Expert Systems With Applications

New semi-supervised fuzzy C-means clustering with asymmetric deviation constraints and fast algorithm

--Manuscript Draft--

Manuscript Number:	ESWA-D-25-14495R1
Article Type:	Full length article
Section/Category:	1.3Computer vision and pattern recognition
Keywords:	Fuzzy clustering, asymmetric deviation constraint, affinity filtering, membership scaling, accelerated algorithm
Abstract:	Semi-supervised clustering leverages prior information to improve algorithm performance and is widely valued by researchers. This paper analyzes the traditional semi-supervised fuzzy C-means (SFCM) objective function, noting that as a labeled sample's membership degree aligns with its prior information, the impact of this information on the deviation constraint weakens. This reduces its supervisory effect on optimizing the membership partition matrix, especially with a large regularization factor. To overcome this, we propose a novel semi-supervised fuzzy C-means method based on an asymmetric deviation constraint and develop a two-level alternating iterative optimization algorithm, supported by theoretical convergence analysis using Zangwill's theorem and the bordered Hessian matrix. To address the slow convergence and high computational cost typical of semi-supervised fuzzy clustering, we further enhance the algorithm with affinity filtering and a membership scaling scheme for improved efficiency. Experimental results demonstrate that our methods significantly outperform existing state-of-the-art techniques, advancing semi-supervised fuzzy C-means clustering.

Responses to the Editor's and Reviewers' Comments on Paper ESWA-D-25-14495

Dear Editor and Reviewers:

Thank you for your letter and comments concerning our manuscript entitled "New Semi-Supervised Fuzzy C-Means Clustering with Asymmetric Deviation Constraints and its Fast Algorithm" (ID: ESWA-D-25-14495). We are very grateful to the editor and all reviewers. Those comments are very helpful for revising and improving our manuscript.

We have studied comments carefully and have made corrections that we hope they meet with your approval. The revised part is highlighted in the manuscript. We hope the revised manuscript meets the expectations of the Editor and reviewers. Below, we provide detailed responses to each comment.

The main corrections in the revised manuscript and the response to comments are as follows:

To Reviewer #1:

We appreciate your recognition and constructive feedback, which helped improve our manuscript's rigor and clarity. We have carefully addressed all your comments in this revision.

To facilitate your further review, new or revised content is highlighted, and the manuscript has been carefully edited for clarity and readability.

Below, we provide detailed responses to your comments and describe the corresponding revisions made in the relevant sections.

1. Comment: T ***in Table 1 and*** T ***in Equation (4) should use different symbols.***

Response:

$$\min J_{FE}(U,V)=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}d_A^2(x_i,v_j)+T\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}\log u_{ij} \tag{1}$$

Appreciation is extended for the careful observation. It is indeed correct that the symbol T was used inconsistently in the original manuscript. In Table 1, T denotes the number of iterations, whereas in **Eq. (1)**, T. serves as a regularization factor that controls the influence of fuzzy entropy on the objective function. This inconsistency may have caused confusion.

In response, a thorough review of all mathematical symbols and notations throughout the manuscript was conducted to eliminate similar ambiguities and ensure consistency. However, after careful reconsideration and in light of suggestions from other reviewers, the corresponding section involving the fuzzy entropy clustering has been removed.

The main reason for this decision lies in the structural foundation of the paper. The study is built upon the classical FCM objective function, which serves as the baseline for the development of various semi-supervised fuzzy clustering methods through the incorporation of regularization terms. Since the primary contribution of this work is the reconstruction of the semi-supervised regularization term in SFCM to improve the utilization efficiency of prior information, the fuzzy entropy clustering falls outside the scope of this study. Therefore, its inclusion in the related work section was deemed misaligned with the paper's core focus.

To maintain logical consistency and avoid redundancy, the fuzzy entropy clustering content was removed. The efforts to eliminate symbol ambiguity have nonetheless been preserved in other parts of the paper. Sincere thanks are expressed once again for the attentive and thoughtful reading.

2. Comment: ***In section 2.2, the limitation in SMUC should be corrected.***

Response:

Thank you for this important observation. The original description of the limitation in SMUC contained a conceptual inaccuracy and lacked sufficient rigor. Specifically, the statement: "This semi-supervised fuzzy clustering method has a major limitation: $\sum_{j=1}^{c}\overline{f}_{ij}$ cannot equal 1, restricting the SMUC algorithm's widespread application" was inappropriate and misleading.

After careful re-examination of the SMUC algorithm, a more accurate understanding has been incorporated. SMUC avoids the need to manually select the fuzzy weighting exponent m by integrating Mahalanobis distance metric learning and maximum entropy regularization. However, it presents several notable drawbacks:

● It requires frequent computation and updating of the covariance matrix and its inverse, resulting in higher computational complexity than traditional SFCM;

● When dealing with high-dimensional data or small sample sizes, the estimation of the covariance matrix may become unstable, potentially leading to singular matrices.

This correction has been taken into account. Furthermore, following in-depth consideration and in accordance with other reviewers' suggestions, the descriptions of semi-supervised fuzzy entropy clustering methods, including FE, ESFCM, SMUC, and SFCM-EP, have been removed from the revised manuscript to streamline the related work section.

The rationale behind this decision is as follows: Although these semi-supervised fuzzy entropy clustering methods also use symmetric deviation $(u_{ij} - b_i \overline{f}_{ij})^2$ in their semi-supervised regularization terms—which can weaken the constraint on the distance measure $d^2(x_i, v_j)$ when u_{ij} approaches $\overline{f}_{ij}$ and reduce the guiding effect of the prior membership—they fall outside the core scope of this paper. The focus of this study lies in reconstructing the semi-supervised regularization term within the SFCM framework to enhance the utilization efficiency of prior information. As the entropy function is not central to this objective, it has been excluded to maintain the structural and logical coherence of the manuscript. Future studies may consider extending the proposed methodology to the semi-supervised fuzzy entropy clustering domain.

Sincere apologies are expressed for the oversight in the original text, and thanks again for the critical and constructive feedback.

3.1. Comment: ***The work neither clarifies the causes of the problems to be solved nor elaborates on the theoretical mechanism by which the proposed method for the issues to be addressed.***

Response:

Thank you for this insightful and important comment. In response, a new subsection entitled "Motivation" has been added in Section 3.1 to provide a detailed explanation of the underlying causes of the problems addressed in this work. Additionally, the theoretical mechanism of the proposed method has been elaborated upon through a newly added block diagram illustrating the clustering optimization model.

The limitations in the objective functions of existing methods, such as SFCM and SSFCM, have been systematically analyzed. These methods employ a symmetric deviation structure in the semi-supervised regularization term, commonly formulated as $(u_{ij} - b_i \overline{f}_{ij})^2$. However, when the membership degree u_{ij} approaches its prior value $\overline{f}_{ij}$, the magnitude of this regularization term tends toward zero. Consequently, the constraint imposed on the distance measure $d^2(x_i, v_j)$—which depends jointly on the data sample and clustering center—becomes negligible. This reduction weakens the ability of prior information to guide the optimization of clustering centers, thereby compromising clustering accuracy.

The ICBF-MFSFCM algorithm addresses part of this problem by introducing a membership fusion mechanism, which retains the scale of the unsupervised term even when $u_{ij} \approx \overline{f}_{ij}$. Nevertheless, this approach fundamentally relies on the unsupervised component to preserve clustering effectiveness, which deviates from the essence of semi-supervised learning—where the primary goal is to leverage prior information rather than fallback on unsupervised clustering when agreement exists between current and prior memberships.

To address this, the proposed method reconstructs the semi-supervised regularization term by introducing an asymmetric deviation constraint of the form: $(u_{ij}(1+b_i)-3b_i\overline{f}_{ij})^2$. This new formulation ensures that even when $u_{ij} \approx \overline{f}_{ij}$, the regularization term maintains sufficient influence. It thereby preserves the guiding role of prior information throughout the optimization process, significantly enhancing the utilization efficiency of the available labeled information.

The revised manuscript now explicitly presents both the problem causes and the theoretical mechanism in **Sections 3.1** and **3.2**, supported by the added block framework diagram to improve clarity and completeness of the explanation.

Added content

The introduction of SFCM (Pedrycz and Waletzky, 1997) established a theoretical basis for semi-supervised clustering, leading to many related optimization algorithms. However, the SFCM objective function's semi-supervised regularization term is typically a product involving a distance metric. Under the symmetric deviation constraint $(u_{ij}-b_i\overline{f}_{ij})^2$, when labeled samples' membership degrees approach their prior values, the constraint's influence on the distance metric $d^2(x_i, v_j)$ weakens. Since the distance metric depends on distances between samples and clustering centers, the guidance from prior information diminishes, reducing clustering accuracy. Although the ICBF-MFSFCM (Zhang et al., 2024) improves this via membership fusion—integrating $(u_{ij}-\overline{f}_{ij})^2+u_{ij}^2$ and $d^2(x_i, v_j)$, it still underuses prior information when memberships align closely with priors, relying mainly on unsupervised clustering. Therefore, a new semi-supervised regularization term is needed.

(See Section 3.1 "Motivation")

Building on the ICBF-MFSFCM algorithm (Zhang et al., 2024), we introduce an asymmetric constraint to integrate $(2u_{ij}-3\overline{f}_{ij})^2$ and $d^2(x_i,v_j)$ in semi-supervised clustering, enhancing performance with a new optimization model.

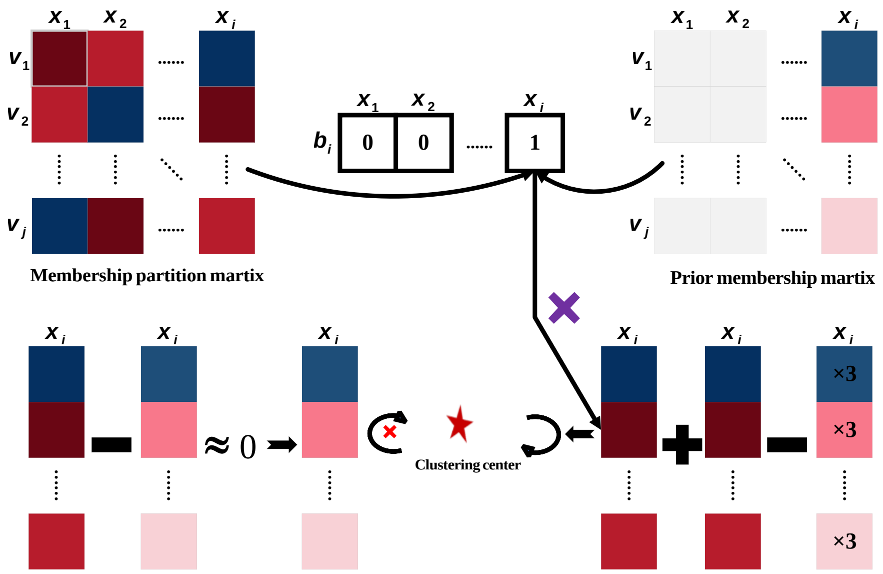


Fig 1. Block diagram of the new clustering optimization model

Fig. 1 illustrates that in the membership partition matrix, sample x_i belongs to the j-th cluster with varying membership u_{ij}. When the binary vector $b_i=0$, the i-th column of the prior membership matrix, lacks prior information, it indicates the sample is unlabeled. Conversely, $b_i=1$ means the i-th sample is labeled with prior membership information. If the current membership degree u_{ij} of sample x_i closely matches its prior membership $\overline{f}_{ij}$, the influence of prior information on guiding cluster centers weakens, potentially reducing clustering performance. To address this, an asymmetric deviation term $(2u_{ij}-3\overline{f}_{ij})^2$ is introduced based on the binary vector b_i. For unlabeled sample x_i, the semi-supervised regularization term reduces to u_{ij}, matching the fuzzy clustering term's magnitude and aligning with unsupervised clustering. For labeled samples with $u_{ij}\approx\overline{f}_{ij}$, the regularization becomes the asymmetric deviation term $(2u_{ij}-3\overline{f}_{ij})^2$, ensuring prior information $\overline{f}_{ij}$ effectively guides clustering. This design preserves the core principle of semi-supervised clustering.

(See the first and second paragraphs in Section 3.2 "Optimization modeling")

3.2. Comment: ***In section 3.1, the author did not provide sufficient elaboration on the research motivation. The author did not discuss the reasons why the terms and in the SFCM algorithm***

lack interpretability, nor did they clarify the asymmetric deviation constraint and how the ADSFCM algorithm enhances interpretability. The author needs to conduct a comprehensive analysis and study of the situational conditions and generation mechanism of the motivation. In principle (rather than just based on experimental results), the author should explain how the proposed method addresses the interpretability issue in this specific situational condition.

Response:

Appreciation is sincerely expressed for this valuable suggestion, which has significantly contributed to improving the clarity and logical coherence of the manuscript. It is acknowledged that the original submission lacked a sufficient explanation of the research motivation, which resulted in an abrupt transition to theoretical formulation and experimental validation without adequately guiding the reader through the reasoning behind the proposed work. This oversight may have weakened the interpretability and logical rigor of the paper, and for that, sincere apologies are extended.

To address this issue, substantial revisions have been made in **Section 3.1**. A new subsection titled "Motivation" has been added, offering a detailed and principled explanation of the motivations behind the proposed method. Specifically, the revised content first analyzes the limitations of the symmetric deviation $(u_{ij} - b_i \overline{f}_{ij})^2$ employed in traditional semi-supervised regularization terms, as used in methods like SFCM. When u_{ij} approaches $\overline{f}_{ij}$, the semi-supervised regularization term diminishes rapidly, thereby weakening the influence of prior information on the optimization of the clustering objective. This behavior reduces the interpretability of the model, as the role of supervision becomes unclear in such conditions.

To overcome this issue, the theoretical mechanism behind the asymmetric deviation constraint introduced in the ADSFCM algorithm has been explicitly formulated. The new constraint term, $(u_{ij}(1+b_i) - 3b_i \overline{f}_{ij})^2$, ensures that the contribution of prior information remains active even when $u_{ij} \approx \overline{f}_{ij}$, thereby preserving the interpretability of the semi-supervised regularization process. Unlike previous approaches, this method does not rely solely on unsupervised terms to stabilize the clustering result when prior and current memberships align. Instead, it provides a theoretical structure where prior information continues to guide the clustering optimization, aligning the algorithm more closely with the essence of semi-supervised learning.

Additionally, a new clustering optimization model block diagram has been introduced to help illustrate the motivation and theoretical foundation of the proposed approach. This visual explanation further supports the textual analysis and clarifies the generation mechanism of the motivation within its situational context.

These changes collectively respond to the concern that the original manuscript lacked theoretical justification for the proposed method's interpretability. The revised content now offers a more comprehensive, principle-based explanation of how ADSFCM addresses the identified limitations in interpretability. Gratitude is once again expressed for this insightful feedback, which helped resolve a critical logical gap in the manuscript.

Added content

The introduction of SFCM (Pedrycz and Waletzky, 1997) established a theoretical basis for semi-supervised clustering, leading to many related optimization algorithms. However, the SFCM objective function's semi-supervised regularization term is typically a product involving a distance metric. Under the symmetric deviation constraint $(u_{ij} - b_i \overline{f}_{ij})^2$, when labeled samples' membership degrees approach their prior values, the constraint's influence on the distance metric $d^2(x_i, v_j)$ weakens. Since the distance metric depends on distances between samples and clustering centers, the guidance from prior information diminishes, reducing clustering accuracy. Although the ICBF-MFSFCM (Zhang et al., 2024) improves this via membership fusion—integrating $(u_{ij} - \overline{f}_{ij})^2 + u_{ij}^2$ and $d^2(x_i, v_j)$, it still underuses prior information when memberships align closely with priors, relying mainly on unsupervised clustering. Therefore, a new semi-supervised regularization term is needed.

(See Section 3.1 "Motivation")

Building on the ICBF-MFSFCM algorithm (Zhang et al., 2024), we introduce an asymmetric constraint to integrate $(2u_{ij} - 3\overline{f}_{ij})^2$ and $d^2(x_i, v_j)$ in semi-supervised clustering, enhancing performance with a new optimization model.

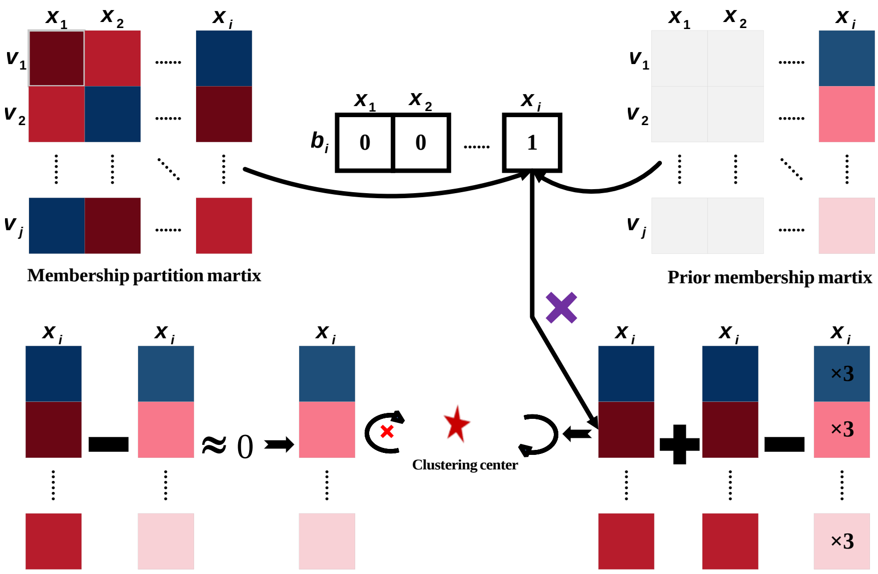


Fig 1. Block diagram of the new clustering optimization model

Fig. 1 illustrates that in the membership partition matrix, sample x_i belongs to the j-th cluster with varying membership u_{ij}. When the binary vector $b_i = 0$, the i-th column of the prior membership matrix, lacks prior information, it indicates the sample is unlabeled. Conversely, $b_i = 1$ means the i-th sample is labeled with prior membership information. If the current membership degree u_{ij} of sample x_i closely matches its prior membership $\bar{f}_{ij}$, the influence of prior information on guiding cluster centers weakens, potentially reducing clustering performance. To address this, an asymmetric deviation term $(2u_{ij} - 3\bar{f}_{ij})^2$ is introduced based on the binary vector b_i. For unlabeled sample x_i, the semi-supervised regularization term reduces to u_{ij}, matching the fuzzy clustering term's magnitude and aligning with unsupervised clustering. For labeled samples with $u_{ij} \approx \bar{f}_{ij}$, the regularization becomes the asymmetric deviation term $(2u_{ij} - 3\bar{f}_{ij})^2$, ensuring prior information $\bar{f}_{ij}$ effectively guides clustering. This design preserves the core principle of semi-supervised clustering.

(See the first and second paragraphs in Section 3.2 “Optimization modeling”)

3.3. Comment: ***In section 3.4, the author did not explain the reasons for the negative impact of non-affinity centers on the optimization of the membership matrix. The author needs to conduct a comprehensive analysis and study of the situational conditions and generation mechanism of the motivation. In principle (rather than just based on experimental results), the author should explain how the proposed method addresses the interpretability issue in this specific situational condition.***

Response:

We sincerely thank you for your constructive suggestion. We acknowledge that the original version of the manuscript did not provide a sufficient theoretical explanation for the negative impact of non-affinity centers on the optimization of the membership matrix. In response, we have conducted a comprehensive analysis of the underlying conditions and generation mechanisms that motivate the proposed optimization strategy. Let $\mathcal{P}_i$ denote the set of non-affinity centers of x_i, where $i = 1, 2, \cdots, n$. For the proposed ADSFCM algorithm, when $j \in \mathcal{P}_i$, several convergence and efficiency issues arise in the optimization process:

- As shown in **Eq. (2)**, during the membership update phase, if $d^2_{j\in\mathcal{P}_i}(x_i, v_j) \gg d^2_{j\notin\mathcal{P}_i}(x_i, v_j)$, the membership of the data sample to the non-affinity clustering center should ideally be zero. However, under the constraint $\sum_{j=1}^{c} u_{ij} = 1$, a small residual value τ remains, compressing the membership of the data sample to its affinity center to $1-\tau$, and thereby weakening the discriminative capacity of the membership values.

$$u_{ij} = \frac{1}{\beta + \alpha(1+b_i)^2} \cdot \frac{\beta + \alpha(1+b_i)^2 - 3\alpha b_i(1+b_i)\sum_{k=1}^{c} \bar{f}_{ik}}{\sum_{k=1}^{c} d^2(x_i, v_j) / d^2(x_i, v_k)} + \frac{3\alpha b_i(1+b_i)}{\beta + \alpha(1+b_i)^2} \bar{f}_{ij} \tag{2}$$

- As shown in **Eq. (3)**, during the clustering center update, the presence of a residual value τ introduces an deviation term $\tau(x_i - v_j)$, even when $d^2_{j\in\mathcal{P}_i}(x_i, v_j) \gg d^2_{j\notin\mathcal{P}_i}(x_i, v_j)$. Since the convergence condition is defined by $\left\| V^{(t+1)} - V^{(t)} \right\| < \varepsilon$, this persistent deviation leads to slower convergence of clustering centers and decreases overall iteration efficiency.

$$v_j = \frac{\beta\sum_{i=1}^{n} u_{ij}^2 x_i + \alpha\sum_{i=1}^{n} (u_{ij}(1+b_i) - 3b_i \bar{f}_{ij})^2 \cdot x_i}{\beta\sum_{i=1}^{n} u_{ij}^2 + \alpha\sum_{i=1}^{n} (u_{ij}(1+b_i) - 3b_i \bar{f}_{ij})^2} \tag{3}$$

- As shown in **Eq. (4)**, the distance $d^2_{j\in\mathcal{P}_i}(x_i, v_j)$ is theoretically expected to be large. However, due to the constraint $\sum_{j=1}^{c} u_{ij} = 1$, the membership degree u_{ij} is forcibly assigned a small positive value close to zero. As a result, a large number of ineffective terms $\sum_{j\in\mathcal{P}_i} u_{ij}^2 d^2(x_i, v_j)$ accumulate in the objective function. This accumulation gradually degrades the convergence behavior from exponential to linear, significantly reducing the overall convergence speed.

$$\min J(U,V)=\beta\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{2}d^{2}(x_i,v_j)+\alpha\sum_{j=1}^{c}\sum_{i=1}^{n}(u_{ij}(1+b_i)-3b_i\bar{f}_{ij})^{2}d^{2}(x_i,v_j) \tag{4}$$

● In the mid-to-late stages of iteration, although most membership degrees have stabilized, distances between samples and non-affinity clustering centers $d_{j\in\mathcal{P}_i}^{2}(x_i,v_j)$ still require frequent computation. This results in considerable computational overhead and increased resource consumption. Therefore, the effective identification and optimization of non-affinity clustering centers are critical for accelerating convergence and improving the overall efficiency of the proposed algorithm.

We have revised **Section 3.4** to incorporate these theoretical insights highlight them, highlighting the practical need to identify and exclude non-affinity centers during the optimization process. This provides a clearer motivation for the acceleration mechanism and reinforces its theoretical foundation.

It is worth noting that the proposed fast framework is a general-purpose acceleration strategy applicable to various semi-supervised fuzzy clustering algorithms. In this paper, we demonstrate its effectiveness by integrating it with the proposed ADSFCM algorithm. The interpretability of ADSFCM itself has been further elaborated in the preceding sections of the manuscript.

Added content

For the proposed ADSFCM algorithm, when $j\in\mathcal{P}_i$, convergence and efficiency issues arise during optimization: According to **Eq. (2)**, if $d_{j\in\mathcal{P}_i}^{2}(x_i,v_j)\gg d_{j\notin\mathcal{P}_i}^{2}(x_i,v_j)$, a data sample's membership to a non-affinity clustering center should be zero. However, with the constraint $\sum_{j=1}^{c}u_{ij}=1$, a small residual value τ remains, compressing the membership to its affinity center as $1-\tau$ and weakening membership discriminability. In **Eq. (3)**, the residual τ introduces a deviation term $\tau(x_i-v_j)$ during clustering center updates, even when $d_{j\in\mathcal{P}_i}^{2}(x_i,v_j)\gg d_{j\notin\mathcal{P}_i}^{2}(x_i,v_j)$. Since the convergence condition is defined by $\left\|V^{(t+1)}-V^{(t)}\right\|<\varepsilon$, this deviation slows convergence and reduces iteration efficiency. Per **Eq. (4)**, the distance $d_{j\in\mathcal{P}_i}^{2}(x_i,v_j)$ should be large, but $\sum_{j=1}^{c}u_{ij}=1$ forces the membership degree u_{ij} to a small positive value near zero, causing many ineffective terms $\sum_{j\in\mathcal{P}_i}u_{ij}^{2}d^{2}(x_i,v_j)$ to accumulate in the objective function. This accumulation degrades convergence from exponential to linear, significantly slowing the process. In later iterations, although most memberships stabilize, distances between samples and non-affinity centers $d_{j\in\mathcal{P}_i}^{2}(x_i,v_j)$ still require frequent calculation, increasing computational overhead. Thus, optimizing non-affinity clustering centers is crucial to speed convergence and improve efficiency.

(See the third paragraph in Section 3.5 "Accelerated semi-supervised fuzzy clustering with membership scaling")

---

4. Comment: ***In section 4, the description provided by the author here should incorporate statistical test analysis to verify whether the differences in performance metrics between the proposed algorithm and other algorithms are statistically significant.***

Response:

Thank you for your valuable suggestion. We fully agree that incorporating statistical analysis is essential to validate whether the performance differences among algorithms are statistically significant. In response to your comment, we have added a new subsection titled "Statistical test analysis" in **Section 4.3** of the revised manuscript.

In this subsection, we conducted a rigorous statistical evaluation based on the experimental results obtained from nine numerical datasets and twelve image datasets. Specifically, we employed the Friedman test to assess whether the differences in performance metrics among the seven compared algorithms, including the proposed ADSFCM, are statistically significant. By calculating the corrected F-distribution and P-values, we found that for all evaluation metrics, the null hypothesis — stating that there is no significant difference in algorithm performance — can be rejected.

Furthermore, we performed post-hoc analysis to determine which pairs of algorithms exhibit statistically significant differences. To provide a clear and intuitive visualization of the results, we also plotted the critical difference (CD) diagrams.

These statistical test results provide strong evidence that the performance differences between ADSFCM and the other algorithms are statistically significant, thereby reinforcing the effectiveness and reliability of the proposed method.

Added content

To evaluate whether performance metrics differences between the ADSFCM algorithm and others are statistically significant, the Friedman test (Demšar, 2006) is applied to nine datasets and twelve images; results appear in **Table 1**.

For clarity, the variables in the statistical test are defined as follows: k is the total number of algorithms compared; N is the total number of datasets evaluated; r_i^j is the rank of the j-th algorithm on the i-th dataset, with 1 as best and k as worst; R_j is the average rank of the j-th algorithm across all datasets; $df_1 = k-1$ and $df_2 = (k-1)(N-1)$ are the degrees of freedom; χ_F^2 and F_F are the test statistics based on the chi-square distribution and the corrected F-distribution, respectively, calculated as follows:

$$\chi_F^2 = \frac{12N}{k(k+1)}\left(\sum_{i=1}^{k} R_j^2 - \frac{k(k+1)^2}{4}\right) \tag{5}$$

$$F_F = \frac{(N-1)\chi_F^2}{N(k-1)-\chi_F^2} \tag{6}$$

The P-value measures the result randomness and tests sample differences. With a significance level $a = 0.05$. $P \le a$ leads to rejecting the null hypothesis of no algorithm performance difference. It is calculated by **Eq. (7)**.

$$P = F_a(df_1, df_2) \tag{7}$$

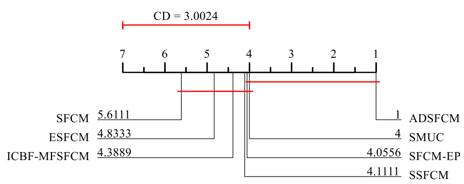
Table 1 F_F and P value for Friedman's test

	ACC1	PE	RE	ARI	NMI	ACC2	Jaccard	mIoU
F_F	6.6002	5.8802	6.6057	6.1972	6.6505	42.9514	40.1664	60.1587
P value	0.000039	0.000117	0.000039	0.000072	0.000036	6.37×10^{-18}	2.11×10^{-17}	1.07×10^{-20}

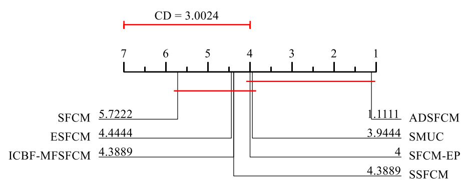
*ACC1 and ACC2 denote clustering and segmentation accuracy, respectively.

Table 1 shows all P-values meet $P \le 0.05$, rejecting the hypothesis and allowing post-hoc tests to identify significant algorithm differences.

If the difference between two algorithms' average ranks is within the critical difference (CD), their performance difference is not statistically significant. If the difference equals or exceeds the CD, the performance difference is significant (Arbelaez et al., 2010). The CD is defined as $CD = q_a\sqrt{\frac{k(k+1)}{6N}}$, where q_a depends on the significance level a. Based on this, CD diagrams for various metrics of seven algorithms across nine numerical datasets and twelve images are shown in **Fig. 2**. In these diagrams, algorithms ranked farther to the right generally indicate better clustering or segmentation performance.



