# SAFE SEMI-SUPERVISED FUZZY CLUSTERING AND ITS APPLICATION IN AGRICULTURAL IMAGE SEGMENTATION

**Phung The Huan, Le Thu Trang***

*TNU - University of Information and Communication Technology*

**ARTICLE INFO**

**Received:** 03/03/2025

**Revised:** 28/03/2025

**Published:** 28/03/2025

**KEYWORDS**

Clustering

Fuzzy clustering

Picture fuzzy set

Agriculture

Safe

**ABSTRACT**

Image segmentation is an important technique in image analysis and processing in agriculture, aiding in the monitoring of crop health, yield estimation, crop mapping, and resource management. Traditional image segmentation methods face challenges due to the need for large labeled datasets and the complexity of the agricultural environment. To address this issue, the paper proposes combining the advantages of fuzzy clustering and semi-supervised learning, allowing the use of both labeled and unlabeled data to improve segmentation accuracy. The paper also introduces picture fuzzy sets, an extension of fuzzy sets, which provides a more detailed representation of uncertainty and effectively handles noisy regions in the data. When combined with safe semi-supervised learning techniques, the proposed method ensures high accuracy in agricultural image segmentation, particularly in noisy and uncertain environments. This approach offers significant advantages over existing methods, minimizing the negative impact of labeled data and optimizing the clustering process. The results show that the safe semi-supervised fuzzy clustering method with picture fuzzy set can be effectively applied to agricultural image segmentation with complex and noisy agricultural data.

# PHÂN CỤM BÁN GIÁM SÁT MỜ AN TOÀN VÀ ỨNG DỤNG TRONG PHÂN ĐOẠN ẢNH NÔNG NGHIỆP

**Phùng Thế Huân, Lê Thu Trang***

*Trường Đại học Công nghệ thông tin và Truyền thông – ĐH Thái Nguyên*

**THÔNG TIN BÀI BÁO**

**Ngày nhận bài:** 03/03/2025

**Ngày hoàn thiện:** 28/03/2025

**Ngày đăng:** 28/03/2025

**TỪ KHÓA**

Phân cụm

Phân cụm mờ

Tập mờ bức tranh

Nông nghiệp

An toàn

**TÓM TẮT**

Phân đoạn ảnh là một kỹ thuật quan trọng trong việc phân tích và xử lý hình ảnh trong nông nghiệp, giúp giám sát sức khỏe cây trồng, ước tính năng suất, lập bản đồ cây trồng và quản lý tài nguyên. Các phương pháp phân đoạn ảnh truyền thống gặp phải thách thức về yêu cầu dữ liệu gắn nhãn lớn và tính phức tạp của môi trường nông nghiệp. Để giải quyết vấn đề này, bài báo đề xuất kết hợp ưu điểm của phân cụm mờ và học bán giám sát, cho phép sử dụng cả dữ liệu có nhãn và không có nhãn để cải thiện độ chính xác phân đoạn. Bài báo cũng sử dụng tập mờ bức tranh, một tập mờ mở rộng từ các tập mờ gốc, giúp biểu diễn sự không chắc chắn một cách chi tiết hơn và xử lý hiệu quả các vùng nhiễu trong dữ liệu. Khi kết hợp tập mờ bức trannh với các kỹ thuật học bán giám sát an toàn, phương pháp đề xuất đảm bảo tính chính xác cao trong phân đoạn ảnh nông nghiệp, đặc biệt trong môi trường có nhiễu và không chắc chắn. Phương pháp này mang lại những ưu điểm vượt trội so với các phương pháp hiện có, giúp giảm thiểu tác động tiêu cực của dữ liệu có nhãn và tối ưu hóa quá trình phân cụm. Kết quả cho thấy phương pháp phân cụm mờ bán giám sát an toàn với tập mờ bức tranh có thể ứng dụng hiệu quả trong phân đoạn ảnh nông nghiệp với dữ liệu phức tạp và chứa nhiễu.

