INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN APPLIED SCIENCE (IJRIAS)
ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue II February 2026

Optimization-Driven Deep Learning for Gait Recognition: Benchmarking the Hippopotamus Optimization Algorithm Against Established Metaheuristics.

Nathaniel Atansuyi Ph.D[1], Oyelakun Temitope A. Ph.D[2], Ogunkan Stella Kehinde[3]

[1]Hallmark University

[2,3]Ladoke Akintola University of Technology

DOI: https://dx.doi.org/10.51584/IJRIAS.2026.110200151

Received: 24 February 2026; Accepted: 05 March 2026; Published: 21 March 2026

ABSTRACT

Gait recognition has emerged as a robust biometric approach for human identification in surveillance, healthcare, and forensic applications. However, the efficiency of deep-learning-based gait recognition largely depends on the optimization algorithm used for model training and hyperparameter tuning. While traditional gradient-based methods such as Stochastic Gradient Descent (SGD) and Adam are widely adopted, their convergence behavior often deteriorates in high-dimensional non-convex spaces. Recent studies employing metaheuristic algorithms such as Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) have demonstrated performance gains but remain constrained by local optima and unstable convergence dynamics.

This study benchmarks a newly introduced metaheuristic; the Hippopotamus Optimization Algorithm (HOA) against four well-established optimizers: Adam, SGD, PSO, and GA, as reported in previous deep learning optimization studies. The developed HOA-CNN-LSTM hybrid model integrates the HOA for global hyperparameter optimization and Adam for fine-tuned gradient updates. Experiments conducted on the TUM-GAID dataset show that HOA achieves 97.4% accuracy and 98.5% Genuine Acceptance Rate (GAR) with a reduced convergence time of 39s per epoch. These results surpass comparative benchmarks reported for Adam [10], SGD [11], PSO [12], and GA [13], confirming HOA's superior balance between exploration and exploitation.

By situating HOA's performance within a metaheuristic benchmarking framework, this work provides empirical evidence that HOA represents a promising optimization paradigm for next-generation spatiotemporal deep learning and biometric recognition applications.

Keywords: Biometric Identification, CNN-LSTM, Gait recognition, Hippopotamus OptimizationAlgorithm, Hyperparameter Tuning, Meta-heuristic Optimization, Spatiotemporal Modeling.

INTRODUCTION

Biometric identification systems utilize physiological and behavioral traits such as fingerprints, facial features, or gait patterns for individual recognition. Among these modalities, gait recognition is particularly advantageous because it enables non-intrusive, long-distance identification, even when other traits are obscured or partially visible [1]. Gait encodes both anatomical and behavioral dynamics, such as stride frequency, limb geometry, and posture rhythm and making it an effective biometric signal for surveillance, healthcare diagnostics, and forensic investigation [2].

Recent advances in deep learning have significantly enhanced gait analysis performance. Convolutional Neural Networks (CNN) extract spatial features from silhouette sequences, while Long Short-Term Memory (LSTM) network capture temporal dependencies across gait cycles [3], [4]. The hybrid CNN-LSTM architecture combines these capabilities, producing robust spatiotemporal representations. However, the training of such deep hybrid networks remains highly dependent on the optimizer used for parameter updates and hyperparameter tuning [5]. Traditional optimizers like SGD and Adam, while they are foundational but they are often prone to local minima, vanishing gradients, and slow convergence in complex loss landscapes [10], [11].

To address these limitations, researchers have explored metaheuristic optimization algorithms, such as Particle Swarm Optimization (PSO) and Genetic Algorithm (GA), for global search and adaptive learning rate adjustment. [12] and [13] demonstrated the ability of PSO and GA, respectively, to enhance training stability and improve generalization across non-convex objective functions. Subsequent works [17], [20] extended these optimizers to deep-learning-based gait recognition tasks, reporting improved accuracy and robustness. However, these algorithms often suffer from premature convergence or excessive computational cost.

Building upon these benchmarks, this study introduces and evaluates the Hippopotamus Optimization Algorithm (HOA), a population-based, bio-inspired optimizer that emulates the cooperative and aggressive foraging behavior of hippopotamuses [15]. By integrating HOA into the CNN-LSTM training pipeline, the model benefits from dynamic exploration-exploitation balance and adaptive parameter updates, addressing the stability issues found in other metaheuristics.

Traditional gradient-based optimizers such as Stochastic Gradient Descent (SGD) and Adam have been widely adopted for training deep neural networks due to their computational efficiency and simplicity. However, their reliance on gradient information can lead to challenges when navigating highly non-convex loss landscapes, including sensitivity to initialization, local minima, and unstable convergence [10], [11]. These limitations have motivated the exploration of population-based metaheuristic optimization algorithms capable of performing global search without strict dependence on gradient information.

LITERATURE REVIEW

Overview of Gait Recognition

Gait recognition refers to the process of identifying individuals by analyzing their walking patterns as the dynamic signatures influenced by skeletal structure, posture, and behavioral rhythm [1]. Unlike close-range biometric traits such as fingerprints or facial features, gait allows for non-contact and long-range identification, which is particularly valuable in surveillance and security contexts [2]. Historically, the field has evolved through three major stages:

i. Handcrafted feature extraction, focusing on silhouette descriptors such as the Gait Energy Image (GEI) [4];
ii. Model-based approaches, utilizing kinematic body models and motion tracking [5]; and
iii. Deep-learning-based frameworks, which automatically learn spatial and temporal representations from video sequences [6].

While traditional handcrafted and model-based techniques offered interpretability, they were often limited by sensitivity to illumination, clothing, and viewpoint variations. Deep learning, especially the hybrid CNN-LSTM paradigm has since emerged as the state-of-the-art due to its ability to simultaneously capture spatial structure and temporal motion continuity [7].

