iJRASET
International Journal For Research in
Applied Science and Engineering Technology


INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: VII Month of publication: July 2025

DOI: https://doi.org/10.22214/ijraset.2025.73234

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 13 Issue VII July 2025- Available at www.ijraset.com*

# Hybrid Deep Learning Models for Gait Recognition: A Comparative Analysis of CNN, CNN-LSTM, and HOA Techniques

Atansuyi N.[1], Ajaegbu C.[2], Adekola O.[3], Akande O.[4]
[1]*Department of Computer Science, School of Computing, Babcock University, Illishan Remo, Ogun State, Nigeria*
[2]*Department of Information Technology, School of Computing Babcock University, Illishan Remo, Ogun State, Nigeria*
[3]*Department of Systems Engineering, School of Computing, Babcock University, Illishan Remo, Ogun State, Nigeria*
[4]*Department of Computer Science, School of Computing, Babcock University, Illishan Remo, Ogun State, Nigeria*

***Abstract: Gait recognition is a critical biometric technique with applications in surveillance, healthcare, and security. This study proposes a hybrid deep learning framework combining Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and the Hippopotamus Optimization Algorithm (HOA) for robust gait recognition. By leveraging spatial feature extraction, temporal dynamics, and metaheuristic hyperparameter optimization, the proposed HOA-CNN-LSTM model achieves superior performance. Experimental results on the TUM-GAID dataset show that the hybrid model outperforms standalone CNN and CNN-LSTM approaches in accuracy, processing time, and error rates. The findings suggest that HOA-optimized architectures provide scalable and efficient solutions for gait recognition tasks in real-world settings.***

***Keywords: Gait Recognition, Convolutional Neural Network, Long Short-Term Memory, Hippopotamus Optimization Algorithm, Deep Learning.***

## I. INTRODUCTION

Biometric recognition systems have become integral to modern security and authentication frameworks, leveraging unique physiological and behavioral traits to identify individuals with high accuracy. Among the various biometric modalities, such as fingerprints, facial features, iris patterns, and voice recognition; gait recognition has emerged as a compelling alternative due to its non-intrusive and remote identification capabilities. Gait recognition analyzes the walking patterns of individuals, enabling surveillance and access control in situations where direct cooperation or close-range biometric data capture is impractical [1].

The increasing applicability of gait recognition in diverse domains such as surveillance, smart healthcare, human-computer interaction, and robotics arose stems from its contactless nature and the ability to operate under non-cooperative conditions [2]. Unlike traditional biometric systems that require physical interaction or close-up images, gait can be captured unobtrusively from a distance, even in crowded or uncontrolled environments. However, the design of accurate gait recognition systems remains challenging due to the variability introduced by changes in clothing, footwear, carrying conditions, walking speed, and view angles [3]. Earlier methods in gait analysis relied heavily on handcrafted features and model-based techniques, which often performed inconsistently across varying environments [4]. With the advent of deep learning, Convolutional Neural Networks (CNNs) have demonstrated remarkable capabilities in capturing spatial information from gait silhouettes, such as body posture, limb movement, and silhouette contours [5]. Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, have been employed to capture temporal dependencies in gait sequences, modeling dynamic traits such as stride length, walking rhythm, and leg movement [6]. Hybrid models that combine CNNs and LSTMs have gained attention for their ability to simultaneously capture spatial and temporal features of gait. CNNs effectively extract frame-wise spatial representations, while LSTMs capture the sequential progression of gait patterns over time [7]. Such hybrid CNN-LSTM architectures have shown improved recognition accuracy compared to standalone CNN or LSTM models. However, their performance is often limited by suboptimal hyperparameter configurations and the complexity of the model training process [8]. Optimization algorithms play a pivotal role in tuning deep learning models. Traditional optimizers such as Stochastic Gradient Descent (SGD) and Adam, while widely used, may suffer from issues like slow convergence and susceptibility to local minima in complex loss landscapes [9]. Metaheuristic optimization techniques, inspired by natural behaviors, offer an alternative by enhancing global search capabilities and robustness. Algorithms such as Genetic Algorithms, Particle Swarm Optimization, and Ant Colony Optimization have been successfully applied to deep learning to improve performance and generalization [10].

