Received: 21 August 2025 | Revised: 26 January 2026 | Accepted: 22 June 2026 | Published online: 10 July 2026

RESEARCH ARTICLE

Artificial Intelligence and Applications
2026, Vol. 00(00) 1–16
DOI: 10.47852/bonviewAIA62027314

A Multi-Instant Machine Learning Algorithm Model for Infant Fingerprinting Verification

Olatayo Moses Olaniyan[1], Kazeem Aderemi Bello[2,*], Rendani Wilson Maladzhi[2], Johnson Olugbenga Joseph[3], Adebimpe Esan[1], Quadri Ademola Mumuni[1,3], Bolaji Abigael Omodunbi[1] and Chinedu Michael Chiejine[4]

[1] *Computer Engineering Department, Federal University of Oye-Ekiti, Nigeria*

[2] *Mechanical Engineering Department, Durban University of Technology, South Africa*

[3] *Electronic and Computer Engineering Department, Lagos State University, Nigeria*

[4] *Department of Electrical and Electronics Engineering, Admiralty University of Nigeria, Nigeria*

Abstract: Understanding the challenges associated with infant biometrics, like immature ridge formation, very small fingerprints, and their fast physiological development, this research implements a multi-instance machine learning technique to enhance accuracy. Multiple instances of fingerprints for each infant were recorded, which enhanced the robustness of the system by dealing with inter-class variability. The datasets used 500 pixels-per-inch fingerprints of infants less than 12 months old that were recorded during two different instances. After preprocessing techniques, feature extraction was performed using a convolutional neural network (CNN), and matching scores were computed. The resulting scores were analyzed using a combination of three learning algorithms, CNN, CNN + SVM, and CNN + RF. Metrics showed that CNN + RF model outperformed all others with regard to its performance, having achieved verification accuracy of 81% compared to 75% and 82%, respectively (the previous multi-instance learning technique had achieved highest accuracy of only 69.05%), training accuracy of 98% compared to 58% and 50%, respectively, validation accuracy of 85% compared to 84% and 83%, respectively, and minimum training loss of 0.2 while having considerably high validation loss of 2.5. Although higher computational cost, CNN + RF proved most active for real-world applications in healthcare and identity management. Upcoming research should explore larger datasets to further enhance model precision.

Keywords: machine learning, digital identity, infant fingerprint, confusion matrix, validation

1. Introduction

It is estimated that there are about 140 million births annually, yet only 237 million of these babies have a birth certificate, while about 25% of children under five years old, some 166 million of them, do not have birth certificates at all [1]. The widespread lack of formal documentation for children is quite evident in developing countries because the lack of proper infrastructure and policy issues affects their ability to obtain any official identification [2]. Identification cards are essential to enable access to social programs such as healthcare, education, and other social protections, in addition to providing legal identity [3]. Lack of identification can leave a child vulnerable to exclusion from social programs, treatment, and immunizations, especially in the crucial period of 1000 days [4, 5].

The use of dependable biometric systems becomes very important in the case of regions where the rates of registration at birth are very low, for example, in Nigeria. For minimizing errors, avoiding duplication, and ensuring accuracy in the process of registering, it is imperative that more advanced biometric systems be employed [6]. Besides this use, there is also a potential of using biometric technology, particularly fingerprint recognition, for reuniting children who were separated from their parents, validating parent–child relationships, and preventing child trafficking [7, 8].

One of the solutions in this regard could be implementing advanced machine learning approaches, namely, multi-instance learning (MIL) and contingent learning. As stated in [9], through MIL, the computer would analyze multiple fingerprints (instances) of an individual and make predictions using this information. This becomes especially useful in the case of biometrics of newborns, where fingerprints obtained during scanning could be of varying quality.

Additionally, through contingent learning, a model can account for the influence of factors such as growing up, environmental impact, and variations in fingerprints [10, 11]. It is, therefore, possible to develop advanced fingerprint recognition systems using the two techniques [12].

***Corresponding author:** Kazeem Aderemi Bello, Mechanical Engineering Department, Durban University of Technology, South Africa. Email: kazeemb@dut.ac.za

The feature extraction via convolutional neural network (CNN), classification via hybrid support vector machines (SVMs) or random forest (RF), and match score fusion are some of the major aspects of the intelligent algorithm [13]. There have been other related studies done on Mask-Refined Region-Based Convolutional Neural Network (R-CNN) and feature pyramid networks [14, 15]. Furthermore, the resource-constrained environment in the healthcare domain requires that efficient and lightweight algorithms be developed to address issues such as the scarcity of power supply and connectivity.

2. Literature Review

One way of improving the identification process in RF fingerprinting is the introduction of a dual-channel parallel CNN. The technique was introduced by [16], where the process uses two parallel channels in the CNN to exploit the benefits of both spectral and spatial information of RF fingerprints to improve the accuracy of identification of fingerprints for each RF device. The tests conducted using the method showed an increase in the accuracy rate of RF fingerprint identification compared to traditional CNN techniques. Nevertheless, one of the shortcomings in the technique is that it has not been evaluated for practical purposes, such as in real time and how it performs in different and heterogeneous RF environments.

A hybrid machine learning algorithm for RF fingerprinting, which uses Deep Belief Networks (DBNs) for extracting features and SVM for classification [17]. In the proposed algorithm, deep and hierarchical features are extracted by utilizing DBNs from raw RF signal data. These features can be classified through an SVM classifier for recognizing different types of wireless devices depending on their RF fingerprints. In experiments, the proposed DBN-SVM classifier outperforms conventional shallow learning algorithms in both accuracy and robustness due to the existence of complex interference. Nevertheless, there remains a problem in relation to the feasibility and scalability of this algorithm, since it may not apply efficiently to a large wireless network, particularly under dynamic RF channels. More investigations need to be conducted for optimizing the computational speed of the algorithm and testing with other real-world RF signal datasets [18, 19].

Proposed a CNN method for an accurate infant fingerprint recognition to aid cybercrime detection processes [20]. This approach includes preprocessing fingerprint images to increase the clarity of ridges and then using a CNN model to automatically extract spatial features and classify individual fingerprints. This model was experimented on benchmark fingerprint databases, and it produced encouraging results regarding the accuracy and efficiency of recognition. As such, the proposed system can be implemented successfully in the area of forensics and cybersecurity. The model exhibited high classification accuracy and also had high tolerance to partial prints and image noise. The paper identified a number of gaps in the research that can limit its implementation in practice. These included issues related to generalization to cybercrime situations, especially those involving poor or altered quality images. Furthermore, the scalability and cross-dataset performance of the model were not well investigated in the paper. It was suggested that such gaps can be addressed through using various data augmentation strategies and combining the model into a larger multimodal biometric system [21].

This study hence suggests that there is a need for the development of a contingent machine learning algorithm for infant fingerprint identification. The technology not only makes an important contribution toward biometric security research but will also support global initiatives for improving child welfare, healthcare, and identification management systems [1]. The technology has the capability to help decrease child mortality rates, increase vaccination programs, and ensure protection against identity theft and trafficking of children, particularly in regions like Nigeria, where such facilities are lacking.

3. Proposed Methodology

3.1. Approvals and materials

This section discusses the procedure followed to acquire the fingerprint images, including but not limited to ethics approval and acquisition protocol.

3.1.1. Ethics approval

To obtain the necessary ethical and social approvals for this study in Lagos State, a structured process was followed in line with official protocols. First, we wrote a formal letter of request for social approval, addressed to the Permanent Secretary, Ministry of Health, Alausa Secretariat, Ikeja. As the Health Research Unit received the letter, we were contacted and guided to proceed with the next steps. Then, a soft copy of our detailed research proposal, along with other relevant documents, was submitted via WhatsApp to the designated contact. Additionally, we were to secure a letter of introduction from our affiliated institution (FEDERAL UNIVERSITY OYE-EKITI, FUOYE) to support our application.

Subsequently, we obtained ethical approval from a recognized health research ethics committee in Lagos. This included institutions such as the Lagos State Health Research Ethics Committee, Lagos University Teaching Hospital, Lagos State University Teaching Hospital, or the Nigerian Institute of Medical Research. Alternatively, such approval could also be sought from the National Health Research Ethics Committee in Abuja.

Following this, we submitted a letter of request for administrative/social approval, attaching all the aforementioned documents to the Permanent Secretary, Public Service Office, under the Office of the Head of Service, Lagos State Secretariat, Alausa. This office, in turn, issued a letter of introduction to the Ministry of Health for further review and final approval.

Finally, we completed the process by making the required payment of ₦5,000 into the Lagos State Government account, as instructed upon reaching that stage of the approval process. The approval is available on request.

3.1.2. Study site

The fingerprints are being acquired mainly in the immunization clinic of the Ita-Elewa Health Centre, Ikorodu, Lagos State, Nigeria.

3.1.3. Study population

Neonates and infants who were attending the immunization clinic of the Ita-Elewa Health Centre, Ikorodu, Lagos State, Nigeria, were recruited for the study.

