Computers, Materials & Continua

Tech Science Press

Doi:10.32604/cmc.2025.065358

ARTICLE

# Relative-Density-Viewpoint-Based Weighted Kernel Fuzzy Clustering

**Yuhan Xia[1], Xu Li[1], Ye Liu[1], Wenbo Zhou[2] and Yiming Tang[1,3,*]**

[1]School of Computer and Information, Hefei University of Technology, Hefei, 230601, China
[2]School of Mechanical Engineering, Hefei University of Technology, Hefei, 230601, China
[3]Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6R 2V4, Canada

*Corresponding Author: Yiming Tang. Email: tym608@163.com

Received: 10 March 2025; Accepted: 28 April 2025; Published: 09 June 2025

**ABSTRACT:** Applying domain knowledge in fuzzy clustering algorithms continuously promotes the development of clustering technology. The combination of domain knowledge and fuzzy clustering algorithms has some problems, such as initialization sensitivity and information granule weight optimization. Therefore, we propose a weighted kernel fuzzy clustering algorithm based on a relative density view (RDVWKFC). Compared with the traditional density-based methods, RDVWKFC can capture the intrinsic structure of the data more accurately, thus improving the initial quality of the clustering. By introducing a Relative Density based Knowledge Extraction Method (RDKM) and adaptive weight optimization mechanism, we effectively solve the limitations of view initialization and information granule weight optimization. RDKM can accurately identify high-density regions and optimize the initialization process. The adaptive weight mechanism can reduce noise and outliers' interference in the initial cluster centre selection by dynamically allocating weights. Experimental results on 14 benchmark datasets show that the proposed algorithm is superior to the existing algorithms in terms of clustering accuracy, stability, and convergence speed. It shows adaptability and robustness, especially when dealing with different data distributions and noise interference. Moreover, RDVWKFC can also show significant advantages when dealing with data with complex structures and high-dimensional features. These advancements provide versatile tools for real-world applications such as bioinformatics, image segmentation, and anomaly detection.

**KEYWORDS:** Fuzzy clustering; fuzzy c-means; feature weighting; information granule

## 1 Introduction

Clustering is an unsupervised machine learning algorithm that groups similar data points in a dataset by calculating similarity while ensuring that the data points between clusters are as different as possible [1–3]. The method aims to divide the data set into subsets so that data points within the same subgroup have the greatest similarity while significant differences remain between different subsets. This process reveals the data's inherent structure and provides powerful support for applications in areas such as image processing, speech recognition, and market analysis [4]. This study addresses a critical gap in clustering research by proposing a novel method that enhances robustness and scalability for complex datasets.

Clustering algorithms can handle large data sets without prior training, so clustering algorithms are widely used in various industries. Early clustering algorithms mainly use complex clustering methods; each data point is strictly assigned to a specific cluster, and each point belongs to only one cluster. The DPC (Density Peak Clustering) algorithm proposed by Rodriguez and Laio [5] is a typical complex clustering method, which can effectively identify the peak density of data sets and thus achieve efficient clustering.

In 1969, Ruspini introduced fuzzy sets into cluster analysis and promoted the development of fuzzy clustering methods. Fuzzy clustering represents the degree of data points belonging to different clusters by membership value, which provides a more flexible way of classification. Among all kinds of fuzzy clustering algorithms, the FCM (Fuzzy C-Means) algorithm [6,7] is the most extensively utilized. By minimizing the objective function, FCM effectively calculates the cluster centre and membership degree of each data point, minimizes the intra-cluster distance, and maximizes the inter-cluster separation degree. However, FCM's sensitivity to initial clustering centres and vulnerability to noise points remain significant limitations. For this reason, Krishnapuram and Keller proposed the Probabilistic C-Means (PCM) algorithm [8], which alleviates the membership constraint of FCM and enhances the robustness of noise points. Although PCM solves the overlapping problem of cluster centres to a certain extent, it still faces the challenge of initialization sensitivity [9].

At the same time, with the wide application of kernel methods in machine learning, kernel-based clustering methods are also developing. The KFCM (Kernel Fuzzy C-Means) algorithm utilizes kernel functions to map data into high-dimensional spaces [10,11], thus improving clustering performance. For example, it improves important clustering effect evaluation indicators such as classification accuracy (ACC) [11]. However, choosing a suitable kernel function remains a major challenge. To solve this problem, Huang et al. proposed the MKFC (Multi-Kernel Fuzzy Clustering) algorithm [12], which uses the weight of each kernel function to enhance the clustering results. In addition, Yang and Benjamin proposed the FWPCM (Feature-weighted Probability C-Mean algorithm) algorithm [13], which combines probability clustering and subspace clustering and enhances the clustering effect by applying feature weighting.

In recent research on fuzzy clustering, introducing domain knowledge into clustering algorithms has become a very popular method for improving clustering quality. Hussain et al. introduce novel weighted multi-view k-means algorithms using L2 regularization [14], designed specifically for clustering multi-view data. Pedrycz and colleagues developed the Viewpoint-based Fuzzy C-Means (V-FCM) algorithm [15], which uses domain knowledge to guide the clustering process to optimize the results from a specific perspective. However, the selection of viewpoints in V-FCM relies on subjective input, which limits its automation and flexibility. The algorithm still has difficulties in dealing with noisy points. In order to solve these problems, Tang et al. proposed the DVPFCM (Density View-induced Probabilistic Fuzzy C-Means) algorithm [16], which uses density peaks as viewpoints to improve the stability of cluster centre initialization. Despite the progress, DVPFCM still faces challenges in making full use of the intrinsic structure of the data for viewpoint optimization. Subsequently, Tang et al. [17] proposed the Kernel-based Hypersphere Density Initialization (KHDI) algorithm as a certain premise and proposed the density Viewpoint-based Weighted Kernel Fuzzy Clustering (VWKFC) algorithm.

In practice, although the VWKFC algorithm shows good results, we believe that the algorithms based on KHDI and VWKFC still have the potential to be further optimized to achieve better results in various data sets. We note that the information granularity factor in the new algorithm plays an important role in membership adjustment and optimization, but there are still two problems to be solved.

The suitability of information granularity factors for adjusting fuzziness has not been properly evaluated. It is not clear whether the adjustment to ambiguity has been efficiently implemented. Therefore, we try to apply the double fuzzy fusion of information granularity factors and further optimize some steps and algorithms in the initialization process of KHDI in the VWKFC algorithm.

In order to solve these problems, this paper proposes a double fuzzy fusion strategy for information granularity factors and further optimizes the key steps in the initialization process of KHDI and the clustering algorithm under the framework of VWKFC.

Unlike previous studies, this paper proposes a knowledge extraction algorithm based on relative density. The algorithm selects the initial viewpoint by identifying the relatively high-density region in the data, provides an ideal initial clustering centre for the clustering algorithm, and improves the accuracy of the initial clustering process. We introduce the adaptive weight mechanism into the fuzzy clustering framework and build the information granularity weight model. By dynamically adjusting the importance of data points in the clustering process, the algorithm effectively reduces the influence of noise and outliers, thus improving the robustness and accuracy of clustering. In addition, we also designed a dynamic optimization algorithm, which adaptively adjusts the weights according to the changes in data distribution and further improves the flexibility and robustness of the algorithm in different feature spaces and dimensions.

The contributions of this paper are as follows:

Cluster initialization method based on relative density: We propose an innovative knowledge extraction algorithm based on relative density, which takes the relatively high-density region in the data as the clustering initialization viewpoint, and significantly improves the quality of the initial clustering process. This algorithm uses the local density feature of data to locate the potential cluster centre accurately, reduces the error caused by traditional random initialization, and enhances stability and accuracy.

Adaptive information granularity weighting mechanism: We introduce an adaptive information granularity weighting model, which dynamically adjusts the influence of data points according to their importance in the clustering process. The proposed mechanism effectively reduces the influence of noise and outliers, especially when the data is unevenly distributed or contains outliers, thus improving the accuracy and robustness of the clustering results.

Dynamic weight optimization algorithm: A dynamic optimization algorithm is designed to adjust the weight of data points based on the change in data distribution. By combining the local characteristics of the data, the algorithm has strong adaptability and flexibility, which ensures robust clustering performance in different feature Spaces and high-dimensional environments.

A new adaptive weighted clustering algorithm: By integrating the above innovations, we propose the RDVWKFC algorithm to combine the advantages of the above different algorithms and methods. An excellent linkage algorithm mechanism is formed, which provides a comprehensive and effective solution to the complex data clustering problem.

## 2 Related Work

In this section, we will review the previous works and concisely introduce some algorithms concerning the new algorithm proposed in this paper.

Referring to [18], a maximum entropy regularized weighted fuzzy C-means algorithm (EWFCM) is proposed, which breaks through the limitation that the traditional Euclidean distance as a measure of membership cannot identify and distinguish these attributes unrelated to clustering. The EWFCM algorithm achieves feature selection by providing weights for the features. Increasing the maximum entropy regularization term effectively affects the feature weight distribution. However, the clustering effect of this algorithm depends very much on the initialization of the cluster centre. The appearance of noise points in the data will seriously interfere with the clustering results.

Oskouei et al. [19] proposed an intuitionistic fuzzy c-means-based clustering algorithm for multi-view clustering (VCoFWMVIFCM) to artificially solve the problems of noise sensitivity, outliers' influence, and different perspectives, features, and sample importance. The algorithm combines membership degree, non-membership degree, and hesitation degree to design intuitive fuzzy C-mean (IFCM) loss terms, which greatly enhances the processing ability of the algorithm for uncertain data. At the same time, by weighting three

different levels of features, views, and samples, the low coupling between classes and high aggregation within classes are more significant, and the clustering effect is further enhanced. Moreover, Oskouei et al. also designed a neighbourhood agreement term to enhance the stability and robustness of the local structure of the cluster by punishing the membership difference of domain samples.

Hashemzadeh et al. [20] proposed a New fuzzy C-means clustering method based on feature-weight and cluster-weight learning (FWCW-FCM). The algorithm solves two problems of traditional FCM by simultaneously optimizing feature weight and cluster weight: feature importance difference and initialization sensitivity. The innovation of the algorithm is to independently adjust the feature weight in each cluster with the help of the local feature weight to improve the quality of clustering. The cluster weight was dynamically adjusted to punish the clusters with large intra-cluster distances to reduce the dependence of the algorithm on initialization. This algorithm significantly improves the clustering performance of traditional FCM by jointly optimizing the feature and cluster weights and reducing the initialization sensitivity. Dynamically adjusting the importance of features and clusters is suitable for complex real datasets. This algorithm outperforms the traditional FCM algorithm in terms of feature weighting and resistance to initialization jitter, but its performance is highly dependent on parameter tuning and data type.

Nie et al. [21] proposed a new graph Clustering model called Fast Clustering with Anchor Guidance (FCAG), which significantly solved the high computational cost and difficulty in adjusting hyperparameters. FCAG introduces anchor technology to construct a bipartite graph, which significantly reduces the size of the similarity matrix and makes it suitable for large-scale data sets. By designing a parameterless optimization objective function, the model effectively avoids the complex parameter tuning problem caused by the regularization term in traditional methods and the trivial solutions of "all samples belong to the same class" or "uniform distribution" in the clustering results.

Hu et al. [22] proposed a novel clustering algorithm called self-regulating possibilistic C-means (PCM) with high-density points (SR-PCM-HDP), which combines Density Peak Clustering (DPC) and possibilistic C-means (PCM) to determine the optimal number of clusters. The density-based knowledge extraction method (DBKE) was introduced to estimate the initial number of clusters without a predefined density radius and extract high-density points as knowledge guidance. Through the adaptive mechanism and knowledge guidance, the clustering efficiency and robustness are improved so that it can be effectively applied to complex data scenarios.

In order to better solve the sensitive problem of cluster initialization, Tang et al. [17] proposed the VWKFC algorithm. This algorithm mainly puts forward the concept of information granules, which mainly contains two aspects: for one thing, it gives a feature matrix that assigns different weights to different features of any data. For another, it provides sample weights for each data point in the dataset. Thus, in the process of clustering, the raw data is transformed into weighted information granules. In the research of initialization methods for clustering analysis, traditional density peak detection methods and initialization strategies based on Euclidean distance often face problems with sensitivity to noise and insufficient processing ability of high-dimensional data. Tang et al. proposed a Kernel-based Hypersphere Density Initialization (KHDI) method. VWKFC achieves efficient clustering for complex data through kernel density initialization, weight mechanism, and viewpoint guidance and is obviously ahead of the traditional fuzzy clustering algorithms in noise robustness, high-dimensional processing, and convergence speed. The idea of combining domain knowledge with data-driven methods provides a new optimization paradigm for fuzzy clustering.

## 3 The Proposed RDKM Method

In previous clustering algorithms, the initialization process of the cluster centres is highly sensitive, and this problem has been significantly improved and solved in the previously proposed KHDI algorithm [17].

