A Derived Multi-population Genetic Algorithm For Adaptive Fuzzy C-Means Clustering

Weiping Ding, *Senior Member*, *IEEE,* Zhihao Feng, Javier Andreu-Perez, *Senior Member, IEEE,* and Witold Pedrycz, *Fellow, IEEE*

***Abstract*—Fuzzy C-Means (FCM) is a common clustering algorithm, and scholars usually use Multiple Population Genetic Algorithm (MPGA) to optimize the clustering centers. Still, MPGA has insufficient global search capabilities and lacks self-adaptability, so the optimized clustering centers are not optimal and eventually tend to converge prematurely. Therefore, this paper proposes an adaptive FCM clustering algorithm, referred to as DMGA-FCM, based on Derived Multi-population Genetic Algorithm (DMGA). The first proposed derivation operator for DMGA-FCM considers the problem of insufficient optimization ability among populations. This operator performs derivative operations on the initial population to improve the algorithm's optimization-seeking capability. At the same time, the adaptive probability fuzzy control operator dynamically adjusts the genetic probability to improve the algorithm adaptability, thereby enhancing the global optimization capability of the DMGA algorithm and avoiding premature convergence problems. Finally, this is integrated into the FCM algorithm. The analysis of simulation experiments and MRI brain map application results shows that DMGA-FCM outperforms other competitive methods in medical imaging segmentation and clustering.**

***Index Terms*—Fuzzy C-Means clustering, Clustering centers, Multi-population genetic algorithm, Derivative operator, Fuzzy control**

I. INTRODUCTION

THE rapid development of computer technology has an increasing influence on medicine. Health practitioners frequently use unsupervised learning technology to divide the medical data into groups and leverage medical data processing [1]. However, real medical data has ambiguities in terms of classification, and the effect of analyzing medical data based on a rigid demarcation clustering method is not effective. To address the above issues, L. A. Zadeh proposed a soft cluster analysis based on fuzzy set theory, that is, fuzzy cluster analysis [2]. The most widely used algorithm is Fuzzy C-Means Clustering (FCM). FCM is a fuzzy-set unsupervised clustering method. It uses a membership function grade to establish the association of each data point with all cluster categories. Data instances can be done by analyzing their fuzzy relationship to the centroids. However, FCM is sensitive to the initial clustering center and has poor self-adaptability.

In view of the shortcomings of the FCM algorithm, many scholars and experts have proposed solutions. Remarkably, Ines Lahmar et al. [4] proposed an integrated approach (SAFCM-CES) based on CF adaptive Fuzzy C-Means change across all the paper (SAFCM) clustering that uses kappa metric to determine good clustering and applies the Ncut spectral clustering model to the similarity matrix to obtain labelled results. This approach managed to obtain higher clustering quality, but the random initialization of the cluster centers resulted in the FCM algorithm lacking strong adaptivity. Arslan et al. [5] calculated the distance by Chebyshev distance function to alleviate the shortcomings of the Euclidean distance. The way they optimize the clustering centers is the Bayesian optimization algorithm (BOA), but the disadvantage of earlier convergence of the BOA algorithm is not well considered. Quang-Thinh Bui et al. [6] proposed a shape-based fuzzy C-Means (SFCM) algorithm, which mainly plays the points of mapper algorithm[7]. This approach showed the same clustering ability as the traditional algorithm, and in addition, highlights some physical relationships of the data. However, concerning the clustering center and robustness of the FCM algorithm, there is still room for improvement.

The work of Venkat et al. [8] proposes an enhanced FCM for clustering, involving Gravitational Search Algorithm (GSA) to relax the restrictions and obtain higher clustering accuracy. However, spatial location information, necessary in some applications such as medical image data, was not considered. Vinodha et al. [9] controlled the weight nonlinear change process by assigning the weights of the FCM algorithm to a local PI controller. In this work, an optimized fuzzy membership function was used to improve the adaptability of the FCM. Still, the cluster center initialization problem was not solved. Lei Tao et al. [10] proposed an improved FCM algorithm based on morphological reconstruction and affiliation filtering (FRFCM) , which can pre-determine the number of categories while optimizing the center. In these previous studies, all these improvements enhanced the clustering ability of FCM. However, most studies still neglected to de-minimize the effect of initial centroids and an adaptive nature for the algorithm.

This work was supported in part by the National Natural Science Foundation of China under Grant 61976120.

W. Ding, and Z. Feng is with the School of Information Science and Technology, Nantong University, Nantong 226019, China (E-mail: dwp9988@163.com; fzhze86@163.com)

Javier Andreu-Perez is with the Center for Computational Intelligence, School of Computer Science and Electronic Engineering, Colchester, University of Essex, United Kingdom (E-mail: javier.andreu@essex.ac.uk)

W. Pedrycz is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6R 2V4, Canada (E-mail: wpedrycz@ualberta.ca)

In summary, traditional evolutionary algorithms generally lack adaptivity and are prone to local optima, resulting in poor results for optimizing clustering centers. To the best of our knowledge, some scholars have made breakthroughs using evolutionary algorithms. For example, Anter *et al*. [11] combined chaos theory and fuzzy C-Means algorithm to design a hybrid crow search optimization algorithm (CFCSA) [12]. The proposed CFCSA algorithm utilizes a global optimization approach to avoid local extreme value traps. Anter et al. [13] also proposed a version of the FFCM, that uses the crow search optimization algorithm (CSA) to find the center of mass of the clusters during the clustering process to obtain more accurate results. ElSoud et al. [14] proposed a new subset feature selection method performed by the new Social Spider Optimizer algorithm (SSOA) that is able to find the optimal region of a complex search space by the interaction of individuals in the population. Shi Y et al. [15] propose a novel update equation and an improved dimension-selection strategy for bee colony optimization to get a more balanced search for superiority. Cui Z et al. [16] consider that NSGA-III has poor convergence in many cases. In this paper, a new selection-elimination operator is designed to balance convergence and diversity.Shang R et al. [17] presents a multi-objective artificial immune algorithm for fuzzy clustering based on multiple kernels (MAFC). MAFC combines multicore learning and multi-objective optimisation in a joint clustering framework that maintains the geometric information of the dataset. The multi-core approach uses kernel functions to map the data from feature space to kernel space. In addition, the introduction of multi-objective optimisation helps to optimise both inter-cluster separation and intra-cluster tightness.

Table 1 Summary of current related work

Type of problem	Reference	Technique	Objective
Research and improvement on fuzzy C-Means clustering	[4]	- kappa metric & the Ncut spectral clustering model	- obtain higher clustering quality
	[6]	- mapper algorithm	- highlights some physical relationships of the data
	[8]	- Gravitational Search Algorithm	- obtain higher clustering accuracy
	[9]	- local PI controllers	- improve the adaptability of the FCM
	[10]	- morphological reconstruction & affiliation filtering	- improvements enhanced the clustering ability of FCM
Application of various evolutionary algorithms for FCM optimization	[11]	- chaos theory & fuzzy C-Means algorithm	- avoid local extreme value traps
	[14]	- Social Spider Optimizer algorithm	- find the best area of a complex search space
	[15]	- bee colony optimization	- get a more balanced search for superiority
Combination of evolutionary algorithm and fuzzy c-mean clustering	[5]	- the Chebyshev distance function & Bayesian Optimization Algorithm	- optimize the initial clustering centers
	[13]	- the crow search optimization algorithm	- find clustering centers that produce more accurate results
	[17]	- combine multicore learning and multi-objective optimisation & the introduction of multi-objective optimisation	- avoid falling into local optima

With respect to state-of-the-art evolutionary algorithms to optimize FCM clustering centers, we aim to enhance both the accuracy of FCM and the adaptiveness of the algorithm. This paper proposes an adaptive FCM clustering algorithm based on a derived multi-population genetic algorithm (DMGA) to solve the initial clustering center sensitivity problem and poor robustness. The derivation operator is generated according to genetic evolution. The adaptive fuzzy control operator continuously adjusts the genetic probability to improve global searchability. This way, the initial search for the FCM clustering centre is more rigorous and thus enhances the clustering quality.

The **main highlights of this study** are as follows:

1. We propose a derived multigroup genetic algorithm to

initialize the clustering centers of the FCM by finding the optimal clustering center to reduce the effect of the initial clustering center on the clustering effect.

2. We use a derivative operator that mimics the species derivation approach and an adaptive probabilistic fuzzy control operator that incorporates the concept of fuzzy control, thereby the global search range of the algorithm is expanded and enhancing its ability to break through local peaks to avoid the prematureness problem.
3. The FCM adaptive clustering algorithm based on the genetic evolution of derived multigroup is applied to MRI brain map segmentation problem, denoting more accurate results.

The rest of this paper is organized as follows. The principle of fuzzy C-mean clustering and its initialized centroid-sensitive problem is described in Section II. In Section III, the principles and improvement points of MPGA are given and the related definitions are given. Then, we present the DMGA-FCM algorithm and describe the strategy in detail in Section IV. Section V evaluates the experimental results, and Section VI applies the algorithm to the segmentation of MRI brain tumors. Finally, Section VII concludes the paper and provides an outlook for future research.