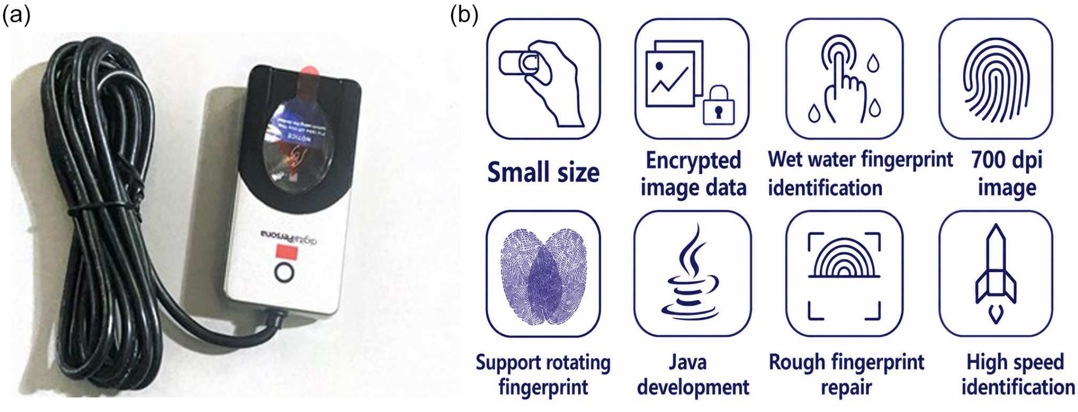
3.1.4. Subject inclusion and exclusion criteria

Fingerprint samples were acquired from infants under 12 months of age, and infants who were 12 months and older or unwell were excluded from the study.

3.1.5. Fingerprint reader acquisition

Sadly, most of the hospitals lack a working fingerprint scanner for conducting the process of collecting the required data, and thus, there was no alternative but to procure two new

Figure 1
(a) Purchase fingerprint reader. (b) Snapped fingerprint reader after removal from the container



fingerprint scanners from overseas, which were then made available to the healthcare professionals for further handling. Figure 1(a) represents the purchased fingerprint scanner. Figure 1(b) represents the cracked fingerprint scanner after it has been removed from the packaging box. Figure 2 represents the features of the Digital Persona fingerprint recognition software. Figures 3(a) and 3(b) represent some of the various images collected during the fingerprint data acquisition process.

3.2. Proposed algorithm

Here, the algorithm explaining the step-by-step approach of the whole methodology is presented and iterated. To begin with, two fingerprint readers are purchased for capturing the infants' fingerprint data at different hospitals after obtaining social approval to proceed with a study in Lagos State, through the Health Research Unit.

The proposed algorithm is itemized below:

3.2.1. Data acquisition

This entails capturing about 500 pixels-per-inch (PPI) fingerprint images from infants in multiple sessions so as to create a robust and reliable dataset for the research.

3.2.2. Preprocessing

In terms of actual processing of the captured fingerprints, each fingerprint image is normalized by dividing it by 255. The pixel intensities of the digital fingerprint image usually have a range from 0 to 255 for an 8-bit grayscale image. Each pixel will be divided by 255, making the range of pixel values now [0, 1]. The normalized value of the image X will become Equation 1:

$$X' = \frac{X}{255} \tag{1}$$

3.2.3. Feature extraction

To extract the salient features of the Infant-Prints, convolutional layers are applied to obtain the feature maps according to Equation 2:

$$F_{i,j,k} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} X_{i+m,j+n} \times W_{m,n,k} + b_k \tag{2}$$

where $F_{i,j,k}$ is the output feature map at position (i, j) for the kth filter (channel) result of convolution and $X_{i+m,j+n}$ is the input image or feature map value at position $(i + m, j + n)$. This is the receptive field of the input being scanned over by the filter. $W_{m,n,k}$ is the weight of the convolutional filter (kernel) at position (m, n) for the kth output channel, b_k is the bias term for the kth filter and it's added once after the weighted sum, $M \times N$ is the size of the convolutional filter (i.e., kernel height × width), i, j are the indices for scanning over the height and width of the output feature map, and m, n are the indices for scanning over the height and width of the filter (kernel). Thereafter, there is the application of the Rectified Linear Unit (ReLU) activation function to the output of a convolution operation according to Equation 3:

$$F'_{i,j,k} = \max\ (0,\ F_{i,j,k}) \tag{3}$$

where $F'_{i,j,k}$ is the results after applying the ReLU function to $F_{i,j,k}$. It rectifies the output by removing negative values, which helps introduce non-linearity into the model and speeds up convergence during training, and ReLU is computed as:

$$\mathrm{ReLU}(x) = \begin{cases} x \text{ if } x > 0 \\ 0 \text{ if } x \le 0 \end{cases} \tag{4}$$

The last operation in step 3 is to carry out the max pooling on the activated ReLU using the computation in Equation 5:

$$P_{i,j,k} = \max_{(m,n)\in W} F'_{i+m,j+n,k} \tag{5}$$

Figure 2
Functionalities and specifications of the reader

Figure 3
Fingerprint acquisition on the data form

where $F'_{i+m,j+n,k}$ is the value from the activated feature map (after applying ReLU) at position $((i+m, j+n)$ in the kth channel, W is the pooling window or receptive field (e.g., 2×2 or 3×3 region) that slides over the feature map, $(m, n) \in W(m, n)$ indicates all positions (m, n) within the window W, max is the maximum value within the pooling window, and $P_{i,j,k}$ is the resulting value in the pooled feature map at position (i, j) in channel k.

3.2.4. Classification

Here, the dataset is ready for usage to train CNN, SVM, and RF classifiers. In addition, the computation of the SOFTMAX probabilities is carried out using Equation 6:

$$P\ (c_i) = \frac{e^{Y_i}}{\sum_j e^{Y_j}} \tag{6}$$

where $P\ (c_i)$ is the predicted probability of class c_i, Yi is the logit (raw score or output) from the model for class iii, and $\sum_j e^{Y_j}$ is the sum of exponentials of all logits over all possible classes.

3.2.5. Multi-instance fusion

Fusing of match scores using a weighted sum approach is implemented using Equation 7:

$$S_f = \alpha S_{CN}^{N} + \beta S_{SV}^{M} + \gamma S_{RF} \tag{7}$$

where S_f is the final fused score or output (e.g., classification score, confidence score, decision metric), S_{CN} is the score or feature from component CN (could be a model, method, or data source). SS_{SV} is the score or feature from SV (another source or method); SR_{RF} is the score or feature from RF (possibly random forest, or another method/data); α, β, and γ are weighting coefficients for each component, and they determine the importance of each source in the final score; and N and M are the exponents, often used to amplify or normalize the influence of certain components.

3.2.6. Performance evaluation

For the performance evaluation of the proposed model, three metrics are considered in section 4, which are the verification accuracy, equal error rate (EER), and AUC-ROC (area under the curve-receiver operating characteristic) for model validation.

To compute the verification accuracy, Equation 8 is used:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

where TP (true positive) is the model correctly predicted positive cases, TN (true negative) is the model correctly predicted negative cases, FP (false positive) is the model incorrectly predicted positive for a negative case (Type I error), and FN (false negative) is the model incorrectly predicted negative for a positive case (Type II error). The EER is the point at which FAR = FRR, or as close as possible. It's the threshold τ where the false acceptances and false rejections are balanced. Lower EER values indicate better performance; that is, the system makes fewer mistakes in both rejecting genuine users and accepting impostors. For the computation of the EER, Equation 9, written below, is used:

$$EER = \min_{\tau} |FAR(\tau) - FRR(\tau)| \tag{9}$$

where EER is the equal error rate, τ is the threshold used for accepting or rejecting a match, $FAR(\tau)$ is the false acceptance rate at threshold τ, and $FRR(\tau)$ is the false rejection rate at threshold τ. Third, the formula for AUC-ROC isn't a simple closed-form expression like many statistical metrics, but it's typically computed numerically from the ROC curve, which plots the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings.

But true positive rate (TPR or recall:

$$TPR = \frac{TP}{TP + FN} \tag{10}$$

False positive rate (FPR):

$$FPR = \frac{FP}{FP + TN} \tag{11}$$

For a set of ROC points (FPR1, TPR1), (FPR2, TPR2), . . . , (FPRn, TPRn), sorted by FPR, one can approximate AUC using the trapezoidal rule as written in Equation 12:

$$AUC = \sum_{i=1}^{n-1} (FPR_{i+1} - FPR_i) \cdot \frac{TPR_{i+1} + TPR_i}{2} \tag{12}$$

Interpretation: AUC = 1.0 → perfect classifier, AUC = 0.5 → no better than random guessing, AUC < 0.5 → worse than random (likely inverted predictions).

3.2.7. Feedback and tuning

Optimize hyperparameters using grid search. Refine feature selection to improve performance. Output: Verification result Y.

3.3. Flowchart of the proposed method

Figure 4 depicts the flowchart diagram of the detailed algorithms presented in the section above.

