Comparative Study of Distance Measures for the Fuzzy C-means and K-means Non-Supervised Methods Applied to Image Segmentation

Martín Vélez-Falconí[a,b], Josué Marín[a,b], Selena Jiménez[a] and Lorena Guachi-Guachi[a,b,c]

[a] *Yachay Tech University, Hacienda San José, Urcuquí 100119, Ecuador*
[b] *SDAS Research Group http://www.sdas-group.com*
[c] *Department of Mechatronics, Universidad Internacional del Ecuador, Av. Simon Bolivar 170411, Ecuador*

Abstract

Recent studies have revealed that the performance of the FCM and K-means is completely related to the distance measures. However, the literature does not provide evidence that the distance used for data-clustering is useful for image segmentation. Therefore, a comparative study of the performance of different distance measures applied to image segmentation, using the mentioned clustering methods is proposed in this work. The selection of the distance measures was based on a literature study of their benefits. As a consequence, the selected distances to be tested are Euclidean, Manhattan, Canberra, and Spearman. Since our principal goal is to compare the effectiveness of the distance, the experiment had been evaluated according to two centroids selected by the user. According to primary results, the best-rated distance employed for image segmentation is the Canberra distance.

Keywords

Clustering, Image Segmentation, Non-Supervised Algorithms

1. Introduction

An essential objective in the computational analysis of images is the segmentation [1]. As consequence, many segmentation algorithms have been developed by using certain mathematical and theoretical tools, such as fuzzy logic, genetic algorithm, neural network [1], pattern recognition, wavelet, and so on. This article is focused on two specific methods for image segmentation: K-means and Fuzzy C-means (FCM). K-means even thought to be an older algorithm, presents an advantage as computationally faster for a large number of variables [2]. On the other hand, the FCM is a semi-automatic segmentation algorithm since it combines both manual and automatic segmentation. The manual process is performed by selecting the initial centroids. Each data point resides in the clusters with a degree of membership. The mem-

ICAIW 2020: Workshops at the Third International Conference on Applied Informatics 2020, October 29–31, 2020, Ota, Nigeria
martin.velez@yachaytech.edu.ec (M. Vélez-Falconí); josue.marin@yachaytech.edu.ec (J. Marín); selena.jimenez@yachaytech.edu.ec (S. Jiménez); lguachi@yachaytech.edu.ec (L. Guachi-Guachi)
0000-0003-0946-1885 (M. Vélez-Falconí); 0000-0002-1109-0465 (J. Marín); 0000-0003-2796-788X (S. Jiménez); 0000-0002-8951-8150 (L. Guachi-Guachi)

bership function is suitable for real-world applications where boundaries between clusters are not well-defined [3]. In general, FCM gives better performance than the k-means algorithm in different applications such as [4, 5]. However, in the field of image segmentation, it is not clear which is the best method. Research as [6] explains that both algorithms are helpful for segmentation purposes. On the other hand, experiments as [7] show superior results in FCM.

Some works have explored the distance efficiency of the methods as K-means [8] and FCM [9] in existing data sets. However, there is no current research on the comparison of the distance efficiency between different methods of non-supervised clustering for the segmentation of images. This work tries to fill this gap by introducing a comparative study of three different distances applied to K-means [8] and FCM [9] algorithms.

The rest of the paper is divided as follows: Section 2 describes the most relevant related works. The methodology used in this work is described in Section 3. Experimental results are presented and discussed in Section 4. Finally, Section 5 deals with the concluding remarks.

2. Related Works

In the last years, clustering algorithms have been explored in several fields including but not limited to computer sciences, medicine, economics, social sciences, and earth sciences [10, 11]. Some applications of clustering algorithms in computer sciences domain include image segmentation such as: brain tumor identification [12], mammography image segmentation [13], satellite image retrieval [14], among others.

Since the main goal of this work is to evaluate the influence of distance measures on the effectiveness achieved by K-means [8] and FCM [9] algorithms for image segmentation purposes, we review the literature mainly based on types of distance measures often used for constructing clustering algorithms. For instance, [15] discusses the use of Manhattan and Euclidean distances. It shows that the Manhattan distance outperforms the Euclidean distance in the number of iterations k>7. Other distances such as Manhattan, Euclidean distance, and Cosine distance have been explored and compared in [16], demonstrating that the computed centroid with minimum distance is sensible to the distance method and the type of data.

