Scientific Journal of Computer Science
futuristech PT. Teknologi Futuristik Indonesia
Received: March 18, 2026; Revised: April 30, 2026; Accepted: May 2, 2026; Published: May 15, 2026; Current version: May 15, 2026. *e-ISSN: 3110-3170*
Digital Object Identifier 10.64539/sjcs.v2i2.2026.463

Article

Deep Learning for Venomous and Non-Venomous Snakes Classification

Yakubu Abubakar Lidani[1,*], Abdullahi Musa Yola[1], Abu Tasiu[2], Nura Muhammad Sani[2], Sulaiman Muhammad Gidado[2]

[1] Department of Computer Science, Federal University Kashere, Gombe state, 771103, Nigeria; e-mail: lidaniyakubu@gmail.com (Y. A. Lidani).
[2] Department of Computer Science, Federal Polytechnic Kaltungo, Gombe State, 770101, Nigeria
* Correspondence

The authors received no financial support for the research, authorship, and/or publication of this article.

Abstract: Snakes are a major health threat in various communities, specifically where human and snake encounters are frequent. When a snake is not identified correctly, healthcare providers often administer the wrong treatment, this can worsen patient recovery outcomes or even prove fatal to the victim. Therefore, a fast, proper and accurate distinction between venomous and nonvenomous snakes is vital for proper antivenom administration. This study proposes a hybrid deep learning system combining a CNN and an LSTM model for snake image classification through feature extraction from visual data. The CNN extracts key spatial features such as colour and scale patterns, texture, and body shape, whereas the LSTM captures sequential dependencies across these features, by helping distinguish visual similarity amongst the species. The model was trained and evaluated on a dataset of 6,798 snake images from diverse sources. The system achieved a performance of 97% accuracy, 97% precision, 96% recall, an F1-score of 97%, and a ROC-AUC of 0.97. These results demonstrate that integrating CNN and LSTM is moderately effective for snake classification. The proposed system has practical applications in the area of emergency healthcare, wildlife management, as well as mobile based identification tool. With 97% accuracy, this model can improve emergency responders first aid, enhance a safer treatment administration and help make safer decisions on the use of antivenom, by reducing treatment delays and improving patient survival prognosis. This model has the potential to save lives and minimize the consequences of snakebite envenoming.

Keywords: Snakes; Classification; Deep Learning; Image Data; Long Short-Term Memory.

1. Introduction

Snakes play a vital role in global ecosystems, but their diversity also poses a serious threat to human lives. Of the 3,000 to 4,000 species worldwide, only about 600 are venomous, leading to frequent snakebite incidents in tropical and subtropical regions, as illustrated in Figure 1. According to the World Health Organization (WHO), 5.4 million people are bitten annually, resulting in 1.8–2.7 million cases of envenoming and 81,410–137,880 deaths each year [1]. survivors often suffer from permanent disabilities, with WHO estimating around three times as many amputations and other permanent disabilities caused by snakebites annually [2]. Identifying venomous snakes quickly is critical for effective treatment, yet reliable identification remains a major challenge, particularly in resource-limited regions such as Africa, Asia, and Latin America, where the burden is high and misidentification can lead to serious consequences [3]. Conversely, it is safe to know that non-venomous snakes do not pose threats to human life [4]. The traditional identification methods relying on herpetologist expertise, physical examination of bite marks, or the 20-minute whole-blood clotting test are slow and require trained personnel, who are also prone to error. Misclassification is common, especially in rural areas where experts are scarce. Administering the wrong antivenom due to incorrect identification can worsen patient outcomes, deplete scarce supplies, and even cause death. This is not merely a technical issue but a critical patient safety risk.

Most current automated approaches include convolutional neural networks (CNNs), which extract spatial features (e.g., colour, texture, shape) but ignore sequential or contextual patterns among image regions. Moreover, many studies report only accuracy, neglecting class imbalance and other metrics like precision and recall. High computational demands further limit real-world deployment, particularly in low-resource clinics where fast, reliable decisions are essential. Automated classification and identification of snake species have become highly important in several fields, including snakebite epidemiology, wildlife conservation, and ecological research [5].

To overcome these challenges, this paper introduces a hybrid deep learning approach that integrates a CNN with a long short-term memory (LSTM) network. Spatial features are extracted using the CNN, while sequential dependencies are learned by the LSTM, enabling better discrimination between visually similar species. Unlike prior CNN-only systems, our method jointly optimizes multiple evaluation metrics such as accuracy, precision, recall, F1 Score, and ROC AUC. Using data augmentation (random blurring, rotation, contrast changes) and training on 6,798 images, our model achieved 97% accuracy, 97% precision, 96.55% recall, and a 97% F1-score in just 15 epochs. The key contribution is a balanced, efficient, and interpretable snake identification system that can be deployed in emergency healthcare, wildlife management, or as a mobile app, directly supporting safer antivenom administration.

The paper is organized as follows: Section 2 reviews related work on snake identification and deep learning; Section 3 describes the dataset, preprocessing, and model architecture; Section 4 presents experimental results and comparisons; Section 5 concludes with a discussion of limitations and future directions.

2. Literature Review

The literature review shows substantial progress in using machine learning to classify snakes by species. [6] and [7] relied only on conventional machine learning classifiers and CNN. Their models achieve high accuracy but often lack robustness and interpretability. Their studies also did not use a comprehensive evaluation. They mainly emphasized accuracy as the primary metric, which provides limited insight into precision, recall, F1-score, and threshold-based performance. These are critical for sensitive applications in snakebite management. Later studies use deep learning and transfer learning. [8], [9], and [10] show that pretrained CNNs such as MobileNetV2, VGG16, and variants of DenseNet can improve classification performance. In some cases, these models achieve over 97% accuracy. Reviews highlight that data augmentation, object detection, and architectural changes enhance model training. However, many approaches rely only on one method. More so, [11] integrate the use of IoT and AI for automation. Six distinct organisms are classified by eliminating taxonomical features from the images by [12].

However, the approaches mostly rely on spatial feature mining rather than on integrating contextual or sequential relationship feature representations. Recent studies conducted between 2024 and 2026 have broadened the scope of snake species classification modelling by leveraging CNNs and transfer learning. Papers presented by [13] and [14] demonstrate that an efficient CNN model with a specialized preprocessing expert technique can achieve high accuracy while reducing computational cost. Likewise, [15] and [16] demonstrated the potential of self-supervised learning on both limited and large-scale global datasets. [17] employed CNN and using 415 images to classify 5 distinct snake species. Despite advances in these fields, transformer-based approaches have proven sensitive to dataset shifts, leading to poor performance in real-world settings. Several studies have explored the integration of machine learning for image-based species identification [18], [19]. Large datasets are analyzed and combined with sophisticated neural networks, enabling deep learning models to extract meaningful patterns from images. This makes the approach a promising method for snake species classification [20].

Across the reviewed studies, various commonalities and trends appeared. First, there is a strong reliance on CNNs, with the model's primary focus on spatial feature extraction. Secondly, various studies report high training and validation accuracy but reveal limited generalization performance on real-world datasets. The third objection is that most research has emphasized species classification over binary classification, which is vital for emergency decision-making. Finally, the evaluation metrics identified in the studies remain limited, indicating insufficient information on precision, recall, F1-score, confusion metrics, and ROC AUC.

