ISSN1330-3651 (Print), ISSN 1848-6339 (Online) https://doi.org/10.17559/TV-20230701000778
Original scientific paper

Clustering Optimized Portrait Matting Algorithm Based on Improved Sparrow Algorithm

Xiang WU*, Yuanhao MA, Hao LIAN, Xiang FANG, Tianfei CHEN

Abstract: As a result of the influence of individual appearance and lighting conditions, aberrant noise spots cause significant mis-segmentation for frontal portraits. This paper presents an accurate portrait segmentation approach based on a combination of wavelet proportional shrinkage and an upgraded sparrow search (SSA) clustering algorithm to solve the accuracy challenge of segmentation for frontal portraits. The brightness component of the human portrait in HSV space is first subjected to wavelet scaling denoising. The elite inverse learning approach and adaptive weighting factor are then implemented to optimize the initial center location of the K-Means algorithm to improve the initial distribution and accelerate the convergence speed of SSA population members. The pixel segmentation accuracy of the proposed method is approximately 70% and 15% higher than two comparable traditional methods, while the similarity of color image features is approximately 10% higher. Experiments show that the proposed method has achieved a high level of accuracy in capricious lighting conditions.

Keywords: elite backward learning strategy; K-means; portrait segmentation; sparrow search algorithm; wavelet proportional shrinkage

1 INTRODUCTION

Image segmentation [1, 2] has drawn a lot of attention as an essential area in image processing, with portrait segmentation [3] being a particularly significant technology in image segmentation. With the continuous development of science and information technology, especially the arrival of the 5G era [4], its high-speed data transmission capability provides a powerful boost. Nowadays, many emerging technologies related to digital image processing, such as intelligent medical diagnosis, automatic driving, expression detection, smart home, etc., have been applied to people's activities, providing great convenience to their daily life. The application of portrait segmentation technology is also applied more and more frequently in scenarios such as video background virtualization, intelligent search and rescue, document background replacement, hybrid display, portrait recognition, data privacy [5], etc. As it tries to extract high-quality portraits of people from natural photographs, it becomes crucial to find a solution for accurately segmenting the portrait's foreground.

At present, the positioning of the clustering center in the portrait and the accuracy of portrait segmentation are currently hampered by the uncertainty of natural lighting, artificial lighting, lighting facilities, and parameter settings of the acquisition equipment. The traditional image processing techniques cannot segment the portrait accurately. Scholars domestically and internationally have suggested several optimum segmentation strategies to increase the accuracy of human portrait segmentation. Wang et al. [6] suggested a social particle swarm to optimize the initial population of the particle swarm to improve the accuracy of fuzzy C-mean clustering, Ahmad Khan et al. [7] suggested a genetic algorithm to optimize the clustering overlap to improve the convergence effect, and Varshali Jaiswal et al. [8] suggested a genetic algorithm in using the Lab color space to minimize the overall distance of clusters to improve the segmentation efficiency; these approaches can improve segmentation. Deep learning has recently been steadily used for portrait segmentation as well. To compensate for information loss, Ronneberger et al. [9] proposed the U-NET model, which combines the deconvolution layer and the feature layer. However, the depth of the constructed network is constrained due to gradient disappearance. Chen et al. [10] improved the U-NET model and designed Refine Block to obtain edge retention through the network filter, so that the segmentation accuracy of portrait edges is on a higher level. However, the problem is that the gradient disappeared.

The center of the portrait affected by noise will be shifted during clustering, and the center of the clusters cannot be located effectively, making it difficult to achieve accurate portrait segmentation. In conclusion, clustering segmentation [11] is still the simplest and fastest method in portrait extraction. To solve the issue of accurate portrait segmentation, this study investigates the K-Means method [12], which is based on wavelet proportional shrinkage and combined with enhanced sparrow algorithm [13] optimization.

2 ALGORITHM PRINCIPLE AND IMPROVEMENT

In this paper, the traditional K-Means segmentation algorithm is improved in the following two steps. Firstly, the color space is converted from RGB to HSV, and the wavelet proportional shrinkage denoising method is used to handle the abnormal noise points in the luminance component of the image; secondly, the SSA algorithm optimized by the elite reverse strategy is introduced to determine the optimal initial clustering central [14].

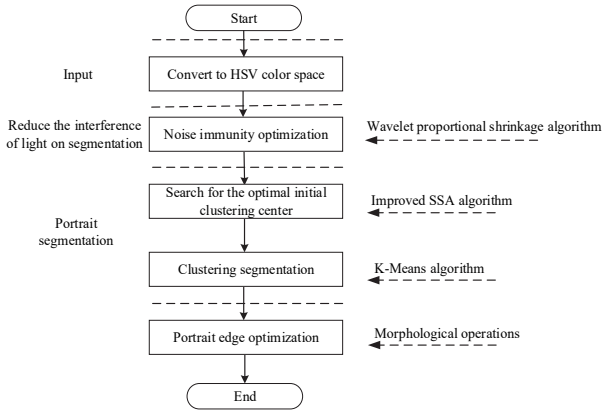


Figure 1 Block diagram of the algorithm

The portrait color map is then cluster segmented using the K-Means technique, and iteration ends when the criteria function converges. The portrait segmentation edges are then further optimized using morphological techniques to produce a superior portrait segmentation map. A brief description is shown in Fig. 1.

2.1 Wavelet Proportional Shrinkage Denoising Algorithm

Wavelet threshold denoising is the process of passing a signal through a wavelet transform [15] and separating the signal from noise based on the size of its wavelet coefficients, either through hard threshold quantization or soft threshold quantization [16]. While the denoising effect is not ideal, assuming that the local wavelet coefficients can be analyzed and the algorithm can be adjusted in real-time to improve the adaptivity of the wavelet denoising algorithm, the effect will be significantly improved. In the current wavelet proportional shrinkage method, local wavelet coefficients are processed differently, putting a lot of emphasis on local noise in the algorithm. The technique used in this study deals with portraiture, which is significantly impacted by lighting. As a result, the system performs denoising on the luminance component of the image.

(i) Firstly, the wavelet coefficients are obtained by doing orthogonal wavelet transform on the image $y(i, j)$ with the image size of $m \times n$.

