An Automated Image Segmentation Framework Using Fractional Calculus and Improved Pigeon Swarm Optimization with 2D Otsu Thresholding

Hua Qiang
Department of Basic Courses Teaching, Yinchuan University of Energy, Yinchuan, 750105, China
E-mail: huaqianghua@outlook.com

Keywords: fractional calculus, image segmentation, pigeon swarm algorithm, chaotic mapping, unsupervised learning

Received: October 14, 2025

Research has proposed an automated image segmentation algorithm AIS-IPTSA-FC, which combines the U-Net active contour model enhanced by fractional calculus (FC), the improved traditional pigeon swarm algorithm (IPTSA), and the two-dimensional Otsu algorithm (2D-OA). This method has been validated on the SiftFlow and OASIS-3 datasets. The experimental results demonstrated that this hybrid algorithm achieved a segmentation accuracy of over 0.99 on the SiftFlow dataset, surpassing mainstream algorithms. In practical applications, this method performed well: in natural landscape images, the average values of structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), and mean squared error (MSE) were 0.36, 10.79, and 5501.61, respectively. In medical images, the corresponding indicators were 0.0086, 6.84, and 7976.47, respectively. The above results demonstrate that the research method can effectively achieve image segmentation under complex conditions and provide a reliable foundation for the development of multi domain intelligent systems.

Povzetek: Raziskava predstavlja nov hibridni algoritem AIS-IPTSA-FC za segmentacijo slik, ki dosega zelo visoko natančnost in je primeren za kompleksne inteligentne sisteme.

1 Introduction

In the field of computing, image segmentation (IS) plays a crucial role. IS involves decomposing a complex scene or object into subregions or objects that are easier to process and analyze. This is a fundamental support for subsequent tasks such as target detection, tracking, and recognition [1-3]. According to the data, the market for picture segmentation services is expected to grow at a compound annual rate of 9.9%, reaching $1.865 billion by 2031. Moreover, the Chinese IS market is expected to exceed RMB 58 billion by 2025, with a CAGR of about 18% [4-5]. IS is not only a fundamental task in computer vision, but its applications also directly support multiple sustainable development goals. For example, accurate medical IS can improve the efficiency of early disease diagnoses, such as Alzheimer's disease. This is significant for ensuring medical equity in areas with limited resources. In the construction of smart cities, optimizing energy use through traffic monitoring and infrastructure analysis can promote sustainable urbanization. In addition, IS plays a crucial role in helping the system better understand user intent and behavior. More specifically, the system can more accurately identify and understand key information in an image by segmenting it into multiple regions with similar features. For example, IS in intelligent monitoring systems can help distinguish between the foreground and background. This allows the system to more accurately detect and track target objects, such as pedestrians or vehicles. In the auto drive system, IS can help vehicles identify roads, traffic signs and obstacles, so as to navigate more safely. Through these applications, IS not only improves the automation level of the system, but also provides users with a more natural and smooth service experience. For example, intelligent monitoring systems can more accurately identify abnormal behavior and issue alerts in a timely manner [6]. Although deep learning (DL) technology has advanced significantly in the field of IS, situations involving complicated backgrounds, shifting illumination, and small targets still require improvements in IS accuracy. The efficiency of picture segmentation in applications is additionally constrained by its high computational resource requirements [7]. Overall, the current research questions focus primarily on enhancing edge preservation and noise suppression capabilities in complex images, efficiently determining segmentation thresholds, and achieving excellent performance across domains. To address the above problems, the study introduces fractional calculus (FC) and constructs an improved traditional pigeon swarm algorithm (IPTSA) based on FC. The study also combines an automated IS based on the IPTSA-FC algorithm (AIS-IPTSA-FC). The aim of the research is to reduce the time-consuming and costly work of the IS method during training, and to facilitate the understanding and analysis in subsequent applications. The innovation of the research mainly lies in

two aspects. On the one hand, it proposes an IPTSA-FC algorithm, which is essentially an unsupervised optimization strategy. Unlike U-Net's feature extraction (FE), which is supervised, this component does not require data labeling to optimize the search process itself. This can reduce costs associated with manual threshold adjustment or optimization. Second, the AIS-IPTSA-FC algorithm is intended to enhance IS performance and address the issue of complicated background interference. The contribution of this research is the proposed IPTSA-FC algorithm. It is an unsupervised learning strategy that does not require supervision of the labeled data, thus significantly reducing the cost of data labeling. By using the two-dimensional Otsu algorithm (2D-OA) to determine the optimal threshold, the anti noise performance of IS has been improved, which can effectively cope with complex background interference.

2 Related work

FC is a mathematical tool that extends the traditional integer order calculus and its ability to deal with systems that are memristive and non localized gives it an advantage over traditional methods in describing complex features of images. Numerous scholars have analyzed and discussed this in depth. Ozarslan [8] introduced the Stancu variant of the modified Bernstein-Kantorovich operator and analyzed its simultaneous approximation properties. Then, a new operator was constructed using the above auxiliary operator, and the main theorem of FC operator approximation was proven. This theorem yielded approximation results for various FC algorithms. Abd-Elmageedet al. [9] gave a similar definition of k-order τ-Gaussian hypergeometric matrix functions and analyzed their analytical properties. Moreover, the study developed new k-order differential integral operators for matrix functions based on the FC tachyphena. Experimental results verified the special case of this hypergeometric matrix function generalization. Yang et al. [10] merged the Cauchy problem from nth order difference equations with a new fractional difference with an exponential kernel function based on FC to formulate propositions and numerical schemes for one another. Finally, the fractional order linear difference equation was given and the exact solution was given using the new discrete Mittag-Leffler function. Simsek et al. [11] designed a new modeling and forecasting method which uses FC with multi-depth evaluation method. The data of air passengers, air cargo, available seat kilometers, number of flights, and destinations of the eight countries with the busiest airports between 2011 and 2019 were also selected for the experiment. The experimental results indicated that the largest error was the air transport cargo coefficient for Germany.

One of the fundamental technologies in the field of computer vision, IS is crucial for many applications. Furthermore, the expressive capacity of picture segmentation and classification has greatly increased with the advancement of DL technology. In response to the shortcomings of the arithmetic optimization technique, which was ineffective and prone to falling into local optimal solutions (LOS), Li et al. [12] added cosine factors and inertia weights to improve the algorithm's stability. Furthermore, chaotic mapping at the Xunze stage enhanced the convergence speed. The outcomes revealed that the research method was more robust and had better segmentation performance compared with other heuristic algorithms. In an effort to address the issue of the clustering algorithm's sensitivity to noise and propensity to fall into local optimum, Li et al. [13] developed an IS technique based on the super-pixel method and sparrow search algorithm. Additionally, the research introduced the chaos strategy for optimization in an attempt to increase the quality of the outcomes. Real-world image experiments showed that the research approach can successfully cut down on time. When using the Skyhawk optimization technique to solve increasingly complicated optimization problems, Guo et al. [14] sought to address the drawbacks of low solution accuracy and local optimal stagnation. Consequently, a hybrid chaotic mapping mechanism with a simulated annealing technique for optimization was presented in the study. The CEC2005 test function's experimental results confirmed the study method's viability and efficacy as well as its generally strong segmentation performance. To improve the chimpanzee optimization technique and prevent local extremes, Li et al. [15] used chaotic initialization with Cauchy mutation. This improved the starting population's diversity and uniformity of distribution. The studied approach exhibited superior segmentation outcomes in IS, according to experimental results based on visual and particular metrics. From this, a summary table of relevant work can be obtained, as shown in Table 1.

