Proceedings of the International MultiConference of Engineers and Computer Scientists 2015 Vol I,
IMECS 2015, March 18 - 20, 2015, Hong Kong

A Comparative Study on CT Image Segmentation Using FCM-based Clustering Methods

Chih-Hung Wu, Xian-Ren Lo, and Chen-Sen Ouyang

***Abstract*—Identifying specific CT-image regions is an important process in medical diagnosis. Clustering is a simple and useful means for automatic image segmentation. However, clustering results vary with the features of image pixels and the settings of parameters of the clustering methods. This study compares the results of CT image segmentation using FCM-based clustering algorithms running with intensity- and texture-based image features. Three types of image features, grayscale, LBP, and grayscale+LBP, are investigated. KM, FCM, and their medoid-variations are tested with various parameter settings. The results show that FCM and the grayscale+LBP feature can produce reasonable and satisfactory clustering results for CT-image segmentation.**

***Index Terms*—image feature, image segmentation, local binary pattern (LBP), FCM-based clustering.**

I. INTRODUCTION

CT scan is an imaging modality which uses X-rays to obtain structural and functional information about the human body [1]. Because animal tissues have various degrees of X-ray absorption, they can be imaged in a CT scan as pixels with different intensity. For example, dense tissues such as bones are white in a CT image, soft tissues such as brain or liver are gray, tissues filled of air or cavity may be black, etc. With the help of the CT scan technology, medical diagnosis advances effectively and more accurately. The investigation of CT images usually relies on human medical doctors or experts, which is time-consuming and error-prone. Automated analysis of CT images can reduce human's efforts and provide summarized information for fast diagnosis and has received increasing attention[2].

Automatically identifying specific image regions that may represent healthy tissues or suspicious nidus is an important process for CT image analysis. The technique of image segmentation is to partition a given image into homogeneous and meaningful regions with specific features and is a useful tool for CT image analysis. Among various techniques that are developed for image segmentation, clustering is one of the commonly used methods. A clustering algorithm is an unsupervised learning process that collects data points with homogeneous features into the same cluster and discriminates clusters by data points with heterogeneous features [3]. For image segmentation, features associated with image pixels, such as color intensity, textures, pixel positions, are calculated for clustering validity. Usually, the Euclidean-distance is used for discriminating homogeneity/heterogeneity of clusters. Clustering results as well as the information obtained from clusters may vary with the use of different image features and parameters settings of clustering algorithms.

CT images are usually monochromatic grayscale so that image pixels form one-dimensional data for clustering. Clustering on one-dimensional data may not obtain correct image segments due to few discriminative information. Texture describes the variation among pixels and reserves structural information of image regions and serves as an effective feature for image clustering. Some texture-based clustering have been studied, such as [4], [5], [6]. The local binary patterns (LBP) is a widely used texture encoding for image clustering because of its robustness to illumination and pose variations and low computational complexity [7], [8], [9], [10].

This study compares the results of CT image segmentation using intensity- and texture-based image features and clustering algorithms. Three features are used for clustering: grayscale intensity values (one-dimensional), LBP textures extracted from grayscale values (one-dimensional), and grayscale+LBP (two-dimensional). The clustering algorithms used in this study are the fuzzy c-means clustering (FCM) algorithm and its variations. There are some parameters effecting the performance of FCM, such as the selection of centroids, the stopping criteria, and the degree of fuzziness. This paper presents the clustering results of CT image segmentation using various settings of image features and FCM-based algorithms.

The remaining part of this paper is organized as follows. Section II reviews FCM and its variations. The extraction of features from CT images is discussed in Section III. The performance of CT clustering is presented and discussed in Section IV. Finally, conclusions are given in Section V.

II. FCM-BASED CLUSTERING

The following terms are used for describing the clustering algorithms.

- N: number of data points for clustering
- m: fuzzifier which determines the level of cluster fuzziness
- x_i: the i-th, $1 \leq i \leq N$, data point
- K: number of clusters
- C_k: the k-th, $1 \leq k \leq K$, cluster
- $\|C_k\|$: number of data points in C_k
- v_k: centroid of C_k
- $\bar{v}$: centroid of all data points, i.e., $\bar{v} = \frac{1}{N}\sum_{i=1}^{N} x_i$
- $|x - y|$: distance between a pair of data points, a pair of cluster centroids, or an object and a centroid, x and

This work was supported in part by the Ministry of Science and Technology, Taiwan, under grant MOST 103-2221-E-390-016.

Chih-Hung Wu and Xian-Ren Lo are with the Department of Electrical Engineering, National University of Kaohsiung, No. 700, Kaohsiung University Rd., Nan-Tzu District, Kaohsiung 811, Taiwan. e-mail: (johnw@nuk.edu.tw)(my0936@gmail.com).