From a data mining perspective, authors in [17] use K means method to evaluate the overall performance of Euclidean, Chebyshev, Minkowski, and Manhattan, demonstrating that, Euclidean distance outperforms when a cluster center is selected randomly.

The distances measures used in clustering algorithms are summarized in Table 1.

An effective choice of a distance measure to express the distance between data and centroids is one important feature of clustering algorithms. Thus, the distances widely used in clustering algorithms, such as Euclidean, Manhattan, Canberra, and Chebyshev are selected and compared in this work for segmentation tasks.

3. Proposed Methodology

The proposed methodology for evaluating the performance of the Euclidean, Manhattan, Canberra, and Chebyshev distances in clustering algorithms applied to image segmentation is schematized in Figure 1. First, the process reads an RGB image. Second, the RGB input image

Table 1
Overview of distance measures used in clustering algorithms.

Distances	Advantages	Disadvantages
Minkowski [18]	1. Gives the best result when the data set is distinct or well separated from each other. 2. Fast, robust, and easier to understand.	1. This algorithm does not work well for categorical data, it is applicable only when the meaning is defined. 2. Algorithm fails for the non-linear data set.
Euclidean [19, 20]	1.- The distance within two objects is not affected by the addition of new objects to the analysis, which may be outliers. 2.- Even if points may be in opposite directions, they may fall into the same cluster, if the distance of both points from the centroid is the same [20].	1.- The distances can be greatly affected by differences in scale among the dimensions from which the distances are computed. 2.- If one of the input attributes has a relatively large range, then it can overcome the other attributes
Manhattan [21]	1.-This distance method is not squared and is less sensitive to noise. 2.- The hierarchical search architecture enables a high-speed search in a large database.	1.-This distance does not deal in a deterministic way when the dataset is high level 2.- The one disadvantage is that it depends upon the rotation of the coordinate system.
Canberra [22]	1.- This distance takes color vectors in the RGB reference system is considered to actually made the computation.	1.- It is sensitive to a small change when both coordinates are near to zero.
Spearman [23]	1.- In comparison, its easier to calculate than Euclidean distance. 2.- It offers the best cluster separation and compactness.	1.- The disadvantage is that there is a loss of information when data are converted to rank.
Chebyshev [23]	1.- The advantage is that it takes less time to decide the distances between data sets.	1.-The disadvantage is that if the position centers are near, they will not be optimal.
Cosine [23]	1.- It produces a simple measure that can be used to differentiate between similar texts, rank order them by similarity, or use the scores as a dependent variable in a regression model	1.- It is unable to provide information on the magnitude of the differences. 2.- It is also invariant to scaling 3.- Cosine is not invariant to shifts.

is converted to HSV color space. In this step, a graphical user interface is displayed to allow the user to select the color pixels that represent the object of interest (airplanes and animals) and the background. The pixels selected refer to the centroids that will be used in the clustering algorithms. Third, a multidimensional matrix is built to merge the tuple pixels positions with the color model matrices. Fourth, K-means of FCM is applied to the multidimensional matrix using

