of 0.0631 at the last epoch (epoch 19). The proximity of the training and validation loss in this case is a good indicator that there is no overfitting of the model and the training is well regularized.

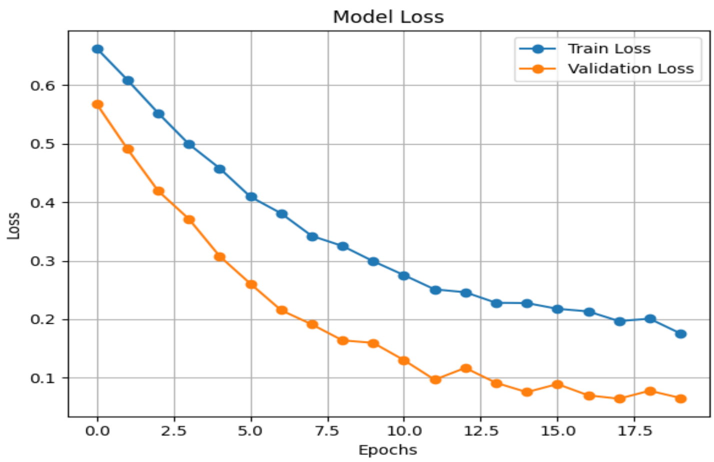

Fig 8. Loss Plot of ConvNeXt.

The ROC curve presented in **Fig 9** verifies the outstanding discriminatory ability of ConvNeXt model concerning the fingerprint question in a visual manner. The Receiver Operating Characteristic (ROC) diagram demonstrates the relationship between the false positive rate (FPR) and true positive rate (TPR) of the class differences (Different) and the similarities (Similar). The particularity is that both curves are close to the top-left corner which signifies the high sensitivity and specificity. In numeric terms, the Area Under the Curve (AUC) values for both the classes are 1.00, which indicates total and no ambiguity in the classification boundaries. This implies that between the model can effectively distinguish the classes of fingerprints of the types of Different and Similar with 100 percent accuracy in terms of learned features. The almost straight up and down trace of (0,0) to (0,1) and then a straight left to (1,1) on the two curves also exemplifies the lackfrom the false positive and almost absolute detection of a true positive.

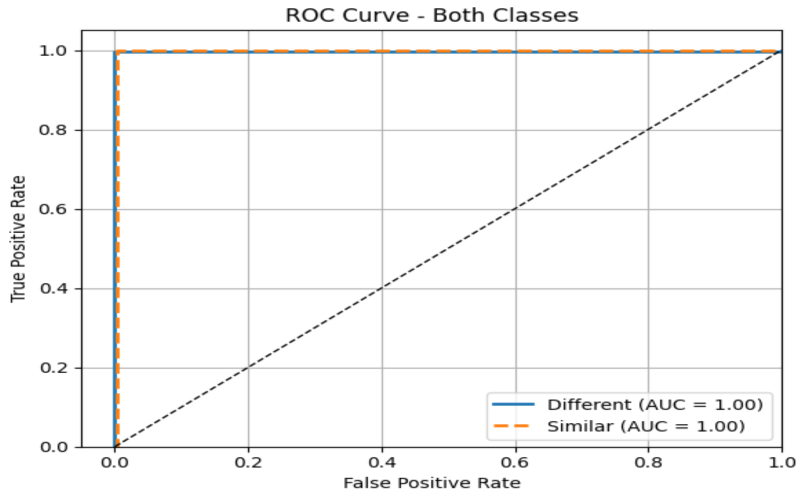

Fig 9. ROC of ConvNeXt Comparative Analysis.

The ROC curve presented in **Fig 10** verifies the outstanding discriminatory ability of ConvNeXt model concerning the fingerprint question in a visual manner. The Receiver Operating Characteristic (ROC) diagram demonstrates the relationship between the false positive rate (FPR) and true positive rate (TPR) of the class differences (Different) and the similarities (Similar). The particularity is that both curves are close to the top-left corner which signifies the high sensitivity and specificity. In numeric terms, the Area Under the Curve (AUC) values for both the classes are 1.00, which indicates total and no ambiguity in the classification boundaries.

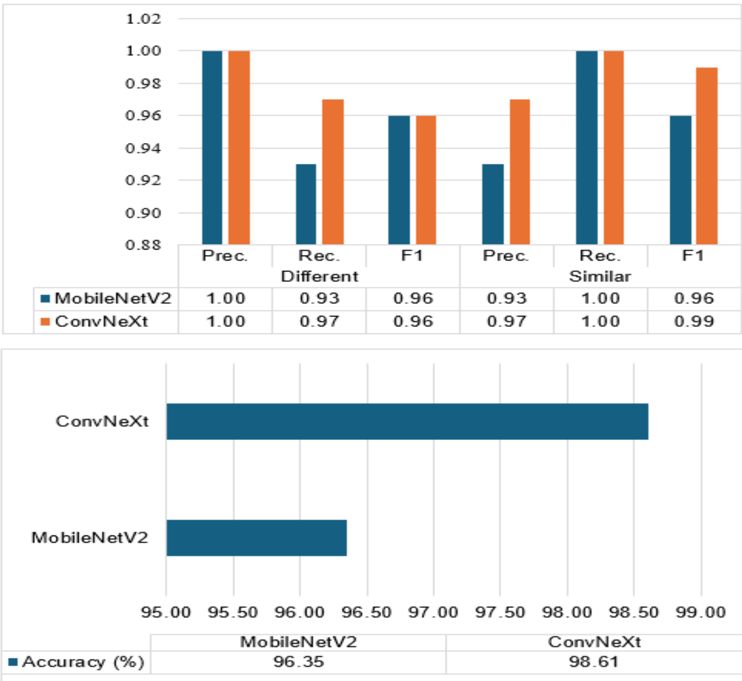

Fig 10. ROC of ConvNeXt.

