REPPORT

Fuzzy C-Means Algorithm

AMJAD Hadia – Master 2 Computer Science - VMI

16 November 2025

GitHub: Fuzzy_C-Means

Université
Paris Cité

Image Segmentation with Fuzzy C-Means (FCM)

1. Goal

Implement the Fuzzy C-Means (FCM) clustering algorithm and apply it to image segmentation, producing fuzzy (soft) segmentation maps and hard segmentation derived from maximal membership. Both grayscale and RGB images are processed.

2. Algorithm

FCM minimizes the objective

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \left\| x_i - c_j \right\|^2 \quad , \quad 1 \le m < \infty$$

With updates

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m \cdot x_i}{\sum_{i=1}^{N} u_{ij}^m} \qquad u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\left\| x_i - c_j \right\|}{\left\| x_i - c_k \right\|} \right)^{\frac{2}{m-1}}}$$

A termination test stops when the change in membership matrix Ufalls below εor after a max iteration count.

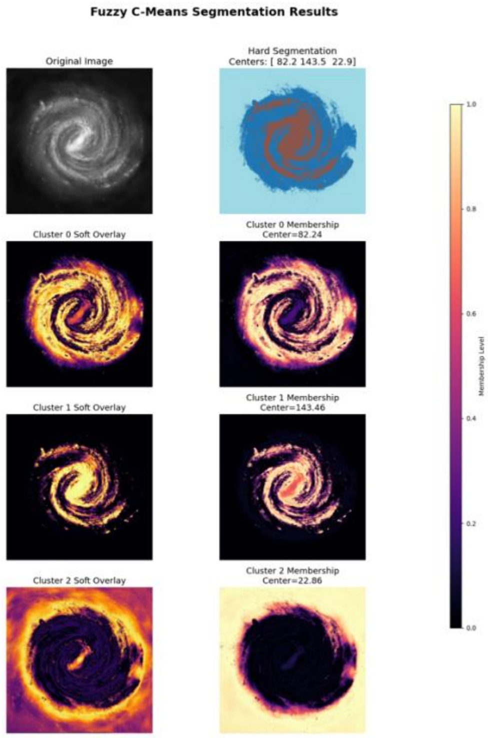
3. Implementation highlights

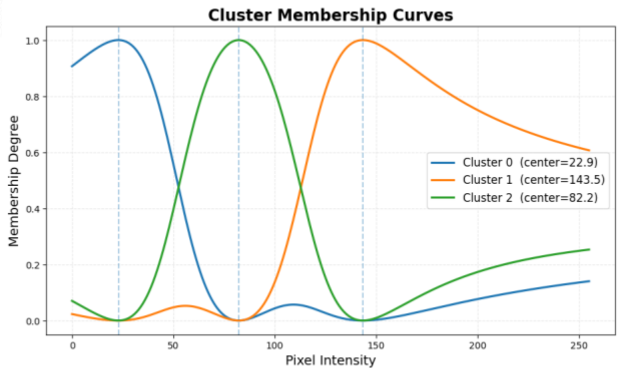
- Language / environment: Python, executed on Google Colab.
- Input handling:
    - Grayscale: pixels treated as scalar intensities (N = H×W data points).
    - RGB: pixels treated as 3-D vectors (N × 3 data matrix).
- Initialization: random / or seeded membership matrix (noted in the notebook comments).
- Stopping criterion: change in membership matrix norm $\| U^{(k+1)} - U^{(k)} \| < \varepsilon$(standard).
- Visual outputs produced:
    - Soft membership heatmaps per cluster (grayscale case).
    - Hard segmentation (argmax over membership).
    - Membership-versus-intensity curves (grayscale).
    - 3D scatterplots of RGB pixel positions colored by membership per cluster.