3.4. CNN-based mathematical model

3.4.1. Convolutional layer operations

Each convolutional layer applies a filter W over the input image X to produce feature maps F of Equation 13 written below:

$$F_{i,j,k} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{i+m,j+n} \times W_{m,n,k} + b_k \tag{13}$$

where $W_{m,n,k}$ represents the kernel weights, b_k is the bias term, (i, j) represents the spatial location, and k represents the filter index.

$$\text{The activation function: } ReLU\ (x) = \max(0,\ x) \tag{14}$$

is applied as;

$$F'_{i,j,k} = ReLU\ (F_{i,j,k}) \tag{15}$$

3.4.2. Pooling operation

Max pooling reduces spatial dimensions by selecting the maximum value in each window using Equation 16:

$$P_{i,j,k} = \max_{(m,n)\in W} F'_{i+m,j+n,k} \tag{16}$$

3.4.3. Fully connected layer and SOFTMAX activation

The output from the final convolutional layer is flattened and connected to dense layers according to Equation 17, which is a linear transformation commonly used in neural networks and machine learning models:

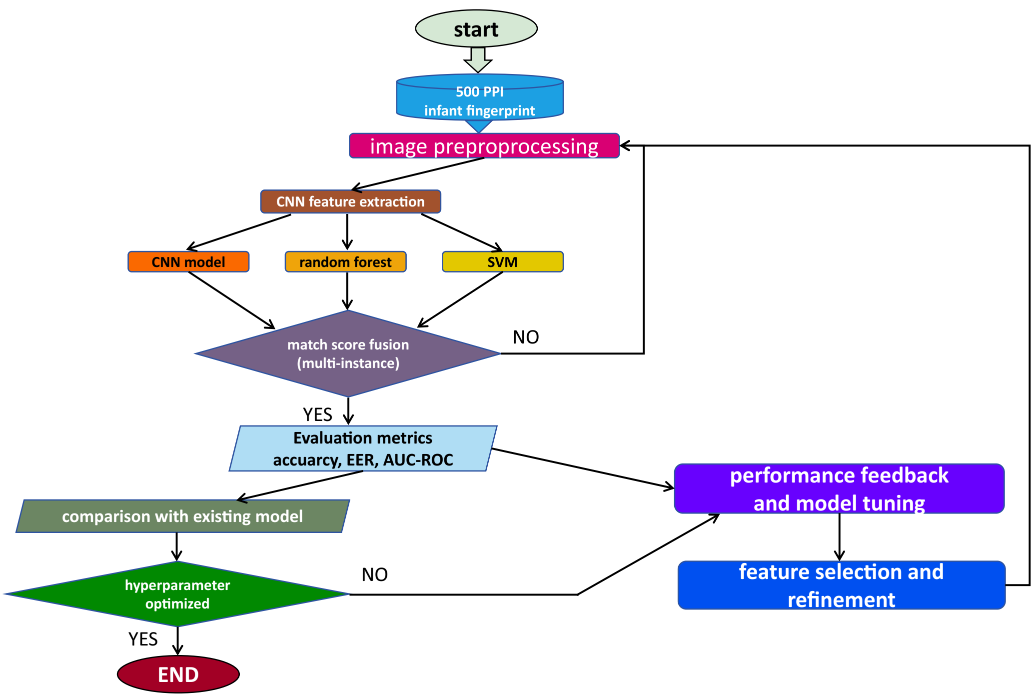
$$Y_i = \sum_{j=1}^{N} P_j W_{ji} + b_i \tag{17}$$

where Y_i is the output or activation input of the ith neuron (before applying any activation function), N is the number of inputs or previous layer neurons, P_j is the input value from the jth node or feature, W_{ji} is the weight connecting the jth input to the ith output neuron, and b_i is the bias term added to the output of the ith neuron. This is typically used in fully connected (dense) layers of neural networks. It calculates the weighted sum of inputs P, applies weights W, and adds bias b, which is then often passed through an activation function like ReLU, sigmoid, or SOFTMAX. Then the SOFTMAX activation converts logits into class probabilities using Equation 18:

$$P\ (c_i) = \frac{e^{Y_i}}{\sum_j e^{Y_j}} \tag{18}$$

where c_i represents class i.

Figure 4
Operation flowchart



3.4.4. Loss function

The categorical cross-entropy loss is defined as:

$$L = -\sum_{i=1}^{C} y_i \log(P(c_i)) \tag{19}$$

This is the categorical cross-entropy loss function, commonly used in classification tasks, where L is the loss value, which is a measure of how well the model's predictions match the true labels, C is the number of classes, and y_i is the true label for class iii. It's usually one-hot encoded, meaning $y_i = 1$ if the instance belongs to class iii, $y_i = 0$ otherwise, and $P(c_i)$ is the predicted probability that the input belongs to class iii, usually output by a SOFTMAX function. For instance, if the true class is class ii, and the predicted probabilities for three classes are $P(c_1) = 0.2$, $P(c_2) = 0.5$, $P(c_3) = 0.3$

Then, $y = [0, 1, 0]$, and $L = -(0 \cdot \log(0.2) + 1 \cdot \log(0.5) + 0 \cdot \log(0.3)) = -\log(0.5)$

3.4.5. Optimization (gradient descent)

Using Adam optimizer, weight updates follow:

$$W_{t+1} = W_t - \eta \frac{m_t}{\sqrt{v_t} + \epsilon} \tag{20}$$

This is the Adam (Adaptive Moment Estimation) optimization algorithm, a popular algorithm used for training deep learning models, where W_t is the current weight at time step t, W_{t+1} is the updated weight for the next iteration, η is the learning rate (step size), m_t is the first moment estimate (the exponentially decaying average of past gradients), v_t is the second moment estimate (the exponentially decaying average of past squared gradients), and ϵ is the small constant to avoid division by zero (e.g., 10–8). Adam adapts the learning rate for each parameter based on the first and second moments of the gradients. This leads to faster convergence, better handling of sparse gradients, and reduced need for fine-tuning the learning rate manually. However, m_t and v_t are computed using Equations 21 and 22, respectively. First moment estimate (mean of gradients) is:

$$m_t = \beta_1 m_{t-i} + (1 - \beta_1) g_t \tag{21}$$

where m_t is the first moment (like momentum) at time t, β_1 is the decay rate for the moving average of gradients (typically $\beta_1 = 0.9$), g_t is the current gradient at time t, and the second moment estimate (uncentered variance):

$$v_t = \beta_2 m_{t-i} + (1 - \beta_2) g_t^2 \tag{22}$$

where v_t is the second moment (squared gradients) at time t, β_2 is the decay rate for the moving average of squared gradients (typically $\beta_2 = 0.999$), and g_t^2 is the element-wise square of the current gradient.

For clarity's sake, m_t helps to smooth out noisy gradients and provides direction (like momentum), and v_t helps to scale the learning rate for each parameter based on the magnitude of past gradients, preventing large updates.

3.5. Hybrid model enhancements: CNN and support vector machine (SVM)

In this case, the output of the CNN feature extractor is passed to an SVM classifier. The SVM objective function is mathematically presented in Equation 23:

$$\min_{w,b,\xi} \frac{1}{2}||w||^2 + C \sum_{i=1}^{N} \xi_i \tag{23}$$

$$\text{Subject to; } y_i\left(w^T x_i + b\right) \geq 1 - \xi_i,\ \xi_i \geq 0 \quad \text{for all } i = 1, \ldots, N \tag{24}$$

where w^T is the weight vector, b is the bias term, $||w||^2$ is the norm (magnitude) of the weight vector, used for margin maximization, ξ_i is the slack variables that allow misclassifications, C is the regularization parameter that balances margin size vs classification error, N is the number of training samples, and x_i and y_i are the training data points and corresponding labels. Note that $||w||^2$ minimizes the weight norm, which maximizes the margin between classes. The term $C \sum_{i=1}^{N} \xi_i$ penalizes violations of the margin (i.e., misclassified or borderline samples). The slack variables ξ_i let some points be within the margin or misclassified, enabling better generalization on noisy data.

3.6. Hybrid model enhancements: CNN and random forest (RF)

The feature set extracted by CNN is fed into an RF classifier, where decision trees aggregate predictions:

$$P(c) = \frac{1}{T} \sum_{t=1}^{T} P_t(c) \tag{25}$$

The formula in Equation 25 is a temporal or ensemble averaging of class probabilities across time steps or models.

where $P(c)$ is the final averaged probability of class ccc, T is the total number of time steps or tree (e.g., in a sequence model) or total number of models (in an ensemble), and $P_t(c)$ is the predicted probability of class c at time step or from model t. Figure 5 shows some samples of the collected infants' fingerprints used for training and testing the model.

The methodology is summarized in Figure 6.