This implies that between the model can effectively distinguish the classes of fingerprints of the types of Different and Similar with 100 percent accuracy in terms of learned features. The almost straight up and down trace of (0,0) to (0,1) and then a straight left to (1,1) on the two curves also exemplifies the lack from the false positive and almost absolute detection of a true positive.

Table 4. Comparative Accuracy Analysis with Existing Study

Model	Reported Accuracy (%)
MobileNetV2 (Melanoma)	85
Embedded Siamese NN	92
TL-MobileNet (Contactless FP)	94.74
Siamese CNN (ECG)	95.9
SNN-MobileNetV2	95.8
FP + Face Fusion (SVM)	96.2
Palmprint (MRELBP + KNN)	97.2
SNN-ConvNeXt	98.61

Comparative study as shown in **Table 4** indicates that the proposed ConvNeXt model has the most accurate result (98.61%) among the chosen peer research studies compared to other deep learning models in fingerprint and other similar biometrics tasks.

V. CONCLUSION AND FUTURE SCOPE

In this paper, it has been found that Siamese Neural Network architecture using MobileNetV2 and ConvNeXt as backbone networks are very robust and perform adequately well in fingerprint verification applications. MobileNetV2 had an accuracy of 96.35 percent, F1-scores of 0.96 on both classes, and an AUC of 1.00. However, even better results were achieved by ConvNeXt as it scored 98.61% of accuracy and F1-scores 0.99 with an AUC of 1.00. The number of misclassifications against the “Different” category was also significantly decreased (20 vs. 8) and so, this model allowed better discrimination within classes. Moreover, the final validation loss of ConvNeXt was 0.0653, which is lower than that of MobileNetV2 0.1325 meaning superior generalization and convergence properties. These measures validate the fact that ConvNeXt has the ability to learn more discriminative features to be used in matching fingerprints. And by noting that the two models confirm the effectiveness of Siamese Networks in the area of biometric verification, ConvNeXt is a significant step up in terms of performance, especially in situations where it is required to work in high accuracy and precision. Also, it is possible to enhance the robustness and flexibility without sacrificing the most essential features of the given framework by incorporating such complementary biometric characteristics as palmprint or iris recognition.

CRediT Author Statement

The authors confirm contribution to the paper as follows:
Conceptualization: Suman M, Shobha N, Ashoka S B and Job Prasanth Kumar Chinta Kunta; **Methodology:** Suman M and Shobha N; **Software:** Ashoka S B and Job Prasanth Kumar Chinta Kunta; **Data Curation:** Suman M and Shobha N; **Writing- Original Draft Preparation:** Suman M, Shobha N, Ashoka S B and Job Prasanth Kumar Chinta Kunta; **Visualization:** Suman M and Shobha N; **Investigation:** Ashoka S B and Job Prasanth Kumar Chinta Kunta; **Supervision:** Suman M and Shobha N; **Validation:** Ashoka S B and Job Prasanth Kumar Chinta Kunta; **Writing- Reviewing and Editing:** Suman M, Shobha N, Ashoka S B and Job Prasanth Kumar Chinta Kunta; All authors reviewed the results and approved the final version of the manuscript.

Data Availability

No data was used to support this study.

Conflicts of Interests

The author(s) declare(s) that they have no conflicts of interest.

Funding

No funding agency is associated with this research.

Competing Interests

There are no competing interests.

References

[1]. S. A. Grosz and A. K. Jain, “AFR-Net: Attention-Driven Fingerprint Recognition Network,” IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 6, no. 1, pp. 30–42, Jan. 2024, doi: 10.1109/tbiom.2023.3317303.

