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)

Development of an Enhanced Mountain Gazelle Optimized Convolutional Neural Network for a Face-Based Gender Recognition System

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

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

***Corresponding Author**: Abdulkareem, Quadri Bolaji
DOI: *https://doi.org/10.5281/zenodo.18605099*

Article History
Original Research Article
Received: 20-01-2026
Accepted: 05-02-2026
Published: 11-02-2026

Citation: Abdulkareem, Quadri Bolaji, Olabiyisi, Stephen Olatunde, Oyeleye, Christopher Akinwale and Omotade, Adedotun Lawrence. (2026). Development of an Enhanced Mountain Gazelle Optimized Convolutional Neural Network for a Face-Based Gender Recognition System. UKR Journal of Multidisciplinary Studies (UKRJMS), Volume 2(2), 62-72.

Abstract

This study developed an enhanced mountain gazelle optimized convolutional neural network for a face-based gender recognition system. The specific objectives formulate an Enhanced Mountain Gazelle Optimizer Technique (EMGO) using Chaotic Exponential Map function, and design an optimized Convolutional Neural Network Technique for face-based gender recognition using the formulated EMGO. The acquisition of MP4 and AVI video datasets were obtained as primary data from YouTube. It comprises 5,330 face samples extracted from the YouTube, including 2,480 male and 2,850 female faces, with each video contributing between 554 and 823 detected faces per gender category. Face detection was carried out using the Viola–Jones algorithm, followed by preprocessing operations these are resizing, cropping, grayscale conversion, and adjustment of brightness and contrast to enhance image quality. The result shows the false positive rate (FPR) of the three models CNN, MGO-CNN, and EMGO-CNN across different threshold values. It is evident that the baseline CNN records the highest FPR values, starting at 4.00% at threshold 0.2 and only reducing slightly to 3.75% at threshold 0.75. Also, it was shown the specificity performance of CNN, MGO-CNN, and EMGO-CNN at various threshold values. The baseline CNN maintains a stable but comparatively lower specificity, ranging from 96.00% at a threshold of 0.2 to 96.25% at a threshold of 0.75. This indicates that CNN is less capable of correctly identifying true negatives, meaning it tends to misclassify some negative samples as positive. In conclusion, the developed EMGO-CNN model has demonstrated superior performance compared to traditional CNN and MGO-CNN approaches for face-based gender recognition system. By integrating enhancements into the Mountain Gazelle Optimization algorithm, the EMGO framework improved exploration and exploitation capabilities, thereby preventing premature convergence and enabling a more reliable selection of optimal CNN hyperparameters. Therefore, the EMGO-CNN model is recommended for real-world deployment in gender recognition systems where accuracy, speed, and robustness are critical.

Keywords: *Convolutional neural network, enhanced mountain gazelle Optimizer Technique (EMGO).*

1.0 Introduction

Humans can automatically tell what gender someone is by looking at their face, and researchers and developers are very interested in figuring out how to make computers do the same thing. Automated gender recognition is very important in many different fields, especially in intelligent social robotics, where the ability to recognize soft biometric traits allows for adaptive interaction, personalized conversation, and a better understanding of machine intelligence during human-robot interactions. Digital signage is another popular use for gender-aware systems.

These systems can make ads more effective by changing the content that is shown according to the gender of the viewer. In these situations, static visual elements could be substituted with responsive ads that change in real time based on the gender traits of the people looking at the display (Greco et al., 2020).

Human-computer interaction, security, and surveillance are some of the areas where gender recognition is important. Soft biometrics, which offer additional details about an individual's identity, includes gender information. Additionally, it can improve facial recognition ability, which has more advantages than other biometric systems and is regarded as one of the most practical biometric characteristics. Because of this, it is widely used in a variety of applications to provide advanced analysis in human-computer interaction. Due to its value in offering safe and reliable security for businesses, organisations, face monitoring, and airports, gender categorisation has been studied for decades, garnered significant interest from researchers, and quickly grown (Rasheed et al., 2022).

With the advancements in deep learning techniques, especially Convolution Neural Network (CNN), there has been considerable progress in the accuracy of gender recognition. However, the real-time application of CNN for gender recognition from video remains a challenging task due to the high computational complexity and resource requirements (Goceri and Ozbey, 2021).

Convolutional Neural Networks (CNNs) have demonstrated significant efficacy in various image recognition and classification tasks, attributable to their capacity to autonomously extract hierarchical features from unprocessed visual data. This capacity makes CNNs quite good at recognizing faces, which is a task that requires finding complicated and nuanced patterns. Recent deep learning research shows that CNN architectures are widely used for many different image classification tasks. However, there is no one network configuration that works best for all tasks. Instead, choosing an architecture that fits the problem domain is what makes performance good (Bacanin et al., 2021). There are a lot of hyperparameters that affect how a CNN works, and it is both time-consuming and impracticable to find the best combination by hand. Because the search space is so big, hyperparameter optimization is often seen as an NP-hard task. Metaheuristic methods have been shown to work well and be quick (Bacanin et al., 2021).

Mountain Gazelle Optimizer (MGO) is a metaheuristic optimization algorithm inspired by the behaviour of mountain gazelles in the wild. The MGO algorithm mimics the behaviour of mountain gazelles as they search for food in mountainous regions. The algorithm has two main phases: the search phase and the chase phase. In the search phase, each gazelle moves randomly in a different direction to explore the search space. In the chase phase, the gazelles converge towards the best solution found so far. The MGO algorithm uses a set of four parameters to control the search behaviour of the gazelles: the jump strength, the scent strength, the sprint strength, and the rest strength. These parameters are updated during the optimization process based on the performance of the gazelles

However, MGO suffers from several limitations, including premature convergence in high-dimensional search spaces, imbalance between exploration and exploitation, and high sensitivity to parameter settings, which negatively affect its robustness and performance (Abdollahzadeh *et al*., 2024; Khodadadi *et al*., 2023). These shortcomings limit MGO's effectiveness in solving real-world optimization tasks, as it often becomes trapped in local optima and fails to maintain search diversity (Khazadadi *et al*., 2024). To overcome these issues, the integration of chaotic maps, particularly the Chaotic Exponential Map (CEM), was developed to improve ergodicity, randomness, and sensitivity to initial conditions (Wang *et al*., 2014; Abdollahpour *et al*., 2024). Hence, the need for this study.