3. Proposed Methodology

The proposed approach integrates two machine learning components: deep learning architectures that combine CNNs and LSTM networks. The CNN model was used to automatically extract discriminative features from snake images, while the LSTMs were harnessed to model sequential dependencies over these features. This is particularly valuable in the snake classification domain, as visual identification of many species relies on patterns that exhibit continuity along the snake's elongated body, such as colour banding, scale arrangement, and overall shape. By capturing the sequential flow and spatial continuity of these features, the LSTM component strengthens the model's ability to recognize domain-specific patterns that span across the length of the snake, even within a single image. This integration can enhance the model's classification accuracy while improving generalization in real-world

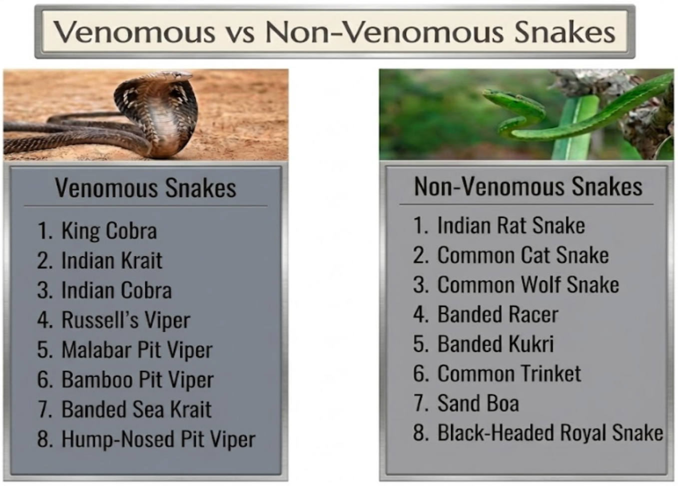

Figure 1. Snake classification under venomous and non-venomous categories.

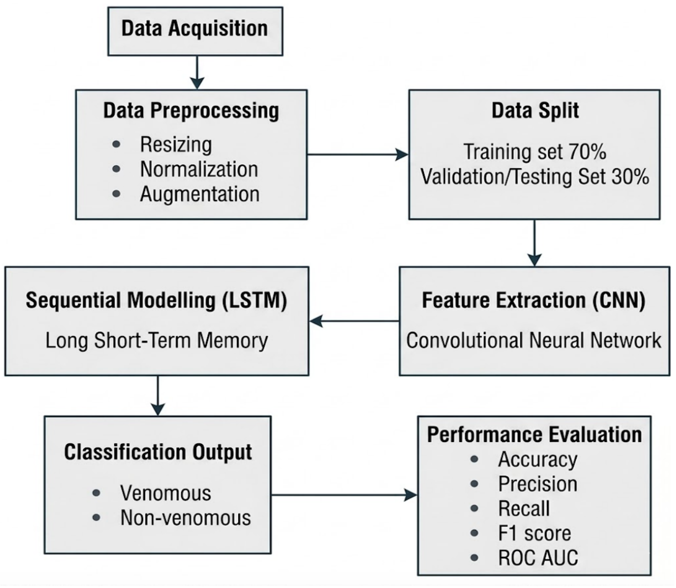

Figure 2. Architecture of the proposed methodology.

conditions, as illustrated in Figure 2. The figure explains the proposed hybrid methodology.

3.1. Dataset Description

The dataset used in this study was collected from [10]. It consists of 6,798 snake images, categorized into two main classes: venomous and non-venomous. There are 3,399 images of venomous snakes and 3,399 of non-venomous snakes, resulting in a well-balanced class distribution before training. The images were gathered from diverse sources and exhibit variations in species, pose, background, illumination, and image size, reflecting real-world conditions. The dataset is well-suited for binary classification, providing sufficient diversity for efficient training and evaluation of the proposed model.

3.2. Data Preprocessing

Data preprocessing plays a crucial role in the proposed method. One important technique used is data augmentation, also referred to as oversampling, which increases the dataset size by generating additional image variations through various image manipulation methods [21]. This process helps maintain data consistency, improves learning efficiency, and strengthens the model's ability to generalize.

The dataset consists of raw snake images collected from diverse sources. Preprocessing was used to standardize the input and minimize noise before training. First, all photos were resized to a specific size compatible with the CNN+LSTM hybrid model's input requirements. This resizing ensured uniformity across the dataset and reduced computational overhead during training. Pixel values were then normalized to stabilize gradient updates and accelerate model convergence. To further enhance robustness and prevent overfitting, data augmentation was also applied during training.

The augmentation techniques include horizontal and vertical flipping, random rotation, scaling, and slight zooming of the augmented dataset. Data augmentation can increase the diversity of training samples by enabling the model to learn invariant features across different conditions. In addition to that, the dataset was shuffled to ensure that the samples from both classes were evenly distributed during the model's training. The preprocessing pipeline played a crucial role in improving feature learning and contributed to the high classification performance achieved by the proposed model.

3.3. Dataset Splitting

The dataset is divided into three subsets, consisting of 70% for training, 15% for validation, and 15% for testing [22]. This split is applied to enable the model parameters to be learned while the weights of the proposed hybrid architecture are optimized. The validation set is utilized for hyperparameter tuning, model selection, and early stopping, which are implemented to reduce the risk of overfitting during the training process. Meanwhile, the testing set is reserved and used for final evaluation, ensuring that an unbiased assessment of the model's generalization ability on unseen data is provided.

3.4. Feature Extraction and Selection

The feature extraction part was performed automatically by the Convolutional Neural Network components layers. The convolutional layers learn graded spatial features from the uploaded images. They start by extracting low-level features, such as edges and corners, and gradually move to higher-level features, such as textures, shapes, and distinctive visual patterns relevant to distinguishing between venomous and non-venomous classes. This hierarchical learning feature enables the model to capture the distinctive visual pattern without manual feature engineering.

The mathematical convolutional procedure can be expressed as:

$$F(i,j,k) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} X(i+m, j+n)\cdot W(m,n,k) + b(k), \tag{1}$$

where F (i, j, k) signifies the output feature value at spatial location (i, j) for the k-th filter, Xdenotes the input image or feature map, W (m, n, k) is expressed as the convolutional kernel of size $M \times N$, and $b(k)$ is identified as the bias term. This equation enables the network to learn local spatial features, such as edges, contours, and texture patterns, which are pivotal for distinguishing venomous from non-venomous species. Also, a nonlinear function is employed to introduce nonlinearity. In this study, the Rectified Linear Unit (ReLU) activation function is applied and is defined as: ReLU(x) = max (0, x).

ReLU enhances learning by allowing positive feature values to propagate while suppressing negative ones, thereby mitigating the vanishing-gradient problem. To further reduce spatial dimensionality and computational complexity, a max pooling operation is applied. The Max-pooling selects the maximum value within a defined pooling window and is expressed below as:

$$\boldsymbol{P(i,j) = \max_{(m,n)\in\Omega} F(i+m, j+n)}, \tag{2}$$

where Ω denotes the pooling region, this operation enhances the translational invariance while preserving the most relevant features. Feature selection is implicitly achieved during training, as the CNN automatically suppresses the less important features while amplifying highly discriminative ones through weight optimization. The high-level feature representations are compact, informative, and robust. Those features are passed to the LSTM component for sequential modelling, ensuring efficient and accurate classification.

3.5. CNN+LSTM Model

The model integrates a CNN with a LSTM network, forming a hybrid ML architecture for classifying snake species using images. The CNN network learns rich spatial features from the uploaded images, while the LSTM component models sequential dependencies among these features. The LSTM treats the features as parts of an ongoing sequence by interpreting them in relation to one another. This hybrid model was designed to jointly exploit spatial and temporal information, resulting in improved classification performance compared to standalone CNN models. Firstly, the CNN processes the uploaded images X through multiple convolutional/pooling layers to extract high-level features and maps. These maps encode spatial patterns, such as textures, shapes, and structural details, that are important for distinguishing venomous from non-venomous snakes. After the convolutional operations, the resulting feature maps are flattened or reshaped into a sequence of feature vectors suitable for temporal modelling.