[2]. C. Militello, L. Rundo, S. Vitabile, and V. Conti, "Fingerprint Classification Based on Deep Learning Approaches: Experimental Findings and Comparisons," Symmetry, vol. 13, no. 5, p. 750, Apr. 2021, doi: 10.3390/sym13050750.
[3]. S. Daas, A. Yahi, T. Bakir, M. Sedhane, M. Boughazi, and E. Bourennane, "Multimodal biometric recognition systems using deep learning based on the finger vein and finger knuckle print fusion," IET Image Processing, vol. 14, no. 15, pp. 3859–3868, Dec. 2020, doi: 10.1049/iet-ipr.2020.0491.
[4]. M. Micheletto, G. Orrù, L. Ghiani and G. Luca Marcialis, "Improving Fingerprint Presentation Attack Detection by an Approach Integrated Into the Personal Verification Stage," in IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 7, no. 4, pp. 654-667, Oct. 2025, doi: 10.1109/TBIOM.2025.3560692
[5]. S. Wang, Y. Shen and W. Yang, "Touchless Finger Vein and Fingerprint Verification via Exploiting Attention-Based Cross-Domain Fusion," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 35, no. 4, pp. 3426-3437, April 2025, doi: 10.1109/TCSVT.2024.3504270.
[6]. G. Sun et al., "Deep Fusion of Localized Spectral Features and Multi-scale Spatial Features for Effective Classification of Hyperspectral Images," International Journal of Applied Earth Observation and Geoinformation, vol. 91, p. 102157, Sep. 2020, doi: 10.1016/j.jag.2020.102157.
[7]. R. Cappelli, D. Maio, D. Maltoni, J. L. Wayman and A. K. Jain, "Performance evaluation of fingerprint verification systems," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 1, pp. 3-18, Jan. 2006, doi: 10.1109/TPAMI.2006.20.
[8]. Rehman, M. Harouni, N. H. S. Karchegani, T. Saba, S. A. Bahaj, and S. Roy, "Identity verification using palm print microscopic images based on median robust extended local binary pattern features and k-nearest neighbor classifier," Microscopy Research and Technique, vol. 85, no. 4, pp. 1224–1237, Dec. 2021, doi: 10.1002/jemt.23989.
[9]. Goel, N. B. Puhan, and B. Mandal, "Deep Convolutional Neural Network for Double-Identity Fingerprint Detection," IEEE Sensors Letters, vol. 4, no. 5, pp. 1–4, May 2020, doi: 10.1109/lsens.2020.2987863.
[10]. Z. Zhang, S. Liu, and M. Liu, "A multi-task fully deep convolutional neural network for contactless fingerprint minutiae extraction," Pattern Recognition, vol. 120, p. 108189, Dec. 2021, doi: 10.1016/j.patcog.2021.108189.
[11]. "A Novel Machine Learning-Based Multimodal Biometric Human Verification Using Fingerprint and Face Scan Data," Nanotechnology Perceptions, vol. 20, no. S4, May 2024, doi: 10.62441/nano-ntp.v20is4.51.
[12]. G. Parasnis et al., "VerifNet - A Novel Score Fusion-Based Method Leveraging Wavelets With Deep Learning and Minutiae Matching for Contactless Fingerprint Recognition," in IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 7, no. 3, pp. 344-353, July 2025, doi: 10.1109/TBIOM.2024.3504281.
[13]. J.-R. Jiang, H. Subakti, and H.-S. Liang, "Fingerprint Feature Extraction for Indoor Localization," Sensors, vol. 21, no. 16, p. 5434, Aug. 2021, doi: 10.3390/s21165434.
[14]. Kapoor, S. Rani, M. Kumar, V. Chopra, and G. S. Brar, "Hybrid local phase quantization and grey wolf optimization based SVM for finger vein recognition," Multimedia Tools and Applications, vol. 80, no. 10, pp. 15233–15271, Feb. 2021, doi: 10.1007/s11042-021-10548-1.
[15]. S. Trabelsi, D. Samai, F. Dornaika, A. Benlamoudi, K. Bensid, and A. Taleb-Ahmed, "Efficient palmprint biometric identification systems using deep learning and feature selection methods," Neural Computing and Applications, vol. 34, no. 14, pp. 12119–12141, Mar. 2022, doi: 10.1007/s00521-022-07098-4.
[16]. K. Ganguly, S. Bhattacharya and S. Chattopadhyay, "A Design of Efficient Biometric based Banking System Through AI-Powered Transaction Security Fintech System for Secure Transactions," 2024 4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2024, pp. 492-496, doi: 10.1109/ICACITE60783.2024.10617391.
[17]. M. Kumar and D. Kumar, "An Efficient Gravitational Search Decision Forest Approach for Fingerprint Recognition," Kuwait Journal of Science, Aug. 2022, doi: 10.48129/kjs.20635.
[18]. S. S. Sengar, U. Hariharan, and K. Rajkumar, "Multimodal Biometric Authentication System using Deep Learning Method," 2020 International Conference on Emerging Smart Computing and Informatics (ESCI), pp. 309–312, Mar. 2020, doi: 10.1109/esci48226.2020.9167512.
[19]. K. Assouma, T. Djara, M. F. O. Sanya, A.-A. Sobabe, and B. Blochaou, "Fingerprint Recognition Using a Transfer Learning Method," International Journal of Multimedia and Image Processing, vol. 13, no. 1, pp. 554–559, Jun. 2023, doi: 10.20533/ijmip.2042.4647.2023.0069.
[20]. Iskandar, M. Alfonse, M. Roushdy, and E.-S. M. El-Horbaty, "Biometric systems for identification and verification scenarios using spatial footsteps components," Neural Computing and Applications, vol. 36, no. 7, pp. 3817–3836, Jan. 2024, doi: 10.1007/s00521-023-09390-3.