The KHDI algorithm uses a Gaussian kernel to calculate the local density and determine whether the data points belong to the same cluster based on a fixed radius. However, choosing this fixed radius may lead to inaccurate density estimates, especially in data sets with uneven distribution or high dimensionality, thus affecting the quality of clustering results.

To improve the KHDI algorithm, in this study, we propose a relative density-based knowledge Extraction method (RDKM) and introduce an adaptive density radius calculation method.

By designing an appropriate function, the DPC algorithm dynamically adjusts the radius to ensure that the average density remains within 2% to 3% of the total number of data points. This adaptive mechanism allows the local density calculation to be optimized according to the actual distribution of the datasets rather than relying on a fixed radius parameter. This innovation significantly improves the robustness of the DPC algorithm in datasets with non-uniform density and avoids the error caused by the fixed radius.

In some cases, it is difficult to directly define an appropriate similarity measure function in the original feature space, especially when the data distribution is complex. In this case, using Euclidean distance often fails to produce desirable clustering results.

We introduce the Gaussian kernel function as an effective tool for constructing a clustering model to try to solve this problem. By normalizing each sample feature, we will constrain the eigenvalues in a standardized range, thus greatly enhancing the stability in the subsequent kernel function calculation. This preprocessing step not only improves the numerical stability of the model but also enhances the adaptability and robustness of the algorithm in high-dimensional space. The data points are mapped from the original feature space to the higher-dimensional space through the kernel function, where the clustering structure becomes clearer and clearer. This method makes good use of its advantages, overcomes the limitations of traditional clustering algorithms in processing high-dimensional data, and provides a new perspective for revealing internal data patterns.

The selection of initial clustering centres is particularly important in the process of fuzzy clustering, which directly affects the convergence of the algorithm and the quality of clustering results. The traditional random initialization method is inefficient and may cause the algorithm to converge to a suboptimal solution. In addition, the density calculation based on Euclidean distance shows limited performance when dealing with high-dimensional or noisy data. In order to solve these problems, this paper proposes an initial clustering centre selection method based on the DPC algorithm and kernel density optimization. Adaptability of the RDKM method to complex data distribution by designing a reasonable density radius and viewpoint selection mechanism.

The local density of a point is defined by the following formula:

$$\rho_i = \sum_{j=1}^{N} \chi(d_{ij} - r), \chi(x) = \begin{cases} 1, & \text{if } x < 0 \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

here, $d_{ij}$ represents the kernel distance between the representing point $x_i$ and point $x_j$, and $r$ is the density radius. The kernel distance is calculated as follows:

$$d_{ij} = \| \phi(x_i) - \phi(x_j) \|^2 = K(x_i, x_i) - 2K(x_i, x_j) + K(x_j, x_j) \tag{2}$$

The kernel function is a mapping function used to map the data from the original space to the high-dimensional feature space. Gaussian Radial Basis Function (RBF) is used as the kernel function $K(x, y)$.

$$K(x, y) = exp\left(-\frac{\| x - y \| 2}{2\sigma^2}\right) \tag{3}$$

here, $\sigma$ is the width parameter of the Gaussian kernel, which controls the sensitivity of the kernel function to distance variations. Through the introduction of the normalization formula, its key role in data preprocessing is revealed. By using the range transformation technique, the effect of diverse variables (features) in the data is productively eliminated. This process maps features of different orders of magnitude to a unified [0, 1] interval, ensuring that the entire set of features has equal importance in the cluster analysis.

$$x_{jl} = \frac{x_{jl} - \min(x_{kl})}{\max(x_{kl}) - \min(x_{kl})} \tag{4}$$

Another advantage of normalization is that the transformed data is restricted to the range [0, 1], which simplifies the choice of parameters.

Using the Radial Basis Function (RBF), the optimal result can be obtained by setting $\sigma = 1$ for normalized features.

Based on the distance matrix, $D = [d_{ij}]$, the density radius $r$ is initialized as follows:

$$r = [max(d_{ij}) + min(d_{ij})]/2 \tag{5}$$

This formula makes the density distribution more stable by smoothing the interference of noise and outliers on the local density calculation, especially for the processing of high-dimensional data.

The selection of the density radius significantly impacts local density estimation and global clustering. A large radius can cause uniform density estimates, reducing the algorithm's ability to capture local structural features, while a small radius leads to excessive sensitivity to noise, affecting the clustering accuracy and effectiveness. To address this complex issue, we present a method designed to dynamically optimize the density radius:

(1) Density sorting: All data points are sorted by their density value $\rho_i$ to form a density sequence.
(2) Average density calculation: Select all the data points and calculate their average density.

$$\overline{\rho} = \frac{1}{N}\sum_{i=1}^{N}\rho(i) \tag{6}$$

(3) Density radius optimization: We adjust the density radius r so that the average density meets the following conditions. Character $p$ is the percentage of the total data volume (e.g., $p = 2\%$ to 3%):

$$\overline{\rho} = p \times N \tag{7}$$

Therefore, we can consider the local density estimation to be consistent with the overall characteristics of the data distribution. By dynamically adjusting the density radius, the new algorithm can capture the local distribution characteristics of the data adaptively, so as to avoid the limitation of fixed parameters.

According to the results of kernel density calculation, the cluster centres and viewpoints can be further determined:

For each data point, we identified all the points with a density greater than. nd computed the minimum distance between $x_j$ and these points, and then we obtained the standard distance $\delta_i$.

$$\delta_i = min_j d_{ij}, \text{ where } \rho_j > \rho_i \tag{8}$$

For each data point, we define a co-weight parameter $\tau_j$ that integrates the density and pivot distance.

$$\tau_i = \rho_i \times \delta_i \tag{9}$$

In the process of selecting the cluster centre, the larger the parameter value $\tau_i$, the more likely the corresponding data point $x_j$ is to be the cluster centre. Specifically, if $x_j$ is not a cluster centre, its local density $\rho_j$ may be relatively high, but its pivot distance $\delta_j$ is small. On the contrary, for noisy points or outliers, their pivot distance is larger, but their local density is usually lower. Therefore, by introducing a selection mechanism that favours larger values, the potential cluster centres can be effectively distinguished while significantly resisting the interference of noise and outliers.

Based on this, we sort the data points $\tau_i$ in descending order and select potential cluster centres from this ranked list. To prevent the selected initial cluster centres from being too close to each other, a reasonable strategy is to ensure that the distance between any two selected cluster centres exceeds the following critical distance $d_c$.

$$d_c = \frac{\max\left(d_{ij}\right)}{aC}, (i, j \in 1, 2, \ldots, N) \tag{10}$$

The $\max\left(d_{ij}\right)$ represents the maximum distance between any two data points in the dataset, $C$ represents the number of cluster centres, and a is the parameter used to adjust the minimum distance between cluster centres. This constraint not only ensures the reasonable distribution of the initial clustering centres in the feature space but also improves the stability of the clustering algorithm.

Users usually directly provide the selection of "viewpoints" to guide the clustering process. However, a new viewpoint selection strategy is used in this study. Specifically, after obtaining the initial cluster centres, we select the data point $\tau_j$ with the largest computed value (that is, the first data point in descending order) as the "viewpoint $x_d$" selection. The rationale is that this data point has a high local density $\rho$ and a large pivot distance $\delta$, which makes it a true cluster centre.

With the help of the setup of viewpoint $x_d$, we observe the structure of the whole dataset starting from core locations with high density and large critical distance. This method not only avoids the subjective deviation caused by the traditional viewpoint selection but also improves the rationality and scientific validity of the viewpoint, which lays the foundation for the subsequent clustering process.

In the cluster centre, the data point $\tau_j$ that has the largest value is selected as the viewpoint $x_d$, that is,

$$x_d = \arg\max_j \tau_i \tag{11}$$

The pseudocode for this algorithm is shown below. We mainly follow the principles of modularity and simplicity when writing this part of the pseudocode. The detailed flow of the RDKM algorithm is given in Table 1.

**Table 1:** Kernel-based hypersphere density initialization method (RDKM)

**Inputs:** A collection of N pieces of data $X = \left\{x_j\right\}_{j=1}^{N}$, a cluster number $C$, and the width control parameter σ of the Gaussian kernel function.
**Outputs:** Clustering centre matrix $V = \{v_i\}_{i=1}^{C}$, and the density viewpoint $x_d$.
procedure RDKM (Data $X$, Number $C$)
V = [ ];
The original data are normalized;
The density radius $r$ is determined by the binary search method;
The local density $\rho_j$ of data points is computed by (5);

(Continued)

**Table 1 (continued)**

The minimum distance $\delta_j$ of the data points is calculated by (8);
The parameters $\tau_j$ of the data points are computed by (9);
The distance $d_c$ between centres is calculated by (10);
Arrange $\tau = \left\{\tau_j\right\}_{j=1}^{N}$ in descending order, and get the corresponding data set $X'$ according to the descending order of $\tau_j$;
Select the data point $x_1^{'}$ corresponding to $\tau_1$, and take it as the first clustering centre $v_1$, and let $V = V \cup v_1$;
Take the first selected cluster centre as the viewpoint, i.e., $x_d = v_1$;
Set $num = 1$, $index = 2$;
repeat
while $\left\|x_k^{'} - V\right\| < d_c$
$index = index + 1$;
$V = V \cup x_k^{'}$;
$num = num + 1$;
until $num = C$
return $V$, $x_d$;
end procedure

## 4 The Proposed RDVWKFC Algorithm

In this paper, a Relative-Density-Viewpoints-based Weighted Kernel Fuzzy Clustering Algorithm with Adaptive Weights (RDVWKFC) is proposed. The objective function is as follows:

$$J = \sum_{i=1,i\neq q}^{C}\sum_{j=1}^{N}\alpha_j^{'}u_{ij}^{m}\sum_{l=1}^{L}\omega_{il}\parallel\phi\left(x_{jl}\right)-\phi\left(v_{il}\right)\parallel^2+\sum_{j=1}^{N}\alpha_j^{'}u_{qj}^{m}\sum_{l=1}^{L}\omega_{ql}\parallel\phi\left(x_{jl}\right)-\phi\left(x_{dl}\right)\parallel^2+\eta\sum_{i=1}^{C}\sum_{j=1}^{N}u_{ij}lnu_{ij}$$
$$+\gamma^{-1}\sum_{i=1}^{C}\sum_{l=1}^{L}\omega_{il}ln\omega_{il}+\lambda\sum_{i=1}^{C}\sum_{j=1}^{N}u_{ij}ln\frac{u_{ij}}{\frac{1}{K}} \tag{12}$$

Among them, the constraints on the parameters are as follows:

$$\sum_{i=1}^{C}u_{ij}=1 \tag{13}$$

$$\sum_{l=1}^{L}\omega_{il}=1,\omega_{il}>0 \tag{14}$$

$$\alpha_j^{'}=\frac{1}{C}\sum_{i=1}^{C}u_{ij}\cdot\exp\left(-\parallel x_j-v_i^{(0)}\parallel\right) \tag{15}$$

In Eq. (12), $C$ denotes the number of clusters, and $N$ denotes the number of data points to be processed. $L$ denotes the feature dimension of the data and $q$ denotes the row position of the viewpoint. $u_{ij}$ denotes the degree of membership of a data point $x_j$ to cluster $i$, its range is [0, 1]. m is the fuzzy coefficient, and its range is $(0, +\infty)$. In general, the optimal range of m is between 1.5 and 2.5. The feature weight $\omega_{il}$ can represent the significance of the characteristics of the data to the cluster $i$. The $x_{dl}$ denotes the feature $l$ of the viewpoint $x_d$ obtained by the RDKM algorithm. The $\alpha_j^{'}$ is defined as the sample weight and the $v_i^{(0)}$ represents the i-th

initial cluster centre obtained by RDKM. The γ parameter is the regularization parameter, and its value is an adjustable parameter. By choosing an appropriate γ, a balance between the first two terms and the third term can be achieved to achieve better clustering performance.

In Eq. (12), $\alpha_j$ represents the weight $x_j$ of a data point, reflecting its importance. $v_i^{(0)}$ refers to the centre of the initial cluster i obtained by the RDKM algorithm. The initial cluster centres obtained by the methods are generally close to the reference cluster centres. Therefore, the initial cluster centre can be used to measure the importance of the data point $x_j$. The closer $x_j$ is to the initial clustering centre $V^{(0)}$, the larger the weight $\varphi_j$ is, indicating that the data point has a greater impact on the clustering results. In this way, the weights of noise points far away from the initial clustering centre will be minimized, which significantly reduces the influence of noise points on the clustering results.

