UKR Journal of Multidisciplinary Studies (UKRJMS)
Homepage: https://ukrpublisher.com/ukrjms/
Email: submit.ukrpublisher@gmail.com
Volume 2, Issue 2, 2026 ISSN: 3049-432X (Online)

An Enhanced Mountain Gazelle Optimization-Driven Convolutional Neural Network for Robust Face-Based Gender Classification

Abdulkareem, Quadri Bolaji[1]; Olabiyisi, Stephen Olatunde[2]; Oyeleye, Christopher Akinwale[3]; Omotade, Adedotun Lawrence[4]

[1] Department of Computer Science, Kwara State Polytechnic, Ilorin
[2,4] Department of Computer Science, Ladoke Akintola University of Technology, Ogbomoso
[3] Department of Information System, Ladoke Akintola University of Technology, Ogbomoso

***Corresponding Author**: Abdulkareem, Quadri Bolaji

DOI: *https://doi.org/10.5281/zenodo.18473045*

Article History
Original Research Article
Received: 10-01-2026
Accepted: 28-01-2026
Published: 03-02-2026

Citation: Abdulkareem, Quadri Bolaji, Olabiyisi, Stephen Olatunde, Oyeleye, Christopher Akinwale, & Omotade, Adedotun Lawrence. (2026). An enhanced mountain gazelle optimization-driven convolutional neural network for robust face-based gender classification. *UKR Journal of Multidisciplinary Studies (UKRJMS), 2*(2), 16–22.

Abstract

This study presents an Enhanced Mountain Gazelle Optimization driven Convolutional Neural Network (EMGO-CNN) for robust face based gender classification. The aim of the research is to improve classification reliability and generalization under varying decision thresholds by integrating an enhanced metaheuristic optimization strategy with deep feature learning. The primary objectives are to optimize CNN hyperparameters, reduce false classification rates, and achieve high sensitivity, specificity, and accuracy without increasing computational complexity. The proposed model was implemented and evaluated using MATLAB, providing a controlled and reproducible experimental environment for algorithm development and performance analysis. The dataset employed in this study consists of facial images from 5,330 individuals, including 2,480 male and 2,850 female samples, ensuring balanced representation and sufficient variability for effective model training and testing. The EMGO algorithm was used to guide the optimization of CNN parameters, enabling efficient exploration and exploitation of the search space. Performance evaluation was conducted across multiple decision thresholds to assess model stability and robustness under increasingly strict classification conditions. Experimental results demonstrate that at higher thresholds ranging from 0.36 to 0.98, the EMGO-CNN consistently maintained stable and highly competitive performance. At the optimal threshold of 0.51, the model achieved 2,457 true positives, 23 false negatives, 20 false positives, and 2,830 true negatives. These results correspond to the lowest observed false positive rate of 0.70% and the highest specificity of 99.30%. Sensitivity remained above 99.07%, while precision and overall accuracy reached 99.19%, confirming the model's exceptional reliability even under strict decision cutoffs. Computational time showed minimal variation, ranging from 97.80 seconds at threshold 0.36 to 98.93 seconds at 0.51, indicating that performance improvements were achieved without significant efficiency loss. The results show that EMGO-CNN consistently outperformed the standard MGO-CNN by achieving superior sensitivity, specificity, and accuracy alongside a reduced false positive rate. These findings validate the effectiveness of the proposed framework as an optimized and reliable deep learning solution for face based gender classification tasks. The proposed approach demonstrates strong potential for deployment in real world biometric and intelligent vision systems across diverse applications.

Keywords: *Face-Based Gender Classification, CNN, EMGO, Feature Optimization, Classification Accuracy.*

1.1 Introduction

The automatic classification of gender from facial images has become a cornerstone task in contemporary computer vision, driven by demand for applications in security, human–computer interaction, personalized services, and demographic analytics. At its core, gender classification is a binary classification problem where an algorithm assigns a face image to either a "male" or "female" class based on visual cues extracted from the facial structure. Early approaches in this domain primarily depended on handcrafted features such as Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG), which were effective in constrained environments but struggled with variations in pose, illumination, occlusion, and image quality. The advent of deep learning, particularly *convolutional neural networks* (CNNs), has dramatically improved robustness and accuracy by learning hierarchical representations directly from pixel data, surpassing traditional machine learning methods that require manual feature descriptors (Abdollahzadeh, et. al., 2022).