3.7. Contributions

The following are the expected contributions to knowledge aimed to achieve at the completion of this work:

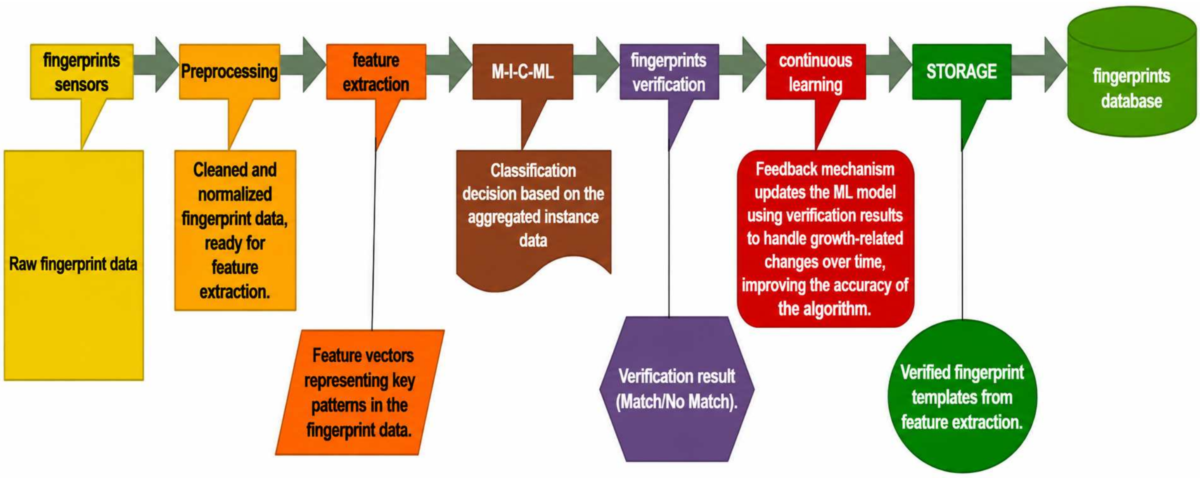
1) Development of a unique multi-instance contingent machine learning model tailored specifically for verifying infants' fingerprints.
2) Acquisition of 500 PPI fingerprint data for infants less than 12 months old in two sessions.

4. Experimental Section

This section presents and discusses the results obtained from each module of the conceptual framework of this study. These include the CNN model performance, the CNN + SVM model

Figure 5
Fingerprint samples of infants

Figure 6
Block diagram of the model



performance, the CNN + RF model performance, and the performance evaluation of the developed system, which will be presented herein.

4.1. Lagos State Government neonate and infant fingerprint (LASG-NIF) dataset

This acquired dataset comprises fingerprint samples from five hundred (500) newborns and infants, including both males and females. Participants were categorized into two age groups: newborns (1–28 days old) and infants (29 days–10 months old). Detailed statistical summary of the dataset, along with a bar chart illustrating the distribution of subjects across age groups at enrolment (Session 1), is provided in Tables 1 and 2 and Figure 7, respectively.

4.2. CNN model performance

The confusion matrix shown in Figure 7 visualizes the prediction performance of the CNN model. It shows the number of correct and incorrect classifications per class. From the 500 fingerprint images labeled with corresponding classes 0 to 9, the preprocessed images are resized to 64 × 64 resolution before normalizing the pixels to the range [0,1]. The datasets are then augmented to encode the labels using one-hot encoding or as integers and divided into 80% for training and 20% for testing. Keras from the Jupyter coding environment (JCE) is then used to define the proposed CNN model. Optimizer in the JCE is deployed to compile the model and fit it to the training data.

The accuracy of the CNN (Attention-Based Convolutional Neural Network (ACNN)) model from the Confusion Matrix of Figure 7 is calculated below:

$$\text{Accuracy} = \frac{2+6+7+10+9+3+13+4+8+13}{100} = \frac{75}{100} = 0.75$$

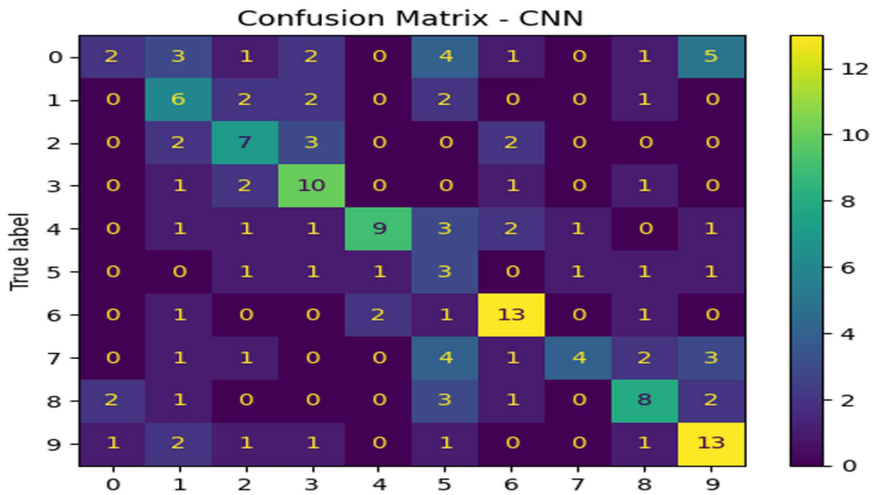
Table 1
LASG-NIF dataset statistics

Parameters	Frequency
Number of infants	500
Age at enrolment	1 day–10 months
Number of subjects with 1 session	297
Number of subjects with 2 sessions	111
Number of subjects with 3 sessions	92
Male: female	38%: 62%

Table 2
Summary of the numerical values from the accuracy and loss curves for CNN model

Metric	Epoch 0	Epoch 5	Epoch 10	Epoch 17.5
Train loss	3.4	2.3	1.7	1.2
Validation loss	2.3	2.3	2.2	1.7
Train accuracy	10%	13%	31%	58%
Validation accuracy	50%	46%	68%	84%

Figure 7
Confusion matrix for CNN model



Therefore, ACNN has an accuracy of 75%.

In addition, Figures 8 and 9 represent the accuracy curves and the loss curves, respectively, for the CNN model over multiple epochs. They help in understanding the training progress and convergence behavior. Figure 8 shows that the validation and train accuracies increase almost proportionally from Epoch 0 to Epoch 17.5, signaling a significant improvement, with train accuracy starting from 10% accuracy and terminating at an increased accuracy of approximately 62%. While the validation accuracy also increased from an accuracy of exactly 50% to an accuracy of approximately 83%, depicting a high performance on training data.

On the other hand, Figure 9 shows that the train loss sporadically reduced from a loss of 3.4 to about 1.1 while the validation loss reduced from 2.3 to about 1.7, signaling a significant minimization of the validation loss on the dataset.

Summarily, the numerical values obtained from Figures 8 and 9 are presented in Table 2.

Figure 8
Training accuracy curve for CNN

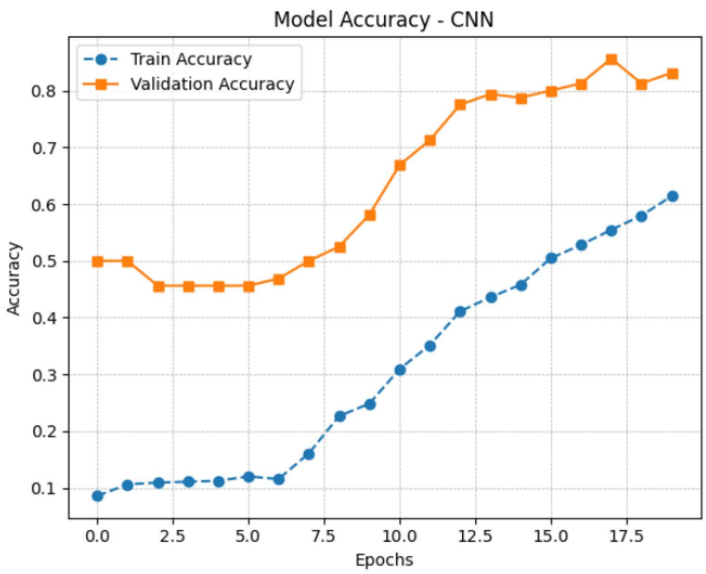
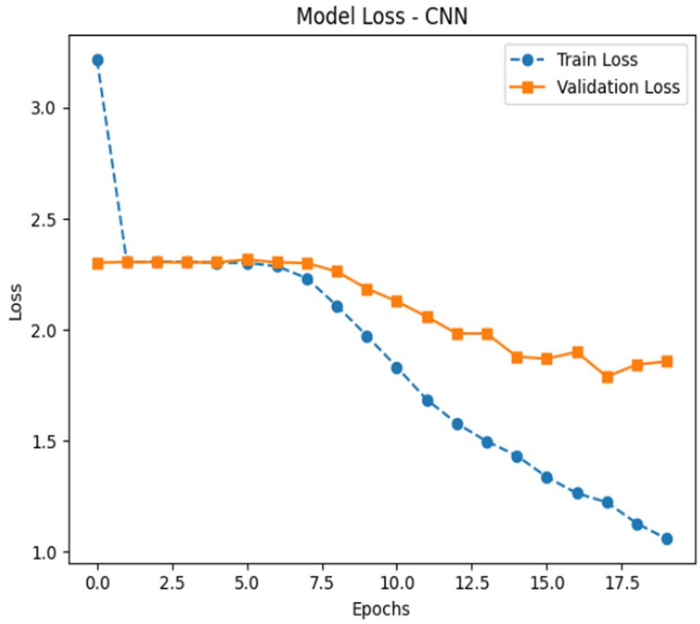


Figure 9
Training loss curve for CNN



4.3. CNN + SVM hybrid model performance

The confusion matrix for CNN + SVM, as shown in Figure 10, illustrates the model's classification accuracy. It helps in analyzing misclassification trends. It shows the number of correct and incorrect classifications per class.