II. FCM CLUSTERING ALGORITHM

Fuzzy C-Means clustering (FCM), proposed by Bezdek originally proposed. by Bezdek. [18] uses a concept of membership of data to cluster [19]. FCM has some advantages of better fuzzy tolerance and can retain more original data information. The generic idea of the FCM algorithm can be outlined as follows:

Let given data set $X=\{x_1,x_2,x_3,\ldots,x_n\}$ is composed of $\mathrm{c}(2\le \mathrm{c}\le n)$ groups (clusters) and set $V=\{v_1,v_2,\ldots,v_c\}$ be the c cluster centers of the dataset X . u_{ij} represents the membership of the i-th data point to the j-th class , taking the value $u_{ij}\in[0,1]$, and the sum of the membership of each data point to all classes should be 1 [20, 21]. Thus, the partition matrix U is formally represented as follows:

$$\sum_{i=1}^{\mathrm{c}} u_{ij} = 1, \forall j = 1,2,\cdots,c \tag{1}$$

where u_{ij} represents the degree to which data point i belongs to class j, when $u_{ij}=0$ represents that it not at all, and when $u_{ij}=1$ represents that it belongs entirely to the category. The objective function of FCM is:

$$J(U,V)=\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{m} d_{ij}^{2} \tag{2}$$

where $d_{ij}=\left\|x_j-y_i\right\|$ is the Euclidean distance between the clustering center i and the data points j $m\in[1,\infty)$ represents the fuzzy weighting parameter. using the technique of Lagrange multiplier we have the following solution

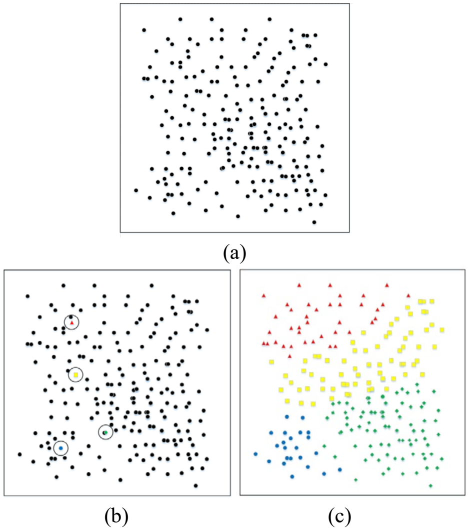
$$V_i=\sum_{j=1}^{n} u_{ij}^{m} x_j \Big/ \sum_{j=1}^{n} u_{ij}^{m}, i=1,2,\cdots,c \tag{3}$$

$$u_{ij}=1\Big/\sum_{k=1}^{c}\left[d_{ij}/d_{ik}\right]^{2/(m-1)} \tag{4}$$

The algorithm is run iteratively by Equations (3) and (4). The clustering centers V_i and partition matrices U are updated until the objective function's minimum value has obtained. The specific steps of the algorithm are as follows:

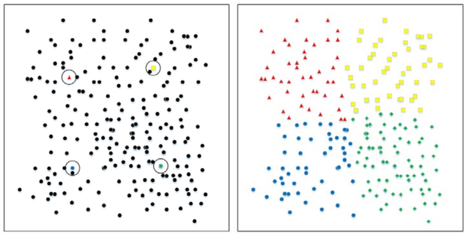
```
Algorithm1: FCM
Input:   Data set $X=\{x_1,x_2,x_3,\ldots,x_n\}$ ,number of clusters c ,fuzzy
         fuzzification coefficient $m$ ,stop iteration condition $\delta$ ,
Output:  clustering result $(U,V)$.
Step 1:  First divide the dataset $X$ into $c(2\le c\le n)$ groups, and
         initialize the set $V=\{v_1,v_2,\ldots,v_c\}$ to be c a cluster center,
Step 2:  Use Equations (3) and (4), and calculate
         $u_{ij}(i=1,2,\ldots,c, j=1,2,\ldots,n)$
Step 3:  Use Equation (2), and calculate $V_i(i=1,2,\ldots,c)$ ,
Step 4:  If $\|v_i\text{-}v_{i-1}\|\le\delta$ , then stop iteration and skip to step 5,
         otherwise, $v_{i-1}=v_i$ , skip to step 3,
Step 5:  Output the clustering result $(U,V)$.
```

When the FCM algorithm deals with uniformly distributed data and the categories are not obvious, it often obtains clustering results that are difficult to interpret due to the random effects of the centroids' initialization. To demonstrate the central sensitivity problem of the FCM algorithm, a set of uniformly distributed and category data is constructed and analyzed. Figure 1(a) shows a set of uniformly distributed data with no natural structure in this dataset. If the cluster centroids are set as in Figure 1(b), the FCM algorithm will produce the clustering results in Figure 1(c); while when the cluster centers are set as in Figure 1(d), the clustering results will be different from those in Figure 1(c), as shown in Figure 1(e). This fact indicates that the initial state of the clustering center is crucial.



(a)

(b) (c)



(d) (e)

Fig. 1. Example of FCM algorithm sensitive to initialized clustering centers: (a) Uniformly distributed data set; (b) Initial location of clustering center I; (c) Clustering results I; (d) Initial location of clustering center II; (e) Clustering results II.

In response, this paper proposes an improved derived multi-population genetic evolution algorithm that is more suitable for handling multi-feature data among the existing evolutionary algorithms [22]. The proposed approach enhances the inter-population finding ability by adding a derivative operator. Instead of empirically setting the fuzzy operators' probability parameters, our proposed method employs an adaptive probabilistic operator via fuzzy control to enable the algorithm's adaptiveness.

III. DERIVED MULTI-POPULATION GENETIC EVOLUTION ALGORITHM

This section introduces various swarm genetic algorithms and the two proposed operators. This section addresses the algorithm's shortcomings as the derivative and adaptive operators via fuzzy control.

A. MPGA

The multiple population genetic algorithm (MPGA) was proposed by Potts et al. [23]. MPGA uses multiple populations to differentiate from single population SGA[24] and increase the search capability of the algorithm. While each subpopulation evolves unidimensionally, the added migration operator allows the interpopulation to retain relatedness as well. However, MPGA may still fail to find the optimal solution in complex optimization environments[25,26].

The flowchart of MPGA's merit search is shown in Figure 2.