Furthermore, the LSTM is then used to capture sequential relationships from the extracted features. The LSTM is a type of recurrent neural network (RNN) designed to mitigate the vanishing gradient problem by using gated memory units. At each step t, the LSTM uses the following equations to update its internal states:

$$\boldsymbol{ft = \sigma(Wf \cdot [ht-1, xt] + bf)} \tag{3}$$

$$\boldsymbol{it = \sigma(Wi \cdot [ht-1, xt] + bi)} \tag{4}$$

$$\boldsymbol{Ct = tanh(Wc \cdot [ht-1, xt] + bc)} \tag{5}$$

$$\mathbf{Ct = ft \odot Ct-1 + it \odot Ct} \tag{6}$$

$$\mathbf{ot} = \boldsymbol{\sigma}\,\mathbf{(Wo \cdot [ht-1, xt] + bo)} \tag{7}$$

$$\mathbf{ht = ot \odot tanh\,(Ct)} \tag{8}$$

where xt denotes the input feature vector at time step t, ht represents the hidden state, and Ct corresponds to the cell state. Meanwhile, ft, it, and ot refer to the forget, input, and output gates, respectively. The learnable parameters are represented by W as weight matrices and b as bias vectors. The sigmoid activation function is denoted by σ (·), while ⊙ indicates element-wise multiplication. By processing the CNN-extracted features as sequences, the LSTM network learns dependencies and contextual relationships from the feature vectors, thereby enhancing the model's distinctive power. Lastly, the LSTM output is passed to a fully connected layer, followed by a SoftMax activation function to generate class probabilities for venomous and non-venomous snake categories. The hybrid CNN+LSTM architecture combines the strengths of both networks by enabling effective spatial feature learning and sequential modelling, thereby enhancing robustness and generalization in snake classification tasks.

3.6. Architecture of the Hybrid Model.

The architectural pipeline of the proposed model was configured as a sequential pipeline to systematically transform raw uploaded snake pictures into high-level feature representations for final classification, as shown in Figure 3. The model's architecture comprises multiple interconnected pipelines: input acquisition, data preprocessing, convolutional feature extraction, temporal. feature

Table 1. Hyperparameter configuration and training setup.

Configuration Area	Setting / Value
Implementation	Python with a deep learning framework (TensorFlow / PyTorch)
Hardware	GPU-enabled environment (acceleration used to reduce training time)
Optimizer	Adam
Learning Rate	0.001
Loss Function	Categorical cross-entropy (suitable for binary probabilistic classification)
Number of Epochs	15 (with monitoring of training and validation performance)
Batch Size	32 (balance between memory, stability, and generalization)
Early Stopping	Based on validation loss – stops training when no further improvement is observed
Consistency	Same hyperparameters used across all experiments (ensures fair comparison & reproducibility)
Task Type	Binary image classification (venomous vs. non-venomous)

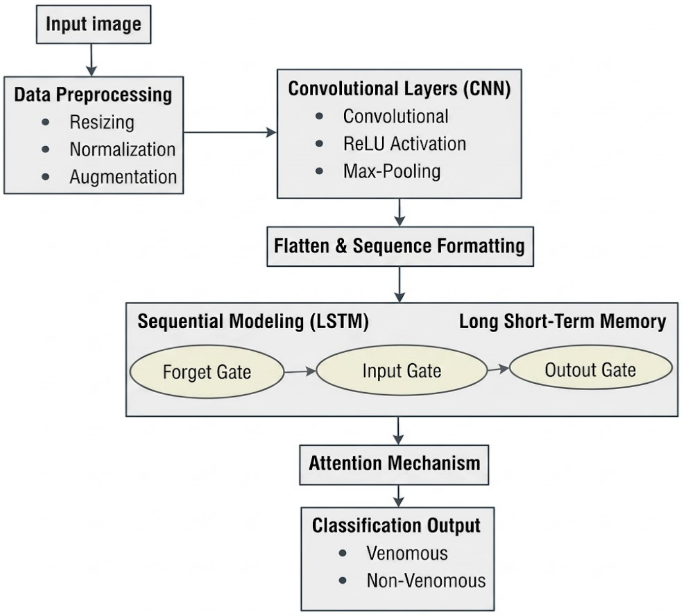

Figure 3. Architecture of the Hybrid Model.

modelling, attention-based refinement, and binary classification. This structural design was intended to enhance the effectiveness of machine learning for both spatial and sequential patterns in snake images. The process begins at the input layer, where preprocessed images are resized to fixed dimensions and fed into the pipeline. After the input process, the pictures are passed through a sequence of convolutional layers, each equipped with learnable filters that extract local spatial features such as edges, textures, and shape patterns. ReLU activation functions are used after convolutional layers to introduce nonlinearity and enhance feature discrimination.

Furthermore, a max-pooling layer is applied to reduce spatial dimensionality, minimize computational cost, and preserve the most relevant features. After convolution and pooling, the feature maps are flattened and restructured into a series of feature dimensions. This particular transformation enabled the extracted spatial features to be treated as sequential inputs, which are appropriate for processing by the LSTM network. The LSTM layer then captured long-term dependencies while simultaneously modelling contextual relationships among feature vectors by allowing the model to learn how spatial patterns interact across different regions of the image. To further enhance the robustness of the model's feature relevance, an attention mechanism was integrated after the LSTM layer. The attention mechanism can assign greater weights to more informative feature representations while simultaneously suppressing less relevant details. This technique sharpens the model's focus by using visually critical cues to distinguish venomous from non-venomous snakes.

The refined feature representations will then be passed to one or more fully connected (dense) layers, where they will combine the studied features through high-level reasoning. Finally, a SoftMax output layer produces high-probability scores for each class, enabling binary classification into venomous and non-venomous categories.

3.7. Experimental Setup

The experiments for the proposed hybrid model were implemented in Python using a deep learning framework and executed on a GPU-enabled environment to accelerate training and reduce computational time. GPU acceleration was vital given the network's depth and the large number of training images. Optimization was performed using the Adam optimizer with a learning rate of 0.001, a commonly adopted value for stable, efficient convergence. Adam was selected for its ability to adapt the learning rate and its robustness when training deep neural networks with complex architectures.

In classification tasks, the categorical cross-entropy loss function was used, as it is well-suited for binary image classification with probabilistic outputs. The loss function enables effective optimization by minimizing the difference between predicted class probabilities and actual class labels. The dataset was trained over 15 epochs, with both training and validation performance monitored. A batch size of 32 was used to balance memory efficiency and training stability, avoid overfitting, and ensure optimal generalization. An early stopping strategy based on validation loss automatically terminated training when no further improvement in validation loss was observed. All experiments used the same hyperparameter configuration, ensuring fair evaluation and reproducible results.

3.8. Evaluation Metrics

The model was evaluated using the following metrics:

1) **Accuracy:** Measures the model's overall correctness.

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \quad (9)$$

2) **Precision:** Indicates how many of the predicted venomous snakes were actually venomous.

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \quad (10)$$

3) **Recall:** Reflects how many of the actual venomous snakes were correctly identified.

