Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

MACHINE LEARNING-DRIVEN FUZZY C-MEANS CLUSTERING FOR MEDICAL IMAGE SEGMENTATION

DIKSHA MALIK
Department of Mathematics, Chaudhary Charan Singh University, Meerut, U.P, India.

TARUN KUMAR
Department of Mathematics, Chaudhary Charan Singh University, Meerut, U.P, India.

SHUBHAM KUMAR
Department of Mathematics, Chaudhary Charan Singh University, Meerut, U.P, India.

AJENDRA SHARMA
Department of Mathematics, N. A. S. College, Chaudhary Charan Singh University, Meerut, U.P, India.

M. K. SHARMA *
Department of Mathematics, Chaudhary Charan Singh University, Meerut, U.P, India.
* Corresponding Author Email: drmukeshsharma@gmail.com
ORCID ID: 0000-0003-3071-5931

Abstract

In the contemporary landscape of machine learning, medical image analysis has experienced monumental leaps forward. Spearheading this progression are cutting-edge clustering and segmentation methods having reshaping our analytical capabilities. This piece delves into the prowess of the Fuzzy C-Means (FCM) clustering technique: a machine learning-centric strategy. By scrutinizing five diverse case studies, we unravel the tangible benefits and the expansive potential of FCM. To ensure a comprehensive view, we also navigate through other prominent image segmentation methodologies, including Thresholding, Watershed, and K-means clustering. To evaluate the resultant images from these methods, we have adeptly employed the fuzzy inference system (FIS). Our analytical juxtaposition underscores FCM's distinctive edge over other techniques, demonstrating its finesse in producing intricate and superior outcomes. Reinforcing the marriage of advanced technology with in-depth research, all our examinations and simulations were seamlessly executed by utilizing MATLAB's robust arsenal.

Keywords: Machine Learning, K-means Clustering, Fuzzy C-Means Clustering, Image Processing, Medical Image Segmentation.

1. INTRODUCTION

Over the past few decades, the realm of medical imaging has experienced a monumental metamorphosis. As we find ourselves in an era where precision and accuracy are paramount, the segmentation of these images emerges as a linchpin in ensuring impeccable diagnosis and enhancing patient care. The proliferation of technological advancements, especially in the age of machine learning, has bequeathed the medical fraternity with a plethora of techniques dedicated to image segmentation. The rich tapestry of literature in this domain is a testament to the relentless efforts of researchers and professionals aiming to refine these techniques for optimal outcomes. This article endeavours not only to traverse the expanse of these advancements but also to shed light on their intricate nuances and potential implications in the ever-evolving landscape of medical imaging.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

In the earlier stages, fuzzy technology found its relevance in healthcare and medicine (Abbod et al., 2001). This utility was further extended to the realms of fuzzy image processing schemes (Chacon et al., 2002) and its robust implementation in medical image segmentation, as demonstrated by Chuang et al. (2006). The evolution of fuzzy clustering for image segmentation has been meticulously studied by Naz et al. (2010) and corroborated by numerous others including Chaira (2011), Mohammed & Al-Ani (2020), Xu et al. (2021), and Dhal et al. (2023). The confluence of various algorithms has been explored in recent years, such as the integration of the watershed transform with fuzzy c-means clustering by Saikumar et al. (2012) and Bahadure et al. (2016). Anter & Hassenian (2019) took this a step further by incorporating neutrosophic sets in CT liver tumor segmentation. The convolutional neural network's efficacy in medical X-ray image segmentation has been presented by Bullock et al. (2019). Furthermore, the utility of MATLAB for image segmentation has been discussed by Abdulrahman & Varol (2020) and Ijemaru et al. (2021). Kumar et al. (2020) highlighted the significance of semi-supervised OTSU in dental radiograph segmentation.

In the last decade, there has been a shift towards intuitionistic methods, such as the one proposed by Chowdhary et al. (2020), and multiple thresholding techniques, as reviewed by Pare et al. (2020). Ramesh et al. (2021) offers a holistic review of medical image segmentation algorithms, and other reviews have delved into the nuances of various segmentation techniques (Kheradmandi & Mehranfar, 2022; Sarma & Gupta, 2021; Wala'a & Rana, 2021; Salpea et al., 2022; Csurka et al., 2022; Wang et al., 2022; Grewal et al., 2023; Yu et al., 2023).

Modern advancements in segmentation techniques like the spatial context model in fuzzy c-means clustering have been put forth by Xu et al. (2021). Meanwhile, Kaur et al. (2022) have demonstrated the integration of the watershed segmentation technique in detecting breast cancer masses in mammograms. Jardim et al. (2022) provide insights into the application of k-means clustering in graphical image region extraction. Recently, meta-heuristic optimization algorithms have shown promise in multilevel thresholding image segmentation, as evidenced by Abualigah et al. (2023). Nawaz et al. (2023) provided an innovative approach by combining fast fuzzy C-mean clustering-based maps for object detection and segmentation. Moussaoui et al. (2023) offer a compelling technique for brain tumour detection using Birch and Marker Watershed.

With such a rich tapestry of methodologies and approaches spanning over two decades, this article aims to weave a narrative around the advancements, applications, and future trajectories in medical image segmentation. This article delves into the nuanced application of the Fuzzy C-Means (FCM) clustering method within this sphere. Through the careful examination of four distinct case studies, we aim to furnish a thorough understanding of its practical ramifications. To provide a balanced perspective, we have also dissected other prominent image segmentation techniques, including thresholding, region growing, and watershed segmentation. By placing these methods side by side with the outcomes of the FCM clustering algorithm, a comparative analysis emerges. Our study strongly indicates the pre-eminence of the FCM clustering algorithm, showcasing its consistent ability to yield enhanced and intricate segmentation results. It is pivotal to

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

highlight that all analytical processes and simulations integral to our case studies were executed using MATLAB’s comprehensive toolkit.

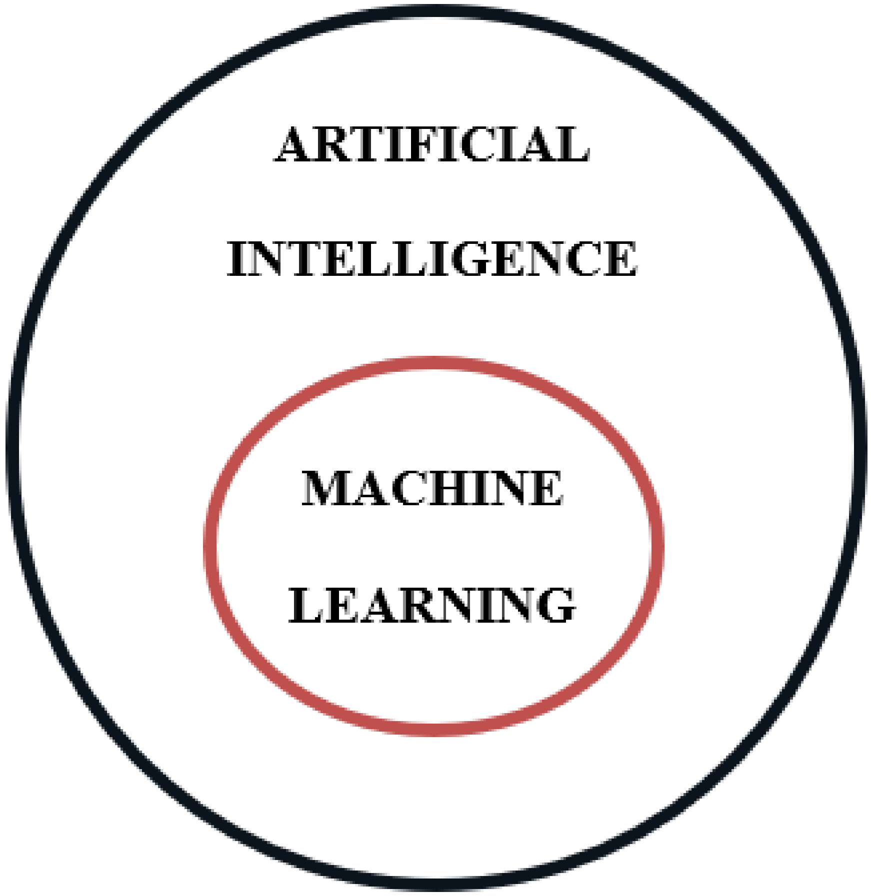
The architecture of this article is meticulously crafted into six detailed sections. Kicking off with the first section, we present a comprehensive introduction that delves into the backdrop and significance of medical imaging and its intersection with machine learning, setting the stage for the subsequent explorations. In the second section, we elaborate on and provide an in-depth explanation of the fundamental definitions pertinent to our study. Progressing to the third section, we demystify the underpinnings of our research by elaborating on the mathematical modelling integral to image segmentation techniques. In the fourth section, we pivot to an empirical dimension, presenting a series of case studies that exemplify real-world scenarios of medical image analysis, drawing heavily from the practical implications we have discussed earlier. The fifth section delves deeper, offering a granular interpretation of these case studies, highlighting challenges, successes, and key takeaways. Finally, in the sixth section where we synthesize our findings, discussions, and insights into a cohesive conclusion, providing a holistic perspective on the advancements and potential future trajectories in the domain of medical image segmentation.