(a)



(b)

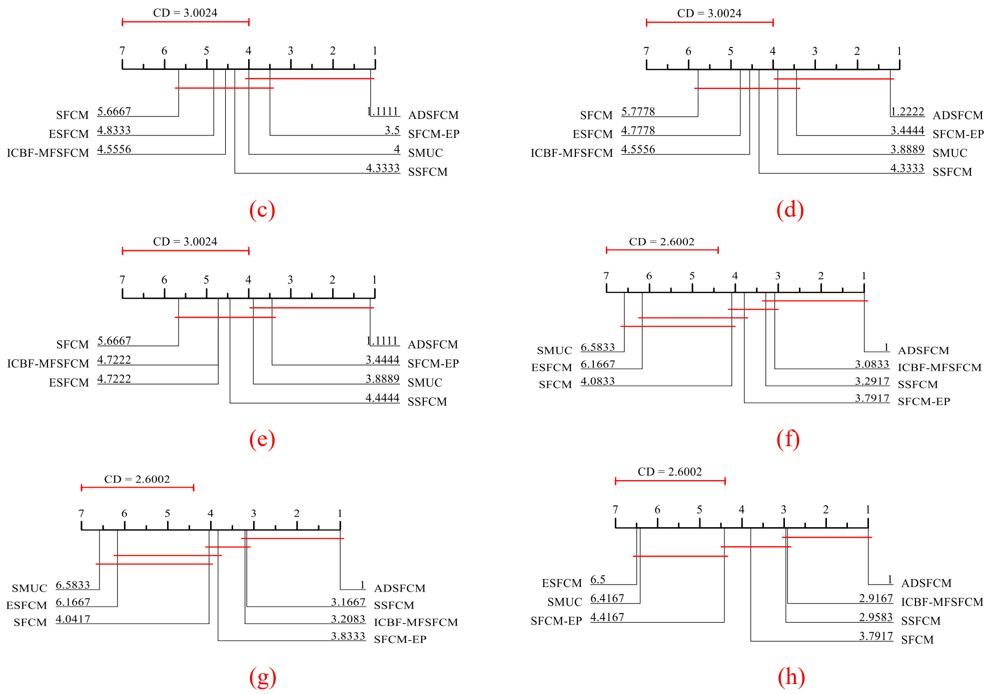


Fig. 2. CD diagrams of seven algorithms on nine numerical and twelve image datasets. (a) ACC1; (b) PE; (c) RE; (d) ARI; (e) NMI; (f) ACC2; (g) Jaccard; (h) mIoU.

Metrics ACC1, PE, and RE assess label consistency, while ARI and NMI evaluate clustering quality. **Figs. 2(a)–(c)** show ADSFCM ranks highest across metrics, and **Figs. 2(d)–(e)** confirm its superior clustering quality. ACC2 qualitatively evaluates segmentation results, and the Jaccard index and mIoU quantify quality. **Figs. 2(f)–(h)** demonstrate ADSFCM's superior performance.

Results indicate ADSFCM’s average rank differences mostly exceed the critical difference, confirming its statistically significant superior performance.

(See Section 4.3 “Statistical test analysis”)

---

5. Comment: ***In section 4.2, the author only analyzed the clustering performance and robustness of the ADSFCM algorithm under different datasets, but did not discuss the correlation between the improvement of algorithm performance and the enhancement of interpretability. Since the research motivation proposed in Section 3.1 is to address the interpretability issue of the SFCM algorithm, the experimental results should have the corresponding experimental motivation. The author should also provide relevant conclusions regarding the interpretability of the algorithm.***

Response:

We sincerely appreciate your valuable suggestion regarding the interpretability aspect of our work.

We acknowledge that the interpretability was not sufficiently demonstrated and reflected in the subsequent experiments, which is indeed a shortcoming of our original manuscript.

In response, we have added a new subsection titled **(3) Interpretability analysis** in **Section 4.2**, which is clearly highlighted in the revised manuscript. The added analysis systematically evaluates the theoretical interpretability and advantages of the proposed asymmetric deviation constraint term $(u_{ij}(1+b_i)-3b_i\overline{f}_{ij})^2$ compared to the symmetric deviation term $(u_{ij}-b_i\overline{f}_{ij})^2$ in SFCM and the membership fusion term $b_i((u_{ij}-\overline{f}_{ij})^2+u_{ij}^2)$ in ICBF-MFSFCM.

Considering that numerical datasets generally possess clear distribution structures and flexible control of prior information, which facilitate an in-depth analysis of the effects of different supervision mechanisms on clustering performance, three representative datasets were selected: the uniformly distributed dataset Wdbc, the multimodally distributed dataset Lsun, and the non-spherically distributed dataset Zelnik6. A systematic evaluation of clustering performance differences among the three algorithms was conducted.

Specifically, for each dataset, the prior membership degree $\overline{f}_{ij}$ of labeled samples x_i belonging to the j-th clustering center was fixed at 1, with prior membership degrees for other clusters set to 0. The percentage of labeled samples was gradually increased, with labeled samples evenly distributed among classes, to evaluate the efficiency of each algorithm in utilizing prior information. ACC curves for the three algorithms were plotted on the Wdbc, Lsun, and Zelnik6 datasets under these conditions.

When the percentage of labeled samples was fixed at 20%, with samples evenly distributed across classes, another set of experiments simulated the gradual approach of labeled samples' membership degrees to ideal prior values. Specifically, the prior membership $\overline{f}_{ij}$ of labeled samples x_i belonging to the j-th cluster was increased stepwise from 0.5 to 1.0 with increments of 0.05, while other clusters' prior memberships remained at zero. Corresponding ACC performance curves varying with prior membership degrees were also plotted for the three algorithms on all three datasets, providing a comprehensive evaluation of interpretability and algorithmic advantages.

Added content

To evaluate the benefits and interpretability of the asymmetric deviation constraint $(u_{ij}(1+b_i)-3b_i\bar{f}_{ij})^2$ versus the symmetric term $(u_{ij}-b_i\bar{f}_{ij})^2$ in SFCM and the membership fusion term $b_i((u_{ij}-\bar{f}_{ij})^2+u_{ij}^2)$ in ICBF-MFSFCM, experiments used three datasets—uniform Wdbc, multimodal Lsun, and non-spherical Zelnik6—to compare clustering performance.

Fixing the prior membership degree, the proportion of prior information is adjusted: Based on standard class labels, the prior membership degree $\bar{f}_{ij}$ for a labeled data point x_i belonging to the j-th cluster is set to 1, and to 0 for all other clusters. Labeled samples increase evenly across classes to evaluate algorithm efficiency. **Fig. 3** presents ACC curves for three algorithms on Wdbc, Lsun, and Zelnik6 datasets with varying labeled data.

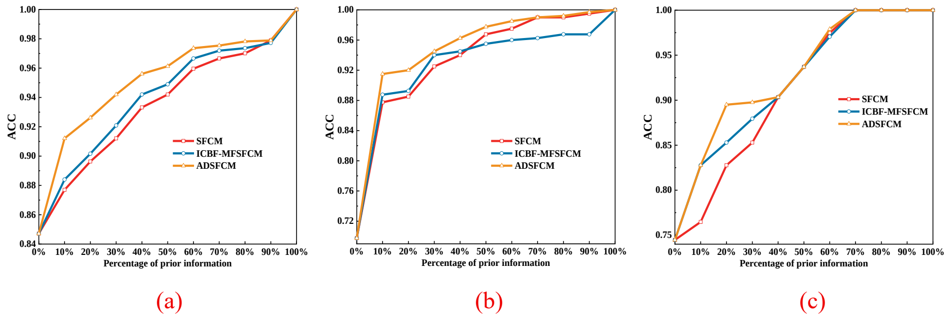


Fig. 3. ACC curves of three algorithms at varying prior information levels. (a) Wdbc; (b) Lsun; (c) Zelnik6.

Fig. 3 shows that without prior information, SFCM, ADSFCM, and ICBF-MFSFCM perform poorly, similar to standard fuzzy C-means. Increasing labeled samples improves all methods, with ADSFCM showing more stable gains at low prior levels due to its asymmetric deviation constraint. While all stabilize at high prior levels, ADSFCM's superior performance with limited labeled data is more practical given labeling costs.

With the prior information percentage fixed, the prior membership degree is adjusted: With 20% of samples labeled and evenly distributed, the membership degree $\bar{f}_{ij}$ of a labeled sample x_i in the j-th cluster rises from 0.5 to 1.0 in 0.05 steps, while others remain zero. **Fig. 4** shows ACC performance curves of three algorithms on Wdbc, Lsun, and Zelnik6 datasets as prior membership degrees change.

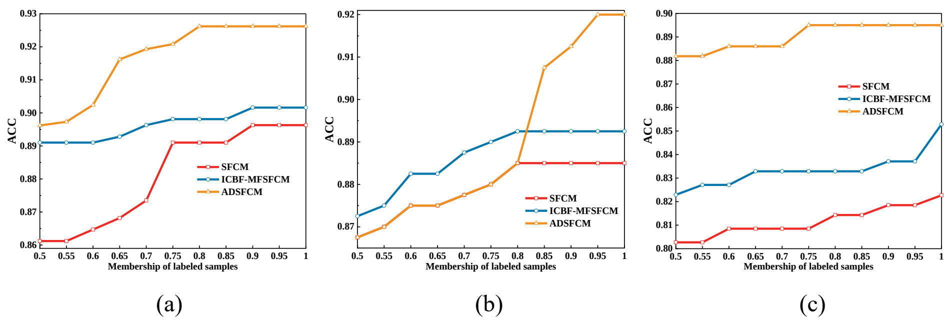


Fig. 4. Accuracy curves of three algorithms at varying prior membership degrees. (a) Wdbc; (b) Lsun; (c) Zelnik6.

Fig. 4 shows that in the uniformly distributed Wdbc dataset, some boundary data points have uncertain memberships, resulting in low membership values to their cluster centers. Setting the prior membership degree of labeled points to 0.65 significantly improves ADSFCM's clustering performance, indicating that the asymmetric deviation constraint effectively guides optimization even when actual memberships only approximate the prior. In the multimodal Lsun dataset, most points lie in well-separated classes with high membership degrees. Setting the prior membership to 0.85 again enhances ADSFCM's performance, confirming the constraint's effectiveness when memberships approach prior values. In the non-spherical Zelnik6 dataset, membership degrees are uneven due to the complex structure. The asymmetric deviation constraint leverages prior information to uncover structural features, enabling ADSFCM to outperform SFCM and ICBF-MFSFCM, demonstrating superior structural adaptability.

In summary, the asymmetric deviation constraint prevents prior information failure when $u_{ij} \approx \overline{f}_{ij}$, adapting to various data distributions and ensuring stable, robust clustering even with limited prior information.

(See subsection (3) Interpretability analysis in Section 4.2 "Performance testing and analysis")

---

6. Comment: ***In section 4.5, the author only selected the SFCM and ADSFCM algorithms as the algorithm carriers for the acceleration methods, but did not include the comparison algorithms such as SSFCM, ESFCM, SMUC, SFCM-EP, and ICBF-MFSFCM adopted in Section 4.2. The author should provide the specific basis for making this selection.***

Response:

The comment is greatly appreciated, and it is acknowledged that the original manuscript lacked a clear explanation for the selection criteria of the algorithm carriers in **Section 4.5**. This oversight has been carefully addressed in the revised version.

Specifically, the aim of the proposed acceleration strategies based on non-affinity center identification and membership scaling is to provide a generalizable mechanism applicable to a broad class of semi-supervised fuzzy clustering algorithms. To validate the feasibility and universality of the proposed method, the SFCM algorithm was selected due to its foundational role in the semi-supervised fuzzy clustering domain; most subsequent algorithms, including SSFCM, ESFCM, and others you mentioned, are essentially extensions or modifications of the SFCM framework. Consequently, using SFCM as a representative baseline ensures both clarity and generalizability in demonstrating the effectiveness of the proposed acceleration strategies.

Additionally, due to limitations in manuscript length and the complexity of comparative presentation, an extensive list of algorithms was not incorporated in Section 4.5. This rationale has now been explicitly added and highlighted in the revised version to clarify the selection basis.

Appreciation is again extended for your insightful feedback, which helped improve the manuscript's clarity and completeness.

Added content

The proposed accelerated semi-supervised fuzzy C-means clustering based on affinity filtering and membership scaling has been theoretically validated, showing strong feasibility and broad applicability to various semi-supervised clustering frameworks. Since SFCM serves as the foundational framework for most subsequent enhancements, these derived methods essentially preserve its core structure. Accordingly, this paper employs both SFCM and ADSFCM as the principal algorithmic carriers for implementing the proposed acceleration strategy. The accelerated version based on SFCM is denoted as AMSFCM, while the accelerated variant of ADSFCM is referred to as AM-ADSFCM.

(See the second paragraph of subsection (1) Numerical data sets in Section 4.7 "Testing of fast semi-supervised algorithms".)

---

7. Comment: ***The effectiveness of the author's membership degree scaling mechanism depends on the accuracy of the prior membership degrees. When there are deviations or errors in the prior membership degrees, this mechanism may have an adverse impact on the algorithm optimization process. Therefore, it is recommended to clearly state in the text the applicable scenarios and application domains of this mechanism in practical applications.***

Response:

We greatly appreciate your insightful comment on the potential limitation of the proposed membership degree scaling mechanism. Your observation is indeed valid and important: the effectiveness of this mechanism depends on the accuracy of the prior membership degrees. When deviations or errors exist in the prior information, the acceleration mechanism may exert a negative impact on the optimization process. We acknowledge that this issue was not sufficiently addressed in the original version of the manuscript.

In response, we have revised the conclusion section to explicitly describe both the advantages and the limitations of the proposed method. In particular, we have clarified that this mechanism is more applicable in scenarios where the confidence level of the prior information is relatively high, and data complexity or computational cost is significant. The relevant revisions have been clearly highlighted in the revised manuscript.

Added content

To enhance the utilization of prior information in semi-supervised clustering when membership degrees approach prior values, we redesign the SFCM regularization term by replacing the symmetric deviation constraint with an asymmetric one. Using non-affinity center identification and membership scaling, we introduce a novel acceleration framework for diverse semi-supervised clustering models. Experiments show ADSFCM's asymmetric deviation constraint enhances prior information use and adaptability, while its acceleration mechanism speeds convergence and stabilizes clustering, benefiting large or complex datasets.

ADSFCM and its acceleration strategy have limitations, including reliance on accurate prior information, difficulty with complex high-dimensional data, and potential negative effects from imprecise prior membership on optimization.

(See the first and second paragraphs of Section 5 “Conclusions and outlook”.)

To Reviewer #2:

We sincerely appreciate your recognition and attention to this research, as well as the valuable comments you provided. The issues you raised are of significant importance for further improving the structure, content, and expression of the manuscript.

To clearly present our revisions, we have highlighted all the changes in the revised manuscript. Additionally, we have conducted comprehensive language polishing and logical optimization throughout the paper, striving to enhance its overall quality and readability.

Below, we will respond to each of your specific comments in detail and explain the modifications and improvements made in the corresponding sections.

1. Comment: ***Research related to semi-supervised fuzzy C-means clustering is quite diverse, with various approaches. The paper should provide a more detailed analysis of the limitations of previous studies and approaches regarding the objective function in semi-supervised fuzzy C-means clustering. Based on this, it should propose an improved objective function idea related to the two-level optimization approach. The main academic contributions of the paper should be summarized concisely.***

Response:

We sincerely appreciate your insightful comments on this important issue. We fully agree that the previous version of the manuscript did not sufficiently highlight our core ideas, nor did it provide a detailed analysis of the limitations associated with existing objective functions in semi-supervised fuzzy C-means clustering. To address this concern, we thoroughly revised the introduction section to improve both the logical organization and the clarity of our research motivation.

Specifically, in the introduction, we conducted a more in-depth analysis of the limitations of traditional semi-supervised fuzzy clustering objective functions, including those used in SFCM and entropy-based methods such as ESFCM and SMUC, particularly in terms of the construction of semi-supervised regularization terms. Based on this analysis, we briefly reviewed existing optimization strategies in SFCM and naturally introduced our motivation for designing an improved objective function that incorporates a two-level optimization framework.

In addition, following your suggestion, we have streamlined and refined the description of our academic contributions. The revised version now highlights our key innovations, including the introduction of the asymmetric deviation constraint and the development of the accelerated semi-supervised fuzzy C-means clustering based on new affinity filtering and membership scaling. The remaining contributions have been presented more concisely to enhance clarity and maintain focus.

These revisions appear in the introduction section, with modifications clearly highlighted for your convenience.

After revision

Clustering groups data based on similarity, with traditional unsupervised methods often ignoring prior information, causing suboptimal results (Singh and Singh, 2024). Machine learning uses unsupervised, supervised, and semi-supervised approaches to analyze data (Mendonca et al., 2024). Labeling data remains challenging, limiting the accuracy of supervised learning, while unsupervised learning may waste partial labels. Semi-supervised learning (Engelen and Hoos, 2020) combines limited labeled and abundant unlabeled data to improve clustering. The SSC algorithm applies this, resulting in different semi-supervised clustering types (Cai et al., 2023).

Fuzzy C-means (FCM) clustering (Bezdek, 1984) and fuzzy entropy clustering (FE) (Tran et al., 2000) are widely used partition-based unsupervised algorithms. Many researchers have developed semi-supervised clustering methods by incorporating prior membership into objective functions via regularization terms. Pedrycz (1985) emphasized integrating prior membership within the FCM framework by adding penalty terms to improve performance. Pedrycz and Waletzky (1997) advanced this by using discrete vectors to represent labeled and unlabeled data, formulating membership metrics through matrix representation. Yasunori et al. (2009) proposed the semi-supervised fuzzy C-means (SSFCM) clustering algorithm, integrating prior membership into the unsupervised objective function, and later developed the entropy-regularized semi-supervised fuzzy C-means (ESFCM) by incorporating maximum entropy principles. Building upon ESFCM, Yin et al. (2012) introduced SMUC, which combines metric learning and entropy regularization using Mahalanobis distance instead of Euclidean distance. However, these methods share a limitation: the semi-supervised regularization term, often a product involving distance metrics, loses influence as membership degrees approach prior values. Since distance metrics are based on sample-to-center distances, this weakens the semi-supervised information's guidance on clustering center positions, reducing clustering accuracy. Therefore, further research is required to optimize the semi-supervised regularization term, enhance the utilization of prior information, and improve clustering accuracy.

Based on this, three main optimization directions have emerged in subsequent research. First, optimizing distance metrics or constraint conditions in semi-supervised regularization: Kanzawa (2017) improved entropy regularization in kernel fuzzy C-means clustering with soft pairwise constraints. Salehi et al. (2021) developed the SMKFC-ER algorithm, replacing geometric distance metrics with relative entropy divergence. Antoine et al. (2022) introduced an adaptive distance metric for semi-supervised possibilistic C-means clustering. Wang et al. (2021) proposed SSFPC by integrating probabilistic fuzzy constraints. Zhu et al. (2024) designed a new distance metric in

SSFKM to better assign labeled samples and enhance label discrimination. Zhang et al. (2023) developed a semi-supervised fuzzy clustering algorithm with soft constraints to resolve conflicts under varying constraint quantities. Ibrahim et al. (2020) incorporated penalty and reward components into constraints. Second, optimizing fusion strategies for prior information: Tuan et al. (2022) proposed MCSSFC-P and MCSSFC-C algorithms using defuzzification coefficients and unique fuzzification factors. Zhang et al. (2024) and Zhang et al. (2024) introduced a member fusion mechanism improving cerebral infarction lesion segmentation with limited labels. Xu et al. (2024) developed a fuzzy C-means clustering algorithm with KL-divergence (Ichihashi et al., 2001) using a prior membership matrix reflecting expert preferences. Yu et al. (2024) created FW-SSPCM, enhancing adaptability to imbalanced features via supervised centers and weighted features. Third, optimizing the reliability of prior information: Gan et al. (2018) proposed LHC-S3FCM using localized chi-squared congruence and Gaussian kernel similarity to align labeled instances with local clusters. Gan (2019) introduced S3FCM, penalizing discrepancies with traditional FCM. Gan et al. (2019) enhanced this with CS3FCM by weighting prior information based on trustworthiness. Tuan et al. (2022) developed TS3MFCM, assigning defuzzification values per data point. Huan et al. (2022) created TS3FCM, reducing computation by decoupling labeling from clustering. Gan et al. (2024) proposed DaS3FCM, integrating discriminative learning and perceptual safety, using contour coefficients to estimate cluster confidence and FCM to correct labels, thereby enhancing prior information reliability.

Overall, although optimizing distance metrics, constraint conditions, prior information fusion strategies, and their reliability can improve semi-supervised clustering algorithms (Daneshfar et al., 2024), most methods rely on symmetric deviation between membership and prior membership in the regularization term. This weakens the guidance of prior information when current membership closely matches prior membership, degrading clustering performance. To address this, we propose a semi-supervised fuzzy C-means clustering with asymmetric deviation constraints (ADSFCM), extending the SFCM framework and incorporating collaborative fuzzy clustering principles (Pedrycz, 2022; Shen and Pedrycz, 2017). By introducing asymmetric deviation into the regularization term, ADSFCM overcomes structural limitations and enhances prior information utilization. Additionally, inspired by Kmita et al. (2024), a weighted adjustment factor is added to the fuzzy clustering term to better balance unsupervised and partially supervised information.

This paper's main contributions are:

● This paper enhances semi-supervised fuzzy C-means clustering by adding asymmetric deviation constraints to better utilize a priori information.

● The algorithm's convergence is confirmed by Zangwill's theorem, ensuring a solid theoretical foundation.

● An accelerated semi-supervised fuzzy C-means algorithm speeds convergence by filtering non-affinity centers and scaling membership, promoting its widespread application.

(See the Introduction.)

----

2. Comment: ***The section on related studies of the two algorithms—Fuzzy C-means and Semi-supervised Fuzzy C-means—should be more concise, focusing on analyzing the limitations of the objective function and boundaries. Additionally, the paper should include an analysis of the shortcomings and existing issues of Semi-supervised Fuzzy C-means, thereby proposing an improvement idea based on the optimization approach.***

Response:

Thank you for the constructive suggestion. The related work section has been thoroughly reconsidered and substantially revised to enhance clarity and focus.

Firstly, the description of the Fuzzy C-means (FCM) objective function framework has been retained as the foundation, highlighting that most semi-supervised clustering methods extend from this framework by incorporating semi-supervised regularization terms. Subsequently, a focused discussion has been added on the limitations of the objective functions in semi-supervised fuzzy C-means methods, SFCM and SSFCM, specifically addressing the use of symmetric deviation terms $(u_{ij} - b_i \overline{f}_{ij})^2$ in their semi-supervised regularization term. This construction tends to significantly weaken the constraint on the distance measure $d^2(x_i, v_j)$ when the membership u_{ij} approaches the prior membership $\overline{f}_{ij}$, thereby reducing the guiding effect on clustering centers and ultimately impairing clustering accuracy.

Following this, the ICBF-MFSFCM algorithm is introduced, emphasizing its membership fusion mechanism, which effectively mitigates the inherent shortcomings in the semi-supervised regularization term of traditional fuzzy clustering. The improvements proposed in this paper build upon this algorithm to further address the limitations in the objective function design of traditional semi-supervised clustering methods.

Regarding fuzzy entropy clustering and its semi-supervised variants, although they also adopt a symmetric deviation constraint, their entropy function is beyond the primary scope of this paper and bears limited relevance to the current work. Therefore, the original manuscript's discussions on FE, ESFCM, SMUC, and SFCM-EP have been removed to streamline the related work section and improve focus, while these topics remain potential directions for future research.

The revised related work section reflects these changes and aims to present a clearer, more concise analysis aligned with the paper's objectives.

After revision

2.1. Fuzzy C-means clustering and variants

The FCM algorithm (Bezdek, 1984) is a soft clustering method that assigns data points to clusters with varying membership degrees by iteratively minimizing a weighted sum of squared Euclidean distances between samples and centroids. The optimization model for FCM is as follows:

$$\min J_{FCM}(U,V;X)=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}d^{2}(x_i,v_j) \tag{1}$$

s.t. (a) $u_{ij}\in[0,1],\forall i,j$; (b) $\sum_{j=1}^{c}u_{ij}=1,\forall i$; (c) $0<\sum_{i=1}^{n}u_{ij}<n,\forall j$.

where $m>1$ (commonly set to 2) controls the fuzziness degree in clustering, with $m=1$ corresponding to hard C-means clustering. Let x_i be the i-th data sample in dataset $X=\{x_i\,|\,1\le i\le n\}$, v_j the j-th clustering center; and u_{ij} the membership degree of the i-th data sample x_i belongs to the j-th cluster. $d^2(x_i,v_j)=\left\|x_i-v_j\right\|_2^2$ is the squared Euclidean distance between the i-th data sample x_i and the j-th clustering center v_j. $U=[u_{ij}]_{c\times n}$ and $V=[v_j]_{c\times d}$ represent the fuzzy partition and clustering center matrices, respectively.

Using the Lagrange multiplier method, the unconstrained Lagrange function for minimizing **Eq. (1)** with constraint $\sum_{j=1}^{c}u_{ij}=1$ is formed, and iterative formulas for membership u_{ij} and the clustering center v_j are obtained by zeroing partial derivatives.

$$u_{ij}=\frac{1}{\sum_{k=1}^{c}\left(d^2(x_i,v_j)/d^2(x_i,v_k)\right)^{\frac{1}{m-1}}} \tag{2}$$

$$v_i = \frac{\sum_{i=1}^{n} u_{ij}^m \cdot x_i}{\sum_{i=1}^{n} u_{ij}^m} \tag{3}$$

The FCM objective function captures sample-cluster uncertainty for soft partitioning and, being differentiable, allows integrating prior information through regularization without losing its iterative optimization.

2.2. Semi-supervised fuzzy C-means clustering

Unsupervised fuzzy clustering is simple and effective but lacks prior information use. Pedrycz and Waletzky (1997) addressed this by adding a semi-supervised regularization term to the FCM objective in the SFCM algorithm. The optimization model for SFCM is as follows:

$$\min J_{SFCM}(U,V) = \sum_{j=1}^{c}\sum_{i=1}^{n} u_{ij}^m d^2(x_i, v_j) + \alpha \sum_{j=1}^{c}\sum_{i=1}^{n} (u_{ij} - b_i \bar{f}_{ij})^m d^2(x_i, v_j) \tag{4}$$

s.t. (a) $u_{ij} \in [0,1], \forall i, j$; (b) $\sum_{j=1}^{c} u_{ij} = 1, \forall i$; (c) $0 < \sum_{i=1}^{n} u_{ij} < n, \forall j$; (d) $\bar{f}_{ij} \in [0,1], \forall i, j$.

where α is a regularization factor balancing unsupervised and partially supervised information, $\bar{f}_{ij} \in F$ is the prior membership degrees of x_i in the j-th cluster, $d^2(x_i, v_j)$ is the squared Euclidean distance between the i-th data sample and the j-th clustering center. Additionally, b_i is a binary vector with defined constraints.

$$\begin{cases} b_i = 1, & \text{if } x_i \text{ is a labeled sample} \\ b_i = 0, & \text{otherwise} \end{cases} \tag{5}$$

Using the Lagrange multiplier method, the unconstrained Lagrange function for minimizing **Eq. (4)** with constraint $\sum_{j=1}^{c} u_{ij} = 1$ is formed, and iterative formulas for membership u_{ij} and the clustering center v_j are obtained by zeroing partial derivatives.

$$u_{ij} = \frac{1}{1+\alpha^{1/(m-1)}} \cdot \frac{1+\alpha^{1/(m-1)}(1 - b_i \sum_{j=1}^{c} \bar{f}_{ij})}{\sum_{k=1}^{c} (d^2(x_i, v_j) / d^2(x_i, v_k))^{1/(m-1)}} + \frac{\alpha^{1/(m-1)}}{1+\alpha^{1/(m-1)}} \cdot b_i \bar{f}_{ij} \tag{6}$$

$$v_j = \frac{\sum_{i=1}^{n} (u_{ij}^m + \alpha(u_{ij} - b_i \bar{f}_{ij})^m) x_i}{\sum_{i=1}^{n} (u_{ij}^m + \alpha(u_{ij} - b_i \bar{f}_{ij})^m)} \tag{7}$$

where $m = 2$.

Yasunori et al. (2009) introduced the Semi-Supervised Fuzzy C-Means (SSFCM) clustering algorithm, integrating prior membership information into the FCM objective function. Its optimization model is:

$$\min J_{SSFCM}(U,V) = \sum_{j=1}^{c}\sum_{i=1}^{n} (u_{ij} - b_i \bar{f}_{ij})^m d^2(x_i, v_j) \tag{8}$$

The model's constraints match those in **Eq. (4)**.

The clustering center v_j is iteratively derived from the unconstrained Lagrangian function in **Eq. (8):**

$$v_j = \frac{\sum_{i=1}^{n} (u_{ij} - b_i \bar{f}_{ij})^m x_i}{\sum_{i=1}^{n} (u_{ij} - b_i \bar{f}_{ij})^m} \tag{9}$$

When $m \neq 1$ and $1 - \sum_{j=1}^{c} \bar{f}_{ij} \geq 0$, the iterative formula for membership u_{ij} is:

$$u_{ij} = b_i \bar{f}_{ij} + \frac{1 - b_i \sum_{j=1}^{c} \bar{f}_{ij}}{\sum_{k=1}^{c} \left(d^2(x_i, v_j) / d^2(x_i, v_k) \right)^{1/(m-1)}} \tag{10}$$

When $m = 1$, the iterative formula for membership u_{ij} is derived:

$$u_{ij} = \begin{cases} b_i \bar{f}_{ij} + 1 - b_i \sum_{j=1}^{c} \bar{f}_{ij}, & j = \arg\min_{l} \{ d^2(x_i, v_l) \} \\ b_i \bar{f}_{ij}, & \text{otherwise} \end{cases} \tag{11}$$

When membership u_{ij} approaches prior supervised information $\bar{f}_{ij}$ under constraint symmetric deviation $(u_{ij} - b_i \bar{f}_{ij})$, semi-supervised regularization in SFCM and SSFCM weakens constraints of distance metric $d^2(x_i, v_j)$, reducing clustering accuracy.