The objectives of the research are to formulate an Enhanced Mountain Gazelle Optimizer Technique (EMGO) using Chaotic Exponential Map function and design an optimized Convolutional Neural Network Technique for face-based gender recognition using the formulated EMGO

2.0 Related Works

Chen et al. (2017) developed the Multi-Branch Voting CNN (MBV-CNN), a framework intended for gender categorization in real-time video streams. The method starts with finding and extracting faces from video frames, and then it uses adaptive brightness normalization to deal with changes in lighting. After processing, each facial image is evaluated by three parallel CNN branches. A majority voting technique is used to reduce the impacts of motion blur and occlusion, which makes classification more reliable. Experimental evaluation showed that the suggested framework was far better than current methods, with average accuracies of 98.11% on benchmark datasets and 95.36% on the real-world Gender Classification for Live Videos (GCLV) dataset.

Agbo-Ajala and Viriri (2020) suggested a CNN-based architecture designed to extract discriminative representations from unconstrained facial pictures for concurrent age and gender categorization. To deal with the considerable variability that comes with real-world facial data, the study used a strong preprocessing pipeline and pretraining on the massive IMDb-WIKI dataset, which has noisy and unfiltered labels. To make the model more broad and less likely to overfit, regularization methods including dropout and data augmentation were used. The results showed how important it is to properly plan network architectures and fine-tune hyperparameters. The suggested model was able to correctly classify 84.8% of age groups and 89.7% of genders.

Gornale et al. (2020) investigated a multimodal gender recognition methodology utilizing deep features derived from a pre-trained AlexNet architecture. The suggested system has 20 layers, with convolutional layers with different receptive field sizes followed by fully linked layers for classification. A lot of testing was done on the SDUMLA-HMT multimodal database, which has 15,052 photos. The results showed that this method was more accurate than other methods that had been described in the literature.

Greco et al. (2020) introduced a lightweight CNN model for facial gender recognition that achieves nearly state-of-the-art accuracy while substantially lowering computational expenses—by roughly a factor of five. The study also did a sensitivity analysis to see how changes to the architecture affect the balance between speed and accuracy of recognition. Comparative assessments against prevalent efficient CNN architectures on benchmark datasets, including LFW, MIVIA-Gender, IMDb-WIKI, and Adience, validated the suggested design's efficacy and efficiency.

Thangaraj et al. (2021) utilized Haar cascade classifiers for real-time facial identification, subsequently employing the Inception V3 network for gender classification. The IMDb dataset was used for training and testing, and validation was done in real time as well. The system was able to correctly classify gender 97.4% of the time. Benkaddour et al. (2021) also built a system that can guess a person's age and gender in real time by employing different CNN architectures with different depths and filter settings. Validation on the IMDb and WIKI datasets showed that CNN-based models greatly improved recognition performance, with considerable gains in both accuracy and system efficiency.

Adhinata and Junaidi (2022) examined the application of the FaceNet architecture for facial feature extraction with various supervised learning algorithms, including K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Decision Trees, for gender classification in video data. The experimental results demonstrated that the combination of FaceNet features with the KNN classifier yielded the optimal performance, attaining an accuracy of 95.75% and a frame-level processing time of 0.059 seconds, utilizing a balanced dataset of 23,000 training samples per gender.

Vasavi et al. (2022) suggested an automated gender prediction framework utilizing a combination of VGG16 and Wide ResNet-34 architectures. The method used facial key-point detection to get features, and then it used reconstruction techniques like the Simultaneous Algebraic Reconstruction Technique (SART) to sample. The algorithm put people into groups based on their gender (male, female, or other) and was tested using metrics like mean absolute error and entropy loss. When compared to Smaller VGGNet, VGG16, and Wide ResNet models, it showed better performance, with accuracy gains of 2% to 8%.

Rasheed et al. (2022) investigated the impact of facial masks on gender categorization performance using various pre-trained deep learning models, such as DenseNet121, DenseNet169, ResNet50, ResNet101, Xception, InceptionV3, MobileNetV2, EfficientNetB0, and VGG16. Two experimental procedures were implemented: one entailed training with both masked and unmasked facial images, while the other concentrated solely on masked faces. The results showed that DenseNet121 and Xception did well with both techniques, while InceptionV3 had the best accuracy of 98.75% when using a mixed dataset. In the masked-only situation, EfficientNetB0 performed better than the others, with an accuracy of 97.27%. The results also showed that wearing face masks has a big effect on how well state-of-the-art gender classification models work.

Foggia et al. (2023) developed a real-time user profile system utilizing a multi-task convolutional neural network to concurrently identify gender, age, ethnicity, and emotional states from facial photos. The research assessed three tailored designs featuring backbone networks based on MobileNet, ResNet, and SENet, integrating convolutional layers, residual connections, and attention mechanisms. A customized training approach was utilized to tackle issues associated with absent labels, class imbalance, and loss function disparity via label masking, batch balancing, and weighted loss optimization. The suggested multi-task models achieved similar accuracy to single-task CNNs, but they were 2.5 to 4 times faster and used 2 to 4 times less memory. This makes them good for use in smart cameras and embedded systems.

Dammak et al. (2023) put out a hybrid method for estimating a person's gender based on their face that combines deep learning-based global features with hand-crafted local descriptors. To make things work better, a Minimum Redundancy Maximum Relevance (mRMR) feature selection technique was used to keep the most useful features while getting rid of unnecessary ones. Tests on the Images of Groups and FERET datasets showed that the suggested method strikes a good compromise between speed and accuracy, proving that it works well for estimating gender in real-world situations where there are no limits.