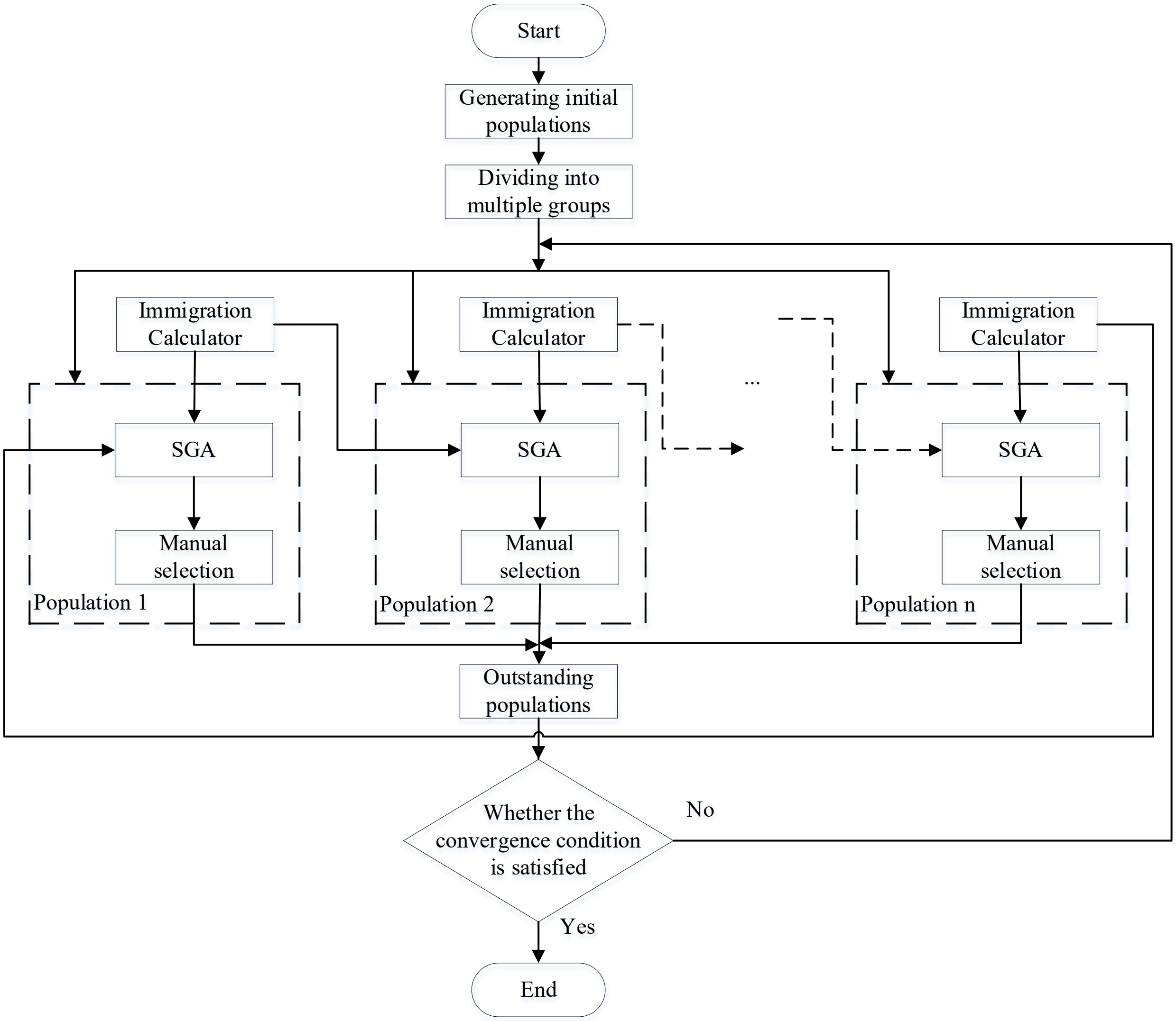


Fig. 2. Flowchart of MPGA's search for the best.

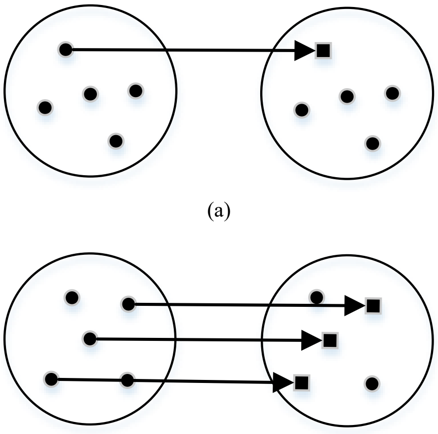
The Derivative Multi-population Genetic Algorithm (DMGA) is proposed for the first time, which first initializes the population by adding a novel derivative operator and then each sub-population individually completes the selection, crossover and mutation operations. Each of its probabilities is dynamically selected by the adaptive probability fuzzy control operator. It addresses the problems of insufficient global merit-seeking ability and the lack of adaptivity of the multi-population genetic algorithm [27]. The global search capability and self-adaptability of the multiple swarm genetic algorithm are Further improved to make the initial centroids found better.

B. Derivative operator

Traditional multiple population genetic algorithms lose the diversity of individual populations after many iterations of evolution within the population, and the global search ability is reduced, and eventually, only locally optimal results are obtained [28]. To overcome these problems, it is argued in this paper that the global search capability obtained from individuals by variational and crossover operators alone cannot completely avoid the emergence of locally optimal solutions. Therefore it is proposed an inter-population-based derivative operator. Since the operator is an evolution of populations, which is population-to-population, that is, based on inter-population, it has a significant evolutionary effect and can search the global optimal solution over a large area and break through the local peak, thus improving the global search ability of the algorithm and finding the optimal solution.

Figure 3 (a) demonstrates that the operator uses a lower derivation probability when the fitness value is large, and only a small number of individuals are found to evolve operations. A large number of individuals are reserved. Figure 3(b) demonstrates that the derivation probability appears to have higher values when the population is not adaptive, deriving most individuals and preserving a small number of outstanding individuals for derivation purposes.

The variation process of the inter-population derivative operator is shown in Figure 3.



(a)

(b)

Fig. 3. Schematic diagram of the structure of the derivative operator:
(a) Schematic diagram of the derivation of highly adaptive populations, and
(b) Schematic diagram of the derivation of low adaptation populations.

After initializing the individuals, all individuals are grouped in order of fitness. Then, each population is selected for derivation with the derivation probability p_d of individuals within the population, selected as d=1, and vice versa as d=0. After derivation, the algorithm obtains a new population from the primary population. It adjusts the derivation probability based on an adaptive probability fuzzy control operator to calculate the mean value of individual fitness for each population. The population with a high mean value has a high derivation probability, and vice versa. The definition of the derivation operator is expressed as follows

$$p_{i+1} = \begin{cases} p_i & ,d=1 \\ \max\begin{pmatrix} f(p_i^j + (1-p_d)(p_i^j - a)), \\ f(p_i^j + p_d(p_i^j - a)), \end{pmatrix} & ,d=0 \end{cases} \tag{5}$$

where p_i represents the chromosome i, p_i^j represents the coding value j of the chromosome i, p_d represents the derivation probability, a is a threshold of a changeset in the range (-128, 127), f represents the adaptation calculation, and $\max$ represents the result of the larger adaptation.

The operator is *O(n)* in its own time complexity, but in the whole algorithm because it is not added to the recursive algorithm, it does not affect the time complexity of the whole algorithm, so it is still *O(n^2)*; its impact on the space complexity, need to apply for extra double the physical space, the space complexity of all the algorithm becomes *T(n)*.

C. Adaptive Probabilistic Fuzzy Control Operator

The new population generated by the derivation operator is mainly controlled by the probability parameter of this operator to control the degree of population derivation, i.e., to control the degree of individual change within the population. The value of the derivation operator is related to the retention of good individuals and the stability of the good population, which will lose the original good individuals if it is too large and will have poor searchability if it is too small. Within the population, the new individuals generated are mainly controlled by the crossover operator and the variation operator. The magnitude of the values of the operator probabilities greatly influences the population selection. At the begining of the algorithm, the larger the value of the crossover operator is, the faster the reproduction of new individuals, which can effectively increase the diversity of populations and promote the generation of optimal solutions. However, in the later stage, the population tends to converge, and the optimal solution has been found, which will destroy the better model.

This algorithm uses an adaptive probabilistic fuzzy control operator to dynamically adjust the operator probabilities, considering the genetic algorithm's characteristics and the change of fitness values in the population [29]. We construct a fuzzy control rule base for this operator with respect to crossover probability p_c, variation probability p_m, and derivation probability p_d. The fuzzification of the input quantities $X(t)$ and $Y(t)$ are these vectors is calculated as shown in Equation (6) and (7), respectively:

$$X(t) = F_{\max} - F_{avg} / F_{\max} - F_{\min} \tag{6}$$

$$Y(t) = F_{avg} - F_{avg'} / F_{avg} \tag{7}$$

where t represents the index of genetic generations, $F_{\max}$ $F_{\min}$ represents the maximum, average, and minimum values of fitness at the t-th generation, respectively, $F_{avg'}$ represents the average value of fitness at the $t-1$-th generation, $X(t)$ and $Y(t)$ is the input of the fuzzy control operator [30].

The structure of the operator is shown in Figure 4, $X(t)$ and $Y(t)$ are the fuzzy control operator (Inputs), which are first fuzzified (Fuzzifier), then inferred (Inference) using the Rule Base, and finally, the inferred results are defuzzified (Defuzzifier) and outputted. The outputs are the three genetic probability values of p_c, p_m and p_d. The above steps process the operator to achieve the fuzzy control effect on p_c, p_m and p_d.

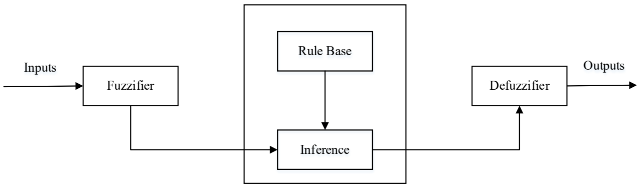


Fig. 4. Structure of adaptive probabilistic fuzzy control operator.

IV. ADAPTIVE FCM CLUSTERING ALGORITHM BASED ON DERIVED MULTI-POPULATION GENETIC ALGORITHM

The improved DMGA algorithm based on the above two operators uses binary encoding to generate a genetic individual with encoding length, as shown in Equation (8) below.

$$L = C \times N \tag{8}$$

where L is the coding length of genetic individuals, C is Number of cluster centers, and N is the feature dimension.

Let the data points $x \in (-b, b)$ be encoded, y denote the binary of x, and the calculation is shown in Equation (9) below:

$$y = (b + x) \times (2^{16} - 1) / 2b \tag{9}$$

where x is the decimal form of any single digit of the clustering center V_i, y is the 16-bit coding result, and b is the threshold of the definition domain of x.