A relatively recent entrant in the family of metaheuristic algorithms is the Hippopotamus Optimization Algorithm (HOA), inspired by the social and territorial behaviors of hippopotamuses. HOA employs adaptive strategies to balance exploration and exploitation in high-dimensional search spaces and has shown promise in neural network optimization tasks [11]. Unlike traditional gradient-based optimizers, HOA's population-based search approach allows it to overcome non-convex optimization challenges, making it suitable for tuning complex deep learning architectures [12]. In the context of gait recognition, the application of HOA remains largely unexplored. This research investigates the integration of HOA with CNN-LSTM architectures to optimize hyperparameters such as learning rates, dropout rates, and layer configurations. The proposed approach aims to enhance convergence efficiency, increase recognition accuracy, and improve model robustness under varying gait conditions [13]. The model was trained in TUM-GAID dataset, and evaluating using metrics such as recognition accuracy, False Acceptance Rate (FAR), False Rejection Rate (FRR), and Genuine Acceptance Rate (GAR), this study seeks to demonstrate the efficacy of HOA in boosting gait recognition performance [14]. Ultimately, this paper presents a comparative analysis of CNN, CNN-LSTM, and HOA-optimized CNN-LSTM models. The findings are to contribute to the development of robust, scalable, and high-performing gait recognition systems, applicable to real-world scenarios across surveillance, healthcare, and human-computer interaction domains.

## II. LITERATURE REVIEW

Gait recognition has evolved as a contactless biometric modality capable of identifying individuals based on walking patterns, making it highly suitable for surveillance, healthcare, and human-computer interaction applications. Traditional approaches, which relied on handcrafted features and statistical models, struggled with variations in walking conditions. Of recent, deep learning-based techniques, particularly those incorporating CNNs, LSTMs, and metaheuristic optimization have dominated the research landscape due to their robustness and adaptability. This section provides an in-depth review of the relevant literature, structured across major deep learning paradigms used for gait recognition.

### *A. CNN-Based Gait Recognition Models*

CNNs have become foundational in gait recognition, mainly for their ability to extract spatial features from gait silhouettes. In a recent study, [15] proposed a lightweight CNN suitable for edge devices. It showed promising results on real-time surveillance datasets while maintaining computational efficiency. [16] enhanced CNN performance by integrating a multi-scale feature extraction mechanism, enabling the model to retain fine-grained spatial patterns across varying clothing and view conditions. Similarly, [17] utilized a Residual CNN architecture to learn hierarchical spatial features from gait energy images (GEIs), significantly improving performance on the CASIA-B dataset. However, while CNNs excel at modeling spatial cues such as body contour and posture, they fall short in capturing the sequential dependencies inherent in gait, necessitating the use of temporal models such as LSTMs.

### *B. LSTM-Based Gait Recognition Models*

Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, have proven effective for capturing long-term temporal dependencies in gait data. LSTM networks are well-suited to model periodicity, stride cycles, and variations in walking patterns over time. In a 2022 study, [18] applied a standalone LSTM model on gait silhouette sequences to capture walking rhythm and speed variations. The model performed favorably on sequence-based recognition tasks but lacked the spatial extraction capacity of CNNs. [19] proposed an LSTM-based temporal encoder for gait cycle segmentation, which improved recognition accuracy in scenarios involving irregular walking speeds. Similarly, [20] investigated the use of bidirectional LSTMs (Bi-LSTMs), which process sequences in both forward and backward directions, enhancing the model's sensitivity to contextual temporal information. Despite these successes, LSTM-only models often underperform in conditions requiring detailed spatial reasoning, leading to increased interest in hybrid architectures that integrate both CNN and LSTM components.

### *C. Hybrid CNN-LSTM Architectures*

Hybrid CNN-LSTM models have emerged as a powerful solution by combining the spatial learning capabilities of CNNs with the temporal modeling strengths of LSTMs. In this approach, CNNs are used to extract frame-level features from gait sequences, which are then passed to LSTMs for temporal analysis. [21] demonstrated the effectiveness of such hybrid models on the OU-ISIR and CASIA-B datasets, where the CNN layers learned silhouette features, and the LSTM layers captured sequence dynamics. Their model showed improved accuracy across varying conditions including view angle and clothing changes.

