MINISTRY OF HIGHER EDUCATION

AND SCIENTIFIC RESEARCH

KASDI MERBAH UNIVERSITY OUARGLA

Faculty of Applied Sciences

Department of Electronics and Telecommunications

Memory

ACADIMIC MASTER

Field Science and Technology

Branch: Automation

Specialty: Automation and Systems

Presented by:

Abdennour Djari

Theme:

Medical Image Segmentation Using Multikernel Method

The Jury composed of:

MR.MANSEUR	A/GHANI	MAA (UKMO)	PRESIDENT
MR.KADRI	SALIM	MAA (UKMO)	EXAMINER
MR.NASRI	NADGIB	MAA (UKMO)	SUPERVISOR

Academic Year 2023/2024

ACKNOWLEDGEMENTS.

I would like to express my deepest gratitude to everyone who contributed to the completion of this memoir.

First and foremost, I sincerely thank my memoir supervisor, Nasri Nadjib, for his guidance, expertise, and invaluable advice throughout this project. Your support and availability have been indispensable, and I have learned a great things from your insightful direction.

I also wish to extend my gratitude to The Jury of KASDI MERBAH UNIVERSITY OUARGLA for the excellent education I received. Your passion for your fields of expertise has been a source of inspiration for me, and I have greatly benefited from your extensive knowledge.

I am profoundly grateful to my loved ones, family, and friends who supported me throughout this academic journey. Your moral support, constant encouragement, and understanding during stressful times have been essential to my success. I am extremely thankful for your unwavering love and support.

Finally, I wish to express my gratitude to everyone who shared their knowledge and contributed in any way to the completion of this thesis. Your ideas, suggestions, and discussions have broadened my perspective and deepened my understanding of the subject.

DEDICATION

I dedicate this memoir to my family, who supported me throughout this academic journey. Your love, encouragement, and unwavering support were essential in overcoming challenges and accomplishing this important milestone in my life.

I also wish to express my gratitude to my friends and loved ones who supported and encouraged me throughout this project. Your presence and moral support were invaluable and kept me motivated during this process.

A big thank you to my professors and supervisors for their guidance, advice, and valuable expertise. Their support and constructive feedback greatly contributed to the improvement of this memoir.

This memoir is dedicated to all of you. Your unwavering support has been the key to my success, and I am eternally grateful to you.

Table of contents

Chapter I: Clustering Techniques in Medical Image Analysis

Chapter II: Kernel Methods

Chapter III: Simulation Results and Discussions

General Conclusion

References

List of figures and tables

ABSTRACT:

"This work introduces the Multi-Kernel Fuzzy C-Means (MKFCM) algorithm for medical image segmentation, demonstrating its superior performance over traditional Fuzzy C-Means (FCM) and Kernel Fuzzy C-Means (KFCM). By integrating multiple kernels, MKFCM effectively handles complex data distributions, noise, achieving higher accuracy and robustness. Quantitative evaluations using Dice Coefficient and Intersection over Union (IOU) scores confirm MKFCM's enhanced segmentation capabilities, making it a highly effective tool for precise medical imaging."

Keywords: *Multi-Kernel Fuzzy C-Means, Segmentation, Medical Image.*

Résumé :

"Ce travail présente l'algorithme Multi-Kernel Fuzzy C-Means (MKFCM) pour la segmentation d'images médicales, démontrant sa performance par rapport aux algorithmes Fuzzy C-Means (FCM) et Kernel Fuzzy C-Means (KFCM) traditionnels. En intégrant plusieurs noyaux, MKFCM traite efficacement les distributions de données complexes, atteignant une précision et une robustesse accrues. Les évaluations quantitatives utilisant le coefficient de Dice et l'Intersection sur l'Union (IOU) confirment les capacités de segmentation améliorées de MKFCM, en faisant un outil très efficace pour l'imagerie médicale."

Mots clees*: Multi-Kernel Fuzzy C-Means (MKFCM), Medical Image Segmentation.*

الملخص:

"تقدم هذه الدراسة خوارزمية التكتل الضبابي متعددة النواة (MKFCM) لتقسيم الصور الطبية، مما يُظهر أداءها المتفوق على الخوارزميات التقليدية (FCM) و(KFCM). من خلال دمج نوى متعددة، تتعامل MKFCM بفعالية مع توزيعات البيانات المعقدة، والضوضاء، والتشويش، محققة دقة وقوة أكبر. تؤكد التقييمات الكمية باستخدام معامل Dice وتقاطع الاتحاد (IOU) على قدرات التقسيم المحسنة لـ MKFCM، مما يجعلها أداة فعالة للغاية لتصوير طبي دقيق."

الكلمات المفتاحية *:خوارزمية التكتل الضبابي متعددة النواة (MKFCM), تقسيم الصور الطبية.*

General Introduction

General Introduction

The need to precisely identify anatomical structures and diseased regions for better diagnosis and treatment planning, has made medical picture segmentation a crucial area of study, in recent years. Our work explores the novel use of a multikernel method for medical image segmentation, to improve the precision and effectiveness. We start by going over basic ideas like K-Means Algorithm and fuzzy set theory. Next, we go over more complex approaches like the Fuzzy c-Means Algorithm (FCM) and Kernel Methods. Finally, we introduce an algorithm called Multi-Kernel Based Fuzzy c-Means (MKFCM).

The fuzzy c-Means (FCM) algorithm is an effective clustering approach that assigns membership levels to each data point in relation to cluster centers. Because of its effectiveness in handling the inherent ambiguity and uncertainty present in medical imaging, this method has a long history and is widely used.
In the present work, we will go over FCM's algorithmic structure, historical evolution, and advantages and disadvantages. Especially when dealing with high-dimensional feature spaces, where linear differentiation of data points becomes an issue.
Kernel methods provide the solution to this shortcoming. Indeed, they convert the input data into a higher-dimensional space, which allows linear separation and more precise grouping.

Our work is divided into three parts, in the first chapter; we began by presenting unsupervised methods. Next, we introduce the k-means algorithm, the fuzzy set theory and FCM algorithm.

In the second chapter, we explored the Kernel version of FCM algorithm. The Multi-Kernel Based Fuzzy c-Means (MKFCM) algorithm is also introduced. By using numerous kernels to capture different data features, the latter algorithm provides a more flexible and reliable clustering framework. A comparison using a single kernel has been shown.

In the third chapter, we presented the results of all algorithms, applied on medical images.

To validate the efficacy of each algorithm, we employed standard metrics in medical image segmentation such as Dice coefficient and Intersection over union (IOU) score. These metrics provide numerical evaluation of the accuracy of segmentation, comparing the segmented output with ground truth data.

Chapter I: Clustering Techniques in Medical Image Analysis

I.1. Introduction

Medical image analysis leverages clustering techniques to enhance diagnostic accuracy. This chapter delves into clustering methodologies, focusing on their application in medical imaging. It explores the basics of medical image analysis, the principles of clustering in this context, and details the K-means and Fuzzy C-means algorithms, highlighting their significance and implementation in medical diagnostics.

I.2. Medical Image Analysis

Medical imaging has developed exponentially in the last few years in terms of technological advance and wide-spread use. High-resolution, three-dimensional anatomical information can now be obtained in a routine manner with magnetic resonance imaging (MRI) and X-ray computer-aided tomography (CT). These two modalities provide complementary information; CT shows detail of bony structures and some contrast between hard and soft tissues while MRI shows detail of soft tissue structures, with almost no detail of bony structures.

CT imaging, like all X-ray techniques, exposes the patient to a dose of X-rays, thus, incurring some health risks. MRI does not expose the patient to radiation, but uses the magnetic properties of the patient’s tissues to provide contrast in the image, and as far as we know at present it is completely harmless.