The significance of the objective function design is as follows:

- First item: In the case that the cluster centre is not the viewpoint, the sample weight $\alpha_j^{'}$ is combined with the feature weight $\omega_{ij}$ to form a weighted information granule. Then, the weighted information granule was adjusted and optimized adaptively and dynamically by combining the membership degree, and the adaptive, optimized weighted information granule was obtained.
- Second item: In the case that the cluster centre is a viewpoint, the importance of the viewpoint is emphasized, and the number of clusters need not be included.
- Third item: This term enhances the fuzziness of the clustering and prevents the membership from going over degree extremes.
- Fourth item: Negentropy regularization is applied to optimize the distribution of feature weights, preventing the clustering process from relying on only a few dimensions.
- Fifth item: By adjusting the weight of KL (Kullback-Leibler) divergence, the distribution of membership entropy can be controlled, and the dispersion of membership can be promoted.

Based on the mathematical concept of weighted information granule, we propose an objective function structure of adaptive optimization weighted information granule, which includes the following three main stages:

- Feature weight matrix construction: We assign different weights to different features of the data to achieve the effect of reducing the adverse influence of irrelevant features.
- Sample weight assignment: We assign varying weights to the features of the data to reduce the effect of influence of irrelevant features.
- Adaptive adjustment of information granules: We dynamically adjust the weighted information granules to optimize their adaptability and improve the clustering performance.

This paper introduces the concept of applying weighted information granularity in fuzzy clustering, which mainly includes two research aspects:

In the clustering process, the weighted information granules are dynamically optimized to obtain better clustering results. We name this structure the weighted information granule for adaptive optimization.

Moreover, we can combine the first and second terms in the objective function to unify the representation of viewpoint and non-viewpoint cluster centres into a single form. Finally, the reduced objective function can be expressed as follows.

$$J_{RDVWKFC} = \sum_{i=1}^{C}\sum_{j=1}^{N}\alpha_j^{'}u_{ij}^{m}\sum_{l=1}^{L}\omega_{il} \parallel \phi\left(x_{jl}\right) - \phi\left(g_{il}\right) \parallel^2 + \gamma^{-1}\sum_{i=1}^{C}\sum_{l=1}^{L}\omega_{il}ln\omega_{il} + \eta\sum_{i=1}^{C}\sum_{j=1}^{N}u_{ij}lnu_{ij} + \lambda\sum_{i=1}^{C}\sum_{j=1}^{N}u_{ij}ln\frac{u_{ij}}{\frac{1}{K}} \tag{16}$$

among them:

$$g_{il} = \begin{cases} x_{dl} & i = q \\ x_{il} & i \neq q \end{cases} \tag{17}$$

We describe this procedure in detail below.

Firstly, the RDVWKFC algorithm was used to obtain the $V^{(0)}$ and $x_d$. To determine the row position $d$ of the viewpoint $x_d$ in the cluster centre matrix $V$, we need to find the cluster centre $x_d$ in the matrix that is closest to the viewpoint $v_q$. Specifically, suppose the row position of the viewpoint in the cluster centre matrix is $q$ and the cluster centre is $v_q$. By calculating the kernel distance, the cluster centre closest to the viewpoint can be found.

$$d_{qd} = \left\| \varphi(x_d) - \varphi(v_q) \right\|^2 \tag{18}$$

$d_{qd}$ reflects the shortest distance between the viewpoint and the cluster centre. At this point, the viewpoint replaces the location of the cluster centre.

In the process of algorithm execution, because the value of the viewpoint will shift constantly, its position in the cluster centre matrix will also shift. A key aspect of this process is the calculation of the kernel distance, which ensures that viewpoints can adjust their position according to the density of the data distribution, thus reflecting the data structure more accurately.

On this framework, the proposed RDVWKFC algorithm aims to achieve more accurate and robust clustering analysis by considering the local density and feature weights of the data. The algorithm improves the stability and accuracy of clustering results by dynamically adjusting the density viewpoint.

The mechanism and derivation of the objective function of the new algorithm are described in detail above.

Finally, the running process of the new algorithm is summarized as shown in Table 2.

**Table 2:** The proposed RDVMKFC algorithm

**Inputs:** Data $X = \{\mathrm{x}_j\}_{j=1}^{N}$, cluster number C, m, $\sigma$, y, iteration stop threshold e and maximum number of iterations iM.

**Outputs:** Membership matrix $\mathrm{U} = \{\mathrm{u}_{ij}\}_{i,j=1}^{C,N}$, cluster centre matrix $\mathrm{G} = \{g_i\}_{j=1}^{C}$, and feature weight matrix $\mathrm{W} = \{\omega_{il}\}_{i,l=1}^{C,L}$

The cluster centres G(0) is initialized by running RDKM, and the density viewpoint $x_d$ is obtained;
Set $\omega_{il}$ uniformly as 1/L;
Let iter = 0;
repeat
    iter = iter + 1;
Compute $u_{ij}$ by (A8), and update U(iter);
Calculate $g_i$ by (A12), and renew G(iter);
Figure up $\omega_{il}$ by (A17), and update W(iter):
until ||G(iter) − G(iter − 1)|| <e or iter> iM;
return U(iter), G(iter), W(iter);
end procedure

## 5 Experiments

To verify the effectiveness of the proposed RDVWKFC algorithm, we conduct a series of experiments to observe the clustering results. In addition, nine other algorithms were selected for comparison, including FCM, KFCM, V-FCM, EWFCM, FRFCM, DVPFCM, VWKFC, FCAG, and SR-PCM-HDP. The experiments were performed on the Windows 10 platform and the programming was done in Matlab 2022b. During the experiments, the algorithm was run on 13 datasets, including 5 synthetic datasets (Data-a, Data-b, Data-c, Data-d, and Data-e) and 8 UCI datasets. The UCI datasets used include Iris, Wisconsin Breast Cancer dataset, Letter recognition (A, B) dataset, SPECT heart dataset, Wine dataset, Landsat, seed and Wifi localization dataset.

To quantitatively evaluate the clustering performance, we used five different evaluation metrics: classification accuracy (ACC), Normalized Mutual Information (NMI) [23], Calinski-Harabasz index [24] (CH), ARI (Adjusted Rand Index) Expansion Index (EARI) [25,26], and Xie-Beni (XB) index [27]. Among them, ACC, NMI, and CH are hard clustering metrics that are applicable to both hard and soft clustering algorithms. EARI and XB indices are soft clustering metrics. EARI and XB index are only suitable for soft clustering algorithms because a membership degree is required for calculation.

### *5.1 Experiments on Synthetic Datasets*

We first conduct an experimental evaluation on five synthetic datasets (Data-a to Data-e) to verify the performance of the proposed algorithm. These datasets are generated by Python's scikit-learn library and have clear cluster structure characteristics. We focused our analysis on the results of Data-a and b. The datasets used here are all two-attribute datasets, which can be visualized as scatter plots. Among them, Data-a and Data-b, as benchmark sets, both contain 7 clusters and 1000 two-dimensional data points, which are characterized by moderate overlap between clusters and tight internal structure. Data-d, on the other hand, presents a higher complexity, containing 12 clusters and 1500 data points, where the clusters are closely linked and there is some degree of overlap between the classes. Although Data-e has the same number of clusters (7) and Data size (1000 points) as Data-a and Data-b, its spatial distribution pattern is more complex. In contrast, Data-c, as the simplest test set, contains only three clusters and 150 data points, but its remarkable feature is that there is a greater degree of overlap between clusters, which puts forward higher requirements for the robustness of the clustering algorithm. A greater degree of overlap between classes poses a more significant challenge to clustering algorithms.

In Table 3, we present the evaluation metric values obtained after running the proposed algorithm and the comparison algorithm on different artificial datasets. Here, "(+)" means that the evaluation criterion is "bigger is better", which indicates higher values, while "(−)" means "smaller is better", which indicates lower values. It can be seen that our proposed new algorithm achieves the best results in all evaluation metrics, among which the ACC, NMI, and EARI values on the datasets are the highest.

The new algorithm uses the RDKM method to initialize the cluster centres, which avoids the problem of convergence to a local optimum caused by improper initialization of the cluster centres. In the clustering process, feature selection is carried out by feature weights to make the clustering results more accurate. In addition, the introduction of high-density viewpoints reduces the influence of outliers and noise points on the clustering results. As shown in Table 3, the accuracy of the new algorithm on the artificial dataset Data-b is the highest, which is 0.9880, and the CH index value reaches 15.1141. Table 3 shows the results for different synthetic datasets.

**Table 3:** Performance of comparative algorithms on synthetic datasets

| | | FCM | KFCM | V-FCM | EWFCM | FRFCM | DVPFCM | VWKFC | FCAG | SR-PCM-HDP | RDVWKFC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Data-a | ACC(+) | 0.9940 | 0.9940 | 0.9940 | 0.9690 | 0.9810 | 0.9670 | 0.9930 | 0.9890 | 0.7590 | 0.9950 |
| | CH(+) | 14.4732 | 14.4552 | 14.3819 | 13.0286 | 13.9121 | 14.0504 | 14.4247 | 13.8007 | 0.2378 | 14.4866 |
| | NMI(+) | 0.9831 | 0.9838 | 0.9823 | 0.9343 | 0.9580 | 0.9344 | 0.9807 | 0.9702 | 0.7653 | 0.9861 |
| | EARI(+) | 0.9932 | 0.9940 | 0.9918 | 0.9566 | 0.9795 | 0.9658 | 0.9928 | 0.9885 | 0.7078 | 0.9941 |
| | XB(−) | 0.1096 | 0.2278 | 0.1148 | 0.1754 | 0.1230 | 1.9737 | 0.1952 | 0.6190 | 0.3412 | 0.0902 |
| | RECALL(+) | 0.8589 | 0.9090 | 0.9066 | 0.8883 | 0.6784 | 0.9487 | 0.9710 | 0.9782 | 0.8278 | 0.9900 |
| | FSCORE(+) | 0.7978 | 0.8223 | 0.8216 | 0.8661 | 0.6049 | 0.9481 | 0.9708 | 0.9781 | 0.6708 | 0.9900 |
| Data-b | ACC(+) | 0.7880 | 0.9880 | 0.9870 | 0.9670 | 0.9600 | 0.7040 | 0.9880 | 0.9870 | 0.8630 | 0.9880 |
| | CH(+) | 14.2019 | 14.9013 | 14.8926 | 13.7571 | 14.7787 | 14.836 | 14.9467 | 14.2088 | 12.1857 | 15.1141 |
| | NMI(+) | 0.8997 | 0.9671 | 0.9659 | 0.9295 | 0.9295 | 0.9029 | 0.9671 | 0.9655 | 0.9130 | 0.9671 |
| | EARI(+) | 0.9467 | 0.9841 | 0.9856 | 0.9531 | 0.9581 | 0.8980 | 0.9868 | 0.9864 | 0.8904 | 0.9870 |
| | XB(−) | 0.5334 | 0.2036 | 0.1059 | 0.3288 | 0.1487 | 0.9030 | 0.1055 | 0.6128 | 0.1272 | 0.0990 |
| | RECALL(+) | 0.8443 | 0.8690 | 0.8729 | 0.8471 | 0.6780 | 0.8942 | 0.9666 | 0.9743 | 0.9569 | 0.9802 |
| | FSCORE(+) | 0.7503 | 0.7727 | 0.7705 | 0.7644 | 0.5945 | 0.8129 | 0.9665 | 0.9742 | 0.8564 | 0.9801 |
| Data-c | ACC(+) | 0.7133 | 0.7667 | 0.9667 | 0.6467 | 0.9533 | 0.9600 | 0.9933 | 0.9800 | 0.7067 | 0.9933 |
| | CH(+) | 8.3890 | 10.9977 | 10.2349 | 10.2172 | 10.9952 | 6.3406 | 12.0128 | 11.5764 | 1.1610 | 12.3104 |
| | NMI(+) | 0.6380 | 0.6829 | 0.8997 | 0.6433 | 0.8342 | 0.8702 | 0.9702 | 0.9188 | 0.7154 | 0.9702 |
| | EARI(+) | 0.8108 | 0.9446 | 0.9523 | 0.8682 | 0.9306 | 0.9486 | 0.9911 | 0.9736 | 0.4787 | 0.9911 |
| | XB(−) | 2.0545 | 1.6796 | 0.2849 | 0.8822 | 0.2138 | 2.1823 | 0.2003 | 0.2716 | 0.2034 | 0.1907 |
| | RECALL(+) | 1.0000 | 1.0000 | 0.9616 | 1.0000 | 0.9600 | 0.9499 | 0.9739 | 0.9605 | 0.9282 | 1.0000 |
| | FSCORE(+) | 1.0000 | 1.0000 | 0.9605 | 1.0000 | 0.9599 | 0.9479 | 0.9733 | 0.9596 | 0.7347 | 1.0000 |
| Data-d | ACC(+) | 0.5940 | 0.5000 | 0.6420 | 0.5400 | 0.736 | 0.4793 | 0.7293 | 0.6027 | 0.7133 | 0.8347 |
| | CH(+) | 7.3025 | 7.4783 | 7.3953 | 7.5401 | 6.9706 | 7.3105 | 7.3671 | 6.8499 | 6.3598 | 7.5672 |
| | NMI(+) | 0.7094 | 0.7096 | 0.7113 | 0.6581 | 0.7548 | 0.6559 | 0.7958 | 0.7643 | 0.7409 | 0.8011 |
| | EARI(+) | 0.7005 | 0.7072 | 0.6652 | 0.6457 | 0.7833 | 0.4786 | 0.7954 | 0.7108 | 0.7822 | 0.8177 |
| | XB(−) | 0.0608 | 0.1036 | 0.0537 | 0.8228 | 0.1127 | 0.0701 | 0.0548 | 1.9735 | 0.4685 | 0.0529 |
| | RECALL(+) | 0.5927 | 0.5674 | 0.6825 | 0.7162 | 0.7117 | 0.7011 | 0.7406 | 0.7081 | 0.6877 | 0.7425 |
| | FSCORE(+) | 0.5131 | 0.6457 | 0.6623 | 0.6698 | 0.5212 | 0.6873 | 0.7003 | 0.6436 | 0.6246 | 0.7412 |
| Data-e | ACC(+) | 0.7740 | 0.7710 | 0.7710 | 0.8790 | 0.8140 | 0.9620 | 0.9700 | 0.9670 | 0.7670 | 0.9720 |
| | CH(+) | 14.8241 | 14.8379 | 15.8842 | 14.5288 | 15.7315 | 15.5855 | 15.8434 | 15.1690 | 12.8991 | 15.8847 |
| | NMI(+) | 0.8906 | 0.8906 | 0.9439 | 0.8558 | 0.8894 | 0.9336 | 0.9414 | 0.9379 | 0.8490 | 0.9439 |
| | EARI(+) | 0.9581 | 0.9645 | 0.9747 | 0.9210 | 0.9518 | 0.9625 | 0.9746 | 0.9661 | 0.9687 | 0.9744 |
| | XB(−) | 0.7266 | 1.6370 | 0.2105 | 2.5234 | 0.2657 | 1.7740 | 0.2691 | 2.4422 | 0.0049 | 0.2034 |
| | RECALL(+) | 0.9021 | 0.8980 | 0.9014 | 0.8923 | 0.8610 | 0.9317 | 0.9008 | 0.9387 | 0.8554 | 0.9468 |
| | FSCORE(+) | 0.8143 | 0.8100 | 0.8844 | 0.8480 | 0.8504 | 0.9292 | 0.8920 | 0.9376 | 0.7652 | 0.9464 |