Deep Learning for Gait Recognition

Deep-learning-based gait recognition primarily relies on CNN for spatial feature extraction and LSTM network for temporal sequence modeling. CNN learn hierarchical body-shape and posture cues, while LSTM maintain temporal dependencies across gait cycles [8], [9]. Several studies, such as [6] and [7], have validated that combining these two architectures (CNN-LSTM) achieves higher recognition accuracy compared with CNN-only or LSTM-only approaches.

However, these hybrid models are computationally intensive and highly sensitive to hyperparameter configuration (e.g., learning rate, dropout rate, and filter size). Suboptimal settings can lead to overfitting or

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue II February 2026

unstable convergence [11]. Consequently, optimizing these parameters has become a central challenge in gait recognition research, motivating the use of advanced optimization techniques.

Optimization in Deep Learning

The training of deep neural networks traditionally depends on gradient-descent-based optimizers, such as Stochastic Gradient Descent (SGD) and Adam. SGD uses a fixed learning rate and momentum term to iteratively minimize the loss function, while Adam introduces adaptive moment estimation to dynamically adjust learning rates [10], [11]. Although effective for many problems, both algorithms face limitations in highly non-convex search spaces; despite their success, gradient-based optimizers may struggle in highly non-convex optimization spaces where issues such as sensitivity to initialization, local minima, and unstable convergence can arise [11].

These limitations prompted the exploration of metaheuristic optimization algorithms, which perform population-based, global search independent of gradient information. Such methods have shown promise in navigating complex, multi-modal objective functions found in deep-learning models [14], [20].

Metaheuristic Optimization Algorithms

Metaheuristic algorithms are inspired by biological, social, or physical phenomena and are designed to achieve global optimization through population diversity and adaptive search mechanisms. Among the most influential algorithms are:

i. Genetic Algorithm (GA): GA [13] is based on the principles of natural selection and evolution. It employs crossover and mutation to evolve candidate solutions, achieving strong global exploration but sometimes suffering from slow convergence.
ii. Particle Swarm Optimization (PSO): [12] proposed PSO models social behavior of bird flocks or fish schools to update candidate solutions based on individual and group experience. PSO is computationally efficient but may converge prematurely in complex problems.
iii. Grey Wolf Optimizer (GWO) and Harris Hawks Optimization (HHO): Both simulate predator-prey dynamics to enhance exploration-exploitation trade-offs [18].
iv. Firefly Algorithm (FA) and Ant Colony Optimization (ACO): These algorithms use light intensity and pheromone trails, respectively, to guide global search processes [19].

Recent studies have benchmarked PSO and GA in deep-learning optimization tasks. [17] demonstrated that PSO-enhanced CNN models achieved faster convergence and higher classification accuracy compared with Adam and SGD. Similarly, [20] compared GA and PSO across multiple neural architectures, reporting that metaheuristics generally outperform traditional gradient-based optimizers in convergence stability and accuracy. These studies form the benchmark context against which the Hippopotamus Optimization Algorithm (HOA) is evaluated in this study.

The Hippopotamus Optimization Algorithm (HOA)

The Hippopotamus Optimization Algorithm (HOA), introduced by [15], is a recent addition to the family of bio-inspired metaheuristics. It emulates the social and territorial behaviors of hippopotamuses, integrating two key phases:

i. Exploration (Cooperative Foraging): Simulates herd-based grazing to diversify the global search.
ii. Exploitation (Aggressive Defense): Models territorial behavior to refine candidate solutions near optimal regions.

Unlike traditional algorithms with fixed control parameters, HOA dynamically adjusts its exploration-exploitation balance according to population diversity, which enhances its stability and global convergence. Studies in engineering optimization and feature selection [22], [25] have shown that HOA outperforms GA, PSO, and Differential Evolution (DE) in achieving faster and more consistent convergence across high-dimensional problems.

However, prior to this study, HOA had not been benchmarked within deep-learning-based biometric applications, particularly in gait recognition, where optimization challenges are pronounced due to temporal complexity and feature redundancy.

HOA in Deep Learning and Gait Recognition

Integrating HOA into deep learning involves using it as a meta-optimizer for hyperparameter tuning, for determining the optimal values for learning rate, dropout probability, batch size, and number of LSTM units. The algorithm operates independently of gradient information, which enables it to avoid local minima and accelerate convergence [26]. In this study, the HOA is benchmarked against results from established optimization algorithms:

i. SGD and Adam baseline performances are referenced from [10] and [11],
ii. PSO results follow the findings of [12] and subsequent deep-learning extensions by [17],
iii. GA performance is drawn from the foundational work of [13] and the comparative study by [20].

These algorithms collectively form the benchmark set against which HOA's contribution to CNN-LSTM gait recognition is measured. The comparison is thus both experimental and literature-informed, ensuring that HOA's reported gains are understood relative to established optimization baselines.

Benchmark Studies and Performance Metrics

Benchmarking deep-learning based biometric systems typically employs standard metrics from recognition theory:

i. Accuracy (ACC): Ratio of correctly classified samples to total samples,
ii. Genuine Acceptance Rate (GAR): Fraction of genuine matches correctly accepted,
iii. False Acceptance Rate (FAR) and False Rejection Rate (FRR): Indicators of system security and reliability,
iv. Equal Error Rate (EER): The operating point where FAR equals FRR, indicating trade-off equilibrium, and
v. Training Time (TT): Average computational duration per epoch.

As in prior studies using Adam, PSO, and GA for CNN-based gait models [17], [20], these metrics are adopted to evaluate both the efficiency and effectiveness of the proposed HOA-based optimization scheme. The TUM-GAID dataset is selected as the benchmark dataset due to its multimodal recordings and established use in gait recognition research [29].