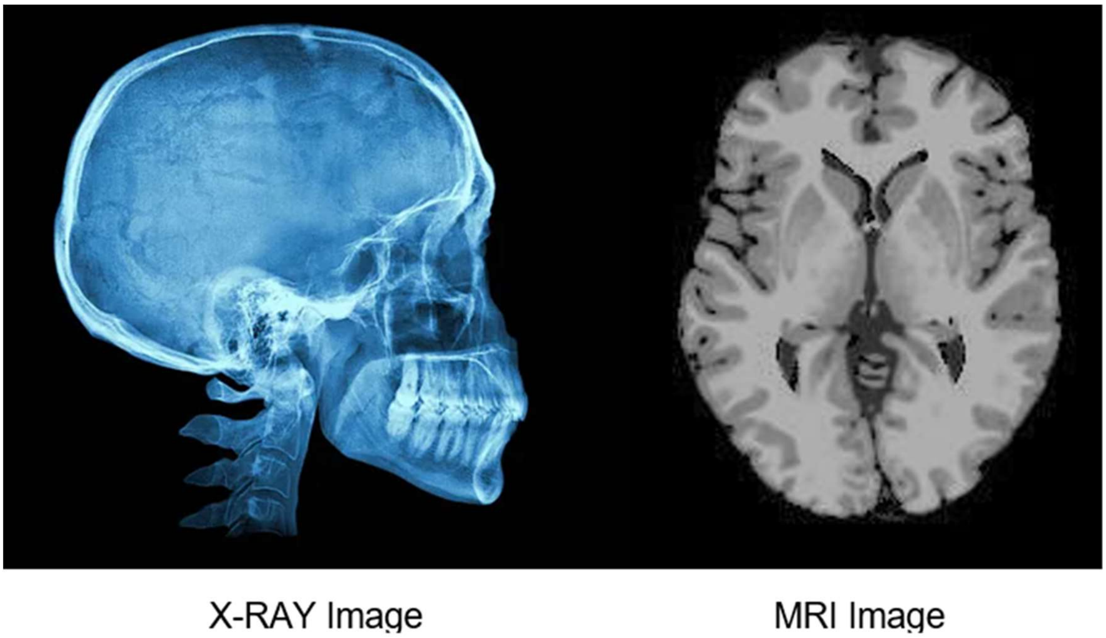


Figure I.1: *Medical Image Example*

I.3. Clustering of Medical Image:

Research on Clustering is well-established; it dates back to the 1950s and is widely reported in various journals. The research problem is concerned with discovering a grouping structure within a number of objects.

Typically, a set of numeric observations, or features, are collected of each object. The collected feature-sets are aggregated into a list which then acts as the input to a chosen computational clustering algorithm. This algorithm then provides a description of the grouping structure which it has discovered within the objects. The clusters would be identified by labels usually supplied by the user.

In this way, a large number of seemingly disparate objects, once a number of features are extracted of them, can be organized into groups of approximately shared features. Data clustering gained initial formal treatment as a sub-field of statistics. Systematic methods of clustering were required to be developed because the data may be large in size and therefore cumbersome to analyse and visualize.

Dealing with grayscale images in medical image segmentation presents significant challenges. The lack of color information makes distinguishing between different tissues and structures more difficult, often resulting in lower contrast and less defined boundaries. This increases the complexity of accurately segmenting regions of interest, necessitating advanced algorithms and preprocessing techniques to enhance image quality and segmentation precision.

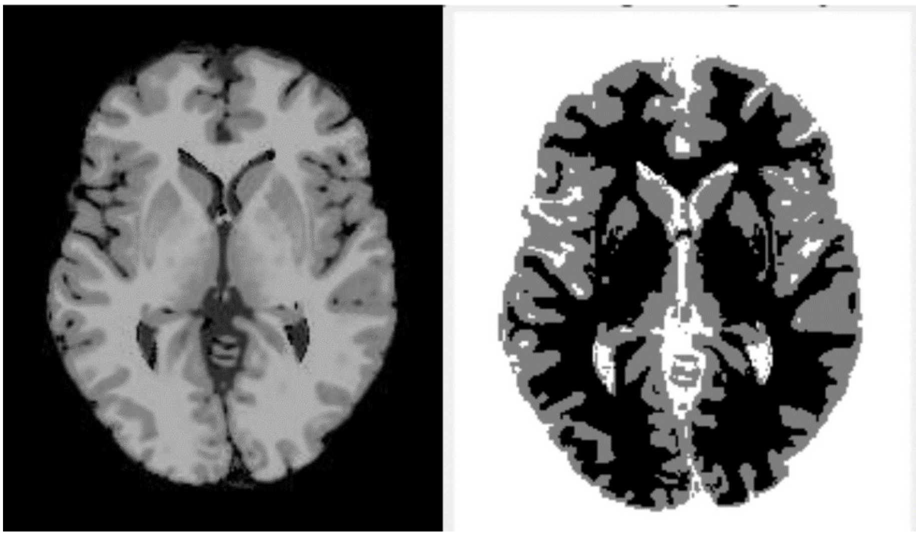


***Figure I.2**: Example of Image Clustering*

I.3.1. Clustering Methods:

Medical image clustering utilizes both supervised and unsupervised methods to organize and interpret complex imaging data. In supervised clustering, algorithms are trained on labeled datasets where the desired output is known, making it easier to identify and classify specific structures within the images. Techniques like Convolutional Neural Networks (CNNs) are commonly used, leveraging labeled medical images to improve diagnostic accuracy and segmentation precision.

On the other hand, unsupervised clustering does not rely on pre-labeled data. Methods such as K-means and Fuzzy C-means are popular, as they group similar pixel intensities to differentiate between tissues and structures without prior knowledge . Hierarchical clustering builds nested clusters by successively merging or splitting them, while density-based clustering identifies clusters based on data density, useful for detecting anomalies.

Both supervised and unsupervised clustering methods play a crucial role in enhancing the interpretability of medical images, aiding in accurate diagnosis, treatment planning, and research.

I.4. K-Means Algorithm:

K-means is one of the simplest unsupervised learning algorithms that solves the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori [1].

The main idea is to define k centroids, one for each cluster. These centroids should be placed in a clever way because of different locations cause different results. So, the better choice is to place them as far away as possible from each other. The next step is to take each point belonging to a given data set and associate it with the nearest centroid [2].

At this point, we need to re-calculate k new centroids of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. A loop has been generated. As a result of this loop, we may notice that the k centroids change their location step by step until no more changes are made.

Finally, this algorithm aims at minimizing an objective function, The objective function

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x^{(j)} - c_j \right\|^2 \qquad (I.1)$$

Where $||x_i^{(j)} - c_j||$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster center c_j, is an indicator of the distance of the *n* data points from their respective cluster centres.

I.4.1. Visualizing K-means Clustering

K-means clustering produces a very nice visual, so here is a quick example of how each step might look.

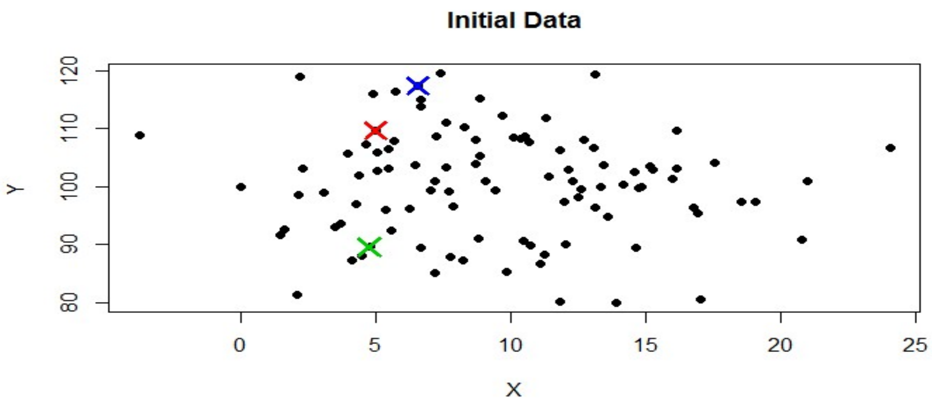


Figure I.3: *Initial 50 Data Points*

Here's 50 data points with three randomly initiated centroids.

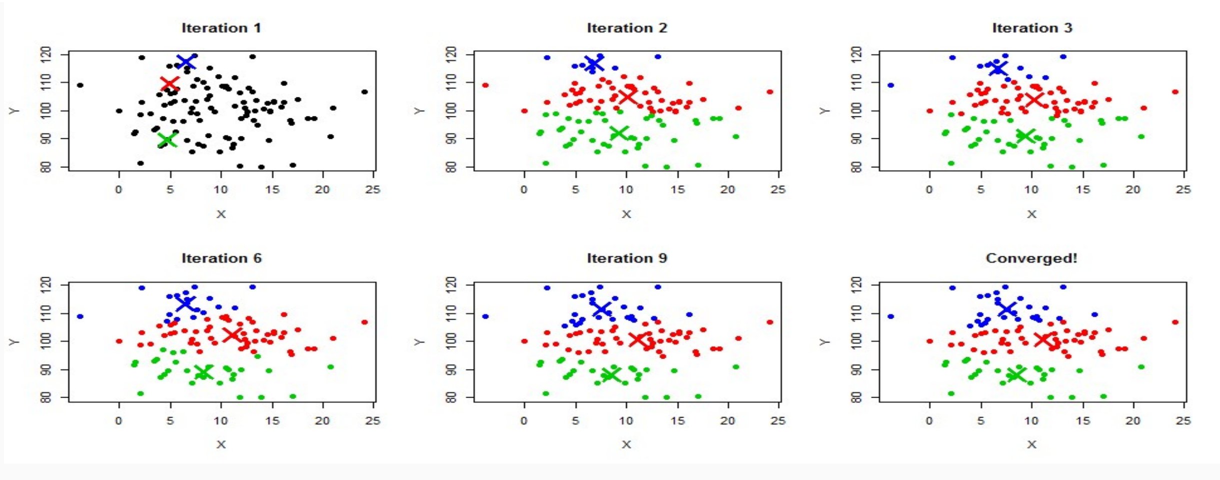


***Figure I.4:** 9 Iterations of K-means*

- Iteration 2 shows the new location of the centroid centers.
- Iteration 3 has a handful more blue points as the centroids move.
- Jumping to iteration 6, we see the red centroid has moved further to the right.
- Iteration 9 shows the green section is much smaller than in iteration 2, blue has taken over the top, and the red centroid is thinner than in iteration 6.
- The 9th iteration's results were the same as the 8th iteration, so it has "converged.”.