Table 1: Summary table of related work

Reference number	Research method	Performance results & metrics	Advantages	Shortcomings
[8]	ariant of modified Bernstein-Kantorovich operators for approximating fractional calculus operators.	Proved simultaneous approximation properties and provided the specific order of approximation.	Provides a solid mathematical foundation for the numerical approximation of fractional operators.	Focuses on pure mathematical derivation, not on engineering applications like image processing.
[9]	Defined k-analogue of τ-Gauss hypergeometric	Experimentally verified the	Extends the theoretical	Research is centered on abstract

	matrix functions and constructed associated fractional calculus operators.	correctness of the proposed hypergeometric matrix function generalization.	framework of FC in the domain of matrix functions.	mathematics, with no demonstrated potential for solving practical problems.
[10]	Proposed a new fractional difference with an exponential kernel and corresponding numerical schemes.	Provided exact solutions for fractional-order linear difference equations and demonstrated asymptotic stability of the numerical scheme.	Provides effective theoretical tools for the analysis and computation of discrete fractional systems.	The focus is on mathematical modeling and numerical analysis of discrete systems, disjoint from image processing tasks.
[11]	Proposed a modeling/forecasting method combining a multi-deep assessment methodology and fractional calculus.	Largest forecasting error was observed for key transportation metrics.	Provides a case study for applying FC to time-series forecasting.	Prediction accuracy needs systematic improvement, and the method does not involve image data analysis.
[12]	Introduced cosine factors, inertia weights, and chaotic mapping into the arithmetic optimization algorithm for multi-threshold image segmentation.	Segmentation accuracy ~0.91; showed better robustness and segmentation performance compared to baseline algorithms.	Effectively improves the convergence speed and stability of the original algorithm.	Lacks sufficient robustness for complex backgrounds and strong noise, and lacks an advanced feature representation mechanism.
[13]	Combined the superpixel method with a sparrow search algorithm employing a chaos strategy for image segmentation.	F-measure ~0.73; significantly reduced computation time.	Effectively improves computational efficiency via superpixel pre-processing.	Sensitive to noise, prone to local optima, and weak in detail preservation for complex scenes.
[14]	Proposed a falcon optimization algorithm with a hybrid chaotic mapping mechanism and simulated annealing.	PSNR ~28.5 dB; demonstrated powerful global search capability on test functions.	High solution accuracy and good segmentation performance.	Prone to loss of details when handling complex textures and edges; poor adaptation to intensity inhomogeneity in medical images.
[15]	Used chaotic initialization and Cauchy mutation to improve the chimp optimization algorithm.	SSIM ~0.75; showed superiority in both visual and quantitative metrics.	Significantly enhances population diversity and the algorithm's ability to escape local optima via improved initialization.	The inherent robustness of the segmentation model itself against noise and intensity inhomogeneity is not fundamentally improved.

In summary, the current research results of FC are mainly concentrated in the field of mathematics, and the IS algorithms have also made great progress. However, it is difficult for the relevant methods to take into account both the noise suppression and the preservation of the key features of the image at the same time. Therefore, the research introduces FC and chaotic mapping to make the IS algorithm better adapted to the image features, and thus the AIS-IPTSA-FC algorithm is constructed.

3 Automated IS algorithm based on IPTSA-FC

To effectively solve the problems of noise, blur, uneven lighting, and target complexity in IS, this study proposes an automated IS algorithm based on FC called

AIS-IPTSA-FC. The algorithm includes multiple key steps. First, an active contour model is constructed based on FC to improve the representation of image features. Second, the traditional pigeon swarm algorithm (PSA) is optimized to enhance its global search capability and convergence speed. Finally, the 2D-OA is used to further enhance the noise resistance and robustness of IS. These steps are interrelated and together form a complete IS framework aimed at improving the accuracy and efficiency of IS.

3.1 FC-based active contour modeling

IS is an important task in the field of computer vision, which is mainly performed by dividing an image into multiple domains to represent objects, scenes or other features. It can help related people to better understand the content in the image, and then perform more accurate object recognition, scene understanding and other computer vision tasks [16-18]. However, IS tasks face great challenges. First, the high dimensionality of image data, which contains a large amount of spatial and color information, makes IS very complex [19-20]. The second is the category imbalance, in practice specific categories of objects or features are much less than other categories of objects or features, which can cause the classifier to be biased towards recognizing more categories. In contrast, convolutional networks for biomedical IS (U-Net) is a DL network for IS, which requires a small quantity of training images to produce accurate classification results. Although U-Net uses prior knowledge from supervised pre-training, it can perform the entire segmentation process for new images without manual labeling. Thus, it constitutes an unsupervised optimization framework. Figure 1 roughly depicts the U-Net segmentation network's structure.

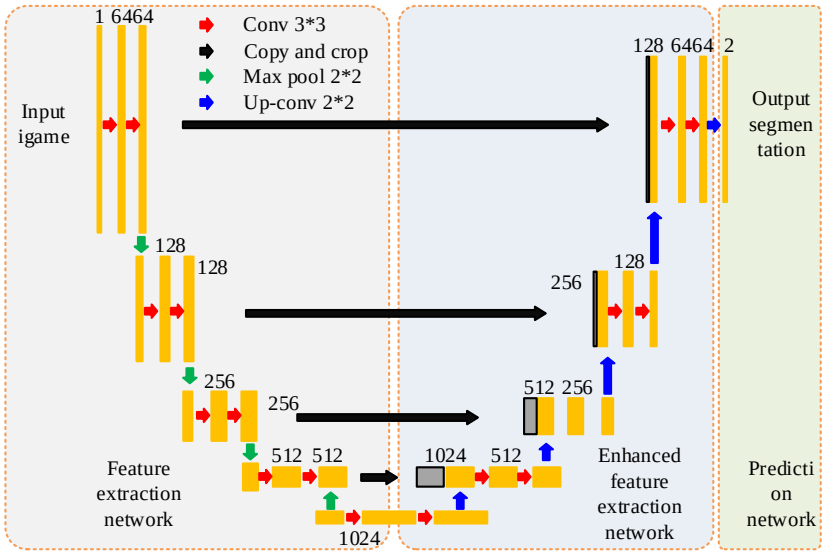

Figure 1: Schematic diagram of U-Net segmentation network structure

In Figure 1, the network contains an encoding-decoding structure with basic symmetry and contains the following three main parts. The backbone extraction portion, which comes first, adheres to the stacking of convolution and maximal pooling principles of the standard convolutional network architecture. The backbone FE portion can be used to obtain five initial effective feature layers. The extremely powerful FE portion is the second section. To create a feature layer with the same number of channels as the first preliminary effective feature layer, the five preliminary effective feature layers that are previously generated are progressively up-sampled, spliced, and fused. The third part is the prediction part. Convolution operation is performed on the feature layer obtained in the second part to classify each pixel point, and the final IS result map can be obtained.

In U-Net, the number of channels of the convolutional layers is twice the number of input channels, except for the first convolutional kernel channel used for the convolutional layer connecting the inputs, which is 64. The above setting is to compensate for the loss due to downsampling. The first four preliminary effective feature layers acquired in the first part are cropped in order to be spliced with the feature layers acquired by up-sampling. The U-Net network structure is designed to preserve image detail information. However, in practical applications, the obtained images are heterogeneous due to the non-uniformity of sensor signals and light response, as well as the complexity of lighting in the environment. This further increases the difficulty of IS due to the decrease in quality. However, the single U-Net network model mentioned above cannot cope with these problems. Compared to traditional integer-order calculus, FC can better handle complex structures and noise in images due to its non-local properties and memory effects. This improves the accuracy and robustness of feature representation, which is crucial for subsequent IS tasks by preserving the edge and texture details of the image. Therefore, the study is based on the U-Net, which has been improved by FC. Thus, an active contour model based on FC can be obtained. Its structure is schematically

shown in Figure 2.

