Journal of Engineering and Technology for Industrial Applications

ITEGAM-JETIA

Manaus, v.11 n.56, p. 168-179. November/December, 2025.
DOI: https://doi.org/10.5935/jetia. v11i56.2262

ISSN ONLINE: 2447-0228

RESEARCH ARTICLE OPEN ACCESS

A NOVEL WAVELET-BASED MULTI-TASK CONVOLUTIONAL NEURAL NETWORK AND GENETIC ALGORITHM APPROACH FOR ENHANCED FACIAL RECOGNITION

Khaled Merit[1]

[1]Laboratory of TIT, Department of Electrical Engineering, Tahri Mohammed University of Bechar.

[1]https://orcid.org/0000-0002-7762-1898

E-mail: merit.khaled@univ-bechar.dz

ARTICLE INFO

Article History
Received: June 24, 2025
Revised: October 20, 2025
Accepted: December 1, 2025
Published: December 31, 2025

Keywords:
Denoising,
Multi task CNN,
Feature extraction,
Genetic algorithm,
Wavelet transform.

ABSTRACT

A new system based on wavelet transform, a multi task convolutional neural network and feature extraction technology, as well as a genetic algorithm model, is put forward to address the insufficient accuracy and high data noise in current Facial Recognition (FR). The new system uses a wavelet transform based face detection algorithm to denoise the noisy data of facial images. Afterwards, feature extraction and genetic algorithm are used to improve the model and enhance the accuracy of facial image recognition. The study results demonstrated that the approach used in the study was 16.0%, 10.0%, and 8.7% higher than other algorithm models in Dataset 1, and 15.3%, 10.1%, and 13.3% higher than other algorithm models in Dataset 2. The highest change value in dataset 1 was 92.3%, which was about 1.1% and 4.7% higher than other models. The recognition accuracy, recall, and regression values of the research method model were the highest at 99.7%, 99.5%, and 99.7%, respectively. The accuracy of the model was 6.00% and 3.33% higher than that of other models. The stability of the research method was the best among traditional algorithms. The approach architecture used in the analysis has improved performance and higher accuracy in FR.

I. INTRODUCTION

Facial Recognition technology has always been of great concern and has been widely applied in many fields, including security systems, social media, finance, healthcare, and more [1]. With the continuous development of computer vision and deep learning fields, significant progress has been made in practical recognition technology, among which deep learning methods have become one of the main research directions currently [2]. The Multi task Cascaded Convolutional Networks (MTCCN) algorithm is a deep learning algorithm for face detection that uses cascaded deep convolutional neural networks (CNN) to achieve multi task processing of faces, including face detection, key point localization, and bounding box regression [3]. MTCNN has achieved an efficient Facial Recognition system by effectively handling the information flow between different tasks. Genetic Algorithm (GA) is a heuristic optimization algorithm suitable for optimizing and searching complex problems. In the social recognition system, GA can be used to optimize the extraction and matching process of facial features, thereby improving recognition accuracy and performance [4].

This study is put forward to achieve the accuracy and overall performance evolution of current Facial Recognition by using the MTCNN algorithm model and feature extraction technology of wavelet transform, as well as the GA model to improve the system model. By researching and developing medical recognition systems based on MTCNN and GA, we aim to provide more reliable and efficient medical recognition solutions for various application scenarios, thereby promoting the development and application of medical recognition technology. This study is categorized as 4 components. The initial one is an overview of the current domestic and international research content. The second part is an analysis of the algorithm and the construction of the system model. The third part is performance testing of the system model through experiments. The fourth is planned as a summary analysis of the current research content.

II. RELATED WORKS

Under the context of artificial intelligence, social recognition technology has become a research hotspot, widely used in security authentication, monitoring systems, intelligent interaction and other fields. Zhi Yang Wang et al. worked and put forward a low rank representation solution to address the issue of neglecting specific local structures and noisy samples from different views, which leads to decreased recognition ability in multi view real recognition. The new method was implemented through a layered Bayesian approach, which constrains and match it with a linear combination. The experiment outcomes said that compared with the most advanced classification and clustering methods of the same kind, this method was more effective [5]. Chen et al. worked and put forward a feature extraction solution supported by neighborhood weighted average and central symmetry to solve the problem of Facial Recognition under complex lighting conditions. A feature fusion algorithm was proposed, which combines the advantages of directional gradient histograms. The experiment outcomes said that compared with other latest algorithms, this algorithm has more reliable effectiveness under challenging lighting conditions [6].

McGugin R W et al. worked and put forward a method based on 7T ultra-high resolution magnetic resonance imaging to investigate the relationship between Facial Recognition and vehicle recognition ability with cortical thickness. The research results indicated that individuals with strong Facial Recognition abilities have relatively thin cortical areas in the selective brain regions of the face, while those with strong vehicle recognition abilities have relatively thick cortical areas in the same regions [7]. Xue S et al. proposed a robust prototype dictionary and robust change dictionary construction method to improve the accuracy of single sample factual recognition per person. The new method utilized dictionary learning methods to obtain atoms, and selects effective atoms through suggested function indexing methods to construct robust prototype dictionaries and robust change dictionaries. The experimental results showed that RPRV had strong robustness to Facial Recognition in unconstrained environments and was superior to state-of-the-art SRC based methods [8].

He M et al. worked and put forward a novel occlusion simulation method for the robustness of Facial Recognition to occlusion, which involves discarding in carefully selected channels. This method simulates real occlusion through spatial regularization and local perception channel removal, and designs a module to improve the contribution rate of non-occluded areas. The experiment outcomes said that the put forwarded one has significant improvements on various benchmarks [9]. Liu Y et al. proposed a multi factor joint normalization network based on generative adversarial networks to address the challenges faced by unconstrained Facial Recognition. This method could normalize multiple factors simultaneously, including posture, lighting, and facial expressions. The experiment outcomes said that the proposed solution could synthesize multi factor normalization results while preserving identity [10]. Xiaoqian Y et al. conducted an experiment to investigate at what spatial resolution the human brain could recognize familiar faces from unfamiliar individuals, and whether this process depends on the distance between the observer and the face.

The results showed that in blurred images with increased spatial resolution, the neural response in the occipital temporal region appeared and quickly reached saturation at approximately 6.3-8.7 cycles, while the neural response disappeared when resolution decreased [11]. Hao M et al. proposed a hyperspectral real recognition solution to address the issues of band misalignment, and high data dimensionality in hyperspectral real recognition. The experiment outcomes said that the algorithm has achieved excellent results on three popular hyperspectral face databases, outperforming most other ones [12]. In summary, most of the current fields of social recognition suffer from issues such as insufficient accuracy, high noise in facial image data, invasion of personal privacy by social recognition, and low recognition accuracy due to occlusion. These have become key issues in social recognition. How to increase the precision of current Facial Recognition and reduce data noise is the main issue of current research. This study proposes a new wavelet transform MTCNN and GA (WMTCNN-GA) based Facial Recognition system to reduce the impact of noise in facial data recognition. At the same time, the GA model is used to improve the feature compensation and expression.

III. FACIAL RECOGNITION MODEL FOR IMPROVING DEEP CNN ALGORITHMS

The main focus of the research is to build a new Facial Recognition model and analyze the main framework and structure of the model. Afterwards, the existing problems in the system were improved and analyzed, and GA models and feature extraction techniques were added to improve the accuracy and precision of facial image recognition.