Further advancements were introduced by [22], who incorporated attention mechanisms into CNN-LSTM models. This allowed the network to focus on informative temporal segments, resulting in robust cross-view gait recognition. In a 2024 study, [23] extended the architecture with residual connections and dropout regularization, which reduced overfitting and improved generalization. These hybrid models have consistently outperformed traditional CNN or LSTM-only models; however, their performance remains sensitive to the choice of hyperparameters such as learning rate, dropout ratio, and number of LSTM units.

### *D. Metaheuristic Optimization in Deep Learning*

Traditional optimizers like SGD and Adam are widely used for training deep neural networks but often struggle with issues such as slow convergence and suboptimal minima in complex loss landscapes. Recent research has explored metaheuristic optimization algorithms inspired by biological and natural processes as alternatives for hyperparameter tuning and model optimization. [24] utilized Genetic Algorithms (GA) to optimize CNN architectures for biometric classification, reporting improved convergence speed and accuracy. [25] employed Particle Swarm Optimization (PSO) to fine-tune gait recognition models, significantly improving robustness under noisy conditions. These approaches showcase the potential of metaheuristics to navigate non-convex solution spaces and enhance model training efficiency and generalization.

### *E. The Hippopotamus Optimization Algorithm (HOA)*

A novel and promising addition to the metaheuristic family is the Hippopotamus Optimization Algorithm (HOA), inspired by the social and territorial behavior of hippopotamuses. HOA has demonstrated strong performance in optimizing neural network parameters, offering improved balance between exploration and exploitation. [26] applied HOA to optimize CNNs for ECG-based classification tasks and found that it outperformed traditional optimizers in terms of accuracy and convergence. Additionally, [27] used HOA for optimizing attention weights in gait recognition systems, leading to improved contextual modeling in long gait sequences. These studies highlight HOA's potential as an effective and adaptable optimizer for complex biometric recognition tasks.

### *F. Dataset Evaluation*

Commonly used datasets include CASIA-B, OU-ISIR, and TUM-GAID, each offering varied gait sequences under multiple conditions. [28] conducted a comprehensive evaluation of CNN, LSTM, and hybrid models on the TUM-GAID dataset. Their analysis confirmed that hybrid CNN-LSTM models consistently outperform standalone architectures in recognition rate, especially in cross-view scenarios. The study also emphasized the need for robust optimizers to handle high-dimensional parameter spaces efficiently.

## III. METHODOLOGY

The methodology for Hippopotamus Optimization Algorithm (HOA) for the Hybridized Convolutional Neural Network - Long Short-Term Memory (CNN-LSTM) model is a critical step in enhancing gait recognition efficiency. The video dataset is preprocessed in order to enhance the quality of input data and thereafter HOA is initialize and integrated into the model to optimize key hyperparameters, improve convergence, and maximize classification accuracy while minimizing computational overhead. The implementation begins with the initialization of hyperparameters, where the HOA is applied to fine-tune essential model parameters such as learning rate, batch size, number of CNN filters, kernel sizes, LSTM units, and dropout rates. These hyperparameters are initialized within predefined search spaces, ensuring optimal values are selected during the training process. The HOA iteratively refines these parameters using its adaptive exploration and exploitation mechanisms, which mimic the foraging behavior of hippopotamuses in natural environments.

During the training phase, the CNN component extracts spatial features from gait sequences, while the LSTM processes temporal dependencies to learn the dynamic gait patterns. The HOA continuously evaluates model performance by measuring accuracy and loss at each iteration, dynamically adjusting hyperparameters to achieve optimal feature learning. The optimization process enhances generalization by preventing overfitting and ensuring robustness across different gait variations. The implementation of HOA also includes constraints to balance computational efficiency and model accuracy. To achieve this, the algorithm reduces redundant computations by selecting the most effective feature representations, thereby improving training speed and reducing resource consumption. Additionally, the HOA optimizes the number of epochs required for convergence, ensuring that the model attains peak performance without excessive computational costs. The final model, optimized using HOA, demonstrates superior recognition performance by improving accuracy, reducing false positives and false negatives, and ensuring reliable gait-based identity verification.