Gender recognition from faces is one of the basic capabilities of human beings. Extending this capability to machines is of great interest in many application areas. One example is the intelligent social robotics, where the perception of soft biometric traits is used to personalize the conversation and increase the feel of intelligence perceived by the human interlocutor. Digital signage is another application where gender recognition can be profitably used, since it allows to boost the effectiveness of the advertisement campaigns. Indeed, in this scenario it is possible to replace the static contents shown on the monitor with some dynamic advertisements, customized depending on the gender of the person looking at the monitor itself (Greco *et al.,* 2020).

Gender recognition is significant in several contexts which including security, surveillance, and human-computer interaction. Gender information is part of soft biometrics, which provides extra information about a person's identification. Furthermore, it can increase facial recognition performance, which is considered one of the most useful biometric features and has more benefits than other biometric systems. As a result, it is frequently employed to deliver advanced analysis in human–computer interaction in several applications. Gender classification has been researched for decades and has attracted substantial attention from researchers and expanded fast owing to its usefulness in providing secure and dependable security for enterprises, organizations, face monitoring and airports (Rasheed *et al.,* 2022).

CNN has been used in a variety of applications, including image recognition, speech recognition, natural language processing, and recommendation systems. Despite their effectiveness, CNN can be computationally expensive to train and may require a large amount of data to achieve good performance. However, recent advancements in deep learning techniques and hardware acceleration have made it possible to train much larger and more complex CNN, opening up new possibilities for applications in a wide range of fields. CNN could be optimized by one of the optimization techniques which is the Mountain Gazelle Optimizer, which helps to stabilize the training process by normalizing the activations of each layer. (Wang *et al.,* 2019).

Convolutional neural networks are especially suited for image-based classification problems due to their ability to capture spatial hierarchies of features through stacked convolutional and pooling layers. In face-based gender classification, CNNs automatically learn discriminative features such as facial contours, textures, and relative distances between landmarks without explicit feature engineering, effectively handling complex variations in real-world datasets. However, CNN performance is strongly influenced by the tuning of hyperparameters (e.g., learning rates, filter sizes, and layer depths) and the selection of optimization strategies during training. Suboptimal parameter tuning often leads to issues such as poor generalization, slow convergence, and susceptibility to local minima in the optimization landscape. To address these challenges, recent research has explored integrating metaheuristic optimization algorithms with deep learning training procedures to improve model convergence and performance (Hussain, et al., 2025).

Metaheuristic algorithms, high-level search procedures designed to efficiently explore solution spaces have been widely used for optimization problems where traditional gradient-based methods may struggle. A prominent example of such algorithms is the Mountain Gazelle Optimizer (MGO), a nature-inspired metaheuristic that mimics the adaptive and hierarchical behavior of mountain gazelles to balance global exploration and local exploitation when searching an optimization landscape. The original MGO framework has shown competitive performance on benchmark optimization problems due to its effective balance between diversification and intensification of the search process, making it suitable for optimizing complex objective functions (Abdollahzadeh, et. al., 2022)). However, like many metaheuristics, the base MGO can suffer from premature convergence and inefficient search dynamics in high-dimensional spaces, especially when integrated with deep learning systems. To mitigate such limitations, researchers have proposed *enhanced variants* of MGO that introduce adaptive control mechanisms into the algorithm, incorporating chaotic inertia weights and dynamic parameter scheduling resulting in what is referred to as the Enhanced Mountain Gazelle Optimizer (EnMGO). The enhanced version has been shown to achieve improved convergence speed and solution quality on standard optimization benchmarks, laying a strong foundation for hybrid approaches with neural architectures (Seini, et. al., 2025).