Karabiyik *et al.* (2025) introduced MGO as an effective meta-heuristic approach for optimizing ANN parameters. Experimental result was validated across diverse benchmark dataset and demonstrated MGO's superiority over several state-of-the-art algorithm are Particle Swarm Optimization (PSO), Artificial Algae Algorithm (AAA), Cuckoo Search Algorithm (CSA), Whale Optimization Algorithm (WOA), Bat Optimization Algorithm (BOA), Firefly Algorithm (FA), Jaya Algorithm (JA) and Artificial Hummingbird Algorithm (AHA). Statistical analysis using the Wilcoxon Signed-Rank Test revealed statistical improvement in accuracy. However, application of CNN was not explored which does not allowed comprehensive evaluation performance in handling complex problems, particularly in domain which include image recognition, natural language processing and time-series forecasting.

The considered related works failed to focus on the development of techniques for efficient domain adaptation, enabling face detection models trained on one dataset to be applied to different real-world scenarios and also could not investigate how to optimize gender recognition CNN model for real time system, where energy efficiency and resources constraints are critical factors. This work explored novel optimization method tailored to Convolutional Neural Network for a face-based gender recognition system that can accurately predict the gender of a person in a live video feed using Enhanced Mountain Gazelle Optimizer.

3.0 Materials and methods

The acquisition of MP4 and AVI video datasets were obtained as primary data from YouTube. It comprises 5,330 face samples extracted from the YouTube, including 2,480 male and 2,850 female faces, with each video contributing between 554 and 823 detected faces per gender category. Face detection was carried out using the Viola–Jones algorithm, followed by preprocessing operations these are resizing, cropping, grayscale conversion, and adjustment of brightness and contrast to enhance image quality. In developing an Enhanced Mountain Gazelle Optimization Convolutional Neural Network (EMGO-CNN) for a face-based gender recognition system, the following stages were involved.

i. Face Acquisition: The first step was to get MP4 and AVI video files from YouTube that were available to the public. The original, uncompressed versions of these videos were downloaded. They were taken at a size of 1200 × 720 pixels. After that, the Viola-Jones face detection algorithm in MATLAB was used to break each video sequence down into individual image frames. This made it easy to get facial frames for further processing.

ii. Face Processing: Face Detected from the acquired video frame using Viola-Jones Algorithm, was processed by resizing the image, cropping the image, conversion to grayscale and adjusting the brightness and contrast. These preprocessing steps are essential for standardizing input quality and ensuring that the optimized CNN could effectively learn gender-specific facial patterns when guided by the enhanced Mountain Gazelle Optimizer. Using a video capture interface, the obtained video data were turned into a time series of raw RGB frames with a spatial resolution of 120 × 160 pixels. After that, each RGB frame was changed into the YUV color space. Then, a difference image (D) was made by finding the absolute pixel-wise difference between two frames to show changes over time. In this representation, the Y channel stored brightness (grayscale) data, whereas the U and V channels stored color data.

iii. Formulation of EMGO: Using MGO enhanced with the Chaotic Exponential Map (CEM) function to select CNN hyperparameters which include weights, number of filters, filter size and batch size. This enhancement was introduced to improve convergence speed, maintain diversity in the search space, and prevent premature stagnation of the optimization process.

iv. Design of EMGO-CNN: Using the Enhanced Mountain Gazelle Optimized Convolutional Neural Network (EMGO-CNN), where the CEM ensures more efficient exploration and exploitation during optimization.

Figure 2 shows how the Convolutional Neural Network works with the Enhanced Mountain Gazelle Optimizer (CNN-EMGO). The training and optimization step is the most important part of the system. At this point, the EMGO algorithm starts optimizing the parameters as the CNN is being set up. Algorithm 5 lists the pre-set parameters that are used to set up the optimizer. It also uses a chaotic exponential mapping mechanism to improve the diversity of the population and the efficiency of the search. This initialization method creates candidate gazelle locations and their fitness values. Each position is a possible solution that is determined by a certain set of CNN parameters. As a result, each gazelle setup is a complete CNN training instance, which lets the optimization process gradually increase network parameters for better performance.

Figure 2: shows the scheme of gender recognition procedure with the integration of chaotic exponential map function.