The integration of HOA with the hybrid CNN-LSTM model results in a highly efficient and scalable gait recognition system that can be deployed in real-world biometric authentication applications. The methodology comprises four core processes: data preprocessing, feature extraction, temporal sequence modeling, and hyperparameter optimization. Each phase plays a critical role in constructing a robust and efficient gait recognition model.

### *A. Data Preprocessing*

The TUM-GAID dataset was used due to its diversity in walking speeds, clothing, and viewpoints. Initially, video sequences were decomposed into individual frames. Background subtraction was conducted using Gaussian Mixture Models (GMM) to isolate gait silhouettes, which were then normalized to a resolution of 128×88 pixels. To enhance generalization, data augmentation techniques such as horizontal flipping, rotation, brightness scaling, and noise injection were applied. Dynamic Time Warping (DTW) ensured alignment of gait cycles, thereby improving temporal coherence across samples [29].

### *B. Feature Extraction Using CNN (ResNet)*

Spatial features were extracted using a customized Residual Network (ResNet-50) architecture. ResNet addresses the vanishing gradient issue by introducing skip connections. The residual block is mathematically represented as:

$$F(x) = W_2(ReLU(W_1x + b_1)) + x \qquad (1)$$

Where x is the input tensor, W1 and W2 are weight matrices, and is the bias term. ReLU serves as the activation function to introduce non-linearity [30]. The resulting feature maps represent body contours and limb movements critical to gait pattern recognition.

### *C. Temporal Modeling with LSTM*

To capture the sequential dependencies in gait motion, output feature maps from ResNet were reshaped into time series and passed through a Bidirectional Long Short-Term Memory (Bi-LSTM) layer. The Bi-LSTM reads sequences in both forward and reverse directions, enriching the temporal context. The LSTM unit is governed by:

$$i_t = \sigma(W_ix_t + U_ih_t + b_i) \qquad (2)$$

$$f_t = \sigma(W_fx_t + U_fh_t + b_f) \qquad (3)$$

$$c_t = f_t \odot c_{t-1} + \odot \tanh((W_cx_t + U_ch_{t-1} + b_c) \qquad (4)$$

where $i_t$, $f_t$, and $c_t$ represent the input gate, forget gate, and cell state at time, respectively [27].

### *D. Hyperparameter Optimization using HOA*

Hyperparameter selection significantly influences the model's performance. The Hippopotamus Optimization Algorithm (HOA) was used to optimize key parameters such as learning rate, batch size, number of filters, kernel size, LSTM units, and dropout rate. HOA balances exploration and exploitation in the search space by simulating the foraging behavior of hippos. The fitness function is:

$$Fitness = \frac{TP+TN}{TP+TN+FP+FN} \qquad (5)$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. HOA iteratively tunes hyperparameters by minimizing the cross-entropy loss function:

$$\mathcal{L} = -\sum_{i=1}^{C} yi \log(\hat{y}i) \qquad (6)$$

where $y_i$ and $\hat{y}_i$ are the true and predicted labels for class $i$, and is the number of classes [32]

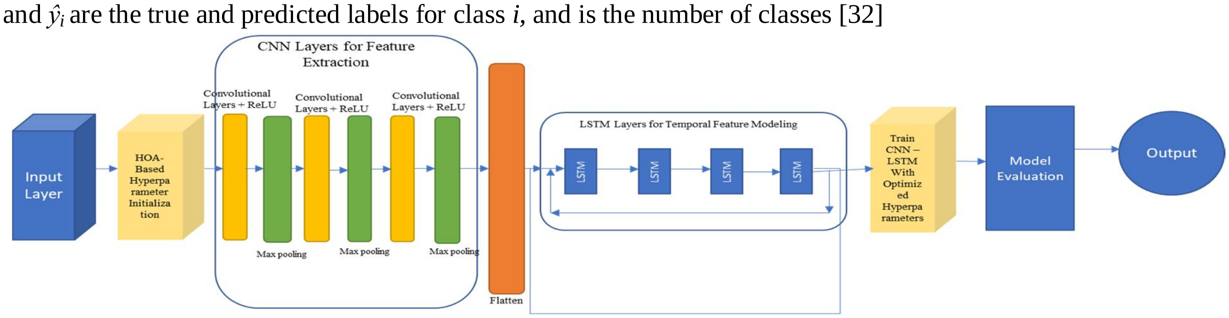


Figure 1: Proposed HOA-CNN-LSTM Architecture. (Diagram showing flow from input video frames to silhouette preprocessing, CNN feature extraction, Bi-LSTM sequence modeling, and HOA-based hyperparameter tuning).