The accuracy of the CNN (ACNN + SVM) model from the confusion matrix of Figure 10 is calculated below:

$$\text{Accuracy} = \frac{5+5+6+14+7+4+14+4+7+16}{100}$$
$$= \frac{82}{100} = 0.82$$

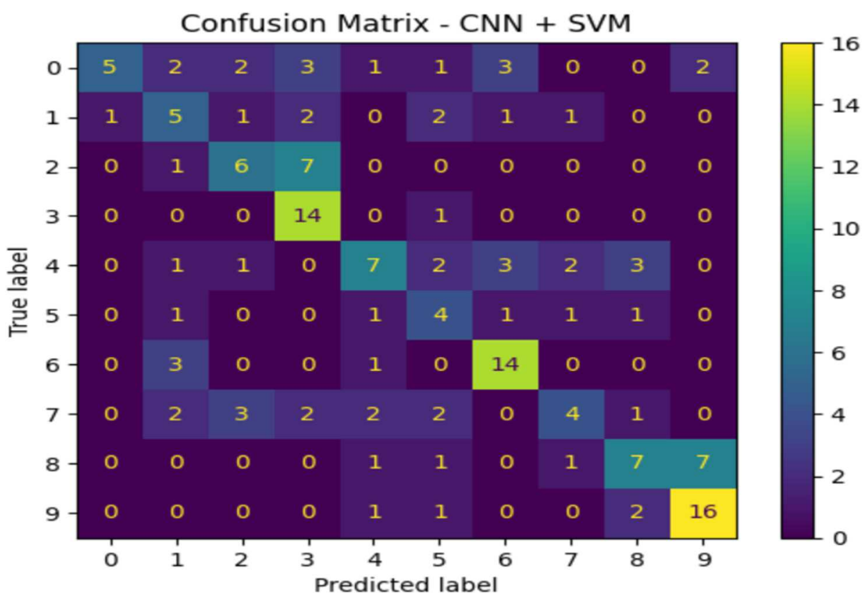
Therefore, ACNN + SVM has an accuracy of 82%.

Figures 11 and 12 illustrate the accuracy and loss curves, respectively, for the hybrid CNN + SVM model across multiple epochs, providing insights into its training progression and convergence behavior.

Figure 11 demonstrates a consistent and nearly proportional increase in both training and validation accuracy from Epoch 0 to Epoch 17.5, indicating substantial model improvement. The training accuracy rises sharply from an initial 10% to approximately 50%, while the validation accuracy shows even more robust growth, climbing from 48% to around 83%. This notable disparity suggests that the model generalizes exceptionally well to unseen data, achieving high performance on the validation set.

However, in Figure 12, the graph shows how there is a steady drop in training loss and validation loss. While

Figure 10
Confusion matrix for CNN + SVM hybrid model



training loss reduces inconsistently from 3.2 to 1.25, the validation loss reduces from 2.28 to about 1.75. The notable loss reduction clearly shows that there is model optimization through the training epochs.

Summarily, the numerical values obtained from Figures 11 and 12 are presented in Table 3.

4.4. CNN + RF model performance

The confusion matrix for CNN + RF, as shown in Figure 13, illustrates the model's classification accuracy. It helps in analyzing misclassification trends. It shows the number of correct and incorrect classifications per class.

Figure 11
Training accuracy curve for CNN + SVM

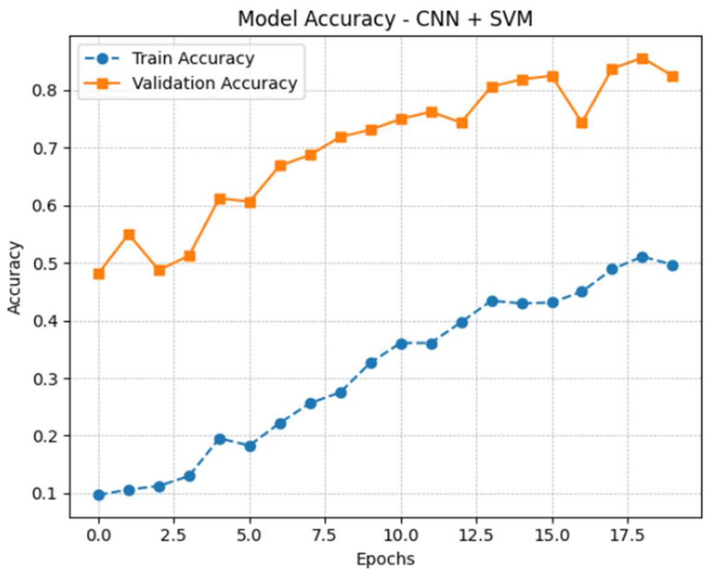


Figure 12
Training loss curve for CNN + SVM

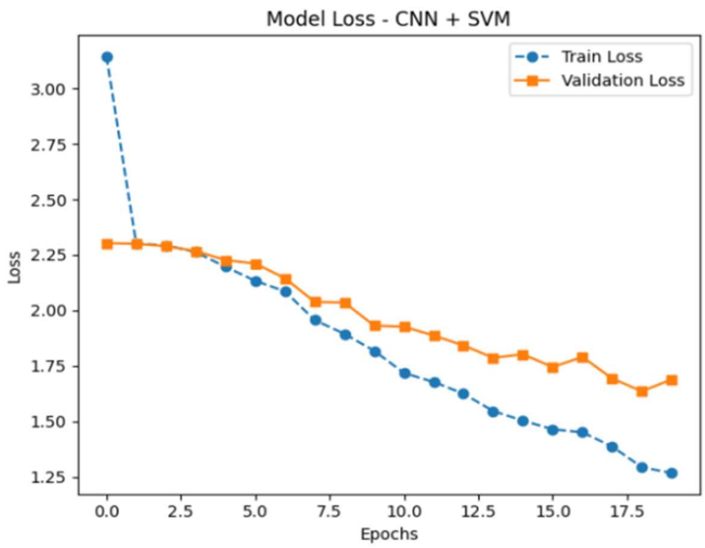


Table 3
Summary of the numerical values from the accuracy and loss curves for the CNN + SVM model

Metric	Epoch 0	Epoch 5	Epoch 10	Epoch 17.5
Train loss	3.2	2.12	1.7	1.3
Validation loss	2.3	2.2	1.9	1.7
Train accuracy	10%	18%	37%	50%
Validation accuracy	48%	60%	75%	83%

The accuracy of the CNN (ACNN + RF) model from the confusion matrix of Figure 13 is calculated below:

$$\text{Accuracy} = \frac{2+6+5+10+7+5+15+6+12+13}{100}$$
$$= \frac{81}{100} = 0.81$$

Therefore, ACNN + SVM has an accuracy of 81%.

Figures 14 and 15 present the accuracy and loss curves, respectively, for the hybrid CNN + RF model across multiple epochs, offering valuable insights into its training dynamics and convergence behavior.

Figure 14 reveals a great and consistent improvement in model performance, with both training and validation accuracy increasing steadily from Epoch 0 to Epoch 17.5. The training accuracy exhibits a dramatic surge, rising from an initial 8% to 98%, while the validation accuracy shows robust growth, climbing from 45% to approximately 85%. This substantial gap between training and validation accuracy highlights the model's strong generalization capability, maintaining high performance on unseen data.

In contrast, Figure 15 tracks the corresponding loss trajectories, demonstrating a clear downward trend in both training and validation losses. The training loss declines sharply from 3.1 to 1.0, whereas the validation loss decreases more modestly from 2.3 to 2.25. This reduction in loss values signifies effective model optimization, though the smaller drop in validation loss suggests potential room for further refinement.

In summary, the key numerical trends extracted from Figures 14 and 15 are consolidated in Table 4 for comprehensive analysis.