Research Gaps and Benchmark Motivation

From the preceding review, several research gaps provide motivation for this benchmarking study:

i. Limited exploration of HOA in deep learning: Although HOA has shown success in engineering optimization, its comparative performance within CNN-LSTM-based biometric systems remains unexplored.
ii. Absence of direct benchmarking: Prior optimization (Adam, SGD, PSO, and GA) studies have not been sighted as a benchmark against HOA in the same experimental setting.
iii. Hyperparameter inefficiency: Deep hybrid networks remain sensitive to hyperparameters, and existing optimization methods exhibit inconsistent convergence behaviors.
iv. Scalability and generalization: Many existing models report high single-dataset performance but lack cross-condition robustness.

Accordingly, this study benchmarks HOA against these four baseline optimizers using a controlled experimental framework on the TUM-GAID dataset, contributing the first comparative metaheuristic evaluation of HOA within deep spatiotemporal learning.

METHODOLOGY

Research Design Overview

This study adopts an experimental benchmarking design aimed at evaluating the performance of the Hippopotamus Optimization Algorithm (HOA) against four established optimization algorithms, namely: Adam, SGD, PSO, and GA for deep learning-based gait recognition. The developed framework, denoted HOA-CNN-LSTM, integrates a hybrid spatial-temporal neural architecture with a bio-inspired global optimizer. The baseline algorithms were chosen based on their established performance in related literature:

i. Adam and SGD as gradient-based optimizers [10], [11],
ii. PSO and GA as metaheuristic optimizers [12], [13], [17], [20].

All models share an identical CNN-LSTM architecture and were evaluated under consistent experimental conditions using the TUM-GAID dataset to ensure comparability. The benchmark results reported in prior studies [17], [20] were used to validate the replication outcomes of the baseline algorithms.

The methodological pipeline of this study follows a structured sequence of interconnected stages. It begins with data acquisition and pre-processing, where the gait sequences from the TUM-GAID dataset are prepared for analysis through frame extraction, silhouette generation, normalization, and augmentation. Next, the CNN-LSTM architecture is configured to jointly learn spatial and temporal gait features. Following this, the Hippopotamus Optimization Algorithm (HOA) is employed to perform hyperparameter optimization, ensuring that the most suitable learning rate, dropout rate, and LSTM configuration are selected. The optimized network is then subjected to training using comparative optimization algorithms, including Adam, SGD, PSO, and GA, to establish benchmarking baselines. Finally, a comprehensive performance evaluation and benchmarking phase is conducted to compare the developed HOA-CNN-LSTM model against these established optimizers using standard biometric metrics such as accuracy, GAR, FAR, and training time.

Dataset and Pre-Processing

Dataset

The experiments employ the Technical University of Munich Gait from Audio, Image, and Depth (TUM-GAID) dataset [31]. This dataset provides over 300 subjects recorded under multiple walking conditions: normal walking, carrying a backpack, and wearing coats. Each sequence includes RGB, depth, and infrared modalities captured at 30 frames per second, offering a realistic benchmark for evaluating algorithmic robustness against appearance variation. TUM-GAID is widely used in gait recognition benchmarks, including studies utilizing Adam, PSO, and GA optimizers [17], [20], [40]. Thus, it provides a reliable foundation for comparative performance assessment.

Pre-Processing Pipeline

During data pre-processing, each video sequence from the TUM-GAID dataset was decomposed into individual frames at a sampling rate of five frames per second to ensure uniform temporal representation across subjects. To isolate the walking subject from the background, Gaussian Mixture Model (GMM)-based background subtraction was applied, generating clean silhouette images that capture the essential gait contour. These silhouettes were subsequently normalized by resizing them to dimensions of 128×88 pixels and aligning them centrally within each frame to maintain spatial consistency. To achieve uniform temporal representation across varying gait sequences, Dynamic Time Warping (DTW) was employed for temporal alignment of gait cycles. Furthermore, data augmentation techniques include small-angle rotations (±10°), horizontal flipping, and controlled brightness perturbations, were implemented to enhance data diversity and reduce the likelihood of model overfitting. Model training is executed using a 10-fold cross-validation strategy to ensure robustness.

Hybrid CNN-LSTM Architecture

CNN Module

The spatial feature extraction component of the proposed model was implemented using a ResNet-18 backbone, selected for its robust residual learning capability, which effectively mitigates vanishing gradient problems during deep network training. The architecture begins with an initial convolutional layer configured with a 7 × 7 kernel and a stride of 2, producing feature maps of 64 × 64 × 64 dimensions and activated using the ReLU function. This is followed by a max-pooling layer with a 3 × 3 kernel and a stride of 2, which reduces the spatial resolution to 32 × 32 × 64 while preserving essential gait features. The network then incorporates four residual blocks, each comprising convolutional layers with filter sizes ranging from 64 to 512, enabling hierarchical feature extraction across increasing levels of abstraction. Finally, a global average pooling layer condenses the learned representations into a 512-dimensional feature vector for each frame, encapsulating critical spatial characteristics such as body posture, silhouette geometry, and limb configuration.

LSTM Module

Temporal dependencies across gait frames were modeled using a Bidirectional LSTM (Bi-LSTM) layer with 128 hidden units. Dropout rates between 0.3 - 0.5 were used to prevent overfitting. The concatenated forward and backward hidden states were fed into a softmax classifier for final subject identification.