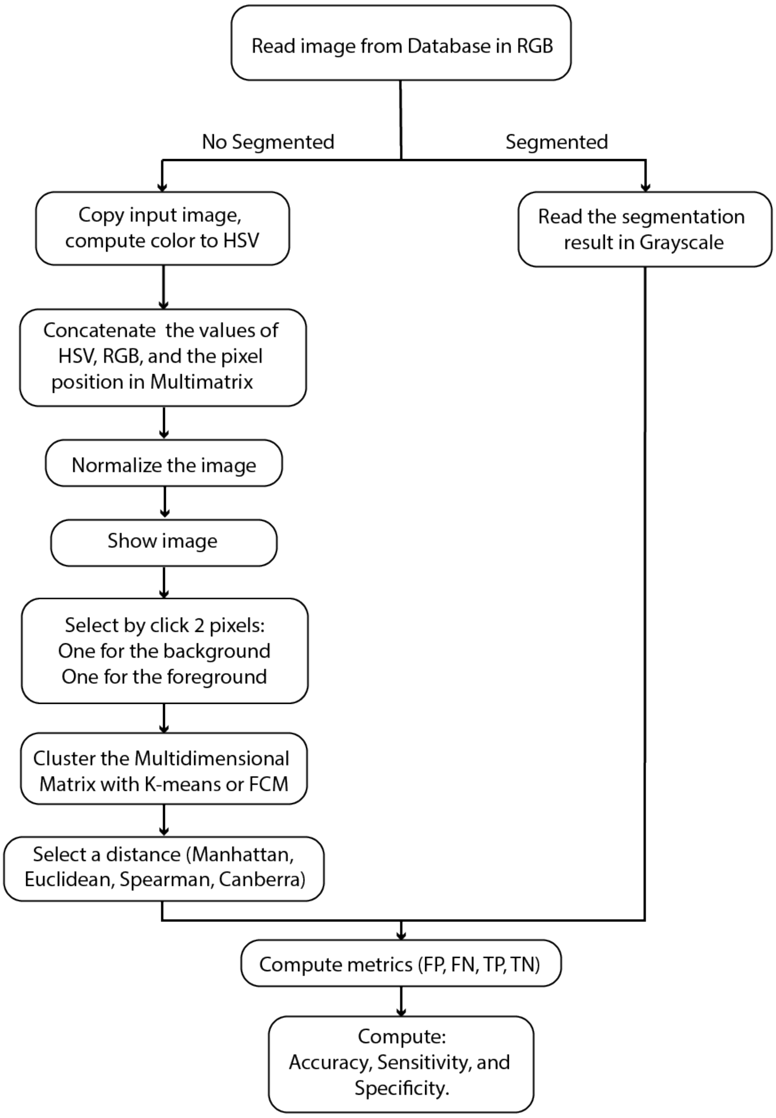


Figure 1: Flowchart of the process for our implementation

the evaluated distance measures. To segment the object of interest in the case of K-means, it classifies each pixel as foreground if the distance from the pixel to the foreground centroid is less than the distance from the pixel to the background centroid. Otherwise, the pixel is classified as a background. It is worth mentioning that the centroids are selected by the user. On The other hand, FCM although it is a similar process it counts with a variable called: the degree of belonging which is linked inversely to the distance from the pixel to the foreground centroid. Finally, the influence of the evaluated distance measures on the overall performance achieved by K-means and FCM methods is determined in terms of accuracy metrics.

3.1. RGB and HSV Color Spaces

Humans perceive colors in screens as a combination of 3 primary colors R (red), G(green), and (blue). The laws of colorimetry state that any color can be derived by the combination of these three primary colors being the combination unique. From RGB other color, representations can be derived such as HSV. Selections of the best color space are one of the challenging tasks in image segmentation [24]. We take RGB into account because it is the most commonly used model for television systems, monitors, cameras, and smartphones. All these devices display color images by modulating the intensity of the three primary colors (red, green, and blue). In spite of RGB being widely used, it is not the most suitable color model for image segmentation purposes.

On the other hand, HSV is one of the color models used in an attempt to use ones closer to how humans perceive color. HSV has three components: hue (H), saturation (S), and value (V). H represents the "color", S is the dominance of that color ranging from unsaturated (shades of gray) to fully saturated (no white component), and V is the brightness. In recent years, HSV color space has gained the ability to support applications in noisy color image segmentation [25]. Also as it is stated in [26], HSV color space can recognize color with high intensities, making it easier to distinguish the objects of interest from the background. Therefore, HSV-based features are used in this work to explored the segmentation clustering algorithms as FCM or K-means.

3.2. Distance Measures

The distance function d is computed for each pixel represented by the vectors x_i and x_j.

- **Euclidean:** The euclidean distance is most commonly used [9], also it derives from Pythagorean Theorem. This is influenced by greater units of measure and this vary with the scale of each variable [8].

$$d(x_i, x_j) = \sqrt{\sum_{k=0}^{n} (x_{ik} - x_{jk})^2}$$

- **Manhattan:** This distance stand out by calculating the absolute differences between coordinates of a pair of objects[9], also it is less noise sensitive [8].

$$d(x_i, x_j) = \sum_{k=0}^{n} |(x_{ik} - x_{jk})|$$

- **Canberra:** This distance examines the sum of series of a fraction differences between coordinates of a pair of objects. This distance is very sensitive to a small change when both coordinates are nearest to zero [27].