### *5.2 Experiments with UCI Datasets*

Subsequently, we conducted experiments using eight UCI datasets. These datasets are as follows: Iris, Wisconsin Breast Cancer, letter recognition (A, B), SPECT heart data, Wine, Statlog (Landsat satellite), Seed, and Wifi localization.

The Iris dataset is one of the most classic datasets in the field of machine learning. It contains 150 data points, and each data point contains four feature attributes: sepal length, sepal width, petal length, and petal width. The Iris dataset can be divided into three classes: Iris-setosa (Iris mountain), Iris-versicolor (Iris changing colors), and Iris-virginica (Iris Virginia). The Wisconsin Breast Cancer dataset is a cancer dataset containing 569 sample data points, which can be divided into two classes: benign and malignant. The Letter Recognition (A, B) dataset is a character recognition dataset. The Spect heart dataset is a heart disease dataset with binary data. Wine is a dataset about wine, containing 178 samples and 13 numerical features. It is divided

into three categories, each with a different sample size. The Stalog (Land Satellite) dataset is a Landsat dataset with 36 feature attributes. The Seed dataset is about wheat seeds, contains 70 data samples, and belongs to three different wheat varieties: Kama, Rosa, Canadian. Table 4 gives the details of the UCI datasets we used in the experiments, which include the number of data samples, the number of categories, and the number of features.

**Table 4:** Details for each UCI dataset

| Name | Instances | Atttibutes | Classes |
|---|---|---|---|
| Iris | 150 | 4 | 3 |
| Wine | 178 | 13 | 3 |
| Breast_cancer Wisconsin | 569 | 30 | 2 |
| SPECT heart data | 267 | 22 | 2 |
| Stalog (Landsat satellite) | 6435 | 36 | 6 |
| Seeds | 210 | 7 | 3 |
| Letter Recognition (A,B) | 1555 | 16 | 2 |
| Wifi localization | 2000 | 7 | 4 |

Table 5 presents the evaluation metrics for different clustering algorithms applied to various UCI datasets.

**Table 5:** Performance results of comparison algorithms on UCI datasets

| | | FCM | KFCM | V-FCM | EWFCM | FRFCM | DVPFCM | VWKFC | FCAG | SR-PCM-HDP | RDVWKFC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Iris | ACC(+) | 0.8867 | 0.8800 | 0.9533 | 0.8933 | 0.9067 | 0.8400 | 0.9467 | 0.9000 | 0.9133 | 0.9733 |
| | CH(+) | 11.6197 | 11.6251 | 11.8362 | 11.5658 | 11.4901 | 11.9263 | 12.0054 | 11.0659 | 3.7412 | 12.0121 |
| | NMI(+) | 0.7353 | 0.7277 | 0.8498 | 0.7496 | 0.7857 | 0.6712 | 0.8322 | 0.7476 | 0.7839 | 0.9011 |
| | EARI(+) | 0.8703 | 0.8868 | 0.9312 | 0.8733 | 0.9172 | 0.7832 | 0.9333 | 0.8980 | 0.9088 | 0.9643 |
| | XB(−) | 0.2903 | 0.3446 | 0.2787 | 0.2806 | 0.3484 | 2.5841 | 0.2961 | 0.5028 | 0.2711 | 0.2546 |
| | RECALL(+) | 0.8190 | 0.8220 | 0.9148 | 0.8299 | 0.8824 | 0.8310 | 0.9238 | 0.8299 | 0.8430 | 0.9483 |
| | FSCORE(+) | 0.8101 | 0.8180 | 0.9117 | 0.8271 | 0.8781 | 0.8123 | 0.9233 | 0.8271 | 0.8415 | 0.9478 |
| Breast Cancer | ACC(+) | 0.9209 | 0.9016 | 0.9297 | 0.8041 | 0.8981 | 0.9016 | 0.9016 | 0.9279 | 0.7610 | 0.9315 |
| | CH(+) | 1.1610 | 17.2804 | 78.259 | 15.3786 | 6.1301 | 21.0100 | 34.4968 | 78.5709 | 9.8154 | 107.3924 |
| | NMI(+) | 0.5947 | 0.5676 | 0.6194 | 0.3377 | 0.5077 | 0.5443 | 0.5208 | 0.6264 | 0.2736 | 0.6293 |
| | EARI(+) | 0.8159 | 0.7777 | 0.8982 | 0.8354 | 0.7636 | 0.8135 | 0.8593 | 0.8890 | 0.4125 | 0.9116 |
| | XB(−) | 85.0710 | 1.4380 | 1.3400 | 5.0580 | 2.8875 | 1.2216 | 1.3478 | 0.6289 | 0.3905 | 0.2974 |
| | RECALL(+) | 0.8752 | 0.8956 | 0.9073 | 0.8859 | 0.8431 | 0.8363 | 0.9082 | 0.8601 | 0.6318 | 0.9115 |
| | FSCORE(+) | 0.8766 | 0.8917 | 0.8929 | 0.8780 | 0.8504 | 0.8447 | 0.8782 | 0.8722 | 0.6483 | 0.8934 |
| Letter_ AB | ACC(+) | 0.9273 | 0.9331 | 0.8887 | 0.9048 | 0.9293 | 0.9305 | 0.9344 | 0.9389 | 0.8051 | 0.9408 |
| | CH(+) | 135.9163 | 166.0395 | 172.2868 | 44.2458 | 177.5480 | 155.0735 | 212.4524 | 204.0585 | 49.9026 | 213.3964 |
| | NMI(+) | 0.6483 | 0.6811 | 0.5021 | 0.5551 | 0.6528 | 0.6955 | 0.7069 | 0.7256 | 0.3487 | 0.7315 |
| | EARI(+) | 0.8517 | 0.8686 | 0.8643 | 0.7868 | 0.8642 | 0.8718 | 0.8952 | 0.9829 | 0.2906 | 0.8995 |
| | XB(−) | 0.3579 | 0.4017 | 0.7316 | 1.0933 | 0.3851 | 0.6098 | 0.4769 | 0.6800 | 0.3678 | 0.3421 |
| | RECALL(+) | 0.8644 | 0.8697 | 0.8792 | 0.8588 | 0.8722 | 0.8823 | 0.8859 | 0.8908 | 0.7361 | 0.8978 |
| | FSCORE(+) | 0.8612 | 0.8667 | 0.8765 | 0.8567 | 0.8689 | 0.8789 | 0.8803 | 0.8858 | 0.7009 | 0.8936 |
| Wine | ACC(+) | 0.9045 | 0.927 | 0.882 | 0.8596 | 0.8876 | 0.9045 | 0.9438 | 0.9213 | 0.6124 | 0.9719 |
| | CH(+) | 9.2506 | 12.1667 | 10.4631 | 14.4481 | 2.0072 | 15.3527 | 10.2622 | 10.1291 | 0.9514 | 14.8109 |
| | NMI(+) | 0.7295 | 0.771 | 0.7047 | 0.6504 | 0.6888 | 0.7255 | 0.7987 | 0.7553 | 0.2662 | 0.8920 |
| | EARI(+) | 0.8012 | 0.8593 | 0.8236 | 0.7848 | 0.7323 | 0.8215 | 0.8813 | 0.9220 | 0.4256 | 0.9738 |
| | XB(−) | 1.2469 | 1.3542 | 0.9811 | 1.4104 | 1.1924 | 3.7708 | 1.4889 | 1.0218 | 0.9978 | 0.9384 |
| | RECALL(+) | 0.8125 | 0.8188 | 0.7840 | 0.7949 | 0.8328 | 0.8627 | 0.8708 | 0.8439 | 0.5841 | 0.9369 |

(Continued)