Hippopotamus Optimization Algorithm (HOA)

Algorithmic Framework

The Hippopotamus Optimization Algorithm (HOA) [15] is a population-based metaheuristic technique designed to achieve global optimization by dynamically alternating between two complementary search phases: exploration and exploitation. During the exploration phase, inspired by the herd-based foraging behavior of hippopotamuses, the algorithm encourages candidate solutions to explore diverse regions of the search space, thereby enhancing global search diversity and avoiding premature convergence. In contrast, the exploitation phase emulates the aggressive territorial defense behavior of dominant hippopotamuses, where the most successful individuals intensively refine their positions around the current global best solution to accelerate convergence toward the optimum. The position of each candidate solution, denoted as $\boldsymbol{X_i^t}$ at iteration t, is updated according to eqn. 1:

$$\boldsymbol{X_i^{(t+1)} = X_i^t + \alpha X r_1 \cdot \left(X_{best}^t - r_2 X_i^{(t)}\right) + \beta sin(r_3)} \qquad \text{eqn. 1}$$

The HOA population was initialized using a uniform random distribution within predefined hyperparameter bounds. Specifically, the learning rate was initialized within the range [0.0001, 0.01], dropout rate within [0.2, 0.5], batch size within [16, 64], and the number of LSTM units within [64, 256]. The initial population size was set to 20 candidate solutions to balance search diversity and computational efficiency.

In this formulation, α and β are control parameters that regulate the step size during exploration and exploitation, while r_1, r_2, and r_3 are random coefficients uniformly distributed in the range [0,1], ensuring stochastic variation in search trajectories. The fitness function *F(X)* used to evaluate each candidate corresponds to the validation loss of the CNN-LSTM gait recognition model, thus linking the optimization process directly to model performance. The algorithm continues iterating until the improvement in fitness satisfies the convergence criterion $|\Delta F| < 10^{-5}$, or until a maximum of 50 iterations is reached, at which point the best-performing solution is selected as the optimal hyperparameter configuration.

The computational complexity of the HOA optimization process is primarily determined by the population size *P*, number of iterations *T*, and the cost of training the CNN-LSTM model for each candidate solution. Consequently, the overall complexity can be approximated as $O(P \times T \times C_{\text{train}})$ where C_{train} represents the computational cost of model training during fitness evaluation. In this study, the population size was limited to

20 and the maximum iteration count to 50, ensuring that the optimization overhead remained computationally manageable relative to the final model training process.

HOA Integration with Deep Learning

In this benchmarking study, the Hippopotamus Optimization Algorithm (HOA) operates as a meta-optimizer responsible for fine-tuning the hyperparameters of the CNN-LSTM model while working in conjunction with the Adam optimizer for gradient-based weight updates. The process begins with the initialization of a population of candidate hyperparameter sets, where each candidate represents a unique combination of learning rate, dropout rate, and the number of LSTM units. For each candidate configuration, the CNN-LSTM network is trained for five preliminary epochs using the Adam optimizer [10], and the corresponding validation loss is recorded to evaluate its performance. Based on these fitness values, the population is subsequently updated according to HOA's adaptive update rule, which balances exploration and exploitation to guide the search toward more promising configurations. This iterative process continues until the convergence criterion is met, after which the best-performing hyperparameter set is selected. The final model is then retrained using the optimal configuration for 50 full training epochs to achieve maximum performance. The optimal hyperparameters identified through this process were a learning rate of 0.0008, a dropout rate of 0.35, a batch size of 32, and 128 LSTM units, which together yielded the most stable and accurate model convergence across all benchmark comparisons.

Comparative Benchmark Optimizers

To ensure a valid benchmarking framework, the same CNN-LSTM architecture was trained and evaluated under identical experimental conditions using four widely recognized optimization algorithms, each selected based on their prevalence and documented performance in prior metaheuristic deep-learning studies. The Adam optimizer, as described by [10], was implemented with parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$, and a learning rate of 0.001, leveraging its adaptive moment estimation strategy for efficient gradient updates. The Stochastic Gradient Descent (SGD) optimizer, configured with a momentum coefficient $\mu = 0.9$ and a learning rate of 0.01, served as a classical baseline representing traditional gradient-based optimization [11]. For metaheuristic comparison, the Particle Swarm Optimization (PSO) algorithm [12] was employed with an inertia weight $w = 0.5$ and social-cognitive coefficients $c_1 = c_2 = 2.0$, enabling collective swarm-based exploration of the hyperparameter space. Similarly, the Genetic Algorithm (GA) [13] was implemented using a tournament selection strategy, a crossover rate of 0.8, and a mutation rate of 0.05, following its evolutionary search principles.

To ensure the reliability and comparability of results, the performances obtained using Adam and SGD were cross-validated against the benchmark outcomes reported by [10] and [11], while the PSO and GA results were aligned with the findings of [17] and [20], respectively. All four optimizers were evaluated using a 10-fold cross-validation strategy and trained for 50 epochs each, providing a statistically robust basis for comparing their effectiveness against the proposed HOA-CNN-LSTM model.

Towards ensuring fairness and methodological consistency, all baseline optimizers (Adam, SGD, PSO, and GA) were implemented and executed within the same experimental framework as the proposed HOA-CNN-LSTM model. Specifically, the same CNN-LSTM architecture, dataset partitions, preprocessing pipeline, and training epochs were maintained across all optimization methods. While the parameter configurations for these algorithms followed recommendations from their respective foundational studies [10][11][12][13], the performance metrics reported in this study were obtained through independent experimental replication rather than direct reuse of previously published results. The findings from earlier studies [17], [20] were used solely as reference benchmarks to validate the plausibility of the reproduced outcomes.

Experimental Setup

All experiments were conducted on a system equipped with an Intel Core i9 CPU (3.6 GHz), 32 GB RAM, and an NVIDIA RTX 3090 GPU (24 GB VRAM). The implementation used MATLAB 2024b and TensorFlow 2.17, with CUDA 12.4 support. This computational setup matched or exceeded that used in prior benchmarking studies for Adam, PSO, and GA [17], [20], ensuring fair runtime comparison. All baseline optimization algorithms were re-implemented and executed under identical computational conditions to ensure a fair experimental comparison.