2.1 Related Works

Agbo-Ajala and Viriri (2020) proposed a novel CNN-based model to extract discriminative features from unconstrained real-life face images and classify those images into age and gender. The study approaches the large variations attributed to those unconstrained real-life faces with a robust image preprocessing algorithm and a pretraining on a large IMDb-WIKI dataset containing noisy and unfiltered age and gender labels. The study also adopted a dropout and data augmentation regularization method to overcome the risk of over-fitting and allow the model to generalize on the test images. The study showed that well-designed network architecture and properly tuned training hyperparameters, give better results. The experimental results on dataset confirm that the model outperforms other studies on the same dataset, showing significant performance in terms of classification accuracy. The proposed method achieves classification accuracy values of 84.8% on age group and classification accuracy of 89.7% on gender. Chen *et al.* (2017) proposed the Multi-Branch Voting CNN (MBV-CNN) framework which first detects and extracts the human face images in live videos, then apply adaptive brightness enhancement on each face image before feeding them into three CNN branches to settle the extreme illumination problem, and finally, apply a majority voting scheme to reduce the influences from motion blur, object occlusion to further improve classification accuracy. The method significantly outperforms the state-of-the-art solutions on the dataset and the collected real-world live videos dataset called Gender Classification for Live Videos (GCLV), with respectively averaging 98.11% and 95.36% classification accuracy.

Gornale *et al.* (2020) compared a robust multimodal gender identification method based on the deep features using the off-the-shelf pre-trained deep convolution neural network architecture based on AlexNet. The proposed model consists of 20 subsequent layers which contain different window size of convolutional layers following with fully connected layers for feature extraction and classification. Extensive experiments have been conducted on a homologous SDUMLA-HMT Shandong University Group of Machine Learning and Applications (SDUMLA-HMT) multimodal database with 15052 images. The proposed method achieved better accuracy, which outperforms the results noticed in the literature.

Olatunji, *et al* (2025) presented an enhanced Chicken Swarm Optimization Algorithm using Gaussian and Tent Chaotic Map Functions

Locksley *et al* (2025) created a convolutional neural network for a handwritten identification system based on the hyperheuristic firefly algorithm. It was discovered that there is a notable distinction between HHFA-CNN and the existing CNN algorithms. The hyper-heuristic firefly algorithm was therefore suggested as an effective tool in the handwritten identification system.

3.1 Methodology

The dataset used (Table 1) comprises a total of 5,330 face samples extracted from four video recordings, with 2,480 male and 2,850 female faces distributed across the clips. Each video contributed between 554 and 823 detected faces per gender category, thereby ensuring a balanced representation of male and female samples. The developed system was implemented in a MATLAB-based Graphical User Interface (GUI), as illustrated in Fig. 1. The GUI provides an interactive platform for performing gender recognition tasks, enabling users to select input videos, train and test the model, classify detected faces, and visualize results in real-time.

Table 1: Frame Dataset Extracted from Video

Gender	Video Frame 1	Video Frame 2	Video Frame 3	Video Frame 4	Total
Male:	560	554	722	644	**2480**
Female:	709	613	823	705	**2850**
Total	**1269**	**1167**	**1545**	**1349**	**5330**