I.4.2. K-means Advantages:

- The K-Means algorithm is very fast.
- It is robust and easier to understand.
- The number of clusters is known.

I.4.3. K-means Disadventages:

- K-Means algorithm requires a priori specification of the number of cluster centers.

- If there are two highly overlapping data then the k-means algorithm will not be able to resolve that there are two clusters and is said to be the use of exclusive assignment.
- Data represented in the form of cartesian coordinates and polar coordinates will give different results.
- It provides the local optima of the squared error function.
- Randomly choosing the cluster center lead to a poor result.
- Applicable only when the mean is defined.
- Unable to handle noisy data and outliers.
- It fails for non-linear data sets.

I.5. Fuzzy C-means

I.5.1. Fuzzy Set Theory:

Fuzzy Set Theory was developed by Lotfi Zadeh 1965 in order to describe mathematically, the imprecision or vagueness that is present in our everyday language. Imprecisely defined classes play an important role when humans communicate and learn.

IN order to deal with these classes, Zadeh introduced the concept of fuzzy set. They are similar to ordinary mathematical sets but are more general than them.

In the fuzzy clustering setting, a cluster is viewed as a fuzzy set in the data set, X . Thus, each feature vector in the data set will have membership values with all clusters membership, indicating a degree of belonging to the cluster under consideration. The goal of a given fuzzy clustering method will be to define each cluster by finding its membership function.

In the general case, the fuzzy sets framework provides a way of dealing with problems in which the source of imprecision is the absence of sharply defined criteria of class membership rather than the presence of random variables. Fuzzy clustering fits well with the rest of the fuzzy sets and systems applications. It has been used with success in, for example, optimizing membership functions for forming fuzzy inferences. [3]

Fuzzy set theory is widely used as a modeling tool in various Pattern Recognition and image analysis problems, [4], because of the relative ease and the robustness to noise, low solution cost and better human computer interaction.

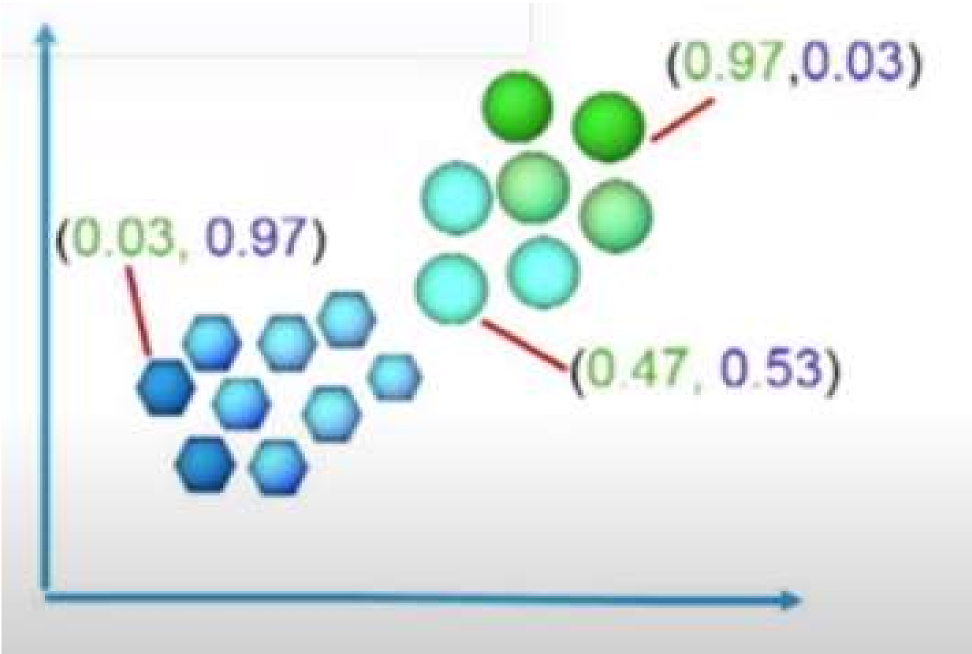


***Figure I.5**: Membership Concept*

I.5.2 Fuzzy c-Means Algorithm:

The FCM algorithm took several names before FCM. These include Fuzzy ISODATA and Fuzzy k-Means. The idea of using fuzzy set theory for clustering is credited to Ruspini 1970. Dunn is credited with the first specific formulation of FCM 1973, but its generalization and current framing are credited to Bezdek 1981.

FCM is a clustering algorithm that provides a fuzzy partition of the input data set. However, there is an infinite range of possible fuzzy partitions. Therefore, FCM aim to to search for the optimal partition according to the chosen objective function. FCM is, thus, first and foremost an objective function.

The way that most researchers have solved the optimization problem has been through an iterative locally optimal technique, called the FCM algorithm. This is not the only way to solve the FCM objective function, for example, in AlSultan & Selim, 1993, it is solved by the Simulated Annealing Optimization technique; in Hathaway & Bezdek, 1995, the problem is reformulated and general optimization methods are suggested for its solution; in Al-Sultan & Fedjki, 1997 it is solved by a combinatorial optimization technique called Tabu Search; in Hall et al., 1999 it is solved by the genetic algorithm, which is an optimization technique based on evolutionary computation; and in Runkler & Bezdek, 1999 it is solved within an alternate optimization framework. In fact, it is not impossible that an exact solution to the problem may be formulated.

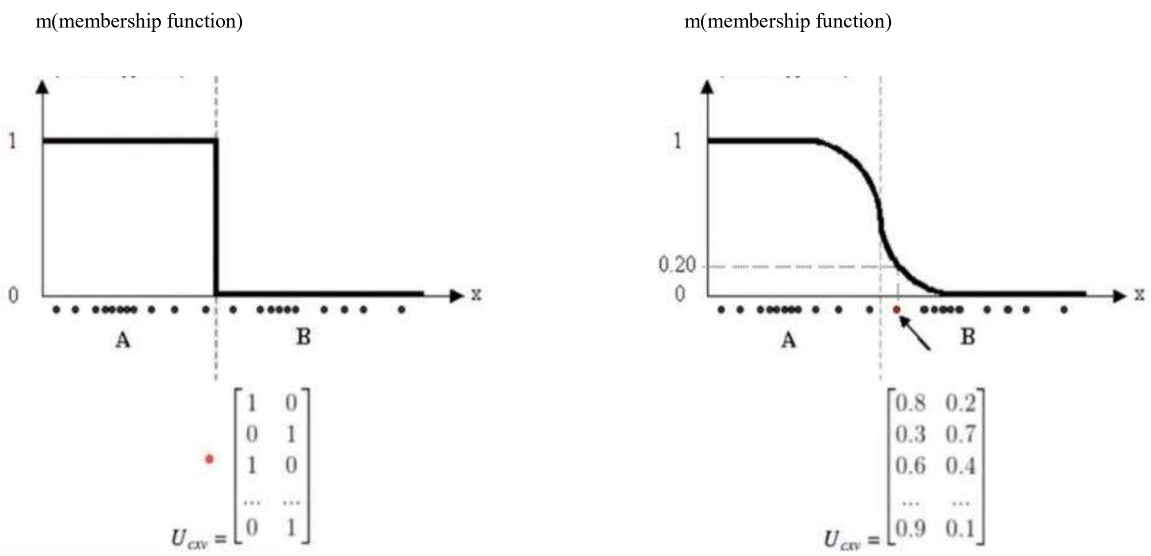


Figure I.6: *The Difference Between K-means and FCM*

I.5.3 The FCM Formulation:

Fuzzy c-means clustering technique is a simplification of the hard c-means algorithm that yields enormously superior results in image region clustering and object categorization. As in hard k-means algorithm, Fuzzy c-means algorithm is based on the minimization of a standard function.

Let a matrix of n data elements (image pixels), each of sizes s (s = 1), is represented as $X=(x_1,x_2,\ldots\ldots,x^n)$ FCM generates the clustering by iteratively minimizing the objective function given in:

- **Objective function:**

$$O_m(U,C)=\sum_{i=1}^{c}\sum_{j=1}^{n} U_y^m D^2(x_j,C_i) \quad \text{(I.2)}$$

- **Constraint**:

$$\sum_{i=1}^{c} U_y = 1;\ \forall j \quad \text{(I.3)}$$

Where U_{ij} is membership of the j^{th} data in the i^{th} cluster C^i , m is fuzziness of the system $(m = 2)$ and D is the distance between the cluster center and pixel.