Chen-Sen Ouyang is with the Department of Information Engineering, I-Shu University, Dashu District, Kaohsiung 840, Taiwan. e-mail: (ouyangcs@isu.edu.tw)

ISBN: 978-988-19253-2-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

y
- μ_{ik}: membership degree of x_i "belonging-to" C_k

A. K-means Clustering

First of all, the K-means (KM) [11] is briefed. KM is a partition-based clustering method that clusters the data set of N data points into k clusters with k known *a priori*. KM performs an iterative process that assigns each data point to a cluster by considering their distance and minimizing the following objective function:

$$\sum_{i=1}^{N}\sum_{k=1}^{K}|x_i - v_k|^2. \tag{1}$$

Initially, KM starts with a given number K of clusters and randomly chooses K centroids. The objective function shown in Eq.(1) is optimized by iteratively updating v_k as

$$v_k = \frac{1}{|C_k|}\sum_{i=1}^{N} x_i. \tag{2}$$

The iteration process stops when v_k does not change. Otherwise, new centroids are calculated according to Eq.(2) and the iteration goes on.

B. Fuzzy C-means Clustering

The FCM algorithm, which is developed by Dunn [12], is a widely used clustering method. FCM can be considered an improved version of KM by considering the membership degree in terms of fuzziness. FCM also uses an iterative process similar to KM that assigns each data point to each cluster with a certain fuzzy membership degree through the minimization of the following objective function:

$$\sum_{i=1}^{N}\sum_{k=1}^{K}\mu_{ik}^{m}|x_i - v_k|^2, m \geq 1. \tag{3}$$

As that in KM, FCM starts with a given number K of clusters and randomly chooses K centroids. The objective function of Eq.(3) is optimized by iteratively updating μ_{ik} and v_k as

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{K}\left(\frac{|x_i - v_k|}{|x_i - v_j|}\right)^{\frac{2}{m-1}}}, \tag{4}$$

$$v_k = \frac{\sum_{i=1}^{N}\mu_{ik}^{m}x_i}{\sum_{i=1}^{N}\mu_{ik}^{m}}. \tag{5}$$

The iteration stops when

$$\|U_{p+1} - U_p\| < \varepsilon, \tag{6}$$

where U_p=[μ_{ik}] is the matrix composed of all μ_{ik}'s, p is the number of iterations, and ε is a threshold given by the user. Otherwise, new centroids are calculated according to Eq.(5) and the iteration goes on.

C. K-Medoids clustering

In both KM and FCM, the position of a centroid, v_k, can be any location in the data space. By investigating Eq.(2) and Eq.(5), it is possible that there is no data point on the location calculated for v_k. In some clustering applications, this may cause ridiculous explanation of data. The K-medoids (KMm) [13] clustering algorithm is designed for solving such a problem. KMm is similar to KM except the method of determining v_k. For allocating v_k, KMm first calculates v_k's location by .(2), then chooses the data point which is the nearest one to the position calculated by Eq.(2). That is, every v_k is allocated on the a data point, i.e., $\exists i$, v_k=x_i. Similarly, the same allocation policy can be applied to FCM. For convenience, we term the medoids-version of FCM as FCMm.

III. Features of CT Images

As mentioned previously, CT-images are grayscale; they are described as a sequence of one-dimensional grayscale intensity values. In this paper, the texture encoded in LBP is also used for describing CT-images.

A. Grayscale Intensity

A $k \times k$ grayscale image can be viewed as a sequence of k^2 integers,

$$\langle g_1, g_2, \ldots, g_{k^2}\rangle.$$

Each g_i, $1 \leq i \leq k^2$, represents the intensity of grayscale. For a b-bit grayscale image, $0 \leq g_i \leq b^2 - 1$. For clustering, the distance between two pixels g_i and g_j, $1 \leq i, j \leq k^2$, is the difference of their grayscale levels, i.e.,

$$\sqrt{(g_i - g_j)^2}. \tag{7}$$

B. LBP Texture

Given a circular neighborhood of radius R centered on pixel g_c in a gray-scale intensity image, the LBP label of g_c is defined as below.

$$LBP_{P,R}(g_c) = \sum_{P=0}^{P-1} s(g_p - g_c)2^P, \tag{8}$$

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{9}$$

where g_p is the gray-scale intensity of P pixels in the circular neighborhood and $s(x)$ is the function which outputs 0 and 1 as the result of comparisons. Fig. 1 presents an LBP encoding example with $P = 8$ and $R = 1$. A $k \times k$ grayscale image encoded by LBP labels can be viewed as a sequence of k^2 integers,

$$\langle p_1, p_2, \ldots, p_{k^2}\rangle.$$

Each p_i, $1 \leq i \leq k^2$, represents the LBP label of g_i, i.e., $LBP_{P,R}(g_i)$. When $P = 8$ and $R = 1$, the range of LBP values in a b-bit grayscale CT image is defined, $0 \leq p_i \leq b^2 - 1$. For clustering, the distance between two pixels g_i and g_j, $1 \leq i, j \leq k^2$, is the difference of their LBP values as

$$\sqrt{(LBP_{P,R}(g_i) - LBP_{P,R}(g_j)^2)}. \tag{10}$$

Fig. 2 illustrates a CT-image described in grayscale and LBP features, respectively.