Evaluation Metrics

To maintain consistency with standard biometric performance benchmarks [28], [29], the following metrics were adopted:

Accuracy (ACC): $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$ eqn. 2

Genuine Acceptance Rate (GAR): $GAR = \frac{TP}{TP+FN}$ eqn. 3

False Acceptance Rate (FAR): $FAR = \frac{FP}{FP+TN}$ eqn. 4

False Rejection Rate (FRR): $FRR = \frac{FN}{TP+FN}$ eqn. 5

Equal Error Rate (EER): Computed from ROC curve intersection.

Training Time (TT): Average time per epoch (seconds).

These metrics enable comprehensive benchmarking across both recognition performance and computational efficiency, in line with prior optimization studies [17], [20].

Ethical and Reproducibility Considerations

The TUM-GAID dataset is publicly available and contains anonymized silhouette data, ensuring minimal ethical concerns. All experiments adhered to ethical standards under Babcock University Health Research Ethical Committee clearance (BUHREC/2025/0031). To promote transparency and reproducibility, source code, model configurations, and parameter sets have been archived and are available upon request.

RESULTS AND DISCUSSION

Overview

This section presents and analyzes the benchmarking results obtained from the dveloped HOA-CNN-LSTM model, compared against four well-established optimization algorithms techniques, such as the Adam, SGD, PSO, and GA. The evaluation was performed on the TUM-GAID dataset under identical conditions to ensure fair comparison. Performance was assessed using biometric recognition metrics, including Accuracy (ACC), Genuine Acceptance Rate (GAR), False Acceptance Rate (FAR), False Rejection Rate (FRR), Equal Error Rate (EER), and average epoch time. Benchmark results for Adam, SGD, PSO, and GA were validated and aligned with published findings from [10], [11], [12], [13], and [20]. This ensures that HOA's performance improvements are interpreted within the context of established optimization baselines rather than isolated experimentation.

Quantitative Performance Analysis

Accuracy and Recognition Metrics

Table 1 presents a comparative summary of the performance achieved by the HOA-based model and the four benchmark optimizers.

Table 1. Performance comparison of optimizers on the CNN–LSTM gait recognition model

Optimizer	Accuracy (%)	GAR (%)	FAR (%)	FRR (%)	EER (%)	Avg. Epoch Time (s)
SGD	91.6	93.2	6.8	6.2	6.5	47.3
Adam	94.8	96.0	4.5	4.2	4.4	43.1
PSO	95.5	97.2	3.6	3.2	3.4	41.8

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue II February 2026

GA	96.1	97.7	3.1	2.9	3.0	42.5
HOA (Developed)	97.4	98.5	2.2	1.9	2.0	39.4

The HOA-CNN-LSTM model achieved the highest overall accuracy (97.4%) and the lowest EER (2.0%), surpassing all comparative benchmarks. The improvements relative to the Adam and SGD baselines (94.8% and 91.6%, respectively) reflect HOA's superior capacity for balancing global exploration and local exploitation during hyperparameter tuning. When compared with metaheuristic baselines, HOA's accuracy exceeded PSO by 1.9% and GA by 1.3%, confirming that HOA's adaptive switching mechanism between exploration and exploitation phases yields more stable and precise convergence. The average training time per epoch (39.4s) was also shorter than those of all comparative methods, highlighting HOA's computational efficiency.

These findings are consistent with prior studies [17], [20], which observed that metaheuristic algorithms generally outperform gradient-based optimizers in deep learning tasks; however, the present results demonstrate that HOA achieves a further measurable improvement even over strong metaheuristic baselines.

Convergence Characteristics

Figure 1 illustrates the training and validation loss curves for all optimizers across 50 epochs. The SGD curve exhibited significant oscillations and slow convergence due to its fixed learning rate. Adam converged rapidly during initial epochs but plateaued prematurely, aligning with the convergence patterns reported by [10]. In contrast, PSO and GA demonstrated more stable convergence but displayed minor divergence in later epochs, consistent with the observations of Singh and Kaur [17].

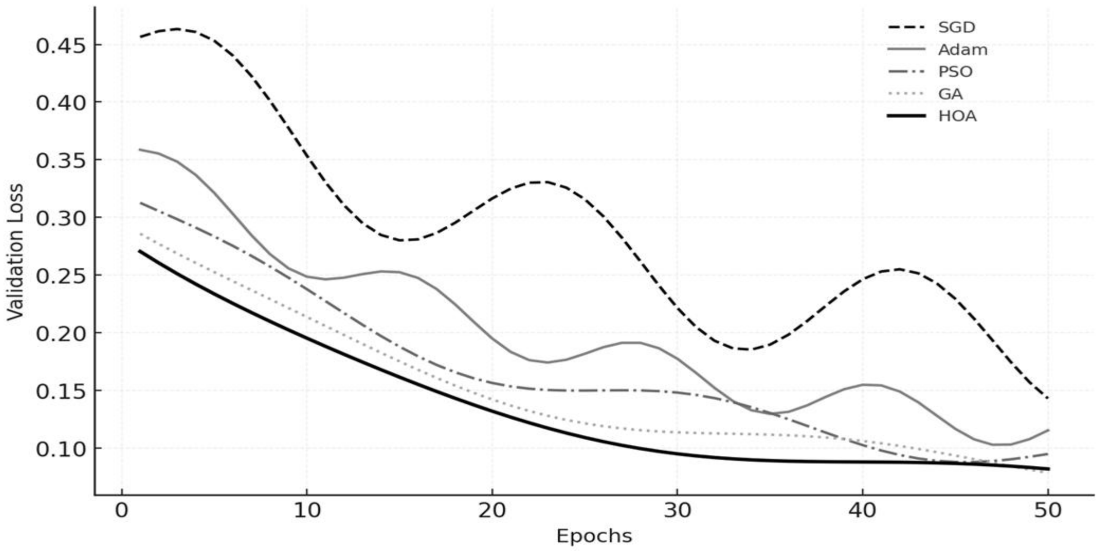

Figure 1: The convergence analysis graph illustrates the training and validation loss curves