After initializing the individuals, all individuals are sorted and grouped into primary subpopulations Pop according to their fitness. Next, the derived populations Pop' are generated by the derivation operator and merged Pop to form the genetic evolutionary initial population, then, by establishing the adaptive probability fuzzy control operator, dynamically adjusting p_c, p_m and p_d, each population is calculated by the crossover operator and the variation operator as shown in Equation (10) and (11) below.

$$\begin{cases} a_{mi} = a_{mi} \times (1 - p_c) + a_{ni} \times p_c \\ a_{ni} = a_{ni} \times (1 - p_c) + a_{mi} \times p_c \end{cases} \tag{10}$$

where a_{mi} is the i-th position of the m-th chromosome, a_{ni} is the i-th position of the n-th chromosome, and p_c is the crossover probability.

$$a_{mn} = a_{mn} + (a_{mn} - a_{\max}) \times p_m \tag{11}$$

where a_{mn} is the n-th position of the m-th chromosome, $a_{\max}$ is the upper bound of the gene, and p_m is the probability of variation.

Next, an artificial selection operator is performed among the various populations to judge the superiority and inferiority of individuals. Then a migration operator is performed to replace the worst individuals with the best individuals in the migrating population. If the algorithm meets the convergence condition, the search for superiority is stopped and the final value is decoded to get the initial of FCM clustering center, and the decoding is calculated as shown in Equation (12). Otherwise the individuals of the elite populations sorted by fitness will replace the poor individuals in each population for re-initialization and re-iterate the calculation.

$$x = -b + 2b \times y / (2^{16} - 1) \tag{12}$$

where x is the decimal form of any one-dimensional number of clustering centers V_{i}, y is the 16-bit binary encoding result, and b is the threshold value of the definition domain of x.

The DMGA search process is shown in Figure 5.

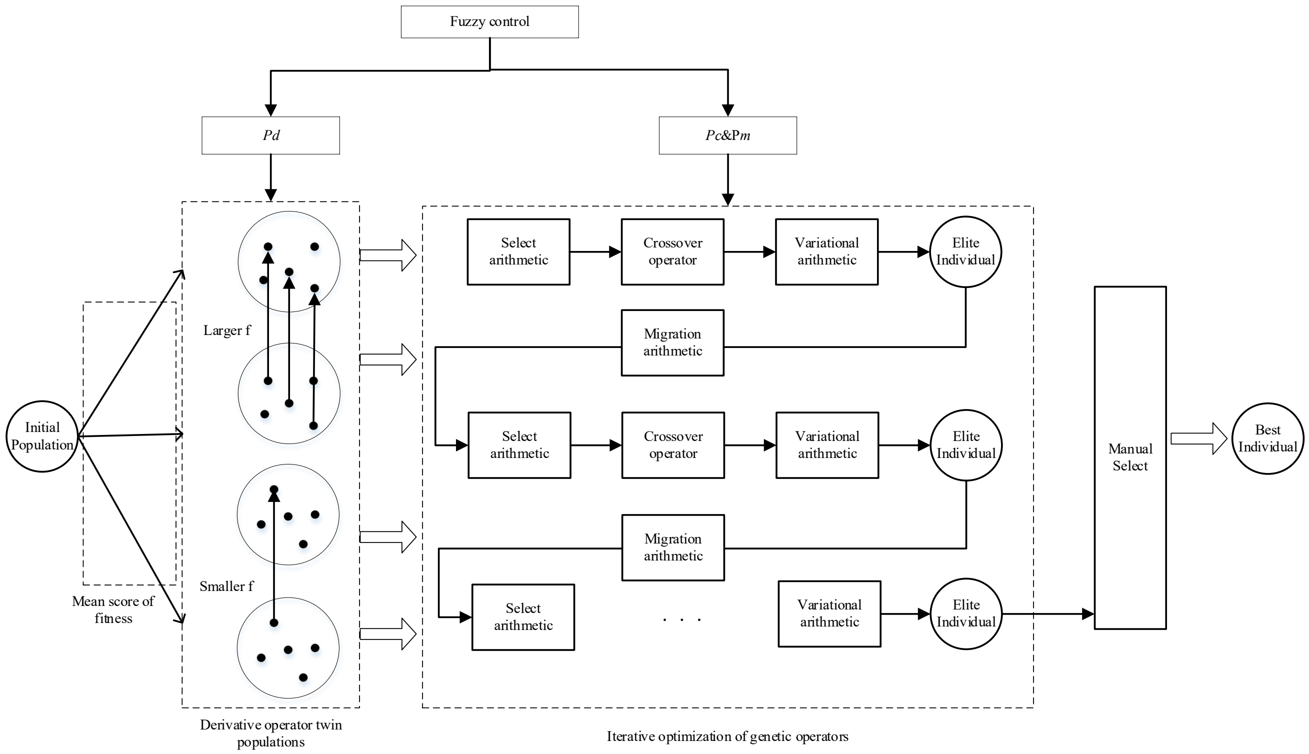


Fig. 5. Flowchart of DMGA optimization search.

The specific steps of the DMGA-FCM algorithm are described as follows:

```
Algorithm2: DMGA-FCM
Input:    Data set $X=\{x_1,x_2,x_3,\dots,x_n\}$, number of clusters $c$, fuzzy weighting parameter $m$, stop iteration condition $\delta$, genetic population size $n$, number of populations $P$.
Output:   The clustering result $(U,V)$.
Step 1:   The random number filling initial parameters, crossover probability $p_c$, variation probability $p_m$, derivation probability $p_d$, number of subpopulations $K$, and number of generations of termination iterations $T$.
Step 2:   Initialize the affiliation matrix with the random function Rand and then use Equation (9) to encode $m$ individuals with encoding length as $L$ to obtain the gene string $b=\{\beta_1,\beta_2,\cdots,\beta_i,\cdots,\beta_L\}$ and randomly generate the population of size $n$.
Step 3:   Cluster all individuals to generate $P$ sub-populations, initialize the DMGA population $Pop$, and then derive the twin population $Pop'$ by the derivation operator.
Step 4:   The adaptive probabilistic fuzzy control operator regulates the three parameters $p_c$, $p_m$ and $p_d$, and the operator selects the cross-variance operator for each subpopulation to find the optimal iteratively.
          (1) Use fitness to evaluate all individuals.
          (2) Use the proportion of the optimal retention strategy to complete the selection of individuals of the parent of this subpopulation, and perform the crossover operation on them.
          (3) Perform mutation operations on individuals within the population according to $p_m$ obtain a new generation population.
          (4) Perform migration operations to select the best individuals to form the elite population.
          (5) If the given number of iterations is not reached, turn Step4, then turn Step5.
Step 5:   If the termination condition is not satisfied, merge all subpopulations, turn Step3, then turn Step6.
Step 6:   Select the best individual from each subpopulation as the optimal solution, and the evolutionary process ends, output the optimal individual $f_a$ and decode using Equation (12) to obtain the initial clustering center $V$ of FCM.
Step 7:   The optimized $V=\{v_1,v_2,\dots,v_c\}$ for c clustering centers are obtained.
Step 8:   Calculate $u_{ij}\left(i=1,2,\dots,c,j=1,2,\dots,n\right)$ using Equation (3) and (4) to update .
Step 9:   Calculate $V_i\left(i=1,2,\dots,c\right)$ using Equation (2) to update.
Step 10:  If $\|v_i\text{-}v_{i-1}\|\le\delta$ , then stop iteration and skip to Step11, then skip to Step8.
Step 11:  Output the clustering results.
```

V. EXPERIMENTS AND COMPARATIVE ANALYSIS

A. Types of Graphics

We use three standard test functions as shown in Equation (13), (14), and (15) are chosen to perform the optimization test to verify the performance of DMGA and its effectiveness. Among them, f_1、f_2、f_3 all three functions have the same search domain, and the global minimum value is also 0. They are also commonly used by scholars to test the global search capability of evolutionary algorithms.

$$f_1(x,y)=0.5-\sin^2\sqrt{(x^2+y^2)}-0.5\big/[1+0.001(x^2+y^2)]^2,x,y\in[-10,10] \tag{13}$$

$$f_2\left(x,y\right)=1/4000\times(\mathrm{x}^2+\mathrm{y}^2)-\cos(\mathrm{x})\times\cos(\mathrm{y}\big/\sqrt{2})+1,x,y\in[-10,10] \tag{14}$$

$$f_3(x,y)=\mathrm{y}*(1+\mathrm{x})+\left|\mathrm{x}+50\mathrm{y}\times(1-2\mathrm{x})\right|+\left|\mathrm{y}+50\times(1-2\mathrm{y})\right|,x,y\in[-10,10] \tag{15}$$

The 3D image of the surface of the test function is shown in Figure 6.