**Table 5 (continued)**

| | | FCM | KFCM | V-FCM | EWFCM | FRFCM | DVPFCM | VWKFC | FCAG | SR-PCM-HDP | RDVWKFC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SPECT heart | FSCORE(+) | 0.8214 | 0.8170 | 0.7920 | 0.8020 | 0.8419 | 0.8699 | 0.8776 | 0.8461 | 0.4964 | 0.9425 |
| | ACC(+) | 0.6030 | 0.7004 | 0.7678 | 0.7940 | 0.6255 | 0.8427 | 0.7865 | 0.7753 | 0.8052 | 0.8543 |
| | CH(+) | 0.0954 | 0.6909 | 30.5773 | 21.9245 | 0.5926 | 33.2145 | 11.9338 | 22.3562 | 1.3601 | 33.2571 |
| | NMI(+) | 0.1029 | 0.1394 | 0.1908 | 0.2518 | 0.1613 | 0.2742 | 0.1305 | 0.1537 | 0.1594 | 0.3209 |
| | EARI(+) | 0.2636 | 0.3730 | 0.8316 | 0.8675 | 0.2893 | 0.8935 | 0.4925 | 0.7635 | 0.1257 | 0.8943 |
| | XB(−) | 24.8697 | 12.5958 | 0.8664 | 1.3100 | 2.6459 | 0.7489 | 0.7165 | 1.2925 | 1.8233 | 0.6948 |
| | RECALL(+) | 0.5169 | 0.5644 | 0.6532 | 0.6337 | 0.5268 | 0.7569 | 0.8844 | 0.6682 | 0.7406 | 0.8922 |
| | FSCORE(+) | 0.5910 | 0.6428 | 0.7174 | 0.6358 | 0.6049 | 0.7833 | 0.7881 | 0.7196 | 0.7596 | 0.8242 |
| Statlog | ACC(+) | 0.9463 | 0.7214 | 0.7759 | 0.8202 | 0.8301 | 0.9445 | 0.9521 | 0.9580 | 0.9302 | 0.9611 |
| | CH(+) | 188.0739 | 262.4992 | 130.0652 | 109.7552 | 60.6775 | 142.9682 | 400.2993 | 379.8112 | 84.6025 | 440.698 |
| | NMI(+) | 0.676 | 0.3165 | 0.3651 | 0.2497 | 0.4027 | 0.713 | 0.7418 | 0.7650 | 0.6626 | 0.7780 |
| | EARI(+) | 0.9035 | 0.4856 | 0.7788 | 0.4376 | 0.6277 | 0.9004 | 0.9407 | 0.9448 | 0.9925 | 1.0000 |
| | XB(−) | 2.1408 | 0.3423 | 2.5743 | 0.6826 | 1.1449 | 1.2239 | 0.3239 | 0.3920 | 0.4012 | 0.3137 |
| | RECALL(+) | 0.9590 | 0.9608 | 0.9586 | 0.7152 | 0.9599 | 0.9558 | 0.9685 | 0.9597 | 0.9400 | 0.9700 |
| | FSCORE(+) | 0.9299 | 0.9333 | 0.9305 | 0.7452 | 0.9365 | 0.9239 | 0.9485 | 0.9313 | 0.8917 | 0.9499 |
| Seeds | ACC(+) | 0.5524 | 0.6841 | 0.8952 | 0.9 | 0.8571 | 0.8952 | 0.8952 | 0.8905 | 0.8571 | 0.9048 |
| | CH(+) | 10.4292 | 10.8944 | 16.046 | 14.8147 | 8.5833 | 15.4141 | 11.9425 | 14.7541 | 8.7259 | 15.6115 |
| | NMI(+) | 0.5446 | 0.5445 | 0.6781 | 0.6842 | 0.6192 | 0.6962 | 0.6395 | 0.6707 | 0.6049 | 0.8150 |
| | EARI(+) | 0.6998 | 0.4813 | 0.858 | 0.866 | 0.716 | 0.8352 | 0.9316 | 0.8743 | 0.8047 | 0.9391 |
| | XB(−) | 0.3916 | 0.1515 | 0.5709 | 0.5649 | 0.2654 | 3.1523 | 0.9233 | 0.5802 | 0.1149 | 0.0686 |
| | RECALL(+) | 0.8188 | 0.8050 | 0.8178 | 0.8141 | 0.8032 | 0.8018 | 0.8083 | 0.8090 | 0.7516 | 0.8210 |
| | FSCORE(+) | 0.8170 | 0.8026 | 0.8173 | 0.8070 | 0.7981 | 0.7991 | 0.8048 | 0.8049 | 0.7460 | 0.8188 |
| Wifi Localization | ACC(+) | 0.6000 | 0.6295 | 0.6830 | 0.6955 | 0.6790 | 0.9240 | 0.9565 | 0.9575 | 0.7720 | 0.9585 |
| | CH(+) | 0.5438 | 35.5301 | 72.3962 | 55.6873 | 18.7185 | 69.5697 | 78.6978 | 74.6102 | 34.9604 | 79.5165 |
| | NMI(+) | 0.5391 | 0.5267 | 0.8049 | 0.5823 | 0.5003 | 0.7928 | 0.8743 | 0.8693 | 0.7485 | 0.8929 |
| | EARI(+) | 0.5039 | 0.4447 | 0.7769 | 0.5344 | 0.4585 | 0.8599 | 0.9600 | 0.9532 | 0.7336 | 0.9655> |
| | XB(−) | 0.8616 | 4.6048 | 0.9243 | 17.3775 | 0.6079 | 0.8697 | 0.4638 | 0.7373 | 1.9549 | 0.4529 |
| | RECALL(+) | 0.8909 | 0.9055 | 0.8953 | 0.9046 | 0.9071 | 0.8914 | 0.9200 | 0.9196 | 0.8955 | 0.9248 |
| | FSCORE(+) | 0.8878 | 0.9050 | 0.8946 | 0.9040 | 0.9035 | 0.8888 | 0.9199 | 0.9180 | 0.7620 | 0.9206 |

From the data in this tables, it can be seen that the proposed algorithm has high accuracy and is better than previous clustering algorithms in other indicators. For the classic Iris dataset, the accuracy of the new algorithm reaches 0.9733. In the datasets with high feature dimensions, such as Wisconsin Breast Cancer dataset, SPECT heart dataset and Statlog dataset, the accuracy of the new algorithm reaches 0.9315, 0.8543 and 0.9611, respectively. This shows that the new algorithm has certain advantages over the previous algorithms when dealing with data sets with more attribute features.

We summarize the performance metrics of the RDVWKFC algorithm on eight UCI datasets in the Appendix A, the proposed RDVWKFC algorithm can obtain the best results on these eight UCI datasets.

### *5.3 Experiments on High-Dimensional Datasets*

In order to verify the performance of the RDVWKFC algorithm in clustering high-dimensional datasets, we conducted several experiments by randomly selecting three classes from the Mfeat dataset with 649 features. The clustering performance is evaluated using two metrics: ACC and EARI. Table 6 presents each algorithm result in Mfeat dataset. The calculation accuracy of the algorithm on the Mfeat dataset is 0.6150, and the EARI value is 0.9555. The experimental results present that the proposed algorithm on the data set has the best clustering performance.

**Table 6:** Evaluation indexes of the new algorithm running on high-dimensional Mfeat dataset

| | FCM | KFCM | V-FCM | EWFCM | FRFCM | DVPFCM | VWKFC | FCAG | SR-PCM -HDP | RDVWKFC |
|---|---|---|---|---|---|---|---|---|---|---|
| ACC(+) | 0.2200 | 0.2917 | 0.4688 | 0.3467 | 0.5733 | 0.4983 | 0.3300 | 0.5033 | 0.4483 | 0.6150 |
| EARI(+) | 0.3073 | 0.2485 | 0.6308 | 0.8766 | 0.4721 | 0.4774 | 0.9555 | 0.8308 | 0.9328 | 0.9555 |
| RECALL(+) | 0.7436 | 0.7398 | 0.7983 | 0.8230 | 0.6376 | 0.7219 | 0.6645 | 0.6439 | 0.5426 | 0.8376 |
| FSCORE(+) | 0.6247 | 0.6647 | 0.7524 | 0.6579 | 0.6252 | 0.5907 | 0.6291 | 0.5776 | 0.4336 | 0.7337 |

Table 6 presents the evaluation metrics for different clustering algorithms applied to the high-dimensional Mfeat dataset.

Through the above experimental results, we can see that the algorithm shows comprehensive advantages in four indicators. The value of ACC is ahead of the second place FRFCM, indicating that its clustering results have the highest matching degree with the true label, and the value of the EARI index is tied for the first place, which verifies its robustness under a complex clustering structure. RECALL and FSCORE both rank first, which indicates that the algorithm performs best in the balance of coverage and precision.

Different from the traditional feature weighted clustering algorithm in the past, the new algorithm uses RDKM method to initialize the clustering centre and introduces the viewpoint, which avoids the local minimum problem in the objective function, which avoids the local minimum problem in the objective function. In contrast to previous studies, our method assigns weights to feature attributes, so that it can show better results while dealing with complex datasets.

Please refer to the Appendix A for the relevant supplementary experiments and discussions.

## 6 Conclusion and Outlook

In this paper, we proposed a new adaptive weighted clustering algorithm and the main contributions and work are summarized as follows:

(i) Faced with the challenges that traditional clustering algorithms often encounter in the initialization phase, the clustering results are highly sensitive to the initial centroid selection and unstable to noise and outliers. We innovatively introduce the Relative Density based Knowledge Extraction Method (RDKM) as the initialization strategy on the original KHDI algorithm. This method determines the initial centroid by identifying relatively high-density regions in the data, which lays a more reasonable foundation for the subsequent clustering process. Compared with the existing methods, RDKM significantly improves the quality of the clustering initialization stage, effectively reduces the risk of clustering results deviation caused by improper selection of initial centroids, and the algorithm has significant advantages in the initial stage.

(ii) In order to alleviate the negative impact of noise and outliers on clustering accuracy, we propose an adaptive weighting mechanism through the construction of the Information Granularity Weight Model (IGWM). We implement the dynamic adjustment of the weights of data points according to their importance in the clustering process. By giving more weight to the data points with higher importance to make them play a key role in the clustering decision, and giving less weight to the points that may be noise or outliers, it effectively weakens their influence on the clustering results. This method improves the reliability of clustering results and significantly enhances the adaptability of the algorithm to complex data environments.

(iii) In order to further improve the performance of the algorithm in different data characteristics and dimensions (the function can adapt to low dimensions, but also can adapt to high and ultra-high

dimensions), we design an adaptive optimization algorithm (AOA). AOA can adjust the weights in real-time according to the dynamic change of the data distribution and always maintain good adaptability of the algorithm. At the same time, when dealing with high-dimensional or feature-diverging data, AOA shows a high degree of flexibility, which enables the algorithm to run stably in various complex scenarios and effectively overcomes the limitations of traditional algorithms in dealing with such data.

(iv) Combining the above innovations, we propose a new adaptive weighted clustering algorithm—RDVWKFC. This algorithm integrates the advantages of each algorithm or method to build an excellent collaborative algorithm mechanism and provides a comprehensive and effective solution for dealing with complex data clustering problems.

(v) Experimental verification results show that the algorithm has achieved a remarkable application effect in the application. In the test of various comparison algorithms, the new algorithm shows superior performance on multiple data sets. For example, ACC, NMI, CH, EARI, and XB are superior to other algorithms on key evaluation indicators. In addition, in terms of algorithm efficiency, the number of iterations required by this algorithm is significantly less than that of the comparison algorithm, indicating that it can quickly converge to stable clustering results and significantly improve computing efficiency. It is worth noting that when the new algorithm is applied to the image segmentation data set for the image clustering task, its clustering effect is superior to the existing methods, which further verifies its effectiveness and superiority in practical applications.

Our study demonstrates that the proposed algorithm significantly improves clustering accuracy and computational efficiency compared to existing methods. Future research will expand the application scope of the new algorithm, especially in fields such as fuzzy inference and fuzzy systems [28]. We also expect that the novel algorithm will exhibit superior performance in these real-world application scenarios, and provide a more effective solution for solving complex data clustering problems.

**Acknowledgement:** Special thanks to the reviewers of this paper for their valuable feedback and constructive suggestions, which greatly contributed to the refinement of this research.

**Funding Statement:** This work was supported by the National Natural Science Foundation of China under Grants 62176083, 62176084, 61877016, and 61976078.

**Author Contributions:** Conceptualization, Yiming Tang and Yuhan Xia; methodology, Xu Li; software, Ye Liu; validation, Xu Li and Wenbo Zhou; formal analysis, Yiming Tang; investigation, Ye Liu; resources, Wenbo Zhou; data curation, Yuhan Xia; writing—original draft preparation, Yuhan Xia and Xu Li; writing—review and editing, Yiming Tang; visualization, Ye Liu; supervision, Yiming Tang; project administration, Xu Li; funding acquisition, Yiming Tang. All authors reviewed the results and approved the final version of the manuscript.

**Availability of Data and Materials:** Not applicable.

**Ethics Approval:** Not applicable.

**Conflicts of Interest:** The authors declare no conflicts of interest to report regarding the present study.

## Appendix A

### *Appendix A.1 Summary of the Experiments*

Fig. A1 presents scatter plots of these five synthetic datasets. These datasets are composed of two-dimensional data. The horizontal and vertical coordinates in Fig. A1 correspond to two dimensions of the data, respectively.

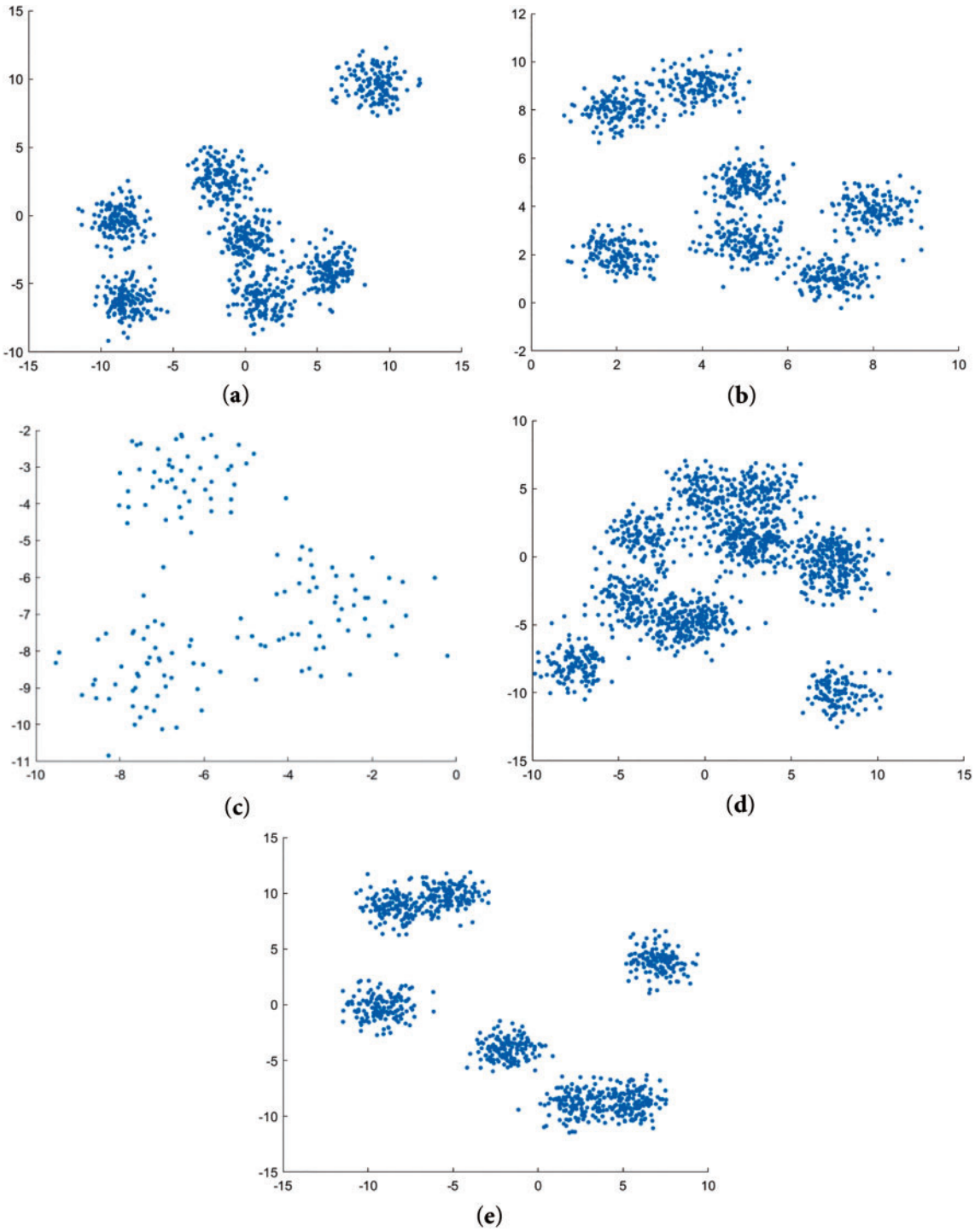


**Figure A1:** Scatter diagrams of artificial data sets. **(a)** Data-a dataset. **(b)** Data-b dataset. **(c)** Data-c dataset. **(d)** Data-d dataset. **(e)** Data-e dataset