III.1 DEEP CNN ALGORITHM FOR FACIAL RECOGNITION

Analyzing and detecting facial data and location information is a crucial step in Facial Recognition. Therefore, when performing Facial Recognition, it is necessary to first detect and analyze facial data. In order to analyze facial data stably, wavelet transform and improved MTCNN face detection algorithm is used. The MTCNN model in facial detection analysis mainly consists of several modules: data loading and preprocessing module, data denoising module, and MTCNN detection module. The model first loads and analyzes the facial images through a data loading module, and then performs denoising on the facial data through a data denoising module. The denoised facial data images are then feature extracted and learned to complete facial data detection. The current system framework structure is shown in Figure 1. From Figure 1, it can be seen that the data processing and loading module for facial images in the four modules includes several operation steps: segmentation of facial images, normalization of data, and balanced distribution of facial images.

Then, in the data denoising process, the image data is first subjected to threshold conversion and processing, and then the threshold is selected and analyzed to transfer the reconstructed facial image data into the algorithm model, The processed facial data can only be input into the simulation model, and the obtained facial feature points complete the construction of the facial image. In the denoising process of the model, it is needed with significant meaning to analyze the original image to reduce its resolution size, and then set a relative threshold size. Simply retain the eliminated noise data with a threshold coefficient, and set the data below the threshold to 0. Finally, use the wavelet transform coefficient to reconstruct the face image through the wavelet inverse transform. At this point, the obtained face data has already eliminated the influence of noise.

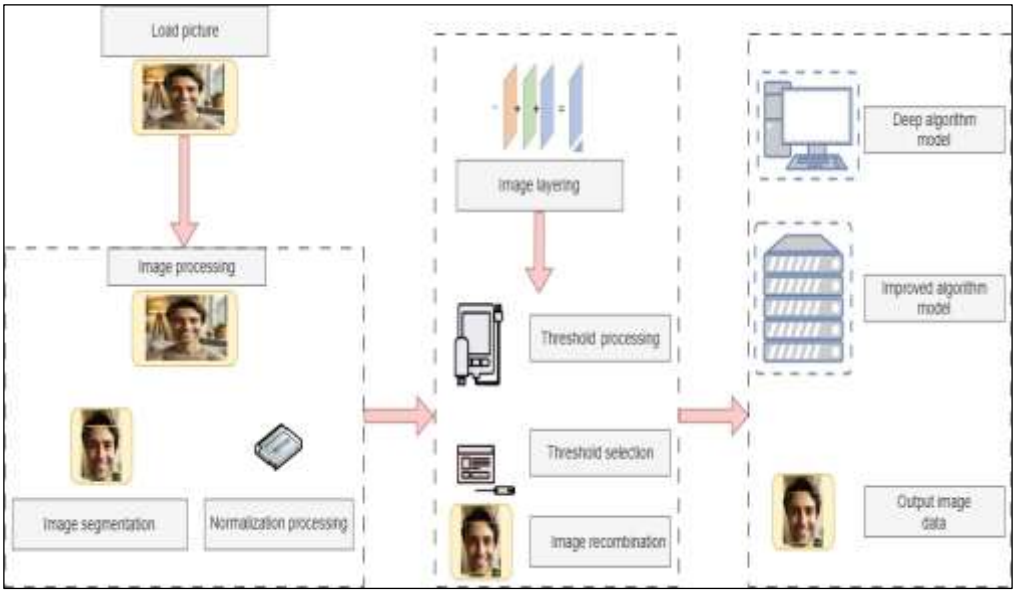

Figure 1: System Model Framework Structure.
Source: Authors, (2025).

MTCNN mainly consists of three parts in facial data recognition and analysis. The initial part is a relatively shallow CNN. It mainly generates facial selection framework images. The second part is a deep convolutional neural work hired to filter the current facial frame image and adjust the frame. The last part is also a more complex CNN model, which mainly detects key information points in the selection box of the face. The schematic diagram of the current MTCNN main stage process framework is shown in Figure 2. From Figure 2, it can be seen that when the detected image information is input, the MTCNN structure will enlarge and shrink the image structure, build a pyramid model of the image, then use the three CNN structures of MTCNN to detect and adjust the face image, and finally output the current required face image data. The largest detection structure in the MTCNN model is the selection of face frames, which calculates the intersection of the human face images based on the detection results, while retaining only the maximum result for the intersection values. At this point, the obtained threshold will not be too low, which will reduce the accuracy of the filtered image. The intersection calculation is shown as in (1) [13].

$$K = \frac{A \cap B}{A \cup B} = \frac{A \cap B}{A+B-A \cap B} \tag{1}$$

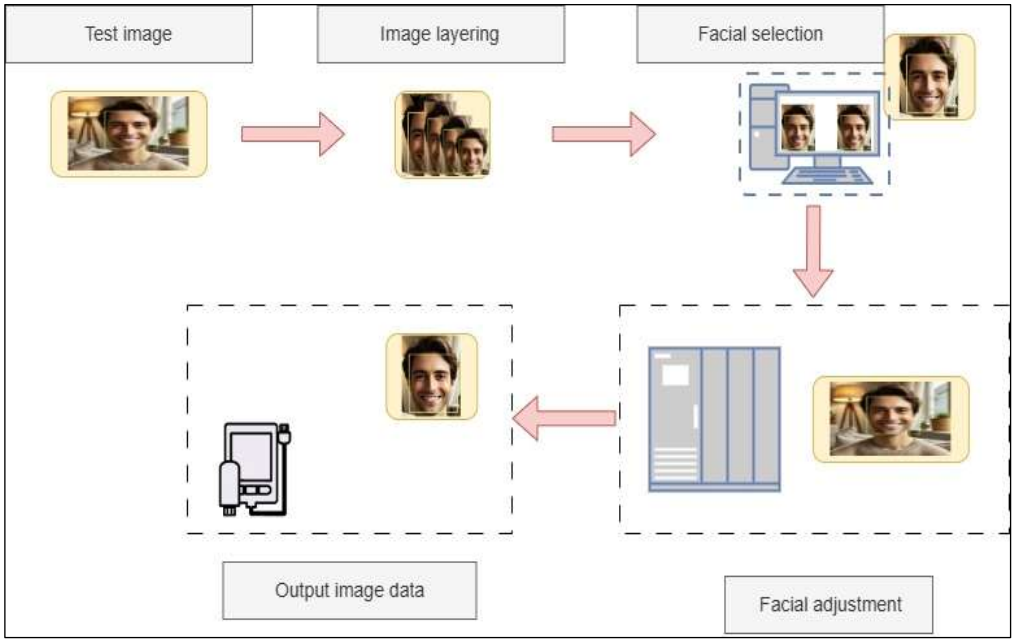

Figure 2: Schematic diagram of MTCNN main stage process framework.
Source: Authors, (2025).

In Equation (1), A represents the area of the trained face selection image, B represents the selection area of the face selection image, and K represents the intersection ratio of the face image. The selection status of the current image is determined by the size of the intersection ratio. The main image processing flow of the MTCNN model CNN is shown in Figure 3.