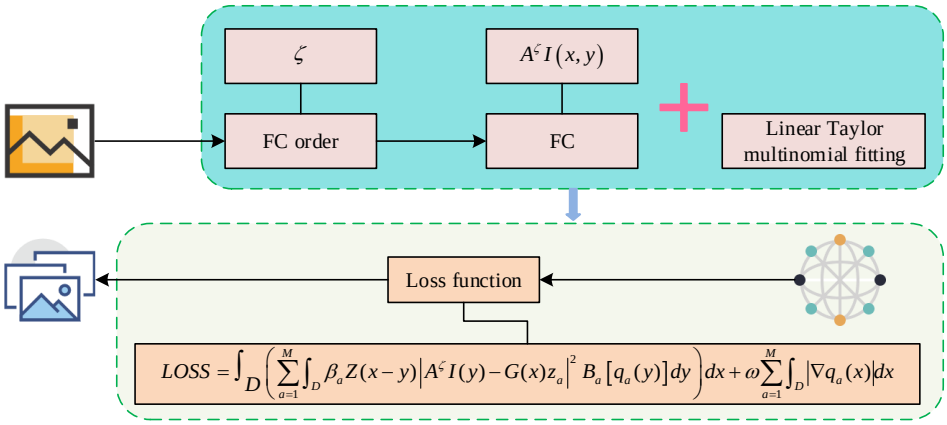

Figure 2: Schematic diagram of activity contour model structure based on FC

In Figure 2, $A^{\zeta}I(x,y)$ denotes the ζ order discretization of image $I(x,y)$. Specifically, the prediction obtained through the U-Net model is processed by the Softmax function to obtain Equation (1).

$$B_a(q_a(\psi))=e^{q_a(\psi)}\Big/\sum_{n=1}^{M}e^{q_k(\psi)},\quad a=1,2,3,4 \quad (1)$$

In Equation (1), $q_a(\psi)$ is the output of the previous layer. $\psi(x,y)\in D$, D, and M are the entire set of pixels and the number of categories in the region where the image is located. The mathematical expression of $A^{\zeta}I(x,y)$ is shown in Equation (2).

$$A^{\zeta}I(x,y)=\sqrt{\left[A_xI(x,y)\right]^2+\left[A_yI(x,y)\right]^2} \quad (2)$$

In Equation (2), $A_xI(x,y)$ and $A_yI(x,y)$ correspond to the FC of $I(x,y)$ in directions x and y. The FC-based loss function $LOSS$ for the active contour model is calculated in Equation (3).

$$LOSS=\int_D\left(\sum_{a=1}^{M}\int_D\beta_a Z(x-y)\left|A^{\zeta}I(y)-G(x)z_a\right|^2 B_a\left[q_a(y)\right]dy\right)dx+\omega\sum_{a=1}^{M}\int_D|\nabla q_a(x)|dx \quad (3)$$

In Equation (3), β_a and ω are the weight parameters as well as the weights of the boundary range constraint terms, respectively. $Z(x-y)$ represents the kernel function. $G(x)$ represents the Taylor-based fitting function. z_a is the indicator function. To further enhance the high-frequency signals while realizing the preservation of the low and medium frequency signals, the study processes the image by the fractional order enhancement filtering method. The study takes the two-dimensional fractional order as an example, which leads to the FC filtering example display map, as shown in Figure 3.

f(k)	⋯	⋯	⋯	f(k)	⋯	⋯	⋯	f(k)
⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
⋯	⋯	⋯	f(1)	f(1)	f(1)	⋯	⋯	⋯
f(k)	⋯	⋯	f(1)	f(0)	f(1)	⋯	⋯	f(k)
⋯	⋯	⋯	f(1)	f(1)	f(1)	⋯	⋯	⋯
⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
f(k)	f(k)	f(k)	f(k)	f(k)	f(k)	f(k)	f(k)	f(k)

Figure 3: Example display of FC filtering

Figure 3 shows the effect of FC filtering. A mask operator for FC filtering is generated by setting the filtering range to 0, 1. The x axis is set to positive at 0°, and the direction is taken every 45°. This filtering method can effectively enhance the high-frequency signal of the image while preserving the low-frequency and mid frequency signals, thus preserving the key details of the image while denoising. This is important because it allows subsequent IS algorithms to more accurately identify and segment target regions, especially in images with complex textures and edges. Specifically, the filtering range is $(2k+1)\times(2k+1)$. The x axis is also set to be 0° in the positive direction, and the direction is taken at 45° intervals. This results in the mask operator $f(k)$ of fractional order, which is then normalized as in Equation (4).

$$f'(k)=f(k)\Big/\sum_{0}^{m}f^{(k)},k=0,1,2,3,\ldots \quad (4)$$

In Equation (4), m represents the maximum value of the template desirable in the filter and $f'(k)$ represents the actual coefficients.

3.2 Improved traditional PSA based on FC

After constructing an active contour model based on FC, the study introduces a chaos algorithm to optimize the traditional PSA. The study also designs the IPTSA-FC algorithm to enrich population diversity and improve the algorithm's global search ability and convergence speed, thereby making it more robust in complex optimization environments. Among these, the classic PSA is an algorithm for group intelligence optimization that draws inspiration from the homing behavior of pigeons in the wild. That is, pigeons will use various navigation tools such as the earth's magnetic field, the sun, landmarks, etc., to find the nest by segmented navigation during the flight. Due to the randomness and adaptive ability in the algorithm, it has strong robustness and can be adapted to different optimization environments. Moreover, its realization process is relatively simple, so it has high practical value. The precise procedure is as follows: based on the number of iterations, the behavior of the pigeon is split into two phases. The proportion of stage 1 to the whole number of iterations is NcRate, which generally takes the value of 0.75. At this time, it is necessary to calculate the pigeon's velocity as well as the new position through the pigeon's position and the target, as in Equation (5).

$$\begin{cases} v_{\text{new},a}^{d} = v_{a}^{d} \cdot e^{-R \cdot iter} + \text{rand}\left(s_{\text{best}}^{d} - s_{a}^{d}\right) \\ s_{\text{new},a}^{d} = s_{a}^{d} + v_{\text{new},a}^{d} \end{cases} \quad (5)$$

In Equation (5), $v_{\text{new},a}^{d}$ and $s_{\text{new},a}^{d}$ are the new velocity and new position of the a th pigeon in the d th dimension, respectively. R denotes a constant, which generally takes the value of 0.2, and is used to control the decay of the velocity update. *iter* and *rand* are the current quantity of iterations and a uniform random quantity of [0, 1], respectively. s_{best}^{d} stands for the current finding of the optimal position in the d th dimension. Stage 2 is relatively complex and requires first sorting the population and dividing it equally into two groups. The better performing group keeps its position unchanged and provides the corresponding position and fitness value as parameters to help the worse performing group determine its new position, as shown in Equation (6).

$$Sc^{d} = \frac{\sum_{a=1}^{N/2}\left(s_{i}^{d} \cdot Fit\left(S_{a}\right)\right)}{\frac{N}{2} \cdot \sum_{a=1}^{N/2} Fit\left(S_{a}\right)} \quad (6)$$

In Equation (6), Sc^{d} is the center position of the pigeon in the better performing group in the d th dimension. $Fit\left(S_{a}\right)$ represents the adaptation value of the a th pigeon. N is the total quantity of pigeons. The updated position expression, as in Equation (7).

$$s_{\text{new},i}^{d} = s_{i}^{d} + \text{rand}\left(Sc^{d} - s_{i}^{d}\right) \quad (7)$$

Lastly, unless the termination condition is satisfied, the previous stages are repeated. Furthermore, the work presents the CA as a means of optimizing the conventional PSA. This can prevent premature convergence, help the classic PSA jump out of the LOS, and somewhat enhance the algorithm's computing performance. Among them, Logistic mapping is a classical chaotic mapping, which is commonly used to study chaotic phenomena. It also shows the transformation from simple rules to complex behaviors by iteration, as shown in Equation (8).