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \quad (11)$$

4) **F1-Score:** The harmonic mean of precision and recall, providing a balanced metric.

$$F1 = 2 \times \frac{Precision\ x\ Recall}{Precision + \ Recall} \quad (12)$$

5) **ROC-AUC:** The Area Under the Receiver Operating Characteristic Curve, which evaluates the model's ability to separate classes. Formula for ROC-AUC:

$$AUC \ = \ \int 0\ (TPR)d(FPR)) \quad (13)$$

Where:

$$TPR = Recall = \frac{TP}{TP + FN} \quad (14)$$

$$FPR = \frac{FP}{FP + TN} \quad (15)$$

4. Results and Discussion

This chapter covers the experimental results of the Hybrid model. The evaluation focused on training and validation loss curves, training and validation accuracy plots, a classification report, a confusion matrix, and ROC-AUC analysis to assess the hybrid model's learning behavior and classification performance.

4.1. Training and Validation Loss

The training and validation loss curves in Figure 4 show a stable, effective learning process for the proposed model over 15 training epochs. At the beginning (Epoch 1), the model records a training loss of 0.6251 and a validation loss of 0.5251. As training progresses, both loss values show a consistent downward trend, indicating effective optimization. By Epoch 10, the training loss declines to about 0.2383, while the validation loss decreases to 0.1465, indicating improved stability. Slight fluctuations in the validation loss are observed at some intermediate epochs, but these variations are minor and do not indicate overfitting.

During epoch 14, the training loss reached a minimum of 0.0894, while the validation loss decreased to 0.0721. Close alignment between the training and validation loss curves, with only minor divergence, confirms that the model generalizes well to unseen data. This behavior indicates effective regularization and demonstrates the robustness and Reliability of the proposed hybrid model.

4.2. Training and Validation Accuracy

The training and validation accuracy curves confirm the model's effectiveness across 15 training epochs (Figure 5). At the first stage, the model attains training and validation accuracies of 66.25% and 84.03%, indicating that a discriminative feature is learned early in training. As training advances, both accuracy curves show a steady upward trend. At Epoch 10, accuracy improves to about 89.65% and validation accuracy to 93.97%, demonstrating stable learning and improved generalization. Minor fluctuations in accuracy are observed, but these variations are minimal and do not suggest performance degradation.

During the final training stage at (Epoch 15), the model achieves a training accuracy of 96.52% and a validation accuracy of 97.13%. The close alignment between the training and validation accuracy curves, with validation accuracy slightly higher than training accuracy, indicates the effectiveness of regularization and robust generalization. These trends confirm that the proposed model achieves reliable classification performance on both the training and unseen validation data.

4.3. Classification Report

Table 1 presents the detailed performance of the classification model, reporting precision, recall, F1-score, and support for both the non-venomous and venomous classes. The results demonstrate that the model achieves balanced performance across both classes, with a recall of 0.98 for the non-venomous class, indicating a strong ability to identify non-venomous species. A precision of 0.96 was achieved with a low false-positive rate. The venomous class achieved 0.98 precision and 0.96 recall, indicating effective detection of venomous specimens with minimal misclassification. The model also achieved 97% overall accuracy, and its macro and weighted-average scores are equal, confirming that it performs well across all classes with no bias. These results demonstrate the hybrid model's strength and Reliability for accurate classification.

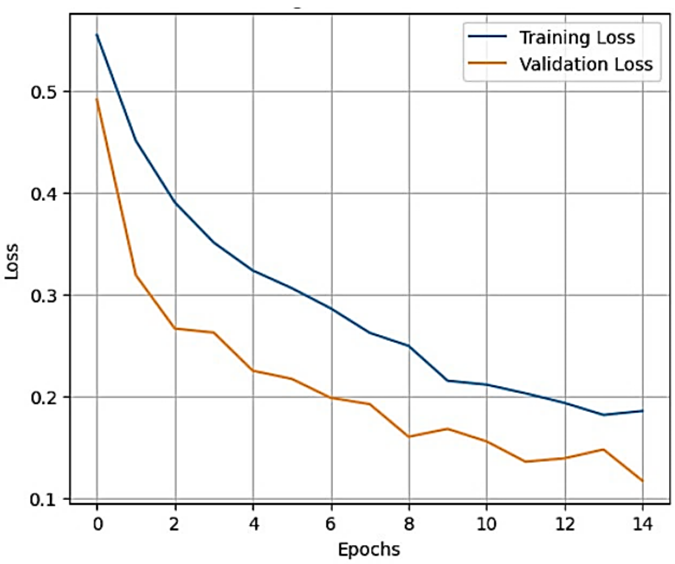

Figure 4. Training and Validation Loss Curve.

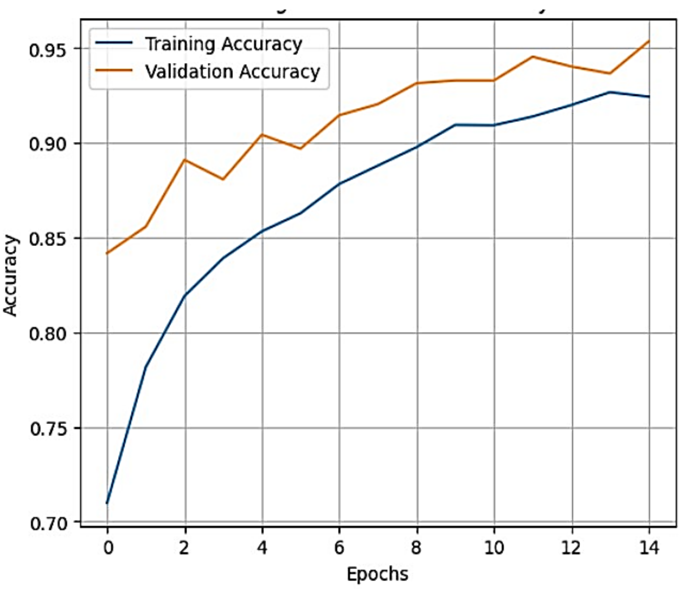

Figure 5. Training and Validation Accuracy.

Table 1. Classification Report of the Proposed Model.

Class	Precision	Recall	F1-score	Support
Non-venomous	0.96	0.98	0.97	674
Venomous	0.98	0.96	0.97	685
Accuracy	0.98	0.96	0.97	1359
Macro Avg	0.97	0.97	0.97	1359
Weighted Avg	0.97	0.97	0.97	1359

Figure 6 shows the classification model's performance on the training dataset. The results illustrate the model's precision, recall, and F1 scores for identifying non-venomous and venomous species. Precision of 0.96, recall of 0.98, and F1-score of 0.97 were recorded for the non-venomous class. For the venomous class, precision was 0.98, recall was 0.96, and F1-score was 0.97. These results show that the model effectively curbs both false positives and false negatives.

Figure 7 shows the model's evaluation metrics. These results indicate an accuracy of 0.97, with 0.97 for both macro average and weighted average scores. The near alignment between macro and weighted averages indicates that model consistency across classes is not affected by class imbalance. Thus, results confirm the Reliability of the proposed model for the venomous classification tasks.

4.4. Confusion Matrix

The confusion matrix in Figure 8 explains the model's classification performance on 1,359 test samples. The model has correctly identified 662 non-venomous snakes and 658 venomous snakes. The 12 non-venomous samples were misclassified as venomous (false positives), and 27 venomous samples were misclassified as non-venomous (false negatives). The 12 and 27 error counts appear modest relative to the total sample size used.