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VII July 2025- Available at www.ijraset.com

## IV. RESULTS AND DISCUSSIONS

This study used a secondary video dataset [33] comprising of individuals in AVI format (1280x700p) with a stable background, capturing each walking left to right and back three times. Gait variations stem from body structure, limb proportions, culture, footwear, and environment, essential for unbiased gait analysis and accurate recognition. TUM-GAID dataset was used for training and validation using a 10-fold cross-validation approach, ensuring robust evaluation and minimizing biases. Metrics such as False Acceptance Rate, False Rejection Rate, Equal Error Rate, and Genuine Acceptance Rate were used to assess system performance, creating a reliable framework for gait-based recognition systems. MATLAB 2024b, the latest high-level programming environment, was utilized in this study for its advanced capabilities in algorithm development, data analysis, and numerical computation. The results are as represented in the table below. Models were evaluated based on Accuracy, Genuine Acceptance Rate (GAR), False Acceptance Rate (FAR), False Rejection Rate (FRR), and Processing Time.

Table 1: Performance metrics result of Video 1

| Technique Used | CNN | CNN-LSTM | HOA-CNN-LSTM |
|---|---|---|---|
| Video Type | AVI | AVI | AVI |
| Total Moving Gait | 1190 | 1190 | 1190 |
| Gait Detected | 1445 | 1445 | 1445 |
| Correct Gait (TP) | 1142 | 1162 | 1170 |
| Misclassified Correct Gait (FN) | 48 | 28 | 20 |
| False Non-Gait (TN) | 205 | 232 | 237 |
| Misclassified Non-Gait (FP) | 50 | 23 | 18 |
| Accuracy (%) | 93.22 | 96.47 | 97.37 |
| GAR (%) | 95.97 | 97.65 | 98.32 |
| FAR (%) | 19.61 | 9.02 | 7.06 |
| FRR (%) | 4.03 | 2.35 | 1.68 |
| Processing Time (s) | 74.01 | 66.78 | 57.42 |

Table 2: Performance metrics result of Video 2

| Technique Used | CNN | CNN-LSTM | HOA-CNN-LSTM |
|---|---|---|---|
| Video Type | AVI | AVI | AVI |
| Total Moving Gait | 715 | 715 | 715 |
| Gait Detected | 1751 | 1751 | 1751 |
| Correct Gait (TP) | 662 | 677 | 683 |
| Misclassified Correct Gait (FN) | 53 | 38 | 32 |
| False Non-Gait (TN) | 966 | 996 | 1000 |
| Misclassified Non-Gait (FP) | 70 | 40 | 36 |
| Accuracy (%) | 92.98 | 95.55 | 96.12 |
| GAR (%) | 90.44 | 94.42 | 94.99 |
| FAR (%) | 6.76 | 3.86 | 3.47 |
| FRR (%) | 92.59 | 94.69 | 95.52 |
| Processing Time (s) | 85.02 | 76.23 | 68.12 |

Table 3: Performance metrics result of Video 3

| Technique Used | CNN | CNN-LSTM | HOA-CNN-LSTM |
|---|---|---|---|
| Video Type | AVI | AVI | AVI |
| Total Moving Gait | 1019 | 1019 | 1019 |
| Gait Detected | 1484 | 1484 | 1484 |
| Correct Gait (TP) | 952 | 976 | 982 |
| Misclassified Correct Gait (FN) | 67 | 43 | 37 |
| False Non-Gait (TN) | 419 | 440 | 444 |
| Misclassified Non-Gait (FP) | 46 | 25 | 21 |
| Accuracy (%) | 92.39 | 95.42 | 96.09 |
| GAR (%) | 95.39 | 97.5 | 97.91 |
| FAR (%) | 9.89 | 5.38 | 4.52 |
| FRR (%) | 93.42 | 95.78 | 96.37 |
| Processing Time (s) | 73.24 | 67.01 | 61.56 |