(ii) Choose the size 3 × 3 filter window and place the estimated wavelet coefficients $Y(i, j)$ in the center of the window. The equation for estimating the variance of the signal wavelet coefficients is shown below.

$$\hat{\sigma}^2(i,j) = \max\left(0, \frac{1}{M^2} \sum_{i,j \in \Omega(i,j)} Y^2(i,j) - \sigma_n^2\right) \tag{1}$$

where σ_n^2 is the noise variance, $\hat{\sigma}^2(i,j)$ is the estimated variance of the signal wavelet coefficients, and $\Omega(i,j)$ is the wavelet coefficient within a window containing $Y(i, j)$ with window size $M \times M$.

(iii) The wavelet coefficients Y_{HH} of the high-frequency information HH part of the wavelet decomposition at level 1 can be used to estimate the noise variance σ_n^2 using the median estimation approach. The noise variance is determined as indicated below using $n \times n$ as the image size.

$$\hat{\sigma}_n^2 = \frac{\sqrt{n} * \text{median}(Y_{HH})}{0.6745} \tag{2}$$

where $\hat{\sigma}^2$ is an estimation of the noise variance, and median () is the function of abstracting the middle value.

(iv) The approximate signal wavelet coefficients $\hat{X}(i,j)$ are computed as follows.

$$\hat{X}(i,j) = \frac{\hat{\sigma}^2(i,j)}{\hat{\sigma}^2(i,j) + \hat{\sigma}_n^2} Y(i,j) \tag{3}$$

where $\hat{\sigma}^2(i,j)$ is the estimated variance of the signal wavelet coefficients, $\hat{\sigma}^2$ is an estimation of the noise variance. The meaning of $Y(i,j)$ is the same as in Eq. (1). Then we get the wavelet coefficient approximation $\hat{X}(i,j)$, and the original image is then reconstructed using this approximation.

2.2 Improved SSA Algorithm

A new technique for locating global optimization called SSA [17] is based on the inference of sparrow foraging and hazard avoidance behavior. In comparison to conventional algorithms, the SSA algorithm has the benefits of a simple structure, simple implementation, and fewer control parameters. The algorithm also performs better on benchmark functions like single-peak and multi-peak than conventional techniques. But the SSA method is like most population algorithms in that it encounters the issue of local optimal solutions late in the iteration as population diversity drops. This study offers an adaptive weight factor to optimize the discoverer update position formula and an elite backward learning approach to optimize the initial population distribution strategy of the SSA algorithm to address the issues.

2.2.1 Elite Reverse Learning Strategies

The population initialization strategy of the SSA algorithm will be improved since it currently uses a simple random strategy to initialize the population distribution within a given range, i.e., by generating random numbers within a given interval. This causes the population members to be unevenly distributed and has negative effects on the algorithm's stability and convergence speed.

Eq. (4) illustrates the elite population technique created based on the population membership using the elite backward learning strategy to initialize the population membership distribution.

$$\bar{X}_{i,j} = \mu(lb_j + ub_j) - X_{i,j} \tag{4}$$

where $X_{i,j}$ is the location information of M d-dimensional sparrow individuals randomly generated by the initial sparrow population, $i \in [1, M]$; $j \in [1, d]$, μ is a random number in the interval [0,1], $\bar{X}_{i,j}$ is the elite inverse solution of population member $X_{i,j}$, and lb_j, ub_j are the upper and lower bounds of the j-th dimensional search region, respectively, whose solution formula is shown in Eq. (5).

$$lb_j = \min(X_{i,j}), ub_j = \max(X_{i,j}) \tag{5}$$

Eq. (6) shows the screening process for the fitness of the population during the search process. By comparing the fitness values corresponding to the elite population ($\bar{X}_i$) and the original population ($\bar{X}_i$), the positions of the

population individuals are directly updated to the side with more fitness values.

$$\boldsymbol{X}_i = \begin{cases} \bar{\boldsymbol{X}}_i, & f\left(\boldsymbol{X}_i\right) < f\left(\bar{\boldsymbol{X}}_i\right) \\ \boldsymbol{X}_i, & \text{Otherwise} \end{cases} \tag{6}$$

The new population has 2*M* persons in it. The top *M* individuals in the new population are chosen as the beginning population members after the fitness values of the individuals in the population are calculated and ordered in ascending order.

2.2.2 Adaptive Weights

The discoverer in the SSA algorithm determines the population's foraging area and direction, so the discoverer's location update iteration method is essential to the algorithm. However, the discoverer's location update iteration method in the early stage, based on the elite reverse strategy to bring the initial distribution of population members closer to the global optimal point, is more likely to succeed than the original random strategy to initialize the population. The range of the global best solution continues to narrow over numerous iterations, necessitating a strengthening of the discoverer's local search capabilities. To enhance the algorithm performance, an adaptive weight factor [18] is added to the discoverer position iteration formula. The weight factor fluctuates as the algorithm iterates, and at the start of an iteration when the algorithm has a stronger global search capability, the weight factor is larger. When the discoverer is close to the optimal solution, the weights are smaller and the algorithm's ability to search locally is improved.

$$x_{ip}^{k+1} = \begin{cases} x_{ip}^{k} + \omega \cdot \alpha \left(x_{ip}^{k} - f_g \right), & R_2 < ST \\ x_{ip}^{k} + Q \cdot L, & R_2 \geq ST \end{cases} \tag{7}$$

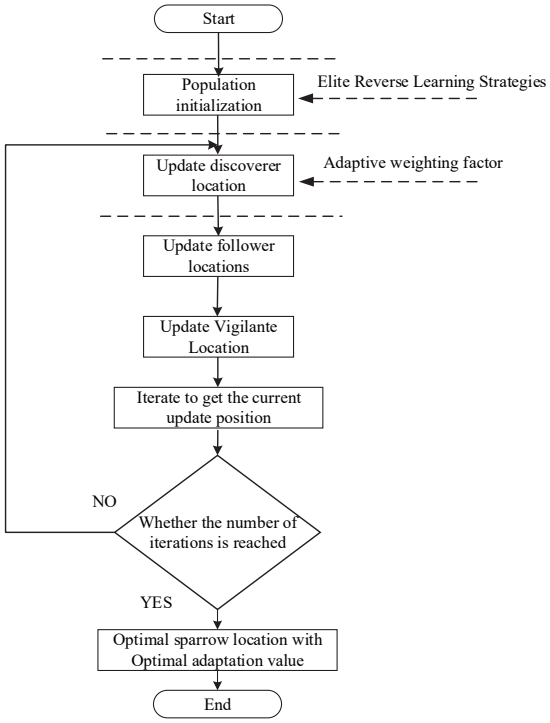


Figure 2 Flow chart of improved SSA algorithm

Which is: the adaptive weight factor $\omega = \cos\left(\left(\pi \cdot k\right)/\left(2 \cdot T\right)\right)$, R_2 represents the magnitude of the warning signal coming from the alerters, *ST* represent the established security threshold, α represent a random number between 0 and 1, *Q* represent a random number that follows the normal distribution, and *L* represent a 1 × *d*matrix with values for all 1 elements. The following figure depicts the flow of the optimization algorithm (Fig. 2):

2.2.3 Basis Function Test

This study chooses the Particle Swarm Optimization (PSO), Gray Wolf Optimization (GWO), and Sparrow Search Algorithm (SSA) to test six benchmark functions for comparison, respectively, to see if the upgraded sparrow algorithm (ESSA) performs better in finding the optimal performance. Taby. 1 contains information on the benchmark functions, where *f*1~ *f*3 represent single-peak benchmark functions and *f*4 ~ *f*6 represent multi-peak benchmark functions.