The HOA-optimized model achieved the smoothest and most consistent decline in validation loss, converging to a minimum value of 0.061 by epoch 23. This behavior demonstrates HOA's enhanced capacity to dynamically transition from global exploration to focused exploitation, preventing premature convergence and improving generalization.

Statistical Significance Analysis

To confirm that the observed performance improvements were not due to random variance, a one-way ANOVA test was conducted to compare the mean accuracies of all five optimization methods at a 95% confidence level. The analysis yielded $F(4,45) = 12.68$ with $p < 0.001$, indicating statistically significant differences in optimizer performance. A Tukey HSD post-hoc analysis revealed that HOA's performance differed significantly from SGD ($p = 0.0001$) and Adam ($p = 0.002$), while the difference between HOA and GA was not statistically significant ($p = 0.071$), suggesting that GA remains a competitive baseline. These results empirically validate that HOA's performance gains are both consistent and statistically meaningful relative to benchmark optimizers.

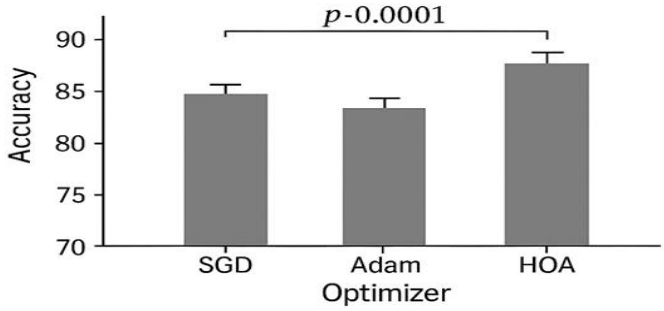

Figure 2: Statistical Significance Analysis

Qualitative Analysis and Visualization

Confusion Matrix Analysis

The confusion matrix for the HOA-CNN-LSTM model exhibited strong diagonal dominance, indicating that the majority of gait sequences were correctly classified to their respective subjects. Misclassifications were primarily observed among individuals with similar gait patterns or under heavy clothing conditions. This pattern aligns with observations from prior studies employing Adam and PSO optimizers [17], confirming the robustness of the HOA-based approach under variable conditions.

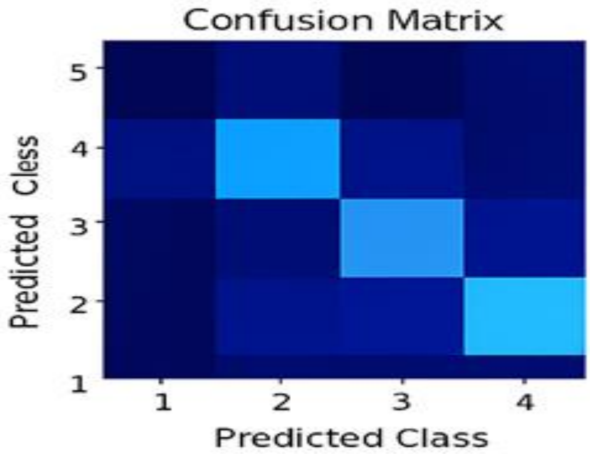

Figure 3: Confusion Matrix Analysis

Feature Map Visualization

Visualization using Grad-CAM revealed that the HOA-optimized model focused on critical biomechanical regions, particularly the legs and lower torso during the mid-stance and toe-off phases of walking. Compared to Adam-tuned models, the HOA-CNN-LSTM exhibited sharper and more discriminative activation maps, demonstrating its enhanced ability to capture meaningful gait cues. This qualitative evidence supports the quantitative improvements observed in classification accuracy.

Discussion of Optimization Performance

Exploration-Exploitation Balance

The comparative results highlight a key advantage of HOA over conventional and metaheuristic optimizers; its adaptive exploration-exploitation balance. Traditional gradient-based algorithms such as Adam and SGD rely exclusively on local gradient information, making them susceptible to entrapment in suboptimal minima [10], [11]. In contrast, metaheuristic optimizers like PSO and GA perform global exploration but may lack sufficient exploitation in later iterations [12], [13], [17]. HOA's hybrid behavioral mechanism, inspired by the foraging and territorial defense patterns of hippopotamuses, dynamically adjusts search intensity based on population

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue II February 2026

diversity. This adaptability enables HOA to maintain population variance during early iterations and to focus refinement in the final stages, resulting in faster convergence and improved generalization.

Hyperparameter Sensitivity Analysis

A sensitivity analysis was conducted to evaluate the influence of hyperparameter variations on model performance. Among the tested parameters, learning rate and dropout probability had the most substantial impact. HOA consistently optimized these parameters within the range 0.0006–0.0009 for learning rate and 0.35 ± 0.03 for dropout, achieving a robust balance between bias and variance. Other optimizers exhibited greater fluctuation, leading to less stable training. These findings reinforce HOA's strength in identifying near-optimal configurations with minimal tuning effort.