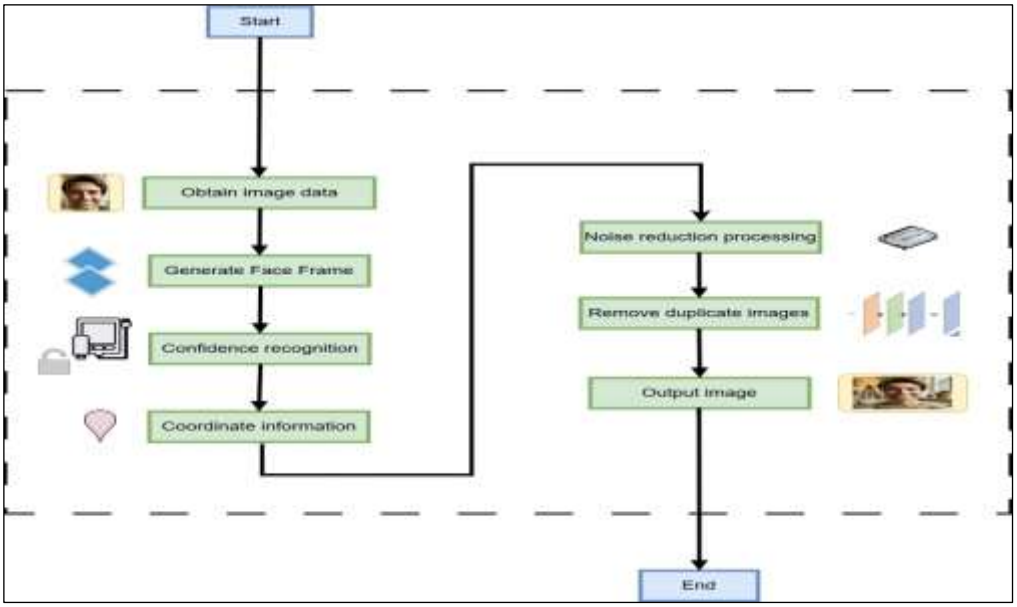

Figure 3: The main image processing flow of the MTCNN model CNN.
Source: Authors, (2025).

From Figure 3, the network module of the MTCNN model first obtains the height and width. Then, if the image data is too large, a face frame is generated, and the coordinate information of the image is obtained through confidence recognition. Then, the image data is denoised and input into the network model. The frame is then calculated for fit while removing duplicate facial images, Finally, obtain the final size of the facial wireframe. The input sample value size of the algorithm model at this time is shown in (2) [14].

$$L^{det} = -\left(y_i^{det}\, log(p_i) + \left(1 + y_i^{det}\right)(1 - log(p_i))\right) \tag{2}$$

In Equation (2), y_i^{det} represents the label value of the face sample, p_i is hired to be the symbol of the network's face probability size, and L^{det} is hired to be the symbol of the input sample size. The loss function value of the facial target wireframe at this time is shown in (3).

$$L^{box} = \left\| \hat{y}_i^{box} - y_i^{box} \right\| \tag{3}$$

In Equation (3), $\hat{y}_i^{box}$ represents the coordinate information of the facial bounding box output by the model, y_i^{box} represents the wireframe boundary coordinates of the real face, and L^{box} represents the loss function value of the facial image. The key points obtained at this point are shown in (4) [15].

$$L^{landmark} = \left\| \hat{y}_i^{landmark} - y_i^{landmark} \right\| \tag{4}$$

In Equation (4), $L^{landmark}$ represents the coordinates of the current facial image keypoints, $\hat{y}_i^{landmark}$ represents the coordinates output by the keypoint network, and $y_i^{landmark}$ represents the true coordinates of the keypoints. Adding up the loss function values of all images yields the weighted loss value, as shown in (5).

$$L = min \sum_{i=1}^{N} \sum_{j \in (det, box, landmark)} a_j \; b_i^j \; L_i^j \tag{5}$$

In Equation (5), N is the symbol of the model training total sample size, a_j is the symbol of the weight size of the model, b_i^j represents the symbol of the sample type, L_i^j represents the weight size of different image wireframes, and L represents the total loss function value of the model. Due to the information loss that occurs when the current model directly selects facial images, reducing its confidence in the algorithm model can reduce the loss of wireframe information, as shown in (6) [16].

$$S_i = \begin{cases} S_i, & K(M, \beta_i) < N_i \\ S_i(1 - K(M, \beta_i)), & K(M, \beta_i) \geq N_i \end{cases} \tag{6}$$

In Equation (6), N_i represents the threshold size of the model, S_i represents the regression wireframe of the model, M represents the wireframe size with the highest confidence, and β_i represents the regression wireframe size for face selection. By using as in (6) to modify the confidence level based on the threshold, the final confidence value can be obtained.

III.2 IMPROVEMENT OF FACIAL RECOGNITION ALGORITHM AND SYSTEM MODEL BUILDING FOR DEEP CNNS

With the purpose to achieve the recognition accuracy improvement of the current system in the MTCNN model, the calculation factor and the method of extracting original image features and GA are added to the system model to improve the algorithm system. The improved model calculates the initial compensation factor of the MTCNN algorithm model to obtain the original image, and then compensates and selects the feature data through GA to obtain the weighted image feature map of the feature model. At this time, the obtained facial feature image is the most accurate feature image, as shown in Figure 4, which is the process of extracting the initial model feature image. From Figure 4, it can be seen that the improved feature image first converts the inserted face data into an image with larger grayscale values. [17] generates a new compensation factor based on the grayscale image.

The optimal compensation coefficient is solved through GA for the compensation factor, and then the obtained optimal compensation coefficient is input into the image recognition system of MTCNN to generate the most original feature description image. The image is then split into multiple sub image data, and a new face image is obtained by calculating the pixel size of the sub images. Finally, the obtained image information is dimensionally reduced to obtain the final high-precision image. Due to the many coefficients that need to be compensated in the improved system model, which can be used to lift the size of the feature values, the obtained feature images can be represented by compensation factors as shown in (7) [18].

$$H = A + S_1 \cdot F_1 + S_2 \cdot F_2 + S_3 \cdot F_3 + \cdots + S_n \cdot F_n \tag{7}$$

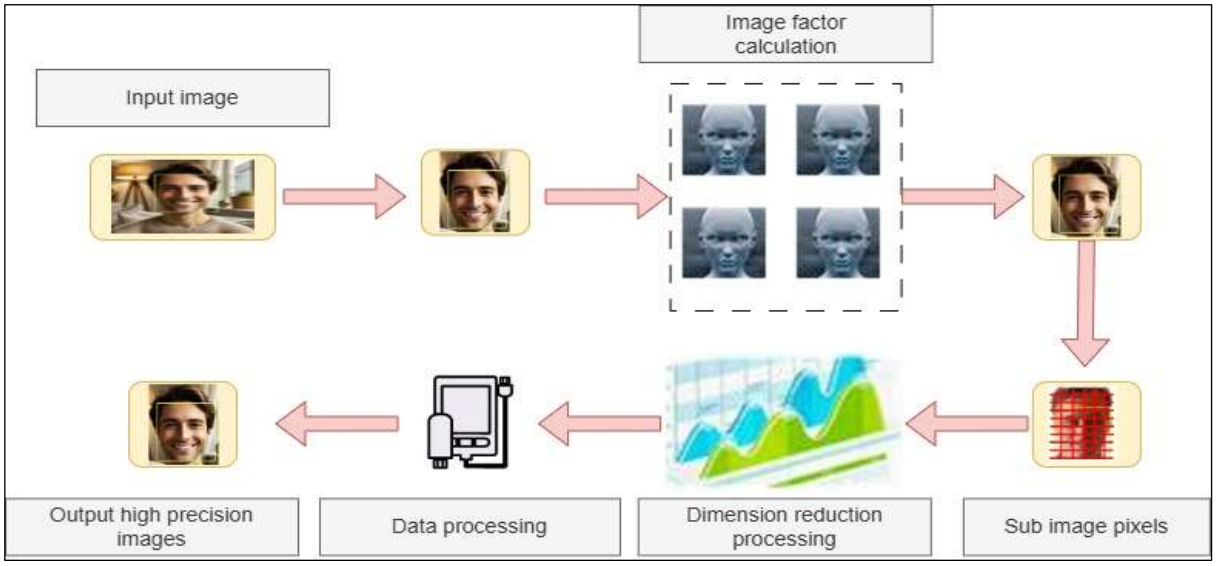

Figure 4: Process of Initial Image Extraction of Model Features.
Source: Authors, (2025).