Table 1 Test function description

SN	Name	Domain	Theoretical Optimum
*f*1	Sphere	[−100, 100]	0
*f*2	Step	[−100, 100]	0
*f*3	Quartic	[−1.28, 1.28]	0
*f*4	Rastrigin	[−5.12, 5.12]	0
*f*5	Levy Function	[−50, 50]	0
*f*6	Levy Function N. 13	[−50, 50]	0

The population size was set to 30, the maximum number of iterations was 200, and the proportions of discoverers and warners were set to 70 percent and 20 percent of the total population of sparrows, respectively. PSO, GWO, SSA, and ESSA were each run 100 times. Tab. 2 shows that the mean value of ESSA on the search is tens of orders of magnitude higher than PSO and GWO, which have higher resilience, and is closest to the theoretical optimum for the single-peaked functions *f*1~ *f*3; ESSA and SSA both have the same standard deviation for the *f*1 function, but ESSA's mean value is much greater than SSA's. Both ESSA and SSA can search for the ideal value in *f*4 for the multi-peaked functions *f*4 ~ *f*6, although ESSA is more accurate at identifying the optimal value than the SSA approach. The experiments demonstrate that ESSA, based on the elite backward learning strategy, increases the diversity of the initial population and improves global search ability in the early iteration and local search ability in the late iteration by adaptive weighting factors to prevent the algorithm from entering local optimum in order for the ESSA algorithm to perform better in terms of search accuracy and speed.

Table 2 PSO, GWO, SSA and ESSA mean results

Algorithm / SN	PSO	GWO	SSA	ESSA
*f*1	$4.86E^{-04}$	$8.49E^{-46}$	$8.30E^{-177}$	$5.12E^{-216}$
*f*2	2.4067	1.2836	$2.43E^{-05}$	$1.96E^{-05}$
*f*3	$2.58E^{-02}$	$7.81E^{-03}$	$7.15E^{-04}$	$7.96E^{-05}$
*f*4	3.6423	$4.04E^{-01}$	0	0
*f*5	3.2513	$2.76E^{-01}$	$6.49E^{-06}$	$2.98E^{-06}$
*f*6	$1.43E^{+01}$	8.6376	$6.88E^{-03}$	$3.17E^{-06}$

Table 3 PSO, GWO, SSA and ESSA deviation results

Algorithm / SN	K-Means	Canopy-K-Means	Algorithm of this paper	ESSA
$f1$	$7.44E^{-05}$	$7.69E^{-46}$	0	0
$f2$	1.8206	$4.123E^{-01}$	$2.58E^{-05}$	$2.09E^{-05}$
$f3$	$1.93E^{-02}$	$4.86E^{-03}$	$6.68E^{-05}$	$7.24E^{-05}$
$f4$	4.4672	$3.12E^{-01}$	0	0
$f5$	3.8820	$8.07E^{-02}$	$9.91E^{-06}$	$3.45E^{-06}$
$f6$	$1.07E^{+01}$	$1.24E^{+01}$	$6.96E^{-03}$	$4.02E^{-06}$

2.3 Combining Morphological Operations for Portrait Foreground Processing

This study suggests using morphological techniques to treat the portrait's foreground in the following way in order to make the segmented image more aesthetically pleasing: To achieve edge denoising of the image's foreground information, first take the segmented mask map, mark the connected domains, calculate the distance between each connected domain and the maximum connected domain, set a threshold to eliminate the connected domains that are too far away, and smooth the image using morphological operations to reduce noise. The image is then morphologically processed to remove edge burr and provide a smoother, more realistic portrait (Fig. 3):

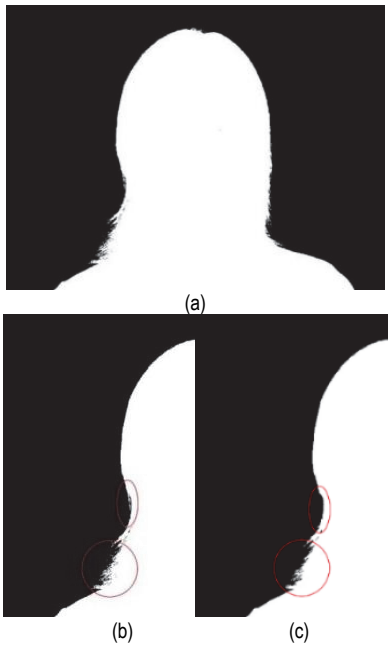


Figure 3 Portrait in the foreground. (a) Original image; (b) Partial original image; (c) Processed image

2.4 K-Means Algorithm

The most popular algorithm for picture segmentation in cluster analysis is K-Means. Iterate the positions of the cluster centrals using the mean until the clustering results are stable, at which point the iteration stops and the results are output [19]. The basic idea of the algorithm is to divide the sample into K cluster centrals for categorization using a given K value and random initial cluster points is carried out, and the Euclidean distance is used to calculate the distance between two pixel spots.

$$D_1 = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{8}$$

Flow of the algorithm is

(i) Enter K clustering centrals and the image to be segmented.

(ii) Randomly selected K initial clustering centrals are used.

(iii) In the K-th iteration, use equation (8) to calculate the distance between any two pixels and the centrals of the K clusters. Then, place the sample in the cluster with the shortest distance to the central.

(iv) Use the mean technique to update the class centroids.

(v) Repeat steps 3 and 4 above for each of the K clustering centrals. If the coordinates of clustering do not change, the iteration is stopped, and the segmentation result is returned.