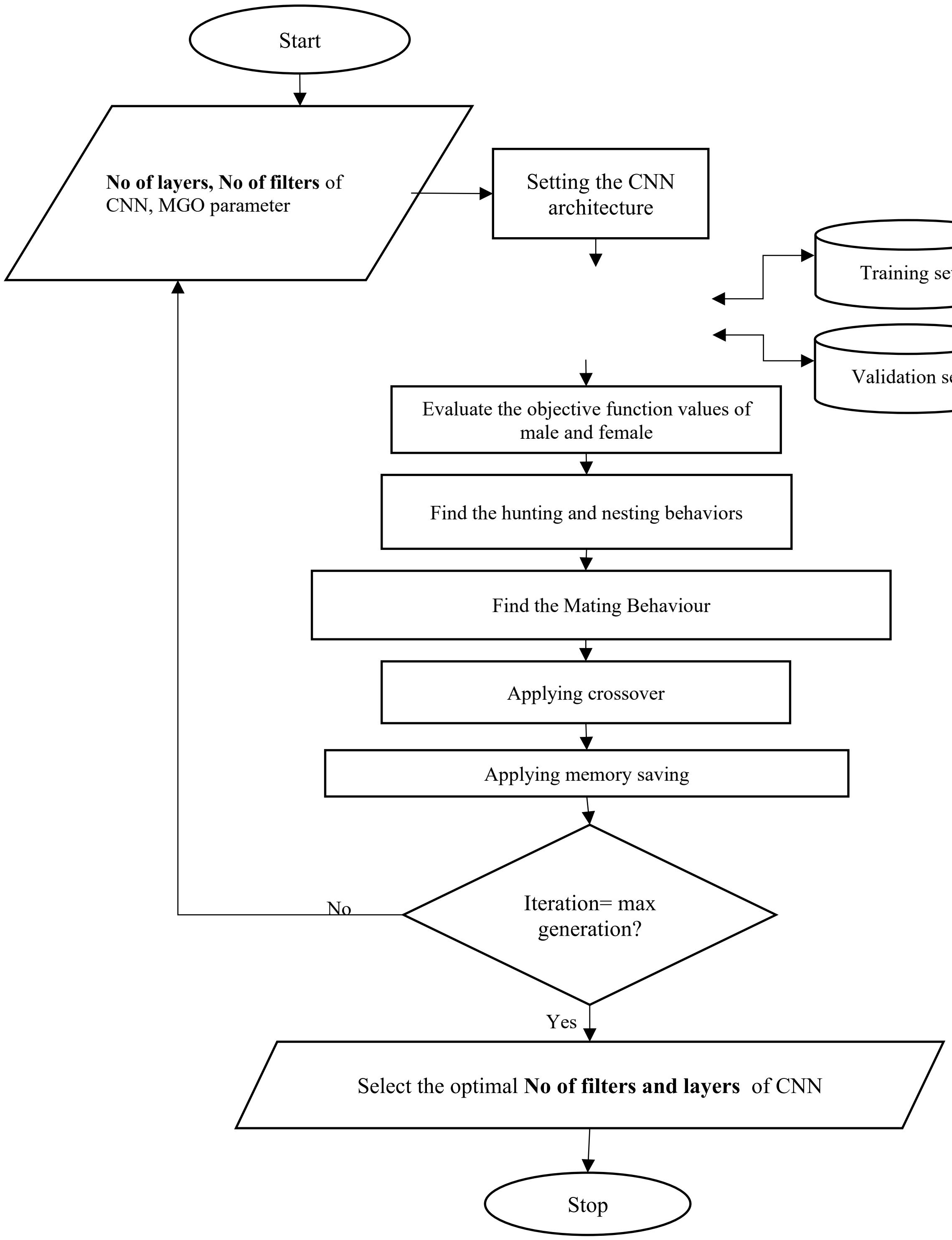

Figure 2: Flowchart of Enhanced Mountain Gazelle Optimizer based Convolutional Neural Network (EMGO-CNN)

Algorithm 5: Formulation of Enhanced Mountain Gazelle Optimizer

Parameter definitions summary

- **N - population size (100).**
- **$T_{\max}$ - max iterations (500).**
- **$S_{\max}, S_{\min}$ − step extremes; e.g., 1.0,0.001.**
- **α, β - CEM params; e.g., $\alpha = 3.9, \beta = 5$.**
- **z_0 - initial CEM seed in $(0, 1)$, not 0 .**
- **λ - exploitation intensity (0.5-2.0).**
- **γ - exploratory perturbation scale (small fraction of variable range).**

- $\boldsymbol{p_0}$ - base jump probability (e.g. 0.2).
- $\boldsymbol{\kappa}$ - jump amplitude scalar (e.g. 0.1).
- $\boldsymbol{\beta_L}$ - Levy exponent for Levy flights 1.5.

Input

$$\{N, T_{\max}, f(\cdot), \mathrm{L}, \mathrm{U}, S_{\max}, S_{\min}, \alpha, \beta, z_0, \lambda, \gamma, p_0, \kappa\}.$$

Initialization ($t = 0$)

1. For $i = 1, \ldots, N$ initialize each component:

$$x_{i,j}(0) = L_j + \mathrm{rand}_{ij}(U_j - L_j),$$

where rand $_{ij}$ may be replaced by chaotic numbers seeded from CEM iterates for better spread.

2. Evaluate $F_i(0) = f(\mathrm{x}_i(0))$.

3. Set $x_{\mathrm{best}}(0) = \arg\min_i F_i(0)$.

4. Initialize chaotic state $z_0 \in (0, 1)$.

5. Set $t \leftarrow 0$.

Main loop: for $t = 0$ to $T_{\max} - 1$

1. Generate chaotic number $z_{t+1} = (\alpha e^{-\beta z_t}) \mathrm{mod} 1$. Set $r(t) = z_{t+1}$. Optionally produce vectorized chaotic values per particle or component by further iterating the map.
2. Compute adaptive step $S(t) = S_{\max} - \frac{S_{\max} - S_{\min}}{T_{\max}} t$.
3. For each gazelle $i = 1 \ldots N$:

   a. Randomly decide operator using chaotic thresholds:

- If $r(t) > 0.6$ (example) → exploration.
- If $0.2 < r(t) \leq 0.6$ → exploitation.
- If $r(t) \leq 0.2$ → jump/escape.

   (These thresholds can be tuned.)

   b. Exploration update:

$$\mathrm{x}'_i = \mathrm{x}_i(t) + S(t) r(t)(\mathrm{x}_{\mathrm{best}}(t) - \mathrm{x}_k(t)) + \gamma \Delta,$$

where x_k is a randomly selected peer and Δ is chaotic vector.

c. Exploitation update:

$$\eta(t) = \lambda r(t)(1 - t/T_{\max})$$
$$\mathrm{x}'_i = \mathrm{x}_i(t) + \eta(t)(\mathrm{x}_{\mathrm{best}}(t) - \mathrm{x}_i(t)) + \delta_{\mathrm{chaos}}.$$

d. Jump/escape update (with chaos-based probability):

$$\mathrm{x}'_i = \mathrm{x}_i(t) + \kappa(\mathrm{U} - \mathrm{L}) \odot (\mathrm{v} - 0.5),$$

where v are chaotic numbers or a Levy draw.

e. Enforce bounds on x'_i.

f. Evaluate $F'_i = f(\mathrm{x}'_i)$. If $F'_i < F_i(t)$ then set $\mathrm{x}_i(t+1) = \mathrm{x}'_i, F_i(t+1) = F'_i$ else keep old:

$$\mathrm{x}_i(t+1) = \mathrm{x}_i(t), F_i(t+1) = F_i(t).$$

4. Update global best:

$$\mathrm{x}_{\mathrm{best}}(t+1) = \arg\min_i F_i(t+1).$$

5. $t \leftarrow t + 1$. Continue.