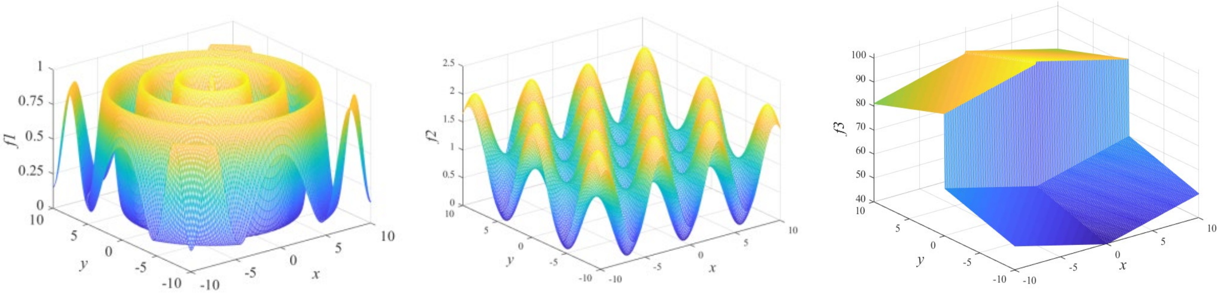


Fig. 6. 3D image of the surface of the function

The function test mainly selects the standard genetic algorithm (SGA), multiple population genetic algorithm [31] (MPGA), and ours algorithm (DMGA) to solve the minimum value of this function. It compares the evolutionary algebra and its optimal solution finding time to verify the optimization-seeking ability of each genetic algorithm. The corresponding parameter settings are shown in Table 2.

Table 2 Parameter settings of each genetic algorithm

ALGORITHM TYPE	POPULATION SIZE	NUMBER OF POPULATIONS	GENERATION GAP	DERIVED PROBABILITIES	CROSSOVER PROBABILITY	MUTATION PROBABILITY	MAXIMUM GENETIC NUMBER OF GENERATIONS
SGA	40	1	0.9	\	0.8	0.005	100
MPGA	40	10	0.9	\	0.8	0.005	100
DMGA	40	10	0.9	Fuzzy	Fuzzy	Fuzzy	100

Table 2 shows the settings of each algorithm parameter, parameter settings of SGA and MPGA are hard-coded. Parameter settings of the DMGA algorithm are set using dynamic fuzzy control. The algorithms are configured to retain the optimal individuals of the superior population for 100 consecutive generations and terminate the algorithm to get the optimal solution.

We use the number of iterations and time consumption be chosen as the comparison index of the algorithm. The number of iterations at convergence indicates that the function finds the extreme value, and the fitness does not change in the following iterations. The running time is the total time spent for each search. The test environment was a 64-bit Windows 10 Professional operating system with Intel Xeon E5-2660 v2 processor.

The test results are shown in Table 3 and Figure 7 below.

Table 3 Table of test results for each algorithm function

ALGORITHM	f_1		f_2		f_3	
	Number of iterations /c	computing overhead /s	Number of iterations /c	computing overhead /s	Number of iterations /c	computing overhead /s
SGA	62	2.282	77	**3.702**	61	1.108
MPGA	57	**1.954**	53	4.187	53	1.517
DMGA	**48**	3.282	**49**	5.358	**39**	**0.956**

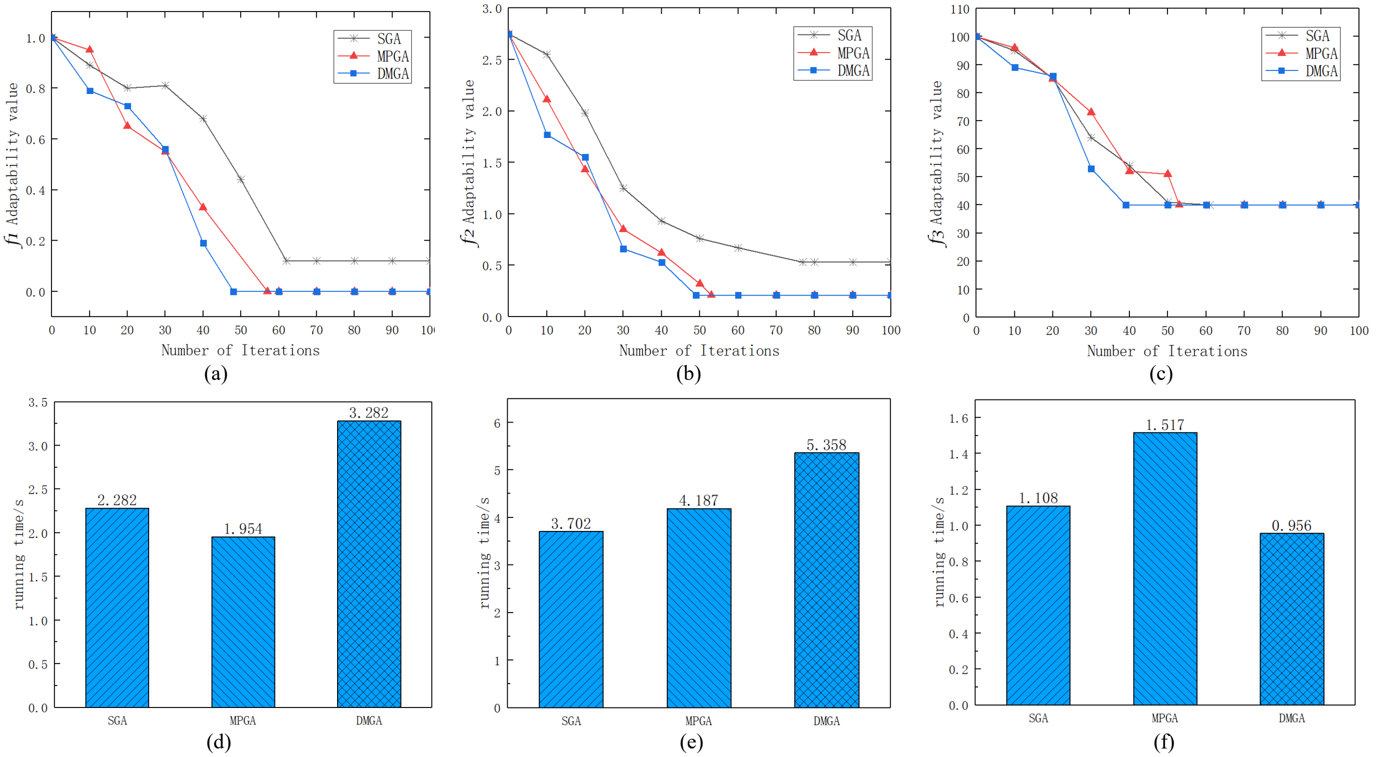


Fig. 7. Comparison of the number of iterations and running time at convergence of each algorithm to the test function: (a) Line graph of the number of iterations of the function f_1, (b) Line graph of the number of iterations of the function f_2, (c) Line graph of the number of iterations of the function f_3, (d) Histogram of the running time of the function f_1, (e) Histogram of the running time of the function f_2, and (f) Histogram of the running time of the function f_3.

The test results in Table 3 show that both MPGA and DMGA can show good optimization results for bowl-shaped functions. The average time taken by DMGA is larger than that of MPGA, mainly because the complexity of the DMGA algorithm is higher than that of MPGA. In addition, SGA's search results performed poorly in comparison. The gradient change around the optimal value of the multi-peak function f_2 is relatively flat compared to f_1. Meanwhile, it is difficult to find the optimal value because of many local peaks. So, DMGA does not see the advantage over MGPA, but like MPGA is better than SGA, and the adaptation value of SGA stops changing above the value of 0, indicating that it is not able to find the global peak. This is mainly attributed to the parallel multiple group computation seeking approach, which increases the algorithm seeking efficiency. For the cliff-like single-peaked function f_3, DMGA converges faster than MPGA and SGA, while consuming less time than the SGA and MPGA, which shows that DMGA is less affected by its high complexity in the simple single-peaked function.

Overall, in Fig. 7 can be seen that the derived multiple population genetic algorithm (DMGA) introduces a derivative operator to ensure the effectiveness of finding the optimal solution. Not this is a complex function of single and multiple peaks, and the operator's calculation is slightly time-consuming. In addition, the huge function gradient of DMGA in simple cliff-like functions narrows the problem of its high complexity. The running time of the experiments here is mainly influenced by the size of the experimental data, the amount of code and the iterations of the algorithm. Here it can be found that the DMGA also has a much larger calculation of individual adaptation values than the other algorithms, and all running times generally increase. Overall, DMGA presents excellent convergence speed and accuracy levels, which can locate global peaks more accurately in complex functions and improve efficiency in simple functions.

B. UCI Medical Dataset Clustering Test

In the experiments, seven UCI datasets were selected to test the traditional Fuzzy C-Means algorithm (FCM) [32], Particle Swarm Optimization optimized fuzzy C-Means (PSO-FCM) [33], MPGA-optimized fuzzy C-Means (MPGA-FCM) [34], and DMGA-FCM, respectively, to evaluate the algorithms more comprehensively. Datasets are from the UCI dataset (https://archive.ics.uci.edu/ml/index.php), and their characteristics are shown in Table 4.