To address this issue, Zhang et al. (2024) introduced the membership fusion semi-supervised fuzzy C-means (ICBF-MFSFCM) clustering algorithm for image segmentation, surpassing SFCM. Its semi-supervised aspect has clear physical significance and can be explained by an electrical analogy. The optimization model for ICBF-MFSFCM is:

$$\min J_{ICBF-MFSFCM}(U, V) = (1-\eta) \sum_{j=1}^{c} \sigma_j \sum_{i=1}^{n} u_{ij}^m d^2(x_i, v_j) + \eta\alpha \sum_{j=1}^{c} \sum_{i=1}^{n} b_i ((u_{ij} - \bar{f}_{ij})^m + u_{ij}^m) d^2(x_i, v_j) \tag{12}$$

where η is the supervision rate and σ_j the interclass balance factor.

The model's constraints are the same as those in **Eq. (4)**.

Applying the Lagrange multiplier method, we derive iterative formulas for membership u_{ij} and the clustering center v_j under condition $m = 2$.

$$u_{ij} = \frac{1 - \eta\alpha b_i \sum_{k=1}^{c} \frac{\bar{f}_{ik}}{(1-\eta)\sigma_k + 2\eta\alpha b_i}}{\sum_{k=1}^{c} \frac{((1-\eta)\sigma_j + 2\eta\alpha b_i) d^2(x_i, v_j)}{((1-\eta)\sigma_k + 2\eta\alpha b_i) d^2(x_i, v_k)}} + \frac{\eta\alpha b_i}{(1-\eta)\sigma_j + 2\eta\alpha b_i} \bar{f}_{ij} \tag{13}$$

$$v_j = \frac{(1-\eta)\sigma_j \sum_{i=1}^{n} u_{ij}^2 x_i + \eta\alpha \sum_{i=1}^{n} b_i ((u_{ij} - \bar{f}_{ij})^2 + u_{ij}^2) x_i}{(1-\eta)\sigma_j \sum_{i=1}^{n} u_{ij}^2 + \eta\alpha \sum_{i=1}^{n} b_i ((u_{ij} - \bar{f}_{ij})^2 + u_{ij}^2)} \tag{14}$$

Without class weighting σ_j, the model simplifies as follows:

$$\min J_{ICBF-MFSFCM}(U,V)=(1-\eta)\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}d^{2}(x_i,v_j)+\eta\sum_{j=1}^{c}\sum_{i=1}^{n}b_i((u_{ij}-\overline{f}_{ij})^{m}+u_{ij}^{m})d^{2}(x_i,v_j) \quad (15)$$

The iterative formulas for membership u_{ij} and cluster center v_j under condition $m=2$, are derived as follows:

$$u_{ij}=\frac{1-\dfrac{\eta b_i}{(1-\eta)+2\eta b_i}\sum_{k=1}^{c}\overline{f}_{ik}}{\sum_{k=1}^{c}d^{2}(x_i,v_j)/d^{2}(x_i,v_k)}+\frac{\eta b_i}{(1-\eta)+2\eta b_i}\overline{f}_{ij} \quad (16)$$

$$v_j=\frac{(1-\eta)\sum_{i=1}^{n}u_{ij}^{2}x_i+\eta\sum_{i=1}^{n}b_i((u_{ij}-\overline{f}_{ij})^{2}+u_{ij}^{2})x_i}{(1-\eta)\sum_{i=1}^{n}u_{ij}^{2}+\eta\sum_{i=1}^{n}b_i((u_{ij}-\overline{f}_{ij})^{2}+u_{ij}^{2})} \quad (17)$$

The ICBF-MFSFCM algorithm uses membership fusion to keep the semi-supervised regularization term comparable to unsupervised membership u_{ij}, ensuring stable and effective clustering.

(See Section 2, "Related work".)

----

3. Comment: ***In the algorithm proposal section (Section 3), the paper should include the conceptual idea and block diagram of the proposed improvements before presenting detailed content. The proposed enhancements to the objective function using the optimization algorithm should be clearly explained.***

Response:

Thank you for your insightful and constructive suggestion. In response, we have made substantial revisions to **Section 3** to improve its structure and clarity.

Specifically, we added a new subsection 3.1 "Motivation", in which we provide a detailed explanation of the inherent limitations in the objective function design of traditional semi-supervised fuzzy clustering algorithms such as SFCM. We also briefly introduce the method ICBF-MFSFCM, which attempts to address this issue but deviates from the core concept of semi-supervised clustering.

Based on this analysis, we further proposed a new optimization model by constructing a semi-supervised regularization term with an asymmetric deviation constraint, which directly addresses the limitations mentioned above.

To better convey the conceptual framework of our improvements, we incorporated a block diagram that visually illustrates the structure of the proposed clustering optimization model. This diagram, along with the accompanying explanation, clearly presents the motivation and advantages of our new semi-supervised regularization term design. Only after this conceptual clarification do we proceed with a detailed discussion of the objective function formulation and the optimization process.

We believe these revisions significantly improve the clarity, completeness, and logical flow of the algorithm section, making it easier for readers to understand the proposed contributions.

Added content

The introduction of SFCM (Pedrycz and Waletzky, 1997) established a theoretical basis for semi-supervised clustering, leading to many related optimization algorithms. However, the SFCM objective function's semi-supervised regularization term is typically a product involving a distance metric. Under the symmetric deviation constraint $(u_{ij} - b_i \overline{f}_{ij})^2$, when labeled samples' membership degrees approach their prior values, the constraint's influence on the distance metric $d^2(x_i, v_j)$ weakens. Since the distance metric depends on distances between samples and clustering centers, the guidance from prior information diminishes, reducing clustering accuracy. Although the ICBF-MFSFCM (Zhang et al., 2024) improves this via membership fusion—integrating $(u_{ij} - \overline{f}_{ij})^2 + u_{ij}^2$ and $d^2(x_i, v_j)$, it still underuses prior information when memberships align closely with priors, relying mainly on unsupervised clustering. Therefore, a new semi-supervised regularization term is needed.

(See Section 3.1 "Motivation")

Building on the ICBF-MFSFCM algorithm (Zhang et al., 2024), we introduce an asymmetric constraint to integrate $(2u_{ij} - 3\overline{f}_{ij})^2$ and $d^2(x_i, v_j)$ in semi-supervised clustering, enhancing performance with a new optimization model.

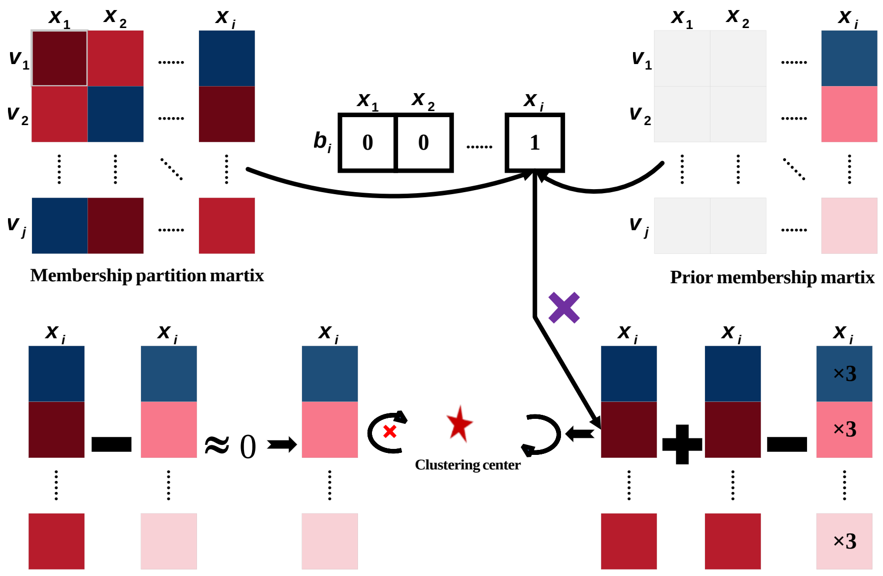


Fig 1. Block diagram of the new clustering optimization model

Fig. 1 illustrates that in the membership partition matrix, sample x_i belongs to the j-th cluster with varying membership u_{ij}. When the binary vector $b_i = 0$, the i-th column of the prior membership matrix, lacks prior information, it indicates the sample is unlabeled. Conversely, $b_i = 1$ means the i-th sample is labeled with prior membership information. If the current membership degree u_{ij} of sample x_i closely matches its prior membership $\overline{f}_{ij}$, the influence of prior information on guiding cluster centers weakens, potentially reducing clustering performance. To address this, an asymmetric deviation term $(2u_{ij} - 3\overline{f}_{ij})^2$ is introduced based on the binary vector b_i. For unlabeled sample x_i, the semi-supervised regularization term reduces to u_{ij}, matching the fuzzy clustering term's magnitude and aligning with unsupervised clustering. For labeled samples with $u_{ij} \approx \overline{f}_{ij}$, the regularization becomes the asymmetric deviation term $(2u_{ij} - 3\overline{f}_{ij})^2$, ensuring prior information $\overline{f}_{ij}$ effectively guides clustering. This design preserves the core principle of semi-supervised clustering.

(See the first and second paragraphs in Section 3.2 "Optimization modeling")

---

4. Comment: ***In the experimental results section, statistical analysis of the experiments should be added***

Response:

Thank you for your thoughtful and constructive suggestion. We appreciate your emphasis on the importance of statistically validating the experimental results. In response, we have included a dedicated subsection titled "Statistical test analysis" in **Section 4.3** of the revised manuscript.

In this analysis, we applied the Friedman test to evaluate whether the observed performance differences across seven algorithms, including our proposed algorithm ADSFCM, are statistically significant. This evaluation was performed based on the results obtained from nine numerical datasets and twelve image datasets. The calculated corrected F-statistics and corresponding P-values consistently indicated that the differences in clustering performance across all evaluation metrics are statistically significant.

To further investigate specific algorithmic differences, we conducted a post-hoc analysis and visualized the outcomes using critical difference (CD) diagrams. These diagrams clearly illustrate the relative ranking and statistical significance of performance differences among the compared methods.

The inclusion of this statistical validation enhances the rigor of our experimental analysis and provides more convincing evidence for the effectiveness of our proposed approach.

Added content

To evaluate whether performance metrics differences between the ADSFCM algorithm and others are statistically significant, the Friedman test (Demšar, 2006) is applied to nine datasets and twelve images; results appear in **Table 1**.

For clarity, the variables in the statistical test are defined as follows: k is the total number of algorithms compared; N is the total number of datasets evaluated; r_i^j is the rank of the j-th algorithm on the i-th dataset, with 1 as best and k as worst; R_j is the average rank of the j-th algorithm across all datasets; $df_1 = k-1$ and $df_2 = (k-1)(N-1)$ are the degrees of freedom; χ_F^2 and F_F are the test statistics based on the chi-square distribution and the corrected F-distribution, respectively, calculated as follows:

$$\chi_F^2 = \frac{12N}{k(k+1)}\left(\sum_{i=1}^{k} R_j^2 - \frac{k(k+1)^2}{4}\right) \tag{18}$$

$$F_F = \frac{(N-1)\chi_F^2}{N(k-1)-\chi_F^2} \tag{19}$$

The P-value measures the result randomness and tests sample differences. With a significance level $a = 0.05$. $P \le a$ leads to rejecting the null hypothesis of no algorithm performance difference. It is calculated by **Eq. (20)**.

$$P = F_a(df_1, df_2) \tag{20}$$

Table 1 F_F and P value for Friedman's test

	ACC1	PE	RE	ARI	NMI	ACC2	Jaccard	mIoU

F_F	6.6002	5.8802	6.6057	6.1972	6.6505	42.9514	40.1664	60.1587
P value	0.000039	0.000117	0.000039	0.000072	0.000036	6.37×10^{-18}	2.11×10^{-17}	1.07×10^{-20}

*ACC1 and ACC2 denote clustering and segmentation accuracy, respectively.

Table 1 shows all P-values meet $P \le 0.05$, rejecting the hypothesis and allowing post-hoc tests to identify significant algorithm differences.

If the difference between two algorithms' average ranks is within the critical difference (CD), their performance difference is not statistically significant. If the difference equals or exceeds the CD, the performance difference is significant (Arbelaez et al., 2010). The CD is defined as $CD = q_a\sqrt{\frac{k(k+1)}{6N}}$, where q_a depends on the significance level a. Based on this, CD diagrams for various metrics of seven algorithms across nine numerical datasets and twelve images are shown in **Fig. 2**. In these diagrams, algorithms ranked farther to the right generally indicate better clustering or segmentation performance.

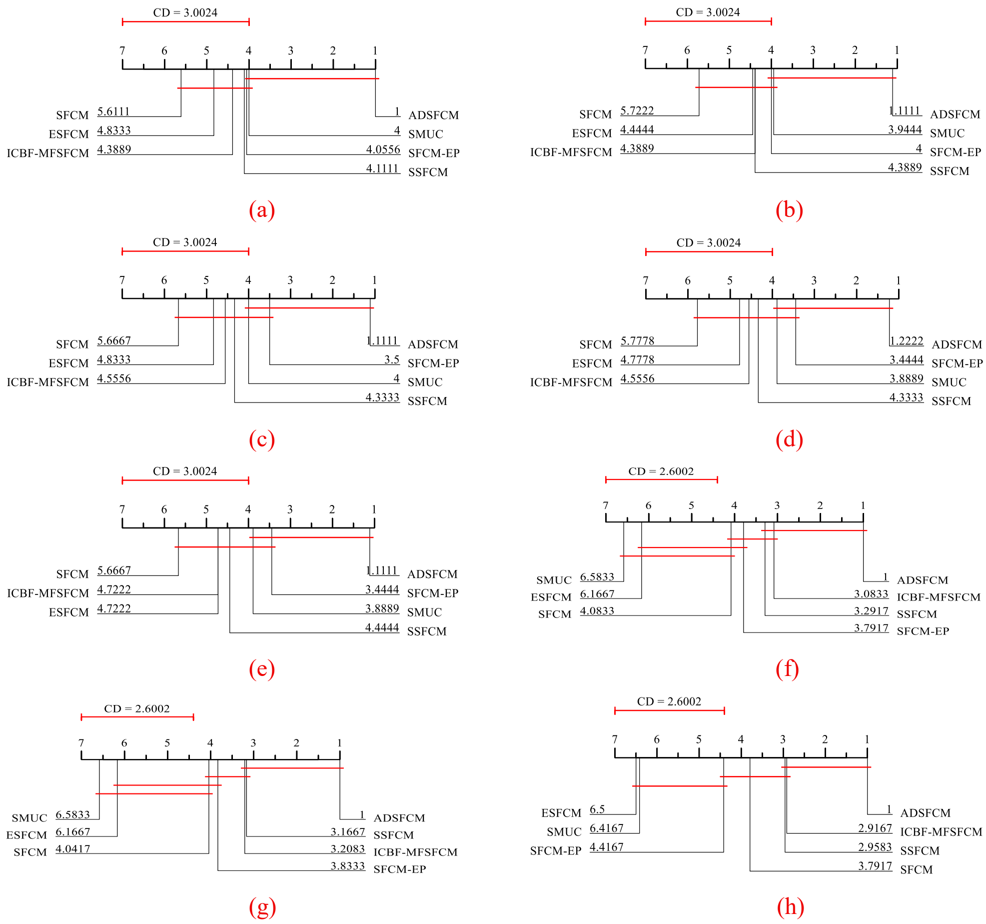


Fig. 2. CD diagrams of seven algorithms on nine numerical and twelve image datasets. (a) ACC1; (b) PE; (c) RE; (d) ARI; (e) NMI; (f) ACC2; (g) Jaccard; (h) mIoU.

Metrics ACC1, PE, and RE assess label consistency, while ARI and NMI evaluate clustering quality. **Figs. 2(a)–(c)** show ADSFCM ranks highest across metrics, and **Figs. 2(d)–(e)** confirm its superior clustering quality. ACC2 qualitatively evaluates segmentation results, and the Jaccard index and mIoU quantify quality. **Figs. 2(f)–(h)** demonstrate ADSFCM's superior performance.

Results indicate ADSFCM's average rank differences mostly exceed the critical difference, confirming its statistically significant superior performance.

(See Section 4.3 "Statistical test analysis")

To Reviewer #3:

We sincerely appreciate your thoughtful evaluation of our work and the valuable constructive feedback. Your insightful suggestions have played a pivotal role in enhancing the technical depth and overall presentation quality of the manuscript. To ensure full transparency, all revisions have been clearly highlighted in the revised manuscript. In addition, we have undertaken a thorough language revision and structural improvements throughout the manuscript to enhance its coherence and clarity.

Below, we provide detailed responses to each of your insightful comments and explain the substantive revisions and enhancements made throughout the manuscript.

1. Comment: ***Please proofread the manuscript carefully to eliminate all typos and grammar errors. For example, "Proposed method" (title of Section 3) should be "The proposed method". "semi-supervise algorithms" (title of Section 4.5) should be "semi-supervised algorithms". More are not listed.***

Response:

Thank you for your valuable suggestion. We sincerely appreciate your careful reading of the manuscript and for pointing out the grammatical inconsistencies. In response to your comment, we have thoroughly proofread the entire manuscript to correct all typographical, grammatical, and stylistic errors. Specifically:

● The section title "Proposed method" has been revised to "The proposed method."

● The section title "semi-supervise algorithms" has been corrected to "semi-supervised algorithms."

● All modifications, including several additional revisions made to improve the overall language quality and ensure grammatical accuracy, have been highlighted in the revised manuscript for your convenience.

2. **Comment:** ***Although the parameter sensitivity is analyzed in experiment, the authors should give out the guidance to choose parameter values in practice.***

Response:

Thank you for your valuable suggestion. We agree that providing practical guidance for parameter selection is important to improve the applicability of the proposed algorithm. In response, we have made the following revisions to the manuscript:

● After performing a detailed parameter sensitivity analysis, we have added explicit guidance on selecting appropriate parameter values in practical applications, based on the observed trends and empirical performance across different data types.

● In **Section 4.5**, subsection (2) Impact of different parameters on clustering efficiency and performance, we revised the textual content to remove redundant expressions, making the sentences more concise and coherent. Furthermore, based on the analysis of how different parameters affect the clustering performance and efficiency of AM-ADSFCM, we also provided practical suggestions for parameter selection.

Added content

Sensitivity analysis suggests setting β to 2–4 and α to 2–10. Stability improves when $\beta = 2\alpha$; thus, $\beta = 4$ and $\alpha = 2$ are recommended. Increasing α stabilizes clustering when adjusting β, while fixing β at 2 enhances robustness when tuning α.

(See the last paragraph of Section 4.5 "Parameter Sensitivity Testing and Analysis".)

Sensitivity analysis indicates that changes in β and α minimally affect AM-ADSFCM's clustering performance and convergence, with iteration variation within ± 2. For stability, set $\beta = 4$ and $\alpha = 2$.

(See the last paragraph of subsection (2) Impact of parameters on clustering efficiency and performance in Section 4.7 "Testing of fast semi-supervised algorithms".)

----

3. **Comment:** ***The authors may simplify the introduction to related work.***

Response:

Thank you for your valuable suggestion. After careful consideration and balancing relevance, we have simplified the related work section in the revised manuscript. Specifically, we removed the discussion of methods such as FE, ESFCM, SMUC, and SFCM-EP, as they are less directly related to the core focus of this paper.

In addition, we reorganized the logical flow of this section to make it more concise and coherent. The revised content now emphasizes a more focused analysis of the limitations of objective functions in existing semi-supervised fuzzy C-means clustering methods. This restructuring not only improves the clarity and compactness of the related work but also provides a smoother and more logical transition into the subsequent sections.

We believe these modifications enhance the readability and thematic focus of the manuscript.

After revision

2.1 Fuzzy C-means clustering and its variants

The FCM algorithm (Bezdek, 1981) is a soft partitioning clustering method that assigns data instances to multiple clusters with varying degrees of association. It uses iterative optimization of an objective function, which is a weighted sum of squared Euclidean distances between data points and cluster centroids. By updating membership degrees and cluster centers to minimize this function, optimal clustering results are achieved. The optimization formulation of the FCM algorithm is as follows:

$$\min J_{FCM}(U,V;X)=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}d^{2}(x_i,v_j) \tag{1}$$

s.t. (a) $u_{ij}\in[0,1],\forall i,j$; (b) $\sum_{j=1}^{c}u_{ij}=1,\forall i$; (c) $0<\sum_{i=1}^{n}u_{ij}<n,\forall j$.

where $m>1$ (commonly set to 2) controls the degree of fuzziness in the clustering outcomes, with $m=1$ aligning with hard C-means clustering. Let x_i be the i-th data sample in the dataset $X=\{x_i\,|\,1\le i\le n\}$, v_j the j-th clustering center; and u_{ij} the membership degree of the i-th data sample x_i belongs to the j-th cluster. $d^{2}(x_i,v_j)=\|x_i-v_j\|_2^2$ is the squared Euclidean distance between the i-th data sample x_i and the j-th clustering center v_j. $U=[u_{ij}]_{c\times n}$ and $V=[v_j]_{c\times d}$ are the fuzzy partition matrix and the clustering center matrix, respectively.

The unconstrained Lagrange function for minimizing **Eq. (1)** under the constraint of $\sum_{j=1}^{c} u_{ij} = 1$ is constructed utilizing the Lagrange multiplier technique. Iterative formulas for membership u_{ij} and the clustering center v_j are derived by computing the partial derivatives and setting them to zero.

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(d^2(x_i, v_j) / d^2(x_i, v_k) \right)^{\frac{1}{m-1}}} \tag{2}$$

$$v_i = \frac{\sum_{i=1}^{n} u_{ij}^m \cdot x_i}{\sum_{i=1}^{n} u_{ij}^m} \tag{3}$$

The objective function of FCM incorporates the membership matrix and fuzzy weighting exponent to effectively capture the uncertainty between samples and clusters, making it particularly suitable for soft partitioning tasks on complex distributions. Its differentiability provides a mathematical foundation for incorporating prior information through regularization terms, thereby enabling the clustering process to align better with prior information while preserving the iterative optimization framework of FCM.

2.2 Semi-supervised fuzzy C-means clustering

Unsupervised fuzzy clustering is popular for its simplicity and effectiveness in certain tasks, but it cannot utilize a priori information. To address this issue, Pedrycz and Waletzky (1997) proposed the SFCM algorithm, which augments the FCM objective function with a semi-supervised regularization term that incorporates prior membership information. The optimization formulation of the SFCM algorithm is as follows:

$$\min J_{SFCM}(U,V) = \sum_{j=1}^{c} \sum_{i=1}^{n} u_{ij}^m d^2(x_i, v_j) + \alpha \sum_{j=1}^{c} \sum_{i=1}^{n} (u_{ij} - b_i \bar{f}_{ij})^m d^2(x_i, v_j) \tag{4}$$

s.t. (a) $u_{ij} \in [0,1], \forall i, j$; (b) $\sum_{j=1}^{c} u_{ij} = 1, \forall i$; (c) $0 < \sum_{i=1}^{n} u_{ij} < n, \forall j$; (d) $\bar{f}_{ij} \in [0,1], \forall i, j$.

where $\alpha > 0$ serves as a regularization factor that balances unsupervised and partially supervised information. $\bar{f}_{ij} \in F$ represents the prior membership degrees of x_i belonging to the j-th cluster. $d^2(x_i, v_j)$ denotes the squared Euclidean distance between the i-th data sample and the j-th clustering center. Additionally, b_i is a binary vector that must satisfy the following constraints.

$$\begin{cases} b_i = 1, \text{ if } x_i \text{ is a labeled sample} \\ b_i = 0, \text{ otherwise} \end{cases} \tag{5}$$

The unconstrained Lagrange function for minimizing **Eq. (4)** under the constraint of $\sum_{j=1}^{c} u_{ij} = 1$ is constructed using the Lagrange multiplier technique. Iterative formulas for membership u_{ij} and the clustering center v_j are derived by setting the partial derivatives of the function to zero.

$$u_{ij} = \frac{1}{1+\alpha^{1/(m-1)}} \cdot \frac{1+\alpha^{1/(m-1)}(1-b_i \sum_{j=1}^{c} \overline{f}_{ij})}{\sum_{k=1}^{c} (d^2(x_i, v_j)/d^2(x_i, v_k))^{1/(m-1)}} + \frac{\alpha^{1/(m-1)}}{1+\alpha^{1/(m-1)}} \cdot b_i \overline{f}_{ij} \tag{6}$$

$$v_j = \frac{\sum_{i=1}^{n} (u_{ij}^m + \alpha(u_{ij} - b_i \overline{f}_{ij})^m) x_i}{\sum_{i=1}^{n} (u_{ij}^m + \alpha(u_{ij} - b_i \overline{f}_{ij})^m)} \tag{7}$$

where $m = 2$.

Yasunori et al. (2009) introduced the Semi-Supervised Fuzzy C-Means (SSFCM) clustering algorithm, integrating prior membership information into the FCM objective function. The optimization formulation of the SSFCM algorithm is as follows:

$$\min J_{SSFCM}(U, V) = \sum_{j=1}^{c} \sum_{i=1}^{n} (u_{ij} - b_i \overline{f}_{ij})^m d^2(x_i, v_j) \tag{8}$$

The constraints of this model are the same as those in **Eq. (4)**.

The iterative formula for the clustering center v_j is derived from the unconstrained Lagrangian function based on **Eq. (8)** as follows:

$$v_j = \frac{\sum_{i=1}^{n} (u_{ij} - b_i \overline{f}_{ij})^m x_i}{\sum_{i=1}^{n} (u_{ij} - b_i \overline{f}_{ij})^m} \tag{9}$$

When $m \neq 1$ and $1 - \sum_{j=1}^{c} \overline{f}_{ij} \geq 0$, the iterative formula for membership u_{ij} is derived as

$$u_{ij} = b_i \overline{f}_{ij} + \frac{1 - b_i \sum_{j=1}^{c} \overline{f}_{ij}}{\sum_{k=1}^{c} \left(d^2(x_i, v_j) / d^2(x_i, v_k) \right)^{1/(m-1)}} \tag{10}$$

When $m = 1$, the iterative formula for membership degree u_{ij} is derived as

$$u_{ij}=\begin{cases} b_i\bar{f}_{ij}+1-b_i\sum_{j=1}^{c}\bar{f}_{ij}, & j=\arg\min_l\{d^2(x_i,v_l)\} \\ b_i\bar{f}_{ij}, & \text{otherwise}\end{cases} \tag{11}$$

Under the influence of the symmetric deviation constraint $(u_{ij}-b_i\bar{f}_{ij})$, when the membership degree u_{ij} approaches its prior membership degree $\bar{f}_{ij}$, the semi-supervised regularization terms in SFCM and SSFCM significantly weaken the constraint on the distance metric $d^2(x_i,v_j)$, thereby reducing the guiding effect on clustering centers and ultimately compromising clustering accuracy.

To address this issue, Zhang et al. (2024) introduced a membership fusion semi-supervised fuzzy C-means (ICBF-MFSFCM) clustering algorithm for image segmentation, outperforming SFCM. Its semi-supervised aspect has clear physical significance and can be interpreted through electrical analogy. The optimization formulation of the ICBF-MFSFCM algorithm is as follows:

$$\min J_{ICBF-MFSFCM}(U,V)=(1-\eta)\sum_{j=1}^{c}\sigma_j\sum_{i=1}^{n}u_{ij}^m d^2(x_i,v_j)+\eta\alpha\sum_{j=1}^{c}\sum_{i=1}^{n}b_i((u_{ij}-\bar{f}_{ij})^m+u_{ij}^m)d^2(x_i,v_j) \tag{12}$$

where η is the supervision rate and σ_j is the interclass balance factor.

The constraints of this model are the same as those in **Eq. (4)**.

Using the Lagrange multiplier method, we derive iterative formulas for membership u_{ij} and the clustering center v_j under the condition of $m=2$, as shown below.

$$u_{ij}=\frac{1-\eta\alpha b_i\sum_{k=1}^{c}\frac{\bar{f}_{ik}}{(1-\eta)\sigma_k+2\eta\alpha b_i}}{\sum_{k=1}^{c}\frac{((1-\eta)\sigma_j+2\eta\alpha b_i)d^2(x_i,v_j)}{((1-\eta)\sigma_k+2\eta\alpha b_i)d^2(x_i,v_k)}}+\frac{\eta\alpha b_i}{(1-\eta)\sigma_j+2\eta\alpha b_i}\bar{f}_{ij} \tag{13}$$

$$v_j=\frac{(1-\eta)\sigma_j\sum_{i=1}^{n}u_{ij}^2x_i+\eta\alpha\sum_{i=1}^{n}b_i((u_{ij}-\bar{f}_{ij})^2+u_{ij}^2)x_i}{(1-\eta)\sigma_j\sum_{i=1}^{n}u_{ij}^2+\eta\alpha\sum_{i=1}^{n}b_i((u_{ij}-\bar{f}_{ij})^2+u_{ij}^2)} \tag{14}$$

Without class weighting σ_j, the model simplifies to the following model.

$$\min J_{ICBF-MFSFCM}(U,V)=(1-\eta)\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^m d^2(x_i,v_j)+\eta\sum_{j=1}^{c}\sum_{i=1}^{n}b_i((u_{ij}-\bar{f}_{ij})^m+u_{ij}^m)d^2(x_i,v_j) \tag{15}$$

The iterative formulas for membership u_{ij} and the clustering center v_j, corresponding to the condition $m=2$, are derived as follows:

$$u_{ij}=\frac{1-\frac{\eta b_i}{(1-\eta)+2\eta b_i}\sum_{k=1}^{c}\overline{f}_{ik}}{\sum_{k=1}^{c}d^2(x_i,v_j)/d^2(x_i,v_k)}+\frac{\eta b_i}{(1-\eta)+2\eta b_i}\overline{f}_{ij} \tag{16}$$

$$v_j=\frac{(1-\eta)\sum_{i=1}^{n}u_{ij}^2x_i+\eta\sum_{i=1}^{n}b_i((u_{ij}-\overline{f}_{ij})^2+u_{ij}^2)x_i}{(1-\eta)\sum_{i=1}^{n}u_{ij}^2+\eta\sum_{i=1}^{n}b_i((u_{ij}-\overline{f}_{ij})^2+u_{ij}^2)} \tag{17}$$

The ICBF-MFSFCM algorithm reconstructs the semi-supervised regularization term through a membership fusion mechanism, ensuring that even when the current membership degrees of labeled samples are consistent with or close to their prior memberships, the semi-supervised regularization term remains on the same order of magnitude as the unsupervised membership degrees u_{ij}. This design guarantees the stability and effectiveness of the clustering results.