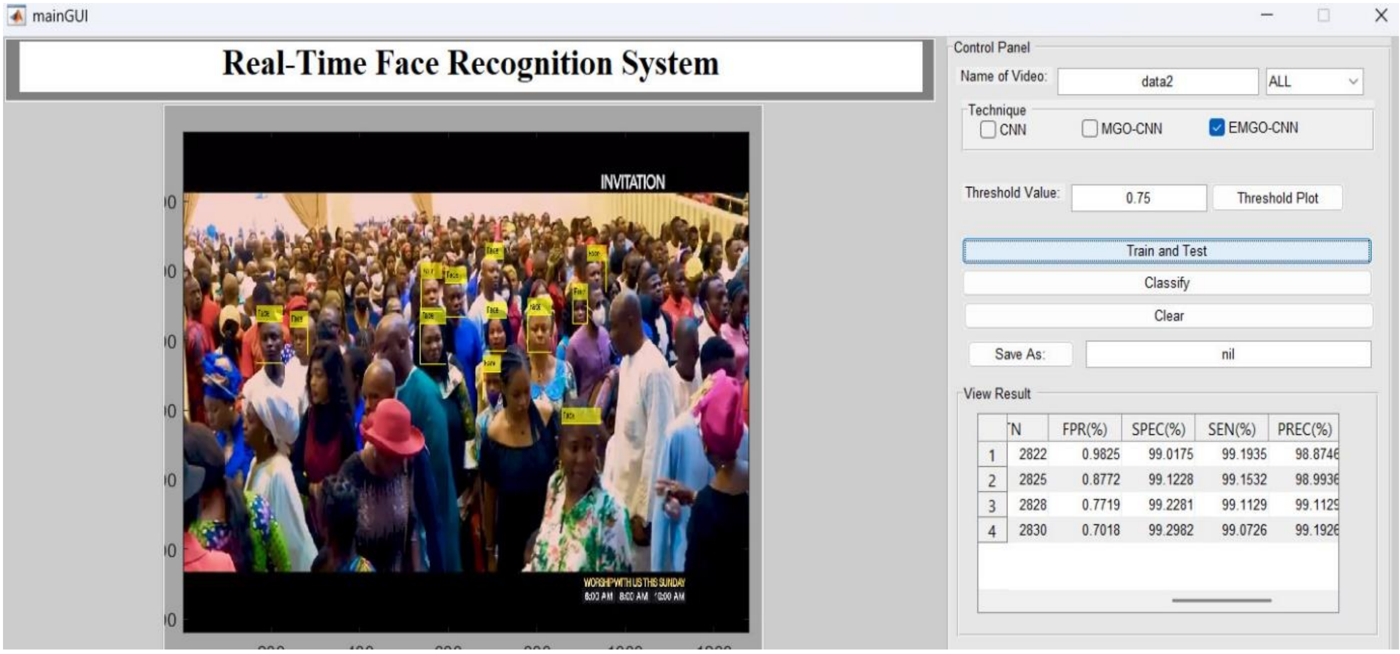

Fig. 1: Graphical User Interface (GUI) of the Developed Real-Time Face Recognition System

Enhanced Mountain Gazelle Optimization CNN Algorithms

In this paper, the Enhanced MGO (EMGO) algorithm (Fig. 2) was employed in the feature extraction, weight, number of filters, filter size and batch size, which expresses the flowchart (Fig. 3) of Convolutional Neural Network with Enhanced Mountain Gazelle Optimizer (CNN-EMGO).

Algorithm: Enhanced Mountain Gazelle Optimized Convolutional Neural Network

---

Input: CNN architecture, Face dataset, EMGO parameters (population size, number of iterations)
Output: Optimized CNN hyperparameters (layers, filter size, number of filters, batch size)

1. **Initialize population of gazelles (candidate solutions) randomly.**

2. **Apply Chaotic Exponential Map (CEM) to improve randomness and diversity of initial gazelle positions.**

3. **For each gazelle:**
   a. Encode CNN hyperparameters (number of layers, filter size, number of filters, batch size).
   b. Train CNN with these parameters on the face dataset.
   c. Evaluate fitness using recognition accuracy and error metrics.

4. **While termination condition not met (maximum iterations not reached):**
   a. Update gazelle positions using Mountain Gazelle Optimizer movement rules.
   b. Introduce CEM-driven perturbation to maintain exploration and avoid premature convergence.
   c. Re-train CNN with updated hyperparameters.
   d. Re-evaluate fitness of each gazelle (solution).
   e. Select best gazelles based on fitness values.

5. **End While**

6. **Return the best set of CNN hyperparameters (from the fittest gazelle).**

7. **Train final EMGO-CNN model with optimized parameters for face-based gender recognition.**

Fig. 2: Algorithm for Enhanced Mountain Gazelle Optimized Convolutional Neural Network