**DOI:** https://doi.org/10.34238/tnu-jst.12195

* Corresponding author. *Email: lttrang@ictu.edu.vn*

## 1. Introduction

Image segmentation is the process of splitting a digital image into distinct regions or objects based on characteristics such as color, brightness, texture, or boundaries of objects [1]. The primary goal of image segmentation is to simplify or alter the image's representation, enhancing its ease and accuracy for analysis and interpretation. In the field of agriculture, image segmentation plays a crucial role in various applications such as crop health monitoring, yield estimation, crop mapping, and resource management. Specifically, in crop health monitoring, image segmentation helps identify areas of disease or damage on plants, thereby aiding in diagnosis and timely intervention [2]. In yield estimation, segmentation can analyze crop images and environmental factors, providing more accurate predictions of crop yields [3]. In crop mapping, this method classifies and maps different plant species based on their image characteristics, supporting agricultural management and production planning. Finally, in resource management, image segmentation can analyze soil images to identify different zones, enabling the efficient allocation of resources such as water and fertilizers [4].

Conventional image segmentation techniques, particularly those using supervised learning, require extensive labeled data, making data collection and processing challenging in complex agricultural environments. Semi-supervised C-Means (SFCM) [5] addresses these issues by combining fuzzy clustering (Fuzzy C-Means) [6] with semi-supervised learning, leveraging both labeled and unlabeled data to enhance segmentation accuracy. SFCM reduces reliance on manual labeling, adapts to various data types, and effectively handles uncertainty. Its ability to incorporate unlabeled data is especially useful in large-scale agricultural monitoring, where acquiring labeled samples is costly. By integrating picture fuzzy sets (PFS) [7], the proposed approach improves data partitioning, enhances robustness against noise, and refines segmentation quality in challenging environments. This method is particularly effective in agricultural applications, where environmental variability and data ambiguity often hinder traditional segmentation techniques.

Data clustering has long been an important research topic, with numerous methods developed to address challenges across various domains [8]. Among these techniques, fuzzy clustering, initially proposed by Bezdek, has gained significant interest due to its capability to manage uncertainty. A well-known algorithm in this domain is Fuzzy C-Means (FCM) [6], which has been extensively applied in image segmentation [9], pattern recognition [10], and bioinformatics [11]. However, FCM is highly sensitive to noise and outliers, which can distort cluster centers and reduce clustering accuracy. To address this issue, several extended methods, such as Robust Fuzzy C-Means [12] and Kernel-based FCM [13], have been proposed to enhance noise robustness through alternative objective functions, distance measures, and spatial information [14]. In the field of agricultural imagery, data collection in natural environments is often subject to various types of noise, complicating the image processing and analysis process. Environmental factors such as variations in sunlight, low-light conditions, rain, fog, and wind-induced motion blur can introduce noise. Additionally, dust accumulation on the lens and plant growth variations further complicate segmentation tasks.

Although fuzzy clustering is effective in handling uncertainty, it often struggles with complex datasets, particularly those containing noise or overlapping clusters. This limitation has prompted the integration of semi-supervised learning (SSL) techniques, which utilize labeled data to assist the clustering process. SSL methods include constraint-based approaches [15], which optimize clusters by applying must-link or cannot-link constraints; distance-based learning [16], which improves similarity measures with labeled data; and graph-based methods [17], which use graphs to model relationships and propagate labels. Recent studies have focused on enhancing both global and local consistency constraints [18] and label propagation methods [19], strengthening the robustness of clustering in noisy environments.

PFS expands conventional fuzzy sets by adding positive, neutral, and negative membership levels, offering a more comprehensive approach to modeling uncertainty. It has been successfully applied in decision-making [20], image processing, and pattern recognition [21]. Recent improvements include optimized distance measures [22] and entropy weighting [23], which enhance clustering performance under uncertain conditions. Despite these advancements, improper integration of labeled data in semi-supervised learning can negatively impact clustering performance compared to unsupervised methods. To address this, safe semi-supervised fuzzy clustering techniques have been proposed to ensure that labeled data does not reduce clustering accuracy. These methods utilize strategies such as reliability weighting, constraint relaxation, and ensemble techniques. Additionally, multi-objective optimization [24] and local density information [25] have been employed to improve safety and stability during the clustering process.

To tackle these challenges, this paper introduces an innovative safe semi-supervised fuzzy clustering approach that leverages PFS to enhance data partitioning in noisy environments. By effectively modeling uncertainty, PFS provides a nuanced approach to handling ambiguous data. This method integrates safe semi-supervised learning techniques, leveraging expert knowledge while ensuring that labeled data does not degrade clustering performance. The proposed method offers several advantages: (1) handling noisy data through PFS's expressive power, (2) integrating expert knowledge to enhance clustering accuracy, and (3) ensuring the safe use of labeled data without compromising performance compared to unsupervised methods. This results in a robust and precise data partitioning approach, particularly in datasets containing noise and uncertainty.