2. BASIC DEFINITIONS

2.1 Machine Learning

Artificial Intelligence (AI) can be described as a human-inspired form of intelligence manifesting in machines and software. It's a groundbreaking domain within computer science that empowers machines to emulate human-like behaviour. The primary goal of AI is not just to simulate human reasoning but also to tackle intricate, real-world challenges more effectively. A crucial subset of AI is Machine Learning (ML). ML provides the framework that equips machines with the ability to learn and adapt. Instead of relying on explicit programming, these systems learn from past experiences, continually refining their approaches. This self-improving nature is instrumental in tasks such as prediction and classification. Primarily data-driven, modern ML aims to both categorize existing data based on established models and forecast future outcomes leveraging those models.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562



Phases of Machine Learning Execution

Machine learning can be streamlined into three distinct phases:

1. Training Phase: Here, the model is rigorously trained using a dedicated dataset, ensuring inputs align with their anticipated outputs. The primary focus during this phase is enabling the model to aptly classify and predict based on the provided data.
2. Validation and Testing Phase: In this stage, the trained model is evaluated using a distinct test dataset. This assessment determines the efficacy of the model, gauging how proficiently it has assimilated its training.
3. Deployment Phase: Once validated, the model is then introduced to real-world scenarios. It's tasked with gleaning insights and producing actionable outputs that can effectively address complex real-world challenges.

2.2 Types of Machine Learning

The following figure 1 display the flowchart of types and component of machine learning.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

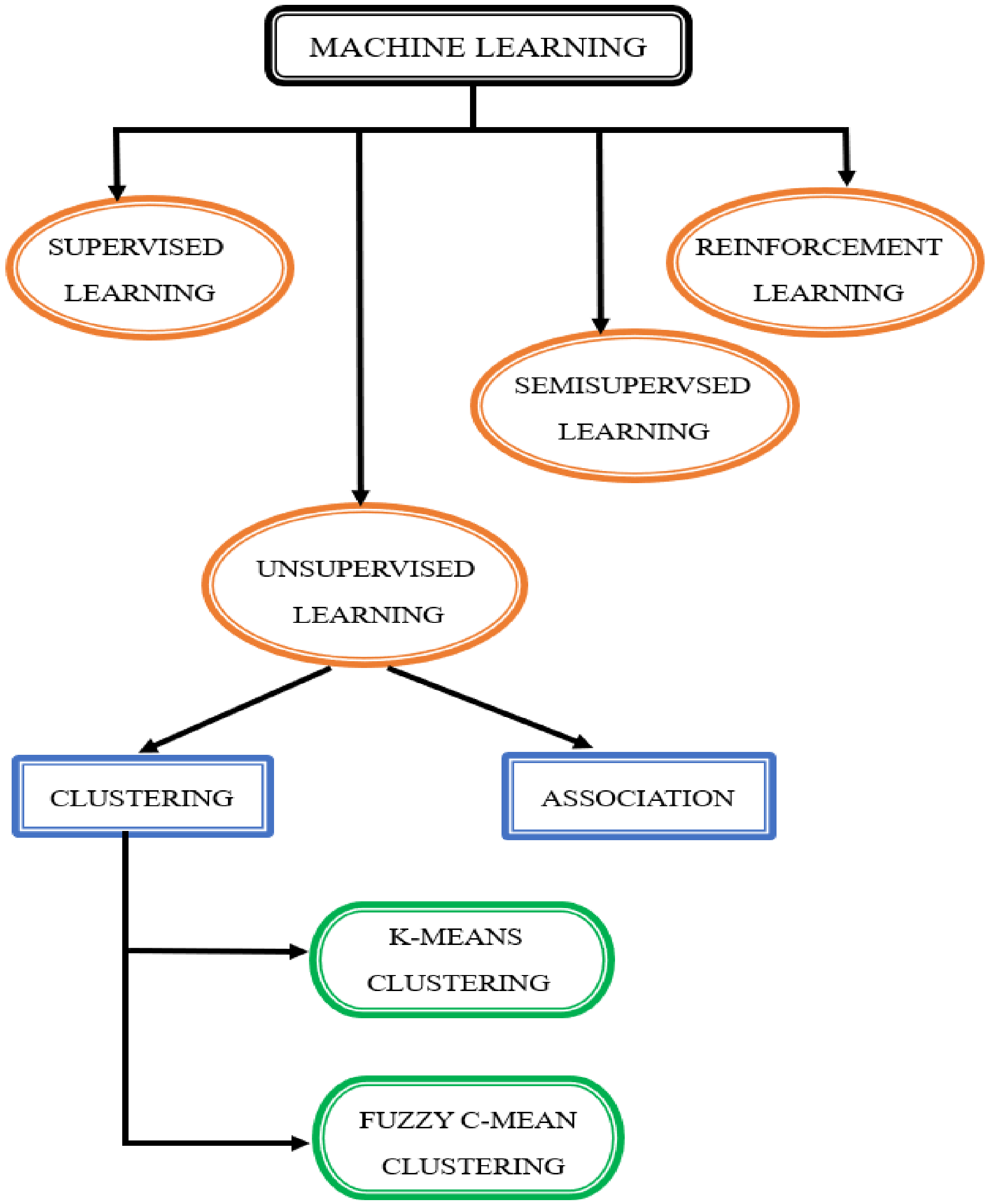


Figure 1: Flowchart of Types of Machine Learning

Machine learning encompasses a diverse array of learning techniques

2.2.1 Supervised Learning

This method revolves around training models using labelled datasets. A labelled dataset is characterized by both its input and the corresponding expected output. This dataset essentially guides the model, acting as its mentor, directing it towards making accurate predictions. The crux of supervised learning is pattern recognition. The model discerns intricate patterns within the data, forming associations with the specified outputs. The

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

foundational logic it develops regarding input-output relationships is determined by the algorithm in place and the quality of the labelled training data provided.

Post-training, the model undergoes a rigorous assessment using a separate labelled dataset, previously unseen during its training. Here, the model's mettle is tested as it endeavours to predict outcomes based on the relational logic it formulated during the training phase. Figure 2 shows the training and testing in supervised learning.

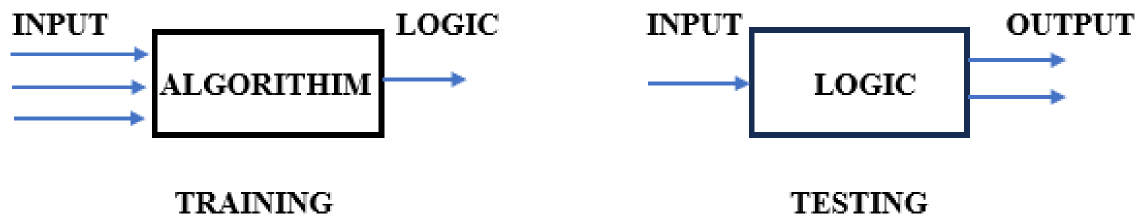


Figure 2: Training and Testing in Supervised Learning

2.2.2 Unsupervised Learning

This approach harnesses algorithms to analyse and cluster unlabelled datasets. Unlike its supervised counterpart, there's no guiding dataset during training. Instead, the model proactively discerns underlying patterns and insights from the data. Trained with unlabelled datasets, the models autonomously act on the data. When presented with new inputs, the algorithm categorizes them based on prior knowledge. Renowned for its capability to identify similarities and variances, unsupervised learning proves indispensable for tasks like exploratory data analysis, customer segmentation, cross-selling strategies, and image recognition.