Critical assessment of false negatives:

The 27 false negatives represent venomous snakes that the model would label as harmless. This yields a false negative rate (FN rate) of:

$$\frac{27}{658 + 27} = 3.94\% \ (\mathit{approximately\ 1\ in\ 25\ venomous\ snakes}) \qquad (16)$$

In settings such as a mobile identification app for rural clinics, a wildlife rescue tool, or a herpetological field guide, a 4% false negative rate is not "very low" from a safety perspective. A user who relies on the model's output could mistakenly handle or approach a dangerous snake, or a bite victim could delay seeking antivenom treatment. Such errors can lead to severe injury or death. Likewise, for safety-critical applications, the acceptable false-negative rate is usually driven as close to zero as technically feasible, or supplemented by mandatory human verification. Therefore, the model's current FN rate of 4% is viewed as a limitation rather than an indicator of Reliability and robustness.

The 12 false positives (1.78% of non-venomous samples) are less consequential. A model predicting a harmless snake as venomous might lead to caution, such as the killing of non-threatened species or the wasting of resources, but it rarely causes harm. Thus, the model demonstrates strong discriminative ability, with moderate precision, recall, and F1-scores for both classes. In practice, a 4% chance of misclassifying a venomous snake as harmless makes the model suitable for standalone deployment in public health or first-response scenarios. Future work must focus on driving false negatives toward zero, even at the cost of increased false positives.

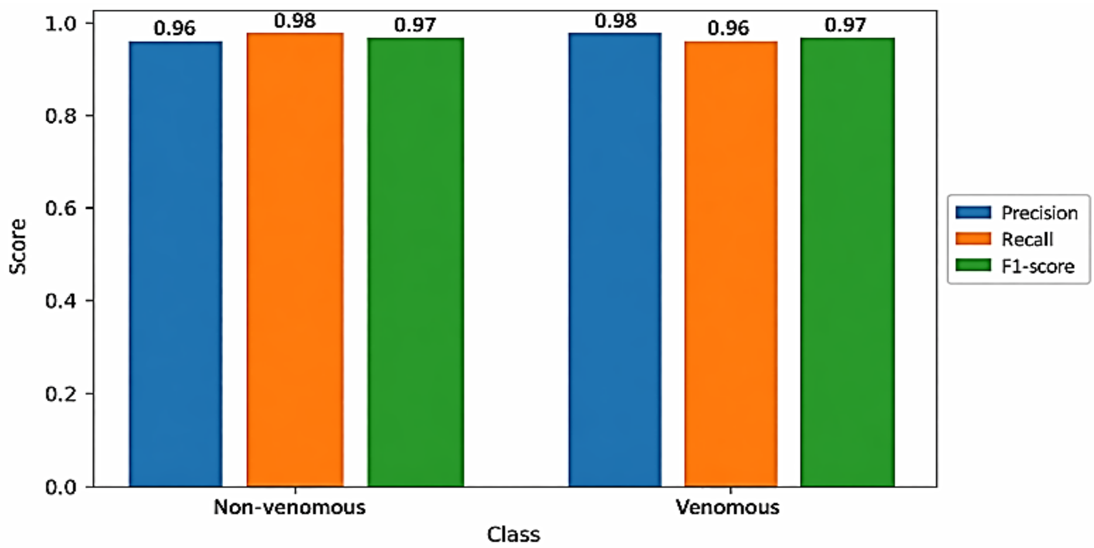

Figure 6. Classification Performance of the Proposed Model for Non-Venomous and Venomous Classes in Terms of Precision, Recall, And F1-Score.

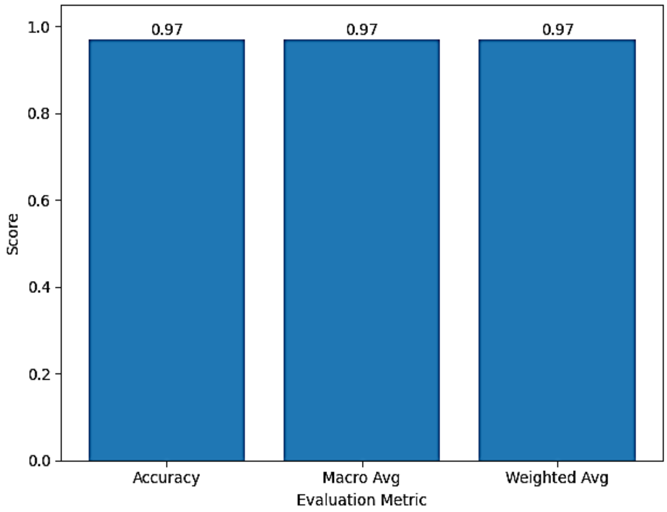

Figure 7. Performance Evaluation of the Proposed Model in Terms of Accuracy, Macro Average, And Weighted Average.

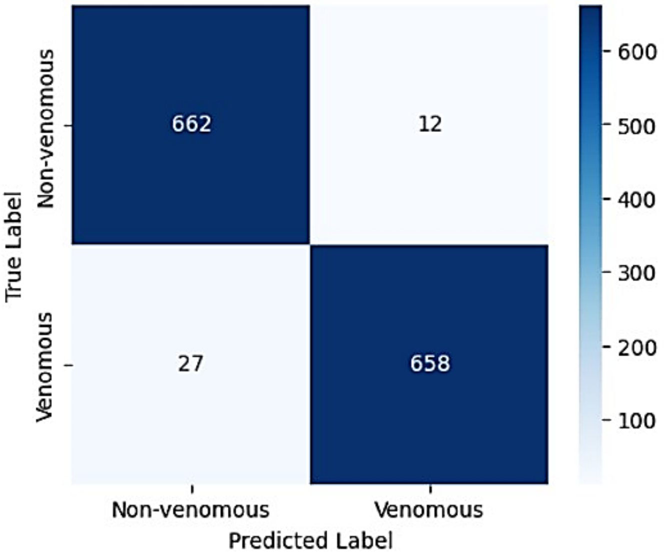

Figure 8. Confusion Matrix.

4.5. ROC-AUC Analysis

Receiver Operating Characteristic (ROC) curves were used to evaluate the proposed model's discrimination across multiple classification thresholds. The ROC plot shows the true positive rate (TPR) against the false positive rate (FPR), revealing the trade-off between sensitivity and specificity. The proposed hybrid model achieves an ROC Curve (AUC) of 0.97, which is close to 1.0 and indicates excellent classification performance, indicating strong potential to distinguish between the two.

The high ROC-AUC value indicates that the model maintains a very robust predictive score across a wide range of decision thresholds. This result further validates the robustness of the proposed technique and its high accuracy, precision, recall, and F1-score, as shown by the classification report and confusion matrix. The overall ROC AUC analysis confirms the Reliability of the proposed model, which is well-suited for real-world classification tasks where accurate discrimination is vital.

The comparative analysis between existing approaches and the proposed hybrid model has been presented in Table 2 based on the dataset's size used for the comparison, the technique used for the model, and key performance metrics for evaluation, as well as accuracy, precision, recall, F1-score, and the number of training cycles (epochs) used for the training. In high-risk regions, even a 1% improvement in classification precision can have a significant impact, potentially avoiding dozens of misclassifications each year. Given that false negatives (venomous snakes misclassified as non-venomous) can lead to improper or delayed antivenom administration, this improvement could directly reduce morbidity and mortality associated with snakebites. For example, with an estimated 5.4 million snakebites globally each year and fatality rates ranging from 1.8 to 2.7 million envenoming cases [1], enhancing model accuracy and precision could

Table 2. Comparative analysis between the existing models and the proposed CNN+LSTM hybrid model.