The results of the proposed RDVWKFC algorithm on the five synthetic datasets are shown in Fig. A2. The horizontal and vertical coordinates in Fig. A2 also correspond to two dimensions of the data, respectively.

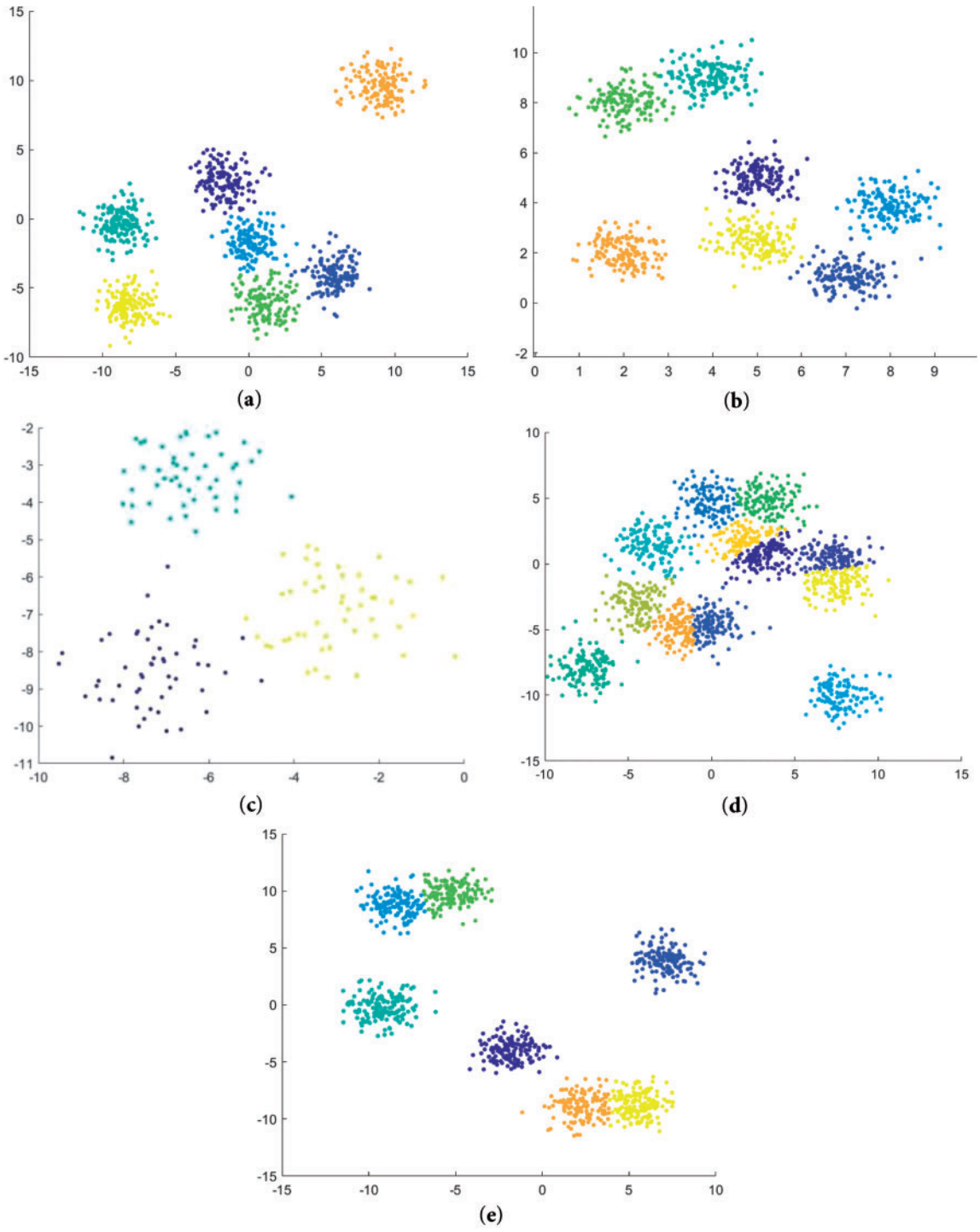


**Figure A2:** Operation results of the RDVWKFC algorithm on artificial data sets. **(a)** Data-a dataset. **(b)** Data-b dataset. **(c)** Data-c dataset. **(d)** Data-d dataset. **(e)** Data-e dataset

As shown in Fig. A3, our proposed new algorithm processes data points from different clusters in different datasets and labels them with easily distinguishable colors and shapes. From the visualized results of the operation, it is not difficult to see that our algorithm shows excellent results on these data sets. The

black dots in the rendering represent the location of the cluster centre obtained by the new algorithm. The proposed algorithm can accurately locate the cluster centre, which is particularly obvious in the dataset Data-e. Even though there is significant overlap between the different twelve categories. The algorithm can still accurately identify 12 different clustering centres.

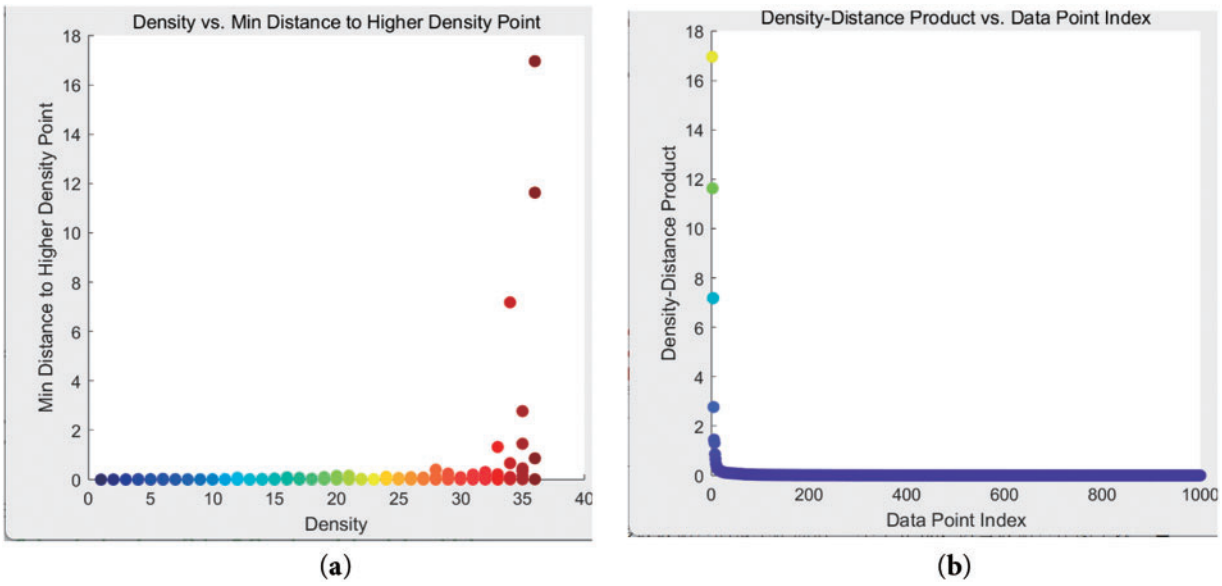


**Figure A3:** The density-distance $(\rho - \delta)$ decision plot on the Data-b dataset. **(a)** Density vs. min distance to higher density point. **(b)** Density-distance product vs. data point index

To further validate the effectiveness of the proposed RDKM method, we examine the density-distance $(\rho - \delta)$ determination plot and the parameter $\tau$ distribution plot after running the RDKM method on the Data-b dataset, as shown in Fig. A3.

Fig. A3a shows the density-distance $(\rho - \delta)$ decision diagram after processing the dataset using the RDKM method. It can be seen from the Fig. A3a that the local density of the data points is evenly distributed throughout the range. Most of the data points have a minimum distance value less than 0.05, and these data points are represented by a "•" in the figure. By examining Fig. A3a, it can be noticed that there are seven data points that are significantly different from the others. These points have relatively large minimum distance values and significantly higher local density values compared to the other data points. In the decision diagram, the seven points are marked with different symbols. Accordingly, Fig. A3b shows the parameter distribution plot for the Data-b dataset. It is evident from the figure that the parameter values of seven data points in Fig. A3b are significantly larger than those of most of the data points. Most of the data points have values close to 0. According to the RDKM method, these seven data points correspond to the seven initial cluster centres in the Data-b dataset. From Fig. A3a,b, we can see that the RDKM method successfully distinguishes the potential cluster centres from the regular data points and effectively identifies the initial cluster centres.

### *Appendix A.2 The Proof of the RDVWKFC Algorithm*

We use the Lagrange multiplier method to derive the update iteration formula for the new algorithm.

Firstly, we construct the Lagrangian function $L$ through Eqs. (16) and (17) as follows:

$$L = J_{RDVWKFC} + \sum_{j=1}^{N} \lambda_j \cdot \left(1 - \sum_{i=1}^{C} u_{ij}\right) + \sum_{i=1}^{C} \beta_i \cdot \left(1 - \sum_{l=1}^{L} \omega_{il}\right) \tag{A1}$$

In order to obtain the minimum value of Eq. (16), the necessary conditions need to be satisfied, and the mathematical constraints are expressed as follows:

$$\frac{\partial L}{\partial u_{ij}} = 0, \frac{\partial L}{\partial g_{il}} = 0, \frac{\partial L}{\partial \omega_{il}} = 0, (i = 1, 2, \ldots, C; j = 1, 2, \ldots, N; l = 1, 2, \ldots, L) \tag{A2}$$

And then, we calculate the result with a partial derivative of 0 with respect to $u_{ij}$, based on Eq. (16), and get the following $(i = 1, 2, \ldots, C; j = 1, 2, \ldots, N)$:

$$\frac{\partial L}{\partial u_{ij}} = \left(\alpha_j \cdot u_{ij}^m + m \cdot u_{ij}^{m-1} \cdot \alpha_j^{'}\right) \cdot \sum_{l=1}^{L} \omega_{il} \parallel \phi(x_{il}) - \phi(g_{il}) \parallel^2 + \eta\left(lnu_{ij} + 1\right) + \lambda\left(lnu_{ij} + 1\right) + \lambda lnK - \lambda_j = 0 \tag{A3}$$

from (A3):

$$\lambda_j - \eta\left(lnu_{ij} + 1\right) - \lambda\left(lnu_{ij} + 1\right) - \lambda lnK = \left(\alpha_j \cdot u_{ij}^m + m \cdot u_{ij}^{m-1} \cdot \alpha_j^{'}\right) \cdot \sum_{l=1}^{L} \omega_{il} \parallel \phi\left(x_{jl}\right) - \phi\left(g_{il}\right) \parallel^2 \tag{A4}$$

Obviously, when $\eta + \lambda = 0$, these two different constraints cancel each other out, preventing the membership from becoming overly concentrated without forcing them to be nearly uniform. This allows membership to more flexibly reflect the clustering structure inherent in the data. The iteration is as follows:

$$u_{ij} = \left(\frac{\lambda_j - \lambda lnK}{\alpha_j\left(\frac{m}{C} + 1\right) \cdot \sum_{l=1}^{L} \omega_{il} \cdot \parallel \phi\left(x_{jl}\right) - \phi\left(g_{il}\right) \parallel^2}\right)^{\frac{1}{m}} \tag{A5}$$