Table 4 Relevant information of the data set

Data set name	Number of data	Number of training data	Number of testing data	Number of features	Number of categories
Cryotherapy	90	67	22	6	2
Cardiotocography	2126	1594	531	23	10
Heart Failure	299	224	74	12	2
Glass	214	161	60	9	6
Cancer-Int	699	524	175	9	2
Breast Cancer	569	426	143	30	2
Heart Disease	303	227	76	13	5

To effectively evaluate the experimental results, the experimental results of different algorithms on seven datasets are evaluated by three widely used evaluation methods. The evaluation methods are as follows:

(1) Accuracy (ACC)

$$ACC = \max_{\pi} 1/n \sum_{ij} X^{t}_{\pi(i)j} X_{ij} \tag{16}$$

where π is the arrangement of the samples, X^t and X are the clustering-accurate and all samples, respectively, if the points j belong to the cluster i , then their i-th entry is equal to 1, otherwise, it is 0.

(2) Normalized mutual information (NMI)

Assume that the computed set of clusters is the true set. Their mutual information is defined as follows:

$$MI(C,C^*) = \sum_{c_i \in C, c_j \in C^*} p(c_i,c_j) \log p(c_i,c_j) / p(c_i)p(c_j) \tag{17}$$

where $p(c_i)$ and $p(c_j)$ are arbitrarily chosen data points from the data set, which belong to the cluster c_i , is the joint probability of an arbitrary choice of data points belonging to both the cluster c_i and c_j . The standardized mutual information (NMI) is defined as follows:

$$NIM(C,C^*) = MI(C,C^*) / \max(H(C), H(C^*)) \tag{17}$$

where $H(C)$ and $H(C^*)$ are the entropy of C and C^* respectively. The larger the NMI, the better the clustering performance.

(3) Adjusted Rand coefficient (ARI)

ARI is defined as follows:

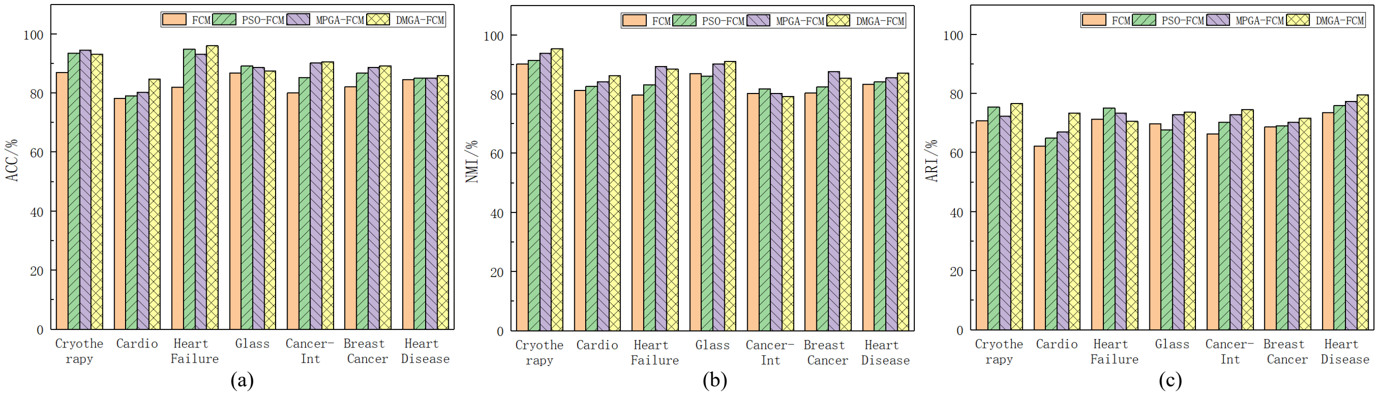
$$ARI = a_{ij} - \left[(a_{11}+a_{01})(a_{11}+a_{10})/a_{00}\right] / \left[(a_{11}+a_{01})(a_{11}+a_{10})/2 - (a_{11}+a_{01})(a_{11}+a_{10})/a_{00}\right] \tag{18}$$

where a_{11} represents assigning samples of the same type to the same set, a_{00} represents assigning samples of different types to different sets, a_{10} represents assigning samples of the same type to different sets and a_{01} represents assigning samples of different types to the same set.

The mean values of ACC, NMI, and ARI are calculated and recorded after 50 experiments are conducted for each dataset. Table 5 and Figure 8 shows the results of the clustering tests.

Table 5 ACC, NMI and ARI of each method in different data sets

Dataset\Data	ACC/%				NMI/%				ARI/%			
	FCM	PSO FCM	MPGA FCM	DMGA FCM	FCM	PSO FCM	MPGA FCM	DMGA FCM	FCM	PSO FCM	MPGA FCM	DMGA FCM
Cryotherapy	87.00	93.42	**94.50**	93.11	90.22	91.42	93.87	**95.36**	70.82	75.41	72.29	**76.64**
Cardiotocography	78.21	79.14	80.23	**84.73**	81.24	82.64	84.24	**86.21**	62.27	65.02	66.96	**73.39**
Heart Failure	82.04	94.83	93.11	**96.04**	79.68	83.24	**89.36**	88.44	71.33	75.11	**73.35**	70.57
Glass	86.76	**89.20**	88.73	87.44	87.00	86.07	90.21	**91.04**	69.71	67.69	72.89	**73.79**
Cancer-Int	80.02	85.24	90.21	**90.65**	80.27	**81.76**	80.23	79.31	66.41	70.28	72.94	**74.57**
Breast Cancer	82.13	86.72	88.65	**89.17**	80.41	82.46	**87.63**	85.37	68.73	69.15	70.29	**71.65**
Heart Disease	84.55	85.10	85.13	**85.98**	83.35	84.15	85.58	**87.18**	73.52	75.93	77.37	**79.61**



(a) (b) (c)

Fig. 8. Comparison of ACC, NMI and ARI for each method in different data sets: (a) Accuracy (ACC); (b) Normalized mutual information (NMI); (c) Adjusted Rand coefficient (ARI).

The bar plot of clustering evaluation indexes from the above test experiments shows that the DMGA-FCM algorithm has different degrees of performance improvement on each data set, indicating that the FCM algorithm optimized with DMGA can indeed improve the clustering performance.

The clustering accuracy (ACC) of each dataset in Figure 8(a), the clustering effect of the seven datasets Cardiotocography, Heart Failure, Cancer-Int, Breast Cancer, and Heart Disease has improved compared to the other three algorithms. However, Cryotherapy and Glass contain few data samples and are not suitable for the complex calculation of DMGA, resulting in poor initialization of clustering centroids. Thus, they yield inaccurate clusters. However, when combined with the normalized mutual information (NMI) index and adjusted Rand coefficient (ARI) index in Figure 8(b) and (c), it can be concluded that the clustering effect of DMGA-FCM significantly improves compared with the other three FCM algorithms.

VI. APPLICATION OF DMGA-FCM IN MRI BRAIN MAP CLUSTERING SEGMENTATION

Since the fuzzy C-means (FCM) clustering algorithm is currently the first segmentation method considered in medical image processing when large data sets are unavailable, numerous scholars have continuously studied it. As a result, a large number of image segmentation algorithms based on FCM have been proposed. However, there are still many challenges in MRI brain map segmentation, such as segmentation accuracy and the handling of image noise [35]. To further verify the clustering performance of DMGA-FCM algorithm on practical applications, the DMGA-FCM algorithm is used to do MRI clustering experiments in simulated human brain MRI images. The dataset is from the Brain Web public repository of brain MRI images [36].

MRI images are characterized by blurring, grayscale inhomogeneity, noise, etc, so, we first pre-process the MRI images, mainly including two major steps of removing parts other than brain tissue by cropping and median filtering and denoising, as shown in Figure 9, first cropping process by creating a binary image threshold for binarization, and then using the 'bwareaopen' function to remove small objects. As shown in the figure, we will get the MRI brain map binary image, followed by filling the image using the 'imfill' function to get the filled image. Finally, the 15*15 pixel size erosion by the erosion function 'imerode', and finally, the desired image area is restored. The pre-processed MRI brain image is output by median filter denoising.