In Equation (7), S_n represents the calculation factor of the image, and F_n represents the size of the compensation coefficient. In order to process and compare the original characteristic information of the compensation more intuitively, the new image obtained after compensation is compared, as shown in Figure 5. From Figure 5, after improving the recognition rate and data layout, the obtained new image data obtains a new histogram data image. At the same time, the pixel value of the generated image also changes, and the image value obtained by feature compensation can also generate new image pixels. At this time, the pixel size of the two images depends on the compensation coefficient size and calculation factor. Therefore, when these two values change, the clarity and accuracy of images can be greatly improved. In order to obtain a better compensation coefficient, a new compensation factor calculation method needs to be added, as shown in (8).

$$S_{t1}(i,j) = \begin{cases} A(i+1,j), i \geq 1, i < m \\ S_{t1}(i-1,j), \ \ i = m \end{cases} \tag{8}$$

In Equation (8), S_{t1} represents the left bias value of the image, and the other parameter sizes are the same as above. The calculation factor obtained at this point is shown in (9) [19].

$$S_1 = |A - S_{t1}| \tag{9}$$

In Equation (9), S_1 represents the initial calculation factor. The formula for calculating the right bias matrix of the model at this time is shown in (10).

$$S_{t2}(i,j) = \begin{cases} A(i-1,j), i > 1, i \leq m \\ S_{t2}(i+1,j), \ \ i = 1 \end{cases} \tag{10}$$

In Equation (10), S_{t2} represents the right bias value of the image, and the other parameters are the same as above. The formula for the upper bias matrix of the model is shown in (11) [20].

$$S_{t3}(i,j) = \begin{cases} A(i-1,j), i > 1, i \le n \\ S_{t3}(i+1,j), \ \ i = 1 \end{cases} \quad (11)$$

In Equation (11), S_{t3} represents the upper bias value of the image, and the other parameter sizes are consistent with the above. The size of the lower bias matrix of the model is shown in (12).

$$S_{t4}(i,j) = \begin{cases} A(i,j+1), j \ge 1, j \le n \\ S_{t4}(i,j-1), \ \ j = n \end{cases} \quad (12)$$

In Equation (12), S_{t4} represents the lower bias value of the image, and the other parameter sizes are consistent with the above. The final size of the model difference matrix obtained is shown in (13).

$$S_5 = (S_1 - S_2) + (S_3 - S_4) \quad (13)$$

In Equation (13), S_5 represents the size of the final calculation factor obtained. In the improved model, GA is used for image dimension processing due to its strong global search ability and dimension processing ability. The main process is to combine the feature compensation of images to generate new feasible combinations, and designate them as the first-generation population. By iterating the population and generating new populations, the population size can be expanded, allowing for more iterations of its feature images and ultimately obtaining better population combinations [21]. Furthermore, the global optimal solution of the system is obtained to accomplish the quality and accuracy improvement of the generated images, as shown in Figure 6, which is the current improved system model process.

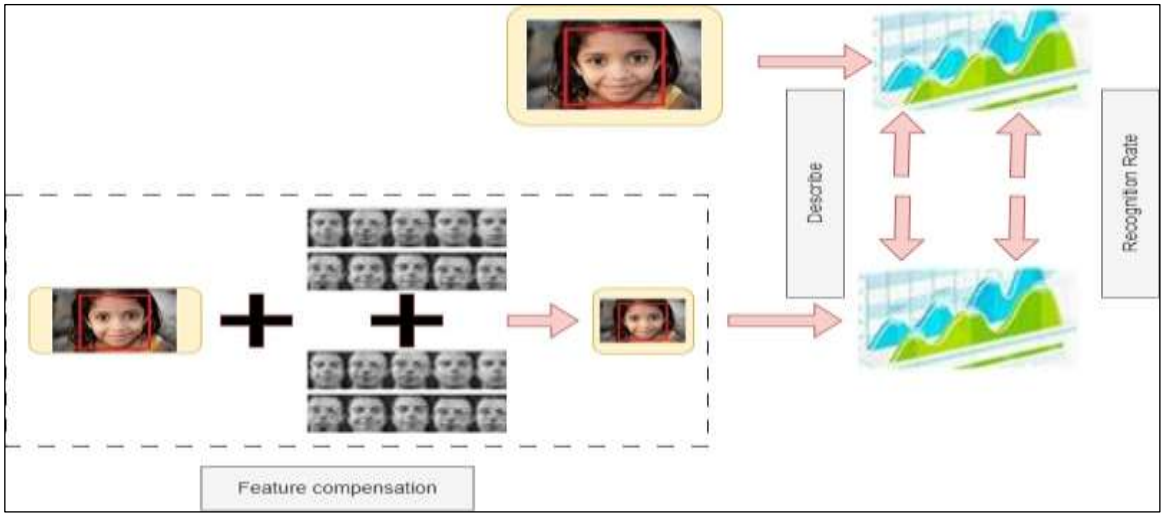

Figure 5: Comparison of new images obtained after compensation.
Source: Authors, (2025).

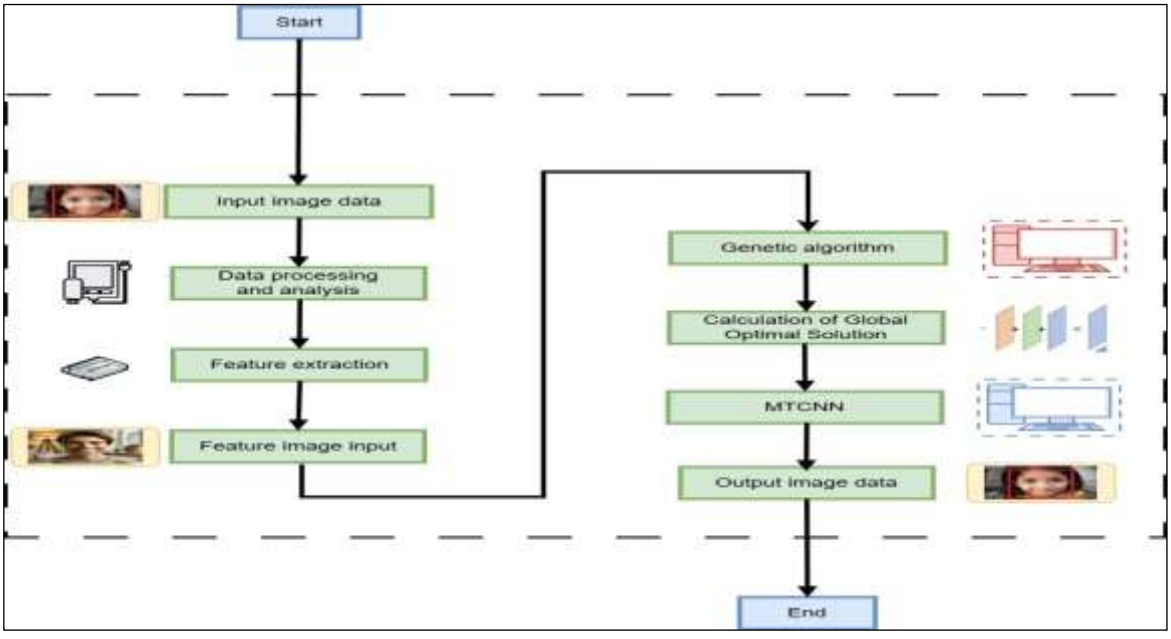

Figure 6: Improved System Model Process.
Source: Authors, (2025).