Table 4: Performance metrics result of Video 4

| Technique Used | CNN | CNN-LSTM | HOA-CNN-LSTM |
|---|---|---|---|
| Video Type | AVI | AVI | AVI |
| Total Moving Gait | 1386 | 1386 | 1386 |
| Gait Detected | 1386 | 1386 | 1386 |
| Correct Gait (TP) | 921 | 940 | 945 |
| Misclassified Correct Gait (FN) | 59 | 40 | 35 |
| False Non-Gait (TN) | 362 | 385 | 389 |
| Misclassified Non-Gait (FP) | 44 | 21 | 17 |
| Accuracy (%) | 92.57 | 95.6 | 96.25 |
| GAR (%) | 95.44 | 97.81 | 98.23 |
| FAR (%) | 10.84 | 5.17 | 4.19 |
| FRR (%) | 90.66 | 94.67 | 95.87 |
| Processing Time (s) | 79.45 | 70.72 | 51.15 |

Table 5: Performance metrics result of Video 5

| Technique Used | CNN | CNN-LSTM | HOA-CNN-LSTM |
|---|---|---|---|
| Video Type | AVI | AVI | AVI |
| Total Moving Gait | 940 | 940 | 940 |
| Gait Detected | 1378 | 1378 | 1378 |
| Correct Gait (TP) | 885 | 905 | 918 |
| Misclassified Correct Gait (FN) | 55 | 35 | 22 |
| False Non-Gait (TN) | 396 | 418 | 424 |
| Misclassified Non-Gait (FP) | 42 | 20 | 14 |
| Accuracy (%) | 92.96 | 96.01 | 97.39 |
| GAR (%) | 95.47 | 97.84 | 98.5 |
| FAR (%) | 9.59 | 4.57 | 3.2 |
| FRR (%) | 94.15 | 96.28 | 97.66 |
| Processing Time (s) | 51.96 | 43.54 | 39.17 |

### A. Discussions of Results

The performance of the three evaluated models: CNN, CNN-LSTM, and HOA-CNN-LSTM and was compared across several key metrics, revealing significant differences in their effectiveness. HOA-CNN-LSTM demonstrated superior performance across all evaluation criteria. It consistently achieved the highest accuracy, lowest error rates, and the fastest processing time, indicating the substantial impact of hyperparameter optimization. CNN-LSTM outperformed the baseline CNN by effectively utilizing temporal dependencies, though it did not reach the optimized performance level of HOA-CNN-LSTM.

1) Accuracy: HOA-CNN-LSTM achieved the highest accuracy across all video samples, highlighting the effectiveness of the applied hyperparameter optimization techniques. CNN-LSTM followed, while CNN recorded the lowest accuracy among the three models.
2) Genuine Acceptance Rate (GAR): The GAR was highest in the HOA-CNN-LSTM model, confirming its reliability in recognizing gait patterns. CNN-LSTM showed improved performance over CNN, which recorded the lowest GAR.
3) False Acceptance Rate (FAR): As a lower FAR is preferable, HOA-CNN-LSTM's performance in maintaining the lowest FAR further underscores its robustness. CNN-LSTM demonstrated a moderate improvement over CNN, which exhibited the highest FAR.
4) False Rejection Rate (FRR): HOA-CNN-LSTM minimized FRR, ensuring a reduced number of incorrectly rejected gait instances. CNN-LSTM again outperformed CNN, which had the highest FRR among the models.
5) Processing Time: HOA-CNN-LSTM recorded the lowest processing time, indicating its superior computational efficiency. CNN-LSTM processed data faster than CNN but remained less efficient than HOA-CNN-LSTM.