$$l_{t+1} = \vartheta \cdot l_{t} \cdot (1 - l_{t}) \quad (8)$$

In Equation (8), l_{t} represents the population proportion of t at the point in time. ϑ represents the growth parameter, which determines the dynamic behavior of the system. The resulting sequence $\{L_{t}\}$ is the chaotic variable, which is ergodic. That is, $\{L_{t}\}$ is traversable over the runnable range when the control parameters are constant. Moreover, it is very sensitive to the parameters and initial values, and small changes can cause completely different dynamic behavior. Therefore, using it to improve the traditional PSA can enrich the population diversity, corresponding to the optimized update calculation of the position variable, as shown in Equation (9).

$$S\left(t+1\right) = S_{\min} + l_{a+1} \cdot \left(S_{\max} - S_{\min}\right) \quad (9)$$

In Equation (9), $S_{\max}$ and $S_{\min}$ are the maximum and minimum values of the position variable in the search space, respectively. This makes it possible to explore the search space more thoroughly, which enhances the algorithm's population diversity and search efficiency. The synthesis of the above content can be obtained IPTSA-FC algorithm flow schematic, as shown in Figure 4.

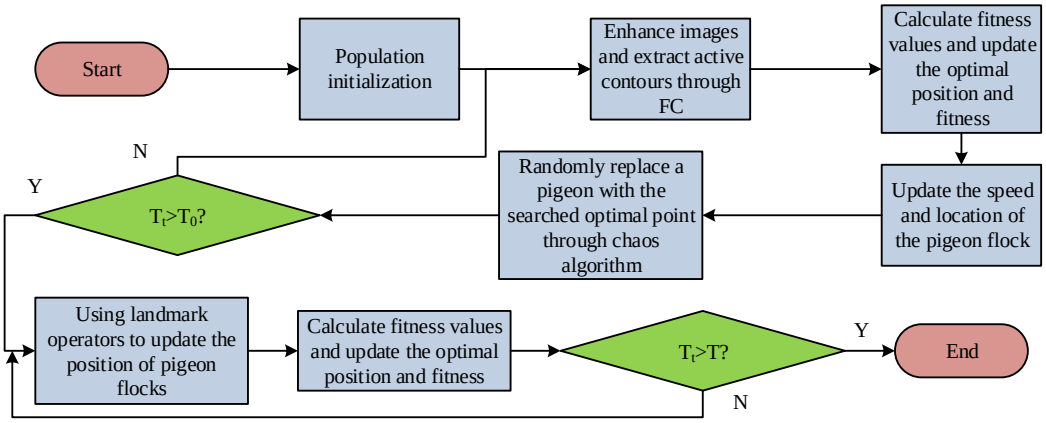

Figure 4: IPTSA-FC algorithm flowchart

In Figure 4, the parameters such as velocity, position, number of iterations, and initial position are first initialized. Second, it is processed by FC-based active contour modeling with FC filtering. Then, the respective velocities and positions are updated and the resulting optimal point is substituted for any pigeon through Logistic mapping. It is determined whether the current iteration number T_t is less than or equal to the iteration number T_0 of the magnetic field model. If it is, the above steps are looped, and vice versa the position can be updated by the landmark operator. Meanwhile, the fitness values of different points are calculated and the optimal fitness value and position are determined. Finally, it is determined whether T_t is greater than the highest iteration T. If it is, the procedure can be stopped, and if it isn't, the landmark operator's position update step for looping must be repeated.

3.3 Automated IS algorithm based on IPTSA-FC

After constructing the IPTSA-FC algorithm, the study presents the 2D-OA for optimization, resulting in the AIS-IPTSA-FC algorithm. This new algorithm enhances the anti-noise performance of the IS approach. The gray value (GV) distribution of pixels and the average GV score of nearby pixels are two of them that the 2D-OA can take into account at the same time. The algorithm's ability to reduce noise can be significantly enhanced by the ideal segmentation threshold that is thus determined. Additionally, the accompanying two-dimensional histogram can enhance the image's segmentation effect by fully utilizing the grayscale and spatial information of the image [21]. Among them, there is the precise procedure: first, the thresholds r_1 and r_2 are set. Meanwhile, the image is divided into three parts, H_1, H_2, and H_3, which need to comply with Equation (10).

$$\begin{cases} H_1 = \{1, 2, \ldots, tr_1\} \\ H_2 = \{r_1 + 1, r_1 + 2, \ldots, r_2\} \\ H_3 = \{r_2 + 1, r_2 + 2, \ldots, R\} \end{cases} \quad (10)$$

In Equation (10), R denotes the maximum GV in the image. The corresponding probability distribution for each part is calculated in Equation (11).

$$\begin{cases} P(H_1) = \sum_{i=1}^{r_1} p_i \\ P(H_2) = \sum_{i=r_1+1}^{r_2} p_i \\ P(H_3) = \sum_{i=r_2}^{R} p_i \end{cases} \quad (11)$$

In Equation (11), p_i represents the probability of a pixel with a GV of i. The relationship of optimal bithreshold i can be obtained from the mean GV of each component and the interclass variance, as shown in Equation (12).

$$\sigma^2\left(r_1^{best}, r_2^{best}\right) = \max_{1 \le r_1 \le r_2 \le R} \left\{\sigma^2\left(r_1, r_2\right)\right\} \quad (12)$$

In Equation (12), σ^2 is the inter-class variance. Finally, the image is divided into $i+1$, then corresponding to I optimal thresholds, from which the calculation of the optimal threshold is obtained, as shown in Equation (13).

$$\sigma^2\left(r_1^{best}, r_2^{best}, \ldots, r_i^{best}\right) = \max_{1 \le r_1 \le r_2 \le \ldots \le r_i \le T} \left\{\sigma^2\left(r_1, r_2, \ldots, r_i\right)\right\} \quad (13)$$

The synthesis of the above leads to the schematic flow of the AIS-IPTSA-FC algorithm, as shown in Figure 5.

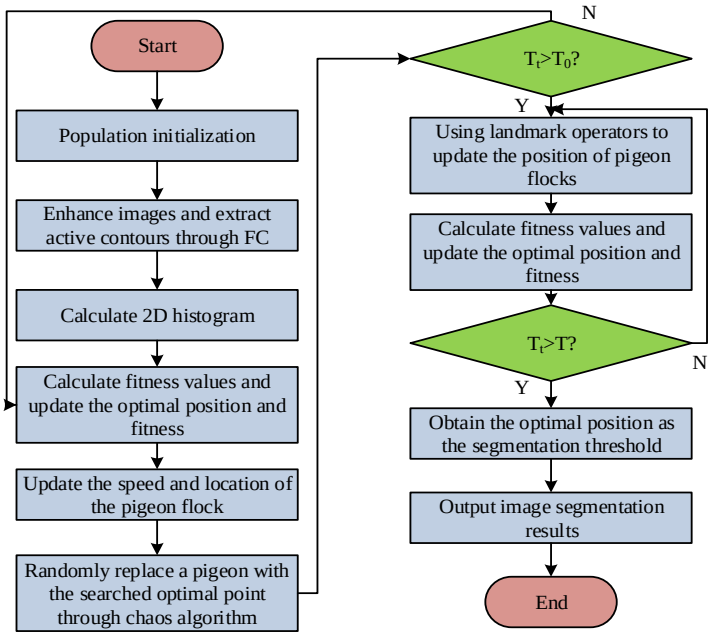

Figure 5: AIS-IPTSA-FC algorithm flowchart

In Figure 5, first, 2D-OA generates the objective function and improves the image quality. Second, entering the operation step of IPTSA-FC algorithm, after satisfying the condition of T, the optimal position can be output as the threshold value for IS, thus obtaining the final IS result. The pseudocode of the key parts of the IPTSA-FC algorithm is shown in Table 2.