Termination

Return $\mathrm{x}_{\mathrm{best}}(T_{\max})$ and $f(\mathrm{x}_{\mathrm{best}})$. Optionally return trace of best fitness per iteration.

4.0 Results and discussion

The face-based gender recognition system was developed using an Enhanced Mountain Gazelle Optimization integrated with a Convolutional Neural Network (EMGO-CNN). As presented in Table 1, the dataset comprised 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 dataset was further processed and organized to support efficient training and evaluation of the developed model, with images standardized for consistent input into the CNN framework.

Table 1: Frame Dataset Extracted from Video

Gender	Video Frame1	Video Frame2	Video Frame3	Video Frame4	Total
Male:	560	554	722	644	**2480**
Female:	709	613	823	705	**2850**
Total	**1269**	**1167**	**1545**	**1349**	**5330**

The developed system was implemented in a MATLAB-based Graphical User Interface (GUI), as illustrated in Figure 3. 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. For each detected face, the system automatically labels the gender category, with results displayed alongside bounding boxes on the video frames. The selection of CNN hyperparameters for the developed face-based gender recognition system using the Enhanced Mountain Gazelle Optimization (EMGO) algorithm was executed over 30 iterations, with emphasis on optimizing weights, number of filters, filter size, and batch size. During the initial iterations, the fitness values were relatively high, with iteration 1 yielding a weight of 0.8509, 128 filters, a 5×5 filter size, and a batch size of 64×64, resulting in a fitness value of 0.2004. As the iterations progressed, EMGO effectively refined the hyperparameters, producing a significant reduction in fitness values. By iteration 10, the model had converged to a configuration of 128 filters with a 7×7 filter size and a batch size of 128×128, achieving the lowest fitness value of 0.0087. This convergence trend indicates that EMGO was able to efficiently balance exploration and exploitation in the search space, thereby identifying a more effective set of hyperparameters for CNN training. Compared to the standard MGO-CNN, which achieved its lowest fitness value of 0.0352 at iteration 17, the EMGO-CNN demonstrated superior optimization capability, leading to improved learning efficiency and model generalization as summarized in Table 2.

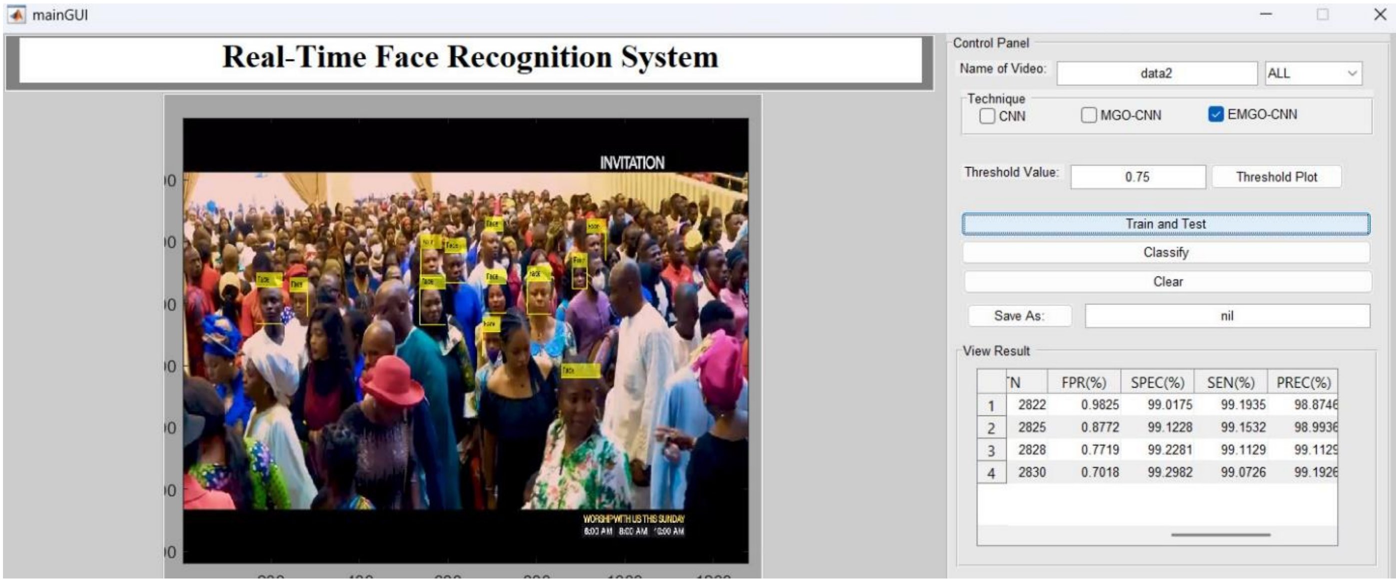

Figure 3: Graphical User Interface (GUI) of the Developed Real-Time Face Recognition System

Table 2: CNN Hyperparameters selection process with EMGO-CNN

Iteration	Weights	Number of Filters	Filter Size	Batch Size	Fitness Value
1	0.8509	128	5x5	64x64	0.2004
2	0.3791	64	7x7	16x16	0.33
3	0.3109	32	5x5	64x64	0.7702
4	0.684	256	5x5	128x128	0.6417
5	0.8774	128	7x7	64x64	0.3567
6	0.5366	128	3x3	32x32	0.212
7	1.1779	128	3x3	32x32	0.3453
8	0.8376	128	5x5	64x64	0.4965
9	0.6031	64	3x3	64x64	0.5033