Due to the constraints of membership $u_{ij}$, we get the following formula:

$$\sum_{i=1}^{C} u_{ij} = \left(\lambda_j - \lambda lnK\right)^{\frac{1}{m}} \cdot \sum_{i=1}^{C} \left(\frac{1}{\alpha_j\left(\frac{m}{C} + 1\right) \cdot \sum_{l=1}^{L} \omega_{il} \cdot \parallel \phi\left(x_{jl}\right) - \phi\left(g_{il}\right) \parallel^2}\right)^{\frac{1}{m}} = 1 \tag{A6}$$

From the above formula, it can be deduced:

$$\left(\lambda_j - \lambda lnK\right)^{\frac{1}{m}} = \frac{1}{\sum_{s=1}^{C} \left(\frac{1}{\alpha_j\left(\frac{m}{C} + 1\right) \cdot \sum_{l=1}^{L} \omega_{sl} \cdot \parallel \phi\left(x_{jl}\right) - \phi\left(g_{sl}\right) \parallel^2}\right)^{\frac{1}{m}}} \tag{A7}$$

Then, the Eq. (A7) is substituted into the Eq. (A5), that is, the updated iterative formula of membership is obtained:

$$u_{ij} = \frac{1}{\sum_{s=1}^{C}\left(\frac{\sum_{l=1}^{L}\omega_{il}\cdot \parallel \phi\left(x_{jl}\right)-\phi\left(g_{il}\right)\parallel^2}{\sum_{l=1}^{L}\omega_{sl}\cdot \parallel \phi\left(x_{jl}\right)-\phi\left(g_{sl}\right)\parallel^2}\right)^{\frac{1}{m}}} \tag{A8}$$

Similarly, calculate the partial derivative of the feature weight and set it to 0 to get the following result ($i = 1, 2, \ldots, C$; $l = 1, 2, \ldots, L$):

$$\frac{\partial L}{\partial \omega_{il}} = \sum_{j=1}^{N}\alpha_j^{'}\cdot u_{ij}^{m}\cdot \parallel \phi\left(x_{jl}\right)-\phi\left(g_{il}\right)\parallel^2 +\gamma^{-1}\cdot\left(1+ln\omega_{il}\right)-\beta_i \tag{A9}$$

After differentiating, we get the following:

$$\omega_{il} = e^{\beta_i\cdot\gamma-1}\cdot e^{-\gamma\cdot\sum_{j=1}^{N}\alpha_j'\cdot u_{ij}^{m}\cdot\parallel\phi\left(x_{jl}\right)-\phi\left(g_{il}\right)\parallel^2} \tag{A10}$$

Further, we get:

$$e^{\beta_i\gamma-1} = \frac{1}{\sum_{l=1}^{L}e^{-\gamma\sum_{j=1}^{N}\alpha_j'\cdot u_{ij}^{m}\cdot\parallel\phi\left(x_{jl}\right)-\phi\left(g_{il}\right)\parallel^2}} \tag{A11}$$

Then we get the updated iterative formula for the feature weights $\omega_{il}$:

$$\omega_{il} = \frac{e^{-\gamma\cdot\sum_{j=1}^{N}\alpha_j'\cdot u_{ij}^{m}\cdot\parallel\phi\left(x_{jl}\right)-\phi\left(g_{il}\right)\parallel^2}}{\sum_{l=1}^{L}e^{-\gamma\sum_{j=1}^{N}\alpha_j'\cdot u_{ij}^{m}cdot\parallel\phi\left(x_{jl}\right)-\phi\left(g_{il}\right)\parallel^2}} \tag{A12}$$

When we use the Gaussian kernel function to calculate the distance between a data point and the cluster centre, the mathematical expression is as follows:

$$\parallel \phi\left(x_{jl}\right)-\phi\left(g_{il}\right)\parallel^2 = 2-2\cdot K\left(x_{jl},g_{il}\right) = 2-2\cdot e^{-\frac{\left(x_{jl}-g_{il}\right)^2}{2\sigma^2}} \tag{A13}$$

according to the equation:

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1L} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nL} \end{bmatrix} \tag{A14}$$

When $i = q$, the cluster centre has the same value as the viewpoint, i.e., $g_{il} = x_{dl}$. When $i \neq q$, it is not a viewpoint, calculate the partial derivative of the cluster centre $g_{il}$, and set the resulting equation to 0, get the following formula ($i = 1, 2, \ldots, C$; $l = 1, 2, \ldots, L$):

$$\frac{\partial L}{\partial g_{il}} = \sum_{j=1}^{N}\alpha_j^{'}\cdot u_{ij}^{m}\cdot\omega_{il}\cdot(-2)\cdot e^{-\frac{\left(x_{jl}-g_{il}\right)^2}{2\sigma^2}}\cdot\frac{\left(x_{jl}-g_{il}\right)}{\sigma^2} = 0 \tag{A15}$$

Therefore, it is not difficult to obtain:

$$g_{il} = \frac{\sum_{j=1}^{N} \alpha_j^{'} \cdot u_{ij}^{m} \cdot \omega_{il} \cdot e^{-\frac{(x_{jl}-g_{il})^2}{2\sigma^2}} \cdot x_{jl}}{\sum_{j=1}^{N} \alpha_j^{'} \cdot u_{ij}^{m} \cdot \omega_{il} \cdot e^{-\frac{(x_{jl}-g_{il})^2}{2\sigma^2}}} \tag{A16}$$

Finally, by combining Formula (17) and formula (A16), the updated iterative formula of cluster centre $g_{il}$ is obtained:

$$g_{il} = \begin{cases} x_{dl}, i = q \\ \frac{\sum_{j=1}^{N} \alpha_j^{'} \cdot u_{ij}^{m} \cdot \omega_{il} \cdot e^{-\frac{(x_{jl}-g_{il})^2}{2\sigma^2}} \cdot x_{jl}}{\sum_{j=1}^{N} \alpha_j^{'} \cdot u_{ij}^{m} \cdot \omega_{il} \cdot e^{-\frac{(x_{jl}-g_{il})^2}{2\sigma^2}}}, i \neq q \end{cases} \tag{A17}$$

### *Appendix A.3 Time Complexity of the Algorithm*

Table A1 shows the time complexity of the proposed algorithm along with the algorithms used for comparison. Let denote the number of iterations and let denote the number of samples in the dataset. The C is the number of clusters, $W = \{w_{il}\}_{i=1,l=1}^{C,L}$ represents the dimension of the data features. Since the new algorithm involves an additional weight matrix to be updated, its time complexity is $O(tNl^2C)$.

**Table A1:** Time complexity of each algorithm

| Algorithm | Time complexity |
|---|---|
| FCM | $O(tNC^2l)$ |
| KFCM | $O(tNC^2l)$ |
| DPC | $O(Nl)$ |
| VFCM | $O(tNC^2l)$ |
| EWFCM | $O(tNl^2C)$ |
| FRFCM | $O(tNC^2l)$ |
| DVPFCM | $O(tNl^2C)$ |
| VWKFC | $O(tNC^2l)$ |
| RDVWKFC | $O(tNl^2C)$ |

Updating the membership values includes distance calculation as the main computational cost. The complexity of each distance calculation is $O(l^2)$ (due to the complexity of the feature map). Therefore, the total complexity of membership update is $O(Nl^2C)$.

For cluster centre updates, the complexity of computing the weighted sum of each cluster $i$ and each feature $l$ is $O(Nl)$, resulting in a total complexity of $O(NlC)$. The complexity of updating the weight parameters for each cluster $i$ and feature$l$ is $O(1)$, so the overall complexity is $O(NlC)$.

Considering that the algorithm needs to iterate, the overall time complexity of retaining the highest order term is $O(tNl^2C)$.

### *Appendix A.4 Iteration Count and Runtime of the Algorithm*

We calculated the average number of iterations of each algorithm on different data sets, and the statistical results are shown in Table A2. We rank the average number of iterations of the different algorithms on the same dataset in ascending order and fill in the brackets of the table indicating their rank.

**Table A2:** Average number of iterations the algorithm runs on different datasets

| | FCM | KFCM | V-FCM | EWFCM | FRFCM | DVPFCM | VWKFC | RDVWKFC |
|---|---|---|---|---|---|---|---|---|
| Data-a | 26(4) | 45(6) | 34(5) | 1000(7) | 6(1) | 1000(8) | 18(3) | 18(2) |
| Data-b | 32(4) | 56(5) | 159(6) | 1000(6) | 3(1) | 1000(8) | 27(3) | 20(2) |
| Iris | 17(4) | 35(6) | 18(5) | 50(7) | 11(2) | 184 | 13(3) | 10(1) |
| Breast_ cancer | 22(5) | 33(7) | 7(2) | 35(8) | 2(1) | 22(4) | 24(6) | 10(3) |
| Letter_AB | 22(6) | 43(8) | 9(4) | 32(7) | 2(1) | 16(5) | 7(3) | 6(2) |
| Spect | 5(3) | 10(6) | 6(4) | 19(8) | 2(1) | 15(7) | 7(5) | 4(2) |
| Wine | 36(6) | 61(7) | 14(2) | 85(8) | 9(1) | 32(4) | 34(5) | 19(3) |
| Stalog | 17(4) | 38(6) | 12(3) | 67(8) | 5(1) | 39(7) | 10(2) | 29(5) |
| Seeds | 19(4) | 31(6) | 26(5) | 53(7) | 8(2) | 60(8) | 12(3) | 8(1) |
| Wifi local- ization | 86(5) | 133(7) | 59(3) | 110(6) | 15(2) | 140(8) | 76(4) | 8(1) |
| The average | 28(4) | 48(6) | 34(5) | 245(8) | 6(1) | 251(7) | 23(3) | 13(2) |

Through the analysis, it is found that the average iteration count of the proposed algorithm is the lowest among all the algorithms on the data sets of Iris, Seeds, heart disease and Wifi localization. On datasets such as Data-a, Data-b, Letter_AB and Spect, its average number of iterations ranks second among all algorithms. However, it is still worth noting that on the Statlog (Landsat) dataset, the average iteration count of the proposed algorithm reaches 29, which is higher than some clustering algorithms such as FCM and V-FCM. This observation indicates that the algorithm may encounter certain limitations when dealing with recognition tasks related to satellite image datasets, which indicates that the algorithm does not show perfect performance on all types of datasets, and also provides research directions and focus for future algorithm optimization and targeted improvement.

However, from the overall experimental results, our proposed algorithm achieves convergence in fewer iterations than the other algorithms. The main reason is that we introduce more effective initial clustering centres as well as the initial viewpoints obtained from the initializing RDKM algorithm to improve the convergence speed of the algorithm. Therefore, the proposed algorithm can quickly reach the steady state in most cases, and reduce the time required for iterative calculation and computing resource overhead, which realizes the superiority of the algorithm in performance.

In Table A3, we calculate the average runtime of each algorithm on different datasets. In doing so, we rank the average running time of each algorithm in ascending order, denoted by the corresponding rank in parentheses.

**Table A3:** Running time of each algorithm on different datasets

| | FCM | KFCM | V-FCM | EWFCM | FRFCM | DVPFCM | VWKFC | RDVWKFC |
|---|---|---|---|---|---|---|---|---|
| Data-a | 0.4020(2) | 1.2500(7) | 1.6040(8) | 0.4450(3) | 0.0150(1) | 0.9050(5) | 0.9400(6) | 0.5940(4) |
| Data-b | 0.4940(3) | 1.4950(7) | 6.0590(8) | 0.4430(2) | 0.0110(1) | 0.9540(5) | 1.0700(6) | 0.6000(4) |
| Iris | 0.0130(3) | 0.0600(6) | 0.0410(5) | 0.0120(2) | 0.0050(1) | 0.0330(4) | 0.1810(8) | 0.0920(7) |
| Breast_cancer | 0.0510(2) | 0.1720(5) | 0.2230(6) | 0.0530(3) | 0.0330(1) | 0.1330(4) | 0.6730(8) | 0.2380(7) |
| Letter_AB | 0.1000(3) | 0.4650(4) | 1.2710(7) | 0.0530(2) | 0.0310(1) | 0.7390(5) | 1.3260(8) | 1.1480(6) |
| spect | 0.0080(1) | 0.0300(4) | 0.0500(6) | 0.0170(3) | 0.0100(2) | 0.0330(5) | 0.1620(8) | 0.1530(7) |
| Wine | 0.0310(4) | 0.1320(7) | 0.0380(5) | 0.0270(3) | 0.0120(1) | 0.0220(2) | 0.2850(8) | 0.0960(6) |
| Stalog | 0.1640(1) | 0.7520(4) | 3.1370(8) | 0.2670(3) | 0.2450(2) | 1.9340(5) | 2.7000(7) | 2.4020(6) |
| Seeds | 0.0220(4) | 0.0750(6) | 0.0550(5) | 0.0190(2) | 0.0060(1) | 0.0210(3) | 0.1420(8) | 0.0920(7) |
| Wifi local-ization | 1.1120(3) | 4.2580(8) | 2.9380(6) | 0.1410(2) | 0.0840(1) | 1.2420(4) | 3.8410(7) | 1.7850(5) |

By observing the experimental data, it is easy to find that the new algorithm needs fewer iterations to reach convergence, but in some cases the running time cost is longer than other algorithms. This can be attributed to the running time overhead caused by the additional computation of the weight matrix in the algorithm. However, given the significant advantages possessed by the proposed algorithm, this additional runtime computational cost is worth the investment. Although the extra computation increases the time cost, we consider the overall performance improvement and clustering result optimization for these advantages, the positive impact is enough to offset the increase in running time. This lays a solid foundation for the algorithm to be used in practice.