$$d(x_i, x_j) = \sum_{k=1}^{n} \frac{|x_i^k - x_j^k|}{|x_i^k| + |x_j^k|}$$

- **Chebyshev:** The Chebyshev distance also is known as the maximum value of distance [9]. It is a defined metric in a vector space when the distance between two points is the maximum of its difference along any of its coordinates dimensions [8].

$$d(x_i, x_j) = max|(x_{ik} - x_{jk})|$$

- **Spearman:** The Spearman Distance is the square of the Euclidean distance within two vectors. Since it is squared, its easier to calculate, hence its computational complexity is reduced [18].

$$d(x_i, x_j) = \sum_{k=1}^{n}(x_{ik} - x_{x_j k})^2$$

3.3. Clustering Algorithms

- **Fuzzy C-Means [9]:** Classical approaches result in strict partitions where each data point can only belong to one cluster. FCM clustering allows data points to belong to more than one cluster [28]. Each cluster is associated with a function that indicates the degree of each data point of belonging to an specific cluster. FCM performs the clustering by iterative searching for a set of fuzzy clusters and their associated centers. On this algorithm the user specify the number of clusters c present in the set of data to be grouped. Then FCM partition the data into c clusters by minimizing the objective function:

$$J_m(U, V) = \sum_{k=1}^{n}\sum_{i=1}^{c}(U_{ik})^m \|X_k - v_i\|^2 , \tag{1}$$
$$1 \leq m \leq \infty$$

  where $\{v_i\}_{i=1}^{c}$ are the centroids of the cluster c and $\|.\|$ is an inner-product norm (this is the parameter that we will compare using distances previously mentioned) from the data x_k to the $i - th$ cluster center.

  The FCM algorithm starts with c random initial cluster centers and, at every iteration, it finds the fuzzy membership of each data point to every cluster using the following equation:

$$U_{ik} = \frac{1}{\sum_{j=1}^{c}(\frac{\|x_k - v_i\|}{\|x_k - v_j\|})^{\frac{2}{m-1}}} \tag{2}$$

  and updating the centroids v_i with:

$$v_i = \frac{\sum_{i=1}^{n} U_{ik}^m \cdot X_k}{\sum_{i=1}^{n} U_{ik}^m} \tag{3}$$

Algorithm 1 Pseudo-code of FCM Algorithm. Adapted from [29].

Fix $c, 2 < c < n$;
Fix ϵ, (e.g., $\epsilon = 0.001$);
Fix maxIterations, (e.g., maxIterations = 100);
Choose any norm distance;
Fix $m, 1 < m < \infty$, (e.g., $m = 2$);
Randomly initialize $V_0 = v_1, v_2, \dots, v_c$ cluster centers;
for $t = 1$ **to** maxIterations **do**
    Update the membership matrix U using Eq. 2;
    Calculate the new cluster centers V^t using Eq. 3;
    Calculate the new objective function J_m^t using Eq. 1;
    if $(abs(J_m^t - J_m^{t-1}) < \epsilon$ **then**
        break;
    else
        $J_m^{t-1} = J_m^t$;
    end if
end for

The pseudo-code of the FCM is shown in Algorithm 1.

This technique is an unsupervised method, which means that no tagged data is needed. Therefore when applying it in image segmentation, it is very useful because we only need the image to the segment as the input data. On the other hand, due to FCM uses the inner-product norm to find the nearest pixels, the sensitivity to noise is reduced. This method is often used in pattern recognition and segmentation of images.

- **K-means [8]:** K-means is a non-supervised and non-hierarchical clustering method. Non-hierarchical methods are characterized by two main steps. First, it assigns a cluster to each datum, and then clusters are recalculated from data assigned to them. Non-hierarchical methods clusters according to two criteria: distance or similarity. Particularly, K-means works with distance criteria [8].

  The algorithm solves a minimization problem, where the function to be minimized is the sum of the quadratic distances of each object to the centroid of its cluster, as the next equation shows:

$$\min_{s} \; E(\mu_i) = \sum_{i=1}^{k} \sum_{x_j \in S_i} \| x_i - \mu_j \|^2 \tag{4}$$

  where k is the number of clusters with their corresponding centroid μ_i and S is the set of data whose elements are the objects x_j represented by vectors, where each of its elements represents a characteristic or attribute. Then, to update the centroids we use the following equation:

$$\mu_i^{(t+1)} = \frac{1}{\|S_i^{(t)}\|} \sum_{x_j \in S_i^{(t)}} x_j \tag{5}$$

Algorithm 2 Pseudo-code of K-Means Algorithm. Adapted from [30]

input: k (number of clusters),
    D (a set of lift ratios);
output: a set of k clusters;
Arbitrarily choose k objects from D as the initial cluster centers;
repeat
    1. Based on the mean value of the objects, assign or reassign each
    object to the cluster.
    2. Calculate the mean value of the objects for each cluster; and
    update the cluster means.
until no change;

The pseudo-code of K-means method is shown in Algorithm 2.