## 2. Method

The NPFS3FCM method introduced combines picture fuzzy sets with safe semi-supervised fuzzy clustering, featuring a novel objective function composed of four essential components. The objective function is expressed as follows:

$$\begin{aligned} F &= \beta\sum_{k=1}^{N}\sum_{j=1}^{C}(\mu_{kj}(2-\xi_{kj}))^2\left\|X_k - V_j\right\|^2 - \delta\sum_{k=1}^{N}\sum_{j=1}^{C}\eta_{kj}(\ln(\eta_{kj})-\xi_{kj}) \\ &+\theta\sum_{k=1}^{L}\sum_{j=1}^{C}\frac{(\mu_{kj}(2-\xi_{kj})-f_{kj})^2}{1+(\overline{\mu}_{kj}-f_{kj})^2}\left\|X_k - V_j\right\|^2 + \theta\sum_{k=L+1}^{N}\sum_{j=1}^{C}(\mu_{kj}(2-\xi_{kj})-\overline{\mu}_{kj})^2\left\|X_k - V_j\right\|^2 \to Min \end{aligned} \quad (1)$$

With the constraint conditions:

$$\begin{cases} \mu_{kj},\xi_{kj},\eta_{kj}\in[0,1] \\ \mu_{kj}+\xi_{kj}+\eta_{kj}\le 1 \\ \sum_{j=1}^{C}\left(\mu_{kj}\right)=1 \\ \sum_{j=1}^{C}\left(\eta_{kj}+\dfrac{\xi_{kj}}{C}\right)=1 \\ (k=\overline{1,N};\ j=\overline{1,C}) \end{cases} \quad (2)$$

$$A_{kj}^{L}=\left(\beta\left(2-\xi_{kj}\right)^2 d_{kj}^{s}+\frac{\theta\left(2-\xi_{kj}\right)^2 d_{kj}^{s}}{1+\left(\overline{\mu}_{kj}-f_{kj}\right)^2}\right) \quad (3)$$

$$B_{kj}^{L}=\sum_{i=1}^{C}\frac{A_{kj}^{L}}{A_{ki}^{L}} \quad (4)$$

$$G_{kj}^{L}=\frac{\theta f_{kj}\left(2-\xi_{kj}\right)d_{kj}^{s}}{1+\left(\overline{\mu}_{kj}-f_{kj}\right)^2} \quad (5)$$

$$A_{kj}^{N}=\left(\beta\left(2-\xi_{kj}\right)^2 d_{kj}^{s}+\theta\left(2-\xi_{kj}\right)^2 d_{kj}^{s}\right) \quad (6)$$

$$G_{kj}^{N}=\theta\overline{\mu}_{kj}\left(2-\xi_{kj}\right)d_{kj}^{s} \quad (7)$$

$$B_{kj}^{N}=\sum_{i=1}^{C}\frac{A_{kj}^{N}}{A_{ki}^{N}} \quad (8)$$

$$d_{kj}^{S}=X_k - V_j^{\,2} \quad (9)$$

In which:

$X=\{X_1, X_2, ..., X_N\}$: Dataset X; $N$: Number of data elements; $L$: The number of data labeled in $X$; $C$: Number of cluster; $\mu_{kj}$, $\eta_{kj}$, $\xi_{kj}$: positive, neutral and refusal degrees of $X_k$.

The optimal solutions are obtained using the Lagrangian method and are presented in equations (10-14).

$$V_j = \frac{\beta\sum_{k=1}^{N}(\mu_{kj}(2-\xi_{kj}))^2 X_k + \theta\sum_{k=1}^{L}\frac{(\mu_{kj}(2-\xi_{kj})-f_{kj})^2}{1+(\overline{\mu}_{kj}-f_{kj})^2}X_k + \theta\sum_{k=L+1}^{N}(\mu_{kj}(2-\xi_{kj})-\overline{\mu}_{kj})^2 X_k}{\beta\sum_{k=1}^{N}(\mu_{kj}(2-\xi_{kj}))^2 X_k + \theta\sum_{k=1}^{L}\frac{(\mu_{kj}(2-\xi_{kj})-f_{kj})^2}{1+(\overline{\mu}_{kj}-f_{kj})^2} + \theta\sum_{k=L+1}^{N}(\mu_{kj}(2-\xi_{kj})-\overline{\mu}_{kj})^2} \tag{10}$$