Semi-supervised Learning: Positioned between supervised and unsupervised learning, semi-supervised learning seeks a middle ground. While supervised learning relies on expensive labelled data and unsupervised learning on unlabelled data, semi-supervised learning leverages both. This technique thrives in scenarios where there's a mix of abundant labelled and unlabelled data, aiming for optimal predictive performance.

2.2.3 Clustering in Machine Learning

Clustering, a fundamental technique in data analysis, serves as a beacon for those navigating the vast sea of data, guiding them towards meaningful groupings and connections. In this paradigm, individual data points coalesce into well-defined groups or clusters, based on their shared attributes. Intriguingly, while these members show profound affinities within their own group, they stand in stark contrast to members of other groups, underscoring the discriminative power of clustering.

Diving deeper, clustering finds its roots in the realm of unsupervised learning, a sector of machine learning where data is allowed to speak for itself, sans any pre-labelled instructions. This branch of learning is less about adhering to predefined labels and more about deciphering hidden narratives, discerning underlying structures, and detecting generative patterns nestled within the data. The impulse driving unsupervised clustering is fuelled by the insatiable curiosity to illuminate these unseen patterns, shedding light on

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

the intricate relationships that tie data points together. Now, when we bring the spotlight onto the dynamic domain of image segmentation, the role of clustering in machine learning becomes even more paramount. Image segmentation is akin to artfully segmenting a vast canvas into meaningful patches, each characterized by unique visual features. In this context, clustering algorithms, like K-means, become the artist's brush, segmenting images based on pixel intensities, textures, and colours, ensuring that similar features cluster together, painting a comprehensive picture.

Furthermore, the symbiosis between clustering and image segmentation transcends mere categorization. It is about interpreting visual data, decomposing images into understandable segments, and thus making them ripe for further analysis, be it in medical imaging to distinguish between healthy tissues and anomalies, or in satellite imagery to identify landforms. In essence, clustering in machine learning, especially in image segmentation, is a testament to the age-old adage - 'unity in diversity', bringing together like pixels while celebrating their distinction from the rest.

2.3 K-means Clustering

The K-means algorithm is a basic unsupervised learning method that addresses the clustering challenge. This algorithm categorizes untagged datasets into groups by determining the central point in the high-dimensional space. It divides 'n' data points into 'k' clusters, ensuring every data point is associated with the cluster whose mean is the closest, acting as its representative. Throughout this repetitive procedure, clusters are determined by calculating the shortest distance, typically the Euclidean distance, between data points and their centroids. In K-means clustering, data points are exclusively assigned to a single cluster, ensuring they are not part of any other group, thus it is often termed as a strict or definitive clustering method. Such algorithms aim to establish a clear separation within a dataset based on set criteria assessing the quality of the separation. This definitive separation implies that every single data point is unequivocally linked to just one cluster in the set.

Let's represent X as a dataset and consider x_i as an individual element within X. A partition P consisting of $\{C_1, C_2, \ldots, C_l\}$ from X is termed "strict" when:

- For every x_i in X, there is a corresponding C_j in P with x_i being a part of C_j.
- If x_i is part of C_j, then x_i is not a part of C_k where k is not equal to j, given both C_k and C_j are in P.

Steps for Implementing the K-means Clustering Method:

1. Decide on the number of clusters, denoted by 'K'.
2. Select 'K' initial centre points (not necessarily from the data points themselves).
3. Assign each data element to the closest centre point, creating 'K' distinct groups.
4. Calculate the mean of all data points within each cluster and redefine it as the new cluster centre.
5. Reallocate data points to the nearest updated centroid.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

6. If reallocations are made, revert to step-4; if not, proceed.
7. Model preparation is complete.

Unlabelled and labelled data shown in figure 3.

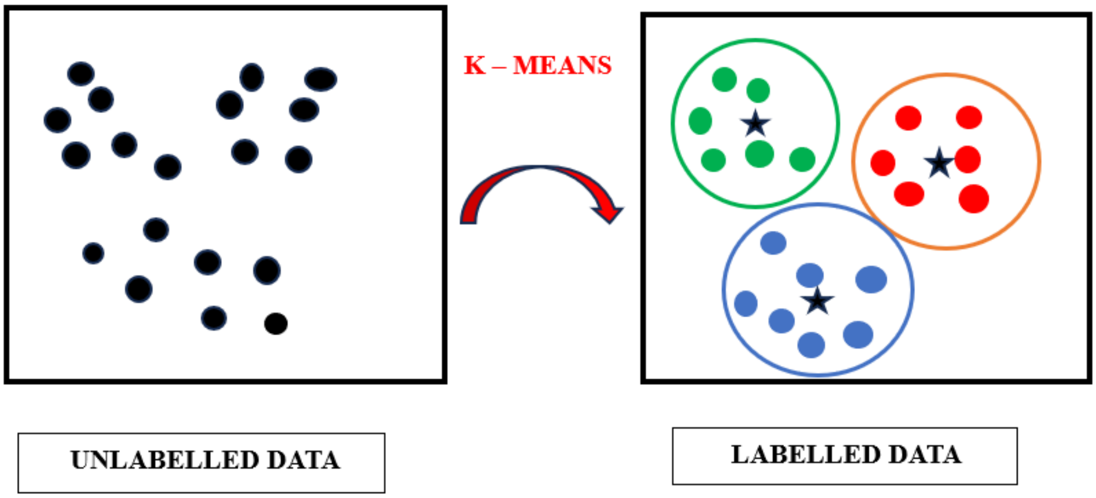


Figure 3: Unlabelled and Labelled Data

Figure 4 display the flowchart of K-means clustering.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