(See Section 2, “Related work”.)

---

4. Comment: ***The convergence validation is supposed in experiments.***

Response:

Thank you for pointing out this important issue. To address your comment, we have added a new subsection titled “Convergence test” in **Section 4.4** of the revised manuscript.

In this section, we conducted convergence validation experiments on both numerical and image datasets of varying characteristics and scales. Specifically, we selected Wdbc from uniformly distributed data; Lsun from multimodally distributed data; Zelnik6 from non-spherically distributed data; Natural image #46316; Medical image Tr-me_0177, and Remote sensing image runway_401. These datasets were chosen to ensure the representativeness and diversity of the evaluation.

Since the stopping criterion of our algorithm is based on the deviation between clustering centers in two successive iterations—where convergence is achieved when this deviation is below a predefined convergence threshold—we first plotted the deviation curves of clustering centers between consecutive iterations for both clustering and segmentation tasks. In addition, we illustrated the monotonic changes in the objective function values during the finite iterations.

Experimental results clearly demonstrate that the objective function values of the ADSFCM algorithm decrease strictly monotonically during iterations, which is consistent with the theoretical convergence analysis provided in **Section 3.4**.

We believe that this addition strengthens the theoretical soundness and empirical reliability of our proposed algorithm.

Added content

This study verifies the ADSFCM algorithm convergence using three datasets and three images. **Figs. 1** and **2** show clustering center deviation and objective function changes over iterations. The algorithm stops when the deviation falls below a 10^{-6} threshold.

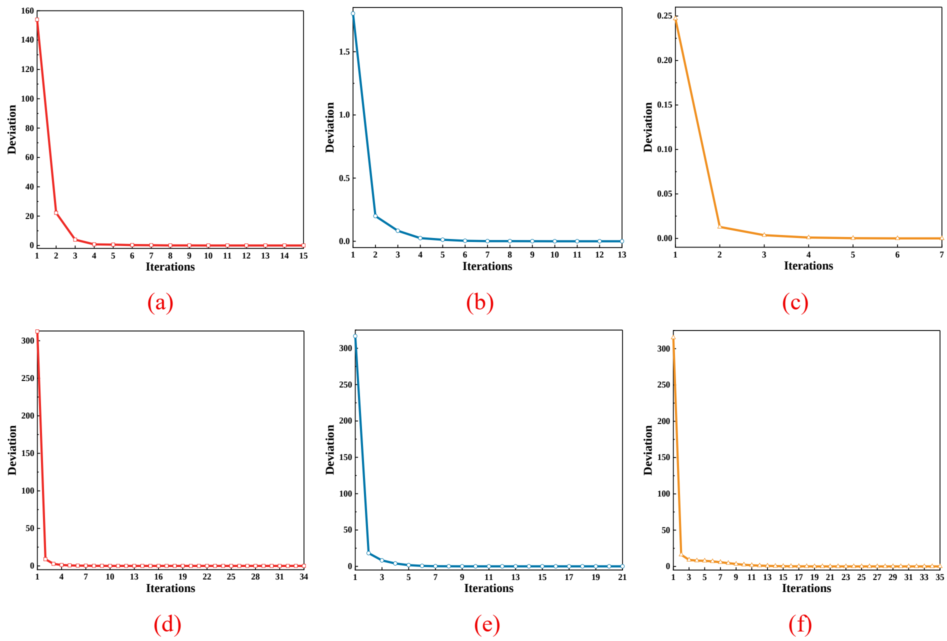
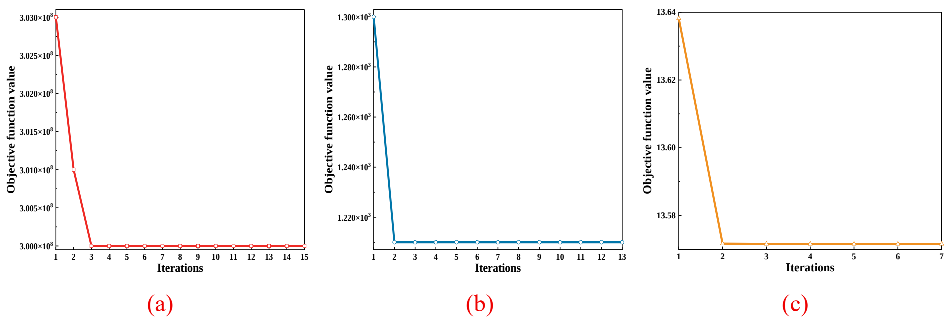


Fig. 1. Deviation of clustering centers between iterations during clustering and segmentation. (a) Wdbc; (b) Lsun; (c) Zelnik6; (d) #46316; (e) Tr-me_0177; (f) runway_401.

Fig. 1 shows that ADSFCM's cluster center deviations decay exponentially, converging within finite iterations. Image segmentation takes longer than clustering, reflecting varied convergence rates and underscoring the need for acceleration strategies to boost efficiency.



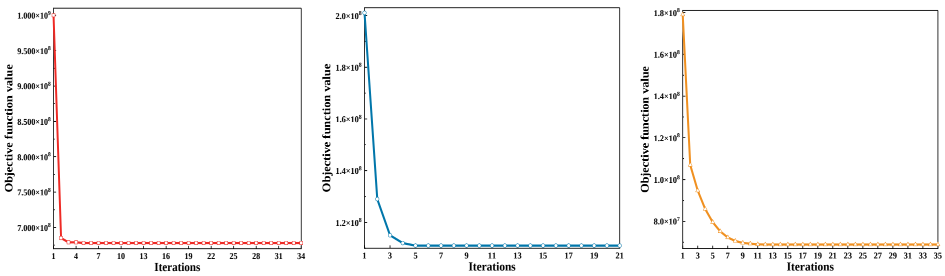

(d) (e) (f)

Fig. 2. Variation in the ADSFCM objective function value across iterations during clustering and segmentation. (a) Wdbc; (b) Lsun; (c) Zelnik6; (d) #46316; (e) Tr-me_0177; (f) runway_401.

Fig. 2 shows ADSFCM's objective function decreases steadily during clustering and more gradually in image segmentation, confirming its monotonic convergence and supporting algorithm improvements.

(See Section 4.4 "The convergence test".)

---

5. Comment: ***The authors are encouraged to (not compulsively) discuss the extension on multi-view data settings, such as [1][2].***

[1] X. Hu, et al. An Efficient Federated Multi-view Fuzzy C-Means Clustering Method, TFS 2023.

[2] J. Liu, et al. Contrastive Multi-view Kernel Learning, TPAMI 2023.

Response:

Thank you for your constructive suggestion. We agree that extending the proposed method to multi-view data settings is a promising and meaningful direction, especially in light of recent advances in federated learning and contrastive representation learning.

In response to your comment, we have included a discussion on this topic in the conclusion section of the revised manuscript. Specifically, we highlight the potential of developing a privacy-preserving distributed semi-supervised clustering system by integrating federated learning with multi-view fuzzy clustering frameworks, enabling secure knowledge sharing across devices. Additionally, we propose incorporating contrastive multi-view learning mechanisms based on kernel methods to enhance the algorithm's representational capacity for complex data structures. These extensions will be the focus of our future research efforts, and the relevant references have been cited accordingly.

Added references

[1] X. Hu, J. Qin, Y. Shen, W. Pedrycz, X. Liu, J. Liu, An efficient federated multiview fuzzy c-means clustering method, IEEE Transactions on Fuzzy Systems, 32(4) (2023) 1886–1899. https://doi.org/10.1109/TFUZZ.2023.3335361

[2] J. Liu, X. Liu, Y. Yang, Q. Liao, Y. Xia, Contrastive multi-view kernel learning, IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(8) (2023) 9552–9566. https://doi.org/10.1109/TPAMI.2023.3253211

Future work includes developing privacy-preserving distributed clustering systems combining federated learning and multi-view fuzzy clustering, alongside integrating multi-view contrastive learning and kernel methods to better represent complex data (Hu et al., 2023; Liu et al., 2023).

(See the third paragraph of Section "5 Conclusions and outlook")

To Associate Editor and all Reviewers:

We sincerely thank the editor and reviewers for their careful review and valuable comments. We would like to emphasize the following points:

(1) We acknowledge that the response letter is relatively lengthy, which may pose some inconvenience for the editor and reviewers. This is due to our intention to provide comprehensive and detailed replies to all comments. We sincerely appreciate your understanding.

(2) We have made efforts to optimize the presentation while preserving the completeness of the content. If the editor deems further compression necessary, we will be glad to make the required adjustments.

(3) We have thoroughly revised the manuscript according to the reviewers' suggestions. While we did not list every minor revision, these changes—focused on improving clarity and structure—have been accurately reflected in the updated version without altering the core content or organization.

We sincerely appreciate the time and valuable feedback provided by the editor and reviewers, and we truly hope that the revised manuscript will meet your expectations. Thank you again for your support and guidance.

Yours sincerely,

Jun Hou

Email: h2927860098@163.com

08/12/2025

https://orcid.org/0000-0001-9736-8052
https://orcid.org/0009-0002-4733-1627

修订稿（标有更改）

Highlights:

(1) A new semi-supervised fuzzy clustering model is proposed.

(2) A semi-supervised fuzzy clustering algorithm's convergence is analyzed.

(3) A fast semi-supervised fuzzy clustering algorithm is presented.

(4) Experimental results show the proposed algorithm's superiority and efficiency.

New semi-supervised fuzzy C-means clustering with asymmetric deviation constraints and fast algorithm

Chengmao Wu, Jun Hou

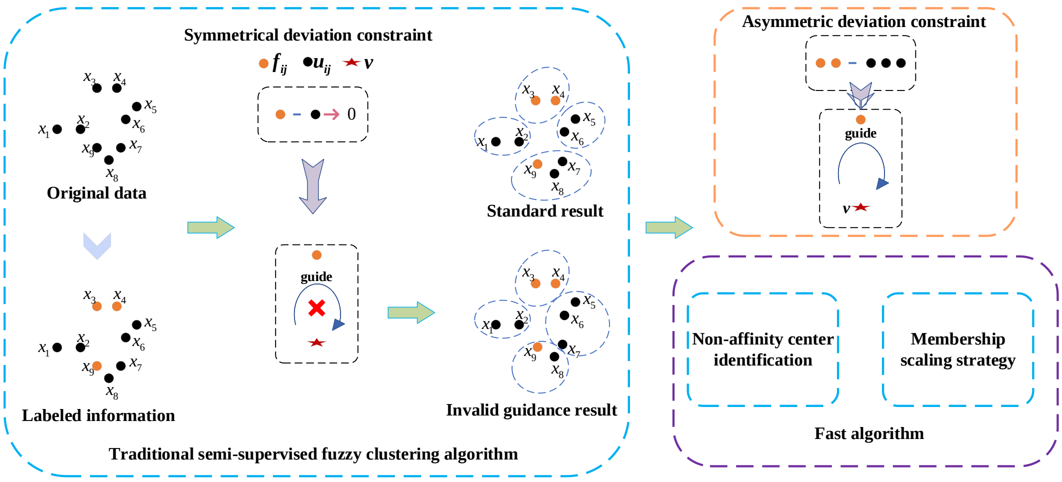
School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, PR China

First author: Chengmao Wu

E-mail: wuchengmao123@sohu.com

Corresponding and second author: Jun Hou

E-mail: h2927860098@163.com



Abstract: Semi-supervised clustering leverages prior information to improve algorithm performance and is widely valued by researchers. This paper analyzes the traditional semi-supervised fuzzy C-means (SFCM) objective function, noting that as a labeled sample's membership degree aligns with its prior information, the impact of this information on the deviation constraint weakens. This reduces its supervisory effect on optimizing the membership partition matrix, especially with a large regularization factor. To overcome this, we propose a novel semi-supervised fuzzy C-means method based on an asymmetric deviation constraint and develop a two-level alternating iterative optimization algorithm, supported by theoretical convergence analysis using Zangwill's theorem and the bordered Hessian matrix. To address the slow convergence and high computational cost typical of semi-supervised fuzzy clustering, we further enhance the algorithm with affinity filtering and a membership scaling scheme for improved efficiency. Experimental results demonstrate that our methods significantly outperform existing state-of-the-art techniques, advancing semi-supervised fuzzy C-means clustering.

Keywords: Fuzzy clustering, asymmetric deviation constraint, affinity filtering, membership scaling, accelerated algorithm

1. Introduction

Clustering groups data based on similarity, with traditional unsupervised methods often ignoring prior information, causing suboptimal results (Singh and Singh, 2024). Machine learning uses unsupervised, supervised, and semi-supervised approaches to analyze data (Mendonca et al., 2024). Labeling data remains challenging, limiting the accuracy of supervised learning, while unsupervised learning may waste partial labels. Semi-supervised learning (Engelen and Hoos, 2020) combines limited labeled and abundant unlabeled data to improve clustering. The SSC algorithm applies this, resulting in different semi-supervised clustering types (Cai et al., 2023).

Fuzzy C-means (FCM) clustering (Bezdek, 1984) and fuzzy entropy clustering (FE) (Tran et al., 2000) are widely used partition-based unsupervised algorithms. Many researchers have developed semi-supervised clustering methods by incorporating prior membership into objective functions via regularization terms. Pedrycz (1985) emphasized integrating prior membership within the FCM framework by adding penalty terms to improve performance. Pedrycz and Waletzky (1997) advanced this by using discrete vectors to represent labeled and unlabeled data, formulating membership metrics through matrix representation. Yasunori et al. (2009) proposed the semi-supervised fuzzy C-means (SSFCM) clustering algorithm, integrating prior membership into the unsupervised objective function, and later developed the entropy-regularized semi-supervised fuzzy C-means (ESFCM) by incorporating maximum entropy principles. Building upon ESFCM, Yin et al. (2012) introduced SMUC, which combines metric learning and entropy regularization using Mahalanobis distance instead of Euclidean distance. However, these methods share a limitation: the semi-supervised regularization term, often a product involving distance metrics, loses influence as membership degrees approach prior values. Since distance metrics are based on sample-to-center distances, this weakens the semi-supervised information's guidance on clustering center positions, reducing clustering accuracy. Therefore, further research is required to optimize the semi-supervised regularization term, enhance the utilization of prior information, and improve clustering accuracy.

Based on this, three main optimization directions have emerged in subsequent research. First, optimizing distance metrics or constraint conditions in semi-supervised regularization: Kanzawa (2017) improved entropy regularization in kernel fuzzy C-means clustering with soft pairwise constraints. Salehi et al. (2021) developed the SMKFC-ER algorithm, replacing geometric distance metrics with relative entropy divergence. Antoine et al. (2022) introduced an adaptive distance metric for semi-supervised possibilistic C-means clustering. Wang et al. (2021) proposed SSFPC by integrating probabilistic fuzzy constraints. Zhu et al. (2024) designed a new distance metric in SSFKM to better assign labeled samples and enhance label discrimination. Zhang et al. (2023) developed a semi-supervised fuzzy clustering algorithm with soft constraints to resolve conflicts under varying constraint quantities. Ibrahim et al. (2020) incorporated penalty and reward components into constraints. Second, optimizing fusion strategies for prior information: Tuan et al. (2022) proposed MCSSFC-P and MCSSFC-C algorithms using defuzzification coefficients and unique fuzzification factors. Zhang et al. (2024) and Zhang et al. (2024) introduced a member fusion mechanism improving cerebral infarction lesion segmentation with limited labels. Xu et al. (2024) developed a fuzzy C-means clustering algorithm with KL-divergence (Ichihashi et al., 2001) using a prior membership matrix reflecting expert preferences. Yu et al. (2024) created FW-SSPCM, enhancing adaptability to imbalanced features via supervised centers and weighted features. Third, optimizing the reliability of prior information: Gan et al. (2018) proposed LHC-S3FCM using localized chi-squared congruence and Gaussian kernel similarity to align labeled instances with local clusters. Gan (2019) introduced S3FCM, penalizing discrepancies with traditional FCM. Gan et al. (2019) enhanced this with CS3FCM by weighting prior information based on trustworthiness. Tuan et al. (2022) developed TS3MFCM, assigning defuzzification values per data point. Huan et al. (2022) created TS3FCM, reducing computation by decoupling labeling from clustering. Gan et al. (2024) proposed DaS3FCM, integrating discriminative learning and perceptual safety, using contour coefficients to estimate cluster confidence and FCM to correct labels, thereby enhancing prior information reliability.

Overall, although optimizing distance metrics, constraint conditions, prior information fusion strategies, and their reliability can improve semi-supervised clustering algorithms (Daneshfar et al., 2024), most methods rely on symmetric deviation between membership and prior membership in the regularization term. This weakens the guidance of prior information when current membership closely matches prior membership, degrading clustering performance. To address this, we propose a semi-supervised fuzzy C-means clustering with asymmetric deviation constraints (ADSFCM), extending the SFCM framework and incorporating collaborative fuzzy clustering principles (Pedrycz, 2022; Shen and Pedrycz, 2017). By introducing asymmetric deviation into the regularization term, ADSFCM overcomes structural limitations and enhances prior information utilization. Additionally, inspired by Kmita et al. (2024), a weighted adjustment factor is added to the fuzzy clustering term to better balance unsupervised and partially supervised information.

This paper's main contributions are:

● This paper enhances semi-supervised fuzzy C-means clustering by adding asymmetric deviation constraints to better utilize a priori information.

● The algorithm's convergence is confirmed by Zangwill's theorem, ensuring a solid theoretical foundation.

● An accelerated semi-supervised fuzzy C-means algorithm speeds convergence by filtering non-affinity centers and scaling membership, promoting its widespread application.

This paper is organized as follows: **Section 2** introduces fuzzy clustering and its semi-supervised variants. **Section 3** presents the semi-supervised fuzzy C-means algorithm with asymmetry deviation, its complexity, convergence, and an accelerated version using affinity filtering and membership scaling. **Section 4** evaluates the algorithm's performance, parameter sensitivity, time complexity, and efficiency against recent methods. **Section 5** concludes and suggests future research directions.

2. Related work

This section introduces fuzzy C-means and improved semi-supervised fuzzy clustering algorithms, supporting the paper's novel methods. **Table 1** lists key variables.

Table 1 Main symbols used in this paper

Notations	Descriptions
n	Number of data samples
c	Number of clusters
p	Number of features per data sample
$d(x_i, v_j)$	Euclidean distance between the sample x_i and the clustering center v_j
m	Fuzzy weighting exponent
b_i	Binary vector of a sample with supervised or non-supervised information
T	Number of iterations
β	Weighted adjustment factor for fuzzy clustering item
α	Regularized factor for semi-supervised clustering item
δ	Offset of clustering center change
$U \in \mathbb{R}^{c\times n}$	Membership partition matrix
$V \in \mathbb{R}^{c\times d}$	Clustering center matrix
$X \in \mathbb{R}^{n\times d}$	Sample dataset
$\overline{f}_{ij} \in \mathbb{R}^{c\times n}$	Semi-supervised membership degree information for samples

2.1. Fuzzy C-means clustering and variants

The FCM algorithm (Bezdek, 1984) is a soft clustering method that assigns data points to clusters with varying membership degrees by iteratively minimizing a weighted sum of squared Euclidean distances between samples and clustering center. The optimization model for FCM is as follows:

$$\min J_{FCM}(U,V;X)=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}d^{2}(x_i,v_j) \tag{1}$$

s.t. (a) $u_{ij}\in[0,1],\forall i,j$; (b) $\sum_{j=1}^{c}u_{ij}=1,\forall i$; (c) $0<\sum_{i=1}^{n}u_{ij}<n,\forall j$.

where $m>1$ (commonly set to 2) controls the fuzziness degree in clustering, with $m=1$ corresponding to hard C-means clustering. Let x_i be the i-th data sample in dataset $X=\{x_i\,|\,1\le i\le n\}$, v_j be the j-th clustering center; and u_{ij} be the membership degree of the i-th data sample x_i belongs to the j-th cluster. $d^2(x_i,v_j)=\|x_i-v_j\|_2^2$ is the squared Euclidean distance between the i-th data sample x_i and the j-th clustering center v_j. $U=[u_{ij}]_{c\times n}$ and $V=[v_j]_{c\times d}$ represent the fuzzy partition and clustering center matrices, respectively.

Using the Lagrange multiplier method, the unconstrained Lagrange function for minimizing **Eq. (1)** with constraint $\sum_{j=1}^{c}u_{ij}=1$ is formed, and iterative formulas for membership u_{ij} and the clustering center v_j are obtained by zeroing partial derivatives.

$$u_{ij}=\frac{1}{\sum_{k=1}^{c}\left(d^2(x_i,v_j)/d^2(x_i,v_k)\right)^{\frac{1}{m-1}}} \tag{2}$$

$$v_i=\frac{\sum_{i=1}^{n}u_{ij}^{m}\cdot x_i}{\sum_{i=1}^{n}u_{ij}^{m}} \tag{3}$$

The FCM objective function captures sample-cluster uncertainty for soft partitioning and, being differentiable, allows integrating prior information through regularization without losing its iterative optimization.

2.2. Semi-supervised fuzzy C-means clustering

Unsupervised fuzzy clustering is simple and effective, but lacks prior information use. Pedrycz and Waletzky (1997) addressed this by adding a semi-supervised regularization term to the FCM objective in the SFCM algorithm. The optimization model for SFCM is as follows:

$$\min J_{SFCM}(U,V)=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}d^{2}(x_i,v_j)+\alpha\sum_{j=1}^{c}\sum_{i=1}^{n}(u_{ij}-b_i\bar{f}_{ij})^{m}d^{2}(x_i,v_j) \tag{4}$$

s.t. (a) $u_{ij}\in[0,1],\forall i,j$; (b) $\sum_{j=1}^{c}u_{ij}=1,\forall i$; (c) $0<\sum_{i=1}^{n}u_{ij}<n,\forall j$; (d) $\bar{f}_{ij}\in[0,1],\forall i,j$.

where α is a regularization factor balancing unsupervised and partially supervised information, $\bar{f}_{ij}\in F$ is the prior membership degrees of x_i in the j-th cluster, $d^2(x_i,v_j)$ is the squared Euclidean distance between the i-th data sample and the j-th clustering center. Additionally, b_i is a binary vector with defined constraints.

$$\begin{cases} b_i = 1, & \text{if } x_i \text{ is a labeled sample} \\ b_i = 0, & \text{otherwise} \end{cases} \tag{5}$$

Using the Lagrange multiplier method, the unconstrained Lagrange function for minimizing **Eq. (4)** with constraint $\sum_{j=1}^{c} u_{ij} = 1$ is formed, and iterative formulas for membership u_{ij} and the clustering center v_j are obtained by zeroing partial derivatives.

$$u_{ij} = \frac{1}{1+\alpha^{1/(m-1)}} \cdot \frac{1+\alpha^{1/(m-1)}(1-b_i \sum_{j=1}^{c} \overline{f}_{ij})}{\sum_{k=1}^{c} (d^2(x_i, v_j)/d^2(x_i, v_k))^{1/(m-1)}} + \frac{\alpha^{1/(m-1)}}{1+\alpha^{1/(m-1)}} \cdot b_i \overline{f}_{ij} \tag{6}$$

$$v_j = \frac{\sum_{i=1}^{n} (u_{ij}^m + \alpha(u_{ij} - b_i \overline{f}_{ij})^m) x_i}{\sum_{i=1}^{n} (u_{ij}^m + \alpha(u_{ij} - b_i \overline{f}_{ij})^m)} \tag{7}$$

where $m = 2$.

Yasunori et al. (2009) introduced the Semi-Supervised Fuzzy C-Means (SSFCM) clustering algorithm, integrating prior membership information into the FCM objective function. Its optimization model is:

$$\min J_{SSFCM}(U, V) = \sum_{j=1}^{c} \sum_{i=1}^{n} (u_{ij} - b_i \overline{f}_{ij})^m d^2(x_i, v_j) \tag{8}$$

The model's constraints match those in **Eq. (4)**.

The clustering center v_j is iteratively derived from the unconstrained Lagrangian function in **Eq. (8):**

$$v_j = \frac{\sum_{i=1}^{n} (u_{ij} - b_i \overline{f}_{ij})^m x_i}{\sum_{i=1}^{n} (u_{ij} - b_i \overline{f}_{ij})^m} \tag{9}$$

When $m \neq 1$ and $1 - \sum_{j=1}^{c} \overline{f}_{ij} \geq 0$, the iterative formula for membership u_{ij} is:

$$u_{ij} = b_i \overline{f}_{ij} + \frac{1 - b_i \sum_{j=1}^{c} \overline{f}_{ij}}{\sum_{k=1}^{c} \left(d^2(x_i, v_j) \big/ d^2(x_i, v_k) \right)^{1/(m-1)}} \tag{10}$$

When $m = 1$, the iterative formula for membership u_{ij} is derived:

$$u_{ij} = \begin{cases} b_i \overline{f}_{ij} + 1 - b_i \sum_{j=1}^{c} \overline{f}_{ij}, & j = \arg\min_l \{d^2(x_i, v_l)\} \\ b_i \overline{f}_{ij}, & \text{otherwise} \end{cases} \tag{11}$$

When membership u_{ij} approaches prior supervised information $\overline{f}_{ij}$ under constraint symmetric deviation $(u_{ij} - b_i \overline{f}_{ij})$, semi-supervised regularization in SFCM and SSFCM weakens constraints of distance metric $d^2(x_i, v_j)$, reducing clustering accuracy.

To address this issue, Zhang et al. (2024) introduced the membership fusion semi-supervised fuzzy C-means (ICBF-MFSFCM) clustering algorithm for image segmentation, surpassing SFCM. Its semi-supervised aspect has clear physical significance and can be explained by an electrical analogy. The optimization model for ICBF-MFSFCM is:

$$\min J_{ICBF-MFSFCM}(U,V)=(1-\eta)\sum_{j=1}^{c}\sigma_j\sum_{i=1}^{n}u_{ij}^{m}d^2(x_i,v_j)+\eta\alpha\sum_{j=1}^{c}\sum_{i=1}^{n}b_i((u_{ij}-\bar{f}_{ij})^m+u_{ij}^m)d^2(x_i,v_j) \quad (12)$$

where η is the supervision rate and σ_j the interclass balance factor.

The model's constraints are the same as those in **Eq. (4)**.

Applying the Lagrange multiplier method, we derive iterative formulas for membership u_{ij} and the clustering center v_j under condition $m=2$.

$$u_{ij}=\frac{1-\eta\alpha b_i\sum_{k=1}^{c}\frac{\bar{f}_{ik}}{(1-\eta)\sigma_k+2\eta\alpha b_i}}{\sum_{k=1}^{c}\frac{((1-\eta)\sigma_j+2\eta\alpha b_i)d^2(x_i,v_j)}{((1-\eta)\sigma_k+2\eta\alpha b_i)d^2(x_i,v_k)}}+\frac{\eta\alpha b_i}{(1-\eta)\sigma_j+2\eta\alpha b_i}\bar{f}_{ij} \quad (13)$$

$$v_j=\frac{(1-\eta)\sigma_j\sum_{i=1}^{n}u_{ij}^2x_i+\eta\alpha\sum_{i=1}^{n}b_i((u_{ij}-\bar{f}_{ij})^2+u_{ij}^2)x_i}{(1-\eta)\sigma_j\sum_{i=1}^{n}u_{ij}^2+\eta\alpha\sum_{i=1}^{n}b_i((u_{ij}-\bar{f}_{ij})^2+u_{ij}^2)} \quad (14)$$

Without class weighting σ_j, the model simplifies as follows:

$$\min J_{ICBF-MFSFCM}(U,V)=(1-\eta)\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}d^2(x_i,v_j)+\eta\sum_{j=1}^{c}\sum_{i=1}^{n}b_i((u_{ij}-\bar{f}_{ij})^m+u_{ij}^m)d^2(x_i,v_j) \quad (15)$$

The iterative formulas for membership u_{ij} and cluster center v_j under condition $m=2$, are derived as follows:

$$u_{ij}=\frac{1-\frac{\eta b_i}{(1-\eta)+2\eta b_i}\sum_{k=1}^{c}\bar{f}_{ik}}{\sum_{k=1}^{c}d^2(x_i,v_j)/d^2(x_i,v_k)}+\frac{\eta b_i}{(1-\eta)+2\eta b_i}\bar{f}_{ij} \quad (16)$$

$$v_j=\frac{(1-\eta)\sum_{i=1}^{n}u_{ij}^2x_i+\eta\sum_{i=1}^{n}b_i((u_{ij}-\bar{f}_{ij})^2+u_{ij}^2)x_i}{(1-\eta)\sum_{i=1}^{n}u_{ij}^2+\eta\sum_{i=1}^{n}b_i((u_{ij}-\bar{f}_{ij})^2+u_{ij}^2)} \quad (17)$$

The ICBF-MFSFCM algorithm uses membership fusion to keep the semi-supervised regularization term comparable to unsupervised membership u_{ij}, ensuring stable and effective clustering.

3. The proposed method

This section presents a semi-supervised fuzzy clustering model with asymmetric deviation constraints, analyzes its complexity and convergence, and introduces an accelerated affinity filtering-based fuzzy C-means method with a membership scaling schedule.

3.1. Motivation

The introduction of SFCM (Pedrycz and Waletzky, 1997) established a theoretical basis for semi-supervised clustering, leading to many related optimization algorithms. However, the SFCM objective function's semi-supervised regularization term is typically a product involving a distance metric. Under the symmetric deviation constraint $(u_{ij} - b_i \bar{f}_{ij})^2$, when labeled samples' membership degrees approach their prior values, the constraint's influence on the distance metric $d^2(x_i, v_j)$ weakens. Since the distance metric depends on distances between samples and clustering centers, the guidance from prior information diminishes, reducing clustering accuracy. Although the ICBF-MFSFCM (Zhang et al., 2024) improves this via membership fusion—integrating $(u_{ij} - \bar{f}_{ij})^2 + u_{ij}^2$ and $d^2(x_i, v_j)$, it still underuses prior information when memberships align closely with priors, relying mainly on unsupervised clustering. Therefore, a new semi-supervised regularization term is needed.