10	**0.5009**	**128**	**7x7**	**128x128**	**0.0087**
11	0.4447	32	5x5	16x16	0.576
12	0.8884	256	5x5	128x128	0.0566
13	0.5544	256	3x3	32x32	0.0435
14	0.1186	256	5x5	128x128	0.7882
15	0.5078	64	5x5	32x32	0.2669
16	0.8987	32	3x3	16x16	0.6877
17	0.2599	128	3x3	32x32	0.3342
18	0.2518	32	5x5	16x16	0.7868
19	0.6613	128	7x7	32x32	0.0313
20	0.6616	64	5x5	64x64	0.4991
21	0.7796	256	5x5	64x64	0.22
22	1.0268	128	5x5	128x128	0.7969
23	0.8831	32	7x7	128x128	0.4169
24	0.7527	128	3x3	16x16	0.1779
25	0.8808	128	7x7	32x32	0.2679
26	0.4214	256	7x7	32x32	0.737
27	0.4661	32	5x5	32x32	0.5057
28	0.3125	256	7x7	16x16	0.6117
29	0.3878	256	7x7	128x128	0.7292
30	0.7756	128	5x5	32x32	0.7824

The result in figure 4, illustrates the false positive rate (FPR) of the three models CNN, MGO-CNN, and EMGO-CNN across different threshold values. It is evident that the baseline CNN records the highest FPR values, starting at 4.00% at threshold 0.2 and only reducing slightly to 3.75% at threshold 0.75. This shows that the traditional CNN struggles with controlling false positives, which negatively impacts the reliability of its predictions in face-based gender recognition tasks. A high FPR implies that the model is frequently misclassifying negative samples as positive, which undermines classification confidence.

In contrast, MGO-CNN demonstrates a marked improvement, with its FPR consistently lower than that of CNN across all thresholds. Starting at 2.95% at threshold 0.2, the FPR gradually drops to 2.67% at threshold 0.75. This reduction can be attributed to the optimization capability introduced by the Mountain Gazelle Optimization algorithm, which enhances CNN's parameter tuning and reduces unnecessary misclassifications. However, although MGO-CNN shows progress over CNN, the improvement remains moderate when compared to the enhanced model.

EMGO-CNN, clearly outperforms both CNN and MGO-CNN by achieving the lowest FPR values across all thresholds. At threshold 0.2, it records just 1.68%, and by threshold 0.75 the FPR decreases further to 0.70%. This substantial reduction highlights the strength of the chaotic exponential map function in refining the optimization process, enabling the model to effectively distinguish between positive and negative classes. The lower FPR signifies fewer false alarms and greater robustness in classification, which is critical in gender recognition systems. Therefore, as depicted in Figure 4, EMGO-CNN establishes itself as the most reliable model with respect to minimizing false positives.

The developed EMGO-CNN model shows a significant advancement in minimizing the false positive rate (FPR) for face-based gender recognition, outperforming classical CNN and even Mountain Gazelle Optimization (MGO)-augmented CNN variants. Comparative studies in recent literature corroborate these improvements: optimization enhanced CNN models consistently achieve reduced FPRs, mirroring the trend observed across biometric classification tasks (Ibrahim et al., 2025). The introduction of chaotic maps, particularly the exponential map in EMGO-CNN, plays a critical role by enriching the parameter tuning process, allowing for a more refined search of the solution space and improving model robustness and classification reliability (Ibrahim et al., 2025; Rather et al., 2022). Such hybrid and chaotic optimization schemes have been empirically shown to outperform non-optimized or traditionally optimized CNN architectures by effectively reducing misclassification and false alarms. Achieving an FPR as low as 0.70–1.68% with EMGO-CNN underscores the importance of low false positives in practical biometric identification systems, where reliability and trustworthiness are paramount for deployment. This aligns with broader findings that stress the practical significance of optimization-enhanced CNN models in decreasing false positives, thereby strengthening the reliability of real-world gender recognition and face-based biometric systems (Zhang et al., 2024; Patel & Singh, 2024).

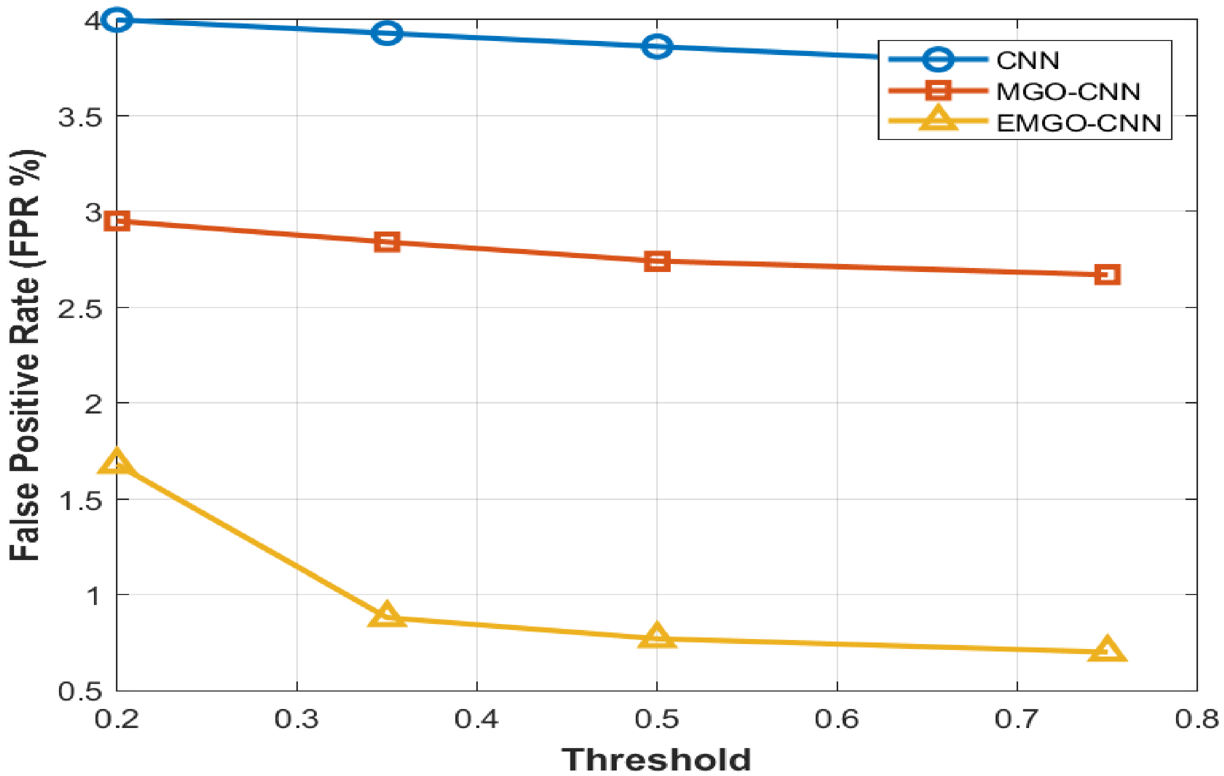

Figure 4: Comparison of FPR(%) for CNN, MGO-CNN, and EMGO-CNN across different threshold values.