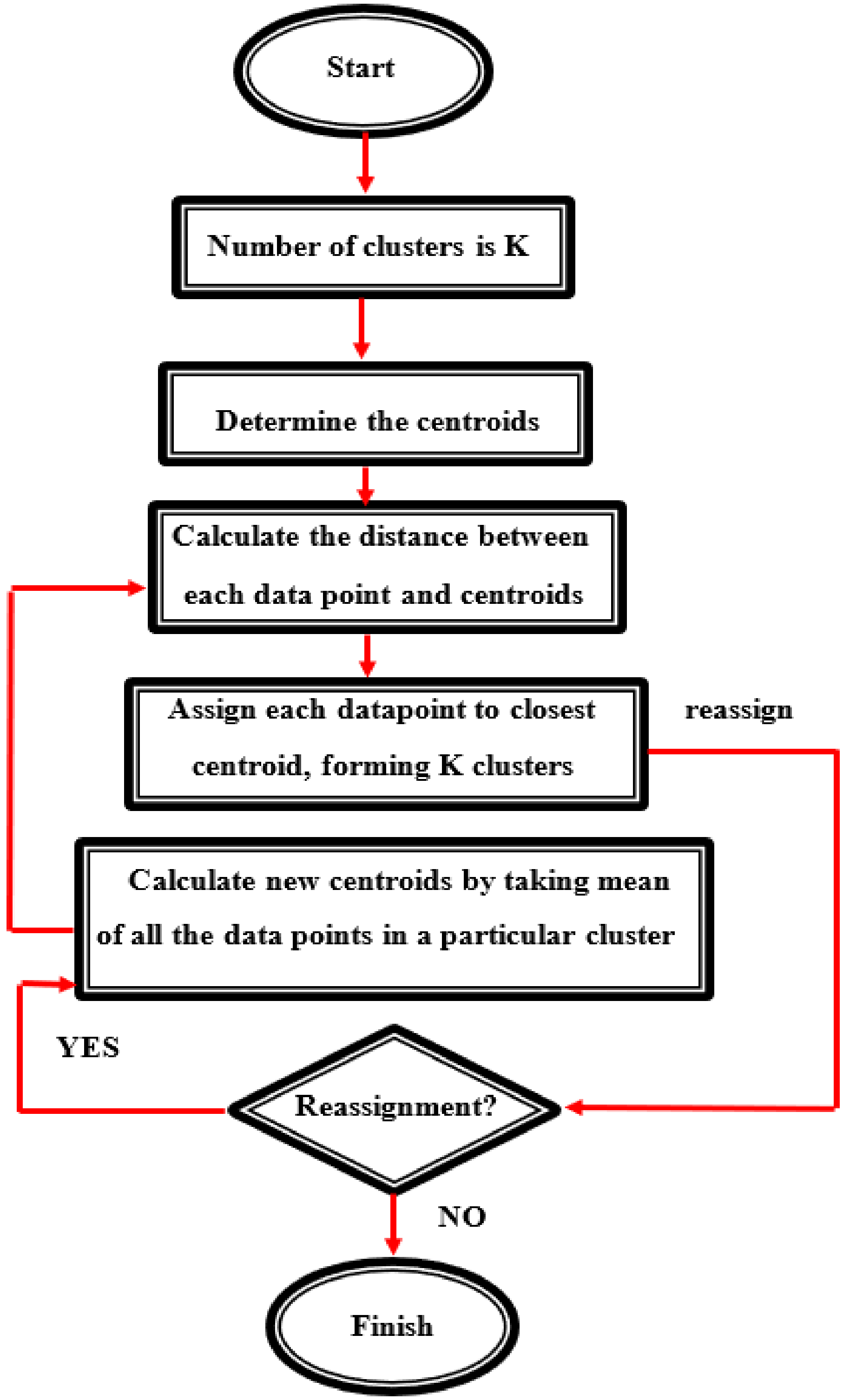


Figure 4: Flowchart of K-means clustering

2.4 Fuzzy c-mean Clustering

In numerous real-world grouping scenarios, certain data points may align with multiple clusters rather than being confined to a singular cluster. This is particularly true when data exhibits ambiguous characteristics, prompting the application of fuzzy clustering methods.

Take, for instance, an apple. Traditional clustering would categorize it as either red or green. Yet, in fuzzy clustering, an apple might simultaneously exhibit degrees of both colours: it could be 50% red and 50% green. While these values are normalized between 0 and 1, they are not probabilities, meaning they do not necessarily sum up to 1.

Such insights led to the inception of the “soft clustering” approach. This method seeks a “soft partition” of a dataset based on specific criteria, allowing data points to have affiliations with several clusters. The idea of a “soft partition” can be described as such:

Consider X as a dataset and x_i as an element within X. A partition P, consisting of {C_1, C_2, …, C_l}, is deemed “soft” if it satisfies the subsequent conditions:

- For every x_i in X and each C_j in P, the value of $\mu_{C_j}(x_i)$ lies between 0 and 1.
- For every x_i in X, there exists a C_j in P where $\mu_{C_j}(x_i)$ is greater than 0.

Here, $\mu_{C_j}(x_i)$ signifies the extent to which x_i is associated with the cluster C_j .

A specific variant of soft clustering ensures that the sum of the membership degrees of an element x across all clusters equals one. This can be represented as: the sum over j of $\mu_{C_j}(x_i)$ is 1 for every x_i in X.

A partition deemed “soft” that adheres to this added stipulation is termed a restricted soft partition. The fuzzy c-means technique, a prominent fuzzy clustering method, yields this kind of partition. It expands upon the c-means method, a traditional hard clustering technique introduced as part of the ISODATA clustering approach. The goal of the c-means method is to pinpoint clusters that are both compact and distinct from one another. Here is how distinct clusters are described:

Consider a partition P, consisting of $\{C_1, C_2, \dots, C_k\}$ from dataset X. This partition exhibits compact separation (CS) for clusters when, within the same cluster, any two points are nearer to each other than they are to points in different clusters. To put it mathematically: for any x and y in C_P, if $d(x, y)$ is less than $d(z, w)$, where z is in C_q and w is in C_r (with j not equal to k), then d stands for a specific distance metric.

Figure 5 display the hard and soft clustering and figure 6 display the flowchart of fuzzy c mean clustering in image segmentation.

3. MATHEMATICAL FORMULATION AND OPERATIONAL PROCEDURE

The fuzzy c-means method (FCM) extends the capabilities of the conventional c-mean algorithm by enabling data points to have fractional memberships across various clusters. As a result, it generates a soft categorization for a given dataset. More precisely, it yields

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

a restricted soft partition. This adaptation from the objective function J_1 of the hard c-mean is realized through:

- Integrating degrees of fuzzy membership within clusters into the equation.
- Introducing a new parameter, m, to act as a weight exponent for fuzzy membership.

The refined objective function, symbolized as J_m, can be expressed as:

$$J_m(\mathrm{P},\mathrm{V}) = \sum_{i=1}^{k} \sum_{x_k \epsilon X} (\mu_{C_i}(x_k))^m \,||x_k - v_i||^2$$

Here, P represents the fuzzy categorization of the dataset X composed of C_1,C_2 …, C_k. The parameter *m* acts as a weighting factor, influencing the impact of data points fractional membership on the final clustering outcome.

Similar to the hard c-means approach, fuzzy c-means aims to identify an optimal partition by seeking prototypes v_i that reduce the value of the objective function J_m. Yet, in contrast to the hard c-means method, fuzzy c-means also endeavors to determine membership functions μ_{C_i} that further minimize J_m. To achieve these goals, a requisite condition for the local minimum of J_m was extracted from J_m. This stipulation, which we detail subsequently, underpins the fuzzy c-means methodology.

3.1 Fuzzy c-Means Theorem

A constrained fuzzy partition encompassing {C_1,C_2, …, C_k} can attain a local minimum of the function J_m if and only if it adheres to these conditions:

$$\mu_{C_i}(x) = \frac{1}{\sum_{j=1}^{k}\left(\frac{\left\|x-v_i\right\|^2}{\left\|x-v_j\right\|^2}\right)^{1/m-1}} \quad 1 \le i \le k, x \text{ in } X \tag{1}$$

$$v_i = \frac{\sum_{x \in X}(\mu_{C_i}(x))^m * x}{\sum_{x \in X}^{n}(\mu_{C_i}(x))^m} \quad 1 \le i \le k \tag{2}$$

Leveraging this principle, FCM iteratively refines the prototypes and the membership function, utilizing the provided equations, until a specified convergence point is achieved.

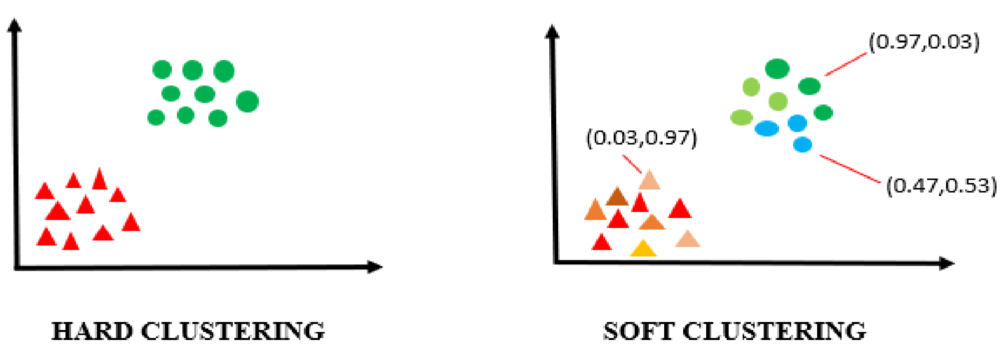


Figure 5: Hard and Soft Clustering

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

EXAMPLE

	F_1	F_2
x_1	2	12
x_2	4	9
x_3	7	13
x_4	11	5
x_6	12	7
x_7	14	4

We have a dataset composed of six data points, as outlined in the previous table. Each of these points is defined by two attributes, F_1 and F_2. To categorize this data into two distinct clusters (taking $c = 2$), we'll employ the Fuzzy C-Means method. We'll fix the parameter m at 2 for this FCM application. As starting points for our cluster centers, we'll utilize the prototypes v_1= (5,5) and v_2 = (10,10).