3.2. Optimization modeling

Building on the ICBF-MFSFCM algorithm (Zhang et al., 2024), we introduce an asymmetric constraint to integrate $(2u_{ij} - 3\bar{f}_{ij})^2$ and $d^2(x_i, v_j)$ in semi-supervised clustering, enhancing performance with a new optimization model.

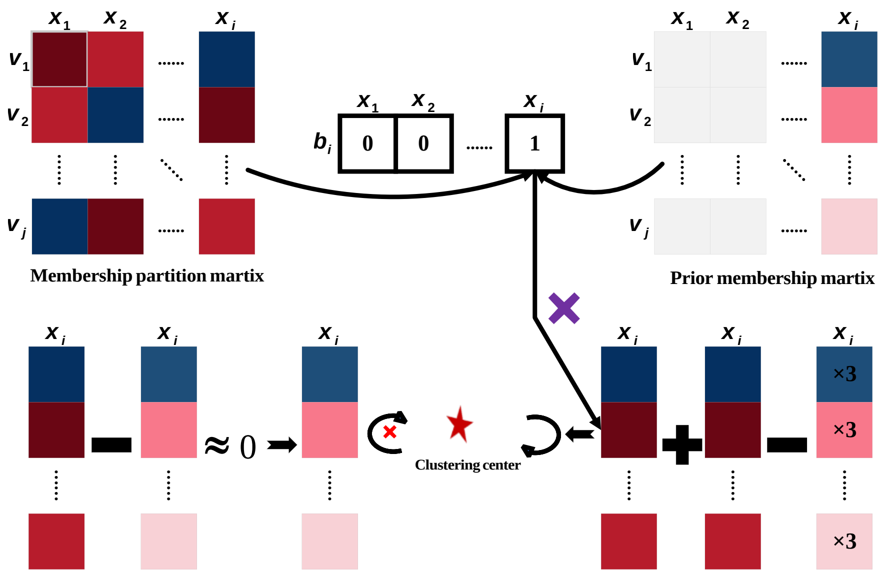


Fig 1. Block diagram of the new clustering optimization model

Fig. 1 illustrates that in the membership partition matrix, sample x_i belongs to the j-th cluster with varying membership u_{ij}. When the binary vector $b_i = 0$, the i-th column of the prior membership matrix, lacks prior information, it indicates the sample is unlabeled. Conversely, $b_i = 1$ means the i-th sample is labeled with prior membership information. If the current membership degree u_{ij} of sample x_i closely matches its prior membership $\bar{f}_{ij}$, the influence of prior information on guiding clustering centers weakens, potentially reducing clustering performance. To address this, an asymmetric deviation term $(2u_{ij}-3\bar{f}_{ij})^2$ is introduced based on the binary vector b_i. For the unlabeled sample x_i, the semi-supervised regularization term reduces to u_{ij}, matching the fuzzy clustering term's magnitude and aligning with unsupervised clustering. For labeled samples with $u_{ij} \approx \bar{f}_{ij}$, the regularization becomes the asymmetric deviation term $(2u_{ij}-3\bar{f}_{ij})^2$, ensuring prior information $\bar{f}_{ij}$ effectively guides clustering. This design preserves the core principle of semi-supervised clustering.

Due to limitations in semi-supervised fuzzy clustering algorithms (Despotovic et al.,2013), the fuzzy weighting exponent is fixed at 2 in the model design. The new clustering optimization model is constructed as follows:

$$\min J(U,V)=\beta\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^2 d^2(x_i,v_j)+\alpha\sum_{j=1}^{c}\sum_{i=1}^{n}(u_{ij}(1+b_i)-3b_i\bar{f}_{ij})^2 d^2(x_i,v_j) \tag{18}$$

The unconstrained Lagrangian combines the objective function and constraints via Lagrange multipliers.

$$\begin{aligned}L(U,V,\lambda)=&\beta\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^2 d^2(x_i,v_j)+\alpha\sum_{j=1}^{c}\sum_{i=1}^{n}(u_{ij}(1+b_i)-3b_i\bar{f}_{ij})^2 d^2(x_i,v_j)\\&+\sum_{i=1}^{n}\lambda_i(1-\sum_{j=1}^{c}u_{ij})\end{aligned} \tag{19}$$

Let $\dfrac{\partial L(U,V,\lambda)}{\partial u_{ij}}=0$,we have

$$u_{ij}=\frac{\lambda_i}{2\beta d^2(x_i,v_j)+2\alpha(1+b_i)^2 d^2(x_i,v_j)}+\frac{3\alpha b_i(1+b_i)}{\beta+\alpha(1+b_i)^2}\bar{f}_{ij} \tag{20}$$

Let $\dfrac{\partial L(U,V,\lambda)}{\partial \lambda_i}=0$,we have

$$\sum_{j=1}^{c}u_{ij}=1 \tag{21}$$

It follows that

$$\lambda_i=\left(1-\frac{3\alpha b_i(1+b_i)}{\beta+\alpha(1+b_i)^2}\sum_{j=1}^{c}\bar{f}_{ij}\right)\cdot\frac{1}{\sum_{j=1}^{c}\dfrac{1}{(2\beta+2\alpha(1+b_i)^2)d^2(x_i,v_j)}} \tag{22}$$

The iterative formula for fuzzy membership u_{ij} is derived by substituting **Eq. (22)** into **Eq. (20)**.

$$u_{ij}=\frac{1}{\beta+\alpha(1+b_i)^2}\cdot\frac{\beta+\alpha(1+b_i)^2-3\alpha b_i(1+b_i)\sum_{k=1}^{c}\bar{f}_{ik}}{\sum_{k=1}^{c}d^2(x_i,v_j)/d^2(x_i,v_k)}+\frac{3\alpha b_i(1+b_i)}{\beta+\alpha(1+b_i)^2}\bar{f}_{ij} \tag{23}$$

The iterative formula for the clustering center v_j is iteratively derived using $\dfrac{\partial L(U,V,\lambda)}{\partial v_j}=0$.

$$v_j = \frac{\beta\sum_{i=1}^{n} u_{ij}^2 x_i + \alpha\sum_{i=1}^{n}(u_{ij}(1+b_i) - 3b_i\bar{f}_{ij})^2 \cdot x_i}{\beta\sum_{i=1}^{n} u_{ij}^2 + \alpha\sum_{i=1}^{n}(u_{ij}(1+b_i) - 3b_i\bar{f}_{ij})^2} \tag{24}$$

The new model can provide clearer interpretations via circuit theory (Bezdek, 1993). Without supervised information $b_i = 0 (i = 1, 2, \cdots, n)$, **Eq. (18)** reduces to the standard FCM algorithm, making FCM a special case of the semi-supervised fuzzy clustering method presented.

3.3. Asymmetric deviation-based semi-supervised fuzzy clustering algorithm

We present a two-level iterative algorithm based on membership and cluster center formulas, **Eq. (23)** and **Eq. (24)**. The pseudocode for the asymmetric deviation-based semi-supervised fuzzy clustering algorithm follows.

Algorithm: ADSFCM

Input: Data set $X = \{x_1, x_2, ..., x_n\}$, cluster number c, convergence threshold ε, weight coefficient β and α, prior membership degree $\bar{f}_{ij}$, binary vector b_i, maximum iterations T.

Output: The clustering center matrix V, membership partition matrix U;

Initialization: The clustering centers $v_j^{(0)} (j = 1, 2, \cdots, c)$ and set t=0;

For $i = 1$ to T **do**

Step 1. Compute $u_{ij}^{(t)}$ with **Eq. (23)**;

Step 2. Update $v_j^{(t+1)}$ with **Eq. (24)**;

Step 3. **if** $\left\|V^{(t+1)} - V^{(t)}\right\|_F^2 > \varepsilon$ and $t > T$ **then**

Step 4. Set t: =t+1;

Step 5. Goto Step 1;

Step 6. **else**

Step 7 Return $V = V^{(t+1)}$;

Step 8 **end if**

end for

Algorithmic complexity analysis (Kolen and Vansteenkiste, 2002) assesses performance by revealing worst-case resource needs and efficiency. For ADSFCM, updating the membership matrix has complexity $O(ncTd)$ (n sample size, c clusters, T iterations, and d features), and updating clustering centers adds a complexity of $O(ncTd)$, resulting in overall complexity $O(ncTd)$.

3.4. Algorithm convergence analysis

Algorithm convergence ensures effectiveness and reliability. Saha and Das (2018) used Zangwill's theorem for center-based clustering convergence analysis; this subsection applies it to the ADSFCM algorithm. Convergence analysis requires the ADSFCM algorithm's objective function to decrease monotonically and be bounded, ensuring convergence to a fixed point by Zangwill's theorem. ADSFCM's continuity and tightness constraints ensure convergence if the objective function decreases.

According to the ADSFCM optimization model, we define the unconstrained objective function as follows:

$$L(U,V,\lambda)=\beta\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{2}d^{2}(x_i,v_j)+\alpha\sum_{j=1}^{c}\sum_{i=1}^{n}(u_{ij}(1+b_i)-3b_i\overline{f}_{ij})^{2}d^{2}(x_i,v_j)+\sum_{i=1}^{n}\lambda_i(1-\sum_{j=1}^{c}u_{ij}) \tag{25}$$

A function $L(U,V,\lambda)$ is decreasing if it meets the following two propositions.

Proposition 1: When clustering center matrix V is given, let $F(U)=L(U,V,\lambda)$, u_{ij}^{*} is a minimum point of $F(U)$ if and only if u_{ij}^{*} comes from the ADSFCM iterative membership formula.

Proof: If **Proposition 1**'s condition holds, then u_{ij}^{*} is a minimum of $F(U)$, and the following equation holds:

$$\frac{\partial F(U)}{\partial u_{ij}}=2\beta u_{ij}d^{2}(x_i,v_j)+2\alpha(1+b_i)(u_{ij}(1+b_i)-3b_i\overline{f}_{ij})d^{2}(x_i,v_j)-\lambda_i=0 \tag{26}$$

Further, we obtain

$$\lambda_i=2u_{ij}d^{2}(x_i,v_j)+2\alpha(1+b_i)(u_{ij}(1+b_i)-3b_i\overline{f}_{ij})d^{2}(x_i,v_j) \tag{27}$$

So that,

$$u_{ij}=\frac{\lambda_i}{2\beta d^{2}(x_i,v_j)+2\alpha(1+b_i)^{2}d^{2}(x_i,v_j)}+\frac{3\alpha b_i(1+b_i)}{\beta+\alpha(1+b_i)^{2}}\overline{f}_{ij} \tag{28}$$

Substituting **Eq. (28)** into the constraint $(1-\sum_{j=1}^{c}u_{ij})=0$, we obtain

$$\lambda_i=\left(1-\frac{3\alpha b_i(1+b_i)}{\beta+\alpha(1+b_i)^{2}}\sum_{j=1}^{c}\overline{f}_{ij}\right)\cdot\frac{1}{\sum_{j=1}^{c}\frac{1}{(2\beta+2\alpha(1+b_i)^{2})d^{2}(x_i,v_j)}} \tag{29}$$

Substituting **Eq. (29)** into **Eq. (28)**, we have

$$u_{ij}^{*}=\left(1-\frac{3\alpha b_i(1+b_i)}{\beta+\alpha(1+b_i)^{2}}\sum_{k=1}^{c}\overline{f}_{ik}\right)\frac{1}{\sum_{k=1}^{c}d^{2}(x_i,v_j)/d^{2}(x_i,v_k)}+\frac{3\alpha b_i(1+b_i)}{\beta+\alpha(1+b_i)^{2}}\overline{f}_{ij} \tag{30}$$

Thus, the Hessian matrix of $F(u_{ij})$ at the minimum point u_{ij}^{*} is

$$\frac{\partial}{\partial u_{ab}}\left(\frac{\partial F(U)}{\partial u_{ij}}\right)=\begin{cases}0, & a\neq i \text{ or } b\neq j\\ 2\beta d^{2}(x_i,v_j)+2\alpha(1+b_i)^{2}d^{2}(x_i,v_j), & a=i \text{ and } b=j\end{cases} \tag{31}$$

Eq. (31) shows that the Hessian matrix is positive definite at fuzzy weighting exponent $m=2$, indicating a minimum point u_{ij}^{*} of the objective function and stability optimization.

Confirming a strict local minimum requires analyzing the Jacobian and Hessian matrices. For constrained optimization, the bordered Hessian matrix (Yang and Tian, 2015), combining second-order derivatives and Lagrange multipliers, provides a complete assessment.

When V is given, the bordered Hessian matrices corresponding to $u_i=[u_{i1},u_{i2},...,u_{ic}]$ and λ_i are:

$$H_L(u_i,\lambda_i)=\begin{bmatrix} 0 & \frac{\partial^2 L}{\partial\lambda_i\partial u_{i1}} & \cdots & \frac{\partial^2 L}{\partial\lambda_i\partial u_{ic}} \\ \frac{\partial^2 L}{\partial\lambda_i\partial u_{i1}} & \frac{\partial^2 L}{\partial u_{i1}\partial u_{i1}} & \cdots & \frac{\partial^2 L}{\partial u_{i1}\partial u_{ic}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial\lambda_i\partial u_{ic}} & \frac{\partial^2 L}{\partial u_{ic}\partial u_{i1}} & \cdots & \frac{\partial^2 L}{\partial u_{ic}\partial u_{ic}} \end{bmatrix} \tag{32}$$

where $\frac{\partial^2 L}{\partial\lambda_i\partial u_{ik}}=-1, 1\le k\le c$.

u_{ij}^* is a local minimum point if the bordered Hessian matrix's principal minors satisfy the conditions.

$$\begin{aligned} \left|\bar{H}_3(u_i^*,\lambda_i^*)\right| &= \det\begin{bmatrix} 0 & \frac{\partial^2 L}{\partial\lambda_i\partial u_{i1}} & \frac{\partial^2 L}{\partial\lambda_i\partial u_{i2}} \\ \frac{\partial^2 L}{\partial\lambda_i\partial u_{i1}} & \frac{\partial^2 L}{\partial u_{i1}\partial u_{i1}} & \frac{\partial^2 L}{\partial u_{i1}\partial u_{i2}} \\ \frac{\partial^2 L}{\partial\lambda_i\partial u_{i2}} & \frac{\partial^2 L}{\partial u_{i2}\partial u_{i1}} & \frac{\partial^2 L}{\partial u_{i2}\partial u_{i2}} \end{bmatrix} \\ &= -\left(\frac{\partial^2 L}{\partial\lambda_i\partial u_{i2}}\right)^2 \frac{\partial^2 L}{\partial u_{i1}\partial u_{i1}} - \left(\frac{\partial^2 L}{\partial\lambda_i\partial u_{i1}}\right)^2 \frac{\partial^2 L}{\partial u_{i2}\partial u_{i2}} \\ &= \left(-2(\beta+\alpha(1+b_i)^2)(d^2(x_i,v_1)+d^2(x_i,v_2))\right)_{u_i=u_i^*,\lambda_i=\lambda_i^*} < 0 \end{aligned} \tag{33}$$

$$\begin{aligned} \left|\bar{H}_4(u_i^*,\lambda_i^*)\right| &= \det\begin{bmatrix} 0 & \frac{\partial^2 L}{\partial\lambda_i\partial u_{i1}} & \frac{\partial^2 L}{\partial\lambda_i\partial u_{i2}} & \frac{\partial^2 L}{\partial\lambda_i\partial u_{i3}} \\ \frac{\partial^2 L}{\partial\lambda_i\partial u_{i1}} & \frac{\partial^2 L}{\partial u_{i1}\partial u_{i1}} & \frac{\partial^2 L}{\partial u_{i1}\partial u_{i2}} & \frac{\partial^2 L}{\partial u_{i1}\partial u_{i3}} \\ \frac{\partial^2 L}{\partial\lambda_i\partial u_{i2}} & \frac{\partial^2 L}{\partial u_{i2}\partial u_{i1}} & \frac{\partial^2 L}{\partial u_{i2}\partial u_{i2}} & \frac{\partial^2 L}{\partial u_{i2}\partial u_{i3}} \\ \frac{\partial^2 L}{\partial\lambda_i\partial u_{i3}} & \frac{\partial^2 L}{\partial u_{i3}\partial u_{i1}} & \frac{\partial^2 L}{\partial u_{i3}\partial u_{i2}} & \frac{\partial^2 L}{\partial u_{i3}\partial u_{i3}} \end{bmatrix} \\ &= -\sum_{k=1}^{3}\left(\frac{\partial^2 L}{\partial\lambda_i\partial u_{ik}}\right)^2 \cdot \prod_{k_1=1,k_1\ne k}^{3} \frac{\partial^2 L}{\partial u_{ik_1}\partial u_{ik_1}} \\ &= \left(-4(\beta+\alpha(1+b_i)^2)^2 \sum_{k=1}^{3}\prod_{k_1=1,k_1\ne k}^{3} d^2(x_i,v_{k_1})\right)_{u_i=u_i^*,\lambda_i=\lambda_i^*} < 0 \end{aligned} \tag{34}$$

Proven by induction:

$$\left|\bar{H}_{c+1}(u_i^*,\lambda_i^*)\right| = \left(-\sum_{k=1}^{c}\left(\frac{\partial^2 L}{\partial\lambda_i\partial u_{ik}}\right)^2 \cdot \prod_{k_1=1,k_1\ne k}^{c}\frac{\partial^2 L}{\partial u_{ik_1}\partial u_{ik_1}}\right)_{u_i=u_i^*,\lambda_i=\lambda_i^*} < 0 \tag{35}$$

Thus, the bordered Hessian matrix theorem states that a local minimum point u_{ij}^* exists when the fuzzy membership satisfies the constraint $(1-\sum_{j=0}^{c}u_{ij})=0$.

Proposition 2: Given a fuzzy partition matrix $U=[u_{ij}]_{n\times c}$, v_j^* is a minimum point of

$F(V)=L(U,V,\lambda)$ if and only if v_j^* comes from the ADFCM clustering center iterative formula.

Proof: If **Proposition 2**'s condition holds, v_j^* is a minimum of $F(V)$, leading to the following equation:

$$\frac{\partial F(V)}{\partial v_j}=2\beta\sum_{i=1}^{n}u_{ij}^2(x_i-v_j)+2\alpha\sum_{i=1}^{n}((1+b_i)u_{ij}-3b_i\bar{f}_{ij})^2(x_i-v_j)=0 \tag{36}$$

It further follows that

$$v_j=\frac{\beta\sum_{i=1}^{n}u_{ij}^2x_i+\alpha\sum_{i=1}^{n}(u_{ij}(1+b_i)-3b_i\bar{f}_{ij})^2\cdot x_i}{\beta\sum_{i=1}^{n}u_{ij}^2+\alpha\sum_{i=1}^{n}(u_{ij}(1+b_i)-3b_i\bar{f}_{ij})^2} \tag{37}$$

Thus, the Hessian matrix of $F(V)$ at the minimum point v_j^* is

$$\frac{\partial}{\partial v_l}\left(\frac{\partial F(V)}{\partial v_j}\right)=\begin{cases}0, & l\neq j\\ -2\beta\sum_{i=1}^{n}u_{ij}^2-2\alpha\sum_{i=1}^{n}((1+b_i)u_{ij}-3b_i\bar{f}_{ij})^2, & l=j\end{cases} \tag{38}$$

With a fuzzy weighting exponent of 2, **Eq. (38)** shows the Hessian matrix is positive definite, having positive diagonal and zero off-diagonal elements.

The ADSFCM algorithm's objective function steadily decreases, ensuring local convergence to a minimum from any start point, supported by continuity and compactness constraints similar to the FCM algorithm by Zangwill's theorem.

3.5. Accelerated semi-supervised fuzzy clustering with membership scaling

In semi-supervised fuzzy C-means clustering, clustering centers and the membership matrix are updated iteratively, involving all data points. If an updated clustering center v_j is not the nearest to a data sample x_i, it becomes that sample's non-affinity center. Updating the membership matrix requires calculating the Euclidean distance between cluster prototypes and data points, which slows convergence. Cai et al. (2009) proposed a fast SFCM algorithm that, despite its effectiveness, lacks theoretical support and violates membership constraints. Antoine et al. (2021) introduced a rapid semi-supervised evidential clustering framework using a heuristic to relax credal partition constraints. Zhou et al. (2021) enhanced accuracy and convergence by leveraging prior pairwise constraints and the IFO algorithm. However, non-affinity centers still demand substantial computation. Building on Li et al.'s (2023) fast FCM algorithm, this paper introduces a non-affinity center identification method and a semi-supervised membership scaling scheme, yielding an accelerated asymmetric deviation constrained SFCM algorithm (AM-ADSFCM).

(1) **Non-affinity center identification**

The modified fuzzy C-means algorithm struggles to identify non-affinity centers in semi-supervised clustering. We propose a new triangular inequality in **Eq. (39)** (Ding et al., 2015) to improve this. Let $\mathcal{P}_i$ denote the set of non-affinity centers of x_i, where $i=1,2,\cdots,n$. **Fig. 2** illustrates the geometric interpretation of these non-affinity centers for sample x_i.

Fig. 2 shows sample x_i randomly selected from the dataset. Clustering centers v_1, v_2, and v_3

(with the number of clusters predefined as $c=3$) represent clustering centers after iterations. Distances from x_i to these centers are arranged in ascending order as v_1, v_2, and v_3. Purple dots mark possible next center positions; δ indicates positional change between centers before and after iteration. Specifically, $\delta_{I_i^*}$ refers to the displacement between the affinity clustering centers of x_i. Red, green, and blue arcs represent upper and lower limits of position changes for v_1 ($\max\delta_{I_i^*}$), v_2 ($\max\delta_2$), and v_3 ($\max\delta_3$), respectively.

For the proposed ADSFCM algorithm, when $j\in\mathcal{P}_i$, convergence and efficiency issues arise during optimization: According to **Eq. (23)**, if $d^2_{j\in\mathcal{P}_i}(x_i,v_j)\gg d^2_{j\notin\mathcal{P}_i}(x_i,v_j)$, a data sample's membership to a non-affinity clustering center should be zero. However, with the constraint $\sum_{j=1}^{c}u_{ij}=1$, a small residual value τ remains, compressing the membership to its affinity center as $1-\tau$ and weakening membership discriminability. In **Eq. (24)**, the residual τ introduces a deviation term $\tau(x_i-v_j)$ during clustering center updates, even when $d^2_{j\in\mathcal{P}_i}(x_i,v_j)\gg d^2_{j\notin\mathcal{P}_i}(x_i,v_j)$. Since the convergence condition is defined by $\left\|V^{(t+1)}-V^{(t)}\right\|<\varepsilon$, this deviation slows convergence and reduces iteration efficiency. Per **Eq. (18)**, the distance $d^2_{j\in\mathcal{P}_i}(x_i,v_j)$ should be large, but $\sum_{j=1}^{c}u_{ij}=1$ forces the membership degree u_{ij} to a small positive value near zero, causing many ineffective terms $\sum_{j\in\mathcal{P}_i}u_{ij}^2d^2(x_i,v_j)$ to accumulate in the objective function. This accumulation degrades convergence from exponential to linear, significantly slowing the process. In later iterations, although most memberships stabilize, distances between samples and non-affinity centers $d^2_{j\in\mathcal{P}_i}(x_i,v_j)$ still require frequent calculation, increasing computational overhead. Thus, optimizing non-affinity clustering centers is crucial to speed convergence and improve efficiency.

Demonstrate that clustering center v_j is a non-affinity center of sample x_i if and only if the equation below holds.

$$d_{ij}^{(t)}-\delta_j^{(t)}\geq D_i^{(1)(t)}+\delta_{I_i^*}^{(t)},\ j\in\{1,2,\cdots,c\} \tag{39}$$

where, $I_i^*=\arg\min_{1\leq j\leq c}\{d_{ij}\}$.

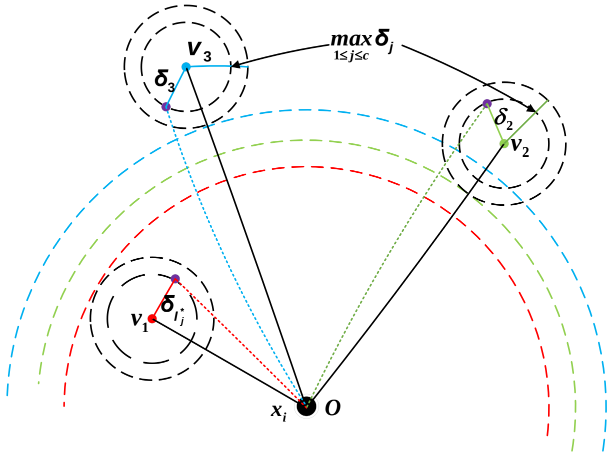


Fig. 2. Geometric interpretation of non-affinity centers for sample x_i

The proof is as follows:

Expanding **Eq. (39)**, for $j \in \{1, 2, \cdots, c\}$ there is

$$d_{ij}^{(t)} - \delta_j^{(t)} = \left\| x_i - v_j^{(t)} \right\| - \left\| v_j^{(t+1)} - v_j^{(t)} \right\| \le d_{ij}^{(t+1)} \tag{40}$$

Similarly, there are

$$D_i^{(1)(t)} + \delta_{I_i^*}^{(t)} = \left\| x_i - v_{I_i^*}^{(t)} \right\| + \left\| v_{I_i^*}^{(t)} - v_{I_i^*}^{(t+1)} \right\| \ge d_{I_i^* i}^{(t+1)} \tag{41}$$

If **Eq. (39)** holds, $d_{ij}^{(t+1)} \ge d_{I_i^* i}^{(t+1)}$. Therefore, it follows that after updating the clustering centers, v_j cannot be the nearest center to sample x_i, so v_j is a non-affinity center of x_i.

(2) **Membership scaling**

Following the above discussion, the set of non-affinity centers for sample x_i can be obtained using **Eq. (39)**. To improve clustering performance and speed up convergence, this paper introduces a semi-supervised membership scaling technique. This method reduces the influence of non-affinity clustering centers during updates by adjusting membership degrees to better reflect the similarity of clustering centers in the original scheme, assuming the non-affinity centers of sample x_i are known. If v_j is a non-affinity center of x_i, then x_i is called a non-affinity sample of v_j. In this method, the membership degree of sample x_i to a non-affinity clustering center v_j is replaced by the prior membership degree or set to zero if sample x_i is unlabeled. Meanwhile, membership degrees for sample x_i affiliated with affinity clustering centers continue to be updated normally. The memberships are then renormalized for the next iteration, significantly improving clustering efficiency. The membership scaling method is outlined below:

$$\hat{u}_{ij}^{(t)} = \begin{cases} \bar{f}_{ij}, & b_i = 1 \text{ and } j \in \mathcal{P}_i^{(t)} \\ u_{ij}^{(t)}, & j \notin \mathcal{P}_i^{(t)} \\ 0, & b_i = 0 \text{ and } j \in \mathcal{P}_i^{(t)} \end{cases} \tag{42}$$

where $u_{ij}^{(t)}$ is the membership value of sample x_i in the j-th cluster during the t-th iteration of a semi-supervised fuzzy clustering algorithm; $\mathcal{P}_i$ is the set of non-affinity centers of x_i, where

$i=1,2,\cdots,n$.

We propose an accelerated semi-supervised fuzzy C-means clustering method that speeds convergence by adjusting membership degrees $u_{ij}^{(t)}$ to $\hat{u}_{ij}^{(t)}$ and renormalizing $\hat{u}_{ij}^{(t)}$ to $\tilde{u}_{ij}^{(t)}$ using **Eq. (43)**. The framework is shown in **Fig. 3**.

$$\tilde{u}_{ij}^{(t)}\begin{cases}\overline{f}_{ij}, & b_i=1 \text{ and } j\in\mathcal{P}_i^{(t)}\\ u_{ij}^{(t)}*\dfrac{1-\sum_{k\in\mathcal{P}_i^{(t)}}\hat{u}_{ik}^{(t)}}{1-\sum_{k\in\mathcal{P}_i^{(t)}}u_{ik}^{(t)}}, & j\notin\mathcal{P}_i^{(t)}\\ 0, & b_i=0 \text{ and } j\in\mathcal{P}_i^{(t)}\end{cases} \tag{43}$$

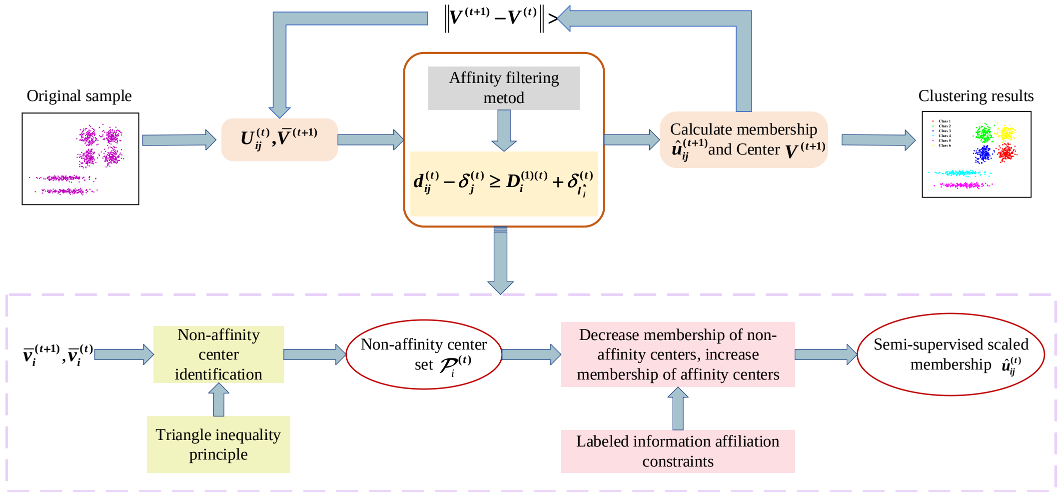


Fig. 3. The main framework of fast semi-supervised fuzzy clustering.