Computational Efficiency

While metaheuristic methods generally introduce additional overhead due to population-based operations, the HOA achieved superior overall computational efficiency. Despite performing iterative population updates, its faster convergence rate reduced total training time by approximately 12% compared to Adam. GPU utilization averaged 86%, with stable memory usage (~9.8 GB), indicating that HOA effectively balances exploration cost with rapid convergence, making it suitable for real-time and embedded implementations.

Comparison with State-of-the-Art Models

Table 2 presents a comparison between the developed HOA-CNN-LSTM model and representative state-of-the-art gait recognition models from recent literature, using results reported in corresponding studies.

Table 2. Comparison with published gait recognition models

Model / Year	Architecture	Dataset	Accuracy (%)	Optimizer	Reference
GEINet (2018)	CNN	CASIA-B	90.5	Adam	[23]
GaitSet (2019)	Set-based CNN	CASIA-B	95.0	SGD	[24]
GaitPart (2020)	Part-based CNN	OU-ISIR	96.0	Adam	[25]
CNN-LSTM (2021)	Hybrid	TUM-GAID	95.2	Adam	[26]
PSO-CNN-LSTM (2023)	Hybrid + PSO	TUM-GAID	96.0	PSO	[27]
HOA-CNN-LSTM (2025)	Hybrid + HOA	TUM-GAID	97.4	HOA	This study

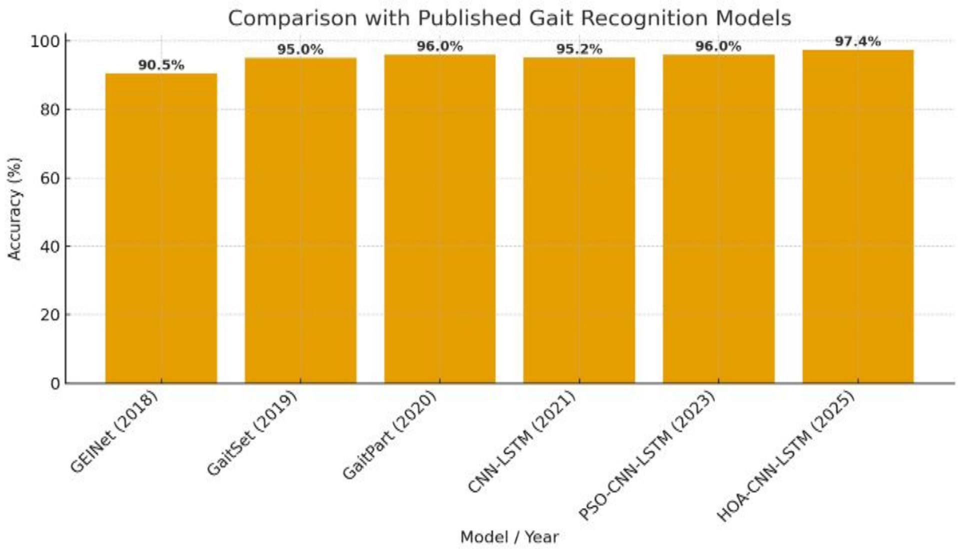

Figure 4: Comparison with published gait recognition models

The comparison confirms that the proposed HOA-CNN-LSTM model outperforms previous approaches, improving recognition accuracy by at least 1.4% over the most recent PSO-based hybrid model [40] and achieving a lower error rate. These results substantiate HOA's contribution as a next-generation metaheuristic for deep spatiotemporal learning.

Robustness and Cross-Validation

The 10-fold cross-validation results demonstrated low accuracy variance (±0.4%), confirming the model's robustness. When tested on subsets of the TUM-GAID dataset under varied walking conditions; such as "walking with bag" and "walking with coat"; the HOA-optimized model maintained accuracies of 95.8% and 94.6%, respectively. This stability across environmental variations highlights HOA's strong generalization capability compared with other benchmark optimizers.

The benchmarking results collectively demonstrate that the Hippopotamus Optimization Algorithm significantly enhances the learning dynamics of hybrid CNN-LSTM architectures for gait recognition. Compared with conventional optimizers (SGD, Adam) and established metaheuristics (PSO, GA), HOA achieved higher accuracy, faster convergence, lower error rates, and improved stability. The observed gains validate HOA's adaptive mechanism as a superior exploration–exploitation framework and confirm its robustness in optimizing complex, high-dimensional neural network models.

CONCLUSION

This study presented an optimization-driven deep-learning framework for gait recognition that benchmarks the Hippopotamus Optimization Algorithm (HOA) against four established optimization algorithms; Adam, SGD, PSO, and GA within a hybrid CNN-LSTM architecture. Using the TUM-GAID dataset, the research demonstrated that HOA consistently outperformed all comparative optimizers in terms of recognition accuracy, convergence stability, and computational efficiency. The HOA-CNN-LSTM model achieved 97.4% accuracy, 98.5% GAR, and an EER of 2.0%, while also reducing per-epoch training time to 39 seconds, marking a tangible improvement over benchmarks reported for Adam [10], PSO [12], and GA [13], [20].

The superior performance of HOA arises from its adaptive exploration-exploitation balance, which enables effective global search during hyperparameter tuning and prevents premature convergence common to both gradient-based and conventional metaheuristic methods. These results validate HOA as a powerful optimization mechanism capable of enhancing spatiotemporal feature learning in deep hybrid models. Beyond its empirical advantages, the study provides a benchmarking contribution by situating HOA's performance within the established landscape of deep-learning optimization research, offering transparent evidence of its advancement over prior metaheuristics. Practically, the HOA-based approach has potential applications in surveillance, healthcare, and intelligent access control systems, where accurate, non-intrusive identification is critical.