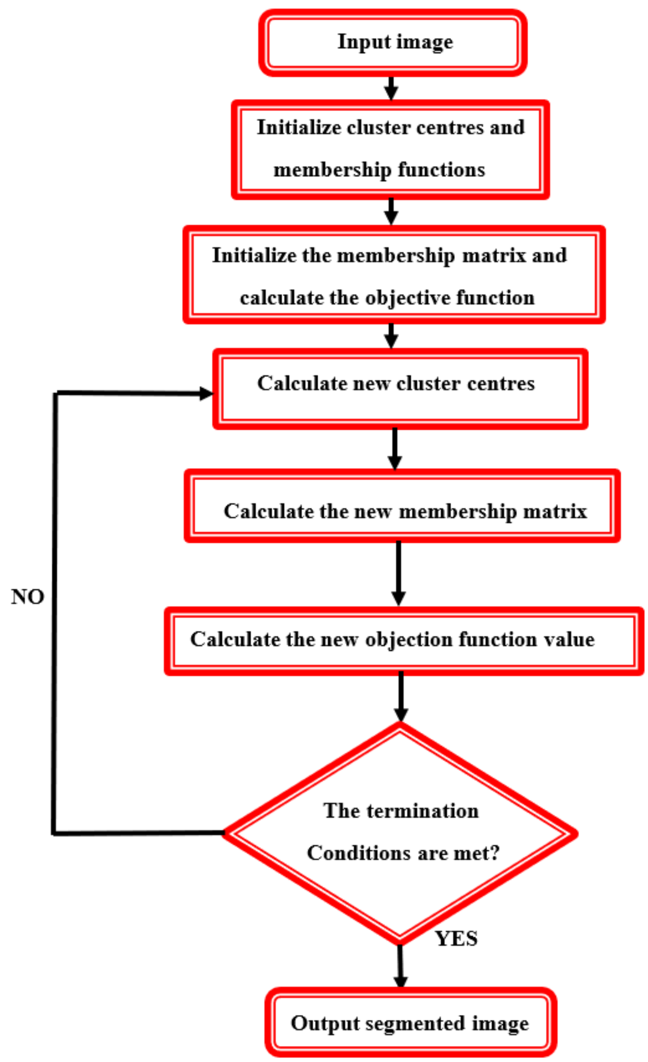


Figure 6: Flowchart of Fuzzy C-mean Clustering

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

Solution: Using the previously mentioned equation (1), we computed the starting membership values for the two clusters.

$$\mu_{C_1}(x_1) = \frac{1}{\sum_{j=1}^{2}(\frac{||x_1 - v_1||}{||x_1 - v_j||})^2}$$

$||x_1 - v_1||^2 = 3^2 + 7^2 = 9 + 49 = 58$

$||x_1 - v_2||^2 = 8^2 + 2^2 = 64 + 4 = 58$

$\mu_{C_1}(x_1) = \frac{1}{\frac{58}{58}+\frac{58}{68}} = \frac{1}{1+0.853} = 0.5397$

In a similar manner, we derive the following:

$\mu_{C_2}(x_1) = \frac{1}{\frac{68}{58}+\frac{68}{68}} = 0.4603$

$\mu_{C_1}(x_2) = \frac{1}{\frac{17}{17}+\frac{17}{37}} = 0.6852$ $\mu_{C_2}(x_2) = \frac{1}{\frac{37}{17}+\frac{37}{37}} = 0.3148$

$\mu_{C_1}(x_3) = \frac{1}{\frac{68}{68}+\frac{68}{18}} = 0.2093$ $\mu_{C_2}(x_3) = \frac{1}{\frac{18}{68}+\frac{18}{18}} = 0.7907$

$\mu_{C_1}(x_4) = \frac{1}{\frac{36}{36}+\frac{36}{26}} = 0.4194$ $\mu_{C_2}(x_4) = \frac{1}{\frac{26}{36}+\frac{26}{26}} = 0.5806$

$\mu_{C_1}(x_5) = \frac{1}{\frac{53}{53}+\frac{53}{13}} = 0.197$ $\mu_{C_2}(x_5) = \frac{1}{\frac{13}{53}+\frac{13}{13}} = 0.803$

$\mu_{C_1}(x_6) = \frac{1}{\frac{82}{82}+\frac{82}{52}} = 0.3881$ $\mu_{C_2}(x_6) = \frac{1}{\frac{52}{82}+\frac{52}{52}} = 0.6119$

Consequently, with the starting prototypes for both clusters, the membership function suggests that x_1 and x_2 align more closely with the first cluster, whereas the other points in the dataset lean more towards the second cluster. Subsequently, the FCM algorithm refines the prototypes based on the previously mentioned equation (2):

$$v_1 = \frac{\sum_{k=1}^{6}(\mu_{C_1}(x_k))^2 * x_k}{\sum_{k=1}^{6}(\mu_{C_1}(x_k))^2}$$

$$= \frac{0.5397^2*(2,12) + 0.6852^2*(4,9) + 0.2093^2*(7,13) + 0.4194^2*(11,5) + 0.197^2*(12,7) + 0.3881^2*(14,4)}{0.5397^2 + 0.6852^2 + 0.2093^2 + 0.4194^2 + 0.197^2 + 0.3881^2}$$

$= (\frac{7.2761}{1.0979}, \frac{10.044}{1.0979})$

= (6.6273, 9.1484)

$$v_2 = \frac{\sum_{k=1}^{6}(\mu_{C_2}(x_k))^2 * x_k}{\sum_{k=1}^{6}(\mu_{C_2}(x_k))^2}$$

$$= \frac{0.4603^2*(2,12) + 0.3148^2*(4,9) + 0.7909^2*(7,13) + 0.5806^2*(11,5) + 0.803^2*(12,7) + 0.6119^2*(14,4)}{0.4603^2 + 0.3148^2 + 0.7909^2 + 0.5806^2 + 0.803^2 + 0.6119^2}$$

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

$= (\frac{22.326}{2.2928}, \frac{19.4629}{2.2928}) = (9.7374, 8.4887)$

The modified prototype v_1 shifts nearer to the midpoint of the cluster comprising x_1,x_2, and x_3. Meanwhile, the refined prototype v_2 gravitates towards the cluster that includes x_4 ,x_5, and x_6.

4. CASE STUDIES

We have meticulously gathered X-ray images from a cohort of five patients, aiming to delve deeper into advanced segmentation techniques. Illustrated below are the results of our comprehensive analysis, which encompassed several state-of-the-art methodologies: thresholding, watershed segmentation, K-means clustering, and Fuzzy C-means clustering. This segmentation showcase is not only a testament to the sophistication of these techniques but also demonstrates their potential applicability in real-world medical imaging scenarios.

Figure 7 to 11 display the x-ray image and output of image segmentation methods of five patients.