Algorithm: Accelerated semi-supervised fuzzy clustering with membership scaling

---

Input: Date set $X=\{x_1,x_2,\cdots,x_n\}$, cluster number c, convergence threshold ε, weight coefficient β and α, prior membership degree $\overline{f}_{ij}$, binary vector b_i, maximum iterations T;

Output: The clustering center matrix V, fuzzy partition matrix U;

Initialization: The clustering centers $V^{(0)}$ and set t=0;

Repeat

Step 1: Calculate $u_{ij}^{(t)}$ with **Eq. (23)**;

Step 2: Update $\overline{V}^{(t+1)}$ with **Eq. (24)**;

Step 3: Calculate $\delta_j^{(t)}=\left\|\overline{v}_j^{(t+1)}-v_j^{(t)}\right\|$ for $j=1,2,\cdots,c$;

Step 4: **for** $i=1$ to n **do**

Step 5: $\mathcal{P}_i^{(t)}=\{1\leq j\leq c\mid d_{ij}^{(t)}-\delta_j^{(t)}\geq D_i^{(1)(t)}+\delta_{I_i^*}^{(t)}\}$;

Step 6: Calculate $\hat{u}_{ij}^{(t)}$ using **Eq. (42)**;

Step 7: Calculate $\tilde{u}_{ij}^{(t)}$ using **Eq. (43)**;

Step 8: **end for**

Step 9: Update $V^{(t+1)}$ with **Eq. (24)**;

Step 10: **If** $\left\|V^{(t+1)}-V^{(t)}\right\|<\varepsilon$ or $t>T$

Step 11: Set t=t+1;

Step 12: Goto Step 1;

Step 13: **else**

Step 14: Return $V = V^{(t+1)}$;

Step 15: **end if**

End

4. Experimental results and analysis

We evaluate ADSFCM's clustering by comparing it with four classical semi-supervised fuzzy clustering algorithms (SFCM (Pedrycz and Waletzky, 1997), SSFCM (Yasunori et al., 2009), ESFCM (Yasunori et al., 2009), SMUC (Yin et al.,2012)) and two recent improvements (SFCM-EP (Xu et al., 2024), ICBF-MFSFCM (Zhang et al., 2024)) across four domains: numerical data, natural images, medical images, and remote sensing. Dataset sources appear in **Table 2**.

Table 2 Dataset and its website

Datasets	Sources
Numerical datasets	https://github.com/milaan9/Clustering-Datasets
BSDS500	https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html#bsds500
Brian-tumor-dataset	https://www.kaggle.com/preetviradiya/brian-tumor-dataset
UC Merced Land Use Dataset	http://weegee.vision.ucmerced.edu/datasets/landuse.html
SIRI_WHU_Dataset	https://figshare.com/articles/dataset/SIRI_WHU_Dataset/8796980

The evaluation employs metrics like Accuracy (ACC), Precision (PE), Recall (RE), Adjusted Rand Index (ARI), Normalized Mutual Information (NMI), Jaccard Index, and Mean Intersection over Union (mIoU) (Mittal et al., 2022) to assess clustering accuracy, consistency, purity, and segmentation quality.

4.1. Parameter setting

This paper ensures experimental validity by using literature-recommended parameter ranges. Tuning parameters for six algorithms are validated from prior studies, with the fuzzy weighting exponent m fixed at 2 for optimal balance.

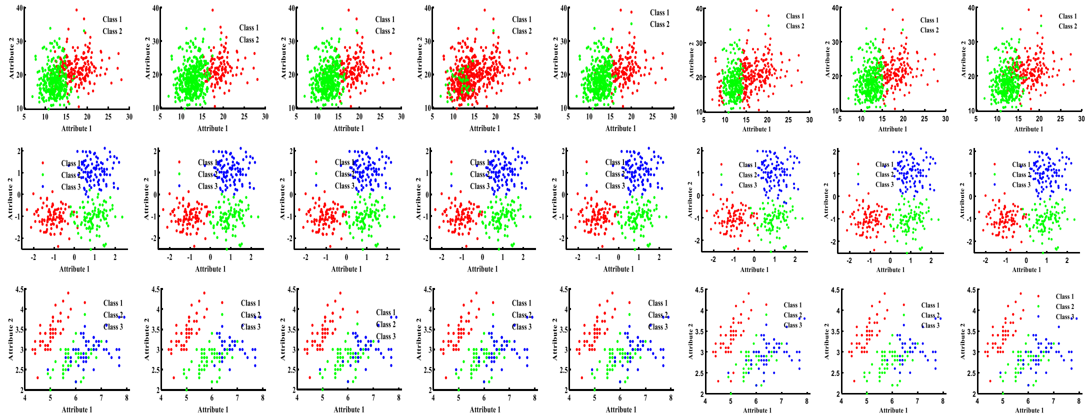
This paper uses variable control and fixed parameters ($\beta = 4$ and $\alpha = 2$) in ADSFCM to ensure unbiased, accurate algorithm comparisons.

Semi-supervised learning algorithms for clustering and image segmentation are evaluated using 10%-30% randomly labeled data points, whose prior membership degrees $\bar{f}_{ij}$ guide the iterative learning.

4.2 Performance testing and analysis

(1) **Numerical data sets**

Uniformly distributed data: Data points are uniformly distributed, showing independence and a lack of distinct clusters. Algorithms are evaluated on Wdbc, Blobs, and Iris datasets, with results in **Fig. 4** and **Table 3**.



(a) (b) (c) (d) (e) (f) (g) (h)

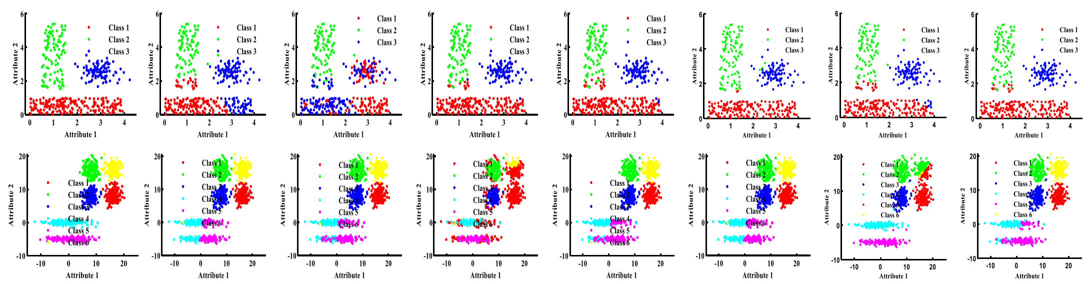
Fig. 4. Clustering results for algorithms on uniformly distributed numerical data. (a) Ground truth; (b) SFCM; (c) SSFCM; (d) ESFCM; (e) SMUC; (f) SFCM-EP; (g) ICBF-MFSFCM; (h) ADSFCM. From top to bottom: Wdbc, Blobs, and Iris.

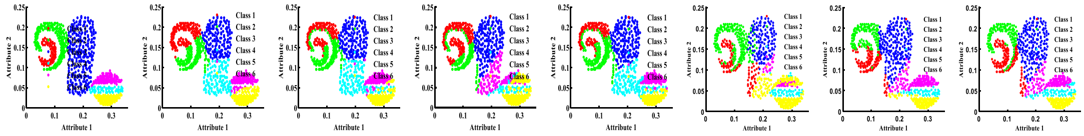
Table 3 Evaluation of seven algorithms on uniformly distributed numerical data

Data set	Indicator	SFCM	SSFCM	ESFCM	SMUC	SFCM-EP	ICBF-MFSFCM	ADSFCM
Wdbc (n=569, d=3 c=2)	ACC	0.8963	0.8928	0.6274	0.9086	0.8383	0.9016	**0.9262**
	PE	0.9267	0.9223	0.7992	0.9302	0.8327	0.9277	**0.9359**
	RE	0.8618	0.8580	0.5000	0.8802	0.8549	0.8698	**0.9153**
	ARI	0.6227	0.6114	-0.0376	0.6636	0.4568	0.6401	**0.7245**
	NMI	0.5697	0.5515	0.0767	0.5883	0.3924	0.5755	**0.6066**
Blobs (n=300, d=2, c=3)	ACC	0.9667	0.9700	0.9700	0.9733	0.9667	0.9700	**0.9833**
	PE	0.9667	0.9701	0.9700	0.9733	0.9683	0.9701	**0.9833**
	RE	0.9667	0.9700	0.9700	0.9733	0.9667	0.9700	**0.9833**
	ARI	0.9028	0.9122	0.9124	0.9217	0.9032	0.9123	**0.9505**
	NMI	0.8600	0.8726	0.8727	0.8828	0.8790	0.8727	**0.9193**
Iris (n=150, d=4, c=3)	ACC	0.9467	0.9267	0.9667	0.9400	0.9600	0.9200	**0.9800**
	PE	0.9508	0.9354	0.9678	0.9455	0.9625	0.9306	**0.9811**
	RE	0.9467	0.9267	0.9667	0.9400	0.9600	0.9200	**0.9800**
	ARI	0.8512	0.8026	0.9038	0.8345	0.8843	0.7874	**0.9410**
	NMI	0.8449	0.8126	0.8851	0.8334	0.8635	0.8031	**0.9306**

Fig. 4 and **Table 3** show that ADSFCM outperforms six algorithms on three datasets, achieving 5% higher accuracy than SMUC and strong results in PE, RE, ARI, and NMI, demonstrating its robustness.

Multimodally distributed data: Data points form distinct clusters with unique distributions, revealing complex structures. This multi-modal distribution aids clustering analysis. **Fig. 5** and **Table 4** present algorithm evaluations on the Lsun, Longsquare, and Ds3c3c6 datasets.





(a) (b) (c) (d) (e) (f) (g) (h)

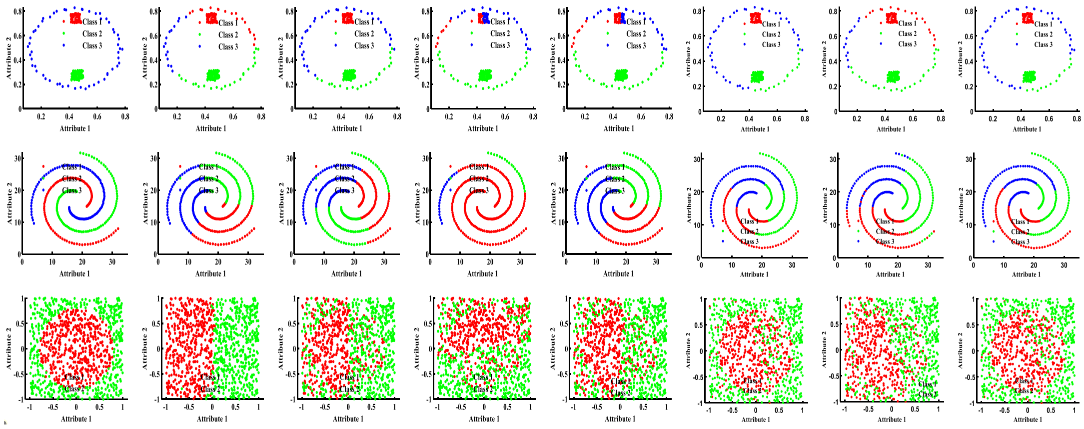
Fig. 5. Clustering results for algorithms on multimodally distributed numerical data. (a) Ground truth; (b) SFCM; (c) SSFCM; (d) ESFCM; (e) SMUC; (f) SFCM-EP; (g) ICBF-MFSFCM; (h) ADSFCM. From top to bottom: Lsun, Longsquare, and Ds3c3c6.

Table 4 Evaluation of seven algorithms on multi-modal distribution data

Data set	Indicator	SFCM	SSFCM	ESFCM	SMUC	SFCM-EP	ICBF-MFSFCM	ADSFCM
Lsun (n=400, d=4, c=3)	ACC	0.9400	0.7050	0.9650	0.9525	0.9400	0.9575	**0.9800**
	PE	0.9482	0.7499	0.9745	0.9653	0.9528	0.9632	**0.9820**
	RE	0.9350	0.6067	0.9550	0.9417	0.9300	0.9533	**0.9783**
	ARI	0.8199	0.2851	0.8956	0.8572	0.8215	0.8699	**0.9371**
	NMI	0.8017	0.3949	0.8715	0.8418	0.8067	0.8459	**0.9128**
Longsquare (n=900, d=2, c=6)	ACC	0.8567	0.8778	0.6533	0.9011	0.8389	0.8956	**0.9522**
	PE	0.8567	0.8778	0.8874	0.9008	0.8903	0.8940	**0.9519**
	RE	0.8567	0.8778	0.6502	0.9011	0.8357	0.8956	**0.9522**
	ARI	0.7962	0.8075	0.3121	0.8262	0.7688	0.8183	**0.8964**
	NMI	0.8632	0.8685	0.5887	0.8735	0.8555	0.8657	**0.9056**
Ds3c3s6 (n=905, d=2, c=6)	ACC	0.8243	0.8254	0.7105	0.8033	0.8055	0.8243	**0.8409**
	PE	0.8238	0.8234	0.6841	0.8065	0.7942	0.8195	**0.8436**
	RE	0.6479	0.6465	0.5882	0.6252	**0.8591**	0.6462	0.6603
	ARI	0.5781	0.5797	0.4732	0,5600	**0.6236**	0.5730	0.6070
	NMI	0.6993	0.6983	0.6012	0.6768	0.7139	0.6943	**0.7361**

Fig. 5 shows that six algorithms have significant clustering errors from incorrect cluster assignments in three multimodal datasets. ADSFCM, despite minor misclassifications, outperforms them all. **Table 4** confirms ADSFCM excels across evaluation metrics, showing superior robustness and generalization with complex multimodal data.

Non-spherically distributed data: Data points form complex, non-spherical structures like rings, chains, and lines. **Fig. 6** and **Table 5** present algorithm clustering and evaluations on the Zelnik6, Spiral, and Circle datasets.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 6. Clustering results of algorithms on non-spherical data. (a) Ground truth; (b) SFCM; (c) SSFCM; (d) ESFCM; (e) SMUC; (f)SFCM-EP; (g) ICBF-MFSFCM; (h) ADSFCM. From top to bottom: Zelnik6, Spiral, and Circle.

Table 5 Evaluation of seven algorithms on non-spherical numerical data

Data set	Indicator	SFCM	SSFCM	ESFCM	SMUC	SFCM-EP	ICBF-MFSFCM	ADSFCM
Zelnik6 (n=238, d=2, c=3)	ACC	0.8277	0.7353	0.7647	0.7647	0.8025	0.8529	**0.8950**
	PE	0.8777	0.8107	0.7612	0.7338	0.7786	0.8961	**0.9333**
	RE	0.7560	0.6765	0.6667	0.6667	0.7316	0.7917	**0.8512**
	ARI	0.6475	0.3839	0.5101	0.5480	0.6113	0.6813	**0.7541**
	NMI	0.6403	0.3494	0.5224	0.5471	0.5789	0.6763	**0.7922**
Spiral (n=312, d=2, c=3)	ACC	0.3397	0.4359	0.4744	0.3397	0.4872	0.3494	**0.4912**
	PE	0.3395	0.4347	0.7975	0.3395	0.4874	0.3493	**0.4896**
	RE	0.3333	0.4371	0.4693	0.3333	0.4875	0.3438	**0.4888**
	ARI	-0.0062	0.0271	0.0323	-0.0059	0.0518	-0.0051	**0.0503**
	NMI	0.0002	0.0303	0.1947	0.0005	0.0514	0.0012	**0.0538**
Cirle (n=1000, d=2 c=2)	ACC	0.5210	0.6990	0.6910	0.7030	0.9020	0.6800	**0.9030**
	PE	0.5210	0.7000	0.6905	0.7037	0.9030	0.6795	**0.9039**
	RE	0.5000	0.7001	0.6906	0.7039	0.9033	0.6796	**0.9043**
	ARI	-0.0004	0.1576	0.1451	0.1640	0.6461	0.1287	**0.6493**
	NMI	0.0004	0.1191	0.1075	0.1237	0.5458	0.0952	**0.5486**

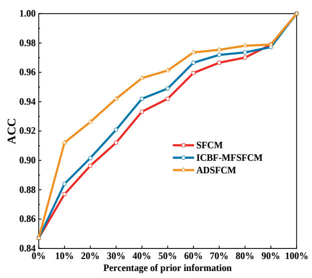
Fig. 6 and Table 5 show that ADSFCM outperforms other algorithms on ring- and chain-shaped datasets by effectively capturing non-spherical data structures. This innovation advances non-spherical clustering, benefiting pattern recognition, data mining, and related fields with strong engineering and industrial potential.

ADSFCM consistently outperforms other algorithms on uniform, multimodal, and non-spherical datasets by efficiently using prior information through an asymmetric deviation in its semi-supervised term. This allows accurate data structure capture, unlike other methods that lose guidance near prior membership values, resulting in weaker clustering. These results demonstrate ADSFCM's superior performance on numerical data.

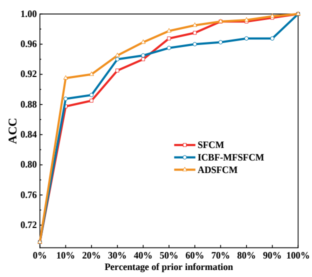
(3) Interpretability analysis

To evaluate the benefits and interpretability of the asymmetric deviation constraint $(u_{ij}(1+b_i)-3b_i\overline{f}_{ij})^2$ versus the symmetric term $(u_{ij}-b_i\overline{f}_{ij})^2$ in SFCM and the membership fusion term $b_i((u_{ij}-\overline{f}_{ij})^2+u_{ij}^2)$ in ICBF-MFSFCM, experiments used three datasets—uniform Wdbc, multimodal Lsun, and non-spherical Zelnik6—to compare clustering performance.

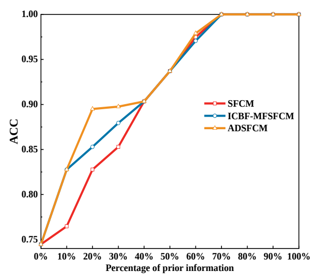
Fixing the prior membership degree, the proportion of prior information is adjusted: Based on standard class labels, the prior membership degree $\overline{f}_{ij}$ for a labeled data point x_i belonging to the j-th cluster is set to 1, and to 0 for all other clusters. Labeled samples increase evenly across classes to evaluate algorithm efficiency. Fig. 7 presents ACC curves for three algorithms on Wdbc, Lsun, and Zelnik6 datasets with varying labeled data.



(a)



(b)

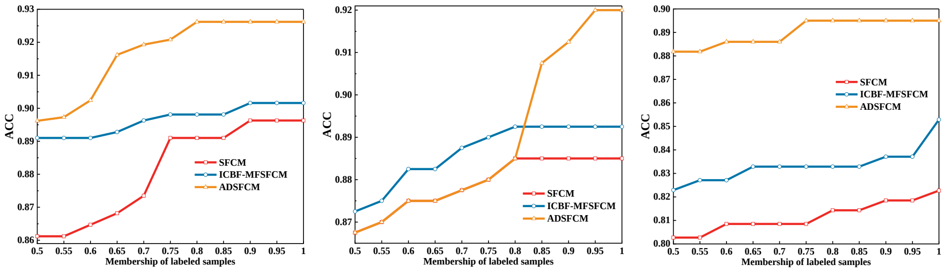


(c)

Fig. 7. ACC curves of three algorithms at varying prior information levels. (a) Wdbc; (b) Lsun; (c) Zelnik6.

Fig. 7 shows that without prior information, SFCM, ADSFCM, and ICBF-MFSFCM perform poorly, similar to standard fuzzy C-means. Increasing labeled samples improves all methods, with ADSFCM showing more stable gains at low prior levels due to its asymmetric deviation constraint. While all stabilize at high prior levels, ADSFCM's superior performance with limited labeled data is more practical given labeling costs.

With the prior information percentage fixed, the prior membership degree is adjusted: With 20% of samples labeled and evenly distributed, the membership degree $\overline{f}_{ij}$ of a labeled sample x_i in the j-th cluster rises from 0.5 to 1.0 in 0.05 steps, while others remain zero. **Fig. 8** shows ACC performance curves of three algorithms on Wdbc, Lsun, and Zelnik6 datasets as prior membership degrees change.



(a) (b) (c)

Fig. 8. Accuracy curves of three algorithms at varying prior membership degrees. (a) Wdbc; (b) Lsun; (c) Zelnik6.

Fig. 8 shows that in the uniformly distributed Wdbc dataset, some boundary data points have uncertain memberships, resulting in low membership values to their clustering centers. Setting the prior membership degree of labeled points to 0.65 significantly improves ADSFCM's clustering performance, indicating that the asymmetric deviation constraint effectively guides optimization even when actual memberships only approximate the prior. In the multimodally distributed Lsun dataset, most points lie in well-separated classes with high membership degrees. Setting the prior membership to 0.85 again enhances ADSFCM's performance, confirming the constraint's effectiveness when memberships approach prior values. In the non-spherically distributed Zelnik6 dataset, membership degrees are uneven due to the complex structure. The asymmetric deviation constraint leverages prior information to uncover structural features, enabling ADSFCM to outperform SFCM and ICBF-MFSFCM, demonstrating superior structural adaptability.

In summary, the asymmetric deviation constraint prevents prior information failure when $u_{ij} \approx \overline{f}_{ij}$, adapting to various data distributions and ensuring stable, robust clustering even with limited prior information.

(4) Natural image segmentation

Natural images, with complex textures and lighting variations, challenge segmentation methods. We test seven algorithms on four representative images, shown in **Fig. 9** alongside ground truth and partial supervision. Table 6 summarizes their segmentation performance.

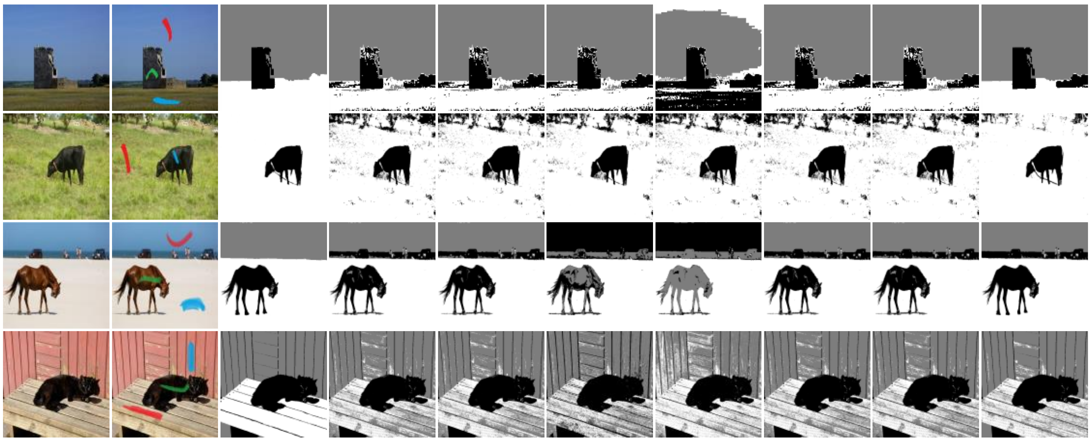


Fig. 9. Segmentation results of algorithms in natural images. (a) Original image; (b) Image with partial supervised information; (c) Ground truth; (d) SFCM; (e) SSFCM; (f) ESFCM; (g) SMUC; (h) SFCM-EP; (i) ICBF-MFSFCM; (j) ADSFCM. From top to bottom: #222, #54053, #83308, and #46316.

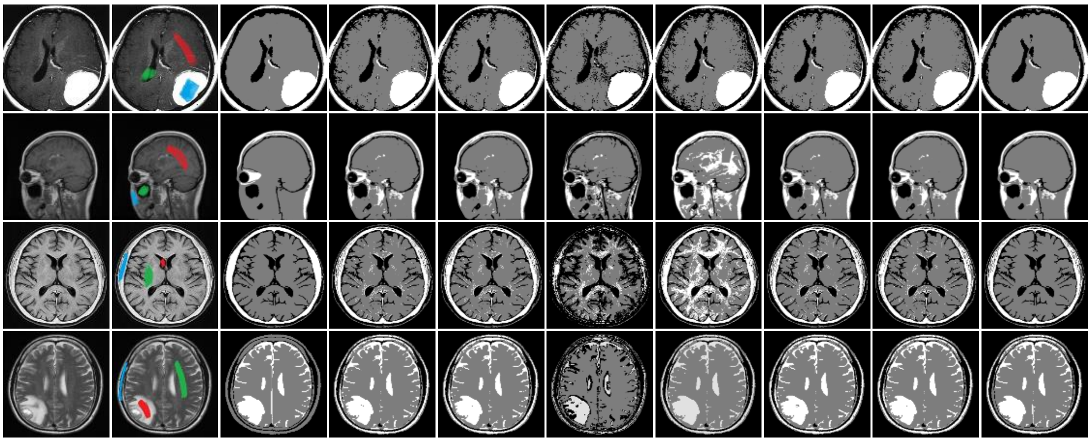
Table 6 Evaluation of seven algorithms applied to four natural images

Image	Indicator	SFCM	SSFCM	ESFCM	SMUC	SFCM-EP	ICBF-MFSFCM	ADSFCM
#222	ACC	0.9552	0.9552	0.9505	0.8007	0.9548	0.9564	**0.9811**
	Jaccard	0.8740	0.8740	0.8617	0.5765	0.8730	0.8773	**0.9450**
	mIoU	29.1305	29.1305	28.7234	21.6315	29.0996	29.2432	**31.4859**
	Time(unit:s)	0.374	0.349	2.919	0.510	3.22	3.914	3.17
	Iterations	13	13	25	12	28	22	27
#54053	ACC	0.9214	0.9215	0.9523	0.9191	0.9550	0.9216	**0.9729**
	Jaccard	0.8541	0.8544	0.8664	0.8503	0.8736	0.8545	**0.9472**
	mIoU	42.6631	42.6754	28.8453	42.4639	29.1036	**42.6808**	47.2724
	Time(unit:s)	0.342	0.433	28.387	1.356	12.937	7.399	6.932
	Iterations	14	14	49	13	24	28	26
#83308	ACC	0.9550	0.9553	0.7525	0.7599	0.9550	0.9553	**0.9573**
	Jaccard	0.8736	0.8744	0.4584	0.4704	0.8734	0.8744	**0.8795**
	mIoU	29.1019	29.1298	15.2739	15.6793	29.0980	29.1302	**29.2797**
	Time(unit:s)	0.464	0.778	0.751	1.254	12.744	13.64	12.982
	Iterations	12	12	32	8	24	25	24
#46316	ACC	0.8989	0.8989	0.8663	0.8829	0.8989	0.8989	**0.9209**
	Jaccard	0.7360	0.7361	0.6652	0.7006	0.7361	0.7361	**0.7873**
	mIoU	24.4383	24.4406	22.0905	23.2639	24.4389	24.4403	**26.1491**
	Time(unit:s)	0.638	0.717	1.824	2.502	22.515	23.025	21.412
	Iterations	13	13	65	14	33	33	32

Fig. 9 shows that noise from lighting and the environment affects natural image segmentation, causing bias. ADSFCM captures details and textures well, closely matching ground truths, while other algorithms lose detail on complex structures. Table 6 confirms ADSFCM's superior accuracy (ACC > 0.9, Jaccard > 0.8, mIoU > 0.29) but at the cost of longer processing times and more iterations, highlighting a trade-off between accuracy and efficiency.

(5) Medical image segmentation

Medical image segmentation aims to accurately outline tissues, organs, or lesions, aiding clinicians. Challenges include complex anatomy, noise, and lesion variability. We evaluate seven algorithms on four representative images, shown in Fig. 10 with original, partially supervised, ground truth, and segmentation results. Table 7 summarizes their segmentation performance.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j)

Fig. 10. Segmentation results of algorithms in medical images. (a) Original image; (b) Image with partial supervised information; (c) Ground truth; (d) SFCM; (e) SSFCM; (f) ESFCM; (g) SMUC; (h) SFCM-EP; (i) ICBF-MFSFCM; (j) ADSFCM. From top to bottom: Tr-me_0177, G_318, ct30, and Y49.

Table 7 Evaluation of seven algorithms in four medical images

Image	Indicator	SFCM	SSFCM	ESFCM	SMUC	SFCM-EP	ICBF-MFSFCM	ADSFCM
Tr-me_0177	ACC	0.9719	0.9719	0.9396	0.9507	0.9718	0.9734	**0.9880**
	Jaccard	0.9189	0.9189	0.8338	0.8620	0.9188	0.9231	**0.9645**
	mIoU	30.5508	30.5508	27.7242	28.6621	30.5499	30.6908	**31.9749**
	Time(unit:s)	0.178	0.456	0.536	0.592	2.766	2.831	2.832
	Iterations	10	10	16	10	20	20	20
G_318	ACC	0.9742	0.9734	0.9277	0.9147	0.9741	0.9740	**0.9774**
	Jaccard	0.9253	0.9232	0.8043	0.7730	0.9251	0.9249	**0.9344**
	mIoU	30.7958	30.7255	26.7707	25.7313	30.7892	30.7830	**31.1110**
	Time(unit:s)	0.654	0.703	0.793	3.770	17.734	17.933	17.117
	Iterations	16	16	20	27	32	32	31
ct30	ACC	0.9726	0.9730	0.8450	0.9076	0.9727	0.9743	**0.9856**
	Jaccard	0.9209	0.9220	0.6223	0.7562	0.9213	0.9258	**0.9576**
	mIoU	30.6114	30.6460	20.6952	25.1422	30.6224	30.7724	**31.8473**
	Time(unit:s)	0.190	0.350	0.273	0.569	2.596	2.675	2.556
	Iterations	12	12	31	21	33	33	33
Y49	ACC	0.9616	0.9603	0.8701	0.9408	0.9600	0.9524	**0.9715**
	Jaccard	0.8910	0.8875	0.6618	0.8242	0.8867	0.8668	**0.9181**
	mIoU	29.6789	29.5639	20.5441	23.8179	29.5347	28.8733	**30.5874**
	Time(unit:s)	0.211	0.384	0.391	0.954	4.939	5.169	4.777
	Iterations	15	15	29	11	29	30	28

Fig. 10 shows that ADSFCM outperforms other algorithms in segmenting complex organs and pathological regions, achieving over 97% accuracy, a Jaccard index above 0.91, and mIoU over 0.30 (Table 7). Despite its precision, ADSFCM's high computational cost requires optimization to enhance efficiency for clinical use.