For labeled data, the membership $\mu$ degree is calculated as:

$$\mu_{kj} = \frac{1}{B_{kj}^L} - \frac{\sum_{i=1}^{C}\frac{G_{ki}^L}{A_{ki}^L}}{B_{kj}^L} + \frac{G_{kj}^L}{A_{kj}^L} \tag{11}$$

For unlabeled data, the membership $\mu$ degree is calculated as:

$$\mu_{kj} = \frac{1}{B_{kj}^N} - \frac{\sum_{i=1}^{C}\frac{G_{ki}^N}{A_{ki}^N}}{B_{kj}^N} + \frac{G_{kj}^N}{A_{kj}^N} \tag{12}$$

The degree of neutral $\eta_{kj}$ and the degree of refusal $\xi_{kj}$ are calculated as follows:

$$\eta_{kj} = \frac{1-\frac{1}{C}\sum_{j=1}^{C}\xi_{kj}}{\sum_{i=1}^{C}\frac{e^{\xi_{ki}}}{e^{\xi_{kj}}}} \tag{13}$$

$$\xi_{kj} = \left(1-\left(\mu_{kj}+\eta_{kj}\right)^{\alpha}\right)^{\frac{1}{\alpha}} \tag{14}$$

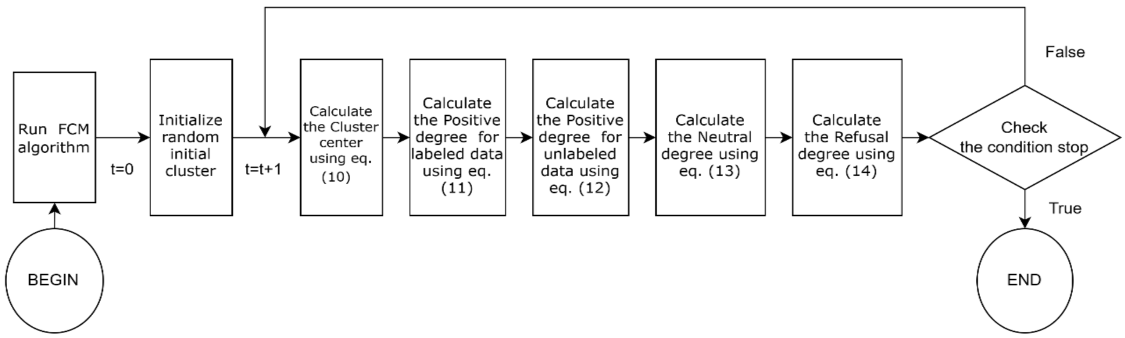



**Figure 1.** *The algorithm diagram of NPFS3FCM*

The algorithm diagram of NPFS3FCM is presented in Figure 1. The algorithm will terminate once the condition $\|F^t - F^{t-1}\| \le \varepsilon$ is met or the iteration count reaches $Maxsteps$.

## 3. Results and Discussion

### *3.1. Evaluation on labeled data*

The proposed approach is assessed by its classification accuracy on labeled data. In experiments with datasets, the auxiliary information for semi-supervised clustering methods consists of pre-labeled data points, accounting for 20% of the total data points. Classification accuracy for NPFS3FCM, FC-PFS, and CS3FCM is computed using labeled data from 15 datasets [26], [27], and the results are presented in Table 1.

**Table 1.** *Classification accuracy for labeled data without noise.*

| METHOD | NPFS3FCM | FC-PFS | CS3FCM |
|---|---|---|---|
| Australian | 0.88044 | 0.59368 | 0.73395 |
| Balance-scale | 0.87801 | 0.47638 | 0.53585 |
| Dermatology | 0.88440 | 0.44638 | 0.47918 |
| Heart | 0.88323 | 0.61693 | 0.74145 |
| Iris | 0.88822 | 0.77855 | 0.85016 |
| Spambase | 0.87500 | 0.69296 | 0.70896 |
| Tae | 0.88861 | 0.53238 | 0.55857 |
| Waweform | 0.87500 | 0.50336 | 0.52359 |
| Wdbc | 0.88031 | 0.65980 | 0.81087 |

Table 1 presents the accuracy of three clustering methods with labeled data in a noise-free environment. NPFS3FCM outperforms all other methods across all datasets, significantly outperforming FC-PFS and CS3FCM. The ability of NPFS3FCM to maintain high accuracy highlights the effectiveness of the safe semi-supervised learning method, ensuring that the integration of labeled data does not degrade clustering quality, but rather enhances clustering results. This result directly addresses the motivation of the study, where the integration of labeled data is identified as an important but challenging aspect of semi-supervised clustering.