Author & Year	**Dataset Used**	**Technique Used**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1Score (%)**	**Epochs**
Progga et at,(2021)	1,766 images from Kaggle	CNN	90%	Not Recorded	Not Recorded	Not Recorded	100
Rajabizadeh et al., (2021)	594 images of six snake species (venomous, nonvenomous, semi-venomous)	LDA + SVM	84%	Not Recorded	Not Recorded	Not Recorded	150
Rajabizadeh et al., (2021)	594 images of six snake species (venomous, nonvenomous, semi-venomous)	MobileNetV2	92%	92.26	89.10	92.30	150
Joshi et al., (2024)	1,766 images	MobileNetV2	88%	Not Recorded	Not Recorded	Not Recorded	Not Recorded
Ahmed et al., (2024)	82,601 images of 45 species from Aircrowd	Salient Object Detection + VGG16	97%	Not Recorded	Not Recorded	Not Recorded	35
Anwarul & Tanwar (2025)	6,798 images (Google + augmented images)	Modified DenseNet169 + augmentation	97%	97.37	96.06	97.35	10
Proposed Model	**6,798 from Anwarul & Tanwar (2025).**	**CNN+LSTM hybrid model**	**97%**	**97.10%**	**96.55%**	**97.12%**	**15**

help save hundreds or thousands of lives annually, depending on the scale of deployment and local incidence rates. This direct human impact underscores the practical importance of striving for even marginal performance gains in automated snake classification models, especially in healthcare and emergency response contexts.

4.6. Comparison of Existing Work and Proposed Model

The comparative analysis between existing approaches and the proposed hybrid model has been presented in Table 2 based on the dataset's size used for the comparison, the technique used for the model, and key performance metrics for evaluation, as well as accuracy, precision, recall, F1-score, and the number of training cycles (epochs) used for the training.

The comparison demonstrates that the proposed model has established a competitive benchmark by incorporating state-of-the-art methods with a relatively small dataset and fewer training epochs than most earlier models. The proposed model achieved an accuracy of 97%, precision of 97%, recall of 96%, F1-score of 97%, and a ROC–AUC of 0.97 with only 15 epochs of training. In comparison, [6] reported 90% accuracy with a CNN trained for 100 epochs; [7] achieved 84.02% accuracy with LDA+SVM and 92.26% with MobileNetV2, both trained for 150 epochs. [8] achieved 88.27% accuracy with MobileNetV2, while [9] reported 97% accuracy with Salient Object Detection using VGG16 over 35 epochs. The model proposed by [10] achieved slightly higher performance with 97.37% accuracy, 97.37% precision, 96.06% recall, and an F1-score of 97.35 using a modified DenseNet169 trained for only 10 epochs on an augmented dataset. Primarily focused on accuracy, the proposed model provides a more comprehensive evaluation metric, including precision, recall, F1 Score, and ROC AUC, demonstrating a well-balanced, dependable classification model. The incorporation of LSTM enhances effective feature learning, improving generalization compared to conventional CNN-based models. Figure 9 compares classification accuracy between the existing models and the proposed hybrid model. The proposed model achieved an accuracy of 97%, which is competitive with recent DL approaches and earlier techniques.

Figure 10 illustrates the precision comparison of the selected existing approaches against the proposed hybrid model, indicating that the proposed model has achieved moderate, balanced precision compared **to** earlier modelling methods. Figure 11 compares recall scores against selected existing approaches and the proposed hybrid model, demonstrating the robustness of the proposed model in properly identifying venomous and non-venomous snake classes. Figure 12 shows the F1-score for the proposed CNN+LSTM model compared with existing models, highlighting well-balanced performance across precision and recall metrics.

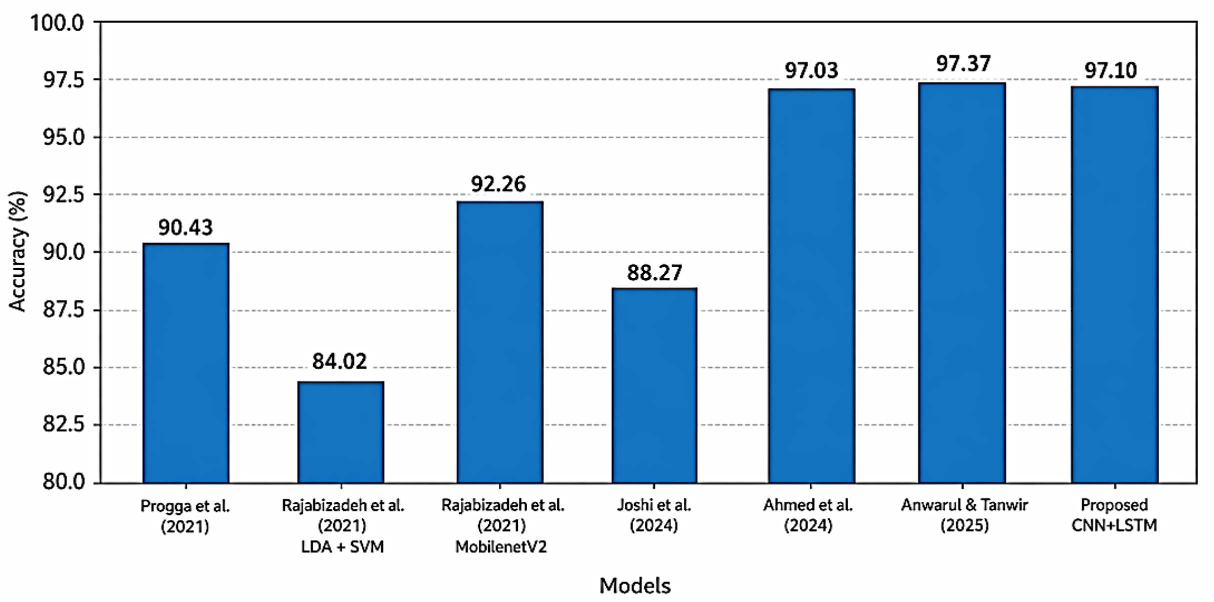

Figure 9. Comparison of Accuracy Across Existing Works and Proposed Model.

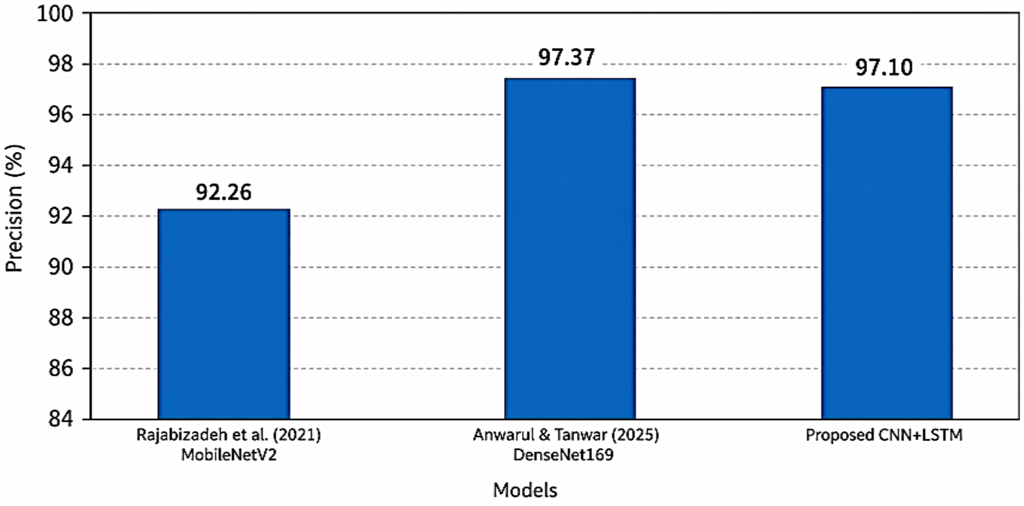

Figure 10. Comparison of precision Across Existing Works and the proposed model.