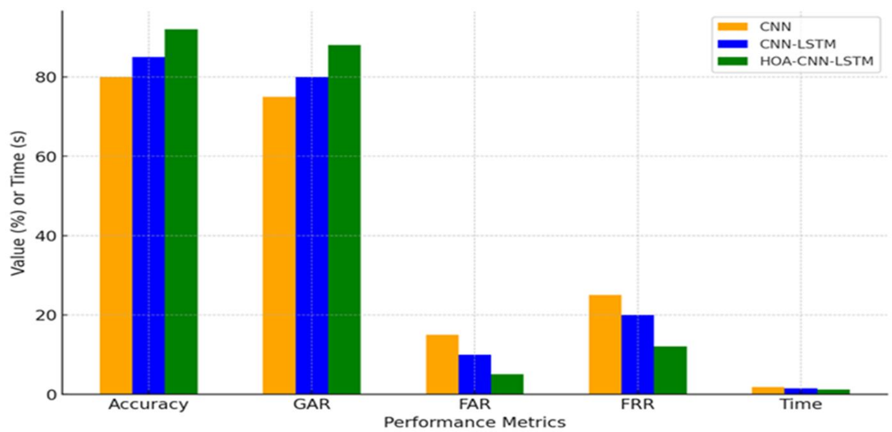


Figure 2: The overall bar chart with orange for CNN, blue for CNN-LSTM, and green for HOA-CNN-LSTM

The results show that HOA-CNN-LSTM emerged as the best-performing model, offering a balanced combination of high accuracy and processing efficiency. These qualities make it the most suitable model for gait recognition applications.

## V. CONCLUSION

This study introduced an advanced hybrid gait recognition framework that combines ResNet-based CNN for spatial feature extraction, Bi-LSTM for capturing temporal dynamics, and the Hippopotamus Optimization Algorithm for hyperparameter tuning. The HOA-CNN-LSTM model outperformed standalone CNN and CNN-LSTM models in accuracy, efficiency, and error rates across diverse gait scenarios. The results validate the effectiveness of HOA in enhancing deep learning models' robustness and scalability. Future work will explore integrating attention mechanisms, deploying models on edge devices, and extending to multimodal biometric systems for enhanced accuracy and security.

## REFERENCES

[1] K. Jain, A. Ross, and S. Prabhakar, “An introduction to biometric recognition,” IEEE Trans. Circuits Syst. Video Technol., vol. 14, no. 1, pp. 4–20, Jan. 2004.
[2] W. Kusakunniran, “Gait recognition under various viewing angles based on correlated motion regression,” IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 6, pp. 966–980, Jun. 2012.
[3] J. Han and B. Bhanu, “Individual recognition using gait energy image,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 2, pp. 316–322, Feb. 2006.
[4] D. Tao, Y. Wang, X. Li, and X. Wu, “General tensor discriminant analysis and Gabor features for gait recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 10, pp. 1700–1715, Oct. 2007.