I.5.3.1 FCM Steps:

The FCM Algorithm is given below.

- **Randomly initialize** the *cluster* centers C_i.

- **The distance *D*** between the cluster center and pixel is calculated by using Eq. (4).

$$D^2(X_j,C_i)=\left\|X_j-C_j\right\|^2 \quad \text{(I.4)}$$

- **The membership** values are calculated by using Eq. (5).

$$U_y = \frac{\left(D(X_j,C_i)\right)^{-1/(m-1)}}{\sum_{k=1}^{c}\left(D(X_j,C_k)\right)^{-1/(m-1)}} \tag{I.5}$$

- **Update the cluster centers.**

$$C_j = \frac{\sum_{j=1}^{n} U_y^m X_j}{\sum_{j=1}^{n} U_y^m} \tag{I.6}$$

- **The iterative process starts:**
  - Update the U_{ij} by using Eq (5).
  - Update the C_i by using Eq (6).
  - Update the D using Eq (4).
  - If $|C_{new} - C_{old}| >$; ($\varepsilon = 0.001$) then go to step 1.
  - Else stop.

This process will assign every pixel to a precise cluster for which the membership value is maximal.

I.5.3.2 FCM Strengths:

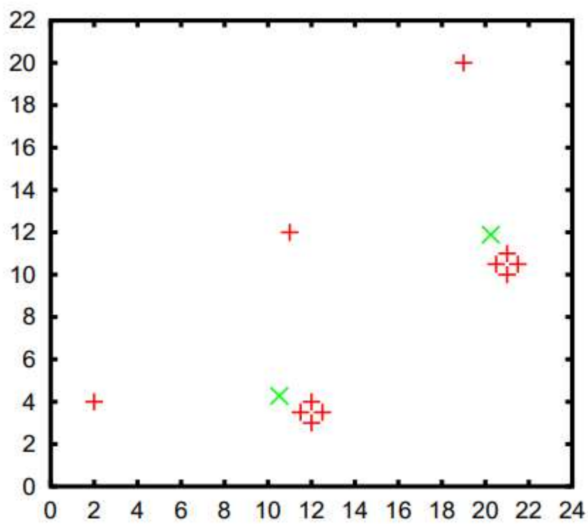
The FCM algorithm has proven to be a very popular method of clustering for many reasons.

- In terms of programming implementation, it is relatively straightforward. It employs an objective function that is intuitive and easy-to-grasp . For data sets composed of hyper spherically-shaped well-separated clusters, FCM discovers these clusters accurately.
- Furthermore, because of its fuzzy basis, it performs robustly: it always converges to a solution, and it provides consistent membership values.

I.5.3.3 FCM Weaknesses:

The shortcomings of FCM, as we have assessed them independently, are:

- It requires the number of clusters to look for to be known a priori.
- It requires initialization for the prototypes, good initialization positions are difficult to assess. If the iterative algorithm commonly employed for finding solutions of the FCM objective function is used, it may find more than one solution, depending on the initialization. This relates to the general problem of local and global optimization.
- Not robust enough with light level noise.



Data		Memberships	
x	y	Cluster 1	Cluster 2
12.0	3.0	0.975	0.025
12.0	4.0	0.983	0.017
11.5	3.5	0.989	0.011
12.5	3.5	0.967	0.033
21.0	10.0	0.028	0.972
21.0	11.0	0.009	0.991
20.5	10.5	0.014	0.986
21.5	10.5	0.021	0.979
2.0	4.0	0.845	0.155
19.0	20.0	0.174	0.826
11.0	12.0	0.588	0.412

Figure I.7: *A Data Set Containing Noise Points.*

I.6. Conclusion:

The FCM algorithm is widely used in image processing, particularly medical ones. Its reduced complexity, Strong elements that have garnered increasing attention include its reduced complexity and ease of implementation, especially for huge data sets, as well as its fuzzy component (integration of the degree of membership).

However, this algorithm does not take into account the spatial information of pixels and considers only the feature vector used. As a result, the neighborhood's characteristics are ignored, making the pixel under consideration more susceptible to noise.

Several modifications are made to FCM in order to improve the quality of clustering, but this remains insufficient.

Chapter II:
Kernel Methods

II.1. Introduction

Kernel methods are a class of algorithms used for solving both classification and regression problems in machine learning.

Unlike traditional linear methods, kernel methods operate in a high-dimensional feature space by implicitly mapping the input data into that space. The key idea behind kernel methods is to find a suitable kernel function that measures the similarity between data points in the transformed feature space.

II.2. Transformation into High-Dimensional Feature Space

A kernel function is used for classification in general. Suppose a set of objects x1,...,xN ∈ Rp should be analyzed, but structure in the original space, Rp is somehow inadequate for the analysis. The idea of the method of kernel functions uses a 'nonlinear' transformation [5].

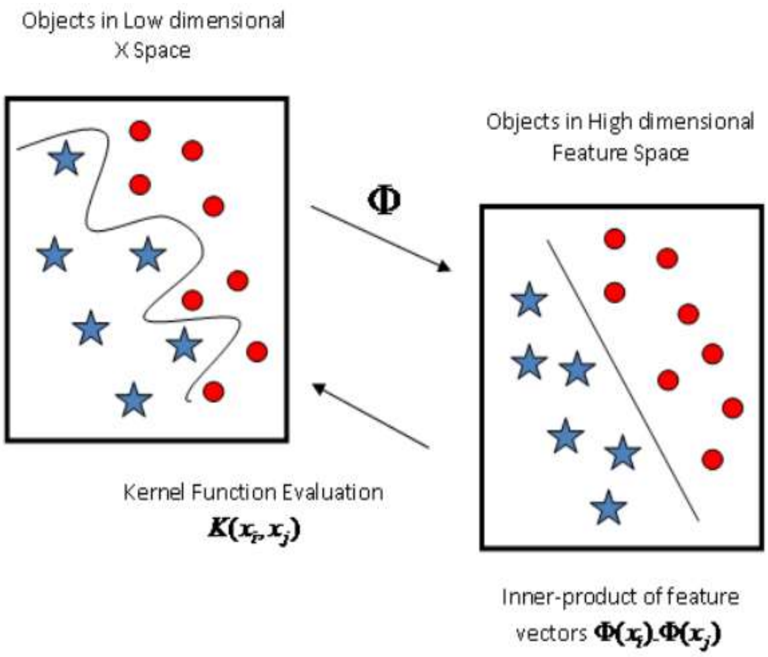


Figure II.1: *An Illustration of Kernel Method*

Φ: Rp → H, where H is a high-dimensional Euclidean space, which is sometimes an infinite dimensional Hilbert space. We can write:

$$\Phi(x)=(\varphi 1(x),\ \varphi 2(x),...). \tag{II.1}$$

It is unnecessary to have a functional form of $\Phi(x)$. Instead, we assume that the scalar product of H is known and is given by a known kernel function K(x, y):

$$K(x, y) = \Phi(x),\ \Phi(y), \tag{II.2}$$

Using the idea of kernel functions, we are capable of analyzing objects in space H instead of the original space Rp; Rp is called data space while H is called a (high-dimensional) feature space. Several types of analysis including regression, principal component, classification, and clustering have been done using the feature space with kernels. In applications, two types of kernel functions are most frequently used [5]:

The first type is:

$$K\ (x,y) = (\{x,y\}+c)^d \tag{II.3}$$

The second type is:

$$K\ (x,y)=exp\ (-\lambda||x-y|^{|\,2} \tag{II.4}$$

Where : x is an input vector in R^p; y is an input vector in R^p; c is a constant; d is the degree of the polynomial; λ is a parameter that controls the width of the kernel.

The former is called a polynomial kernel, and the latter is called a Gaussian kernel.