Table 2: Pseudo code for key parts of IPTSA-FC algorithm

Input: Image I, Max iterations T_{max}, Swarm size N.
Output: Segmented image I_{seg}.
1: Initialize swarm positions X and velocities V.
2:Pre-process I with FC filtering.
3: for $t=1$ to T_{max} do
4:if $t \le T_{max} \times 0.75$ then
5: Update V and X using magnetic operator.
6: Apply chaotic perturbation to half the swarm.
7: else
8: Sort swarm by fitness L.
9: Update worse half positions using landmark operator.
10: end if
11: Evaluate fitness L for all pigeons.
12: Update global best solution X_{best}.
13:end for
14: Obtain optimal threshold from X_{best}.
15: return I_{seg} refined by 2D-Otsu.

4 Result analysis of automated IS algorithm based on IPTSA-FC

The study initially examines the performance of the basic method, the IPTSA-FC algorithm, in order to test the segmentation effect of the research method. Then, the results of the AIS-IPTSA-FC algorithm are validated.

4.1 Analysis of results based on IPTSA-FC algorithm

The experimental setup is as follows: 32GB RAM, Windows 10 operating system, and software Marlab R2024b are used to verify the IPTSA-FC algorithm's performance. The CPU is Intel Xeon Gold 6248R, and the GPU is NVIDIA RTX A6000. The dataset is tested using publicly available datasets, including the SiftFlow dataset

and the Open Access Series of Imaging Study-3 dataset (OASIS -3). The former contains 2,688 images of various outdoor scenes, such as streets, mountains, beaches, and buildings. These images are annotated with 33 semantic categories. The latter contains 2,842 high-quality multimodal neuroimaging datasets focusing on the study of normal aging and Alzheimer's disease. It supports multiple parameter customization with grayscale non-uniformity levels of 0%, 20%, and 40% and noise levels of 0%, 1%, 3%, 5%, 7%, and 9%. The specific preprocessing steps are as follows: all images are uniformly scaled to 256×256 pixels, and the pixel intensity is normalized to the range of [0,1] to ensure stable training. To improve the generalization ability of the model and prevent overfitting, an enhanced training process is adopted, including random horizontal flipping, rotation within ±10 degrees, and slight brightness and contrast adjustments. Meanwhile, to avoid overfitting the research methodology, the study conducts a fivefold cross-validation experiment based on the SiftFlow dataset, which provides a more stable and reliable performance evaluation. Specifically, the 2,688 images in the SiftFlow dataset are divided into five mutually exclusive subsets, each containing 537-538 images, for processing. During each validation run, use four folds for training and the remaining one fold for testing. To ensure that each image is used once as a test set, this process is repeated five times, providing a more stable and reliable performance evaluation. The experimental parameters are set as follows, the initial learning rate, the quantity of iterations and the batch size are set to $3*10^{-4}$, 300, and 32 respectively. The input image size is set to 256*256*3. The optimizer selects the Adam with Weight Decay algorithm, which decouples weight decay from gradient updates. In addition, the pigeon swarm size is set to 50, the magnetic operator stage ratio is 0.75, the velocity decay constant is 0.2, the Logistic chaotic map growth parameter is 4, and the FC order is 0.8. To verify the performance of the research methods more scientifically, the study introduces the existing mainstream methods for comparative experiments, i.e., arithmetic optimization improved based on chaotic mapping embedding algorithm (AOI-CME), multi threshold IS based on falcon optimization algorithm (MTIS-FOA), joint chaos initialization technology, chimpanzee optimization algorithm, and cauchy mutation (CIT-COA-CM). To evaluate the effectiveness of different methods for IS, the study selects the currently used metrics for analysis, namely segmentation accuracy (SA), mean squared error (MSE), structural similarity index measure (SSIM), peak signal-to-noise ratio (PSNR). Six images are randomly selected in the SiftFlow dataset, denoted as Image A-Image F. The SA results are shown in Table 3.

Table 3: The segmentation accuracy results of different algorithms in the SiftFlow dataset

IS algorithm	IPTSA-FC	CIT-COA-CM	MTIS-FOA	AOI-CME
Image A	0.9925±0.0018	0.9745±0.0041	0.9647±0.0053	0.9119±0.0082
Image B	0.9902±0.0021	0.9537±0.0055	0.9412±0.0061	0.9010±0.0077
Image C	0.9913±0.0019	0.9420±0.0062	0.9335±0.0059	0.8919±0.0091
Image D	0.9972±0.0009	0.9666±0.0049	0.9472±0.0050	0.9191±0.0084
Image E	0.9893±0.0023	0.9479±0.0051	0.9384±0.0063	0.9003±0.0088
Image F	0.9887±0.0025	0.9755±0.0039	0.9667±0.0047	0.9314±0.0075
Average execution time per image (seconds)	12.5±1.8	9.8±1.5	8.4±1.2	7.1±0.9

In Table 3, the SA value of the IPTSA-FC algorithm fluctuates up and down at 0.99, with a maximum of 0.9972, which indicates that its segmentation performance is excellent and outperforms the other mainstream algorithms. Although the other mainstream algorithms perform relatively poorly, the SA values are all above 0.89. Compared with the baseline method, the proposed IPTSA-FC algorithm has increased computation time, which is a trade-off between significantly higher accuracy and robustness. This cost increase is primarily due to the more complex optimization process of the IPTSA and FC operations. In terms of memory usage, the peak memory consumption of IPTSA-FC when processing a 256×256×3 image is about 1.8 GB, which is about 20-30% higher than the baseline method (averaging about 1.3-1.5 GB). This is mainly due to the need to store score gradient maps and larger population sizes in IPTSA. This is due to the introduction of FC in the IPTSA-FC algorithm, which can well balance the relationship between denoising and detail preservation. This enables the research method to accurately recognize different regions when processing complex images. Furthermore, U-Net itself acts as a powerful IS network that can handle complex image structures. Table 4 displays the IPTSA-FC algorithm's robustness results on the OASIS-4 dataset.

Table 4: Robustness results of IPTSA-FC algorithm in OASIS-3 dataset

Project	Noise level/%	Uneven gray level/%		
		0	20	40
Normal image	0	0.9604	0.9639	0.9640
	1	0.9381	0.9446	0.9437
	3	0.9406	0.9502	0.9482
	5	0.9057	0.9274	0.9373
	7	0.9111	0.9040	0.9150
	9	0.8923	0.8873	0.8924
Alzheimer's disease imaging	0	0.9472	0.9617	0.9537
	1	0.8452	0.8533	0.8511
	3	0.9456	0.9396	0.9296
	5	0.9252	0.9315	0.9296
	7	0.8989	0.9084	0.9090
	9	0.8936	0.8901	0.8912

In Table 4, the IPTSA-FC algorithm shows excellent results for each noise level and gray level of inhomogeneity in different projects. Although the SA value decreases to some extent, the robustness of the method is still strong in general. Moreover, the noise level has a greater effect on the SA index. At 7% and 9% noise levels, the SA metrics decrease most significantly. The effect produced by the gray scale inhomogeneity, on the other hand, is relatively small. Only when it reaches 40%, the SA value will show some fluctuation. The impact of noise level on SA index is greater than that of grayscale non-uniformity. The algorithm maintains high robustness at low to moderate noise levels (e.g. 1-5%). However, it shows a significant decrease in SA metrics at high noise levels (7% and 9%). This suggests that, while FC components help reduce noise, high-frequency noise that exceeds a certain threshold can severely damage image features and hinder accurate segmentation. On the contrary, the impact of grayscale non-uniformity is relatively small. Even at a high level of non-uniformity of 40%, the SA value only shows slight fluctuations. This proves the efficiency of the 2D Otsu algorithm. By combining the average grayscale value of the region, it has a natural ability to resist spatial intensity variations at low frequencies. Comprehensive ablation experiments are conducted on the SiftFlow dataset to rigorously evaluate the independent contributions of each key component in the proposed algorithm. Moreover, the following model is designed, Model A: replacing chaotic mapping initialization with standard random initialization. Model B: removing the two-dimensional Otsu threshold optimization step. Model C: replacing IPTSA with the standard PSA. The results are shown in Table 5.