Figure 13
Confusion matrix for CNN + RF hybrid model

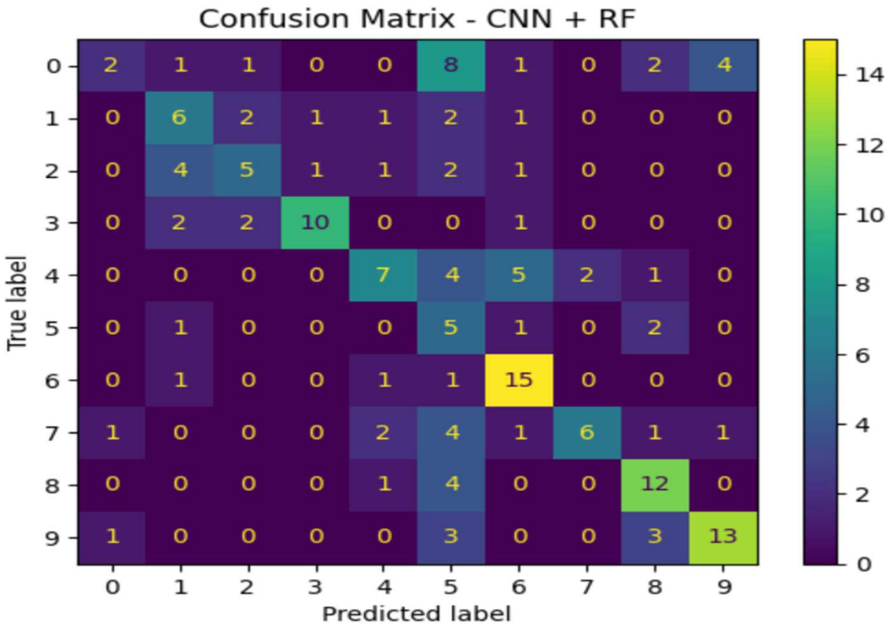


Figure 14
Training accuracy curve for CNN + RF

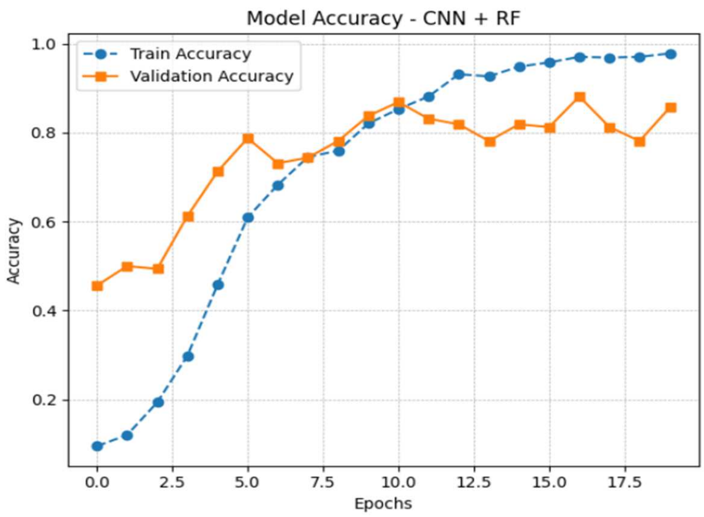


Figure 15
Training loss curve for CNN + RF

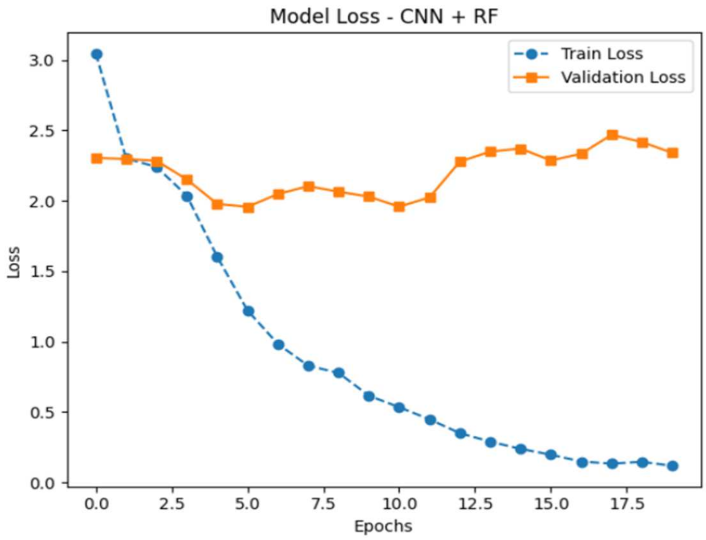


Table 4
Summary of the numerical values from the accuracy and loss curves for the CNN + RF model

Metric	Epoch 0	Epoch 5	Epoch 10	Epoch 17.5
Train loss	3.1	1.25	0.5	0.2
Validation loss	2.3	1.8	1.8	2.5
Train accuracy	5%	60%	83%	98%
Validation accuracy	45%	80%	84%	85%

4.5. Comparison of performance metrics among the three models

The performances of the three developed models (CNN, CNN + SVM, and CNN + RF) are compared here to determine the most accurate model. This is simply done by comparing metrics like the train and validation accuracies, train and validation losses, accuracies from their respective confusion matrices, F1-score, recall, and precision.

4.5.1. Comparison of accuracies from the confusion matrices

As calculated from the confusion matrix of Figure 7, which is for the standalone CNN model, the accuracy is 75%.

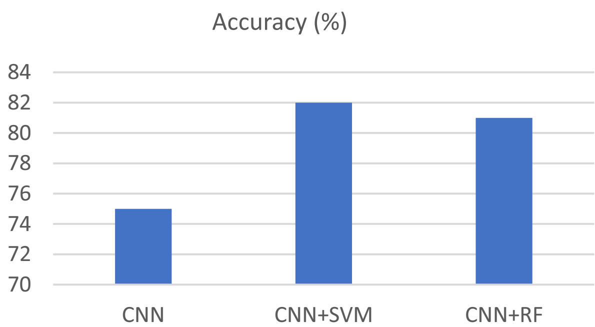
The accuracy obtained from the confusion matrix of Figure 10, which corresponds to that of the hybrid CNN + SVM model, gave 82%.

However, the accuracy gotten from the confusion matrix of Figure 13, corresponding to the hybrid model CNN + RF, gave 81%, and these are summarized in Table 5 and visualized in Figures 4 and 16.

Table 5
Comparison of accuracies for CNN, CNN + SVM, and CNN + RF

Model	Accuracy (%)
CNN	75
CNN + SVM	82
CNN + RF	81

Figure 16
Comparison of the accuracies from the confusion matrices



4.5.2. Comparison of train and validation accuracies

Table 6 compares the train and validation accuracies of CNN, CNN + SVM, and CNN + RF models at epoch 17.5. It helps in evaluating which model performs best in terms of classification accuracy, while Figure 17 depicts the graph of the accuracies for a better visualization.

From Figure 17, it is evident that CNN + RF outperforms the other two models.

4.5.3. Comparison of train and validation losses

Table 7 compares the train and validation losses of CNN, CNN + SVM, and CNN + RF models at epoch 17.5. It helps in evaluating which model performs best by having the lowest losses, while Figure 18 depicts the graph of the losses for a better visualization.

Table 6
Comparison of train and validation accuracies for CNN, CNN + SVM, and CNN + RF

Model	Validation accuracy (%)	Train accuracy (%)
CNN	84	58
CNN + SVM	83	50
CNN + RF	85	98

Figure 17
Comparison of train and validation accuracies for the models

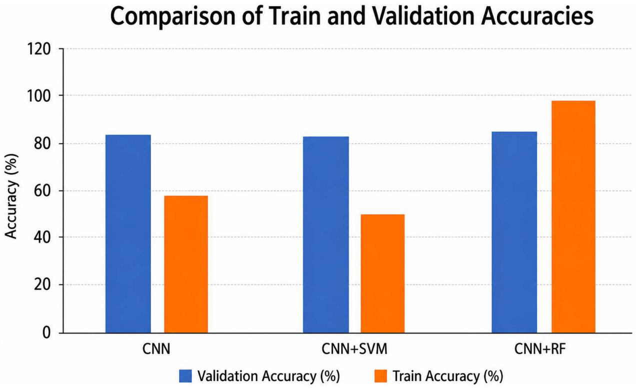
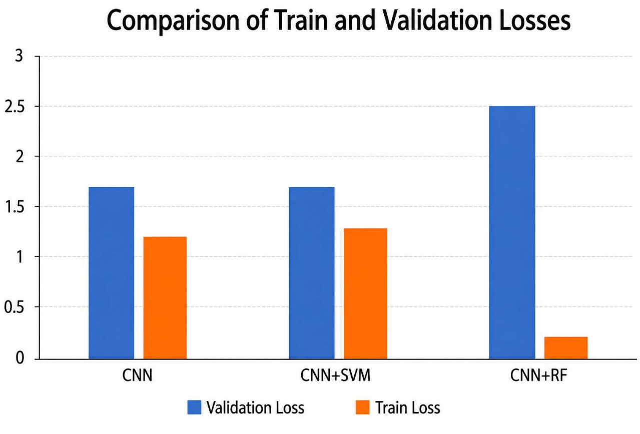


Table 7
Comparison of train and validation losses for CNN, CNN + SVM, and CNN + RF