This method is very effective in producing tighter clusters than hierarchical clustering when the clusters are globular.

3.4. Metrics

Sensitivity ($T_p/(T_p + F_n)$), Specificity ($(Tn/(T_n + F_p)$), and Accuracy ($(T_p + T_n)/(T_p + F_n + T_n + F_p)$) metrics were computed to determine how the different distance measures influence on the performance achieved by clustering algorithms to correctly distinguish pixels of object of interest from background ones. The parameters T_p, T_n, F_p, and F_n are obtained from the confusion matrix, and refer to the foreground pixels correctly classified as foreground, background pixels correctly classified as background, foreground pixels erroneously classified as background, and background pixels erroneously classified as foreground, respectively. Thus, sensitivity quantifies the ratio of pixels correctly classified as foreground over the total foreground pixels. Specificity quantifies the ratio of pixels correctly classified as a background over the total background pixels, and Accuracy quantifies the ratio of pixels correctly classified.

4. Experimental Results

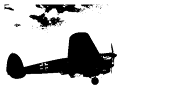
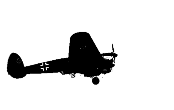
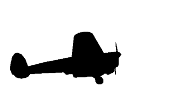
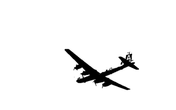
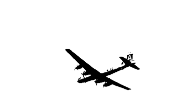
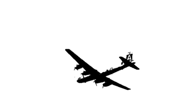
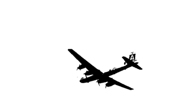
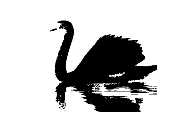
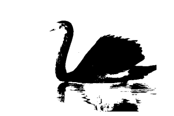
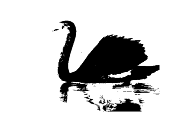
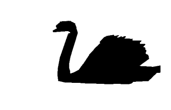
To test the proposed methodology, software routines were implemented in Python 3 using libraries like Numpy, Sys, Opencv, and plplot. Experiments were performed in Ubuntu 18.04 with kernel 4.18 installed in a laptop with an AMD processor with two cores running at 1.6 Ghz.

For experimental purposes, FCM and K-means use two centroids and one iteration. Particularly, FCM uses a fuzzier value of 1.5 based on the experimental setup introduced in [31].

This work uses a subset of 10 images from the Berkeley image dataset [32]. This dataset contains 500 natural images in format jpg along with their corresponding ground-truth (images correctly segmented). Each image of the subset contains one object of interest such as a swan, a plane, and a wildebeest in external environments with a resolution of 481 × 321.

All of the software routines implemented for testing purposes is available online at the

Table 2
Results obtained with different distance measures in FCM algorithm.

Input	Segmented Image with				Ground Truth
	Euclidean	Canberra	Spearman	Mahatan	
					
					
					
					
					

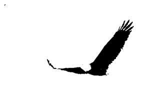
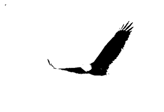
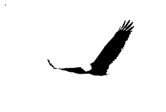
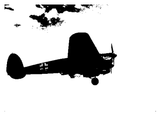
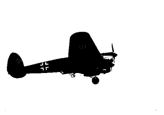
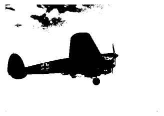
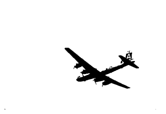
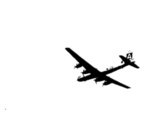
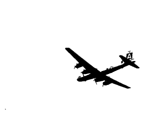
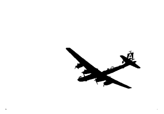
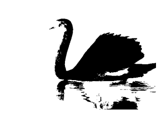
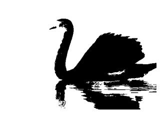
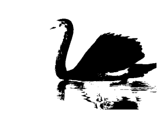
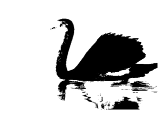
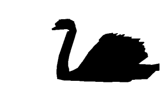
repository: *https://github.com/martinvelezf/Comparative-Study-of-Distance-Measures-for-FCM-K-means-Applied-in-Image-Segmentation*