From Figure 6, after the input of facial image data, the image is first processed and analyzed through the three levels of CNNs of the MTCNN model. After obtaining new image data, the image data information is obtained through feature extraction. The obtained feature image is input into the GA model, and the current image data is calculated for the global optimal solution. Then, the obtained global optimal solution is input into the MTCNN for further calculation. The final image data is then dimensionally reduced, resulting in a clearer and more accurate facial image.

IV. TEST OF THE PROPOSED DEEP CNN FOR FACIAL RECOGNITION

The dataset used in the research experiment includes 32405 facial image data, including 6520 real face instances, including facial images obtained under different environmental conditions. The selected system CPU for the study is an Intel(R) Core (TM) i7-12800H CPU @ 4.0 GHz, with 32 GB of memory, 8 GB Nvidia Quadro RTX A2000. The initial learning rate of the algorithm is set to 0.01, and the weight value is set to 0.0005. Divide the dataset into two equally sized facial image datasets, and to test the feasibility of the current analysis approach model, compare the recognition accuracy of the traditional algorithm model, Local Binary Patterns (LBP), Local Gradient Patterns (LGP), Load Average Algorithm (LAA), and the algorithm model used in the study, as shown in Figure 7. From Figure 7 (a), it can be seen that when comparing the accuracy of the four algorithms, the accuracy of the algorithm goes up with the expansion of sample data. However, when the accuracy of the algorithm model is between 250-2000 image data, except for the algorithm used in the study, the accuracy changes curves of all other algorithm models show significant fluctuations.

However, in Figure 7 (b), the accuracy changes of the four algorithm models show an upward trend, but there is no significant fluctuation. The accuracy values of the research methods used in datasets 1 and 2 are higher than those of the other three algorithm models. The highest accuracy of the research methods used in dataset 1 is 92.3%, and the highest accuracy is 94.6% in dataset 2. In dataset 1, compared to 76.3%, 82.3%, and 83.6% of LBP, LGP, and LAA algorithm models, it is 16.0%, 10.0%, and 8.7% higher. In dataset 2, compared to 79.3%, 84.5%, and 81.3% of LBP, LGP, and LAA algorithm models, it is 15.3%, 10.1%, and 13.3% higher. To test the effectiveness of the current study approach in model ablation testing, the recognition accuracy of the research method was compared with the MTCNN and GA, and the results are shown in Figure 8.

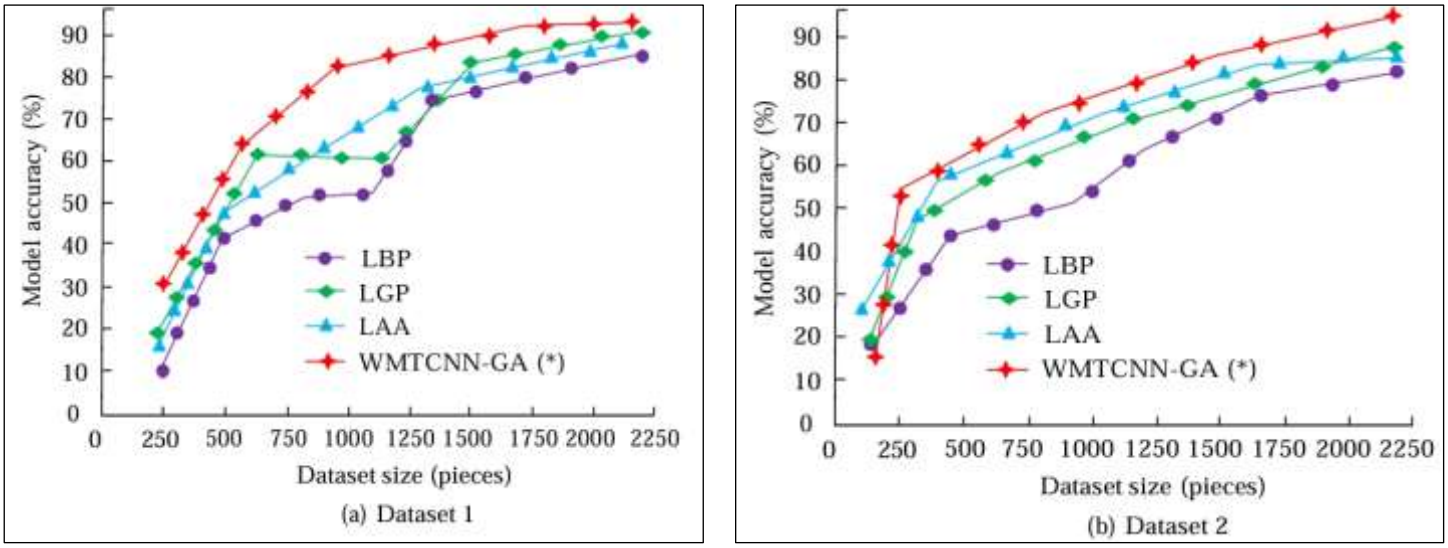

Figure 7: Accuracy of Four Algorithm Models.
Source: Authors, (2025).

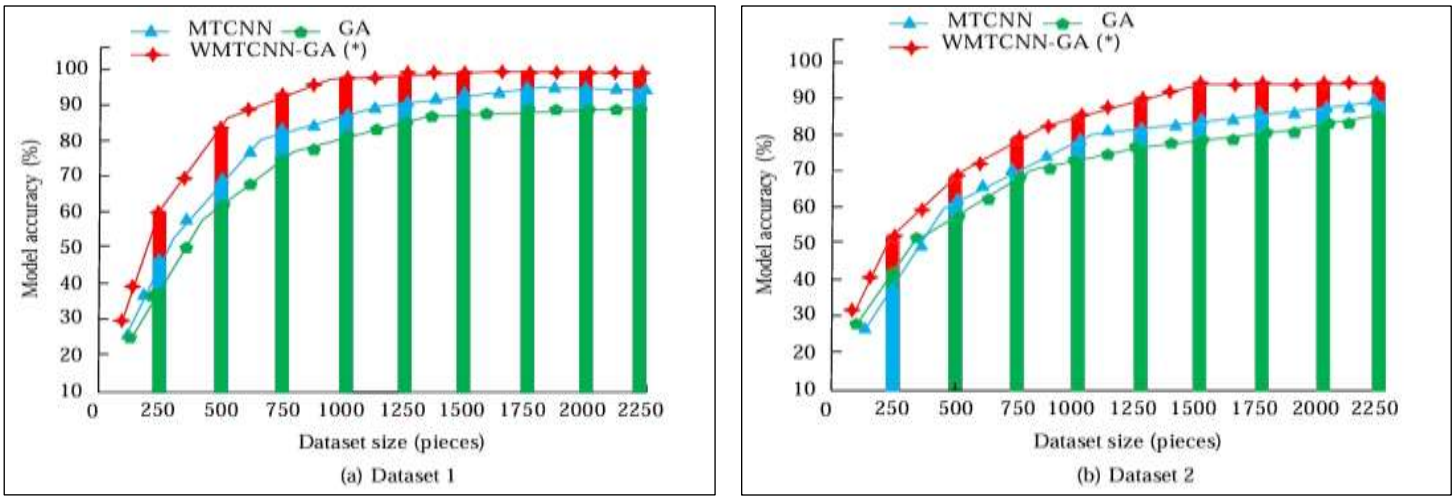

Figure 8: Comparison of accuracy of ablation pressure recognition using algorithm models.
Source: Authors, (2025).