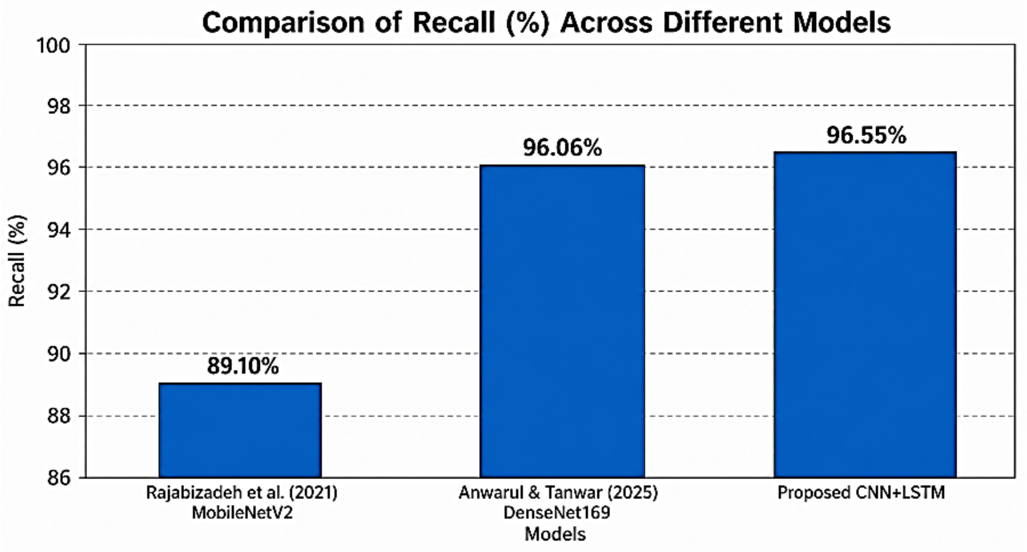

Figure 11. Comparison of recall Across Existing Works and the proposed model.

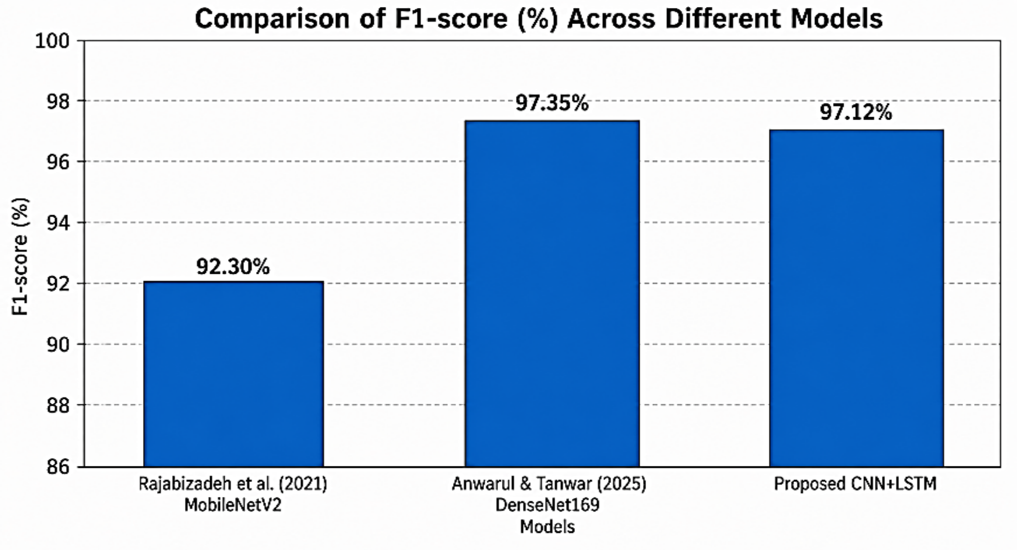

Figure 12. Comparison of f1-score Across Existing Works and the proposed model.

4.7. Discussion

This study investigated the effectiveness of a CNN-LSTM hybrid deep learning model for classifying and identifying venomous and non-venomous snake species. The experiment showed that the proposed hybrid model performs robustly across all evaluation metrics, confirming its suitability for automated snake classification from images. The training and validation loss curves indicate a consistent, stable decline throughout training, demonstrating effective learning and model integration. The lack of a significant gap between the training and validation losses suggests that overfitting is well controlled, due to the hybrid architecture and appropriate training strategy. Steady improvements in the training and validation accuracy curves reflect the model's ability to learn increasingly discriminative features and generalize to unseen data.

The classification report further validated the proposed model's efficiency. Both venomous and non-venomous classes achieved an accuracy, precision, recall, and F1 scores of about 97%, indicating balanced performance and minimal classification bias. This balance is vital in critical applications, where misclassifying venomous snakes can have serious consequences. The confusion matrix analysis confirmed the model's dependability, with very few misclassifications, especially in true positives and true negatives. These results validate that the model can precisely distinguish between the two species, even in sensitive cases. The ROC-AUC value of 0.97 indicates excellent classification performance between the venomous and non-venomous classes across varying decision thresholds, further emphasizing the robustness of the proposed model.

Compared with existing studies, the proposed hybrid model has achieved competitive, and in some areas superior, performance with a moderate dataset size and fewer training epochs than most recent deep learning-based approaches. While previous work primarily focuses on accuracy, our study provides a range of evaluation metrics, including precision, recall, F1-score, and ROC-AUC, to provide a more reliable assessment of model performance. The integration of the LSTM enables effective modelling of sequential feature dependencies, and the attention mechanism helps enhance feature relevance, leading to improved generalization over conventional CNN standalone models.

5. Conclusion and Future Work

The framework was designed to automatically classify snake species from images into two classes: venomous and non-venomous. Integrating the CNN and LSTM networks enhances the model's ability to capture both spatial and temporal features in images. The experimental evaluation demonstrated robust classification performance, achieving an accuracy of 97%, a precision of 97.10%, a recall of 96.55%, an F1-score of 97.12%, and a significant ROC-AUC value, indicating excellent discriminative capability. The training and validation results also demonstrated stable learning, with the accuracy and loss curves closely aligned, suggesting effective generalization and minimal overfitting. The confusion matrix further validates the model's robustness, with only a few misclassifications in both species. These findings show that integrating CNN-based feature extraction with LSTM-based sequence modelling can significantly enhance classification Reliability, particularly in visually complex image recognition tasks such as snake identification.

Despite the model's strengths, several limitations remain. The relatively small amount of data can limit the model's generalizability to a broader range of snake species in more complex environmental settings. The current model is also limited to binary classification, while multiclass species identification could enhance its practical application. These limitations highlight directions for future research. Although the study was initially intended to use data collected from KSBH (Kaltungo Snakebite Hospital), due to time constraints and unforeseen circumstances, pictures from [10] were used. Future work will incorporate a more diverse, locally acquired dataset from additional locations, particularly Kaltungo LGA in Gombe State, Nigeria, to improve generalization and classification. Expanding the dataset will enhance the model's robustness, contextual relevance, and ecological variability, supporting improved local adaptability, clinical Reliability, and effective application in healthcare and community settings in Kaltungo.