The conventional K-Means algorithm continuously updates the K clustering centrals by employing a single mean approach and K initial clustering centrals that are chosen at random, as can be seen from the algorithm flow. This leads to unstable and subpar clustering outcomes. Additionally, as the image's pixel values are sensitive to lighting, when the light intensity changes, the computation of how similar two pixels are may misclassify some pixel points, producing unsatisfactory image segmentation results.

3 SIMULATION RESULTS AND ANALYSIS

The hardware for the simulation trials includes the MATLAB R2020B simulation program, an Intel(R) Core(TM) i5-1135G7 processor running at 2.40 GHz, 16GB of RAM, and Windows 10 as the operating system (64-bit). The Hong Kong Face Sketch Database CUHK Face Sketch Database (CUFS) dataset is used for the experimental photos in this study. The images from the dataset are chosen for segmentation and validation. All images are type *.jpg and their image sizes are 1024 × 768. To demonstrate that the optimal starting clustering points sought by this approach have a noticeable benefit in portrait clustering segmentation, this research proposes to use the traditional K-Means algorithm with the Canopy-K-Means algorithm.

3.1 Analysis of Segmentation Accuracy

The following three photos of short hair types were clustered and segmented using K-Means and Canopy-K-Means with the method in this research to confirm the validity of the algorithm on the effect of portrait segmentation (Fig. 4).

While clustering portraits, the head edge texture of short hair type portraits is more complicated, and the gap of portrait edge clustering is more visible, according to the segmentation impact of the three methods; when clustering backdrops, there are also clear differences in noise resistance: For instance, when segmenting the head edge of pictures with short hair, the accuracy rate is too low and there is a lot of noise in Fig. 4c, which depicts a segmentation mask map based on the K-Means technique. To improve the clustering effect in the K-Means algorithm, Canopy algorithm preprocessing is added to the segmentation mask in Fig. 4d so that the head edge of the portrait has substantially better head edge segmentation,

but there is still some noise in the top right of the head and the left and right shoulder regions when the background is dense. Based on the algorithm in this research, the segmentation mask in Fig. 4e. The accuracy of portrait edge clustering is maintained while the noise issue with the background clustering of the previous two algorithms is resolved in the figure.

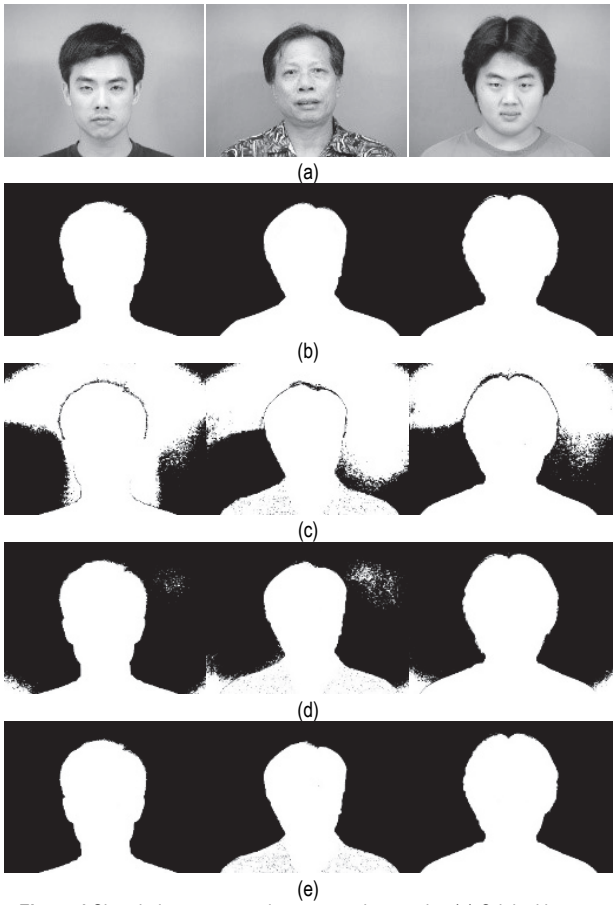


Figure 4 Short hair group portrait segmentation results: (a) Original image; (b) Standard segmentation chart; (c) K-Means algorithm to segment the mask map; (d) Canopy-K-Means algorithm to segment the mask map; (e) The algorithm of this paper divides the mask map

The segmented portrait mask map's pixel accuracy (PA) is introduced [20] to help illustrate how the algorithm in this study differs from other algorithms. PA is determined as follows:

$$PA = \frac{\sum_{x=0, x \in C_i} p_x}{\sum_{i=0}^{m} \sum_{j=0}^{n} p_{ij}} \tag{9}$$

where p_{ij} is the number of pixels in the algorithmic segmentation map portraiture, C_i is the number of classes, and p_x is the number of pixels that the standard segmentation map and the algorithmic segmentation map overlap.

Tab. 4 displays the pixel segmentation accuracy of the portraits segmented by the three segmentation algorithms mentioned above. The improved K-Means algorithm achieves pixel segmentation accuracy levels that are roughly 70% and 15% higher than those of the K-Means algorithm and Canopy-K-Means, respectively. The pixel segmentation accuracy is shown in Tab. 4. The serial numbers appearing in the table are the actual code names in the HKU Facial Sketch Database website (CUFS) dataset.

Table 4 Pixel segmentation accuracy of the three algorithms for the short hair group (%).

Algorithm / SN	K-Means	Canopy-K-Means	Algorithm of this paper
M0-033	22.44	77.32	98.83
M0-038	20.48	74.02	98.29
M0-045	29.44	85.06	98.93

The other nine images (from the CUHK Face Sketch Database) were divided into three groups for segmentation experiments: long hair, overexposure, and underexposure. The results are shown (which are meant to demonstrate the general efficacy of the approach suggested in this paper, Figs. 5, 6, and 7).