From Figure 8 (a), it can be seen that during the ablation experiment, the algorithm approach used in the research had the highest accuracy in Facial Recognition in dataset 1. The accuracy values of the three algorithm models improved with the expansion of facial data, and then tended to stabilize. When the accuracy change is relatively stable, the highest change value of the algorithm model in dataset 1 is 92.3%, which is about 1.1% and 4.7% higher than the 91.2% of the MTCNN model and 87.6% of the GA model. From Figure 8 (b), the accuracy in dataset 2 is 94.6%, which is about 7.2% and 10.4% higher than the MTCNN model's 87.4% and GA model's 84.2%. It can be seen from this that the use of method models in research results in better model performance. To test the other performance of the current research method, such as the number of successful recognitions, recognition accuracy, recall, and regression value, 1000 facial image data were used to compare the research method with the five algorithm models mentioned above. The test outcomes are provided in Table 1.

Table 1: Performance Comparison of Different Algorithm Models.

Algorithm model	Recognition successrate	Precision (%)	Recall (%)	Regression value (%)
GA	976	97.6	96.5	98.4
MTCNN	962	96.2	98.4	96.4
LGP	986	98.6	98.1	98.3
LAA	984	98.4	96.5	99.4
LBP	975	97.5	98.5	98.5
WMTCNN-GA (Ours*)	**997**	**99.7**	**99.5**	**99.7**

Source: Authors, (2025).

From Table 1, the method used achieved the highest recognition success rate of 997 when the face data was the same. At the same time, the model recognition accuracy, recall, and regression values of this method were the highest at 99.7%, 99.5%, and 99.7%, respectively. Compared to other algorithm models, the researched one's accuracy is 3.5% higher than the lowest algorithm model, the recall is 4% higher than the lowest algorithm model, and the regression value is 3.3% higher than the lowest algorithm model. It can be seen that the algorithm performance of the research using method models is better than other algorithm models, and the recognition and detection effect of faces is better. To test the impact of algorithm feature coefficients on the algorithm model used in the current study, 8 different compensation coefficients were selected for testing. The compensation coefficients were randomly combined to generate 10 different combinations. Subsequently, the accuracy of the research method was compared with the MTCNN and GA, as shown in Table 2.

Table 2: Comparison test of accuracy of random compensation coefficients for three algorithms.

/	Accuracy (%)		
Combination	**MTCNN**	**GA**	**WMTCNN-GA**
1	66.30	69.45	71.23
2	67.42	70.32	73.24
3	67.95	70.62	73.95
4	68.23	65.31	71.25
5	60.23	64.85	73.25
6	61.24	67.35	72.35
7	96.32	69.25	73.95
8	63.95	65.23	72.86
9	64.57	68.21	72.68
10	66.84	69.84	73.84

Source: Authors, (2025).

From Table 2, the proposed one has the highest accuracy value among different compensation coefficient combinations. When combining the compensation coefficients of 3, the accuracy of the method model used in the study was 73.95%. Compared to the accuracy of 67.95% in the MTCNN model, it is 6.00% higher, and compared to the accuracy of 70.62% in the GA model, it is 3.33% higher. The algorithm model accuracy will be improved after adding compensation coefficients. To analyze and compare the half error rate and error rate of the current model, GA model, and MTCNN model, as shown in Figure 9, the smaller the value of the half error rate, the better the model. From Figure 9 (a), the error rates of the three algorithm models increase with the rise of sample quantity, and then tend to a relatively stable state. Among the three algorithms, the error rate of the algorithm approach used in the research is the lowest when in a relatively stable state, only 6.2%, which is 1.5% and 1.9% lower than the 7.7% and 8.1% of other algorithm models. It can be seen that the performance of models with relatively lower recognition error rates using algorithmic models is better.

From Figure 9 (b), the half error rate of the three algorithm models decreases with the expansion of sample size. The lowest half error rate of the algorithm model used in the study is 1.9%, which is 0.1% and 0.9% lower than the half error rate values of other algorithm models of 2.0% and 2.8%. It can be seen that the model using the research method has better performance. To test the stability of the current study approach, the loss function changes of five algorithm models were compared and tested, as shown in Figure 10. From Figure 10, the size of the loss function value in the five algorithms decreases with the increase of iteration times. When the iteration number reaches a certain value, the loss function value reaches a relatively stable state. From the graph, it can be seen that the algorithm used in the study has a minimum loss function value of only 1.5 when the loss function is relatively stable. Compared to other algorithm models, the loss function value is smaller, indicating that studying the stability of algorithms using algorithm models is more stable compared to other algorithm models.

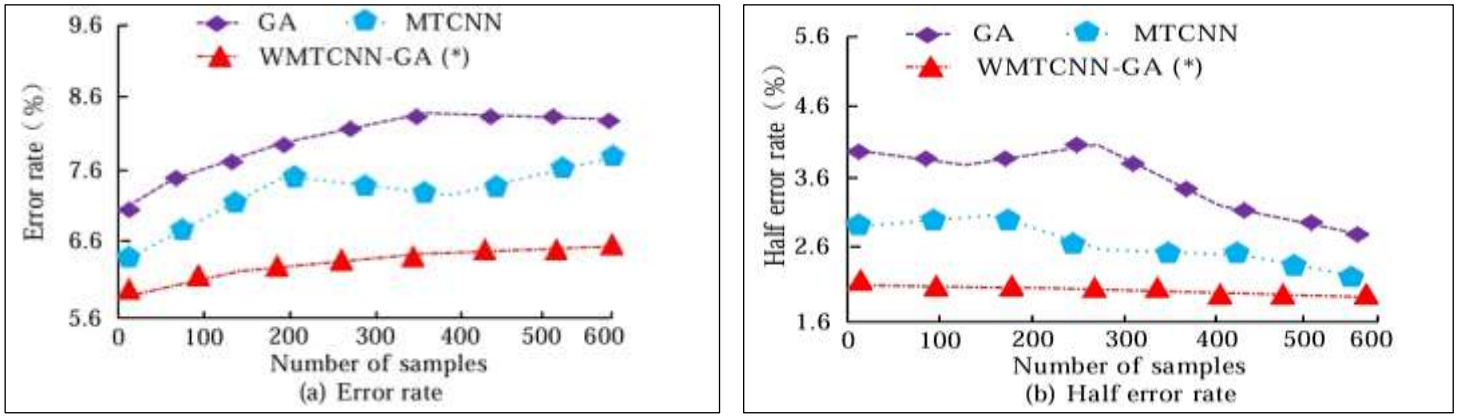

Figure 9: Comparison of Half Error Rate and Error Rate of Three Models.
Source: Authors, (2025).

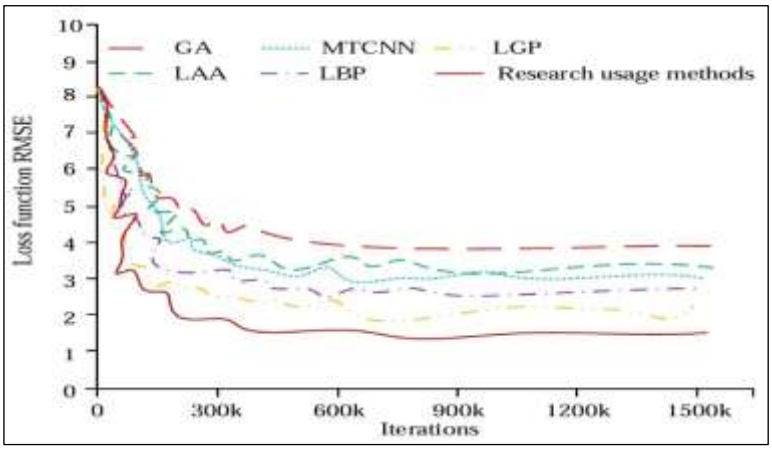

Figure 10: Comparison of loss function changes among five algorithm models.
Source: Authors, (2025).