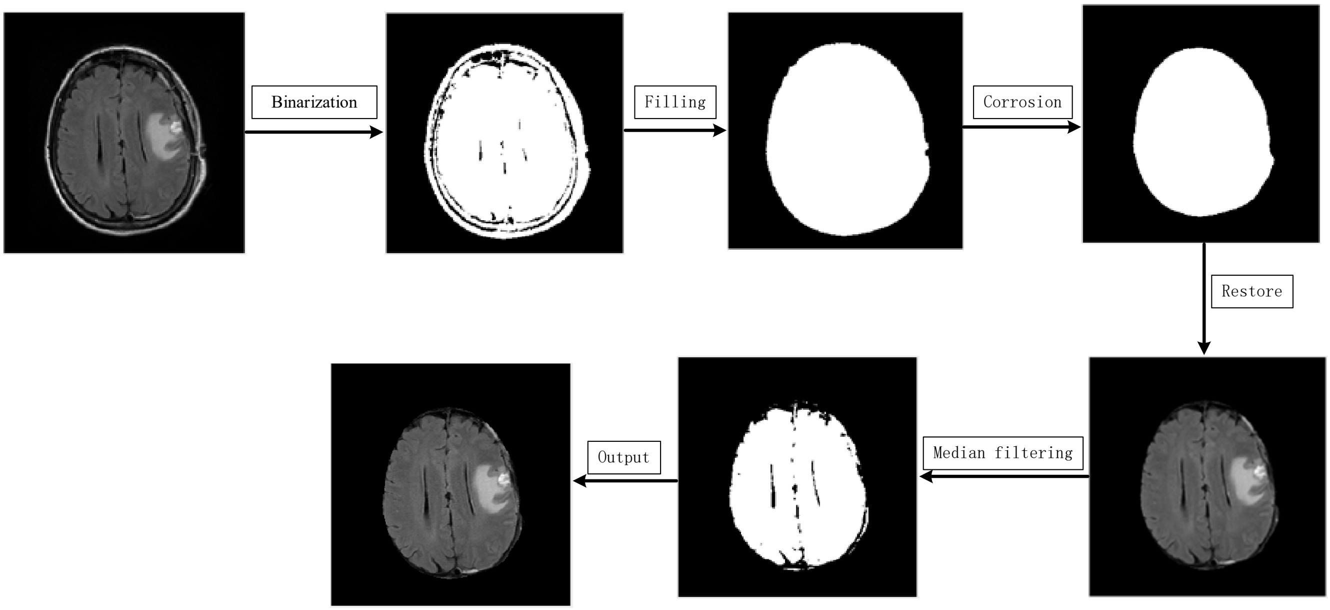
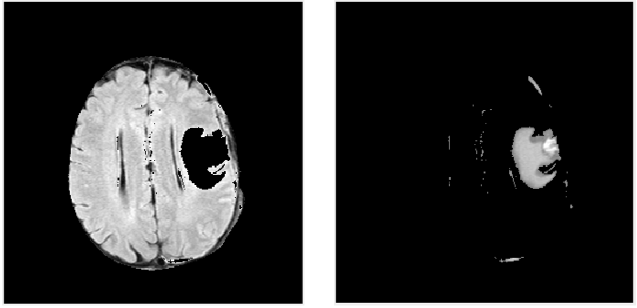
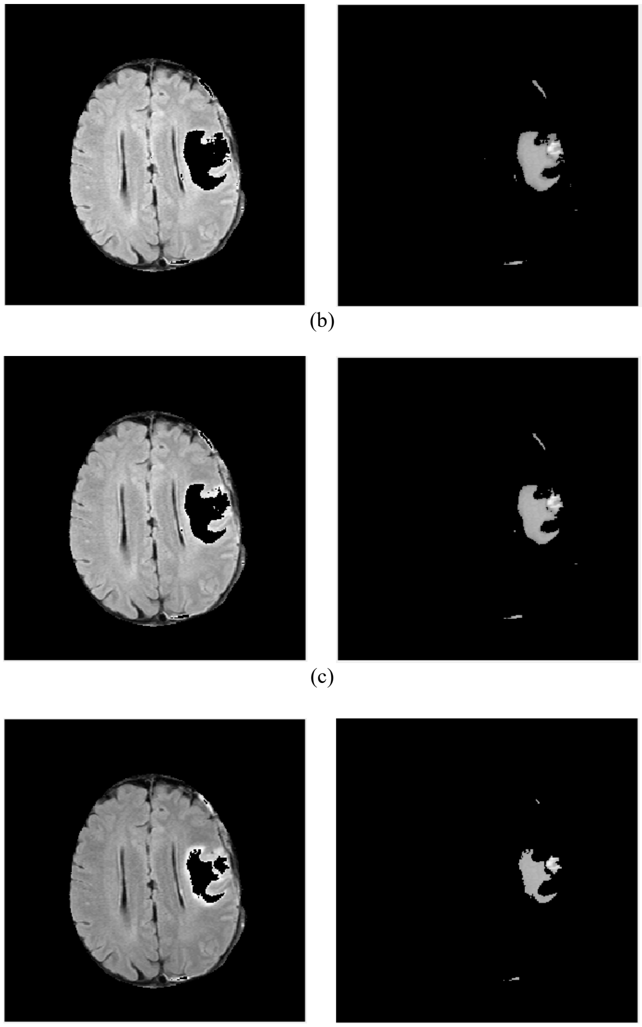


Fig. 9. MRI brain image pre-processing process

The effect of traditional image clustering algorithms on brain tumors MRI images is often less than ideal, as shown in Figure 10 of the related experimental results. It can be observed that the conventional FCM in Figure 10 (a) can identify the tumor region from the whole brain map but obviously segment the region larger than the tumor and there is a certain amount of mis-segmentation; Figure 10 (b) PSO-FCM and (c) MPGA-FCM, two improved algorithms, have improved segmentation accuracy but are not able to distinguish well at the edge of the demarcation between tumor and edema. In contrast, the DMGA-FCM algorithm in Figure 10(d) can effectively distinguish the imaginary boundary of brain tumors, and the segmentation effect is better, and the accuracy is higher than the above algorithms.



(a)



(b)

(c)

(d)

Fig. 10. The effect of clustering MRI images by algorithm: (a) FCM; (b) PSO-FCM; (c) MPGA-FCM; (d) DMGA-FCM.

Based on the MRI image clustering results of each of the above algorithms, the experimental data analysis mainly evaluates the algorithms in terms of Jaccard Similarity and Sensitivity indexes.

1. JS (Jaccard Similarity) indicates the accuracy of segmentation and is calculated as follows:

$$JS(S_1,S_2)=|S_1\cap S_2|/|S_1\cup S_2| \tag{19}$$

2. The sensitivity indicates the degree of sensitivity to the region to be segmented and is calculated as follows:

$$Sensitivity=|S_1\cap S_2|/|S_1| \tag{20}$$

In the above Equation, S1 is the segmentation result to be judged, and S2 is the exact segmentation result.

Table 6 Comparing our evaluation index with different methods

Algorithm	FCM	PSO-FCM	MPGA-FCM	DMGA-FCM
Jaccard Similarity(%)	87.22	90.22	91.37	**93.20**
Sensitivity(%)	83.42	85.44	87.37	**89.23**
Running time(s)	**12.653**	16.714	19.873	21.890

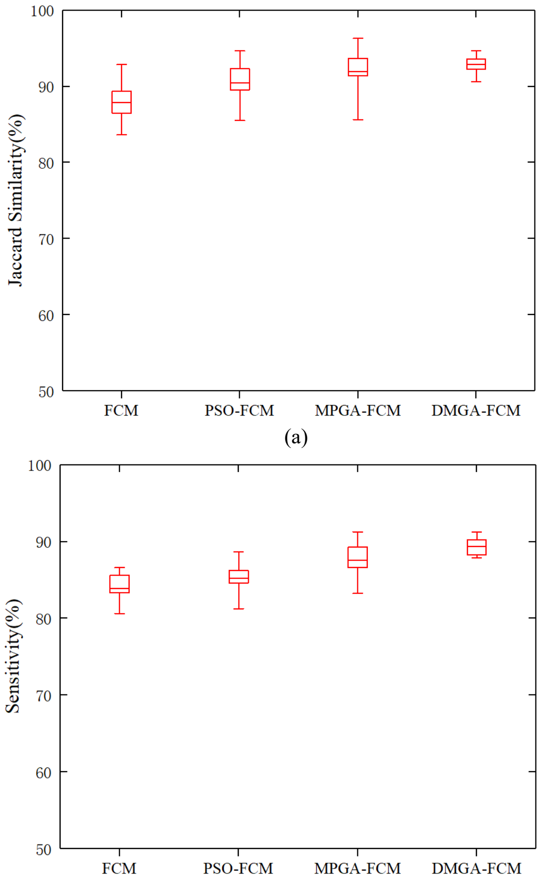


Fig. 11. Box line diagram of indicators of various algorithms in multiple experiments: (a) Box line diagram of Jaccard Similarity; (b) Box line diagram of Sensitivity.

From the segmentation result images, it can be seen that the segmentation of the segmented image regions of the DMGA-FCM algorithm is more accurate, and the algorithm has better performance and higher reliability. From the evaluation index Table 6, it can be seen that the DMGA-FCM algorithm has higher accuracy compared to the comparative algorithm segmentation, and the final average accuracy can reach: Jaccard Similarity index 93.20%, Sensitivity index 89.23%, and the clustering effect is significantly increased compared with the traditional FCM. Meanwhile, compared with two improved FCM algorithms, PSO-FCM and MPGA -FCM, this algorithm improves the Jaccard similarity by 2.98% and 1.83%, respectively. In Sensitivity index improved by 3.79% and 1.86%, DMGA-FCM algorithm has a higher computational effort and higher running time than other comparative algorithms.