Table 2 presents the accuracy for labeled datasets containing noise, where NPFS3FCM continues to lead in accuracy, particularly in complex datasets such as Vertebral and Wine. Unlike FC-PFS, which struggles with noisy labeled data, NPFS3FCM demonstrates its superiority by effectively handling challenges posed by noise without compromising clustering results. This reinforces the objective of the study, ensuring that labeled data can be safely and effectively utilized, even under less-than-ideal conditions that contain noise.

**Table 2.** *Classification accuracy for labeled data with noise*

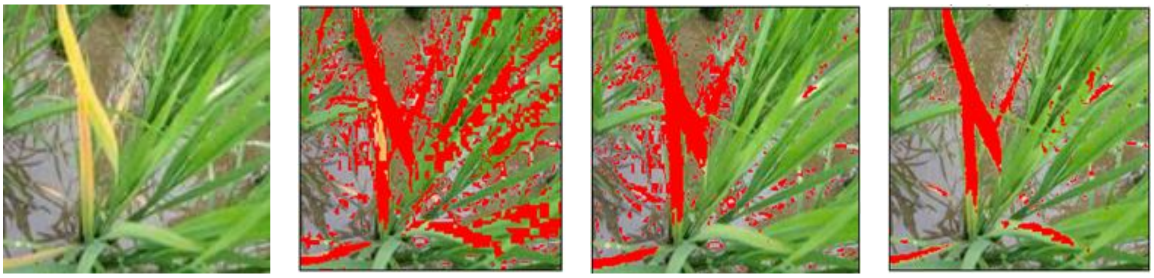
| METHOD | NPFS3FCM | FC-PFS | CS3FCM |
|---|---|---|---|
| Ecoli | 0.88318 | 0.55551 | 0.51386 |
| Glass | 0.62214 | 0.43231 | 0.44376 |
| Yeast | 0.46417 | 0.29044 | 0.35129 |
| Wine | 0.89762 | 0.80023 | 0.82272 |
| Vertebral | 0.76643 | 0.49862 | 0.58464 |
| Ionosphere | 0.81833 | 0.53435 | 0.55682 |

### *3.2. Evaluation on image from Kaggle dataset*

The image dataset is taken from Kaggle [28]. The images produced by applying the three algorithms—CS3FCM, FC-PFS, and NPFS3FCM—on three original images (Figure 2.a, 3.a, and 4.a) from a set of 20 experimental images are shown in Figures 2, 3, and 4. Healthy rice leaves are typically bright green and smooth, indicating good growth without pest infestations or environmental stress. However, when rice leaves dry out or become diseased, they turn yellow or brown, often developing dry streaks along the veins. These symptoms may indicate diseases like rice blast or nutrient and water deficiencies, which can significantly impact plant health. The presence of dry patches reduces photosynthesis efficiency, directly affecting crop yield. In the original images, Figure 2.a clearly differentiates between healthy and diseased leaf areas. However, Figures 3.a and 4.a exhibit similar color tones in these regions, making segmentation more challenging. Therefore, a robust segmentation method is essential for accurately identifying diseased areas and assessing their severity.