Model	Validation loss	Train loss
CNN	1.7	1.2
CNN + SVM	1.7	1.3
CNN + RF	2.5	0.2

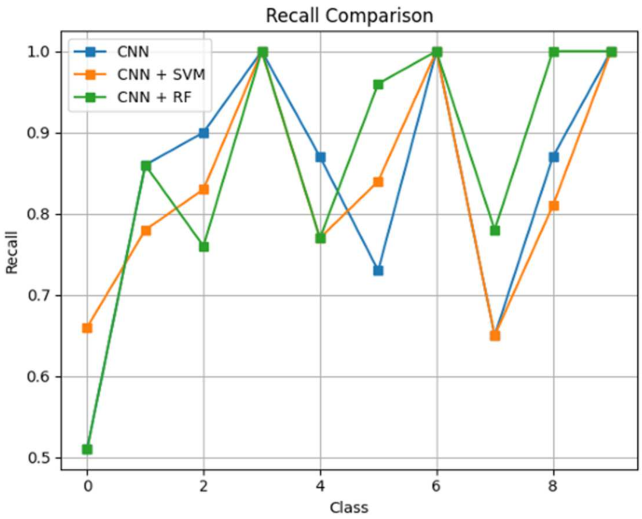
Figure 18
Train and validation losses for the models



4.5.4. Comparison of precision across models

Figure 19 compares the precision for CNN, CNN + SVM, and CNN + RF models. It provides insights into the classification performance of each model across different classes. It shows that CNN, CNN + SVM, and CNN + RF approximately have a precision of 83%, 88%, and 88%. Both hybrid models obviously outperformed the standalone CNN model with respect to precision.

Figure 19
Precision across the models



4.5.5. Comparison of recall across models

Figure 20 compares the recall for CNN, CNN + SVM, and CNN + RF models. It provides insights into the classification performance of each model across different classes. It shows that CNN, CNN + SVM, and CNN + RF approximately have a recall of 84%, 84%, and 87%. The CNN + RF model obviously outperformed the standalone CNN model and the hybrid CNN + SVM with respect to recall.

4.5.6. Comparison of F1-score across models

Figure 21 compares the F1-score for CNN, CNN + SVM, and CNN + RF models. It provides insights into the classification performance of each model across different classes. It shows that CNN, CNN + SVM, and CNN + RF approximately have F1-score of 85%, 86%, and 88%. All three models achieved very high F1-scores, but the hybrid CNN + RF model obviously outperformed the standalone CNN model and the hybrid CNN + SVM with respect to F1-score.

4.5.7. Comparison of execution time across the models

Figure 22 illustrates the execution time required for each model. It helps in assessing the computational efficiency of CNN, CNN + SVM, and CNN + RF.

Figure 20
Recall across the models

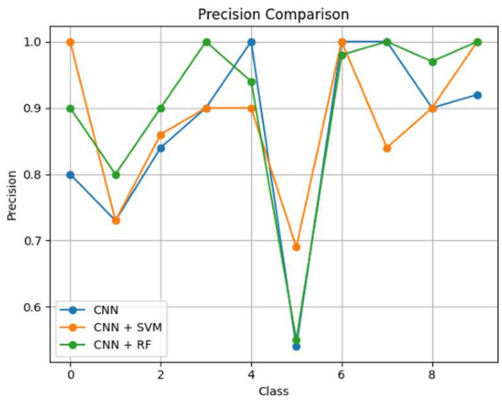


Figure 21
F1-score across the models

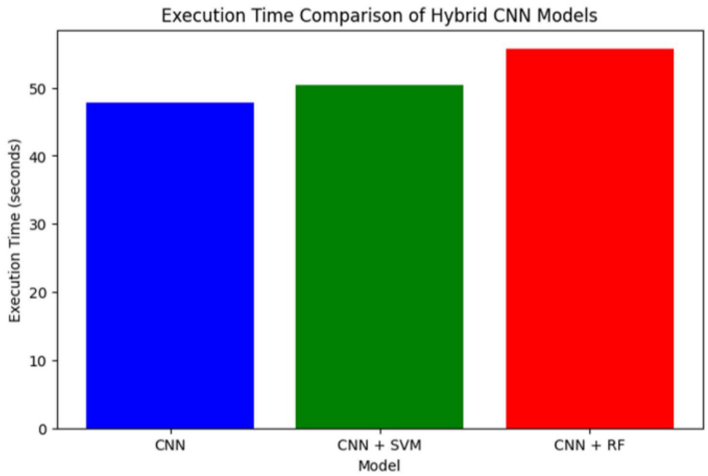
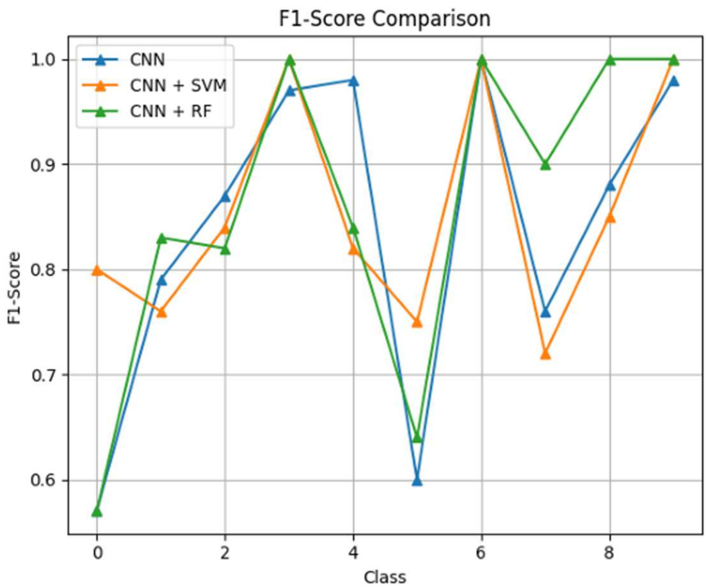


Figure 22
Execution time across the models



4.6. Benchmarking this model with existing models to validate performance improvements

Several works have been done on the verification of infants' fingerprints with performance accuracies ranging from 66.7% [22] for ages between 0 and six months, 65.8% [23] for ages between 0 and three months, to 69.05% [24], as presented in Table 8.

Meanwhile, the three models introduced in this study (CNN, CNN + SVM, and CNN + RF) have accuracies of 75%, 82%, and 81%, respectively, and all outperformed any of the existing methods (they all have accuracies below 70%), the latest being with an accuracy of 69.05% [24]. The visual representation of the benchmark is presented in Figure 23 and in Table 8.

Table 8
Comparison of this proposed method with state-of-the-art methods

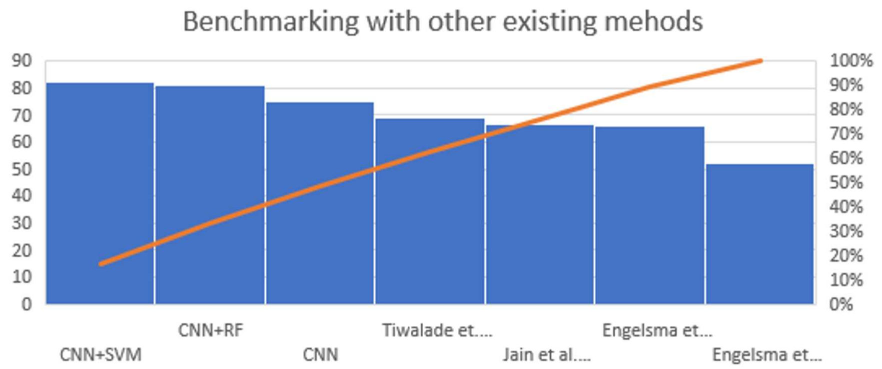
Author(s)	Methodology	Age of subjects at enrolment (month(s))	Verification accuracy (%)
[22]	Multi-sample and multi-instance fusion	0–6	66.70 after a time-lapse of 6 months
[23]	Multi-sample and multi-instance fusion	0–3	65.80 after a time-lapse of 3 months

(*Continued*)

Table 8
(*Continued*)

Author(s)	Methodology	Age of subjects at enrolment (month(s))	Verification accuracy (%)
[23]	Multi-algorithm, multi-sample, and multi-instance fusion	0–2 2–3	35.50 52.10
[24]	Multi-instance contingent	1–4	69.05 after a time-lapse of 3 months
[25]	Hybrid machine learning-based	1–4	CNN = 75 CNN + SVM = 82 CNN + RF = 81

Figure 23
Comparison of the accuracy of this study with existing models



5. Conclusion

An infant fingerprint identification system using a multi-instance dependent machine learning technique was effectively created in this study, providing an effective solution to all the main problems related to infant biometric identification. MIL techniques were used in this research to ensure that the system could use multiple fingerprints from each infant, enhancing accuracy in identification and resistance to biometric changes like immature ridge formation, small fingerprint surface area, and quick changes in physiology resulting from infant growth. The design incorporated complex preprocessing stages, feature extraction using CNN, and contingent learning models with the ability to make optimized choices by relying on the most accurate fingerprint data. A set of 500 PPI images from infants under 12 months old was gathered, and the images went through the process of CNN-based preprocessing and feature extraction. Three model setups, CNN, CNN + SVM, and CNN + RF, were tested in terms of match score fusion involving the two fingerprints. It emerged that CNN + RF performed better than the other models, registering the highest accuracy of 98% and 85% at the levels of training and validation and the lowest training loss of 0.2%. Furthermore, this model exhibited more robust precision, recall, F1-score, and generalization. In spite of being computationally expensive, CNN + RF has been regarded as the most effective and reliable model for implementing fingerprint authentication technology in practice. The use of fingerprints combined with the soft biometric features of gender, blood group, and ethnicity on a large dataset is suggested for further studies.