Table 5: Results of ablation experiment

Model	SA	SSIM	MSE
Complete model	0.9915±0.0029	0.4158±0.0451	5783.42±1180.55
Model A	0.9851± 0.0035	0.3886±0.0483	6520.91± 1254.78
Model B	0.9889±0.0031	0.4012±0.0460	6050.33±1201.64
Model C	0.9743±0.0042	0.3617± 0.0521	7350.19 ±1329.95

Table 5 shows that the decrease in SA and SSIM, as well as the increase in MSE, of Model A indicate that chaos initialization is crucial for enriching population diversity. It prevents premature convergence, allowing the algorithm to explore the solution space more effectively and achieve more accurate, stable segmentation results. Model B's performance is inferior to that of the complete model, which becomes particularly evident at higher MSE. This confirms that two-dimensional Otsu plays a key role in enhancing noise resistance and determining more robust segmentation thresholds by utilizing spatial information. This optimizes the final segmentation output and reduces errors. Model C leads to the most significant performance decline across all metrics. This strongly validates the effectiveness of the improvements introduced in IPTSA. These improvements are crucial for enhancing the algorithm's global search capability, convergence speed, and overall segmentation performance.

4.2 Analysis of results based on the AIS-IPTSA-FC algorithm

To deeply investigate the segmentation effect of the AIS-IPTSA-FC algorithm, the study conducts experiments with two randomly selected images from four categories of real images, namely, people (Category 1), buildings (Category 2), natural landscape (Category 3), and medicine (Category 4). The fitness value is the objective function generated by the 2D Otsu algorithm. The goal of IPTSA is to find the threshold that minimizes the fitness value (maximizes inter class variance). Moreover, the final result represents the cost function

value that is minimized. First, Figure 6 compares the segmentation outcomes of several techniques under Category 1.

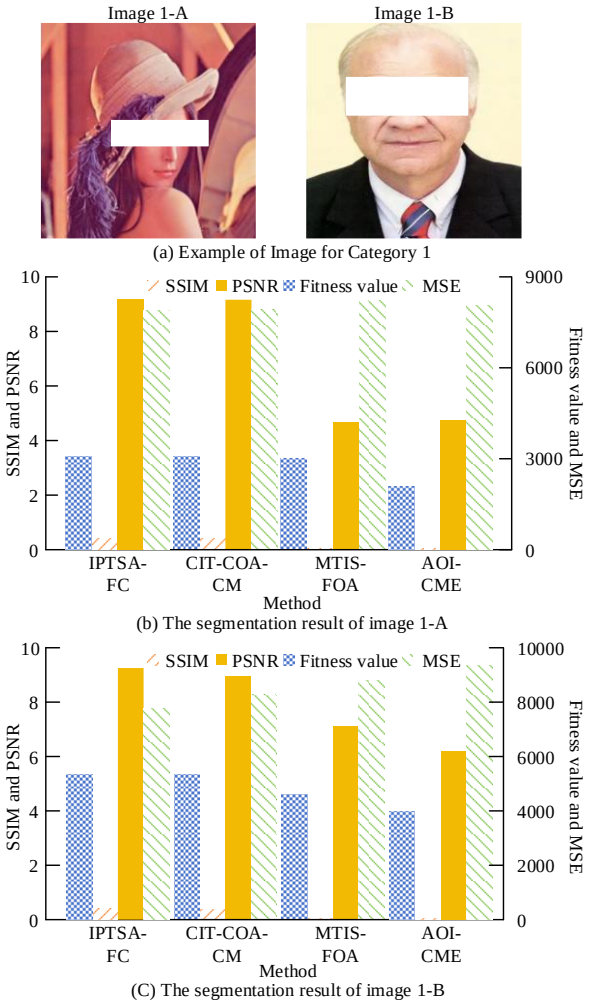

Figure 6: The segmentation results of different methods in Category 1

Figure 6(a) shows two examples of real images corresponding to Category 1, while Figure 6(b) and 6(c) display the segmentation results of different methods in image 1-A and image 1-B. Compared with other mainstream methods, the segmentation results of both AIS-IPTSA-FC algorithms show the best performance in terms of their ability to accurately recognize the contours and details of the people. Specifically, the SSIM and PSNR of the research method are the highest, with an average of 0.4169 and 9.2109. The fitness value and MSE are the lowest, with an average of 4194.0368 and 7840.5750, respectively. This is because the 2D-OA, which can more precisely identify the segmentation thresholds, optimizes the research methodology.

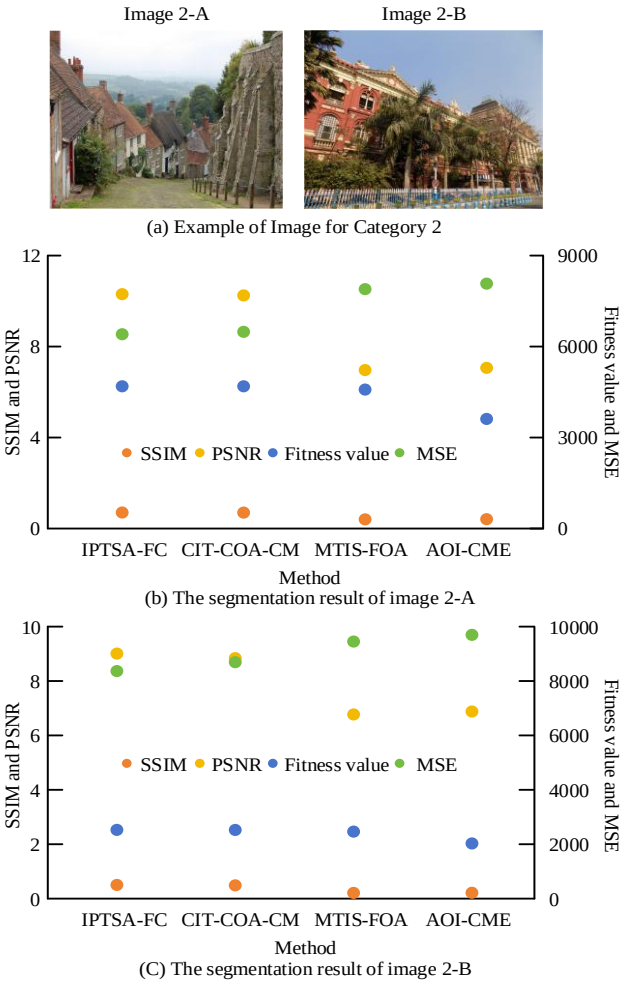

Figure 7: The segmentation results of different methods in Category 2

Figure 7(a)-Figure 7(c) show two examples of real images of Category 2 and the segmentation results of different methods on the corresponding images, respectively. Architectural images usually have regular geometric shapes with clear edges, but they are highly susceptible to the problems of uneven illumination and shadows. The research method can handle the above problems well, while other mainstream methods are prone to contour blurring or mis-segmentation when dealing with this type of complex images. The adaptation value, SSIM, PSNR, and MSE of the research method are 3430.7139, 0.3057, 9.5526, and 7278.9447, respectively. The AOI-CME algorithm, on the other hand, performs relatively poorly, with 3397.7531, 0.0086, 6.8350, and 7976.4734 for each index, respectively.