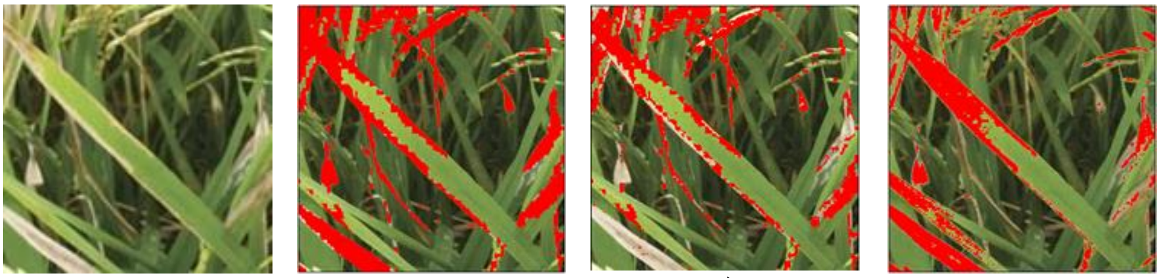
**a)** **b)** **c)** **d)**

**Figure 2.** *Clustering result of Image 2*
*a) Image before processing; b) CS3FCM; c) FC-PFS; d) NPFS3FCM*



**a)** **b)** **c)** **d)**

**Figure 3.** *Clustering result of Image 3*
*a) Image before processing; b) CS3FCM; c) FC-PFS; d) NPFS3FCM*



**a)** **b)** **c)** **d)**

**Figure 4.** *Clustering result of Image 4*
*a) Image before processing; b) CS3FCM; c) FC-PFS; d) NPFS3FCM*

When applying clustering methods such as CS3FCM, FC-PFS, and NPFS3FCM to Figure 2.a, the segmentation results obtained are relatively consistent (for example, the results in Figure 2.b, 2.c, 2.d). However, for Figure 3.a and Figure 4.a, significant differences only appear in the segmentation results, as these images have similar color tones between the diseased leaf area and the rice panicle area, causing noise and complicating the analysis.

In this situation, fuzzy clustering methods such as FC-PFS and NPFS3FCM demonstrate better effectiveness, with Figure 3.d and Figure 4.d serving as evidence of the superiority of the NPFS3FCM method in segmentation. This could be because FC-PFS and NPFS3FCM use a fuzzy set with four membership levels, significantly improving clustering performance. In particular, the NPFS3FCM method adds an objective function compared to FC-PFS, resulting in more optimal clustering performance.

Additionally, the NPFS3FCM method shows improved adaptability to complex images with high noise levels, allowing for a more accurate separation of different regions within the image.

This advantage is particularly crucial for images where color similarities make traditional clustering methods less effective. The ability of NPFS3FCM to handle uncertainty in segmentation ensures a more precise classification of image regions, leading to a clearer distinction between diseased and healthy areas. Furthermore, its robust performance in handling variations in texture and illumination makes it a reliable choice for agricultural image analysis, especially in challenging conditions.

## 4. Conclusion

This study introduces NPFS3FCM, a safe semi-supervised fuzzy clustering method built on the foundation of Fuzzy Image Sets (PFS), aimed at overcoming the limitations of traditional clustering methods when dealing with noisy data, ambiguity, and the problem of integrating labeled data. The uniqueness of NPFS3FCM lies in its ability to leverage the expressive power of PFS, allowing a data point to simultaneously belong to multiple clusters with varying degrees of membership, thus flexibly capturing the inherent uncertainty in real-world data. Furthermore, this method cleverly combines a safe semi-supervised learning technique, enabling the effective use of information from labeled data without disrupting the natural clustering process.

The effectiveness of NPFS3FCM has been validated through experiments on several datasets with varying levels of noise. The results demonstrate that NPFS3FCM achieves superior accuracy. In particular, the harmonious combination of labeled data integration and robust noise resistance has confirmed the high applicability of NPFS3FCM in practice.

In the future, the research will focus on extending NPFS3FCM to handle large-scale data and exploring its potential applications in fields such as image segmentation and bioinformatics. With its outstanding advantages, NPFS3FCM is expected to become a powerful and flexible solution, contributing to the advancement of data mining and machine learning.

## Acknowledgment

This research result is the product of a scientific research project with code T2024-07-04, funded by Thai Nguyen University of Information and Communication Technology (ICTU).

## REFERENCES

[1] S. Minaee, *et al.*, "Image segmentation using deep learning: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3523-3542, 2021.

[2] V. Singh, N. Sharma, and S. Singh, "A review of imaging techniques for plant disease detection," *Artificial Intelligence in Agriculture*, vol. 4, pp. 229-242, 2020.

[3] A. K. Aggarwal and P. Jaidka, "Segmentation of crop images for crop yield prediction," *International Journal of Biology and Biomedicine*, vol. 7, pp. 45-52, 2022.

[4] T. H. Pham, *et al.*, "A new picture fuzzy clustering method to segment the surface water from satellite images," (in Vietnamese), *TNU Journal of Science and Technology*, vol. 227, no. 16, pp. 28-36, 2022.

[5] Y. Endo, *et al.*, "On semi-supervised fuzzy c-means clustering," in *Proc. IEEE International Conference on Fuzzy Systems*, 2009, pp. 1234-1240.

[6] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2, pp. 191-203, 1984.

[7] C. C. Bui, "Picture fuzzy sets," *Journal of Computer Science and Cybernetics*, vol. 30, no. 4, pp. 409-416, 2014.