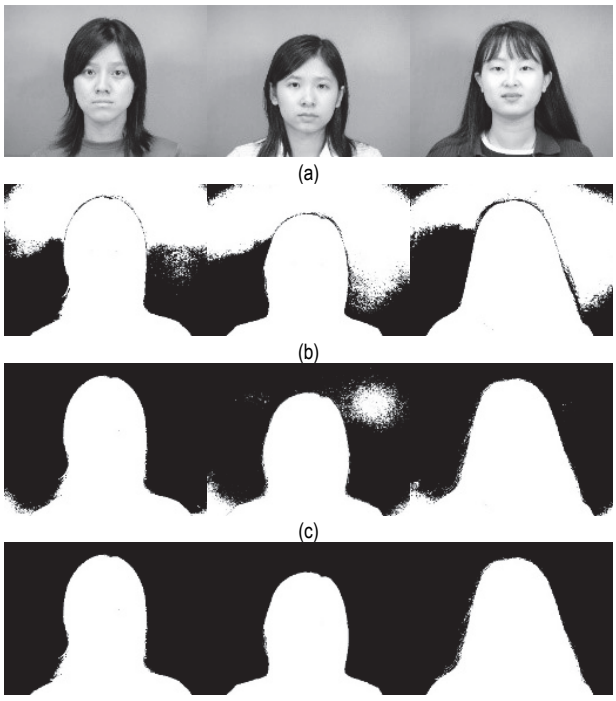


Figure 5 Portrait segmentation results for the long hair group: (a) Original image; (b) K-Means algorithm to segment the mask map; (c) Canopy-K-Means algorithm to segment the mask map; (d) The algorithm of this paper divides the mask map

The segmentation results for the long hair group and the short hair group in Fig. 5 are similar, and because the edges of long hair are smoother than those of short hair, the processing effect of the three algorithms on the edges of the figure is somewhat improved; the segmentation results for the underexposure group in Fig.6 demonstrate that when the brightness is insufficient, the segmentation results of the three algorithms are less accurate. Fig. 7 segmentation results for the group with excessive exposure demonstrate that when the brightness is adequate. The reflection of the portrait subject's clothing and a tiny area of background noise have an impact on the segmentation results of the Canopy-K-Means method, but the segmentation results of the K-Means and algorithm in this study are more stable. The pixel segmentation accuracy is shown in Tabs. 5, 6, 7.

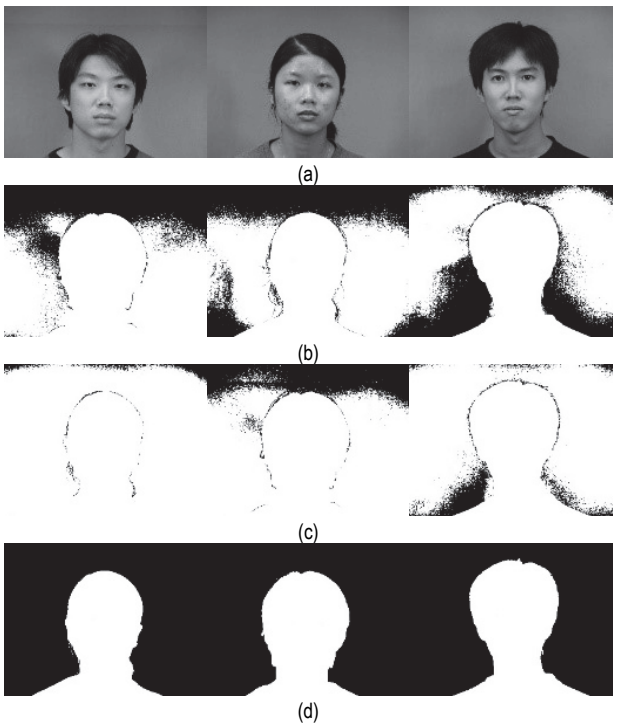


Figure 6 Underexposed group portrait segmentation results: (a) Original image; (b) K-Means algorithm to segment the mask map; (c) Canopy-K-Means algorithm to segment the mask map; (d) The algorithm of this paper divides the mask map

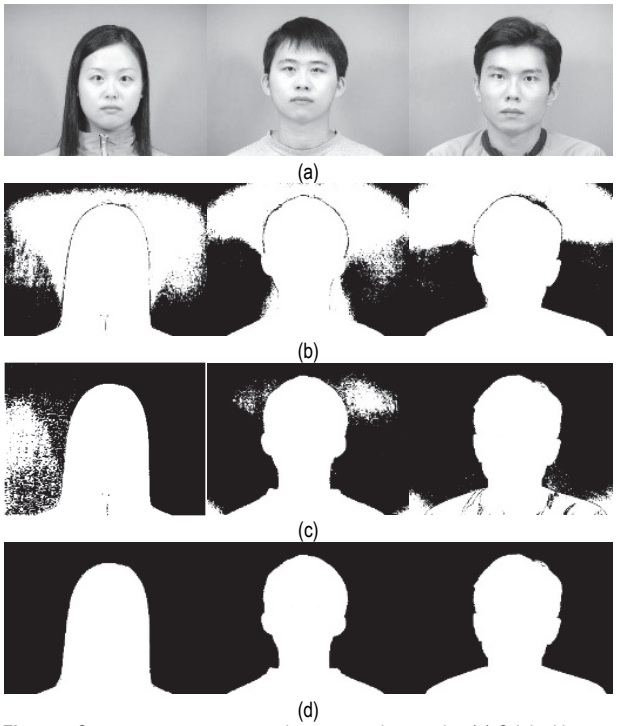


Figure 7 Overexposure group portrait segmentation results: (a) Original image; (b) K-Means algorithm to segment the mask map; (c) Canopy-K-Means algorithm to segment the mask map; (d) The algorithm of this paper divides the mask map

From Tabs. 5 and 7, it is clear that this algorithm's pixel segmentation accuracy for portraits with long hair and overexposure is around 70% and 20% higher than that of K-Means and Canopy-K-Means; from Table 6, we can see that in the underexposure group, the pixel segmentation accuracy of the portrait segmented by this algorithm is about 90% higher than that of K-Means and Canopy-K-Means and Canopy-K-Means by about 90%.

The experimental findings demonstrate that the improved SSA and K-Means portrait clustering algorithm suggested in this paper can successfully reduce the interference of lighting in portrait segmentation. Additionally, the segmentation results are better than those of the other two conventional algorithms, with greater portrait segmentation accuracy.

Table 5 Pixel segmentation accuracy of the three algorithms for the long hair group (%)

SN \ Algorithm	K-Means	Canopy-K-Means	Algorithm of this paper
F0-017	28.17	84.23	98.92
F0-023	26.64	79.85	98.80
F0-035	27.43	82.51	98.86

Table 6 Pixel segmentation accuracy of the three algorithms for the underexposure group (%)

SN \ Algorithm	K-Means	Canopy-K-Means	Algorithm of this paper
F1-015	9.14	5.43	97.14
M1-031	14.57	11.44	97.35
M1-036	5.30	3.78	96.06

Table 7 Pixel segmentation accuracy of the three algorithms for the overexposure group (%)

SN \ Algorithm	K-Means	Canopy-K-Means	Algorithm of this paper
F0-040	20.59	70.37	98.73
M0-075	25.18	81.73	98.95
M1-011	21.79	78.71	98.50

3.2 Analysis of Feature Similarity

One image from each of the four groups (long hair, short hair, overexposure, and underexposure) is selected and analyzed using K-Means, Canopy-K-Means, and this paper's algorithm for color quantization and calculation of its color image FSIM (feature similarity) value in order to visually compare the differences of portrait segmentation effect. The following is the formula for determining the similarity value of color picture attributes.