Patient 1

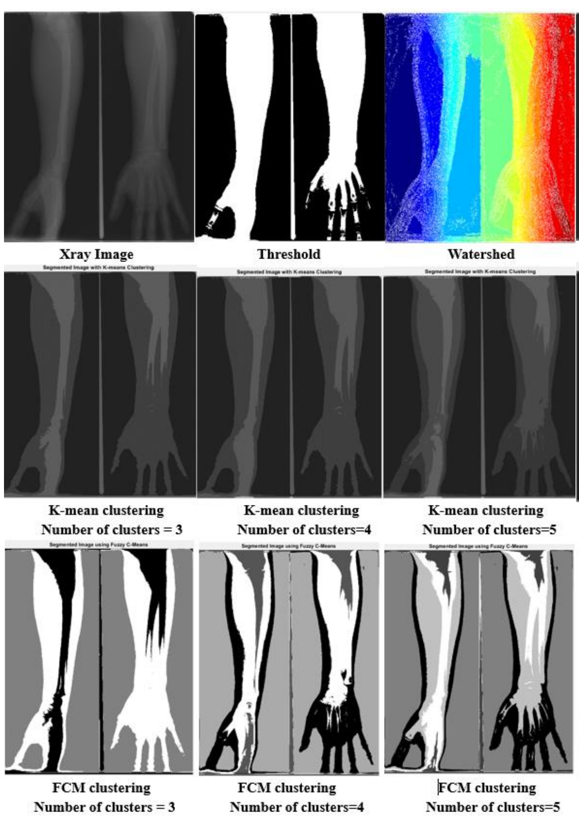


Figure 7: Xray Image and output of segmentation methods

Patient 2

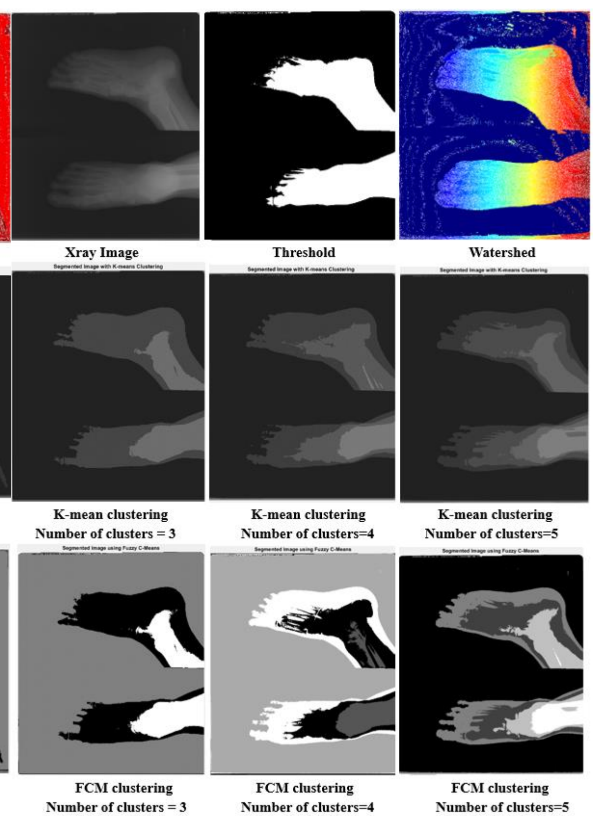


Figure 8: Xray Image and output of segmentation methods

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

Patient 3

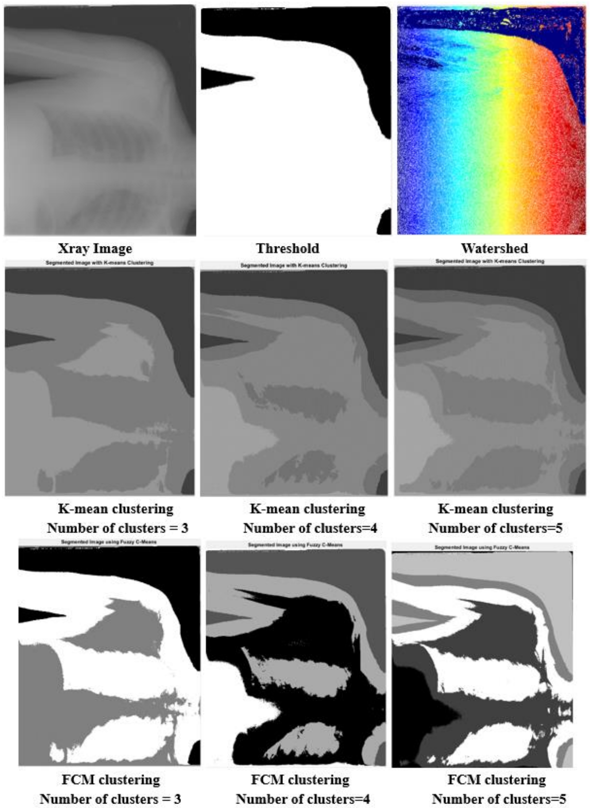


Figure 9: Xray Image and output of segmentation methods

Patient 4

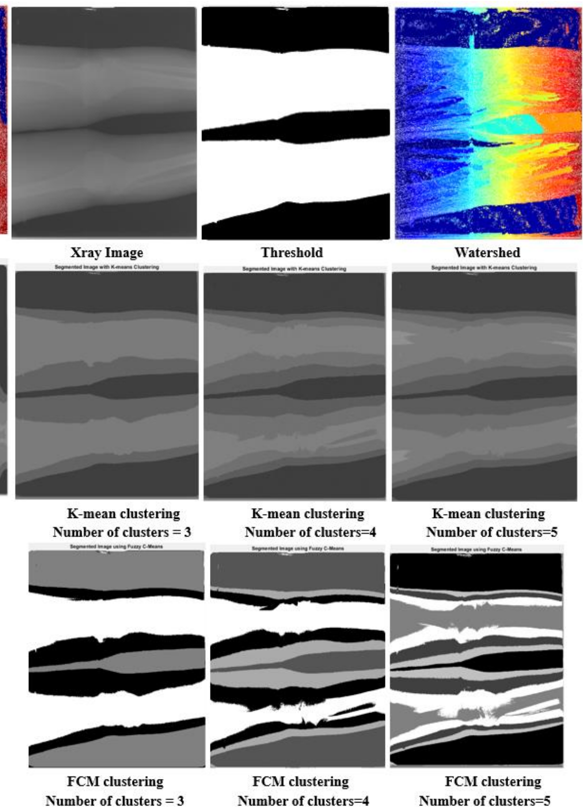


Figure 10: Xray Image and output of segmentation methods

Patient 5

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

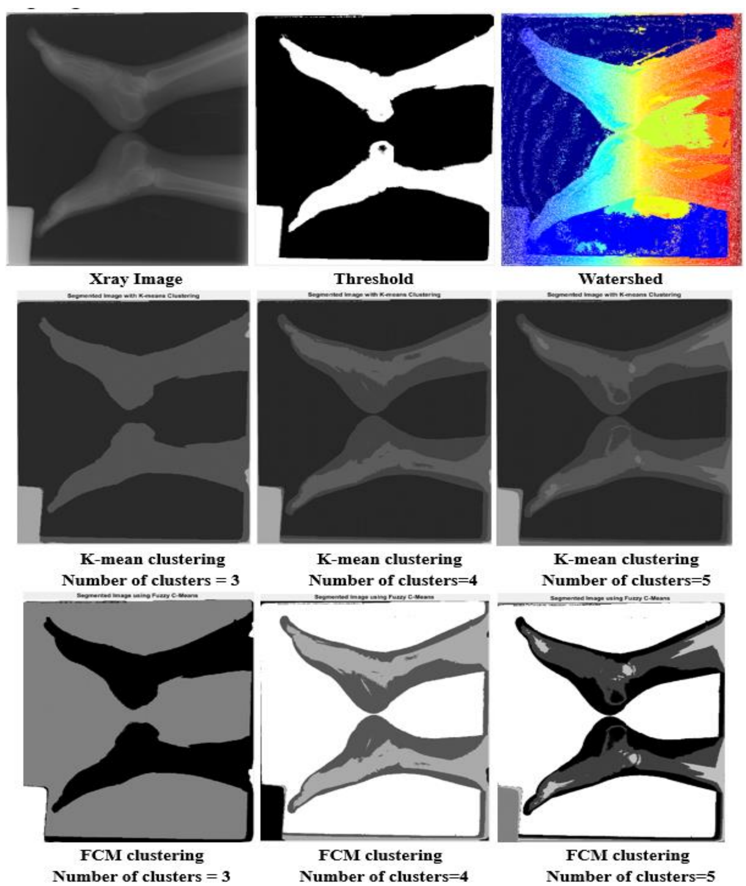


Figure 11: Xray Image and output of segmentation methods

In the comprehensive evaluation of the obtained patient images, we have strategically implemented a fuzzy inference system (FIS). This decision underscores our commitment to precision and depth in analysis. The detailed methodology we followed using this system is presented below:

In Table 1, we present the various input factors, each accompanied by their respective linguistic variables, offering a clear delineation of the parameters under consideration. Moving on to Table 2, we have meticulously constructed fuzzy rules, leveraging these aforementioned input factors. It is important to note that our primary output factor in this analysis is the ‘image improvement score’. Subsequent to our rigorous implementation of the Fuzzy Inference System (FIS) in MATLAB, employing both the designated input factors and our crafted fuzzy rules, we derived a set of image improvement scores. These scores, which evaluate the effectiveness of different image segmentation methods, are comprehensively detailed in Table 3.