In addition to this, in order to more effectively argue the effectiveness of the algorithm improvement, we conducted multiple experiments for all the experiments and drew box line plots, as shown in Figure 11, it can be seen that the difference between the maximum and minimum values of the box line plots of the DMGA-FCM algorithm is smaller, which can indicate that its stability is better than the remaining three algorithms.

VII. CONCLUSION

For the shortcomings of FCM clustering algorithm, This paper proposes an Adaptive FCM clustering algorithm based on a derived Multi-population Genetic Algorithm. First, the algorithm expands the global search space of the algorithm through the derivative operator that mimics the species-derived features; second, it enhances the algorithm adaptiveness through the adaptive probabilistic fuzzy control operator. The latter combines the fuzzy control concept to obtain better-initialized clustering centroids and reduces the influence of poor initialization. Experiments show that the clustering effect of DMGA-FCM algorithm is improved compared with the other three algorithms, especially in the four datasets of Cardiotocography, Heart Failure, Glass and Cancer-Int with accuracies of 84.73%, 96.04%, 95.44% and 90.65%, respectively. In the application of MRI brain maps, more accurate segmentation was achieved, with 2.98% and 1.83% is this difference substantially better higher accuracy than the results obtained when using PSO-FCM and MPGA -FCM.

Due to the addition of a derivative operator and adaptive probabilistic fuzzy control operator and optimization of FCM, the proposed method results in a bit high complexity. We will take the algorithm time optimisation and the optimization of the operator as improvements for our subsequent research.

REFERENCES

[1] Segatori A, Marcelloni F, Pedrycz W. On distributed fuzzy decision trees for big data[J]. IEEE Transactions on Fuzzy Systems, 2017, 26(1): 174-192.

[2] Zadeh L A. Information and control[J]. Fuzzy sets, 1965, 8(3): 338-353.

[3] Zhang W, Hu J, Zhou G, et al. Detection of apple defects based on the FCM-NPGA and a multivariate image analysis[J]. IEEE Access,2020,8: 38833-38845.

[4] Lahmar I, Zaier A, Yahia M, et al. A Self Adaptive FCM Cluster Forests Based Feature Selection[C].2019 IEEE 19th Mediterranean Microwave Symposium (MMS). IEEE, 1-4.

[5] Arslan H, Toz M. Hybrid FCM-WOA data clustering algorithm[C].2018 26th Signal Processing and Communications Applications Conference (SIU). IEEE, 2018: 1-4.

[6] Bui Q T , Vo B , Snasel V , et al. SFCM: A Fuzzy Clustering Algorithm of Extracting the Shape Information of Data[J]. IEEE Transactions on Fuzzy Systems, 2020, PP(99):1-1.

[7] Venkat R, Reddy K S. Dealing Big Data using Fuzzy C-Means (FCM) Clustering and Optimizing with Gravitational Search Algorithm (GSA)[C].2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI). IEEE, 2019: 465-467.

[8] Vinodha R. FCM Scheduled Multiple Model Controller for the Simulated Model of Spherical Tank Process[C].2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT). IEEE, 2019: 1-6.

[9] Lei T, Jia X, Zhang Y, et al. Significantly fast and robust fuzzy c-means clustering algorithm based on morphological reconstruction and membership filtering[J]. IEEE Transactions on Fuzzy Systems, 2018, 26(5): 3027-3041.

[10] Bezdek J C, Ehrlich R, Full W. FCM: The fuzzy c-means clustering algorithm[J]. Computers & geosciences, 1984, 10(2-3): 191-203.

[11] Anter A M, Ali M. Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems[J]. Soft Computing, 2020, 24(3): 1565-1584.

[12] Askarzadeh A. A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm[J]. Computers & Structures, 2016, 169: 1-12.

[13] Anter A M, Hassenian A E, Oliva D. An improved fast fuzzy c-means using crow search optimization algorithm for crop identification in agricultural[J]. Expert Systems with Applications, 2019, 118: 340-354.

[14] ElSoud M A, Anter A M. Computational intelligence optimization algorithm based on meta-heuristic social-spider: case study on CT liver tumors diagnosis[J]. Comput Intell, 2016, 7(4): 466-475.

[15] Shi Y, Pun C M, Hu H, et al. An improved artificial bee colony and its application[J]. Knowledge-Based Systems, 2016, 107: 14-31.

[16] Cui Z, Chang Y, Zhang J, et al. Improved NSGA-III with selection-and-elimination operator[J]. Swarm and Evolutionary Computation, 2019, 49: 23-33.

[17] Shang R, Zhang W, Li F, et al. Multi-objective artificial immune algorithm for fuzzy clustering based on multiple kernels[J]. Swarm and Evolutionary Computation, 2019, 50: 100485.

[18] Liu X, Zhu X, Li M, et al. Multiple kernel k-means with incomplete kernels[J]. IEEE transactions on pattern analysis and machine intelligence, 2019, 42(5): 1191-1204.

[19] Rubio E, Castillo O, Melin P. A new interval type-2 fuzzy possibilistic C-means clustering algorithm[C].2015 Annual Conference of the North American Fuzzy Information Processing Society (NAFIPS) held jointly with 2015 5th World Conference on Soft Computing (WConSC). IEEE, 2015: 1-5.

[20] Naik N, Jenkins P, Savage N, et al. Cyberthreat hunting-part 2: Tracking ransomware threat actors using fuzzy hashing and fuzzy c-means clustering[C].2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). IEEE, 2019: 1-6.

[21] Arshad A, Riaz S, Jiao L. Semi-supervised deep fuzzy C-mean clustering for imbalanced multi-class classification[J]. IEEE Access, 2019, 7: 28100-28112.

[22] Wang G G, Cai X, Cui Z, et al. High performance computing for cyber physical social systems by using evolutionary multi-objective optimization algorithm[J]. IEEE Transactions on Emerging Topics in Computing, 2017, 8(1): 20-30.

[23] Potts J C, Giddens T D, Yadav S B. The development and evaluation of an improved genetic algorithm based on migration and artificial selection[J]. IEEE transactions on systems, man, and cybernetics, 1994, 24(1): 73-86.

[24] Sarajcev P, Jakus D, Vasilj J, et al. Application of genetic algorithm in designing high-voltage open-air substation lightning protection system[J]. Journal of Electrostatics, 2018, 93: 43-51.

[25] Gladkov L A, Gladkova N V, Semushin E Y. Parallel Hybrid Genetic Algorithm for Solving Design and Optimization Problems[C]. The International Symposium on Computer Science, Digital Economy and Intelligent Systems. Springer, Cham, 2019: 249-258.

[26] Babanezhad M, Behroyan I, Nakhjiri A T, et al. Thermal prediction of turbulent forced convection of nanofluid using computational fluid dynamics coupled genetic algorithm with fuzzy interface system[J]. Scientific Reports, 2021, 11(1): 1-12.

[27] [12] Dziwiński P, Bartczuk Ł. A new hybrid particle swarm optimization and genetic algorithm method controlled by fuzzy logic[J]. IEEE Transactions on Fuzzy Systems, 2019, 28(6): 1140-1154.

[28] Han X H, Quan L, Xiong X Y, et al. A novel data clustering algorithm based on modified gravitational search algorithm[J]. Engineering Applications of Artificial Intelligence, 2017, 61: 1-7.

[29] ALKhafaji F S M, Hasan W Z W, Isa M M, et al. A Modified GA based PI controller for DC Motor Performance[C].2019 IEEE International Conference on Smart Instrumentation, Measurement and Application (ICSIMA). IEEE, 2019: 1-4.

[30] Al-Mahturi A, Santoso F, Garratt M A, et al. An intelligent control of an inverted pendulum based on an adaptive interval type-2 fuzzy inference system[C].2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). IEEE, 2019: 1-6.

[31] Tong W, Tao B, Jin X, et al. Design optimization of multipole galatea trap coils by multiple population genetic algorithm[J]. IEEE Transactions on Plasma Science, 2016, 44(6): 1018-1024.

[32] Bose A, Mali K. Type-reduced vague possibilistic fuzzy clustering for medical images[J]. Pattern Recognition, 2021, 112(2):107784.

[33] Pravin K, Sumithra M G, N Saranya. Particle Swarm Optimization (PSO) with fuzzy c means (PSO-FCM)-based segmentation and machine learning classifier for leaf diseases prediction[J]. Concurrency and Computation: Practice and Experience, 2019, e5312.

[34] Hang Z, Xiaoqun D, Jixiang D, et al. Network partitioning for reactive/voltage control based on multigroup improved FCM[J]. Electrical Measurement & Instrumentation, 2015: 07.

[35] Lei Tao, Zhang Xiao, Jia Xiaohong, Liu Shigang, Zhang Yanning. Research progress of image segmentation based on fuzzy clustering[J]. Chinese Journal of Electronics, 2019,47(8):1776-1791. (in Chinese)

[36] Kwan R K S, Evans A C, Pike G B. MRI simulation-based evaluation of image-processing and classification methods[J]. IEEE transactions on medical imaging, 1999, 18(11): 1085-1097.