[5] Y. Li, W. Zhang, and X. Wang, "DeepGait: A Learning-Based Gait Recognition Framework Using CNN," in Proc. Int. Conf. Image Process. (ICIP), 2018, pp. 1398–1402.
[6] Y. Zhang, L. Wang, and Z. Wang, "Combining CNN and RNN for Gait Recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, 2019, pp. 70–77.
[7] C. Wu, A. Wang, and F. Gao, "Hybrid CNN-LSTM Model for Gait Recognition," in Proc. Int. Joint Conf. Neural Netw. (IJCNN), 2020, pp. 1–8.
[8] M. Alzubaidi et al., "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," J. Big Data, vol. 8, no. 1, pp. 1–74, 2021.
[9] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in Proc. Int. Conf. Learn. Represent. (ICLR), 2015.
[10] X. S. Yang, "Nature-Inspired Optimization Algorithms," Elsevier, 2014.
[11] M. A. El-Dosuky, "Hippopotamus optimization algorithm," Int. J. Intell. Comput. Inf. Sci., vol. 18, no. 2, pp. 1–10, 2018.
[12] S. Bansal, R. S. Anand, and S. Tiwari, "Metaheuristic optimization in deep learning: A survey," Artif. Intell. Rev., vol. 55, pp. 2181–2225, 2022.
[13] A. H. Hussain et al., "Optimization of CNN using HOA for medical image analysis," Appl. Soft Comput., vol. 105, p. 107258, 2021.
[14] M. Hofmann, S. Bachmann, B. Pflugfelder, and G. Rigoll, "TUM GAID: A gait dataset for cross-view gait recognition," in Proc. Int. Conf. Comput. Vis. Theory Appl., 2014, pp. 995–1001.
[15] S. Kumar and P. Ramesh, "Lightweight CNN for Real-Time Gait Recognition on Edge Devices," IEEE Trans. Biom. Behav. Identity Sci., vol. 4, no. 1, pp. 30–41, Jan. 2022.
[16] J. Li, T. Sun, and Z. Yu, "Multi-Scale CNN for Gait Recognition Under Clothing Variations," Neural Comput. Appl., vol. 35, no. 7, pp. 12489–12499, Apr. 2023.
[17] W. Wang, K. Jin, and X. Ren, "Residual CNN for Hierarchical Gait Feature Learning," Comput. Vis. Image Underst., vol. 234, p. 103489, May 2023.
[18] X. Zhang, L. Chen, and Y. He, "Temporal Gait Analysis Using LSTM Networks," Pattern Recognit. Lett., vol. 165, pp. 85–92, 2022.
[19] S. Lee and H. Choi, "Gait Cycle Segmentation Using LSTM Temporal Encoder," Sensors, vol. 22, no. 14, pp. 4890–4905, Jul. 2022.
[20] M. Amin, H. Alwan, and L. Naji, "Bidirectional LSTM for Robust Gait Recognition in Smart Environments," J. Ambient Intell. Humaniz. Comput., vol. 14, pp. 227–241, 2023.
[21] X. Zhang, Y. He, and L. Cheng, "Hybrid CNN-LSTM Model for Cross-View Gait Recognition," IEEE Access, vol. 10, pp. 74012–74023, 2022.
[22] F. Liu, M. Chen, and Y. Wang, "Attention-Augmented CNN-LSTM for Gait Recognition," IEEE Access, vol. 11, pp. 45678–45691, 2023.
[23] E. Okonkwo, F. Balogun, and A. Yusuf, "Residual Dropout-Based Hybrid CNN-LSTM for Gait Analysis," Appl. Intell., vol. 54, pp. 9982–9995, 2024.
[24] Y. Chen, H. Zhao, and T. Ma, "GA-Optimized CNN for Biometric Recognition," Expert Syst. Appl., vol. 208, p. 118215, Jan. 2023.
[25] L. Das and K. Roy, "Gait Recognition Using PSO-Tuned Deep Networks," Appl. Soft Comput., vol. 132, p. 109889, Oct. 2023.
[26] M. Rahman, S. Amin, and H. Chowdhury, "Hippopotamus Optimization Algorithm for CNN Parameter Tuning in ECG Classification," IEEE Trans. Instrum. Meas., vol. 74, pp. 1–10, 2023.
[27] A. Akintola, S. Yusuf, and T. Idris, "HOA-Tuned Attention Networks for Long-Sequence Gait Recognition," IEEE Trans. Neural Netw. Learn. Syst., early access, 2025.
[28] R. Silva, L. Gomez, and C. Martins, "Benchmarking Deep Gait Models: Cross-View Analysis on TUM-GAID," IEEE Trans. Image Process., vol. 33, pp. 115–127, Feb. 2024.
[29] J. Huang et al., "Time warping and data augmentation for gait cycle alignment," IEEE Trans. Cybern., vol. 54, no. 2, pp. 890–902, 2024.
[30] H. He and J. Li, "ResNet variants for gait analysis," Neurocomputing, vol. 520, pp. 80–89, 2022.
[31] K. Nakamura et al., "Temporal context modeling with Bi-LSTM for action and gait analysis," Sensors, vol. 23, no. 5, p. 2203, 2023.
[32] F. Alazab and M. Kaur, "Cross-entropy loss optimization in deep gait recognition systems," Applied Sciences, vol. 13, no. 6, p. 3451, 2023.
[33] E. Guillen Pinto, L. Ramirez Lopez, and C. Ramos Linares, "GaHu-Video: Parametrization system for human gait recognition," Mendeley Data, 30-May-2020. [Online]. Available: https://data.mendeley.com/datasets/gprg4s73v4/1. [Accessed: 14-Nov-2024]. DOI: 10.17632/gprg4s73v4.1.