V. COMPARISON WITH RELATED WORKS

Table 3 provides a comprehensive comparison of our results with several significant studies in the field of face detection, all of which have utilized deep learning methodologies. The selected works were chosen based on the similarity of their datasets to ours, in terms of the number of test subjects, with some studies incorporating slightly larger datasets. As illustrated in the table, our proposed method not only matches but, in some instances, surpasses the accuracy of other approaches employing different techniques and datasets. These findings underscore the robustness and efficacy of our method in face detection, highlighting its potential to perform competitively within the current state of the art. The Figure 11 illustrates the comparative performance of various face detection methodologies, highlighting their accuracy across different studies. The results reveal a significant advancement in face detection techniques, showcasing how recent innovations have led to higher accuracy rates. For instance, the method by Long et al. (2023) [22] achieved an impressive accuracy of 99.69%, closely followed by Khan et al. (2024) [23] with 99.70%.

These high-performance rates indicate the effectiveness of contemporary architectures in handling complex face detection tasks. In contrast, Qi et al. (2022) [24] reported a considerably lower accuracy of 86.00%, reflecting the challenges faced by earlier models in adapting to diverse facial conditions. The performance of Jantarasorn et al. (2023) [25] and Wang, S. (2024) [26], with accuracies of 93.98% and 94.80%, respectively, suggests that while these methods have improved upon traditional techniques, there remains room for optimization. Additionally, the proposed method (2024) showcases an outstanding accuracy of 99.70%, further emphasizing the potential of integrating advanced techniques such as wavelet transforms and multi-task learning in enhancing facial recognition systems. Overall, the figure not only illustrates the progressive trends in accuracy across different methods but also underscores the ongoing evolution of face detection technology, paving the way for more effective and robust applications in various fields.

Table 3: The Accuracy (%) score of our proposed model along with Recent Facial Recognition (FR).

Author	Year	Method	Dataset	Accuracy (%)
Long, D. T. [22]	2023	Efficient MTCNN	HOUS22	99.69
Jantarasorn, N. et al.[25]	2023	Multi-Task Cascaded Convolutional Networks (MTCNN)	Self-Acquired: 200 images (10 individuals)	93.98 (Front-facing), 88.91(Side-facing), 83.42 (Maskwearing)
Qi, S. et al. [24]	2022	MTCNN + FaceNet	LFW Dataset	86.0
Khan, S.S. et al. [23]	2024	MTCNN++	Self-Acquired: 113,586 faces (9661 images)	87.7 (12 faces per image), 99.7 (2 faces per image)
Wang, S. [26]	2024	MobileNet SSD	LFW Dataset	94.8

Mamieva, D. et al. [27]	2023	RetinaNet Baseline	WIDER FACE, FDDB	95.6
Pai, G. et al. [28]	2023	Semi-Dense U-Net	FDDB OpenImage	98.97 91.60
Proposed method	**2024**	**Wavelet-Based Multi-Task CNN + GA (WMTCNN-GA)**	**Self-Acquired: 32,405 facial images**	**99.70**

Source: Authors, (2025).

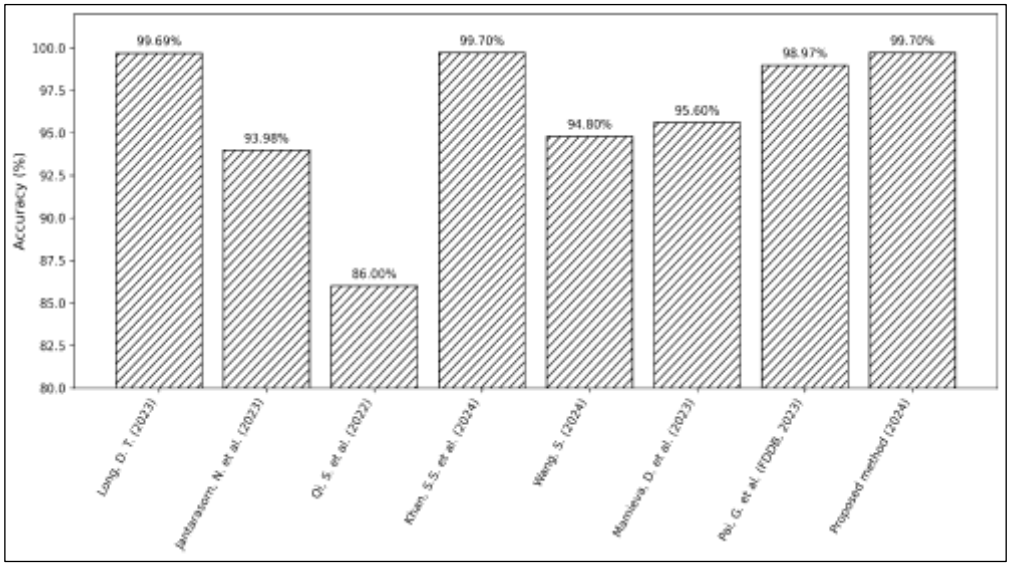

Figure 11: Accuracy Comparison Across Face Detection Methods.
Source: Authors, (2025).

V. CONCLUSIONS

A new system model combining MTCNN model feature extraction technology and GA algorithm model is proposed to address the low accuracy and significant impact of image data noise in current Facial Recognition models. The study first established a system model framework for the MTCNN model, and then added feature extraction and GA algorithms on the basis of the MTCNN model to accomplish the accuracy and precision improvement of the current system model. Finally, the efficiency and recognition performance of the current architecture model were compared through experimental analysis. The study results demonstrated that the precision values of the research methods used in different datasets were higher than those of the other three algorithm models, with the highest accuracy of 92.3% in dataset 1 and 94.6% in dataset 2. In dataset 1, compared to other algorithm models, it was 16.0%, 10.0%, and 8.7% higher, and in dataset 2, it was 15.3%, 10.1%, and 13.3% higher.

The highest change value in dataset 1 was 92.3%, which was about 1.1% and 4.7% higher than other models. The maximum number of successful recognitions achieved by the research using the same facial data was 997, and the recognition accuracy, recall, and regression values of the model were the highest at 99.7%, 99.5%, and 99.7%, respectively. The accuracy of the model was 6.00% and 3.33% higher than that of other models. At the same time, the error rate of using algorithmic models was the lowest at 6.2%, which was 1.5% and 1.9% lower than other algorithmic models. It can be seen that the precision and stability of the architecture developed in the research are superior to other algorithm models. Although this study has achieved significant results, there are still many shortcomings, and future research will involve studying and analyzing more extensive and diverse datasets.

Furthermore, building on these findings, future work could focus on enhancing the robustness of facial recognition systems against challenges such as identity alteration through cosmetic surgery. Investigating how the proposed model can accurately identify and analyze facial features before and after surgical modifications will be crucial. This research direction will require expanding the dataset to include instances of facial alterations, thereby improving the model's capability to recognize individuals despite significant changes in their appearance. Additionally, exploring the ethical implications and privacy concerns associated with the deployment of advanced facial recognition technologies will be essential for responsible use in real-world applications.