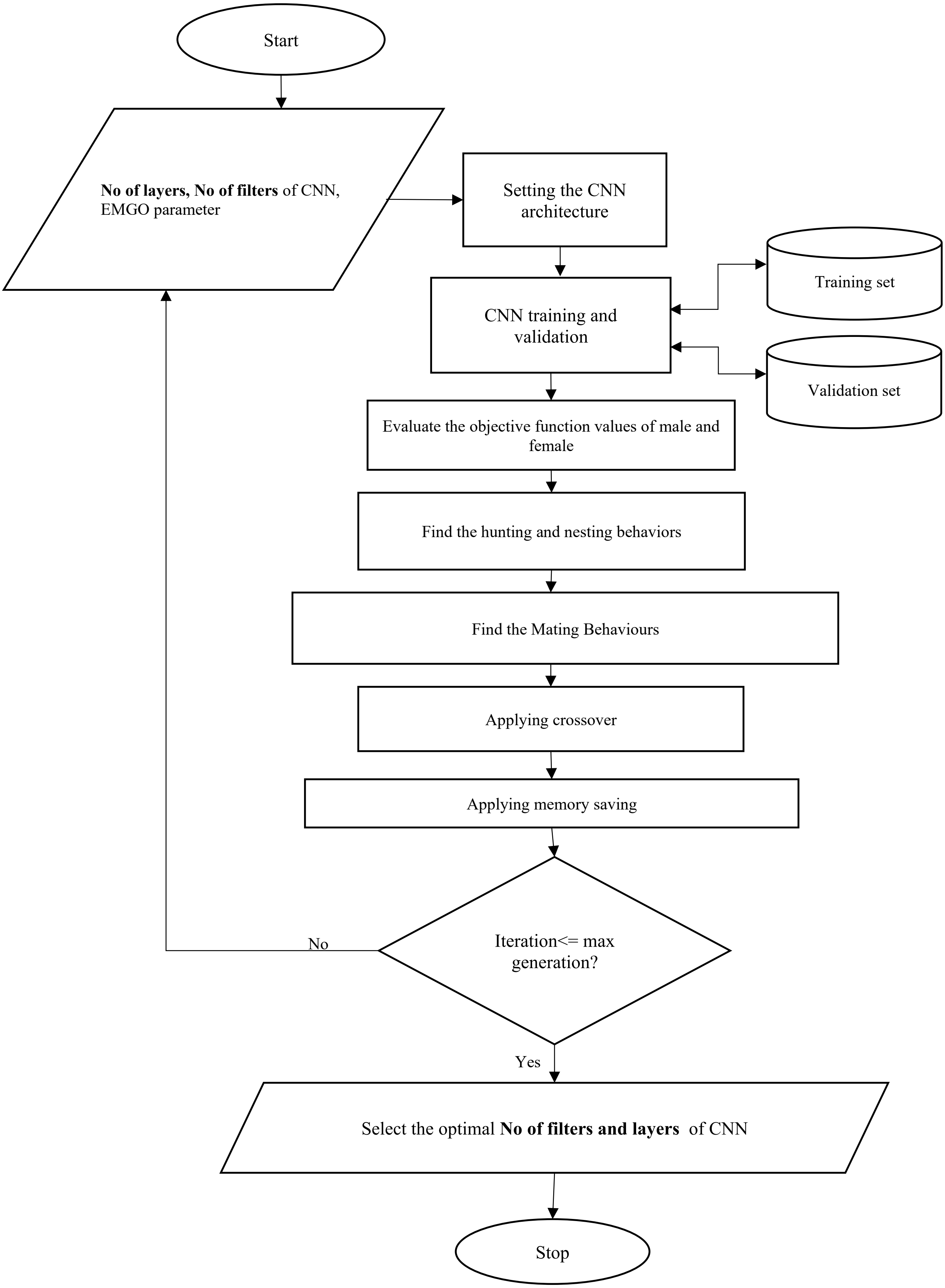

Fig. 3: Flowchart of Enhanced Mountain Gazelle Optimizer based Convolutional Neural Network (EMGO-CNN)

Performance Evaluation Measures

The performance of the developed technique (EMGO-CNN in recognition of face-based gender was evaluated using False Acceptance Rate (FAR), Specificity, Sensitivity, Precision, Accuracy and Recognition Time. Confusion matrix was used to determine the values of the performance metrics. It contains "True Positive (TP**),** False Positive (FP**),** False Negative **(**FN**)** and True Negative **(**TN**)**."