### *Appendix A.5 Discussion*

We verify the performance of the algorithm through a series of experiments. Initially, the algorithm was performed on a manual dataset. The experimental results show that the algorithm has good performance in the artificial data set environment. In order to accurately quantify the clustering performance of the algorithm, we calculate 7 key indicators of the algorithm and the comparison algorithm on the artificial datasets. The experimental results strongly show that the proposed algorithm outperforms other algorithms on all seven indexes.

We then experimented with the new algorithm on eight UCI datasets. The experimental data show that the proposed algorithm has superior performance on all evaluation indexes of each UCI dataset compared with other algorithms. The test results show that it can show more outstanding advantages in many aspects, including precision, stability and anti-interference.

The time complexity of each of the other clustering algorithms is given, along with the average number of iterations and running times on different datasets. It is easy to see that in most cases, the average number of iterations of this algorithm is less than that of other comparison algorithms, which shows that the iterative effect of our designed algorithm is efficient and convergent.

Finally, we use the high dimensional state data set to carry out the experiment. The results show that RDVWKFC algorithm has good performance. Compared with other algorithms, this algorithm mainly shows its advantages in the following aspects:

- **Innovative initialization strategy:** This method is based on Relative Density Knowledge Extraction (RDKM), which can accurately identify relatively high-density regions and thus determine the initial viewpoint. This method overcomes the sensitivity of traditional clustering algorithms to initialization and reduces the risk of result bias. It significantly improves the initial quality of clustering and provides clear advantages over previous methods.

- **Effective anti-interference mechanism:** We introduce the adaptive weight mechanism and construct the Information Granularity Weight Model (IGWM). The model dynamically adjusts the influence of data points according to the importance of data points to reduce the bad interference of noise and outliers on clustering results. It improves the adaptability of the algorithm to a complex data environment and improves the reliability of the clustering results.
- **Outstanding performance:** We design an adaptive optimization algorithm (AOA) to significantly improve the performance of different data characteristics and dimensions. The algorithm can dynamically adjust the weights according to the real-time distribution of data and can run stably in a high-dimensional or feature differentiated data environment, which overcomes the limitations of traditional algorithms. By integrating a variety of innovative elements, the proposed algorithm is superior to the comparison algorithms on multiple data sets and related key evaluation indicators, and realizes the goals with fewer iterations, faster convergence speed and higher computational efficiency. The above performance indicators fully prove its effectiveness and superiority in practical applications.

### *Appendix A.6 Limitations*

Our findings confirm that the performance of the algorithm does depend heavily on a few key parameters. These parameters include fuzzy coefficient, regularization parameter and the width of the Gaussian kernel. The process of tuning these parameters can be quite subjective and may not always produce optimal results.

### *Appendix A.7 Comparison of Different Clustering Algorithms*

The comparison in Table A4 ensures a balance between classical foundational work and recent research, effectively strengthening the context of our study.

**Table A4:** Comparison of different clustering algorithms

| | |
|---|---|
| FCM | FCM algorithm is the most classical algorithm among many fuzzy clustering algorithms, and also the most common the algorithm used. |
| KFCM | KFCM is an improved kernel function based fuzzy C-means clustering algorithm that maps data from original space to high-dimensional feature space by kernel function. |
| V-FCM | V-FCM algorithm is short for fuzzy C-means clustering algorithm based on viewpoint and introduces the concept of viewpoint in the process of clustering for the first time |
| EWFCM | The EWFCM algorithm constructs a dimension matrix and adds an additional attribute weight entropy regularization term. The selection of features is realized by providing weights for features. |
| FRFCM | FRFCM is an improved FCM algorithm based on morphological reconstruction and member filtering, which is faster and more robust than FCM. |
| DVPFCM | DVPFCM proposes a viewpoint-driven possibility fuzzy clustering algorithm and initializes cluster centres by searching density peaks |
| VWKFC | VWKFC is a density-based weighted kernel fuzzy clustering algorithm. Based on the kernel, the hypersphere density initialization algorithm is proposed and the concept of weighted information particles is established. |
| FCAG | FCAG is a fast clustering model based on anchor point guidance. The proposed model avoids trivial solutions and has no additional regularization terms. |

(Continued)

| Table A4 (continued) | |
|---|---|
| SR-PCM-HDP | SR-PCM-HDP is a self-adjusting possibility C-means clustering algorithm which uses high density points to simplify clustering and improve clustering efficiency while determining the number. |
| RDVWKFC | RDVWKFC proposed a weighted kernel fuzzy clustering algorithm based on relative density view, which can capture the intrinsic structure of data more accurately and improve the initial quality of clustering. |

### *Appendix A.8 Abbreviation of Various Clustering Algorithms*

Table A5 shows the abbreviations of clustering algorithms mentioned in the manuscript.

**Table A5:** Abbreviation of various clustering algorithms

| Abbreviation | Full name |
|---|---|
| DPC | Density peak clustering |
| FCM | Fuzzy C-Means algorithm |
| PCM | Possibilistic C-Means algorithm |
| KFCM | Kernelized Fuzzy C-Means |
| FCAG | Fast clustering with anchor guidance |
| SR-PCM-HDP | self-regulating possibilistic C-Means (PCM) with high-density points |
| FWPCM | Feature-Weighted Possibilistic C-Means algorithm |
| V-FCM | Viewpoint-based Fuzzy C-Means (V-FCM) algorithm |
| EWFCM | Entropy regularized weighted fuzzy C-Means algorithm |
| FRFCM | A feature-reduction FCM |
| DVPFCM | Density view-induced probabilistic fuzzy C-Means |
| VWKFC | Viewpoint-based weighted kernel fuzzy clustering |
| RDVWKFC | Relative-density-viewpoint-based weighted kernel fuzzy clustering algorithm |

## References

1. Li XL, Zhang H, Wang R, Nie FP. Multi-view clustering: a scalable and parameter-free bipartite graph fusion method. IEEE Trans Pattern Anal Mach Intell. 2022;44(1):330–44. doi:10.1109/tpami.2020.3011148.
2. Hu XH, Tang YM, Pedrycz W, Jiang JC, Jiang YC. Knowledge-driven possibilistic clustering with automatic cluster elimination. Comput Mater Contin. 2024;80(3):4917–45. doi:10.32604/cmc.2024.054775.
3. Zhang J, Fan R, Tao H, Jiang J, Hou C. Constrained clustering with weak label prior. Front Comput Sci. 2024;18(3):183338. doi:10.1007/s11704-023-3355-7.
4. Hussain I, Sinaga KP, Yang MS. Unsupervised multi-view fuzzy C-means clustering algorithm. Electronics. 2023;12(21):4467. doi:10.3390/electronics12214467.
5. Rodriguez A, Laio A. Clustering by fast search and find of density peaks. Science. 2014;344(6191):1492–6. doi:10.1126/science.1242072.
6. Xu K, Pedrycz W, Li Z, Nie W. Constructing a virtual space for enhancing the classification performance of fuzzy clustering. IEEE Trans Fuzzy Syst. 2019;27(9):1779–92. doi:10.1109/tfuzz.2018.2889020.
7. Pedrycz W, Amato A, Lecce VD, Piuri V. Fuzzy clustering with partial supervision in organization and classification of digital images. IEEE Trans Fuzzy Syst. 2008;16(4):1008–26. doi:10.1109/tfuzz.2008.917287.

8. Krishnapuram R, Keller JM. A possibilistic approach to clustering. IEEE Trans Fuzzy Syst. 1993;1(2):98–110. doi:10.1109/91.227387.
9. Pal NR, Pal K, Keller JM, Bezdek JC. A possibilistic fuzzy c-means clustering algorithm. IEEE Trans Fuzzy Syst. 2005;13(4):517–30. doi:10.1109/tfuzz.2004.840099.
10. Hu J, Wu M, Chen L, Zhou K, Zhang P, Pedrycz W. Weighted kernel fuzzy C-means-based broad learning model for time-series prediction of carbon efficiency in iron ore sintering process. IEEE Trans Cybern. 2022;52(6):4751–63. doi:10.1109/tcyb.2020.3035800.
11. Graves D, Pedrycz W. Kernel-based fuzzy clustering and fuzzy clustering: a comparative experimental study. Fuzzy Sets Syst. 2010;161(4):522–43. doi:10.1016/j.fss.2009.10.021.
12. Huang H, Chuang Y, Chen C. Multiple kernel fuzzy clustering. IEEE Trans Fuzzy Syst. 2012;20(1):120–34. doi:10.1109/tfuzz.2011.2170175.
13. Yang MS, Benjamin JBM. Feature-weighted possibilistic C-means clustering with a feature-reduction framework. IEEE Trans Fuzzy Syst. 2021;29(5):1093–106. doi:10.1109/tfuzz.2020.2968879.
14. Hussain I, Nataliani Y, Ali M, Hussain A, Mujlid HM, Almaliki FA, et al. Weighted multi-view K-means clustering with L2 regularization. Symmetry. 2024;16(12):1646. doi:10.3390/sym16121646.
15. Pedrycz W, Loia V, Senatore S. Fuzzy clustering with viewpoints. IEEE Trans Fuzzy Syst. 2010;18(2):274–84. doi:10.1109/tfuzz.2010.2040479.
16. Tang YM, Hu XH, Pedrycz W, Song XC. Possibilistic fuzzy clustering with high-density viewpoint. Neurocomputing. 2019;329(2):407–23. doi:10.1016/j.neucom.2018.11.007.
17. Tang YM, Pan ZF, Pedrycz W, Ren FJ, Song XC. Viewpoint-based kernel fuzzy clustering with weight information granules. IEEE Trans Emerg Top Comput Intell. 2023;7(2):342–56. doi:10.1109/tetci.2022.3201620.
18. Zhou J, Chen L, Chen CLP. Fuzzy clustering with the entropy of attribute weights. Neurocomputing. 2016;19(8):125–34. doi:10.1016/j.neucom.2015.09.127.
19. Oskouei AG, Samadi N, Tanha J, Bouyer A, Arasteh B. Viewpoint-based collaborative feature-weighted multi-view intuitionistic fuzzy clustering using neighborhood information. Neurocomputing. 2025;617(5):128884. doi:10.1016/j.neucom.2024.128884.
20. Hashemzadeh M, Oskouei AG, Farajzadeh N. New fuzzy C-means clustering method based on feature-weight and cluster-weight learning. Appl Soft Comput. 2019;78:324–45. doi:10.1016/j.asoc.2019.02.038.
21. Nie F, Xue J, Yu W, Li X. Fast clustering with anchor guidance. IEEE Trans Pattern Anal Mach Intell. 2024;46(4):1898–912. doi:10.1109/tpami.2023.3318603.
22. Hu XH, Jiang YC, Pedrycz W, Deng ZH, Gao JW, Tang YM. Automated cluster elimination guided by high-density points. IEEE Trans Cybern. 2025;55(4):1717–30. doi:10.1109/tcyb.2025.3537108.
23. Strehl A, Ghosh J. Cluster ensembles—a knowledge reuse frame work for combining multiple partitions. J Mach Learn Res. 2002;3(1):583–617.
24. Calinski T, Harabasz J. A dendrite method for cluster analysis. Commun Stat. 1972;3(1):1–27.
25. Campello RJGB. A fuzzy extension of the rand index and other related indexes for clustering and classification assessment. Pattern Recognit Lett. 2007;28(7):833–41. doi:10.1016/j.patrec.2006.11.010.
26. Hubert L, Arabie P. Comparing partitions. J Classif. 1985;2(1):193–218.
27. Xie XL, Beni G. A validity measure for fuzzy clustering. IEEE Trans Pattern Anal Mach Intell. 1991;13(8):841–7. doi:10.1109/34.85677.
28. Tang YM, Chen JJ, Pedrycz W, Ren FJ, Zhang L. Universal quintuple implicational algorithm: a unified granular computing framework. IEEE Trans Emerg Top Comput Intell. 2024;8(1):1044–56. doi:10.1109/tetci.2023.3327719.