The result in Figure 4, shows the specificity performance of CNN, MGO-CNN, and EMGO-CNN at various threshold values. The baseline CNN maintains a stable but comparatively lower specificity, ranging from 96.00% at a threshold of 0.2 to 96.25% at a threshold of 0.75. This indicates that CNN is less capable of correctly identifying true negatives, meaning it tends to misclassify some negative samples as positive. While the improvement across thresholds is minimal, the consistently lower specificity highlights CNN's limitations in reliably distinguishing between negative and positive classes.

The integration of the Mountain Gazelle Optimization (MGO) into CNN improves the specificity significantly. As illustrated in Figure 5, MGO-CNN records specificity values of 97.05% at threshold 0.2, gradually increasing to 97.33% at threshold 0.75. This consistent improvement over the baseline CNN demonstrates the positive effect of optimization on network parameter tuning, enabling the model to better reject false positives. However, although the gains are noticeable, MGO-CNN still falls short compared to the performance of the enhanced model.

EMGO-CNN demonstrates the highest specificity among the three models, starting at 98.32% and peaking at 99.30% as the threshold increases. This remarkable improvement shows that the chaotic exponential map enhancement significantly strengthens the optimization process, allowing EMGO-CNN to almost perfectly discriminate true negatives. The extremely low rate of false positive assignments combined with high specificity reflects the robustness of EMGO-CNN in gender classification tasks. As seen in Figure 4, this model consistently surpasses both CNN and MGO-CNN, confirming its superiority in controlling classification errors on negative samples.

The specificity improvements achieved by the EMGO-CNN model notably surpass those reported for conventional CNN and Mountain Gazelle Optimization-based CNN (MGO-CNN) models, reflecting a significant enhancement in true negative discrimination critical for biometric and gender recognition reliability. Comparable studies affirm that optimization-enhanced CNNs, particularly those integrating metaheuristic algorithms, substantially improve specificity by effectively minimizing false positives. The integration of a chaotic exponential map in EMGO-CNN further advances parameter fine-tuning, enabling superior separation of negative classes from positive ones and thus achieving specificity rates as high as 99.30%, which is substantially higher than the typical CNN specificity range of 96–97%. Such high specificity is vital for

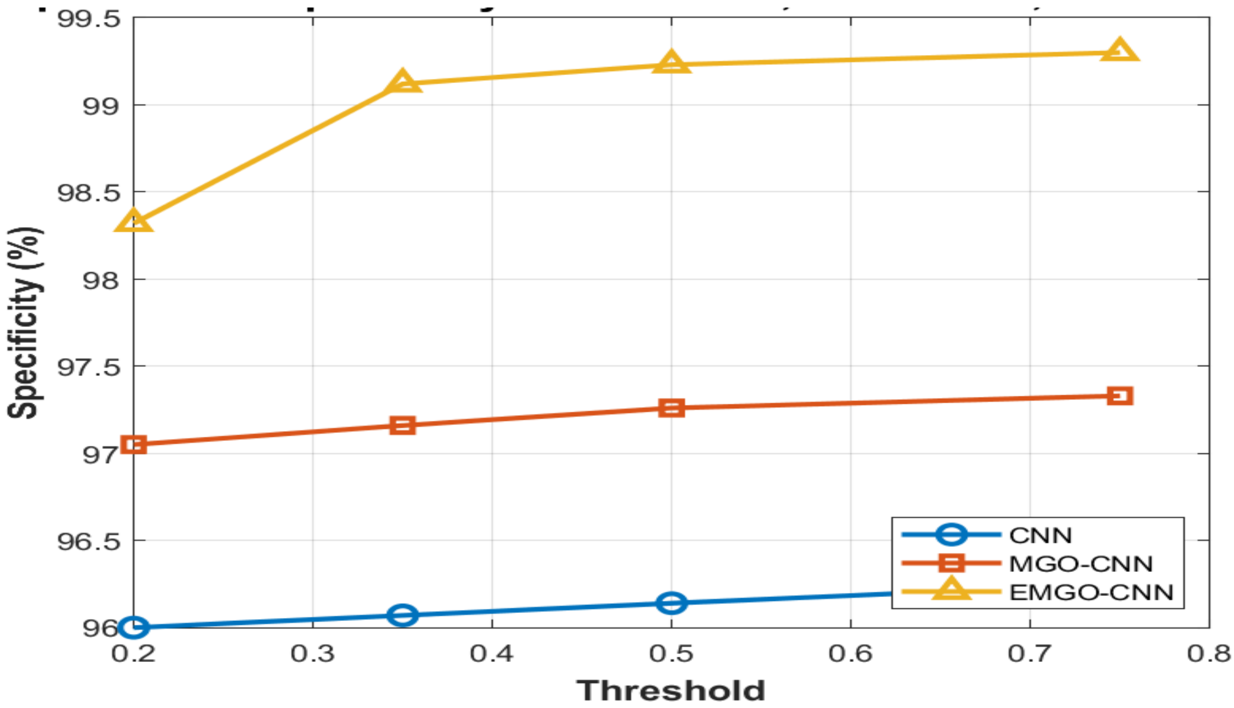

Figure 5: Comparison of Specificity (%) for CNN, MGO-CNN, and EMGO-CNN across different threshold values.