(6) Remote sensing image segmentation

Remote sensing images capture diverse features but face challenges like complex terrain and noise. Clustering algorithms segment these images based on pixel spectra, with varying success. We evaluate seven algorithms on four representative images, shown in Fig. 11, with results summarized in Table 8.

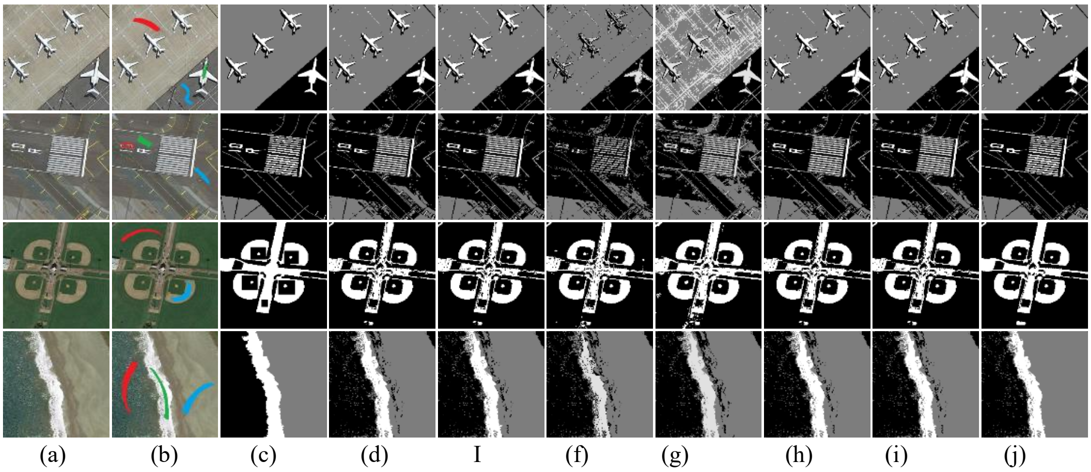


Fig. 11. Segmentation results of algorithms in remote sensing images. (a) Original image; (b) Image with partial supervised information; (c) Ground truth; (d) SFCM; I SSFCM; (f) ESFCM; (g) SMUC; (h)SFCM-EP; (i) ICBF-MFSFCM; (j) ADSFCM. From top to bottom: airplane_225, runway_401, baseball_diamond_057, and beach_072.

Table 8 Evaluation of seven algorithms in four remote sensing images

Image	Indicator	SFCM	SSFCM	ESFCM	SMUC	SFCM-EP	ICBF-MFSFCM	ADSFCM
airplane_225	ACC	0.9852	0.9856	0.9572	0.9018	0.9852	0.9853	**0.9888**
	Jaccard	0.9566	0.9576	0.8770	0.7093	0.9567	0.9570	**0.9670**
	mIoU	31.8727	31.9058	28.0878	18.6961	31.8756	31.8854	**32.2269**
	Time(unit:s)	0.262	0.445	0.302	0.895	2.647	2.914	2.528
	Iterations	22	12	22	24	20	21	19
runway_401	ACC	0.9312	0.9312	0.8888	0.8167	0.9327	0.9334	**0.9816**
	Jaccard	0.8129	0.8129	0.7088	0.5556	0.8135	0.81	**0.9463**
	mIoU	27.0872	27.0872	22.7955	17.4621	27.0855	27.0876	**31.5271**
	Time(unit:s)	0.155	0.429	0.316	0.848	5.882	6.704	5.73
	Iterations	11	22	19	21	44	43	43
baseball_diamond_057	ACC	0.9601	0.9612	0.9589	0.9451	0.9608	0.9603	**0.9737**
	Jaccard	0.9232	0.9241	0.9211	0.8959	0.9228	0.9239	**0.9487**
	mIoU	46.1178	46.1183	46.0108	44.7545	46.1164	46.1158	**47.4064**
	Time(unit:s)	0.157	0.287	0.248	0.610	0.902	0.896	0.890
	Iterations	6	6	7	21	14	14	14
beach_072	ACC	0.9477	0.9490	0.9072	0.9068	0.9478	0.9483	**0.9695**
	Jaccard	0.8545	0.8578	0.7455	0.7423	0.8548	0.8559	**0.9125**
	mIoU	28.4657	28.5778	22.9285	22.4907	28.4779	28.5129	**30.3761**
	Time(unit:s)	0.181	0.354	0.419	0.520	2.722	2.952	2.718
	Iterations	10	11	11	9	20	21	20

Fig. 11 shows that the ADSFCM algorithm excels in remote sensing segmentation, accurately identifying objects like aircraft and runways while preserving details. **Table 8** highlights its low complexity, over 96% accuracy, Jaccard index above 0.91, and mIoU over 30. These results confirm ADSFCM's precision and efficiency, making it suitable for large-scale geographic and environmental applications.

Overall, ADSFCM consistently outperforms other methods in segmenting natural, medical, and remote sensing images. Its asymmetric deviation constraint leverages prior annotations to enhance detail extraction, boosting segmentation accuracy and completeness.

4.3. Statistical test analysis

To evaluate whether performance metrics differences between the ADSFCM algorithm and others are statistically significant, the Friedman test (Demšar, 2006) is applied to nine datasets and twelve images; results appear in **Table 9**.

For clarity, the variables in the statistical test are defined as follows: k is the total number of algorithms compared; N is the total number of datasets evaluated; r_i^j is the rank of the j-th algorithm on the i-th dataset, with 1 as best and k as worst; R_j is the average rank of the j-th algorithm across all datasets; $df_1 = k-1$ and $df_2 = (k-1)(N-1)$ are the degrees of freedom; χ_F^2 and F_F are the test statistics based on the chi-square distribution and the corrected F-distribution, respectively, calculated as follows:

$$\chi_F^2 = \frac{12N}{k(k+1)}\left(\sum_{i=1}^{k} R_j^2 - \frac{k(k+1)^2}{4}\right) \tag{44}$$

$$F_F = \frac{(N-1)\chi_F^2}{N(k-1)-\chi_F^2} \tag{45}$$

The P-value measures the result randomness and tests sample differences. With a significance level $a = 0.05$. $P \le a$ leads to rejecting the null hypothesis of no algorithm performance difference. It is calculated by **Eq. (46)**.

$$P = F_a(df_1, df_2) \tag{46}$$

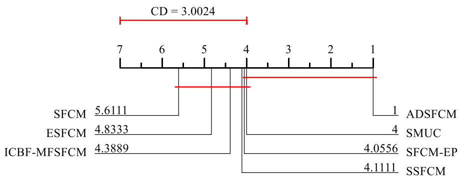
Table 9 F_F and P value for Friedman's test

	ACC1	PE	RE	ARI	NMI	ACC2	Jaccard	mIoU
F_F	6.6002	5.8802	6.6057	6.1972	6.6505	42.9514	40.1664	60.1587
P value	0.000039	0.000117	0.000039	0.000072	0.000036	6.37×10^{-18}	2.11×10^{-17}	1.07×10^{-20}

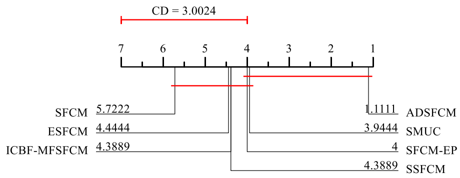
*ACC1 and ACC2 denote clustering and segmentation accuracy, respectively.

Table 9 shows all P-values meet $P \le 0.05$, rejecting the hypothesis and allowing post-hoc tests to identify significant algorithm differences.

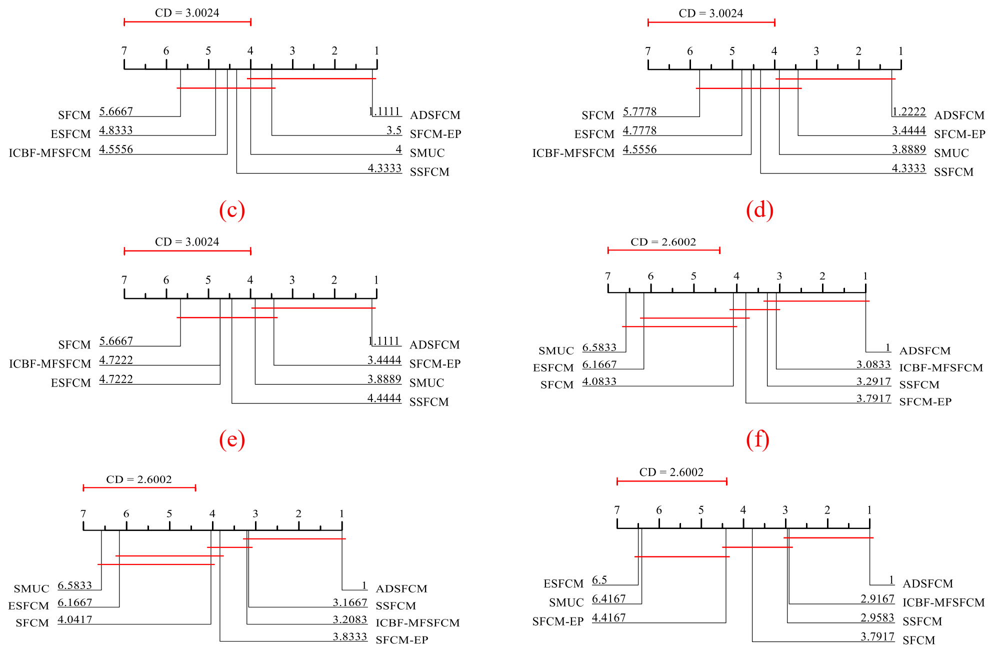
If the difference between two algorithms' average ranks is within the critical difference (CD), their performance difference is not statistically significant. If the difference equals or exceeds the CD, the performance difference is significant (Arbelaez et al., 2010). The CD is defined as $CD = q_a\sqrt{\frac{k(k+1)}{6N}}$, where q_a depends on the significance level a. Based on this, CD diagrams for various metrics of seven algorithms across nine numerical datasets and twelve images are shown in **Fig. 12**. In these diagrams, algorithms ranked farther to the right generally indicate better clustering or segmentation performance.



(a)



(b)



(c) (d)

(e) (f)

(g) (h)

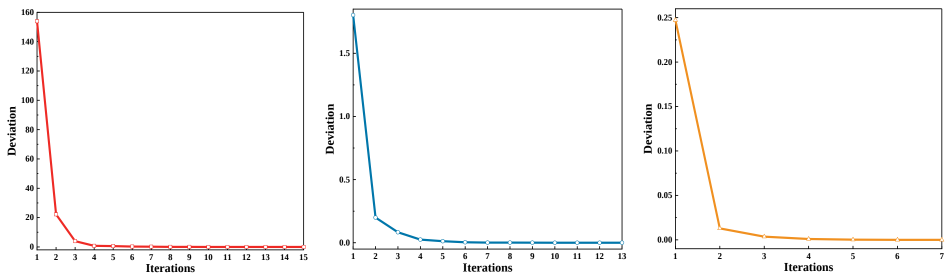
Fig. 12. CD diagrams of seven algorithms on nine numerical and twelve image datasets. (a) ACC1; (b) PE; (c) RE; (d) ARI; (e) NMI; (f) ACC2; (g) Jaccard; (h) mIoU.

Metrics ACC1, PE, and RE assess label consistency, while ARI and NMI evaluate clustering quality. **Figs. 12(a)–(c)** show ADSFCM ranks highest across metrics, and **Figs. 12(d)–(e)** confirm its superior clustering quality. ACC2 qualitatively evaluates segmentation results, and the Jaccard index and mIoU quantify quality. **Figs. 12(f)–(h)** demonstrate ADSFCM's superior performance.

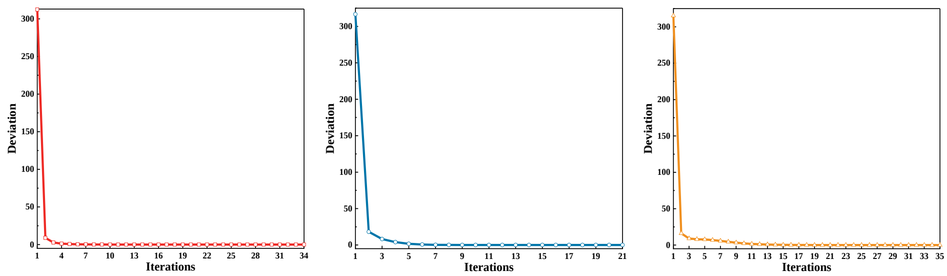
Results indicate ADSFCM's average rank differences mostly exceed the critical difference, confirming its statistically significant superior performance.

4.4. Convergence test

This study verifies the ADSFCM algorithm convergence using three datasets and three images. **Figs. 13** and **14** show clustering center deviation and objective function changes over iterations. The algorithm stops when the deviation falls below a 10^{-6} threshold.



(a) (b) (c)



(d) (e) (f)

Fig. 13. Deviation of clustering centers between iterations during clustering and segmentation. (a) Wdbc; (b) Lsun; (c) Zelnik6; (d) #46316; (e) Tr-me_0177; (f) runway_401.

Fig. 13 shows that ADSFCM's clustering center deviations decay exponentially, converging within finite iterations. Image segmentation takes longer than clustering, reflecting varied convergence rates and underscoring the need for acceleration strategies to boost efficiency.

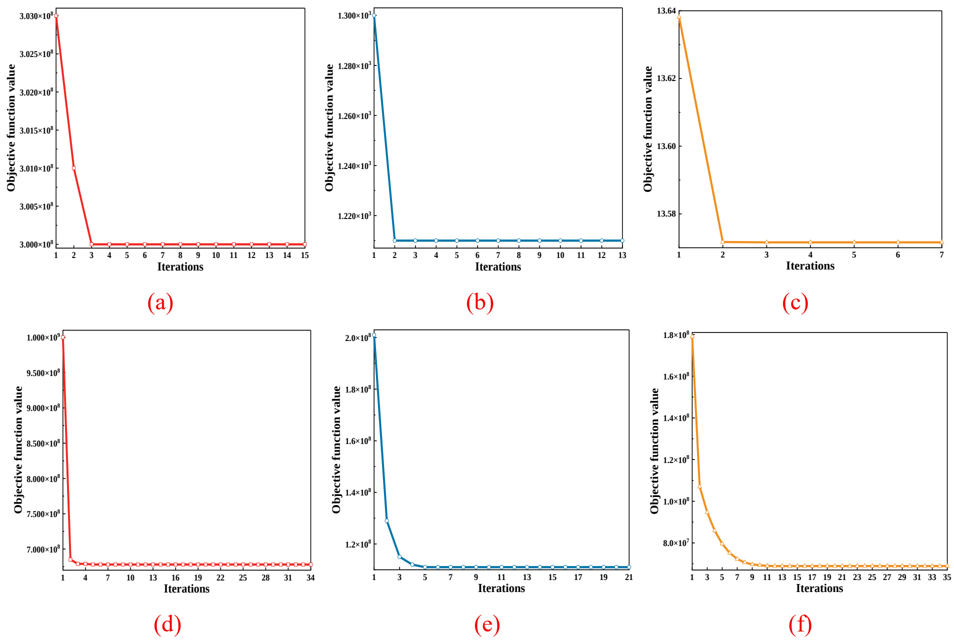


Fig. 14. Variation in the ADSFCM objective function value across iterations during clustering and segmentation. (a) Wdbc; (b) Lsun; (c) Zelnik6; (d) #46316; (e) Tr-me_0177; (f) runway_401.

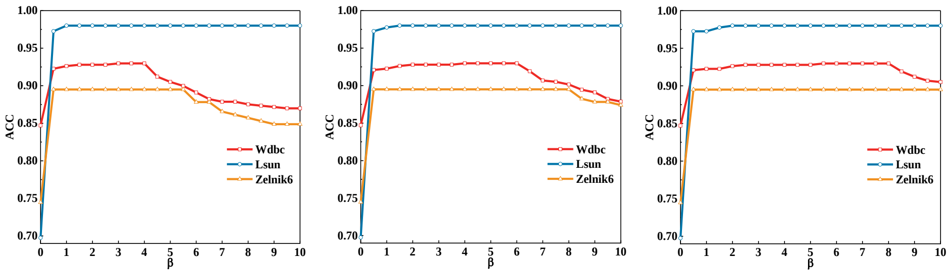
Fig. 14 shows that ADSFCM's objective function decreases steadily during clustering and more gradually in image segmentation, confirming its monotonic convergence and supporting algorithm improvements.

4.5. Parameter sensitivity analysis

The weighted adjustment factor β and regularization factor α balance unsupervised and semi-supervised information, enhancing ADSFCM clustering. Using three datasets—Wdbc, Lsun, and Zelnik6—and varying one parameter at a time, clustering accuracy (ACC) is evaluated to guide practical parameter selection.

(1) Sensitivity analysis for parameter β

Fig. 15 shows ADSFCM's clustering accuracy on three datasets as parameter β varies from 0 to 10, with α fixed, illustrating the algorithm's sensitivity and aiding parameter selection.



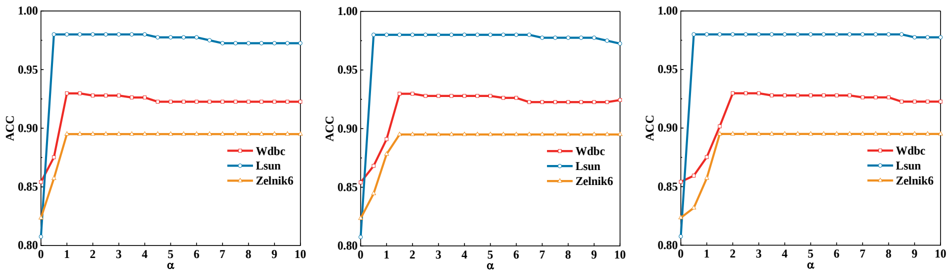
(a) (b) (c)

Fig. 15. ACC trend curves of the ADSFCM algorithm with varying β on numerical data. (a) $\alpha=2$; (b) $\alpha=3$; (c) $\alpha=4$.

Fig. 15 shows that with α at 2 or 3 and $\beta\in[0,4]$, ADSFCM consistently performs optimally across datasets. When $\beta\in[4,10]$, accuracy drops on Wdbc and Lsun, while $\alpha=4$, performance remains strong for $\beta\in[0,8]$ except for slight decline for $\beta\in[8,10]$ on Wdbc. Parameter β is more sensitive to data distribution when α is low. Increasing α reduces feature weight imbalance from high β, enhancing robustness.

(2) Sensitivity analysis for parameter α

Fig. 16 shows ADSFCM's clustering accuracy on three datasets as parameter α varies from 0 to 10, with β fixed, illustrating the algorithm's sensitivity and stability for parameter selection.



(a) (b) (c)

Fig. 16. ACC trend curves of the ADSFCM algorithm with varying α on numerical data. (a) $\beta=2$; (b) $\beta=3$; (c) $\beta=4$.

Fig. 16 shows the experimental system evaluating the effect of parameter α on clustering performance across various numerical datasets. With β fixed at 2, 3, or 4, the ADSFCM algorithm's accuracy rapidly increases and saturates when $\alpha\in[0,2]$, remaining stable for $\alpha\geq 2$. However, the Zelnik6 dataset behaves differently, with ACC consistently at 0.89, showing little sensitivity to changes in α. These results indicate that, when β is fixed, clustering performance is directly influenced by tuning α, though sensitivity varies by dataset.

Sensitivity analysis suggests setting β to 2–4 and α to 2–10. Stability improves when $\beta=2\alpha$; thus, $\beta=4$ and $\alpha=2$ are recommended. Increasing α stabilizes clustering when adjusting β, while fixing β at 2 enhances robustness when tuning α.

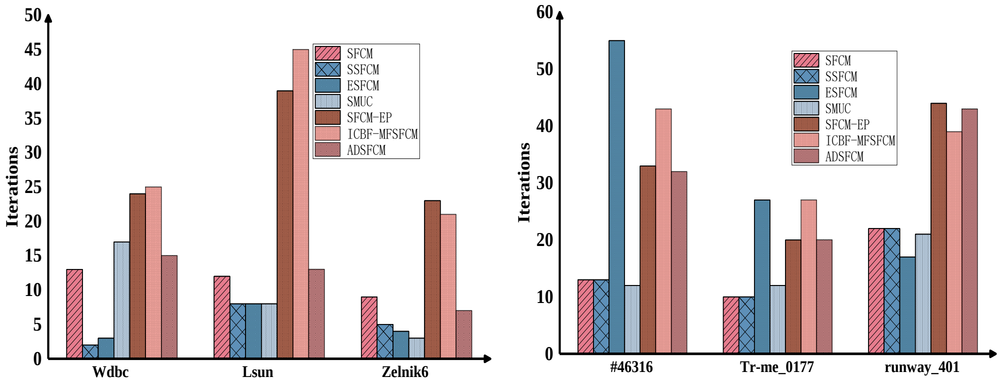
4.6. Time complexity analysis

Time complexity measures clustering algorithm efficiency using parameters n (samples), c (clusters), T (iterations), and d (dimensionality). **Table 10** compares their complexities.

The experiments use three numerical and three image datasets (#46316, Tr-me_0177, and runway_401) to compare ADSFCM's time complexity with other algorithms, establishing a basis for assessing performance on large-scale data.

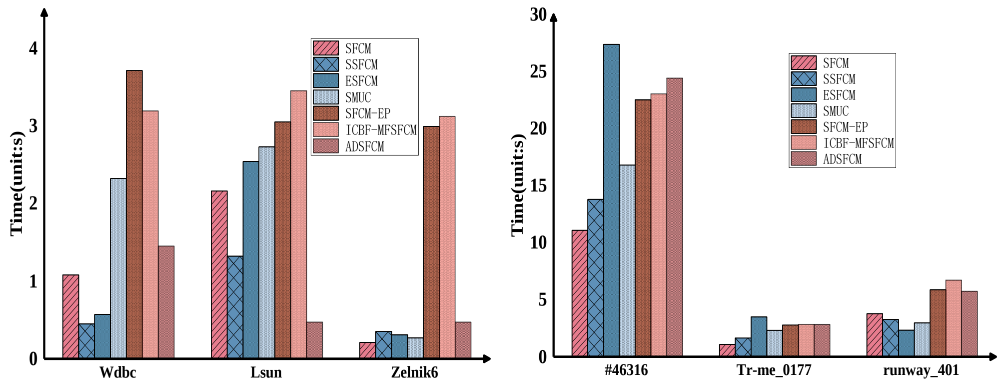
Table 10 Algorithm computational complexity

Algorithms	Time complexity
SFCM	$O(n \cdot c \cdot T \cdot d)$
SSFCM	$O(n \cdot c \cdot T \cdot d)$
ESFCM	$O(n \cdot c \cdot T \cdot d)$
SMUC	$O(n \cdot c \cdot T \cdot d)$
SFCM-EP	$O(n \cdot c \cdot T \cdot d)$
ICBF-MFSFCM	$O(n \cdot c \cdot T \cdot d)$
ADSFCM	$O(n \cdot c \cdot T \cdot d)$



(a) (b)

Fig. 17. Iteration histogram by algorithm. (a) Numerical datasets; (b) Image datasets.



(a) (b)

Fig. 18. Algorithm running time histogram. (a) Numerical datasets; (b) Image datasets.

Figs. 17 and **18** compare the time complexity of various algorithms on numerical and image datasets. The ADSFCM algorithm incurs high computational costs for numerical data due to its complex optimization. In medical image segmentation, its time cost is similar to six other algorithms but increases significantly for natural and remote sensing images. These findings reveal ADSFCM's time complexity limitations: despite good performance in some tasks, its efficiency is suboptimal, likely due to its complex design for improved segmentation accuracy. Future research should focus on reducing computational costs via algorithm optimization, parallel computing, or hardware acceleration while maintaining accuracy to improve practicality and scalability.

4.7. Testing of fast semi-supervised algorithms

(1) Numerical datasets

The proposed accelerated semi-supervised fuzzy C-means algorithm is tested on six datasets (Wdbc, Blobs, Ds3c3s6, Longsquare, Spiral, and Circle) to confirm faster convergence and high clustering performance.

The proposed accelerated semi-supervised fuzzy C-means clustering based on affinity filtering and membership scaling has been theoretically validated, showing strong feasibility and broad applicability to various semi-supervised clustering frameworks. Since SFCM serves as the foundational framework for most subsequent enhancements, these derived methods essentially preserve its core structure. Accordingly, this paper employs both SFCM and ADSFCM as the principal algorithmic carriers for implementing the proposed acceleration strategy. The accelerated version based on SFCM is denoted as AMSFCM, while the accelerated variant of ADSFCM is referred to as AM-ADSFCM. **Table 11** compares their performance across datasets.

Table 11 Test results of two semi-supervised methods and their fast algorithms on numerical data

Data Set	Indicator	SFCM	ADSFCM	AMSFCM	AM-ADSFCM
Wdbc (n=569, d=3, c=2)	ACC	0.8963	0.9262	0.9578	**0.9596**
	PE	0.9267	0.9359	0.9670	**0.9683**
	RE	0.8618	0.9153	0.9444	**0.9467**
	ARI	0.6227	0.7245	0.8370	**0.8435**
	NMI	0.5697	0.6066	0.7661	0.7730
	Iterations	30	15	9	9
	time(unit:s)	1.08	1.45	0.50	0.51
Blobs (n=300, d=2, c=3)	ACC	0.9667	**0.9833**	0.9800	**0.9833**
	PE	0.9667	0.9833	0.9801	**0.9834**
	RE	0.9667	**0.9833**	0.9800	**0.9833**
	ARI	0.9028	0.9505	0.9409	**0.9506**
	NMI	0.8600	0.9193	0.9082	**0.9235**
	Iterations	29	12	8	**6**
	time(unit:s)	2.98	0.49	0.50	0.52
Ds3c3s6 (n=905, d=2, c=6)	ACC	0.8243	**0.8409**	0.8276	0.8210
	PE	0.8238	**0.8436**	0.8258	0.8178
	RE	0.6479	**0.6603**	0.6492	0.6444
	ARI	0.5781	**0.6070**	0.6036	0.5962
	NMI	0.6993	**0.7361**	0.7393	0.7248
	Iterations	143	116	11	11
	time(unit:s)	1.94	0.98	0.59	0.58
Longsquare (n=900, d=2, c=6)	ACC	0.8567	**0.9522**	0.9011	0.9144
	PE	0.8567	**0.9519**	0.9011	0.9147
	RE	0.8567	**0.9522**	0.9011	0.9144
	ARI	0.7962	**0.8964**	0.8261	0.8414
	NMI	0.8632	**0.9056**	0.8751	0.8830
	Iterations	61	32	12	10
	time(unit:s)	2.75	0.64	0.64	0.59

Spiral (*n*=312, *d*=2, *c*=3)	ACC	0.3397	**0.4912**	0.4872	0.4872
	PE	0.3395	**0.4896**	0.4875	0.4875
	RE	0.3333	**0.4888**	0.4875	0.4875
	ARI	-0.0062	0.0503	0.0513	**0.0536**
	NMI	0.0002	0.0538	0.0507	0.0535
	Iterations	158	28	11	13
	time(unit:s)	3.09	0.51	0.50	0.32
Cirle (*n*=1000, *d*=2, *c*=2)	ACC	0.5210	**0.9030**	0.8970	0.9020
	PE	0.5210	**0.9039**	0.8985	0.9030
	RE	0.5000	**0.9043**	0.8985	0.9033
	ARI	-0.0004	**0.6493**	0.6301	0.6461
	NMI	0.0004	**0.5486**	0.5319	0.5458
	Iterations	57	48	17	16
	time(unit:s)	5.43	0.51	0.31	0.29

Table 11 shows that the AM-ADSFCM algorithm greatly improves convergence efficiency through innovative affinity filtering and membership scaling. It reduces iterations by up to 82% for uniformly distributed datasets and 90% for multimodally distributed datasets like Ds3c3s6 and Longsquare, while preserving clustering quality. For non-spherically distributed datasets, it cuts iterations by 95% and enhances clustering performance. For instance, on the Wdbc dataset, it lowers computational costs and boosts accuracy, especially with complex data. This acceleration strategy provides a practical solution for large-scale clustering with significant theoretical and practical benefits.

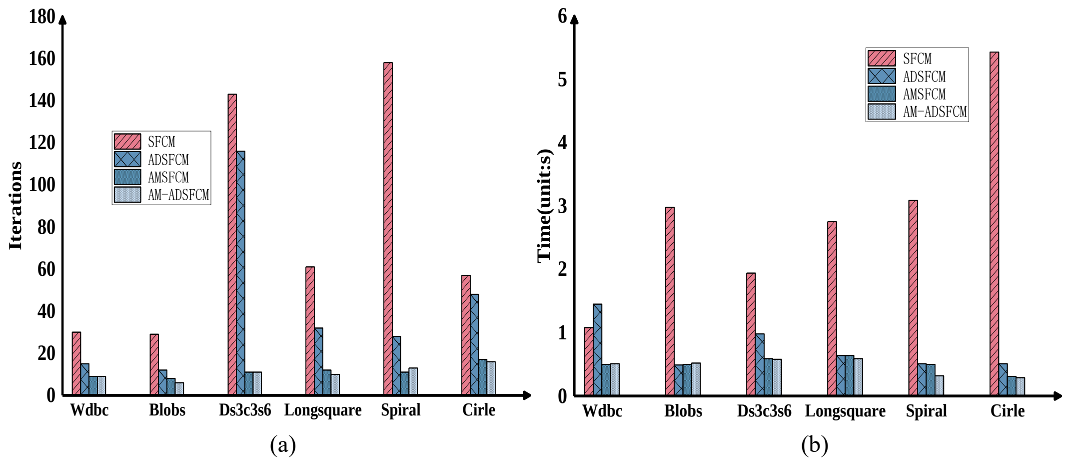


Fig. 19. Histogram of iterations and runtime for each algorithm and its accelerated version. (a) Number of iterations; (b) Running time.