4.1. Preliminary results

Experimental results depicted in the Tables 2 and 3 clearly exhibit that Canberra distance leads to less noise and pixels erroneously classified as foreground and background pixels in the majority of evaluated images. In contrast, in terms of accuracy, Canberra distance is higher by a value of 93.5%, in both FCM and K means implementations. However, based on Table 4, the greater maximum values of sensitivity values are reached by FCM with Manhattan distance by a 74.3%, meanwhile, K-mean outperforms when it uses Spearman one by a 74.2%. The other metrics perform that their maximum value is achieved in Canberra distance for both methods, FCM and K means.

Table 4 presents a comparison between averaged values between FCM and K-means. The column "dif" is the difference between the average, and the column "best" represents the method with the higher average value. Results presented in Figures 2, and 4 show that Canberra distance reaches the highest average value of accuracy and specificity. The performance was better at images where the object of interest contrasted with the background or was in a similar area. Besides, there was a background with a width color spectrum, so that some parts of it were identified as objects of interest.

Table 3
Results obtained with different distance measures in K-means algorithm.

Input	Segmented Image with				Ground Truth
	Euclidean	Canberra	Spearman	Mahatan	
					
					
					
					
					

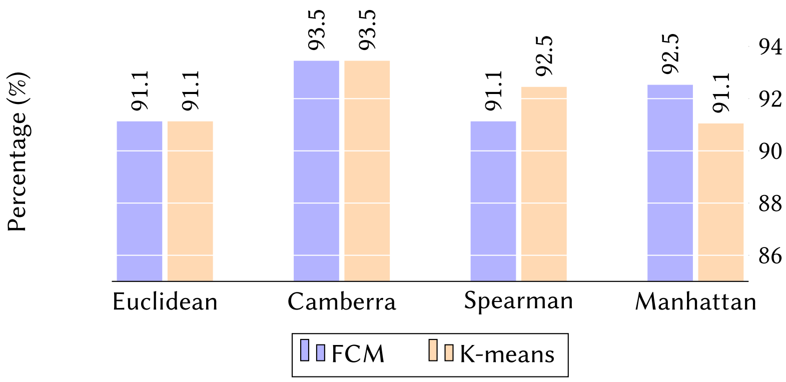


Figure 2: Comparison results: average accuracy of FCM and K-means

5. Conclusions

This work explores the combination of RGB and HSV color spaces for segmentation, however, the segmentation is still a challenging task when the background and foreground are characterized by similar color values. The manual selection of the centroids manually achieves better outcomes than automatic selection approaches.

Table 4
Performance metric values reached by K-means and FCM with different distance measures.

Distance	Metric	AVGFCM(%)	AVGKM(%)	DIF	BEST
Euclidean	Accuracy	91.13	91.13	1.30E-05	K-means
	Sensitivity	70.65	70.65	6.95E-05	K-means
	Specificity	94.07	94.07	7.43E-06	K-means
Camberra	Accuracy	93.45	93.45	1.30E-05	K-means
	Sensitivity	71.20	71.21	6.95E-05	K-means
	Specificity	96.27	96.27	7.43E-06	K-means
Spearman	Accuracy	91.13	92.45	1.32E-02	K-means
	Sensitivity	70.65	74.17	3.52E-02	K-means
	Specificity	94.07	95.13	1.07E-02	K-means
Manhattan	Accuracy	92.53	91.05	1.48E-02	FCM
	Sensitivity	74.26	70.63	3.63E-02	FCM
	Specificity	95.21	93.98	1.24E-02	FCM