VII. AUTHOR'S CONTRIBUTION

Conceptualization: Khaled Merit
Methodology: Khaled Merit
Investigation: Khaled Merit
Discussion of results: Khaled Merit
Writing – Original Draft: Khaled Merit
Writing – Review and Editing: Khaled Merit
Resources: Khaled Merit
Supervision: Khaled Merit
Approval of the final text: Khaled Merit

VIII. REFERENCES

[1] Li, W., Li, J., Cao, D., & Luo, N. (2021). Neural mechanism of noise affecting face recognition. Neuroscience, 468(3), 211–219. https://doi.org/10.1016/j.neuroscience.2021.06.017

[2] Madarkar, J., Sharma, P., & Singh, R. P. (2021). Sparse representation for face recognition: A review paper. IET Image Processing, 15(9), 1825–1844. https://doi.org/10.1049/ipr2.12155

[3] Zaman, F. H. K. (2020). Locally lateral manifolds of normalized Gabor features for face recognition. IET Computer Vision, 14(4), 122–130. https://doi.org/10.1049/iet-cvi.2019.0531

[4] Kar, A., & Neogi, P. P. G. (2020). Triangular coil pattern of local radius of gyration face for heterogeneous face recognition. Applied Intelligence, 50(3), 698–716. https://doi.org/10.1007/s10489-019-01545-x

[5] Wang, Z., Abhadiomhen, S. E., Liu, Z., & Shen, X. J. (2021). Multi-view intrinsic low-rank representation for robust face recognition and clustering. IET Image Processing, 15(14), 3573–3584. https://doi.org/10.1049/ipr2.12232

[6] Chen, T., Gao, T., Li, S., Xi, Z., & Cao, J. (2021). A novel face recognition method based on fusion of LBP and HOG. IET Image Processing, 15(14), 3559–3572. https://doi.org/10.1049/ipr2.12192

[7] Mcgugin, R. W., Newton, A. T., Tamber-Rosenau, B., Tomarken, A., & Gauthier, I. (2020). Thickness of deep layers in the fusiform face area predicts face recognition. Journal of Cognitive Neuroscience, 32(7), 1316–1329. https://doi.org/10.1162/jocn_a_01551

[8] Xue, S., & Ren, H. P. (2022). Single sample per person face recognition algorithm based on the robust prototype dictionary and robust variation dictionary construction. IET Image Processing, 16(3), 742–754. https://doi.org/10.1049/ipr2.12381

[9] He, M., Zhang, J., Shan, S., & Liu, X. (2022). Locality-aware channel-wise dropout for occluded face recognition. IEEE Transactions on Image Processing, 31(5), 788–798. https://doi.org/10.1109/tip.2021.3132827

[10] Liu, Y., & Chen, J. (2021). Multi-factor joint normalisation for face recognition in the wild. IET Computer Vision, 15(6), 405–417. https://doi.org/10.1049/cvi2.12025

[11] Xiaoqian, Y., Goffaux, V., & Bruno, R. (2021). Coarse-to-fine(r) automatic familiar face recognition in the human brain. Cerebral Cortex, 32(8), 1560–1573. https://doi.org/10.1093/cercor/bhab238

[12] Hao, M., Liu, G., & Xie, D. (2021). Hyperspectral face recognition with a spatial information fusion for local dynamic texture patterns and collaborative representation classifier. IET Image Processing, 15(8), 1617–1618. https://doi.org/10.1049/ipr2.12131

[13] Choi, J. Y., & Lee, B. (2020). Ensemble of deep convolutional neural networks with Gabor face representations for face recognition. IEEE Transactions on Image Processing, 29(3), 3270–3281. https://doi.org/10.1109/tip.2019.2958404

[14] Gundogdu, B., & Bianco, M. (2020). Collaborative similarity metric learning for face recognition in the wild. IET Image Processing, 14(9), 1759–1768. https://doi.org/10.1049/iet-ipr.2019.0510

[15] Sadeghzadeh, A., & Ebrahimnezhad, H. (2020). Pose-invariant face recognition based on matching the occlusion free regions aligned by 3D generic model. IET Computer Vision, 14(5), 268–277. https://doi.org/10.1049/iet-cvi.2019.0244

[16] Koc, M., Ergin, S., Gulmezoglu, M. B., & Edizkan, R. (2020). Use of gradient and normal vectors for face recognition. IET Image Processing, 14(10), 2121–2129. https://doi.org/10.1049/iet-ipr.2019.1128

[17] Zhang, Y., & Yan, L. (2022). A fast face recognition based on image gradient compensation for feature description. Multimedia Tools and Applications, 81(26), 26015–26034. https://doi.org/10.1007/s11042-022-12804-4

[18] Yan, X., & Rossion, B. (2020). A robust neural familiar face recognition response in a dynamic (periodic) stream of unfamiliar faces. Cortex, 132(10), 281–295. https://doi.org/10.1016/j.cortex.2020.08.016

[19] Liu, Q., Jiang, B., & Zhang, J. L. (2020). Semi-supervised uncorrelated dictionary learning for color face recognition. IET Computer Vision, 14(3), 92–100. https://doi.org/10.1049/iet-cvi.2019.0125

[20] Mishra, N. K., Kumar, S., & Singh, S. K. (2022). MmLwThV framework: A masked face periocular recognition system using thermo-visible fusion. Applied Intelligence, 53(3), 2471–2487. https://doi.org/10.1007/s10489-022-03517-0

[21] Hebbi, C., & Mamatha, H. (2023). Comprehensive dataset building and recognition of isolated handwritten Kannada characters using machine learning models. Artificial Intelligence and Applications, 1(3), 179–190. https://doi.org/10.47852/bonviewaia3202624

[22] Long, D. T. (2023). Efficient multi-task CNN for face and facial expression recognition using residual and dense architectures for application in monitoring online learning. International Journal of Fuzzy Logic and Intelligent Systems, 23(3), 229–243. https://doi.org/10.5391/ijfis.2023.23.3.229

[23] Khan, S. S., Sengupta, D., Ghosh, A., & Chaudhuri, A. (2024). MTCNN++ : A CNN-based face detection algorithm inspired by MTCNN. The Visual Computer, 40(2), 899–917.
https://doi.org/10.1007/s00371-023-02822-0

[24] Qi, S., Zuo, X., Feng, W., & Naveen, I. G. (2022). Face recognition model based on MTCNN and FaceNet. 2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC), 1–5.
https://doi.org/10.1109/icmnwc56175.2022.10031806

[25] Jantarasorn, N., Whasphuttisit, J., & Jitsakul, W. (2023). Face recognition using deep learning. In 2023 7th International Conference on Information Technology (InCIT) (pp. 470–474).
https://doi.org/10.1109/incit60207.2023.10413069

[26] Wang, S. (2024). A face recognition method based on lightweight neural network and multi-hash recognition degree weighting. IAENG International Journal of Applied Mathematics, 54(3).

[27] Mamieva, D., Abdusalomov, A. B., Mukhiddinov, M., & Whangbo, T. K. (2023). Improved face detection method via learning small faces on hard images based on a deep learning approach. Sensors, 23(1), 502.
https://doi.org/10.3390/s23010502

[28] Pai, G., & Sharmila, K. M. (2023). Semi-Dense U-Net: A novel U-Net architecture for face detection. International Journal of Advanced Computer Science and Applications, 14(6).
https://doi.org/10.14569/ijacsa.2023.0140643