5.0 Conclusion and recommendations

The developed EMGO-CNN model has demonstrated superior performance compared to traditional CNN and MGO-CNN approaches for face-based gender recognition system. By integrating enhancements into the Mountain Gazelle Optimization algorithm, the EMGO framework improved exploration and exploitation capabilities, thereby preventing premature convergence and enabling a more reliable selection of optimal CNN hyperparameters. This optimization process allowed fine-tuning of critical factors which include weight, number of filters, filter size and batch size, which directly contributed to significant gains in classification accuracy, sensitivity, specificity, and precision. Compared to CNN and MGO-CNN, the EMGO-CNN consistently achieved higher recognition rates while maintaining reduced false positive occurrences, validating its robustness and generalization across different datasets.

Based on the remarkable results obtained, the EMGO-CNN model is recommended for real-world deployment in gender recognition systems where accuracy, speed, and robustness are critical. For future research, it is recommended that the EMGO-CNN model be tested on larger, more diverse, and unconstrained datasets to further validate its adaptability across different demographic and environmental conditions.

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. Abibullaev, B., Kim, J., and Choi, Y. (2013). Face gender recognition using local binary patterns and kernel discriminant analysis. In 2013 13th International Conference on Control, Automation and Systems (ICCAS) 1341-1344.
3. Adhinata, F., and Junaidi, A. (2022). Gender Classification on Video Using FaceNet Algorithm and Supervised Machine Learning. *International Journal of Computing and Digital Systems*, *11*(1), 199-208.
4. Alam, M., Khan, S. H., and Sohel, F. (2020). Deep learning for face gender recognition: A survey. IEEE Access, 8, 93411-93428.
5. Baek, N. R., Cho, S. W., Koo, J. H., Truong, N. Q., and Park, K. R. (2019). Multimodal camera-based gender recognition using human-body image with two-step reconstruction network. *IEEE Access*, *7*, 104025-104044.
6. Cao, L., Dikmen, M., Fu, Y., and Huang, T. S. (2008). Gender recognition from body. In *Proceedings of the 16th ACM international conference on Multimedia*, pp. 725-728.
7. Chen, J., Liu, S., and Chen, Z. (2019). Gender classification in live videos. In *2017 IEEE International Conference on Image Processing (ICIP)*, pp. 1602-1606.
8. Chen, J., Liu, S., and Chen, Z. (2017). Gender classification in live videos. In *2017 IEEE International Conference on Image Processing (ICIP)*, pp. 1602-1606.
9. Collins, M., Zhang, J., Miller, P., and Wang, H. (2009). Full body image feature representations for gender

profiling. In *2009 IEEE 12th International conference on computer vision workshops, ICCV workshops*, pp. 1235-1242.

10. Dammak, S., Mliki, H., and Fendri, E. (2023). Gender estimation based on deep learned and handcrafted features in an uncontrolled environment. *Multimedia Systems*, *29*(1), 421-433.

11. Finley, G., and McRoberts, R. (2008). Gender classification using neural networks and facial commensurability ratios. In 2008 IEEE International Conference on Automatic Face & Gesture Recognition (pp. 1-6).

12. Goceri, F., and Ozbey, B. (2021). Real-time gender recognition from live video using multi-stream convolutional neural networks. Multimedia Tools and Applications, 80(24), 38035-38050.

13. 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.

14. Ibrahim, M. Q., Hussein, N. K., Guinovart, D., and Qaraad, M. (2025). Optimizing Convolutional Neural Networks: A Comprehensive Review of Hyperparameter Tuning Through Metaheuristic Algorithms. Archives of Computational Methods in Engineering, 1-38.

15. Karabiyik, M. A., Bahaeddin, T., and Tunc A. (2025). Optimizing Artificial Neural Network using Mountain Gazelle Optimization. IEEACCESS, Vol. 13, 50464-504479

16. Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., and Shah, M. (2021). Transformers in vision: A survey. *ACM Computing Surveys, 54*(10s), 1–41.

17. Khalajzadeh, H., Mansouri, M., and Tavakoli, R. (2014). Face gender classification using mixture of expert neural networks. In 2014 4th International Conference on Computer and Knowledge Engineering (ICCKE) (pp. 596-600).

18. Lanitis, A., Draganova, C., and Christodoulou, C. (2004). Comparing different classifiers for automatic age and gender recognition. In Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition (FGR 2004) (pp. 629-634).

19. Mansanet, J., Albiol, A., and Paredes, R. (2016). Local deep neural networks for gender recognition. Pattern Recognition Letters, 70, 80-86.

20. Martinez, A. M., and Kak, A. C. (2001). PCA versus LDA. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(2), 228-233. doi: 10.1109/34.908974

21. Peng, L. (2005). Face gender recognition based on PCA and Bayesian inference. In Proceedings of the 2005 International Conference on Machine Learning and Cybernetics (Vol. 6, pp. 3448-3452).

22. Piotrowski, T., Laskowski, E., and Grochowski, M. (2020). Face gender classification using convolutional neural networks. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8).

23. 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.

24. Rustam, Z., and Ruvita, A. (2018). Face gender recognition using convolutional neural network (CNN) with transfer learning. In 2018 International Conference on Computer, Control, Informatics and its Applications (IC3INA) (pp. 57-61).

25. Salem, M., Choi, J., and Ravyse, I. (2010). Face gender recognition using local binary patterns and support vector machines. In 2010 IEEE International Conference on Image Processing (pp. 3305-3308). IEEE. doi: 10.1109/ICIP.2010.5652835

26. Shan, C. (2012). Learning local binary patterns for gender classification on real-world face images. *Pattern recognition letters*, *33*(4), 431-437.

27. Singh, D., Kumar, P., and Singh, S. K. (2020). Face gender recognition using convolutional neural networks. In 2020 International Conference on Computational Intelligence and IoT (ICCIIoT) (pp. 145-149).

28. Sundararajan, K., and Damarla, T. (2018). Deep learning for face gender classification. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 5204-5206).

29. Wang, S., Wang, F., and Chen, W. (2019). Deep belief networks: A comprehensive survey. Neural Networks, 110, 166-187.