Compared with existing studies, the proposed model achieves competitive performance with a relatively small dataset and fewer training epochs than most previous deep learning approaches. Unlike earlier studies that primarily focused on accuracy, this study uses a comprehensive evaluation metric that combines multiple performance metrics to ensure a balanced assessment of predictive capabilities. This evaluation has enhanced the credibility and practical implications of the proposed approach. The findings show that the hybrid model provides a reliable and efficient solution for combating snake misclassification. The system has broad potential for enhancing public safety, wildlife management, and emergency medical response by enabling rapid, accurate identification of venomous snake species. This study contributes to expanding the body of knowledge on deep learning-based networks and wildlife management, and provides a blueprint for future research on automated snake species classification.

6. Declarations

6.1. Author Contributions

Yakubu Abubakar Lidani: Conceptualization, methodology, software, validation, formal analysis, writing- original Draft, Data curation**; Abdullahi Musa Yola:** Review and editing, Visualization, Supervision, project administration; **Abu Tasiu:** Formal analysis, investigation Recourses; **Nura Muhammad Sani:** Conceptualization, methodology, software, validation, formal analysis, writing- original Draft, Data curation**; Sulaiman Muhammad Gidado:** Formal analysis, investigation Recourses.

6.2. Institutional Review Board Statement

Not applicable.

6.3. Informed Consent Statement

Not applicable.

6.4. Data Availability Statement

The data presented in the study is available from the original author Anwarul & Tanwar (2025).

6.5. Acknowledgment

Not applicable.

6.6. Conflicts of Interest

The author declares no conflicts of interest.

7. References

[1] World Health Organization, “Snakebite envenoming,” Fact sheet, Geneva, Switzerland: WHO, Sep. 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/snakebite-envenoming.

[2] World Health Organization, “Control of Neglected Tropical Diseases,” 2023, [Online]. Available: https://www.who.int/teams/control-of-neglected-tropical-diseases/snakebite-envenoming/prevalence. Accessed: Sep. 10, 2024.

[3] M. J. Mazerolle, L. L. Bailey, W. L. Kendall, J. A. Royle, S. J. Converse, and J. D. Nichols, “Making great leaps forward: accounting for detectability in herpetological field studies” *Journal of Herpetology*, vol. 41, no. 4, pp. 672-689, 2007. https://www.jstor.org/stable/40060463.

[4] Walk Through India, “India’s 10 Beautiful but Deadliest Venomous Snakes,” 2021. [Online]. Available: https://www.walkthroughindia.com/wildlife/indias-10-beautiful-but-deadliest-venomous-snakes. Accessed: Sep. 15, 2024.

[5] I. Bolon, L. Picek, A. M. Durso, G. Alcoba, F. Chappuis, and R. Ruiz de Castañeda, "An artificial intelligence model to identify snakes from across the world: Opportunities and challenges for global health and herpetology," *PLoS Neglected Tropical Diseases*, vol. 16, no. 8, p. e0010647, 2022. https://doi.org/10.1371/journal.pntd.0010647.

[6] N. I. Progga, N. Rezoana, M. S. Hossain, R. U. Islam, and K. Andersson, "A CNN based model for venomous and non-venomous snake classification," in *Applied Intelligence and Informatics: Proc. AII 2021*, Nottingham, U.K., Jul. 2021, pp. 216–231. https://doi.org/10.1007/978-3-030-82269-9_17.

[7] M. Rajabizadeh and M. Rezghi, "A comparative study on image-based snake identification using machine learning," *Scientific Reports*, vol. 11, no. 1, p. art. No. 19142, 2021. https://doi.org/10.1038/s41598-021-96031-1.

[8] P. Joshi, S. Sati, T. Choudhury, T. Bajaj, K. Kotecha, A. Sar, "Classification of Venomous and Non-venomous Snakes Using Transfer Learning with MobileNetV2." in *International Conference on Universal Threats in Expert Applications and Solutions*, vol 1006, pp 427–438, 2024. https://doi.org/10.1007/978-981-97-3810-6_35.

[9] K. Ahmed, M. A. Gad, and A. E. Aboutabl, "Snake species classification using deep learning techniques," *Multimedia Tools and Applications*, vol. 83, no. 12, pp. 35117–35158, 2024. https://doi.org/10.1007/s11042-023-16773-0.

[10] S. Anwarul and R. Tanwar, "Venomify: Automated classification of venomous and non-venomous snake species using deep learning," *Procedia Computer Science*, vol. 259, pp. 219–229, 2025. https://doi.org/10.1016/j.procs.2025.03.323.

[11] S. Anwarul, T. Misra, and D. Srivastava, "An IoT & AI-assisted framework for agriculture automation" *In 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2022. https://doi.org/10.1109/ICRITO56286.2022.9964567.

[12] A. P. James, B. Mathews, S. Sugathan, and D. K. Raveendran, "Discriminative histogram taxonomy features for snake species identification." *Human-centric Computing and Information Sciences,* vol. 4, no. 1, pp. 1–11, 2014. https://doi.org/10.1186/s13673-014-0003-0.

[13] V. Patel, A. Chudasama, and A. Thakkar, "YOLOv11 Based Classification of Lumbar Spine Degenerative Changes Across Multi-Modal Imaging." in *AIJR Proceedings*, vol. 7, no. 6, pp. 7–18, 2025. https://doi.org/10.21467/proceedings.7.6.2

[14] S. B. Islam, D. Valles, and M. R. J. Forstner, "Herpetofauna Species Classification from Images with Deep Neural Network." *In 2020 Intermountain Engineering, Technology and Computing (IETC),* pp. 1-6, 2020. https://doi.org/10.1109/IETC47856.2020.9249141.

[15] I. S. Abdurrazaq, S. Suyanto, and D. Q. Utama, "Image-based classification of snake species using convolutional neural network." *In 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI),* pp. 97–102. IEEE, 2019. https://doi.org/10.1109/ISRITI48646.2019.9034633.

[16] R. Kamalraj, "Deep learning model for identifying snakes by using snakes' bite marks" *In 2020 International Conference on Computer Communication and Informatics (ICCCI),* pp. 1-4. IEEE, 2020. https://doi.org/10.1109/ICCCI48352.2020.9104200.

[17] L. Jeantet and E. Dufourq, "Improving deep learning acoustic classifiers with contextual information for wildlife monitoring," *Ecological Informatics*, vol. 77, 102256, 2023. https://doi.org/10.1016/j.ecoinf.2023.102256.

[18] D. Lakshmi, R. C. Panda, Amrita, and A. Prakash, "Life-Saving APP: Snake Classification 'Venomous and Nonvenomous' Using fast. ai Based on Indian Species" *In European, Asian, Middle Eastern, North African Conference on Management & Information Systems*, pp. 109-115, 2021. https://doi.org/10.1007/978-3-030-77246-8_11.

[19] S. B. Abayaratne, W. M. K. S. Ilmini, and T. G. I. Fernando, "Automated methods to identify snake species in Sri Lanka: A review," *Scholar Bank Library, University of Sri Jayewardenepura,* 2019. http://dr.lib.sjp.ac.lk/handle/123456789/12078.

[20] S. Anwarul, "An efficient minimum spanning tree-based color image segmentation approach," in *International Advanced Computing Conference*, Cham, Switzerland: Springer International Publishing, 2021, pp. 588–598. https://doi.org/10.1007/978-3-030-95502-1_44.

[21] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, 2019. https://doi.org/10.1186/s40537-019-0197-0.

[22] L. Arcila-Díaz, H. Mejía-Cabrera, and J. Arcila-Díaz, "Estimation of mango fruit production using image analysis and machine learning algorithms," *Informatics*, vol. 11, no. 4, p. 87, 2024. https://doi.org/10.3390/informatics11040087.