Ethical Statement

The study got ethical clearance from the concerned Institutional Review Board/Research Ethics Committee. Informed written consent was taken from the parents or legal representatives of all the infants prior to data collection. All fingerprint data was anonymized, stored safely, and used only for scientific research purposes in adherence to ethical and data protection laws. This paper is devoid of any studies involving human or animal subjects conducted by the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data available on request from the corresponding author upon reasonable request.

Author Contribution Statement

Olatayo Moses Olaniyan: Conceptualization, Methodology, Software, Validation, Investigation, Resources, Data curation, Writing – original draft, Visualization, Supervision, Project administration. **Kazeem Aderemi Bello:** Conceptualization, Methodology, Investigation, Resources, Data Curation, Writing – original draft. **Rendani Wilson Maladzhi:** Conceptualization, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing. **Johnson Olugbenga Joseph:** Conceptualization, Investigation, Resources, Data curation, Writing – original draft. **Adebimpe Esan:** Methodology, Software, Validation, Formal analysis, Writing – review & editing, Visualization, Supervision, Project administration. **Quadri Ademola Mumuni:** Conceptualization, Investigation, Resources, Data Curation, Writing – original draft. **Bolaji Abigael Omodunbi:** Formal analysis, Writing – review & editing. **Chinedu Michael Chiejine:** Software, Project administration.

References

[1] UNICEF. (2019). *Birth registration for every child by 2030: Are we on track?* https://www.unicef.org/media/62981/file/Birth-registration-for-every-child-by-2030.pdf

[2] De Kalaf, E. H (Ed.). (2021). *Legal identity, race and belonging in the Dominican Republic*. UK: Anthem Press.

[3] Ipinnimo, O., Ogbemhe, J., Mumuni, Q., Owoeye, S. O., & Olurombi, A. (2024). Facial recognition-based online voting system with two-factor security. *The Journal of Innovation Science and Technology*, *3*(1), 107–114.

[4] Falch-Eriksen, A., Toros, K., Sindi, I., & Lehtme, R. (2021). Children expressing their views in child protection casework: Current research and their rights going forward. *Child & Family Social Work*, *26*(3), 485–497. https://doi.org/10.1111/cfs.12831

[5] Draper, C. E., Klingberg, S., Wrottesley, S. V., Milner, K., Fisher, J., Lakes, K. D., & Yousafzai, A. K. (2023). Interventions to promote development in the next 1000 days: A mapping review. *Child: Care, Health and Development*, *49*(4), 617–629. https://doi.org/10.1111/cch.13084

[6] Betrand, C. U., Onyema, C. J., Benson-Emenike, M. E., Kelechi, D. A., & Benson-Emenike, M. E. (2023). Authentication system using biometric data for face recognition. *International Journal of Sustainable Development Research*, *9*(4), 68–78. https://doi.org/10.11648/j.ijsdr.20230904.12

[7] Barnert, E., Katsanis, S. H., Mishori, R., Wagner, J. K., Selden, R. F., Madden, D., . . . , & Svetaz, M. V. (2021). Using DNA to reunify separated migrant families. *Science*, *372*(6547), 1154–1156. https://doi.org/10.1126/science.abh3979

[8] Katsanis, S. H. (2023). DNA in immigration and human trafficking. In D. Primorac & M. Schanfield (Eds.), *Forensic DNA applications* (pp. 459–476). CRC Press. https://doi.org/10.4324/9780429019944

[9] Jiang, X., & Ge, Z. (2021). Information fingerprint for secure industrial big data analytics. *IEEE Transactions on Industrial Informatics*, *18*(4), 2641–2650. https://doi.org/10.1109/TII.2021.3104056

[10] Bansal, P., & Ouda, A. (2024). Continuous authentication in the digital age: An analysis of reinforcement learning and behavioral biometrics. *Computers*, *13*(4), 103. https://doi.org/10.3390/computers13040103

[11] Haraksim, R., Galbally, J., & Beslay, L. (2019). Fingerprint growth model for mitigating the ageing effect on children's fingerprints matching. *Pattern Recognition*, *88*, 614–628. https://doi.org/10.1016/j.patcog.2018.12.024

[12] Moolla, Y., De Kock, A., Mabuza-Hocquet, G., Ntshangase, C. S., Nelufule, N., & Khanyile, P. (2021). Biometric recognition of infants using fingerprint, iris, and ear biometrics. *IEEE Access*, *9*, 38269–38286. https://doi.org/10.1109/ACCESS.2021.3062282

[13] Elsedimy, E. I., Elhadidy, H., & Abohashish, S. M. (2024). A novel intrusion detection system based on a hybrid quantum support vector machine and improved Grey Wolf optimizer. *Cluster Computing*, *27*(7), 9917–9935. https://doi.org/10.1007/s10586-024-04458-8

[14] Zhang, Y., Chu, J., Leng, L., & Miao, J. (2020). Mask-refined R-CNN: A network for refining object details in instance segmentation. *Sensors*, *20*(4), 1010. https://doi.org/10.3390/s20041010

[15] Lin, W., Chu, J., Leng, L., Miao, J., & Wang, L. (2024). Feature disentanglement in one-stage object detection. *Pattern Recognition*, *145*, 109878. https://doi.org/10.1016/j.patcog.2023.109878

[16] Wang, H., & Wang, J. (2024). Collaborative radio frequency fingerprint identification using dual-channel parallel CNN. In *2024 International Conference on Ubiquitous Communication*, 351–355. https://doi.org/10.1109/Ucom62433.2024.10695867

[17] Yan, G., Fu, X., Wang, Y., Zhang, Q., & Gui, G. (2025). Radio frequency fingerprint identification towards statistical and deep learning features: Review, recent results and future directions. *Peer-to-Peer Networking and Applications*, *18*(3), 116. https://doi.org/10.1007/s12083-024-01902-9

[18] Yang, Y., & Yan, T. (2022). RF fingerprint recognition method based on DBN-SVM. In *2022 IEEE 10th International Conference on Information, Communication and Networks*, 572–577. https://doi.org/10.1109/ICICN56848.2022.10006500

[19] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, *2*(3), 1–21. https://doi.org/10.1007/s42979-021-00592-x

[20] Krishnakumar, K., Smitha, G. L., Gnana Soundari, A., Dhanalakshmi, R., & Kesavan, R. (2025). Optimized fingerprint crime detection using robust deformed convolutional neural network for 5G network secure smart cities. *Knowledge-Based Systems*, *317*, 113342. https://doi.org/10.1016/j.knosys.2025.113342

[21] Kavitha, T., Patil, B., Saraswathi, S., Rajarajeswari, S., & Patil, A. (2024). Recognition of fingerprint images using CNN for cybercrime detection system. In *2024 Second International Conference on Networks, Multimedia and Information Technology*, 1–6. https://doi.org/10.1109/NMITCON62075.2024.10698773

[22] Engelsma, J. J., Deb, D., Cao, K., Bhatnagar, A., Sudhish, P. S., & Jain, A. K. (2021). Infant-ID: Fingerprints for global good. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(7), 3543–3559. https://doi.org/10.1109/TPAMI.2021.3057634

[23] Engelsma, J., Deb, J., Jain, D., Bhatnagar, A., A, & Sewak Sudhish, P. (2019). Infant-prints: Fingerprints for reducing infant mortality. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 67–74.

[24] Southier, L. F. P., Nunes, G. A. T., Machado, J. H. P., Buratti, M., Trentin, P. H. D. V., Bona, W. A. C. D., . . . , & Casanova, D. (2025). A systematic literature review on neonatal fingerprint recognition. *ACM Computing Surveys*, *57*(12), 1–34. https://doi.org/10.1145/3735551

[25] Jain, A. K., Arora, S. S., Cao, K., Best-Rowden, L., & Bhatnagar, A. (2016). Fingerprint recognition of young children. *IEEE Transactions on Information Forensics and Security*, *12*(7), 1501–1514. https://doi.org/10.1109/TIFS.2016.2639346

How to Cite: Olaniyan, O. M., Bello, K. A., Maladzhi, R. W., Joseph, J. O., Esan, A., Mumuni, Q. A., …, & Chiejine, C. M. (2026). A Multi-Instant Machine Learning Algorithm Model for Infant Fingerprinting Verification. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA62027314