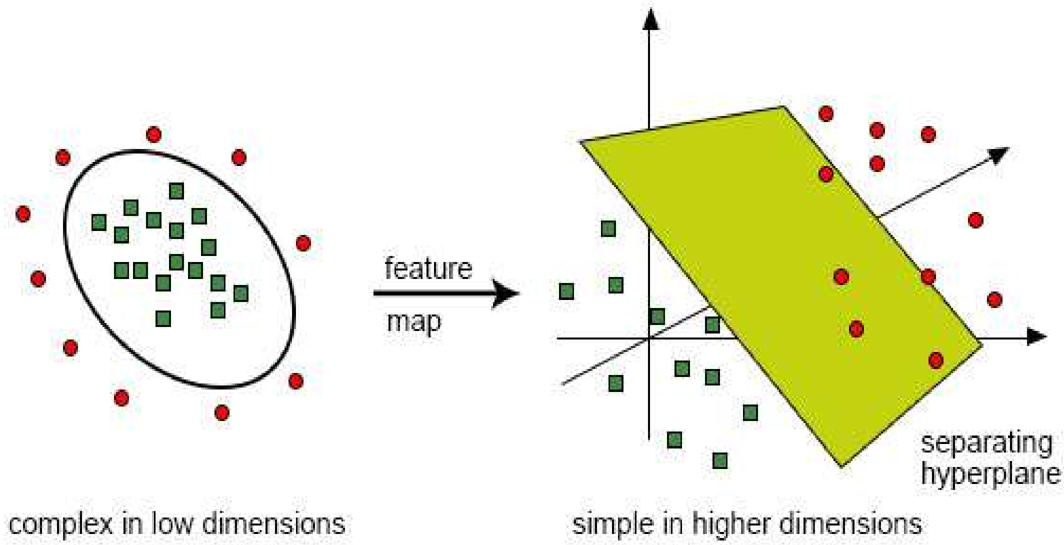


Figure II.2: *Transformation Into High-Dimensional Feature Space*

II.3. Kernel Types

II.3.1. The Gaussian kernel:

The Gaussian kernels are widely used and have been extensively studied in neighboring fields. The proposition in the following kernel is indeed valid.

The Gaussian kernel is defined by:

$$K(x,z) = exp\left(-\frac{||x-z||^2}{2\delta^2}\right) \quad \text{(II.5)}$$

Where : x is an input vector in R^p; z is an input vector in R^p; σ>0 is the bandwidth parameter.

Note that we are not restricted to using the Euclidean distance in the input space. If for example κ1 (x, z) is a kernel corresponding to a feature mapping φ1 into a feature space F1, we can create a Gaussian kernel in F1 by observing that:

$$||\emptyset_1(x) - \emptyset_1(z)||^2 = k_1(x,z) - 2k_1(x,z) + k_1(z,z) \quad \text{(II.6)}$$

giving the derived Gaussian kernel as:

$$k_1(x,z) = exp\left(-\frac{k1(x,x) - 2k1(x,z) + k1(z,z)}{2\delta^2}\right) \quad \text{(II.7)}$$

The parameter σ controls the flexibility of the kernel in a similar way to the degree d in the polynomial kernel. Small values of σ correspond to large values of d since, for example, they allow classifiers to fit any labels, hence risking over-fitting.

In such cases, the kernel matrix becomes close to the identity matrix. Progressively diminish the kernel function and prevent the learning of non trivial classifiers.

The feature space has infinite- dimensions for every value of σ but for large values, the weight quickly decline on the higher-order features.

II.3.2. The Laplacian kernel:

The Laplacian kernel. matrix can also be used in clustering as it frequently possesses more balanced properties than the kernel matrix. It is defined as follows.

The Laplacian matrix L(K) of a kernel matrix K is defined by:

$$L(K) = D - K \quad \text{(II.8)}$$

where D is the diagonal matrix with entries:

$$D_{ii} = \sum_{j=1}^{l} k_{ji} \tag{II.9}$$

Observe the following simple property of the Laplacian matrix. Given any real vector

$$v = (v1,...,v) \in R \tag{II.10}$$

$$\sum_{i,j=1}^{l} k_{ij} (v_i - v_j)^2 = 2\sum_{i,j=1}^{l} k_{ij}\, v_i^2 - 2\sum_{i,j=1}^{l} k_{ij}\, v_i v_j \tag{II.11}$$

$$= 2v`Dv - 2v`Kv = 2v`L(K)v\ .$$

II.3.3. Polynomial kernels:

The derived polynomial kernel for a kernel κ1 is defined as :

$$\kappa\ (x, z) = p\ (\kappa 1\ (x, z))\ , \tag{II.12}$$

where p (·) is any polynomial with positive coefficients. Frequently, it also refers to the special case:

$$k_d(x,z) = (\{x,z\} + R)^{\ d}\ , \tag{II.13}$$

defined over a vector space X of dimension n, where R and d are parameters.

The dimension of the feature space for the polynomial kernel :

$$k_d(x,z) = (\{x,z\} + R)^{\ d}\ \ is\ .^{n+d}_{d} \tag{II.14}$$

II.3.4. ANOVA kernel:

The embedding of the ANOVA kernel of degree d is given by:

$$\emptyset_d: x \rightarrow (\emptyset_A(x))_{|A|=d} \quad \text{(II.15)}$$

where for each subset A the feature is given by:

$$\emptyset_A(x) = \prod_{i \in A} x_i = x^{iA} \quad \text{(II.16)}$$

where iA is the indicator function of the set A:

The dimension of the resulting embedding is clearly $\binom{n}{d}$, since this is the number of such subsets, while the resulting inner product is given by:

$$K_d(x,z) = \emptyset_d(x), \emptyset_d(z)\} = \sum_{|A|=d} \emptyset_A(x) \emptyset_A(z)$$

$$= \sum_{1 \le i_1 \le i_2 \le \dots i_d \le n} (x_{i1} z_{i1})(x_{i2} z_{i2}) \dots (x_{id} z_{id}) \quad \text{(II.17)}$$

$$= \sum_{1 \le i_1 \le i_2 \le \dots i_d \le n} \prod_{j=1}^{d} x_{ij} z_{ij} ,$$

II.3.5. Intersection kernel:

Define the intersection kernel over the subsets of D by:

$$\kappa \cap (A1, A2) = \mu (A1 \cap A2), \quad \text{(II.18)}$$

that is, the measure of the intersection between the two sets [6].

This can be seen to be a valid kernel by considering the feature space of all measurable functions with the inner product defined by:

$$\{f_1 , f_2\}= \int f_1 (a) f_2 (a) d\mu(a) \quad (II.19)$$

The feature mapping is now given by:

$$\varphi : A \longrightarrow IA, \quad (II.20)$$

implying that:

$$\begin{aligned} \mathcal{R} \cap (A_1, A_2) &= \mathcal{U}(A_1 \cap A_2) = \int_D I_{A_1 \cap A_2}(a) du(a) \\ &= \int_D I_{A_1 \cap A_2}(a) du(a) = \langle I_{A_1}, I_{A_2} \rangle \\ &= \langle \emptyset(A_1), \emptyset(A_2) \rangle. \end{aligned} \quad (II.21)$$

II.3.6. Union complement kernel:

Union complement kernel Assuming that μ (D) = 1, we have the union complement kernel [7]:

$$\breve{\kappa}(A_1,A_2) = \mu ((D \setminus A_1) \cap (D \setminus A_2))= 1-\mu(A_1 \cup A_2). \quad (II.22)$$

There are many cases in which we may want to represent data items as sets. There is also the potential for an interesting duality, since we represent an object by the items it contains, while the items themselves can also be described by the set of objects they are contained in.

II.4. The kernelization of Clustering

In real world situations, methods of clustering to deal with datasets having non-linear classification boundaries are more appropriate. It has been shown that the conventional clustering methods cannot separate two overlapping clusters [8].

For this reason, kernel methods to classify non-linear datasets have been studied by many researchers in recent years. Several clustering methods have been modified to incorporate kernels. The use of kernel allows to map implicitly data into a high dimensional space called feature space and computing a linear partition in a non-linear partition space.

Kernel-based clustering techniques based on k-means have been developed in Camastra and Verri (2005), Dhillon et al. (2004) and Kim et al. (2005). The authors of Kim et al. (2005) performed a comparative analysis between k-means and FCM clustering algorithms and their kernelized versions. Kernel-based fuzzy clustering techniques based on FCM have been developed in Wu and Xie (2003), Zhang and Chen (2002), Kim et al. (2005), Chiang and Hao (2003) and Zhang and Chen (2003).

II.5. The Kernel Fuzzy C-Means Clustering (KFCM)

The KFCM algorithm adds kernel information to the traditional fuzzy c-means algorithm, and it overcomes the disadvantage that the FCM algorithm can't handle the small differences between clusters.

There are two major forms of kernel-based fuzzy clustering [8]:

- The first one comes with prototypes constructed in the feature space. These clustering methods will be referred to as KFCM-F (with F standing for the feature space).
- In the second category, abbreviated as KFCM-K, the prototypes are retained in the kernel space, and thus the prototypes must be approximated in the feature space by computing an inverse mapping from the kernel space to the feature space.

The advantage of the KFCM-F clustering algorithm is that the prototypes reside in the feature space and are implicitly mapped to the kernel space through the use of the kernel function.