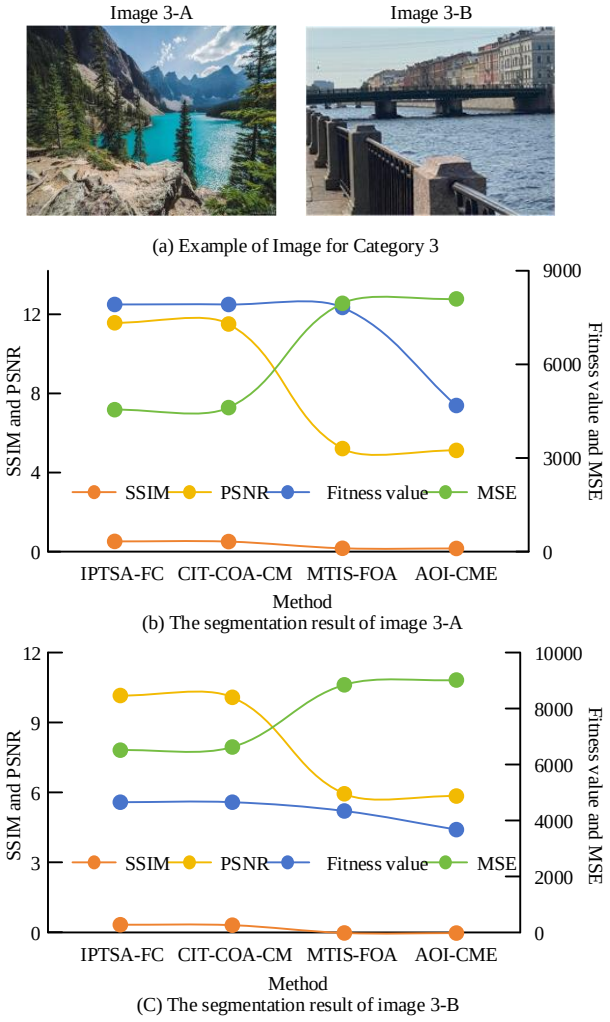

Figure 8: The segmentation results of different methods in Category 3

Figure 8(a) shows two examples of real images corresponding to Category 3. Figure 8(b) and Figure 8(c) then display the segmentation outcomes of different methods in Image 3-A and Image 3-B. The category contains rich texture variations and color variations with complex background. However, the research method is still able to accurately identify and retain the complex texture and color variations while effectively removing the background interference, whereas the mainstream methods suffer from loss of details when dealing with texture. The adaptation values, SSIM, PSNR, and MSE of the AIS-IPTSA-FC algorithm are 6316.3864, 0.3617, 10.7905, and 5501.6114, respectively.

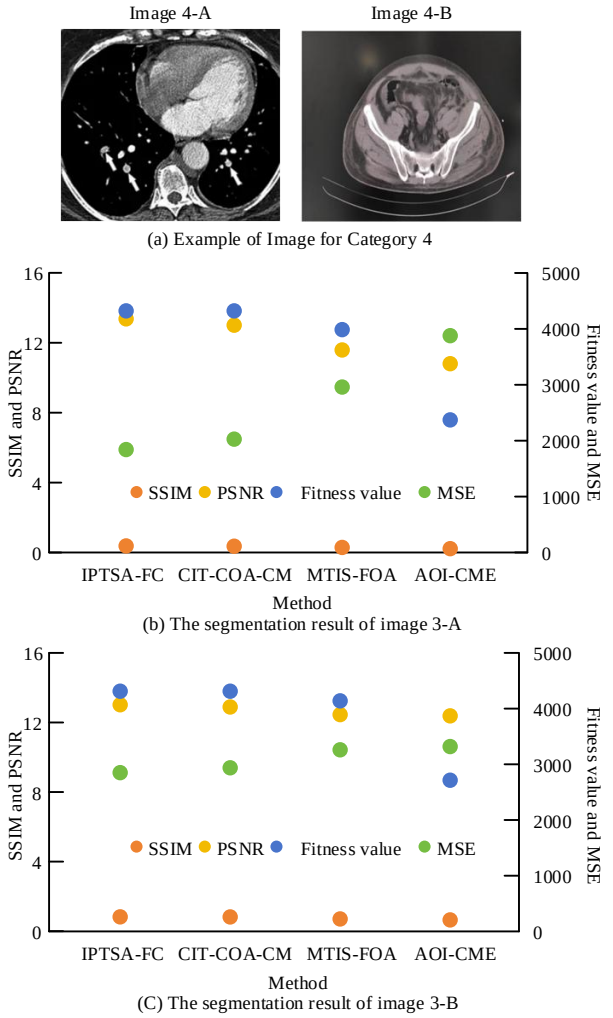

Figure 9: The segmentation results of different methods in Category 4

Figure 9(a)-Figure 9(c) show two examples of real images of Category 4 and the segmentation results of different methods on the corresponding images, respectively. Medical images require high precision segmentation. When compared with the mainstream methods, the AIS-IPTSA-FC algorithm can recognize the important structures in medical images more accurately, and it can deal with the noise and gray scale inhomogeneity well. The average values of stress value, SSIM, PSNR, and MSE of the research method are 4429.4413, 0.7057, 14.3913, and 2421.3004. The aforementioned findings may be explained by the research method's rapid determination of the best segmentation thresholds using the IPTSA-FC algorithm, which increases the algorithm's robustness. Moreover, FC has a greater advantage in denoising and detail preservation.

5 Discussion

5.1. Quantitative and qualitative comparisons

In the quantitative results, on both the SiftFlow and OASIS-3 datasets, the AIS-IPTSA-FC algorithm consistently outperformed all baseline methods (AOI-CME, MTIS-FOA, CIT-COA-CM) in all key metrics. On the SiftFlow dataset, the AIS-IPTSA-FC algorithm achieved a segmentation accuracy of around 0.99, which was significantly higher than the range of 0.89-0.98 achieved by other algorithms. Qualitatively, visual results demonstrated that AIS-IPTSA-FC could achieve clearer boundaries and better preservation of detailed structures. For instance, in architectural images, it effectively overcame the interference of uneven lighting and shadows, while other methods resulted in blurred contours or mis-segmentations. Complex natural landscape images were precisely captured with intricate textures, whereas comparison methods suffered from varying degrees of detail loss.

5.2. Analysis of performance differences

The performance advantage of AIS-IPTSA-FC can be attributed to the synergistic effects of its three core components:

The edge and detail retention capability brought by FC: The introduction of FC enhanced the traditional U-Net model. The non-local characteristics and memory effect of FC enabled the active contour model to better capture subtle transitions in the image, such as edges and textures. The FC filtering process enhanced high-frequency signals, such as edges and details, while preserving low- and medium-frequency information. This capability was essential for navigating complex backgrounds and blurred boundaries, resulting in higher SSIM and lower MSE.

The search diversity and convergence brought by the improved pigeon-inspired optimization algorithm: IPTSA with chaotic mapping optimization significantly enriched population diversity during initialization and iteration, preventing the algorithm from becoming trapped in local optima. This enhanced the algorithm's global search capability, enabling IPTSA-FC to more effectively find better solution spaces when optimizing segmentation parameters. Consequently, convergence speed and final segmentation accuracy improved simultaneously.

The threshold optimality and noise resistance brought by two-dimensional Otsu: The 2D-OA determined the optimal segmentation threshold by considering both the gray-level information of pixels and the average gray level of their neighborhoods. This method fully utilized the spatial correlation of the image, making the algorithm highly robust to noise. This was one of the key reasons why AIS-IPTSA-FC could maintain high segmentation quality under different noise levels and gray-level inhomogeneities in the OASIS-3 dataset.