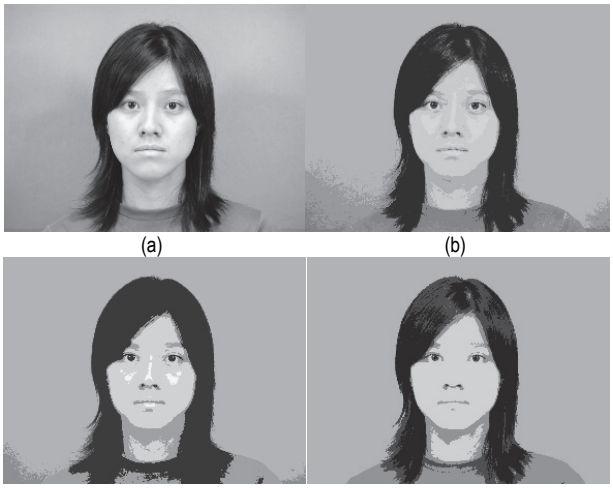
$$S_L = \left[\frac{2PC_1(x)PC_2(x)+T_1}{PC_1(x)^2+PC_2(x)^2+T_1}\right]^{\alpha} \left[\frac{2GM_1(x)GM_2(x)+T_2}{GM_1(x)^2+GM_2(x)^2+T_2}\right]^{\beta} \tag{10}$$

$$PC_m(X) = \max(PC_1(X), PC_2(X)) \tag{11}$$

$$FSIM = \frac{\sum_{x\in\Omega} S_L(x)\cdot PC_m(x)}{\sum_{x\in\Omega} PC_m(x)} \tag{12}$$

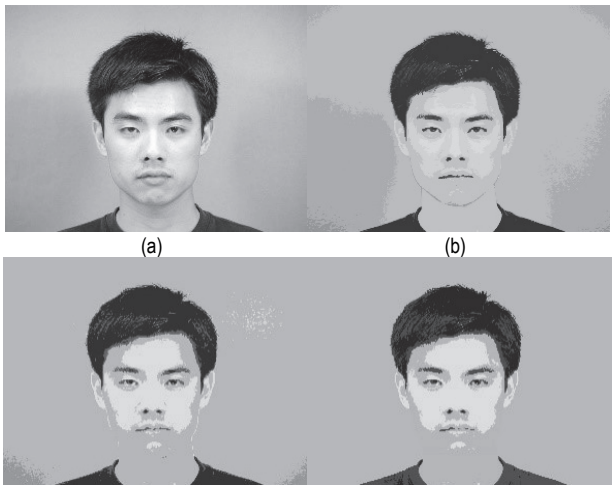
where, $PC(x)$ is the similarity of phase features, $PC_1(x)$ denotes the similarity of phase features of the first image, $PC_2(x)$ denotes the similarity of phase features of the second image; $GM(x)$ is the similarity of gradient features; T_1, T_2 are constants to increase the stability of S_L; $PC_m(x)$ is the maximum value between $PC_1(x)$,

$PC_2(x)$ used as the weight measurement factor; Ω denotes the whole image space domain.



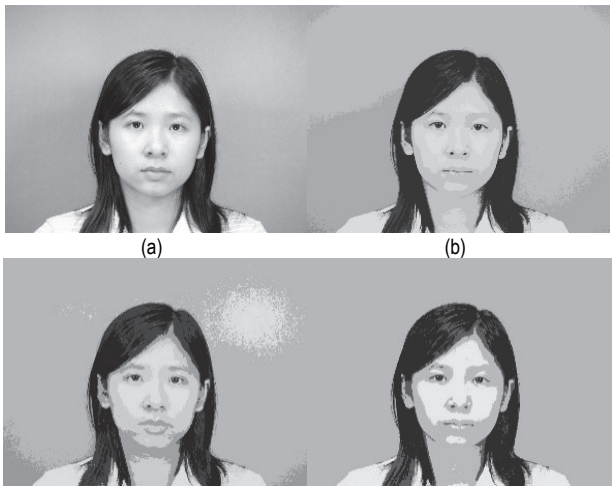
(a) (b) (c) (d)

Figure 8 Long hair: (a) Original image; (b) K-Means; (c) Canopy-K-Means; (d) Algorithm of this paper



(a) (b) (c) (d)

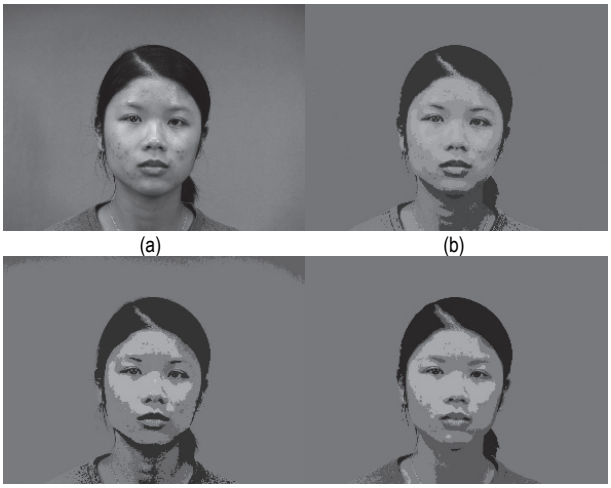
Figure 9 Short hair: (a) Original image; (b) K-Means; (c) Canopy-K-Means; (d) Algorithm of this paper



(a) (b) (c) (d)

Figure 10 Overexposure: (a) Original image; (b) K-Means; (c) Canopy-K-Means; (d) Algorithm of this paper

The clustering effect of the K-Means algorithm on edges is unsatisfactory in all four images of the experiment. According to the quantification results in Figs. 8b, 9b, 10b, and 11b, there is also a lot of noise on the background clustering, primarily because the clustering process of the K-Means algorithm is relatively simple and susceptible to the noise, causing the initial clustering center to fall into the local optimal solution. In comparison to the K-Means algorithm and the Canopy-K-Means algorithm, as shown in Figs. 8d, 9d, 10d, and 11d, the K-Means algorithm combined with the SSA algorithm to find the initial clustering center ensures the clustering accuracy of portrait edges and addresses the issue of noise when clustering the background in portrait segmentation (see Tab. 8):



(a) (b) (c) (d)

Figure 11 Underexposure: (a) Original image; (b) K-Means; (c) Canopy-K-Means; (d) Algorithm of this paper.