4. Experimental results (as shown in the notebook)

Grayscale segmentation

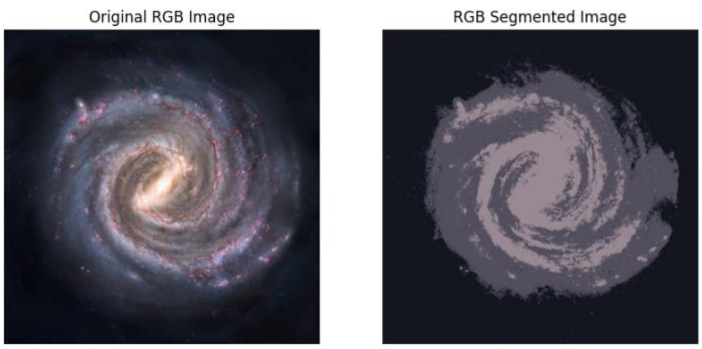
- Convergence: Algorithm converged at iteration 51.
- Final cluster centers (reported): [22.860821, 143.46290739, 82.23758839] (these are the three centers printed; ordering depends on cluster indexing).
- Run time: ~8 minutes on Google Colab for the grayscale pipeline.
- Visuals:

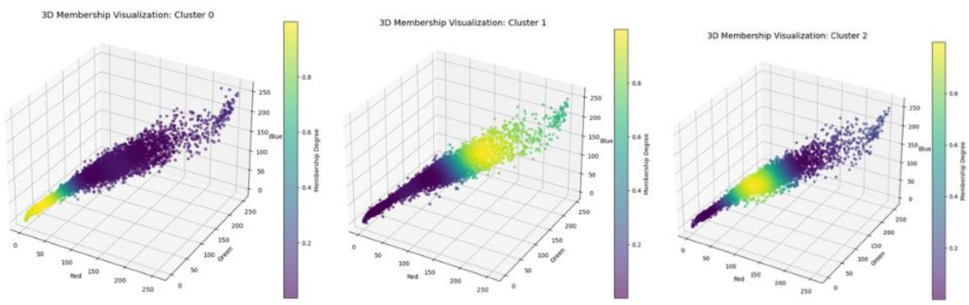




RGB segmentation

- Convergence: Converged at iteration 43
- Run time: ~44 minutes on Google Colab.
- Visuals:
  - Original RGB image and the derived RGB hard-segmented image
  - 3D scatter membership visualizations for each of the clusters (R,G,B axes) showing membership degree as color





The difference in runtime is expected, RGB clustering 3-dimensional features for each pixel ,processes more data and more expensive distance computations than grayscale.

5. Interpretation & discussion

- Convergence behavior: Both experiments converged reliably (iter ~40–50). The iteration counts are moderate; convergence speed depends on initialization, fuzzifier m, tolerance ε, and the image complexity.
- Centers & cluster meaning: The final centers indicate typical intensity (grayscale) or color prototypes (RGB). For grayscale we have one low center ~23, one mid ~82, and one high ~143, which aligns with dark, middle, and bright regions of the galaxy image.
- Soft segmentation advantage: Membership maps show meaningful fuzzy overlaps at region boundaries and gradations useful when pixel classes are ambiguous (e.g., edges, textured regions).
- RGB 3D views: The 3D membership scatterplots clearly show how clusters occupy bands in RGB-space; helpful to understand why some pixels end up with intermediate memberships.

6. Conclusion

The results show that Fuzzy C-Means successfully produces meaningful soft segmentations in both grayscale and RGB images, with grayscale heatmaps capturing smooth transitions across the galaxy structure and membership–intensity curves confirming consistent cluster behavior.

The RGB hard segmentation visually separates the galaxy from its background, while 3D membership scatterplots help illustrate how pixels distribute across color space. The implementation is correct and well-supported by visual diagnostics, though it comes with limitations: the RGB version is computationally expensive (44 minutes), the algorithm ignores spatial relationships between pixels sometimes resulting in small noisy regions and its sensitivity to initialization means different starting points can lead to different cluster outcomes.