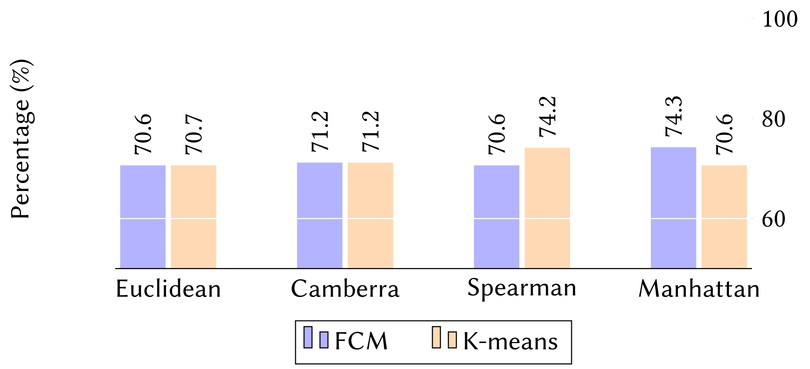


Figure 3: Comparison results: average sensitivity of FCM and K-means

The efficiency of the clustering algorithms, using different distance measures has been evaluated with the accuracy, sensitivity, and specificity metrics. According to the results, the Canberra distance gives the best accuracy and specificity. Furthermore, the best sensitivity was achieved by the Manhattan distance applied in FCM, follow by the Spearman distance in K-means.

A limitation of the proposed implementation is that the number and values of clusters to be segmented out should be selected manually. In order to improve the utility, the proposed implementation should be facilitated with mechanisms that could adaptively determine the appropriate number of clusters to be segmented out. This is an important issue for future work.

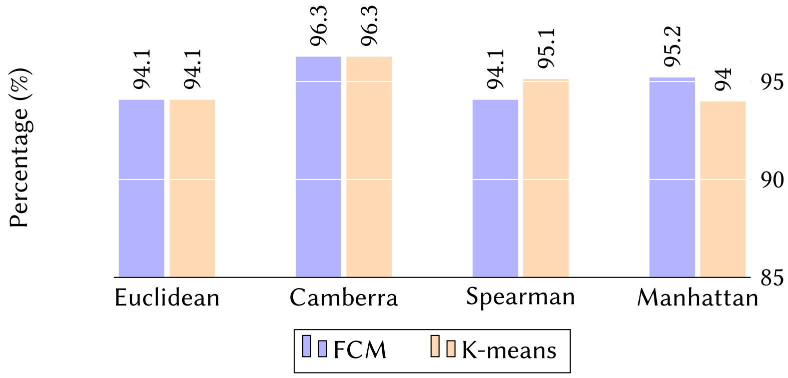


Figure 4: Comparison results: average specificity of FCM and K-means

Acknowledgment

The authors thank the valuable support given by the SDAS Research Group (www.sdas-group.com) and the head of the group Diego Peluffo.

References

[1] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, D. Terzopoulos, Image segmentation using deep learning: A survey, arXiv preprint arXiv:2001.05566 (2020).

[2] N. Dhanachandra, K. Manglem, Y. J. Chanu, Image segmentation using k-means clustering algorithm and subtractive clustering algorithm, Procedia Computer Science 54 (2015) 764–771.

[3] A. Dik, K. Jebari, A. Bouroumi, A. Ettouhami, A new fuzzy clustering by outliers, Journal of Engineering and Applied Sciences 9 (2014) 372–377.

[4] W. Cai, S. Chen, D. Zhang, Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation, Pattern recognition 40 (2007) 825–838.

[5] A. A.-h. Hassan, W. M. Shah, M. F. I. Othman, H. A. H. Hassan, Evaluate the performance of k-means and the fuzzy c-means algorithms to formation balanced clusters in wireless sensor networks., International Journal of Electrical & Computer Engineering (2088-8708) 10 (2020).

[6] N. Singh, A. G. Mohapatra, G. Kanungo, Breast cancer mass detection in mammograms using k-means and fuzzy c-means clustering, International Journal of Computer Applications 22 (2011) 15–21.

[7] S. Madhukumar, N. Santhiyakumari, Evaluation of k-means and fuzzy c-means segmentation on mr images of brain, The Egyptian Journal of Radiology and Nuclear Medicine 46 (2015) 475–479.

[8] A. Singh, A. Yadav, A. Rana, K-means with three different distance metrics, International Journal of Computer Applications 67 (2013).

[9] A. B. Amina Dik, 1Abdelaziz El moujahid, A. Ettouhami, Weighted distances for fuzzy clustering, Applied Mathematical Sciences 8 (2014) 147 – 156.