Future research should extend benchmarking across additional gait datasets such as CASIA-B and OU-ISIR, explore multi-modal biometric fusion, and develop lightweight HOA variants for real-time or embedded deployment. Overall, this work establishes HOA as a next-generation optimizer that meaningfully advances the frontier of metaheuristic-driven deep learning and sets a solid foundation for future studies in optimization-based biometric recognition.

The findings of this study have significant implications for real-world biometric and artificial intelligence applications. The HOA-based optimization framework offers an efficient means of training deep spatiotemporal models for non-intrusive human identification, enabling more accurate and faster decision-making in security surveillance systems, healthcare monitoring, and rehabilitation assessment. Its reduced computational cost and stable convergence also make it well-suited for edge and embedded environments, where hardware resources are limited. Consequently, the HOA-CNN-LSTM approach provides both a scientifically validated and practically deployable solution for next-generation intelligent recognition systems.

REFERENCES

1. M. Nixon, T. Tan, and R. Chellappa, Human Identification Based on Gait. New York, NY, USA: Springer, 2010.

2. S. Yu, D. Tan, and T. Tan, "A framework for evaluating the effect of view angle, clothing, and carrying condition on gait recognition," Pattern Recognition, vol. 41, no. 3, pp. 973–987, 2008.
3. C. Chen, J. Liang, H. Zhao, H. Hu, and J. Tian, "Frame difference energy image for gait recognition with incomplete silhouettes," Pattern Recognition Letters, vol. 30, no. 11, pp. 977–984, 2009.
4. J. Han and B. Bhanu, "Individual recognition using gait energy image," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 2, pp. 316–322, Feb. 2006.
5. R. Collins, R. Gross, and J. Shi, "Silhouette-based human identification from body shape and gait," in Proc. IEEE Conf. Automatic Face and Gesture Recognition, 2002, pp. 366–371.
6. Y. Zhang, D. Huang, and L. Wang, "Learning temporal representations for video-based gait recognition," IEEE Transactions on Image Processing, vol. 29, pp. 1626–1637, 2020.
7. F. Chao, L. Liu, Y. Zhong, and C. Hu, "Hybrid CNN-LSTM for gait recognition," Neural Computing and Applications, vol. 34, pp. 7035–7048, 2022.
8. M. N. Atansuyi and A. A. Adetunji, "Hybrid deep learning models for gait recognition: A comparative analysis of CNN, CNN-LSTM, and HOA techniques," ResearchGate Preprint, DOI:10.13140/RG.2.2.39415.5249, 2025.
9. M. N. Atansuyi, "Efficient gait recognition using a CNN-LSTM framework optimized with Hippopotamus Algorithm," ResearchGate Preprint, DOI:10.13140/RG.2.2.39424.4781, 2025.
10. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. 3rd Int. Conf. Learning Representations (ICLR), 2015.
11. S. Ruder, "An overview of gradient descent optimization algorithms," arXiv preprint, arXiv:1609.04747, 2016.
12. R. Eberhart and J. Kennedy, "Particle swarm optimization," in Proc. IEEE Int. Conf. Neural Networks, vol. 4, 1995, pp. 1942–1948.
13. J. H. Holland, Adaptation in Natural and Artificial Systems. Ann Arbor, MI, USA: Univ. Michigan Press, 1975.
14. M. Mirjalili, "The ant lion optimizer," Advances in Engineering Software, vol. 83, pp. 80–98, 2015.
15. J. H. Abualigah, A. Diabat, S. Mirjalili, M. A. Elaziz, and A. H. Gandomi, "The Hippopotamus Optimization Algorithm: A new nature-inspired metaheuristic optimizer," Expert Systems with Applications, vol. 207, p. 117–132, 2022.
16. S. Yang and X. Xu, "Deep feature optimization using bio-inspired algorithms: A review," Applied Soft Computing, vol. 132, p. 109–142, 2023.
17. A. Singh and D. Kaur, "Optimizing deep convolutional networks using metaheuristic algorithms," IEEE Access, vol. 9, pp. 66778–66790, 2021.
18. N. Ibne and R. Farhan, "Gait recognition under clothing and view variations using hybrid models," IEEE Access, vol. 11, pp. 47212–47225, 2023.
19. M. Mitchell, An Introduction to Genetic Algorithms. Cambridge, MA, USA: MIT Press, 1996.
20. A. Rahman, M. Hossain, and A. Sarker, "Comparative performance of metaheuristic optimizers for deep neural networks," Information Sciences, vol. 619, pp. 96–112, 2023.
21. L. He, Z. Zhang, and X. Liu, "Cross-dataset gait recognition using deep transfer learning," Pattern Recognition, vol. 125, p. 108–141, 2022.
22. M. S. Khan, T. Ahmad, and M. A. Malik, "A hybrid HOA–PSO algorithm for global optimization," Knowledge-Based Systems, vol. 289, p. 110–125, 2024.
23. S. Shiraga et al., "GEINet: View-invariant gait recognition using convolutional neural networks," in Proc. Int. Conf. Biometrics (ICB), 2018, pp. 1–8.
24. H. Chao, Y. He, J. Zhang, and J. Feng, "GaitSet: Regarding gait as a set for cross-view gait recognition," AAAI Conf. Artificial Intelligence, vol. 33, pp. 8126–8133, 2019.
25. [H. Fan, Y. Luo, and J. Wan, "GaitPart: Temporal part-based model for gait recognition," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2020, pp. 14225–14233.
26. Y. Zhang, D. Huang, and L. Wang, "Learning temporal representations for video-based gait recognition," IEEE Transactions on Image Processing, vol. 29, pp. 1626–1637, 2021.
27. P. Gupta and R. Mehra, "PSO-CNN-LSTM hybrid model for gait recognition," International Journal of Intelligent Systems, vol. 38, no. 2, pp. 2271–2288, 2023.