II.6. Kernel Fuzzy C-Means Algorithm

kernel version of the FCM algorithm and its objective function are given below:

$$O_m(U,C)=\sum_{i=1}^{c}\sum_{j=1}^{n}U_{ij}^{m}(1-K(x_j,C_i)) \quad \text{(II.23)}$$

Thus, the revised equations for the essential conditions for minimizing $O_m\,(U,C)$ are given below:

$$C_i=\frac{\sum_{j=1}^{n}U_y^{m}K(X_j,C_i)X_j}{\sum_{j=1}^{n}U_y^{m}K(X_j,C_i)};\, i=1,2\ldots C \quad \text{(II.24)}$$

$$U_y=\frac{\left(1-K(X_j,C_i)\right)^{-1/(m-1)}}{\sum_{k=1}^{c}\left(1-K(x_j,C_i)\right)^{-1/(m-1)}};\ \begin{matrix} i=1,2\ldots c \\ j=1,2\ldots n\end{matrix} \quad \text{(II.25)}$$

We identify the essential conditions for minimizing $O_m\,(U,C)$ are revising the last two equations. only when the kernel function K is selected to be the Gaussian function with :

$$K(x_j,\, C_i)=exp\,(-\,||\,x_j\text{-}C_i\,||^2 \setminus \sigma^2) \quad \text{(II.26)}$$

Different kernels can be selected for replacing the Euclidean distance for different conditions.

II.7. Multi-Kernel Based Fuzzy C-Means (MKFCM)

The multiple-kernel methods provide us a great tool to fuse information from different sources. A multi-kernel is a combination of two or more kernel functions.

Instead of relying on a single kernel to measure the similarity between data points, multiple kernels are combined, often through a weighted sum, to create a composite kernel. The composite kernel aims to leverage the strengths of each individual kernel.

II.7.1. Mathematical Formulation:

Given two kernels, K1 and K2, a simple multi-kernel can be formed as [9]:

$$K_{multi}(x_i,x_j) = \propto K_1(x_i, x_j) + (1-\propto) K_2(x_i,x_j) \backslash \quad \text{(II.27)}$$

where α is a weighting parameter that balances the contribution of each kernel.

II.7.2. MKFCM Algorithm

The Multi-Kernel Fuzzy C-Means (MKFCM) algorithm extends the standard Fuzzy C-Means (FCM) clustering algorithm by incorporating multiple kernels. Here's a step-by-step description of the MKFCM algorithm [8]:

a) <u>Initialize Parameters:</u>

- Number of clusters c.
- Fuzziness parameter m.
- Maximum number of iterations.
- Convergence threshold.
- Kernel weighting parameters $\propto_1, \propto_2, \ldots, \propto_k$ (if combining k kernels).

b) <u>Compute Kernel Matrices:</u>

- For each kernel Ki compute the kernel matrix $K_i(x_p,x_q)$ for all data points x_p and x_q.

c) Combine Kernel Matrices:

– Form the composite kernel matrix K_{multi} using the weighted sum:

$$K_{multi}\ (x_p,x_q) = \sum_{i=1}^{k} \propto_i K_i\ (x_p,x_q) \tag{II.28}$$

– Ensure that the weights αi sum to 1.

d) Initialize Membership Matrix:

– Initialize the membership matrix U randomly such that the sum of memberships for each data point is 1.

e) Iterative Optimization:

Repeat until convergence or maximum iterations reached:

– Compute Cluster V using the current membership matrix U:

$$\text{Vj}=\frac{\sum_{p=1}^{N}.(u_{jp}^{m} k_{multi}(x_p))}{\sum_{p=1}^{N} u_{jp}^{m}} \tag{II.29}$$

– Where u_{jp} is the membership of data point x_p in cluster j, and $K_{multi}\ (x_p,.)$ is the $p\text{-}th$ row of the composite kernel matrix.

– Update the membership values u_{jp}:

$$u_{jp}=\frac{1}{\sum_{k=1}^{c}\left(\frac{k_{multi}\,(xp,xq)-2k_{multi}\,(xp,vj)+\,k_{multi}\,(vj,vj)}{k_{multi}\,(xp,xq)-2k_{multi}\,(xp,vj)+\,k_{multi}\,(vk,vk)}\right)^{\frac{1}{m-1}}} \tag{II.30}$$

f) Convergence Check:

- Check for convergence by evaluating the change in the membership matrix U or the cluster centers V If the change is below a predefined threshold, stop the iterations.

I.8. Conclusion:

In this chapter, We have seen a variant of the fuzz C-meaus technique called Kernel FCM, uses-a Kernel function to change the original data space into higher dimentional space where the clusters are more effectively divided.

We will now use MKFCM, KFCM and FCM on MRI pictures, and the results will be shown in the chapter that follows.

Chapter III: Simulation Results and Discussions

III.1. Introduction

The finding for the various algorithms evolved in the earlier chapter are shown in the section that follows.

We used T1 weighted MRI pictures (181x217x181), and we extracted the slice n=79 for illustration.

III.2. Evaluation Measures:

I will provide an overview of most used evaluation metrics, in medical image segmentation.

The evaluation metrics are:

- Dice coefficient
- Intersection over Union (IoU)

The presented metrics are based on the computation of a confusion matrix for a binary segmentation mask, which contains the number of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) predictions. The value ranges of all presented metrics span from zero (worst) to one (best).

Prediction \ Ground truth	FTU (1)	Background (0)
FTU (1)	TP	FP
Background (0)	FN	TN

Figure III.1: *Evaluation Metrics*

- **TP (True Positive):** represents the number of FTU pixels that have been properly classified as FTU
- **FP (False Positive):** represents the number background pixels being misclassified as FTU
- **FN (False Negative):** represents the number of FTU pixels being misclassified as background
- **TN (True Negative):** represents the number of background pixels that have been properly classified as background

FTU (1) stands for "Foreground Tumor Unit" or "Feature of Tumor Unit." represents the pixels that belong to the tumor (foreground), and Background (0) represents the pixels that do not belong to the tumor (background). This kind of matrix is used to evaluate the performance of a model in distinguishing between tumor and non-tumor regions in medical images.

III.2.1. Dice Coefficient (F-Score):

Dice Coefficient also called F-score: one of the most widespread scores for performance measuring in computer vision and in Medical Image Segmentation.

Dice coefficient is calculated from the precision and recall of a prediction. Then, it scores the overlap between predicted segmentation and ground truth. It also penalizes false positives, which is a common factor in highly class imbalanced datasets like Medical Image Segmentation.

Based on the Dice Coefficient, there are two popular utilized metrics in Medical Image Segmentation:

a) The Intersection-over-Union (IOU), also known as Jaccard index or Jaccard similarity coefficient is the area of the intersection over union of the predicted segmentation and the ground truth:

$$IOU=\frac{TP}{TP+FP+FN} \tag{III.1}$$

We can represent this index in the below figure:

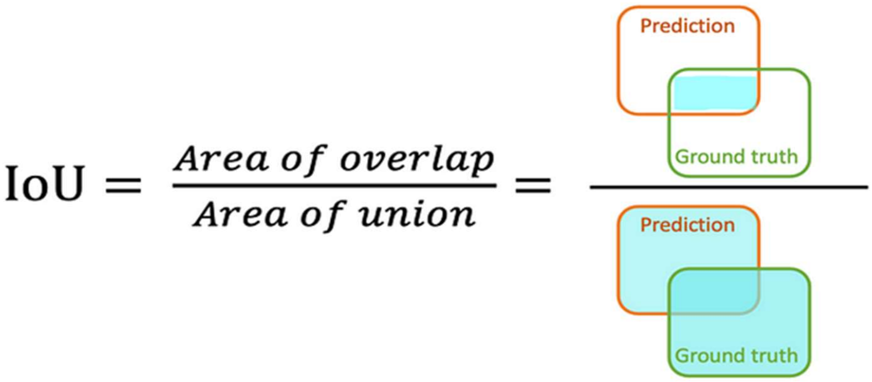


Figure III.2: *IOU Index Illustrated*

b) The Dice similarity coefficient (DSC), also known as F1-score or Sørensen-Dice index: most used metric in the large majority of scientific publications for Medical Image Segmentation evaluation.

The difference between the two metrics is that the IOU penalizes under- and over-segmentation more than DSC.