[8] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognition Letters*, vol. 31, pp. 651–666, 2010.

[9] H. Mittal, A. C. Pandey, M. Saraswat, S. Kumar, R. Pal, and G. Modwel, "A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets," *Multimedia Tools and Applications*, vol. 81, no. 24, pp. 1-26, 2022.

[10] S. Singh and A. H. Ganie, "Applications of picture fuzzy similarity measures in pattern recognition, clustering, and MADM," *Expert Systems with Applications*, vol. 168, 2021, Art. no. 114264.

[11] Z. Li, *et al.*, "Unified K-means coupled self-representation and neighborhood kernel learning for

clustering single-cell RNA-sequencing data," *Neurocomputing*, vol. 501, pp. 715-726, 2022.
[12] Y. Gao, Z. Wang, J. Xie, and J. Pan, "A new robust fuzzy c-means clustering method based on adaptive elastic distance," *Knowledge-Based Systems*, vol. 237, 2022, Art. no. 107769.
[13] J. Hu, M. Wu, L. Chen, and W. Pedrycz, "A novel modeling framework based on customized kernel-based fuzzy C-means clustering in iron ore sintering process," *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 2, pp. 950-961, 2021.
[14] Y. Long, *et al.*, "Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST," *Nature Communications*, vol. 14, no. 1, 2023, Art. no. 1155.
[15] X. Zhao, C. Fang, D. J. Fan, X. Lin, F. Gao, and G. Li, "Cross-level contrastive learning and consistency constraint for semi-supervised medical image segmentation," in *Proc. IEEE 19th Int. Symp. on Biomedical Imaging (ISBI)*, 2022, pp. 1-5.
[16] A. Młodak, "K-means, ward and probabilistic distance-based clustering methods with contiguity constraint," *Journal of Classification*, vol. 38, no. 2, pp. 313-352, 2021.
[17] B. Pourasghar, H. Izadkhah, A. Isazadeh, and S. Lotfi, "A graph-based clustering algorithm for software systems modularization," *Information and Software Technology*, vol. 133, 2021, Art. no. 106469.
[18] X. Yang, *et al.*, "Robust semi-supervised fuzzy clustering algorithm based on pairwise constraints," *Iranian Journal of Fuzzy Systems*, vol. 21, no. 3, pp. 155-175, 2024.
[19] K. Berahmand, S. Haghani, M. Rostami, and Y. Li, "A new attributed graph clustering by using label propagation in complex networks," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 5, pp. 1869-1883, 2022.
[20] T. Shahid, S. Ashraf, and D. S. Mashat, "Enhancing urban development with picture fuzzy sets: A strategic decision support framework," *J. Urban Dev. Manag.*, vol. 2, no. 4, pp. 172-180, 2023.
[21] S. Singh and A. H. Ganie, "Applications of picture fuzzy similarity measures in pattern recognition, clustering, and MADM," *Expert Systems with Applications*, vol. 168, 2021, Art. no. 114264.
[22] T. H. Pham and S. H. Le, "Picture fuzzy clustering: a new computational intelligence method," *Soft Computing*, vol. 20, no. 9, pp. 3549-3562, 2016.
[23] S. H. Le, "DPFCM: A novel distributed picture fuzzy clustering method on picture fuzzy sets," *Expert Systems with Applications*, vol. 42, pp. 51–66, 2015.
[24] T. H. Pham, *et al.*, "Picture-Neutrosophic Trusted Safe Semi-Supervised Fuzzy Clustering for Noisy Data," *Computer Systems Science & Engineering*, vol. 46, no. 2, pp. 1981-1997, 2023.
[25] M. Hu, *et al.*, "Fuzzy system based medical image processing for brain disease prediction," *Frontiers in Neuroscience*, vol. 15, 2021, Art. no. 714318.
[26] UCI Machine Learning Repository, "Data," 2024. [Online]. Available: https://archive.ics.uci.edu/ml/index.php. [Accessed Feb. 16, 2025].
[27] Outlier Detection DataSets (ODDS), "Anomaly detection datasets," 2024. [Online]. Available: http://odds.cs.stonybrook.edu. [Accessed Feb. 16, 2025].
[28] N. Sankalana, "Rice Leaf Disease Images," 2024. [Online]. Available: https://www.kaggle.com/datasets/nirmalsankalana/rice-leaf-disease-image. [Accessed Feb. 16, 2025].