Table 8 FSIM values of color feature similarity for the three algorithms

SN \ Algorithm	K-Means	Canopy-K-Means	Algorithm of this paper
F0-017	0.683	0.713	0.792
M0-033	0.749	0.721	0.836
F0-023	29.44	0.746	0.824
F1-015	0.765	0.752	0.811

According to Tab. 8, which compares the FSIM values of color feature similarity for the images segmented by the three segmentation methods, the FSIM values of the images segmented by the algorithm in this study are roughly 10% higher than those of the other segmentation methods. Compared to the K-Means and Canopy-K-Means techniques, the algorithm in this paper is highest FSIM of 0.836 is closer to the theoretical maximum value.

3.3 Light Supplementation Experiments

In order to further verify the effect of the proposed algorithm for uneven illumination, wavelet proportional denoising is temporarily removed from the algorithm and divided into two parts for comparison experiments before and after illumination correction. A set of low-illumination images from the CUHK Face Sketch Database (CUFS) dataset in the Hong Kong Face Sketch Database website is used for the experiments as shown in Figs. 12 and 13.

The PA (pixel segmentation accuracy) of the portraits segmented by the above three segmentation algorithms is shown in Tab. 9. It is obvious that the pixel segmentation accuracy of the portraits segmented by the algorithm in this paper before low illumination correction is about 65% and 60% higher than that of the K-Means algorithm and Canopy-K-Means, respectively.

Table 9 Accuracy of pixel segmentation before low-light portrait illumination correction (%)

SN \ Algorithm	K-Means	Canopy-K-Means	Algorithm of this paper
F1-015	5.43	9.14	71.32
M1-031	11.44	14.57	79.85
M1-036	3.78	5.30	62.06

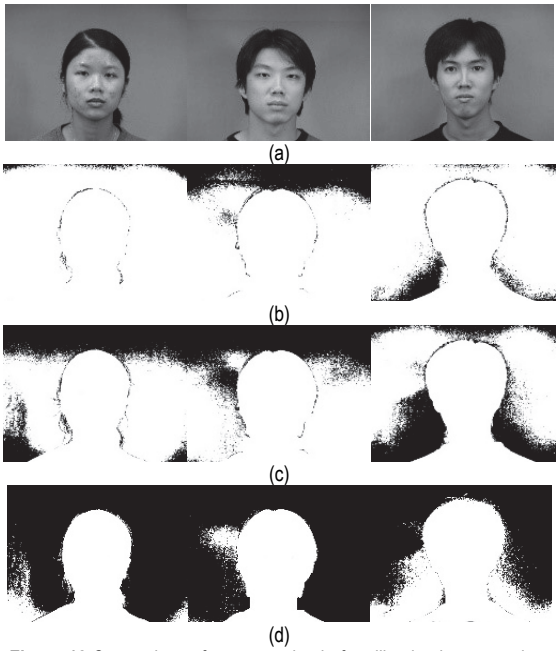


Figure 12 Comparison of segmentation before illumination correction: (a) Original image; (b) K-Means algorithm to segment the mask map; (c) Canopy-K-Means algorithm to segment the mask map; (d) The algorithm of this paper divides the mask map

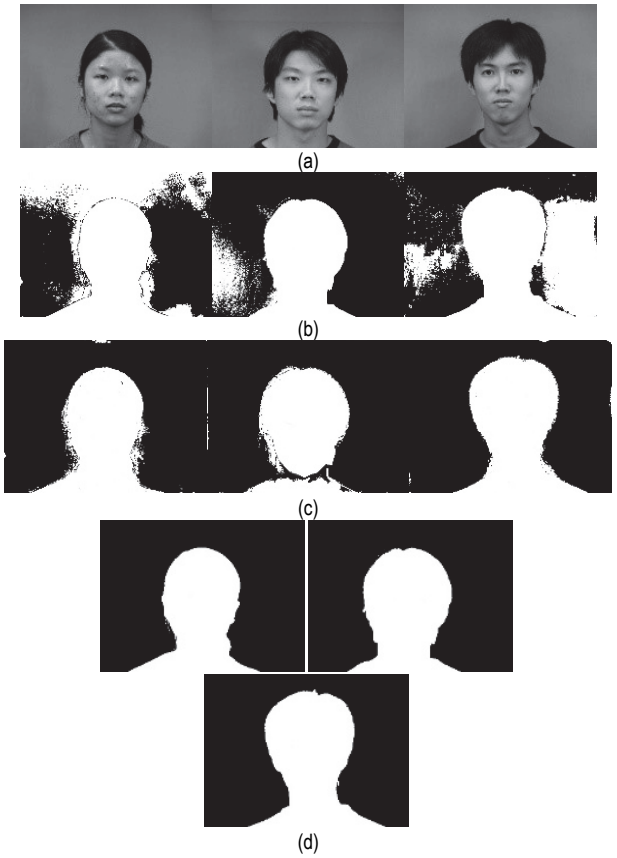


Figure 13 Comparison of segmentation after illumination correction: (a) Original image; (b) K-Means algorithm to segment the mask map; (c) Canopy-K-Means algorithm to segment the mask map; (d) The algorithm of this paper divides the mask map

From Tab. 10, we can see that the pixel segmentation accuracy of the human portrait segmented by the three algorithms after light correction is about 30%, 70%, 29% higher than that of light correction before separately, and the pixel segmentation accuracy of the human portrait segmented by the algorithm in this paper after light correction is about 62% and 17% higher than that of the K-Means algorithm and Canopy-K-Means, respectively.

Table 10 Accuracy of pixel segmentation after low-light portrait illumination correction (%)

SN \ Algorithm	K-Means	Canopy-K-Means	Algorithm of this paper
F1-015	30.59	81.73	98.14
M1-031	45.18	78.17	98.35
M1-036	31.79	84.23	97.06

4 CONCLUSIONS

Aiming at the mis-segmentation of frontal portraits caused by distorted noise spots due to the influence of individual appearance and lighting conditions, in this paper, an accurate portrait segmentation method based on the combination of wavelet proportional shrinkage and enhanced sparrow search (SSA) clustering algorithm is proposed. Firstly, the wavelet proportional shrinkage denoising method is combined with the color space transformation to deal with the abnormal noise in the image to preprocess the image, and secondly, in order to improve the traditional sparrow algorithm which is easy to fall into the local optimal solution, the location of the optimal sparrow is searched through the introduction of the elite inverse-planning population as well as the adaptive weight adjustment factor. Then the clustering and the best initial center are determined using the location of the best sparrow, which minimizes the influence of the anomalies on the subsequent K-Means method segmentation and ensures the correctness of the background clustering and portrait edge segmentation.