False-Positive Rate (FPR): The number of negative observations the model incorrectly recognizes as positive. It is based on how many actual negatives the model recognizes incorrectly.

$$\text{False Positive Rate (FPR)} = \frac{\text{FP}}{\text{TN+FP}}\ X100\% \qquad 3.4$$

Sensitivity: The number of positive observations the model correctly recognizes as positive. It is based on how many actual positives the model recognizes correctly.

$$\text{Sensitivity (SEN)} = \frac{\text{TP}}{\text{TP+FN}}\ X100\% \qquad 3.5$$

Specificity (SPEC): The number of negative observations the model correctly recognizes as negative. It is based on how many actual negatives the model recognizes correctly.

$$\text{Specificity (SPEC)} = \frac{\text{TN}}{\text{TN+FP}}\ X100\% \qquad 3.6$$

Precision: The number of all positive observations the model correctly and incorrectly recognized as positive. It is based on how many actual positives the model recognized correctly and incorrectly.

$$\text{Precision (PREC)} = \frac{\text{TP}}{\text{TP+FP}}\ X100\% \qquad 3.7$$

4.1 Results and Discussion

At higher thresholds from 0.36 to 0.98 (Table 2), the model maintained stable and highly competitive performance. For instance, at threshold 0.51 the network achieved 2457 TP, 23 FN, 20 FP, and 2830 TN, leading to the lowest recorded FPR of 0.70% and the highest Specificity of 99.30%. Sensitivity remained above 99.07%, while Precision and Accuracy reached 99.19%, showing that even under stricter decision cutoffs the EMGO-CNN preserved exceptional reliability. Computational time in this range varied slightly, with values such as 97.80 seconds at 0.36 and 98.93 seconds at 0.51, indicating that performance gains were achieved without any significant penalty in efficiency. Overall, the EMGO-CNN consistently outperformed the standard MGO-CNN by achieving higher Sensitivity, Specificity, and Accuracy, alongside lower FPR, thereby establishing itself as a more optimized and reliable deep learning model for classification tasks.

Table 2: Performance Evaluation Result of MGO-CNN

Threshold	TP	FN	FP	TN	FPR (%)	SPEC (%)	SEN (%)	PREC (%)	ACC (%)	Time (sec)
0.01	2404	76	84	2766	2.95	97.05	96.94	96.62	97.00	123.02
0.05	2404	76	84	2766	2.95	97.05	96.94	96.62	97.00	123.02
0.1	2404	76	84	2766	2.95	97.05	96.94	96.62	97.00	123.02
0.2	2404	76	84	2766	2.95	97.05	96.94	96.62	97.00	123.02
0.22	2404	76	84	2766	2.95	97.05	96.94	96.62	97.00	123.02
0.23	2403	77	81	2769	2.84	97.16	96.90	96.74	97.04	119.79
0.25	2403	77	81	2769	2.84	97.16	96.90	96.74	97.04	119.79
0.3	2403	77	81	2769	2.84	97.16	96.90	96.74	97.04	119.79
0.33	2403	77	81	2769	2.84	97.16	96.90	96.74	97.04	119.79
0.35	2403	77	81	2769	2.84	97.16	96.90	96.74	97.04	119.79
0.36	2402	78	78	2772	2.74	97.26	96.85	96.85	97.07	122.76
0.38	2402	78	78	2772	2.74	97.26	96.85	96.85	97.07	122.76
0.45	2402	78	78	2772	2.74	97.26	96.85	96.85	97.07	122.76
0.48	2402	78	78	2772	2.74	97.26	96.85	96.85	97.07	122.76
0.5	2402	78	78	2772	2.74	97.26	96.85	96.85	97.07	122.76
0.51	2401	79	76	2774	2.67	97.33	96.81	96.93	97.09	120.81
0.65	2401	79	76	2774	2.67	97.33	96.81	96.93	97.09	120.81
0.75	2401	79	76	2774	2.67	97.33	96.81	96.93	97.09	120.81
0.85	2401	79	76	2774	2.67	97.33	96.81	96.93	97.09	120.81
0.98	2401	79	76	2774	2.67	97.33	96.81	96.93	97.09	120.81