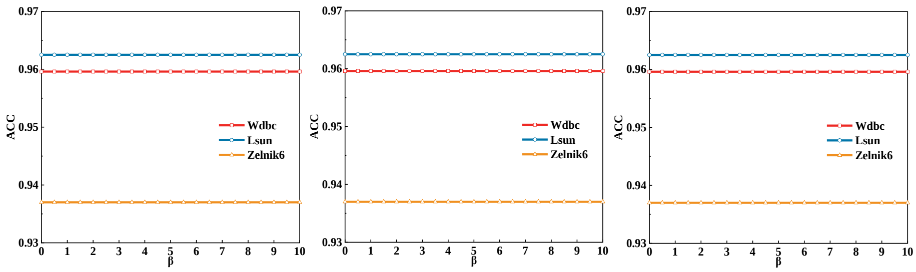
Fig. 19 shows that the AM-ADSFCM algorithm's accelerated versions greatly reduce iterations and runtime for large datasets by using affinity filtering and membership scaling, enhancing clustering efficiency.

The proposed accelerated semi-supervised fuzzy C-means algorithm, featuring affinity filtering and membership scaling, improves clustering accuracy and convergence speed across various datasets. It addresses slow convergence in traditional methods and offers practical solutions for large-scale data clustering, advancing semi-supervised clustering in pattern recognition and data mining.

(2) Impact of parameters on clustering efficiency and performance

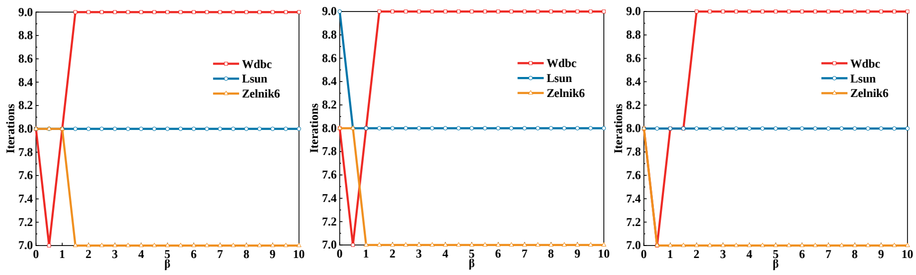
The AM-ADSFCM algorithm's efficiency and robustness are tested on three datasets (Wdbc, Lsun, Zelnik6) by evaluating clustering accuracy (ACC) and iterations across varying β and α parameters.

Figs. 20 and **21** show the AM-ADSFCM algorithm's clustering accuracy and iterations on three datasets, with α fixed and β varying from 0 to 10. This highlights the algorithm's sensitivity and stability, aiding parameter selection.



(a) (b) (c)

Fig. 20. Impact of parameter β on the ACC value of AM-ADSFCM. (a) $\alpha = 2$; (b) $\alpha = 3$; (c) $\alpha = 4$.

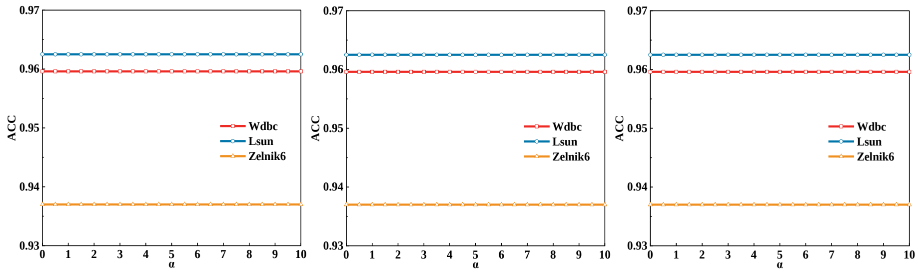


(a) (b) (c)

Fig. 21. Impact of parameter β on iterations of AM-ADSFCM. (a) $\alpha = 2$; (b) $\alpha = 3$; (c) $\alpha = 4$.

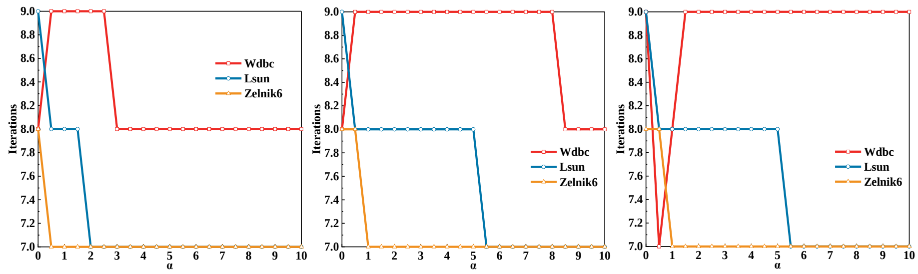
The experiments reveal how the weighted parameter β affects the performance and efficiency of the AM-ADSFCM algorithm. As shown in **Fig. 20**, with α fixed at 2, 3, and 4, the algorithm maintains stable clustering performance across three numerical datasets within the range $\beta \in [0,10]$. **Fig. 21** shows that parameter β significantly impacts convergence efficiency. When $\beta \in [0,2]$, the iteration counts on the uniformly distributed Wdbc and non-spherically distributed Zelnik6 datasets are low but fluctuate. For $\beta \in [2,10]$, iterations on Wdbc increase and stabilize, while on Zelnik6, they drop sharply before stabilizing. For the multimodally distributed dataset, iterations remain stable for $\beta \in [0,10]$, demonstrating robustness to parameter variation in complex distributions.

Figs. 22 and **23** show the AM-ADSFCM algorithm's clustering accuracy and iterations on three datasets, with β fixed and α varying from 0 to 10 in 0.5 increments. This highlights the algorithm's sensitivity and stability, aiding parameter selection.



(a) (b) (c)

Fig. 22. Impact of α on ACC values of AM-ADSFCM. (a) $\beta=2$; (b) $\beta=3$; (c) $\beta=4$.



(a) (b) (c)

Fig. 23. Impact of α on iterations of AM-ADSFCM. (a) $\beta=2$; (b) $\beta=3$; (c) $\beta=4$.

The experiment shows that the AM-ADSFCM algorithm's weighted parameter α regulates performance and efficiency. As shown in **Fig. 22**, with β fixed at 2, 3, and 4, respectively, clustering remains stable across datasets within the range $\alpha \in [0,10]$. **Fig. 23** demonstrates that Parameter α impacts convergence: for $\alpha \in [0,2]$, iterations rise on Wdbc, fall on Zelnik6, and stay stable on Lsun. When $\alpha \in [2,10]$, iterations on Wdbc stabilize and decline, while Zelnik6 and Lsun remain stable or decrease slightly. Thus, parameter α balances clustering quality and convergence across data types.

Sensitivity analysis indicates that changes in β and α minimally affect AM-ADSFCM's clustering performance and convergence, with iteration variation within ± 2. For stability, set $\beta=4$ and $\alpha=2$.

4.8. Comparison with the latest algorithms

This paper confirms ADSFCM's superiority over three recent semi-supervised fuzzy clustering algorithms—LIF-SFC (Zhu et al., 2025), WSPFCM (Samadi et al., 2025), and CapS3FCM (Gan et al., 2025)— in numerical data clustering and image segmentation.

Table 12. Evaluation of four algorithms on numerical data

Data set	Indicator	LIF_SFC	WSPFCM	CapS3FCM	ADSFCM
	ACC	0.8541	0.8735	0.9148	**0.9262**
Wdbc	PE	0.9026	0.9107	0.9538	**0.9359**
(n=569,p=3	RE	0.8052	0.8321	0.8714	**0.9153**
c=2)	ARI	0.4914	0.5502	0.5762	**0.7245**
	NMI	0.4648	0.5018	0.5255	**0.6066**
Blobs	ACC	**0.9833**	0.9821	0.9804	**0.9833**
(n=300,p=2,	PE	**0.9833**	0.9820	0.9803	**0.9833**
c=3)	RE	**0.9833**	0.9821	0.9804	**0.9833**

	ARI	**0.9505**	0.9364	0.9347	**0.9505**
	NMI	**0.9193**	0.8988	0.8973	**0.9193**
Iris (n=150,p=4, c=3)	ACC	0.9733	0.9644	0.9723	**0.9800**
	PE	0.9753	0.9673	0.9753	**0.9811**
	RE	0.9733	0.9644	0.9723	**0.9800**
	ARI	0.9222	0.8784	0.8856	**0.9410**
	NMI	0.9144	0.8672	0.8743	**0.9306**
Lsun (n=400,p=4, c=3)	ACC	0.9550	0.9613	0.9691	**0.9800**
	PE	0.9725	0.9713	0.9792	**0.9820**
	RE	0.9400	0.9522	0.9599	**0.9783**
	ARI	0.8660	0.9051	0.9124	**0.9371**
	NMI	0.8662	0.8950	0.9022	**0.9128**
Longsquare (n=900,p=2, c=6)	ACC	0.9045	0.9177	0.9129	**0.9522**
	PE	0.9044	0.0176	0.9128	**0.9519**
	RE	0.9045	0.9177	0.9129	**0.9522**
	ARI	0.8366	0.8488	0.8383	**0.8964**
	NMI	0.8673	0.8700	0.8754	**0.9056**
Ds3c3s6 (n=905,p=2, c=6)	ACC	0.7912	0.8342	0.8400	**0.8409**
	PE	0.7937	0.8369	0.8427	**0.8436**
	RE	0.6213	0.6550	0.6596	**0.6603**
	ARI	0.5711	0.6022	0.6064	**0.6070**
	NMI	0.6926	0.7302	0.7353	**0.7361**
Zelnik6 (n=238,p=2, C=3)	ACC	0.8319	0.8235	0.8933	**0.8950**
	PE	0.8666	0.8273	0.9110	**0.9333**
	RE	0.7638	0.7575	0.8687	**0.8512**
	ARI	0.6502	0.6355	0.7995	**0.7541**
	NMI	0.6367	0.6137	0.8104	**0.7922**
Spiral (n=312,p=2, c=3)	ACC	0.3429	0.3526	0.4856	**0.4912**
	PE	0.3430	0.3526	0.4871	**0.4896**
	RE	0.3376	0.5214	0.4860	**0.4888**
	ARI	-0.0062	-0.0034	0.2572	**0.0503**
	NMI	0.0003	0.0007	0.2342	**0.0538**
Cirle (n=1000,p=2 c=2)	ACC	**0.9030**	0.7130	0.8843	**0.9030**
	PE	0.9036	0.7172	0.8849	**0.9039**
	RE	0.9041	0.7093	0.8854	**0.9043**
	ARI	**0.6493**	0.1806	0.6359	**0.6493**
	NMI	0.5468	0.1375	0.5355	**0.5486**

Table 12 shows that ADSFCM significantly outperforms the latest methods. On the Wdbc dataset, its accuracy (ACC) reaches 0.9262, 1.24% higher than CapS3FCM, with Adjusted Rand Index (ARI) and Normalized Mutual Information (NMI) improving by 2.57% and 1.54%, respectively, demonstrating its strength in handling complex real-world data. For datasets with complex geometries like Longsquare and Zelnik6, ADSFCM achieves ACCs of 0.9522 and 0.8950 and ARIs of 0.8964 and 0.7541, substantially surpassing other methods, highlighting its adaptability to non-convex data. Notably, on the challenging Spiral dataset, ADSFCM's ACC and ARI improved by 1.15% and 95.53% over the latest algorithms, confirming the effectiveness of the asymmetric deviation constraint in capturing complex structures. Although its performance on Lsun and Circle datasets is slightly lower than some competitors, ADSFCM still outperforms them overall. These results show that ADSFCM effectively balances prior knowledge and data adaptability by integrating asymmetric deviation constraints in membership fusion, offering a robust solution for clustering diverse numerical data.

Table 13. Evaluation of four algorithms in sixteen images

Image	**Indicator**	**LIF_SFC**	**WSPFCM**	**CapS3FCM**	**ADSFCM**
#222	ACC	0.9548	0.9564	0.9678	**0.9811**
	Jaccard	0.8730	0.8773	0.9322	**0.9450**

	mIoU	29.0996	29.2432	31.0591	**31.4859**
#54053	ACC	0.9550	0.9216	0.9347	**0.9729**
	PE	0.8736	0.8545	0.9100	**0.9472**
	RE	29.1036	**42.6808**	45.4163	47.2724
#83308	ACC	0.9550	0.9553	0.9528	**0.9573**
	PE	0.8734	0.8744	0.8754	**0.8795**
	RE	29.0980	29.1302	29.1421	**29.2797**
#46316	ACC	0.9059	0.9174	0.8989	**0.9209**
	PE	0.7745	0.7843	0.7361	**0.7873**
	RE	25.7232	26.0497	24.4389	**26.1491**
#Tr-me _0177	ACC	0.9718	0.9729	0.9736	**0.9880**
	PE	0.9188	0.9198	0.9205	**0.9645**
	RE	30.5499	30.5845	30.6065	**31.9749**
#G_318	ACC	0.9735	0.9726	0.9699	**0.9774**
	PE	0.9246	0.9238	0.9212	**0.9344**
	RE	30.7748	30.7463	30.6610	**31.1110**
#Ct30	ACC	0.9730	0.9739	0.9728	**0.9856**
	PE	0.9220	0.9229	0.9218	**0.9576**
	RE	30.6476	30.6759	30.6413	**31.8473**
#Y49	ACC	0.9617	0.9601	0.9611	**0.9715**
	PE	0.8912	0.8897	0.8906	**0.9181**
	RE	29.6861	29.6367	29.6676	**30.5874**
#airplane_225	ACC	0.9852	0.9826	0.9844	**0.9888**
	PE	0.9566	0.9541	0.9558	**0.9670**
	RE	31.8727	31.7886	31.8468	**32.2269**
#runway_401	ACC	0.9312	0.9354	0.9369	**0.9816**
	PE	0.8128	0.8165	0.8178	**0.9463**
	RE	27.0855	27.2077	27.2513	**31.5271**
#baseball_diamond_057	ACC	0.9601	0.9593	0.9581	**0.9737**
	PE	0.9232	0.9224	0.9213	**0.9487**
	RE	46.1178	46.0794	46.0217	**47.4064**
#beach_072	ACC	0.9476	0.9488	0.9500	**0.9695**
	PE	0.8543	0.8554	0.8565	**0.9125**
	RE	28.4596	28.4956	28.5317	**30.3761**

Table 13 demonstrates that ADSFCM improves image segmentation, with ACC gains up to 4.77% and Jaccard increases up to 15.75% over CapS3FCM. It also boosts RE by 4.05% in low-contrast scenes and raises mIoU by 1.20% for small targets. Overall, ADSFCM delivers average ACC and Jaccard gains of 2.15% and 3.62%, offering a robust method for complex visual scene analysis.

5. Conclusions and outlook

To enhance the utilization of prior information in semi-supervised clustering when membership degrees approach prior values, we redesign the SFCM regularization term by replacing the symmetric deviation constraint with an asymmetric one. Using non-affinity center identification and membership scaling, we introduce a novel acceleration framework for diverse semi-supervised clustering models. Experiments show ADSFCM's asymmetric deviation constraint enhances prior information use and adaptability, while its acceleration mechanism speeds convergence and stabilizes clustering, benefiting large or complex datasets.

ADSFCM and its acceleration strategy have limitations, including reliance on accurate prior information, difficulty with complex high-dimensional data, and potential negative effects from imprecise prior membership on optimization.

The proposed semi-supervised fuzzy clustering methods will expand to collaborative and federated frameworks to boost efficiency and robustness (Salehi et al., 2021; Pedrycz, 2022) to further enhance their performance and computational efficiency. Existing adaptive techniques (Chen et al., 2024) will be employed to further improve the robustness and adaptability of the proposed methods. Future work includes developing privacy-preserving distributed clustering systems combining federated learning and multi-view fuzzy clustering, alongside integrating multi-view contrastive learning and kernel methods to better represent complex data (Hu et al., 2023; Liu et al., 2023).

CRediT authorship contribution statement

Chengmao Wu: Conceptualization, Methodology, Writing-Editing, Supervision

Jun Hou: Methodology, Software, Application to real data, Writing - reviewing and editing

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This study was supported by the Shaanxi provincial department of education special fund of China (24JK0531), the Shann provincial natural science foundation of China (2025JC-YBMS-343).

Reference

Antoine, V., Guerrero, J. A., & Romero, G. (2022). Possibilistic fuzzy C-means with partial supervision. Fuzzy Sets and Systems, 449, 162-186. https://doi.org/10.1016/j.fss.2022.08.003

Antoine, V., Guerrero, J. A., & Xie, J. (2021). Fast semi-supervised evidential clustering. International Journal of Approximate Reasoning, 133, 116-132. https://doi.org/10.1016/j.ijar.2021.03.008

Arbelaez, P., Maire, M., Fowlkes, C., & Malik, J. (2010). Contour detection and hierarchical image segmentation. IEEE transactions on pattern analysis and machine intelligence, 33(5), 898-916. https://doi.org/10.1109/TPAMI.2010.161

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy C-means clustering algorithm. Computers & geosciences, 10(2-3), 191-203. https://doi.org/10.1016/0098-3004(84)90020-7

Bezdek, J. C. (1993). A physical interpretation of fuzzy ISODATA. In Readings in Fuzzy Sets for Intelligent Systems (pp. 615-616). Morgan Kaufmann. https://doi.org/10.1016/B978-1-4832-1450-4.50065-1

Cai, J., Hao, J., Yang, H., Zhao, X., & Yang, Y. (2023). A review on semi-supervised clustering. Information Sciences, 632, 164-200. https://doi.org/10.1016/j.ins.2023.02.088

Cai, J. X., Yang, F., & Feng, G. C. (2009). Fast semi-supervised fuzzy clustering: Approach and application. In 2009 Chinese Conference on Pattern Recognition (pp. 1-5). IEEE. https://doi.org/10.1109/CCPR.2009.5344131

Chen, H. R., Wang, X. P., Wu, J. X., & Wang, H. Z. (2024). Adaptive semi-supervised fuzzy C-means method with local spatial information and pre-clustering for image segmentation. IEEE Access, 12 (2024), 196328-196346. https://doi.org/10.1109/ACCESS.2024.3521595

Daneshfar, F., Soleymanbaigi, S., Yamini, P., & Amini, M. S. (2024). A survey on semi-supervised graph clustering. Engineering Applications of Artificial Intelligence, 133, 108215. https://doi.org/10.1016/j.engappai.2024.108215

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. Journal of Machine learning research, 7, 1-30. https://dl.acm.org/doi/10.5555/1248547.1248548

Despotovic, I., Vansteenkiste, E., & Philips, W. (2013). Spatially coherent fuzzy clustering for accurate and noise-robust image segmentation. IEEE Signal Processing Letters, 20(4), 295-298. https://doi.org/10.1109/LSP.2013.2244080

Ding, Y., Zhao, Y., Shen, X., Musuvathi, M., & Mytkowicz, T. (2015). Yinyang k-means: A drop-in replacement of the classic k-means with consistent speedup. In International conference on machine learning (pp. 579-587). PMLR. https://dl.acm.org/doi/abs/10.5555/3045118.3045181

Fantoukh, N. I., Ismail, M. M. B., & Bchir, O. (2020). Automatic determination of the number of clusters for semi-supervised relational fuzzy clustering. International Journal of Fuzzy Logic and Intelligent Systems, 20(2), 156-167. https://doi.org/10.5391/IJFIS.2020.20.2.156

Gan, H., Fan, Y., Luo, Z., & Zhang, Q. (2018). Local homogeneous consistent safe semi-supervised clustering. Expert Systems with Applications, 97, 384-393. https://doi.org/10.1016/j.eswa.2017.12.046

Gan, H. (2019). Safe semi-supervised fuzzy C-means clustering. IEEE Access, 7, 95659-95664. https://doi.org/10.1109/ACCESS.2019.2929307

Gan, H., Fan, Y., Luo, Z., Huang, R., & Yang, Z. (2019). Confidence-weighted safe semi-supervised clustering. Engineering Applications of Artificial Intelligence, 81, 107-116. https://doi.org/10.1016/j.engappai.2019.02.007

Gan, H., Yang, Z., Shi, M., Ye, Z., & Zhou, R. (2025). Improved safe semi-supervised clustering based on capped $\ell_{2,1}$ -norm. Fuzzy Sets and Systems, 505, 109276. https://doi.org/10.1016/j.fss.2025.109276

Gan, H., Gan, W., Yang, Z., & Zhou, R. (2024). Discrimination-aware safe semi-supervised clustering. Information Sciences, 676, 120798. https://doi.org/10.1016/j.ins.2024.120798

Hu, X., Qin, J., Shen, Y., Pedrycz, W., Liu, X., & Liu, J. (2023). An efficient federated multiview fuzzy C-means clustering method. IEEE Transactions on Fuzzy Systems, 32(4), 1886-1899. https://doi.org/10.1109/TFUZZ.2023.3335361

Huan, P. T., Thong, P. H., Tuan, T. M., Hop, D. T., Thai, V. D., Minh, N. H., Giang, N. L., Son, L. H. (2022). TS3FCM: trusted safe semi-supervised fuzzy clustering method for data partition with high confidence. Multimedia Tools and Applications, 81(9), 12567-12598.

https://doi.org/10.1007/s11042-022-12133-6

Ichihashi, H., Miyagishi, K., & Honda, K. (2001). Fuzzy C-means clustering with regularization by KL information. In 10th IEEE International Conference on Fuzzy Systems. (pp. 924-927). IEEE. https://doi.org/10.1109/FUZZ.2001.1009107

Kanzawa, Y. (2017). Semi-supervised fuzzy C-means algorithms by revising dissimilarity/kernel matrices. In Fuzzy Sets, Rough Sets, Multisets and Clustering (pp. 45-61). Springer. https://doi.org/10.1007/978-3-319-47557-8_4

Kmita, K., Kaczmarek-Majer, K., & Hryniewicz, O. (2024). Explainable impact of partial supervision in semi-supervised fuzzy clustering. IEEE Transactions on Fuzzy Systems, 32(5), 3189-3198. https://doi.org/10.1109/TFUZZ.2024.3370768

Kolen, J. F., & Hutcheson, T. (2002). Reducing the time complexity of the fuzzy C-means algorithm IEEE Transactions on Fuzzy Systems, 10(2), 263-267. https://doi.org/10.1109/91.995126

Li, D., Zhou, S., & Pedrycz, W. (2023). Accelerated fuzzy C-means clustering based on new affinity filtering and membership scaling. IEEE Transactions on Knowledge and Data Engineering, 35(12), 12337-12349. https://doi.org/10.1109/TKDE.2023.3273274

Liu, J., Liu, X., Yang, Y., Liao, Q., & Xia, Y. (2023). Contrastive multi-view kernel learning. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(8), 9552-9566. https://doi.org/10.1109/TPAMI.2023.3253211

Mendonça, M. O., Netto, S. L., Diniz, P. S., & Theodoridis, S. (2024). Machine learning: Review and trends. Signal processing and machine learning theory, 869-959. https://doi.org/10.1016/B978-0-32-391772-8.00019-3

Mittal, H., Pandey, A. C., Saraswat, M., Kumar, S., Pal, R., & Modwel, G. (2022). A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets. Multimedia Tools and Applications, 81(24), 35001-35026. https://doi.org/10.1007/s11042-021-10594-9

Pedrycz, W. (1985). Algorithms of fuzzy clustering with partial supervision. Pattern recognition letters, 3(1), 13-20. https://doi.org/10.1016/0167-8655(85)90037-6

Pedrycz, W., & Waletzky, J. (1997). Fuzzy clustering with partial supervision. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 27(5), 787-795. https://doi.org/10.1109/3477.623232

Pedrycz, W. (2002). Collaborative fuzzy clustering. Pattern Recognition Letters, 23(14), 1675-1686. https://doi.org/10.1016/S0167-8655(02)00130-7

Pedrycz, W. (2021). Federated FCM: Clustering under privacy requirements. IEEE Transactions on Fuzzy Systems, 30(8), 3384-3388. https://doi.org/10.1109/TFUZZ.2021.3105193

Saha, A., & Das, S. (2018). Stronger convergence results for the center-based fuzzy clustering with convex divergence measure. IEEE Transactions on Cybernetics, 49(12), 4229-4242. https://doi.org/10.1109/TCYB.2018.2861211

Salehi, F., Keyvanpour, M. R., & Sharifi, A. (2021a). SMKFC-ER: Semi-supervised multiple kernel fuzzy clustering based on entropy and relative entropy. Information Sciences, 547, 667-688. https://doi.org/10.1016/j.ins.2020.08.094

Salehi, F., Keyvanpour, M. R., & Sharifi, A. (2021b). GT2-CFC: General type-2 collaborative fuzzy clustering method. Information Sciences, 578, 297-322. https://doi.org/10.1016/j.ins.2021.07.037

Samadi, N., Tanha, J., & Jalili, M. (2025). A weighted semi-supervised possibilistic fuzzy C-means algorithm for data stream classification and emerging class detection. Knowledge-Based Systems, 309, 112831. https://doi.org/10.1016/j.knosys.2024.112831

Shen, Y., & Pedrycz, W. (2017). Collaborative fuzzy clustering algorithm: Some refinements. International Journal of Approximate Reasoning, 86, 41-61. https://doi.org/10.1016/j.ijar.2017.04.004

Singh, J., & Singh, D. (2024). A comprehensive review of clustering techniques in artificial intelligence for knowledge discovery: Taxonomy, challenges, applications and future prospects. Advanced Engineering Informatics, 62, 102799. https://doi.org/10.1016/j.aei.2024.102799

Tran, D., & Wagner, M. (2000). Fuzzy entropy clustering. In Ninth IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2000) (pp. 152-157). IEEE. https://doi.org/10.1109/FUZZY.2000.838650

Tuan, T. M., Sinh, M. D., Khang, T. Đ., Huan, P. T., Ngan, T. T., Giang, N. L., & Thai, V. D. (2022). A new approach for semi-supervised fuzzy clustering with multiple fuzzifiers. International Journal of Fuzzy Systems, 24(8), 3688-3701. https://doi.org/10.1007/s40815-022-01363-3

Tuan, T. M., Thong, P. H., Ngan, T. T., & Son, L. H. (2022). An improvement of trusted safe semi-supervised fuzzy clustering method with multiple fuzzifiers. Journal of Computer Science and Cybernetics, 38(1), 47-61. https://doi.org/10.15625/1813-9663/38/1/16720

Van Engelen, J. E., & Hoos, H. H. (2020). A survey on semi-supervised learning. Machine learning, 109(2), 373-440. https://doi.org/10.1007/s10994-019-05855-6

Wang, Z., Wang, S. S., Bai, L., Wang, W. S., & Shao, Y. H. (2021). Semi-supervised fuzzy clustering with fuzzy pairwise constraints. IEEE Transactions on Fuzzy Systems, 30(9), 3797-3811. https://doi.org/10.1109/TFUZZ.2021.3129848

Xu, S., Hao, Z., Zhu, Y., Wang, Z., Xiao, Y., & Liu, B. (2024). Semi-supervised fuzzy clustering algorithm based on prior membership degree matrix with expert preference. Expert Systems with Applications, 238, 121812. https://doi.org/10.1016/j.eswa.2023.121812

Yang, M. S., & Tian, Y. C. (2015). Bias-correction fuzzy clustering algorithms. Information Sciences, 309, 138-162. https://doi.org/10.1016/j.ins.2015.03.006

Yasunori, E., Yukihiro, H., Makito, Y., & Sadaaki, M. (2009). On semi-supervised fuzzy c-means clustering. In 2009 IEEE International Conference on Fuzzy Systems (pp. 1119-1124). IEEE. https://doi.org/10.1109/FUZZY.2009.5277177

Yin, X., Shu, T., & Huang, Q. (2012). Semi-supervised fuzzy clustering with metric learning and entropy regularization. Knowledge-Based Systems, 35, 304-311.

https://doi.org/10.1016/j.knosys.2012.05.016

Yu, H., Xu, X., Li, H., Wu, Y., & Lei, B. (2024). Semi-supervised possibilistic C-means clustering algorithm based on feature weights for imbalanced data. Knowledge-Based Systems, 286, 111388. https://doi.org/10.1016/j.knosys.2024.111388

Zhang, B., Huang, L., Wang, J., Zhang, L., Wu, Y., Jiang, Y., & Xia, K. (2024a). Semi-supervised fuzzy C-means based on membership integration mechanism and its application in brain infarction lesion segmentation in DWI images. Journal of Intelligent & Fuzzy Systems, 46(1), 2713-2726. https://doi.org/10.3233/JIFS-234148

Zhang, B., Jiang, Y., & Xia, K. (2024b). Interclass balance factor-based membership fusion semi-supervised fuzzy clustering algorithm for lesion segmentation in cerebral infarction images, 12 (2024) 107077-107088. IEEE Access. https://doi.org/10.3233/JIFS-234148

Zhang, Z., Yu, X., Tao, R., Zhang, X., Li, H., Lu, J., & Zhou, J. (2023). Knowledge augmentation-based soft constraints for semi-supervised clustering. Applied Soft Computing, 144, 110484. https://doi.org/10.1016/j.asoc.2023.110484

Zhou, R., Liu, Q., Wang, J., Han, X., & Wang, L. (2021). Modified semi-supervised affinity propagation clustering with fuzzy density fruit fly optimization. Neural Computing and Applications, 33(10), 4695-4712. https://doi.org/10.1007/s00521-020-05431-3

Zhu, H., Xie, W., Mu, Y., Xu, J., Wang, F. L., Qu, Y., & Hao, T. (2024). A new semi-supervised fuzzy K-means clustering method with dynamic adjustment and label discrimination. Neural Computing and Applications, 36(9), 4709-4725. https://doi.org/10.1007/s00521-023-09115-6

Zhu, H., Kan, B., Li, Y., Yan, E., Weng, H., Wang, F. L., & Hao, T. (2025). A new semi-supervised fuzzy clustering method based on latent representation learning and information fusion. Applied Soft Computing, 170, 112717. https://doi.org/10.1016/j.asoc.2025.112717