5.3. Cross-domain generalization ability

Experimental results on four types of real images (people, buildings, natural landscapes, and medical) indicated that the performance improvement of AIS-IPTSA-FC had good generalization across different domains. In natural images, it excelled in handling complex textures, color variations, and lighting challenges. In medical images, where precision was of utmost importance, it could accurately segment key anatomical structures while effectively suppressing noise and gray-level inhomogeneities. This cross-domain robustness stemmed from the algorithm's design: FC provided excellent feature representation capabilities, IPTSA ensured effective parameter optimization, and two-dimensional Otsu enhances noise resistance. These components work together to enable the algorithm to adapt to the different feature requirements of various types of images.

6 Conclusion

In computer vision and image processing applications, IS is crucial for complex images with irregular shapes, textures, and overlapping boundaries, and it is beneficial to analyze and understand complex images. Therefore, the study introduced FC and CA to optimize the traditional PSA and designed IPTSA-FC algorithm. Moreover, based on this, 2D-OA was introduced to improve it, and AIS-IPTSA-FC algorithm was obtained. The experimental results revealed that the SA value of the IPTSA-FC algorithm floated up and down at 0.99, with a maximum of 0.9972. The other mainstream algorithms performed relatively poorly, but the SA value was still higher than 0.89. In the robustness experiments, the SA index of the IPTSA-FC algorithm showed a small decrease, but the overall robustness was still strong. In the segmentation results of different categories of real images, the research methods were able to show excellent results. MSE was lowest at 2421.3004. SSIM and PNSR were highest at 0.7057 and 14.3913 respectively. AOI-CME algorithm segmentation was the worst with fitness value, SSIM, PSNR, and MSE of 3397.7531, 0.0086, 6.8350, and 7976.4734, respectively. In summary, the research method can achieve accurate IS and further improve the understanding of image content, which can promote the development of intelligent and automated vision systems. However, this algorithm has certain limitations. It brings higher computational costs compared to baseline methods, making it challenging to extend to high-resolution images or real-time video processing. In addition, this study only involves experimental verification of static images and does not include segmentation of dynamic images or video sequences. In future research, optimization strategies can be implemented. These strategies include IS, more efficient memory management, and potential model distillation. These strategies aim to reduce computational resource usage. At the same time, they try to maintain segmentation accuracy.

References

[1] Houser T E, Mcmillan A, Dong B. Bridging the gap between criminology and computer vision: A multidisciplinary approach to curb gun violence. Security Journal, 2024, 37(4):1409-1429. https://doi.org/10.1057/s41284-024-00423-7

[2] Satric A, Tomasevic I, Djekic S M J. Evaluation of low sodium Kakavalj cheese properties using 3D scanning, scanning electron microscopy and computer vision system. International Journal of Dairy Technology, 2024, 77(2):586-593. https://doi.org/10.1111/1471-0307.13046

[3] Boulkroune A, Hamel S, Zouari F, Boukabou A, Ibeas A. Output - Feedback Controller Based Projective Lag - Synchronization of Uncertain Chaotic Systems in the Presence of Input Nonlinearities. Mathematical Problems in Engineering, 2017, 2017(1): 8045803. https://doi.org/10.1155/2017/8045803

[4] Kanaujia V K, Yadav S P, Kumar A, De Albuquerque V H C, Nascimento C D S. Evaluation of CNN-based computer vision recommended treatments for recognised guava disease. International Journal of Embedded Systems, 2023, 16(5/6):354-363. https://doi.org/10.1504/IJES.2023.141930

[5] Zouari F, Saad K B, Benrejeb M. Robust neural adaptive control for a class of uncertain nonlinear complex dynamical multivariable systems. International Review on Modelling and Simulations, 2012, 5(5): 2075-2103.

[6] Yang D H, Song X Z, Zhu Z P, Pan T, Tian L, Zhu L. An automatic workflow for the quantitative evaluation of bit wear based on computer vision. Petroleum Science, 2024, 21(6):4376-4390. https://doi.org/10.1016/j.petsci.2024.10.005

[7] Bhina M R, Liu K Y, Shen Y T, Tsai C T. Advancing building technology: Utilizing recycled brick sand, geopolymer mortars, robotics, computer vision, and sustainable low carbon emission strategies. MRS Advances, 2024, 9(20):1587-1592. https://doi.org/10.1557/s43580-024-00921-1

[8] Ozarslan M A. Approximating fractional calculus operators with general analytic kernel by Stancu variant of modified Bernstein-Kantorovich operators. Mathematical Methods in the Applied Sciences, 2024, 47(5):3809-3825. https://doi.org/10.1002/mma.9635

[9] Abd-Elmageed H, Hidan M, Abdalla M. Investigation for the k -analogue of τ -Gauss hypergeometric matrix functions and associated fractional calculus. Linear & Multilinear Algebra: An International Journal Publishing Articles, Reviews and Problems, 2024, 72(1/6):737-750. https://doi.org/10.1080/03081087.2022.2161459

[10] Yang G, G.-C. W, Fu H. Discrete fractional calculus with exponential memory: Propositions, numerical schemes and asymptotic stability. Nonlinear Analysis: Modelling and Control, 2024, 29(1):32-52. https://doi.org/10.15388/namc.2024.29.33550

[11] Simsek K, Tugrul N O O, Cam V K E. A Novel Method for Modeling and Predicting Transportation Data via Multi-Deep Assessment Methodology and Fractional Calculus. Transport and telecommunication journal, 2024, 25(2):136-149. https://doi.org/10.2478/ttj-2024-0010

[12] Li H, Zhu X, Li M, Yang Z, Wen M. Multi-threshold image segmentation research based on improved enhanced arithmetic optimization algorithm. Signal, Image and Video Processing, 2024, 18(5):4045-4058. https://doi.org/10.1007/s11760-024-03026-2

[13] Li H, Wen H, Xiao L L. SLIC-SSA: an image segmentation method based on superpixel and sparrow search algorithm. International Journal of Computational Science and Engineering, 2024, 27(2):182-194. https://doi.org/10.1504/IJCSE.2024.137288

[14] Guo H, Wang J, Liu Y. Multi-threshold image segmentation algorithm based on Aquila optimization. The visual computer, 2024, 40(4):2905-2932. https://doi.org/10.1007/s00371-023-02993-w

[15] Li S, Li Z, Cheng W, Qi C, Li L. Multi-Level Image Segmentation Combining Chaotic Initialized Chimp Optimization Algorithm and Cauchy Mutation. Computers, Materials & Continua, 2024, 80(8):2049-2063. https://doi.org/10.32604/cmc.2024.051928

[16] MAO Y, Fan S, Cao W, Li S. Bulk Rate Statistical Method Based on an Optimised Algorithm for Blasted Ore Image Segmentation. Journal of Northeastern University(Natural Science), 2023, 44(5):705-711.

[17] Rigatos G, Abbaszadeh M, Sari B, Siano P, Cuccurullo G, Zouari F. Nonlinear optimal control for a gas compressor driven by an induction motor. Results in Control and Optimization, 2023, 11: 100226. https://doi.org/10.1016/j.rico.2023.100226

[18] Boubellouta A B Z. Adaptive fuzzy control for practical fixed-time synchronization of fractional-order chaotic systems. Sage Publications[2025-11-07].

[19] Choudhuri S, Adeniye S, Sen A. Distribution Alignment Using Complement Entropy Objective and Adaptive Consensus-Based Label Refinement for Partial Domain Adaptation. Artificial Intelligence and Applications, 2023, 1(1): 43-51. https://doi.org/10.47852/bonviewAIA2202524

[20] Zouari F, Saad K B, Benrejeb M. Adaptive backstepping control for a class of uncertain single input single output nonlinear systems. IEEE, 2013: 1-6. https://doi.org/10.1109/SSD.2013.6564134

[21] Merazka L, Zouari F, Boulkroune A. High-gain observer-based adaptive fuzzy control for a class of multivariable nonlinear systems. 2017 6th International Conference on Systems and Control (ICSC).0[2025-11-07]. https://doi.org/10.1109/ICoSC.2017.7958728