Table 1

	Input factors		
Edge Clarity	Low (0-0.4)	Medium (0.2-0.8)	High (0.6 and above)
Region Uniformity	Poor (0-0.4)	Average (0.3-0.7)	Good (0.6 and above)
Contrast Quality	Weak (0-0.5)	Moderate (0.2-0.8)	Strong (0.58 and above)

ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

Table 2

	Input factors			Output
	Edge clarity	Region Uniformity	Contrast Quality	Image Improvement Score
1	Low	Poor	Weak	Poor
2	Low	Average	Moderate	Poor
3	Medium	Poor	Weak	Poor
4	Medium	Average	Moderate	Average
5	High	Good	Strong	Good
6	High	Poor	Weak	Average
7	Medium	Good	Strong	Average
8	Low	Good	Strong	Average
9	High	Average	Moderate	Good
10	Low	Average	Strong	Average

Table 3

Patients → / Methods ↓	Patient 1 (Image Improvement Score)	Patient 2 (Image Improvement Score)	Patient 3 (Image Improvement Score)	Patient 4 (Image Improvement Score)	Patient 5 (Image Improvement Score)
Threshold	0.399	0.332	0.341	0.276	0.283
Watershed	0.263	0.406	0.259	0.234	0.371
KM, Clusters=3	0.500	0.352	0.319	0.329	0.303
KM, Clusters=4	0.500	0.383	0.465	0.416	0.437
KM, Clusters=5	0.565	0.500	0.536	0.535	0.520
FCM, Clusters=3	0.633	0.328	0.506	0.500	0.488
FCM, Clusters=4	0.706	0.490	0.549	0.552	0.500
FCM, Clusters=5	0.828	0.623	0.646	0.668	0.607

5. RESULTS INTERPRETATION AND DISCUSSION

In our meticulous analysis of imaging characteristics, as highlighted in Table 1, we carefully delineated three critical input factors: ‘Edge Clarity’, ‘Region Uniformity’, and ‘Contrast Quality. These were subsequently translated into descriptive linguistic terms using pre-established value ranges. To illustrate, the ‘Edge Clarity parameter is segmented such that values not exceeding 0.4 are labelled as ‘Low’. In contrast, values that ascend beyond the 0.6 mark are distinctly categorized as ‘High’. Progressing to the fuzzy rule matrix exhibited in Table 2, a prominent observation emerges: images marked by a ‘Low’ edge clarity, in conjunction with ‘Poor’ uniformity and ‘Weak’ contrast, are unequivocally assigned a ‘Poor’ score in terms of image improvement. On the other end of the spectrum, images that manifest a congruous amalgamation of ‘High’ edge clarity, complemented by ‘Good’ region uniformity and ‘Strong’ contrast quality, consistently earn a commendable ‘Good’ rating. Venturing further into the nuanced, patient-centric data

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

illustrated in Table 3, a consistent pattern becomes evident: the Fuzzy C-Means (FCM) segmentation approach, notably when utilizing 5 clusters, often eclipses other methodologies in procuring the most favourable image improvement scores across diverse patient samples. Such observations elevate the critical importance of elements like contrast quality and edge clarity as pivotal influencers shaping the overarching success of the chosen segmentation strategy. For a visual representation that juxtaposes the various image segmentation methodologies against the resultant image improvement scores garnered from patient samples, one can refer to the comparative graph in Figure 12.

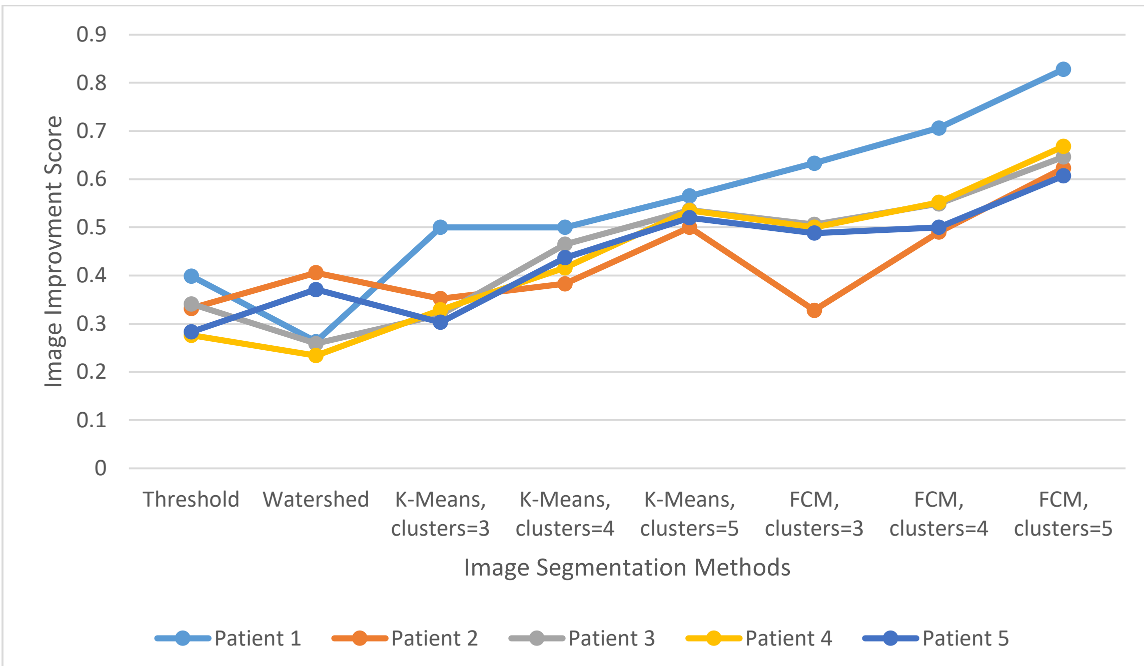


Figure12: Comparison Between Image Segmentation Methods And Image Improvement Scores

Among all the techniques, fuzzy c means clustering algorithm appeared to be the most effective for segmenting x-ray images of the patients under study. The choice of segmentation method and the number of clusters should be tailored based on the specific anatomical structures or pathologies that need to be highlighted in the x-ray images. It is also essential to consider the computational efficiency and time taken by each method, especially if the intention is real-time or near-real-time processing.

6. CONCLUSION

As we journey through the evolving realms of machine learning and its transformative influence on medical image analysis, it becomes abundantly clear that innovative clustering and segmentation methods are setting new benchmarks in the field. Among these, the Fuzzy C-Means (FCM) clustering technique stands out, demonstrating not just its theoretical relevance but also its practical applicability, as evidenced by our deep dive into five varied case studies. While we have explored multiple segmentation methodologies such as Thresholding, Watershed, and K-means clustering, FCM consistently showcased its upper hand, reinforcing its credibility and potential in the medical imaging domain. The fuzzy inference system (FIS) provided an apt evaluation mechanism, further strengthening our analyses. The culmination of our research underscores the pivotal role of FCM in advancing medical image analysis, producing results of exceptional depth and precision. Furthermore, the seamless integration of MATLAB in our study exemplifies the symbiotic relationship between modern technology and rigorous academic research, paving the way for future explorations in this domain.

ACKNOWLEDGEMENTS

The second author acknowledges receiving financial assistance from the University Grants Commission (UGC).

This work has been carried out under the University Grant Research Scheme Ref. Number Dev./1043 dated 29.06.2022.

Conflict of Interest Statement: The authors state that they have no conflicts of interest.

References

1) Abbod, M. F., von Keyserlingk, D. G., Linkens, D. A., & Mahfouf, M. (2001). Survey of utilisation of fuzzy technology in medicine and healthcare. *Fuzzy Sets and Systems*, *120*(2), 331-349. https://doi.org/10.1016/S1386-5056(01)00160-5

2) Chacon, M. I., Aguilar, L., & Delgado, A. (2002, October). Definition and applications of a fuzzy image processing scheme. In *Proceedings of 2002 IEEE 10th Digital Signal Processing Workshop, 2002 and the 2nd Signal Processing Education Workshop.* (pp. 102-107). IEEE. DOI: 10.1109/DSPWS.2002.1231085

3) Chuang, K. S., Tzeng, H. L., Chen, S., Wu, J., & Chen, T. J. (2006). Fuzzy c-means clustering with spatial information for image segmentation. *computerized medical imaging and graphics*, *30*(1), 9-15.