A harmonic mean of precision and recall. In other words, it is calculated by 2*intersection divided by the total number of pixels in both images.

$$Dice = \frac{2TP}{2TP+FP+FN} \qquad \text{(III.2)}$$

We can represent this coefficient in the following figure:

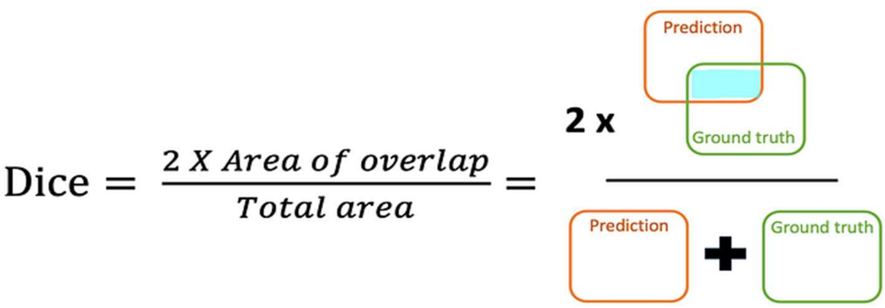


Figure III.3*: Dice Coefficient Illustrated*

III.3. Magnetic Resonance Imaging (MRI):

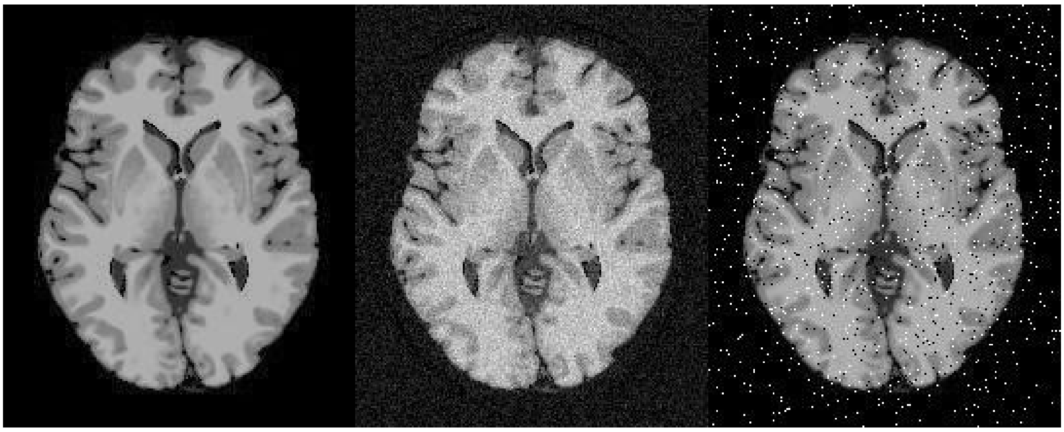
MRI, being a "non-invasive" procedure, has become one of the most popular diagnostic tools in neuroradiology. This technique is widely used in monitoring various conditions, such as Alzheimer's, Parkinson's, and others.

MRI is a medical imaging technique based on the principle of nuclear magnetic resonance (NMR). It provides a 2D or 3D description of a part of the body, especially the brain. In practice, MRI images are provided as a series of 2D slices that, when arranged in the correct geometry, form a three-dimensional image.

III.4. Brain web Database:

The literature recommends using Brainweb a database provided by a powerful simulator made available online by the McConnell Brain Imaging Centre at the Montreal Neurological Institute, McGill University. This institute aims to better understand neurological diseases through various imaging methods [10].

The simulator allows for the creation of images that are very close to reality. Additionally, it lets us choose different parameters such as noise levels, inhomogeneity rates, T1 or T2 modality, resolution, it also offers reference images along with their manual segmentations. These data enable us to evaluate the performance of different algorithms.



(a) (b) (c)

Figure III.4: *MRI Images (a):original, (b):gaussian noise, (c):salt and pepper noise*

III.5. Initialization Step:

The experiments and performance evaluation were performed on medical images, including image of the MR image of brain (Brainweb Database).

- The fuzzy c-means (FCM) clustering, the kernel based fuzzy c-means (KFCM), and multi kernel fuzzy c-means (MKFCM) level set method were implemented with Matlab R2018a (MathWorks) in a Windows 10 System Pro. All the experiments were run on a computer with Intel i3 and 4GB RAM.
- For the FCM, KFCM, and MKFCM algorithms, the initialization of class centers is random.
- The attribute vector consists of the different gray levels constituting the image. The number of clusters is 3, the number of iterations for FCM is 100, and the fuzziness parameter is usually set to 2.
- The same code is used for KFCM, adding a Gaussian kernel with a bandwidth parameter of 10.
- Add to the previous code, KFCM another kernel, the Laplacian kernel, with a bandwidth parameter of 5.

III.6. The results:

We compared clustering results using three different fuzzy clustering methods: Fuzzy C-Means (FCM), Kernel Fuzzy C-Means (KFCM), and Multi-Kernel Fuzzy C-Means (MKFCM). Using various images with different noise levels, the results are presented in the following figures:

a) **Real Brain MRI Image with No Noise:**

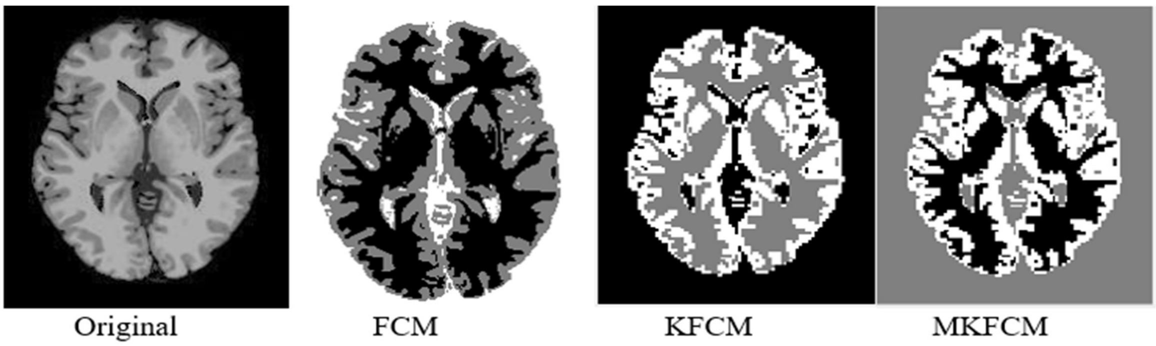


Figure III.5: *Real Brain MRI Image No Noise Results*

The Original Image is real brain MRI image with distinct anatomical structures. The FCM segmentation shows that it identifies major structures but lacks detail and has some misclassifications, while the KFCM Segmentation improves the segmentation detail, capturing more structures. At the end, MKFCM Segmentation offers the best detail and accuracy, closely matching the anatomical features of the brain.

b) Brain MRI with Salt and Pepper Noise:

Now, salt and pepper noise corrupts the original image.

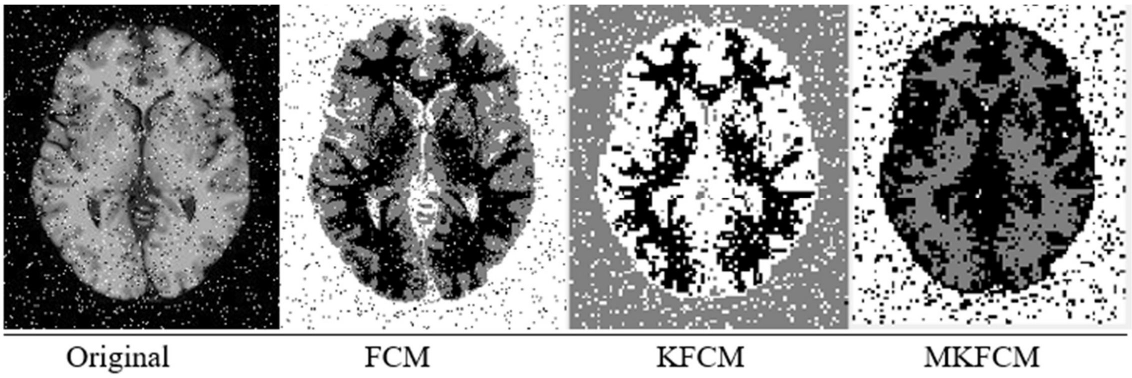


Figure III.6: *Brain MRI with Salt and Pepper Noise Results*

We notice that the noise has a significant impact on the FCM algorithm, resulting a poor segmentation, while the KFCM segmentation shows resilience to noise but still has some artifacts. At the end, MKFCM segmentation effectively handles the noise, providing clean and accurate segmentation.

c) Brain MRI Image With Gaussian Noise:

Here the original image is effected by gaussian noise.

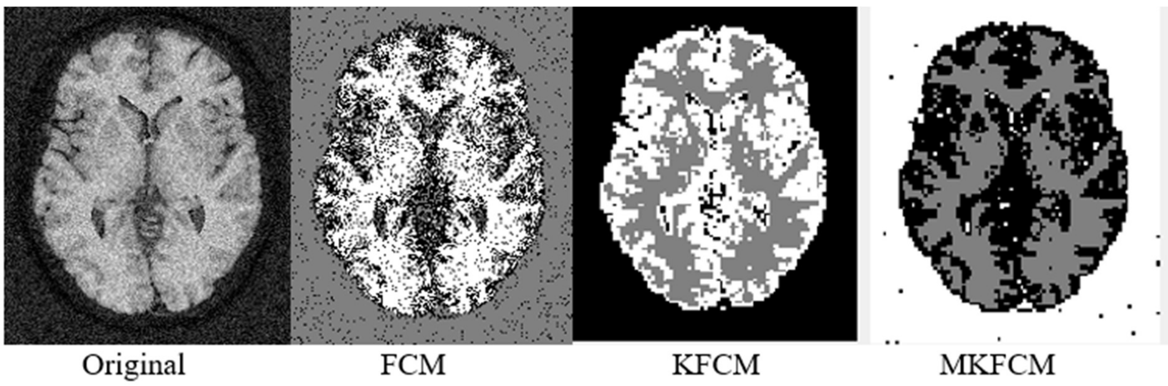


Figure III.7: *Brain MRI Image with Gaussian Noise*

At the first look the FCM Segmentation produces a rough segmentation, failing to distinguish regions accurately. While the KFCM Segmentation improves segmentation accuracy, but still has issues with boundaries. At the end, MKFCM Segmentation demonstrates superior performance, offering the most accurate segmentation with well-defined boundaries.