[10] S. Khanmohammadi, N. Adibeig, S. Shanehbandy, An improved overlapping k-means clustering method for medical applications, Expert Systems with Applications 67 (2017) 12–18.

[11] S. Barak, T. Mokfi, Evaluation and selection of clustering methods using a hybrid group mcdm, Expert Systems with Applications 138 (2019) 112817.

[12] P. K. Mishro, S. Agrawal, R. Panda, A. Abraham, A novel type-2 fuzzy c-means clustering for brain mr image segmentation, IEEE Transactions on Cybernetics (2020) 1–12.

[13] S. Tyagi, S. Malhotra, D. Kumar, V. S. Verma, A. Bhardwaj, Mammographic image segmentation with modified fcm based clustering algorithm, AIP Conference Proceedings 2214 (2020) 020028. URL: https://aip.scitation.org/doi/abs/10.1063/5.0003604. doi:10.1063/5.0003604.

[14] P. Kavitha, P. V. Saraswathi, Segmentation for content based satellite image retrieval using fuzzy clustering, International Journal of Advanced Science and Technology 29 (2020).

[15] R. Loohach, K. Garg, Effect of distance functions on simple k-means clustering algorithm, International Journal of Computer Applications 49 (2012) 7–9. doi:10.5120/7629-0698.

[16] D. Bora, D. Gupta, Effect of different distance measures on the performance of k-means algorithm: An experimental study in matlab, arXiv 5 (2014).

[17] M. K. Arzoo, A. Prof, K. Rathod, K-means algorithm with different distance metrics in spatial data mining with uses of netbeans ide 8. 2, Int. Res. J. Eng. Technol 4 (2017) 2363–2368.

[18] A. Singh, A. Yadav, A. Rana, K-means with three different distance metrics, International Journal of Computer Applications 67 (2013).

[19] A. Román y Zubeldia, Implementación de pruebas para una hipótesis sobre la aplicación de distancia euclidiana para realizar agrupamientos en espacios multidimensionales, 2018.

[20] M. Bora, D. Jyoti, D. Gupta, A. Kumar, Effect of different distance measures on the performance of k-means algorithm: an experimental study in matlab, arXiv preprint arXiv:1405.7471 (2014).

[21] Y. Oike, M. Ikeda, K. Asada, A high-speed and low-voltage associative co-processor with exact hamming/manhattan-distance estimation using word-parallel and hierarchical search architecture, IEEE Journal of Solid-State Circuits 39 (2004) 1383–1387.

[22] D. Androutsos, K. Plataniotiss, A. N. Venetsanopoulos, Distance measures for color image retrieval, in: Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No. 98CB36269), volume 2, IEEE, 1998, pp. 770–774.

[23] V. Kumar, J. K. Chhabra, D. Kumar, Performance evaluation of distance metrics in the clustering algorithms, INFOCOMP 13 (2014) 38–52.

[24] H.-D. Cheng, X. H. Jiang, Y. Sun, J. Wang, Color image segmentation: advances and prospects, Pattern recognition 34 (2001) 2259–2281.

[25] D. Bora, A. Gupta, F. Khan, Comparing the performance of l*a*b* and hsv color spaces with respect to color image segmentation, arXiv (2015).

[26] S. Sural, G. Qian, S. Pramanik, Segmentation and histogram generation using the hsv color space for image retrieval, in: Proceedings. International Conference on Image Processing, volume 2, IEEE, 2002, pp. II–II.

[27] D. Kaur, A comparative study of various distance measures for software fault prediction, arXiv preprint arXiv:1411.7474 (2014).
[28] J. C. Dunn, A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters, Journal of Cybernetics 3: 32-57 (1973).
[29] O. Moh'd Alia, R. Mandava, M. E. Aziz, A hybrid harmony search algorithm for mri brain segmentation, Evolutionary Intelligence 4 (2011) 31–49.
[30] P.-Y. Zhou, K. C. Chan, A model-based multivariate time series clustering algorithm, in: Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer, 2014, pp. 805–817.
[31] K.-L. Wu, Analysis of parameter selections for fuzzy c-means, Pattern Recognition 45 (2012) 407–415.
[32] P. Arbelaez, M. Maire, C. Fowlkes, J. Malik, Contour detection and hierarchical image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2011) 898–916. URL: http://dx.doi.org/10.1109/TPAMI.2010.161. doi:10.1109/TPAMI.2010.161.