The experimental outcomes demonstrate the effectiveness of the enhanced SSA in conjunction with the K-Means portrait clustering algorithm for replacing portrait backgrounds. The segmentation accuracy of the image mask map and the feature similarity of color photos are effectively improved, and this can lead to a more precise segmentation of the target portraits, particularly for color portraits with strong differences. The whole average time of this algorithm is about 30 seconds, which is not instantaneous but is more efficient than the artificial work that results in the same effect. The algorithm in this work increases the segmentation accuracy of the algorithm while sacrificing operation speed to some extent, and the technique in this paper can be further investigated to increase operational efficiency.

Acknowledgments

This paper is supported by The National Natural Science Foundation of China (62173127).

5 REFERENCES

[1] Wang, Z., Wang, E., & Zhu, Y. (2020). Image segmentation evaluation: a survey of methods. *Artificial Intelligence Review*, *53*, 5637-5674.

https://doi.org/10.1007/s10462-020-09830-9
[2] Ralevic, N. & Paunovic, M. (2021). Applications of the Fuzzy Metrics in Image Denoising and Segmentation. *Tehnicki vjesnik-Technical Gazette*, *28*(3), 819-826. https://doi.org/10.17559/TV-20200305075136
[3] Riad, R., Ros, F., Hajji, M. E. et al. (2022). An industrial portrait background removal solution based on knowledge infusion. *Applied Intelligence*, *52*, 1159-11605. https://doi.org/10.1007/s10489-021-03099-3
[4] Babu, E. S., Barthwal, A., & Kaluri, R. (2023). Sec-edge: Trusted blockchain system for enabling the identification and authentication of edge based 5G networks. *Computer Communications*, *199*, 10-29. https://doi.org/10.1016/j.comcom.2022.12.001
[5] Begum, M. B., Deepa, N., Uddin, M., Kaluri, R., Abdelhaq, M., & Alsaqour, R. (2023). An efficient and secure compression technique for data protection using burrows-wheeler transform algorithm. *Heliyon*, https://doi.org/10.1016/j.heliyon.2023.e17602
[6] Le, Z., Wang, J., & An, Z. (2020). FCM fuzzy clustering image segmentation algorithm based on fractional particle swarm optimization. *Journal of Intelligent & Fuzzy Systems*, *38*(4), 3575-3584. https://doi.org/10.3233/JIFS-179580
[7] Khan, A. et al. (2019). Color image segmentation using genetic algorithm with aggregation-based clustering validity index (CVI). *Signal, Image and Video Processing*, *13*, 833-841. https://doi.org/10.1007/s11760-019-01419-2
[8] Jaiswal, V., Varsha, S., & Sunita, V. (2019). An implementation of novel genetic based clustering algorithm for color image segmentation. *Telkomnika (Telecommunication Computing Electronics and Control)*, *17*(3), 1461-1467. https://doi.org/10.12928/telkomnika.v17i3.10072
[9] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference*. https://doi.org/10.1007/978-3-319-24574-4_28
[10] Xi, C., Qi, D., & Shen, J. (2019). Boundary-aware network for fast and high-accuracy portrait segmentation. arXiv preprint arXiv:1901.03814.
[11] Himanshu, M. et al. (2021). A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets. *Multimedia Tools and Applications*, 1-26.
[12] Jiankai, X. & Shen, B. (2020). A novel swarm intelligence optimization approach: sparrow search algorithm. *Systems science & control engineering*, *8*(1), 22-34. https://doi.org/10.1080/21642583.2019.1708830
[13] Rana Muhammad, A. et al. (2021). Prediction of hydraulics performance in drain envelopes using Kmeans based multivariate adaptive regression spline. *Applied Soft Computing*, *100*, 107008. https://doi.org/10.1016/j.asoc.2020.107008
[14] Zikai, W., Huang, X., & Zhu, D. (2022). A Multistrategy-integrated learning sparrow search algorithm and optimization of engineering problems. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/2475460
[15] Manel, R. et al. (2019). Wavelet transform application for/in non-stationary time-series analysis: a review. *Applied Sciences*, *9*(7), 1345. https://doi.org/10.3390/app9071345
[16] Li, G., Gan, Y., & Shi, J. (2022). A novel intelligent denoising method of ecg signals based on wavelet adaptive threshold and mathematical morphology. *Applied Intelligence*, *52*(9), 10270-10284. https://doi.org/10.1007/s10489-022-03182-3
[17] Jie, M., Hao, Z., & Sun, W. (2022). Enhancing sparrow search algorithm via multi-strategies for continuous optimization problems. *Information Processing & Management*, *59*(2), 102854. https://doi.org/10.1016/j.ipm.2021.102854
[18] Weibo, L. et al. (2019). A novel sigmoid-function-based adaptive weighted particle swarm optimizer. *IEEE transactions on cybernetics*, *51*(2), 1085-1093. https://doi.org/10.1109/TCYB.2019.2925015
[19] Sinaga, K. P. & Yang, M.-S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access*, *8*, 80716-80727. https://doi.org/10.1109/ACCESS.2020.2988796
[20] Shengyuan, L., Zhao, X., & Zhou, G. (2019). Automatic pixel-level multiple damage detection of concrete structure using the fully convolutional network. *Computer-Aided Civil and Infrastructure Engineering*, *34*(7), 616-634. https://doi.org/10.1111/mice.12433

Contact information:

Xiang WU
(Corresponding author)
College of Electrical Engineering,
Henan University of Technology,
Zhengzhou, China, 450000
Interdisciplinary Creative Application Research,
GDPOLAR Co-work Space, Zhengzhou, China, 450000
E-mail: xiangw@haut.edu.cn

Yuanhao MA
College of Electrical Engineering,
Henan University of Technology,
Zhengzhou, China, 450000

Hao LIAN
College of Electrical Engineering,
Henan University of Technology,
Zhengzhou, China, 450000

Xiang FANG
College of Electrical Engineering,
Henan University of Technology,
Zhengzhou, China, 450000

Tianfei CHEN
Key Laboratory of Grain Information Processing and Control,
Ministry of Education, Henan University of Technology,
Zhengzhou, China, 450000