III.7. Evaluation of The Algorithms:

We compared the segmentation efficiency using the two different quantitive metrics that we explained earlier in this chapter, the Dice Coefficient and Intersection over Union (IoU). Using various images with different noise levels, the results are presented in the following table:

Table III.1: Results for the Three Algorithms

Algorithm	**FCM**			**KFCM**			**MKFCM**		
Gaussian Noise	3%	5%	7%	3%	5%	7%	3%	5%	7%
IOU (%)	0.89	0.87	0.83	0.93	0.90	0.86	**0.97**	**0.96**	**0.95**
Dice (%)	0.94	0.92	0.91	0.96	0.95	0.92	**0.97**	**0.96**	**0.95**

When comparing the algorithms, we find that MKFCM does reasonably well in noisy environments. At noise levels of 3%, 5%, and 7%, respectively, it achieves a IoU of 0.97, 0.96, and 0.95. While KFCM obtaining a IoU of 0.93, 0.90, and 0.86 respectively and FCM is reaching 0.89, 0.87, and 0.83 respectively for the corresponding noise levels. This demonstrates that FCM perform just slightly worse. The findings are validating by the Dice coefficient calculation.

According to the results, we could say that MKFCM outperforms KFCM and FCM in terms of robustness to noise, with higher values obtained for the majority of evaluation measures.

This could be a result of the way these algorithms evaluate similarity. MKFCM and KFCM measure similarity using a kernel function, whereas FCM is based on the minimization of Euclidean distances

The kernel function included in Kernel versions of FCM enables non-linear transformations of the input data. In situations where linear separation is insufficient, this allows these algorithms to capture non-linear interactions among data points, resulting more accurate clustering. Contrarily, FCM makes the assumption that data points and clusters have linear correlations.

However, the Kernel function allow to MKFCM and KFCM to handle different sized and shaped clusters. Unlike FCM which is limited with the assumption of spherical clusters.

III.8. Conclusion

The results indicate that Multi-Kernel Fuzzy C-Means (MKFCM) consistently outperforms both Fuzzy C-Means (FCM) and Kernel Fuzzy C-Means (KFCM) across various types of images and noise conditions.

MKFCM's use of multiple kernels allows it to handle complex data distributions and noise more effectively, leading to more accurate and robust segmentation. This makes MKFCM a superior choice for image segmentation tasks, especially in medical imaging, where precision is essential.

General Conclusion

General conclusion

Our main goal has been to support physicians in their clinical practice by applying the Kernel versions of Fuzzy C-Means approach to medical image segmentation.

Since they will make it easier to precisely identify and analyze lesions, anatomical structures, and other important regions of interest. By providing correct segmentation of medical images, we contribute in improving treatment planning and diagnosis.

The findings in the latter chapter show that, for a variety of medical image and noise levels, the Multi-Kernel Fuzzy C-Means (MKFCM) method consistently performs better than the Fuzzy C-Means (FCM) and Kernel Fuzzy C-Means (KFCM) algorithms.

Because MKFCM uses several kernels, it can handle complex data distributions and noise more effectively and improves the accuracy and robustness of image segmentation. Because of this enhanced capabilities, MKFCM is a better option for activities involving medical imaging, where accuracy is crucial.

Through the application of quantitative analysis, MKFCM routinely achieves the highest performance metrics, as demonstrated by the Dice Coefficient and Intersection over Union (IoU) scores. This suggests significantly better segmentation performance by integrating two kernels

In summary, we demonstrate the benefits of the MKFCM algorithm for medical image segmentation, opening the door for more accurate and dependable diagnostic instruments in the field of medical imaging.

Future study can concentrate on investigating the integration of sophisticated approaches, such as deep learning. The system can extract complex patterns and characteristics from the data directly through deep learning, which produces more accurate segmentation.

References

References

1. Smith, J., Supervised Clustering in Medical Imaging, 2018.
2. F. Gibou and R. Fedkiw, “A fast hybrid k-means level set algorithm for segmentation,” in Proceedings of the 4th Annual Hawaii International Conference on Statistics and Mathematics.
3. Chi et al., 1996; Chen & Wang, 1999.
4. Rosenfeld, 1979; Philip et al., 1994.
5. Pal, Nikhil R, & Pal, Sankar K. A Review on Image Segmentation Techniques. Pattern Recognition, 1993.
6. Bajcsy, P, & Ahuja, N, Location and density-based hierarchical clustering using similarity analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998.
7. John Shawe-Taylor and Nello Cristianini “Kernel Methods for Pattern Analysis”.
8. Krishnapuram, Raghu, & Keller, James M. Quantitative Analysis of Properties and Spatial Relations of Fuzzy Image Regions. IEEE Transactions on Fuzzy Systems, 1993.
9. Camastra, F. and Verri, A. ‘A novel kernel method for clustering’, IEEE Transaction on Pattern Analysis and Machine Intelligence. 2005.
10. http://mouldy.bic.mni.mcgill.ca/brainweb/,
11. AlSultan, K S, & Selim, S Z. A Global Algorithm for Fuzzy Clustering. 1993.
12. Brown, L., Advances in CNNs for Medical Image Analysis, 2019.
13. White, K., Unsupervised Learning: K-means and Fuzzy C-means, 2017.
14. Green, M., Applications of Unsupervised Clustering in Medical Imaging, 2020.
15. Johnson, R., Hierarchical and Density-Based Clustering Techniques, 2021.
16. Backer, E. Computer-Assisted Reasoning in Cluster Analysis. Prentice Hall. 1995.
17. Barni, M, Cappellini, V, & Mecocci, A. A possibilistic approach to clustering - Comments. IEEE Transactions on Fuzzy Systems. 1996.
18. Bezdek, James. Pattern Recognition with Fuzzy Objective Function Algorithms. 1981.
19. Chen, Mu-Song, & Wang, Shinn-Wen. Fuzzy Clustering Analysis for Optimizing Fuzzy Membership. 1999.

20. Chun-yan Yu, Ying Li, Ai-lian Liu, Jing-hong Liu, “A Novel Modified Kernel Fuzzy C-Means Clustering Algorithm on Image Segmentation”, IEEE International Conference on Computational Science and Engineering.
21. Jain, Anil K., "Data clustering: 50 years beyond K-means." Pattern recognition letters
22. Krishnapuram, Raghu, & Keller, James M. Quantitative Analysis of Properties and Spatial Relations of Fuzzy Image Regions. IEEE Transactions on Fuzzy Systems, 1993.
23. L. Szilagyi, Z. Benyo, S. Szilagyii, and H. Adam, “MR brain image segmentation using an enhanced fuzzy C-means algorithm,” in Proceedings of the 25" Annual International Conference of the IEEE EMBS
24. Lee Song Yeow, “Spatial Kernel-Based Generalized C mean Clustering for Medical Image Segmentation”, School of Computer Sciences, University Sains Malaysia
25. M. Rastgarpour and J. Shanbehzadeh, “Automatic medical image segmentation by integrating KFCM clustering and level set based FTC model,” in IAENG Transactions on Electrical Engineering, Special Issue of the International Multi Conference of Engineers and Computer Scientists World Scientific
26. Miin-Shen Yang, Hsu-Shen Tsai, “A Gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction,” Pattern Recognition Letters
27. S. Chen and D. Zhang, “Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure,” IEEE Transactions on Systems, Man, and Cybernetics,
28. Sadaaki Miyamoto, Katsuhiro Honda, Hidetomo Ichihashi. Algorithms for Fuzzy Clustering Methods in c-Means Clustering with Applications
29. T. Saikumar, B. Shashidhar, V. Harshavardhan, and K. S. Rani, “MRI brain image segmentation algorithm using watershed transform and kernel fuzzy C-means clustering on level set method,” International Journal on Computer Science and Engineering
30. Wu Z, Xie,W.X Yu J.P. “Fuzzy C-means Clustering Algorithm Based on Kernel Method” In: Proceedings of Fifth International Conference on Computational Intelligence and Multimedia Applications
31. X. Munoz, J. Freixenet, X. Cufi, and J. Marti, “Strategies for Image Segmentation Combining Region and Boundary Information,” PatternRecognition Letters