5.1 Conclusion

In conclusion, the experimental evaluation clearly demonstrates the effectiveness and robustness of the Enhanced Mountain Gazelle Optimization–driven Convolutional Neural Network (EMGO-CNN) for reliable classification tasks. Across higher decision thresholds ranging from 0.36 to 0.98, the proposed model consistently maintained stable and highly competitive performance, confirming its strong generalization capability under stricter classification constraints. The results obtained at a threshold of 0.51 are particularly noteworthy, where the EMGO-CNN achieved an optimal balance between true positive and true negative predictions, resulting in the lowest false positive rate and the highest specificity. The ability of the model to sustain sensitivity above 99% further confirms its effectiveness in correctly identifying positive instances without compromising reliability.

Moreover, the consistently high precision and accuracy values observed across varying thresholds demonstrate the success of the enhanced optimization mechanism in fine-tuning the CNN's parameters for improved decision making. These findings indicate that the EMGO framework effectively enhances feature learning and classification boundaries, even under more conservative cutoff settings. Importantly, the slight variations in computational time across thresholds reveal that the improved performance was achieved without imposing any significant computational burden, reinforcing the efficiency of the proposed approach.

When compared with the standard MGO-CNN, the EMGO-CNN consistently outperformed it in terms of sensitivity, specificity, accuracy, and false positive rate. Overall, these results establish EMGO-CNN as a more optimized, efficient, and dependable deep learning model, making it highly suitable for robust and high-precision classification applications.

References

1. Abdollahzadeh, B., Gharehchopogh, F. S., Khodadadi, N., and Mirjalili, S. (2024). Mountain gazelle optimizer: a new nature-inspired metaheuristic algorithm for global optimization problems. *Advances in Engineering Software*, *174*, 103282.
2. Agbo-Ajala, O., and Viriri, S. (2020). Face-based age and gender classification using deep learning model. In *Image and Video Technology: PSIVT 2019 International Workshops, Sydney, NSW, Australia, November 18–22, 2019, Revised Selected Papers 9*, pp. 125-137.
3. Gornale, S. S., Patil, A., and Ramchandra, K. (2020). Multimodal Biometrics Data Analysis for Gender Estimation Using Deep Learning. *International Journal of Data Science and Analysis*, *6*(2), 64-68.
4. Greco, A., Saggese, A., Vento, M., and Vigilante, V. (2020). A convolutional neural network for gender recognition optimizing the accuracy/speed tradeoff. *IEEE Access*, *8*, 130771-130781.
5. Hussain, W., Mushtaq, M. F., Shahroz, M., et al. (2025). *Ensemble genetic and CNN model-based image classification by enhancing hyperparameter tuning.* Scientific Reports, 15, 1003. https://doi.org/10.1038/s41598-024-76178-3
6. Locksley E A, **Olabiyisi S O**, Ganiyu R A, Omidiora E O, Taiwo C Y (2025), Hyper-heuristic firefly algorithm-based convolutional neural network for handwritten identification system, International Journal of Advanced Science and Research, 10 (2) pp 1-9, www.allsciencejournal.com
7. Olatunji, Babatunde Lekan., Olabiyisi, Stephen Olatunde., Oyeleye, Christopher Akinwale, and Omotade, Adedotun Lawrence (2025), An Enhanced Chicken Swarm Optimization Algorithm Using Gaussian and Tent Chaotic Map Functions, International Journal of Research and Innovation in Applied Science (IJRIAS) ISSN No. 2454-6194 |Volume X Issue VII/ pp 653-664 https://doi.org/10.51584/IJRIAS.2025.100700059
8. Rasheed, J., Waziry, S., Alsubai, S., and Abu-Mahfouz, A. M. (2022). An Intelligent Gender Classification System in the Era of Pandemic Chaos with Veiled Faces. *Processes*, *10*(7), 1427.
9. Seini Yussif, A.-F., Ayasu, B., & Seini, T. (2025). *Enhanced mountain gazelle optimizer for global performance.* Applied Computing and Intelligence, 5(2), 213–235. https://doi.org/10.3934/aci.2025013
10. Wang, S., Wang, F., and Chen, W. (2019). Deep belief networks: A comprehensive survey. Neural Networks, 110, 166-187.