4) https://doi.org/10.1016/j.compmedimag.2005.10.001

5) Naz, S., Majeed, H., & Irshad, H. (2010, October). Image segmentation using fuzzy clustering: A survey. In *2010 6th international conference on emerging technologies (ICET)* (pp. 181-186). IEEE. DOI: 10.1109/ICET.2010.5638492

6) Chaira, T. (2011). A novel intuitionistic fuzzy C means clustering algorithm and its application to medical images. *Applied soft computing*, *11*(2), 1711-1717. https://doi.org/10.1016/j.asoc.2010.05.005

7) Saikumar, T., Nagarani, M., Yojana, K., & Shashidhar, B. (2012, March). Image segmentation of an adaptive threshold algorithm using watershed transform and fuzzy c-means clustering on level set method. In *IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM-2012)* (pp. 33-38). IEEE. https://ieeexplore.ieee.org/abstract/document/6215968/

8) Bahadure, N. B., Ray, A. K., & Thethi, H. P. (2016, March). Performance analysis of image segmentation using watershed algorithm, fuzzy c-means of clustering algorithm and simulink design. In *2016 3rd international conference on computing for sustainable global development (INDIACom)* (pp. 1160-1164). IEEE. https://ieeexplore.ieee.org/abstract/document/7724447/

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

9) Anter, A. M., & Hassenian, A. E. (2019). CT liver tumor segmentation hybrid approach using neutrosophic sets, fast fuzzy c-means and adaptive watershed algorithm. *Artificial intelligence in medicine*, *97*, 105-117. https://doi.org/10.1016/j.artmed.2018.11.007

10) Bullock, J., Cuesta-Lázaro, C., & Quera-Bofarull, A. (2019, March). XNet: a convolutional neural network (CNN) implementation for medical x-ray image segmentation suitable for small datasets. In *Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging* (Vol. 10953, pp. 453-463). SPIE. 1812.00548.pdf (arxiv.org)

11) Kumar, A., Bhadauria, H. S., & Singh, A. (2020). Semi-supervised OTSU based hyperbolic tangent Gaussian kernel fuzzy C-mean clustering for dental radiographs segmentation. *Multimedia Tools and Applications*, *79*, 2745-2768. https://link.springer.com/article/10.1007/s11042-019-08268-8

12) Chowdhary, C. L., Mittal, M., P, K., Pattanaik, P. A., & Marszalek, Z. (2020). An efficient segmentation and classification system in medical images using intuitionist possibilistic fuzzy C-mean clustering and fuzzy SVM algorithm. *Sensors*, *20*(14), 3903. https://doi.org/10.3390/s20143903

13) Abdulrahman, A., & Varol, S. (2020, June). A Review of Image Segmentation Using MATLAB Environment. In *2020 8th International Symposium on Digital Forensics and Security (ISDFS)* (pp. 1-5). IEEE.

14) https://ieeexplore.ieee.org/abstract/document/9116191/

15) Pare, S., Kumar, A., Singh, G. K., & Bajaj, V. (2020). Image segmentation using multilevel thresholding: a research review. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, *44*, 1-29.

16) https://link.springer.com/article/10.1007/s40998-019-00251-1

17) Ijemaru, G. K., Nwajana, A. O., Oleka, E. U., Otuka, R. I., Ihianle, I. K., Ebenuwa, S. H., & Obi, E. R. (2021). Image processing system using MATLAB-based analytics. *Bulletin of Electrical Engineering and Informatics*, *10*(5), 2566-2577. https://doi.org/10.11591/eei.v10i5.3160

18) Ramesh, K. K. D., Kumar, G. K., Swapna, K., Datta, D., & Rajest, S. S. (2021). A review of medical image segmentation algorithms. EAI Endorsed Transactions on Pervasive Health and Technology, 7(27), e6-e6. https://doi.org/10.4108/eai.12-4-2021.169184

19) Oskouei, A. G., Hashemzadeh, M., Asheghi, B., & Balafar, M. A. (2021). CGFFCM: Cluster-weight and Group-local Feature-weight learning in Fuzzy C-Means clustering algorithm for color image segmentation. *Applied Soft Computing*, *113*, 108005. https://doi.org/10.1016/j.asoc.2021.108005

20) Zhao, D., Liu, L., Yu, F., Heidari, A. A., Wang, M., Liang, G., ... & Chen, H. (2021). Chaotic random spare ant colony optimization for multi-threshold image segmentation of 2D Kapur entropy. *Knowledge-Based Systems*, *216*, 106510. https://doi.org/10.1016/j.knosys.2020.106510

21) Xu, J., Zhao, T., Feng, G., Ni, M., & Ou, S. (2021). A fuzzy C-means clustering algorithm based on spatial context model for image segmentation. *International Journal of Fuzzy Systems*, *23*, 816-832.

22) https://link.springer.com/article/10.1007/s40815-020-01015-4

23) Sarma, R., & Gupta, Y. K. (2021). A comparative study of new and existing segmentation techniques. In *IOP conference series: materials science and engineering* (Vol. 1022, No. 1, p. 012027). IOP Publishing. https://iopscience.iop.org/article/10.1088/1757-899X/1022/1/012027/meta

24) Wala'a, N. J., & Rana, J. M. (2021). A survey on segmentation techniques for image processing. *Iraqi Journal for Electrical and Electronic Engineering*, *17*, 73-93. (iasj.net)

25) Kheradmandi, N., & Mehranfar, V. (2022). A critical review and comparative study on image segmentation-based techniques for pavement crack detection. *Construction and Building Materials*, *321*, 126162.

26) https://doi.org/10.1016/j.conbuildmat.2021.126162

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 10-2023
DOI: 10.5281/zenodo.8429562

27) Jardim, S., António, J., & Mora, C. (2022). Graphical image region extraction with k-means clustering and watershed. *Journal of Imaging*, *8*(6), 163. https://doi.org/10.3390/jimaging8060163

28) Kaur, A., Rashid, M., Bashir, A. K., & Parah, S. A. (2022). Detection of breast cancer masses in mammogram images with watershed segmentation and machine learning approach. In *Artificial Intelligence for Innovative Healthcare Informatics* (pp. 35-60). Cham: Springer International Publishing.

29) https://link.springer.com/chapter/10.1007/978-3-030-96569-3_2

30) Salpea, N., Tzouveli, P., & Kollias, D. (2022, October). Medical image segmentation: A review of modern architectures. In *European Conference on Computer Vision* (pp. 691-708). Cham: Springer Nature Switzerland.

31) https://link.springer.com/chapter/10.1007/978-3-031-25082-8_47

32) Csurka, G., Volpi, R., & Chidlovskii, B. (2022). Semantic image segmentation: Two decades of research. *Foundations and Trends® in Computer Graphics and Vision*, *14*(1-2), 1-162. https://www.nowpublishers.com/article/Details/CGV-095

33) Wang, R., Lei, T., Cui, R., Zhang, B., Meng, H., & Nandi, A. K. (2022). Medical image segmentation using deep learning: A survey. *IET Image Processing*, *16*(5), 1243-1267. https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/ipr2.12419

34) Abualigah, L., Almotairi, K. H., & Elaziz, M. A. (2023). Multilevel thresholding image segmentation using meta-heuristic optimization algorithms: Comparative analysis, open challenges and new trends. *Applied Intelligence*, *53*(10), 11654-11704. https://link.springer.com/article/10.1007/s10489-022-04064-4

35) Grewal, R., Kasana, S. S., & Kasana, G. (2023). Hyperspectral image segmentation: a comprehensive survey. *Multimedia Tools and Applications*, *82*(14), 20819-20872. https://link.springer.com/article/10.1007/s11042-022-13959-w

36) Nawaz, M., Qureshi, R., Teevno, M. A., & Shahid, A. R. (2023). Object detection and segmentation by composition of fast fuzzy C-mean clustering based maps. *Journal of Ambient Intelligence and Humanized Computing*, *14*(6), 7173-7188. https://link.springer.com/article/10.1007/s12652-021-03570-6

37) Dhal, K. G., Das, A., Sasmal, B., Ray, S., Rai, R., & Garai, A. (2023). Fuzzy C-Means for image segmentation: challenges and solutions. *Multimedia Tools and Applications*, 1-37. https://link.springer.com/article/10.1007/s11042-023-16569-2

38) Yu, Y., Wang, C., Fu, Q., Kou, R., Huang, F., Yang, B., ... & Gao, M. (2023). Techniques and challenges of image segmentation: A review. *Electronics*, *12*(5), 1199. https://doi.org/10.3390/electronics12051199

39) Moussaoui, H., El Akkad, N., & Benslimane, M. (2023). A Brain Tumor Segmentation and Detection Technique Based on Birch and Marker Watershed. *SN Computer Science*, *4*(4), 339. https://link.springer.com/article/10.1007/s42979-023-01802-4