ISBN: 978-988-19253-2-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

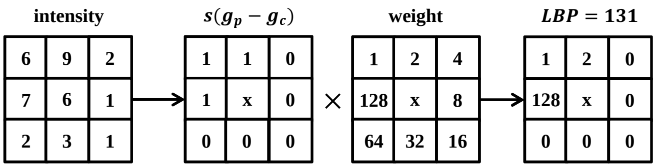
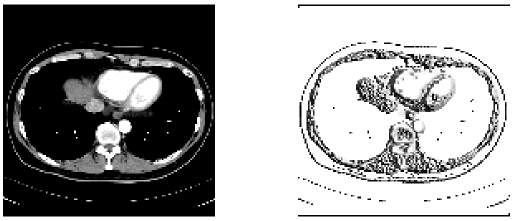


Fig. 1. An LBP encoding example



(a) Grayscale (b) LBP

Fig. 2. A CT image in grayscale and LBP feautres

C. Grayscale + LBP

Combining grayscale and LBP labelling, a $k \times k$ grayscale CT-image can be considered as a sequence of k^2 integer-pairs,

$$\langle (g_1, p_1), (g_2, p_2), \ldots, (g_{k^2}, p_{k^2}) \rangle.$$

For clustering, the distance between two pixels g_i and g_j, $1 \leq i, j \leq k^2$, is calculated by

$$\sqrt{(g_i - g_i)^2 + (LBP_{P,R}(g_i) - LBP_{P,R}(g_j)^2)} \qquad (11)$$

With both grayscale and LBP, a two-dimensional feature space is organized that may provide advanced discriminative information for cluster validity evaluation.

IV. EXPERIMENT

To evaluate the effectiveness of various features and clustering methods for CT-image clustering, the following experiments are performed. Six CT-images are chosen for test, as shown in Fig. 3. These images are in 8-bit grayscale and 256×265 in size. Data sets to be clustered are generated based on the three types of features, as mentioned in Section III. Four clustering algorithms are implemented; they are FCM, KM, FCMm, and KMm. These clustering algorithms are implemented in C++. All experiments are conducted in a personal computer with 8G RAM and a Core i3 CPU running on Windows 7. The clustering algorithms run with various K=2–10 (number of clusters) and the best results are retained.

Fig. 4–Fig. 9 present some selected clustering results of CT-images. In these figures, pixels in the same color belong to the same cluster. It seems that clustering with grayscale+LBP as features can form better recognizable clusters; clustering with grayscale only may not always has satisfactory results. Table I–Table III presents the number of clusters obtained from the four clustering algorithms with fuzziness degree m=2.

TABLE I
FCM AND FCMM WITH GRAY, LBP, AND GRAY+LBP FEATURES

	FCM	FCM	FCM	FCMm	FCMm	FCMm
Image	Gray	LBP	Gray+LBP	Gray	LBP	Gray+LBP
CT52	3	2	3	3	2	3
CT131	3	2	4	4	2	3
CT139	3	2	3	3	2	3
CT155	3	2	4	4	2	3
CT176	3	2	3	3	2	3
CT181	3	2	3	3	2	3

TABLE II
RESULTS OF CT IMAGE CLUSTERING ($m = 2$)

Image	FCM	FCMm	KM	KMm
CT52	3	3	3	3
CT131	3	4	4	4
CT139	3	4	3	3
CT155	3	3	4	4
CT176	3	3	3	3
CT181	3	4	3	3

TABLE III
CLUSTERING RESULTS OF FCM WITH VARIOUS m

m=	1.5	2.0	3.0	4.0	5.0	6.0	7.0	8.0
CT52	3	3	3	3	3	4	4	4
CT131	4	3	3	3	3	3	4	4
CT139	3	3	3	3	3	3	3	3
CT155	3	3	4	3	3	4	3	3
CT176	4	3	3	3	3	3	3	3
CT181	3	3	3	3	3	3	4	4

TABLE IV
CLUSTERING RESULT OF FCMM WITH VARIOUS m

m=	1.5	2.0	3.0	4.0	5.0	6.0	7.0	8.0
CT52	3	3	3	3	3	4	4	4
CT131	4	3	3	3	3	3	4	4
CT139	3	3	4	3	3	3	3	3
CT155	3	3	4	3	3	4	3	3
CT176	4	3	3	3	3	3	3	3
CT181	3	3	4	3	3	3	4	4

V. CONCLUSION

Analysis of CT images by image segmentation is an import task in medical diagnosis. Clustering is a useful method for fast and effective image segmentation. This study compares the results of CT image segmentation using intensity- and texture-based image features and clustering algorithms. The effectiveness of three features, grayscale, LBP, and grayscale+LBP, for clustering are investigated. The performance of four FCM-based clustering algorithms with various parameter settings are discussed. The results show that FCM and grayscale+LBP can produce reasonable and satisfactory clustering results for CT-image segmentation. Parameters associated with clustering algorithms and feature selection, such as the degree of fuzziness m in Eq.(3) and P and R in Eq.(8) should be investigated. These will be included in our future work.

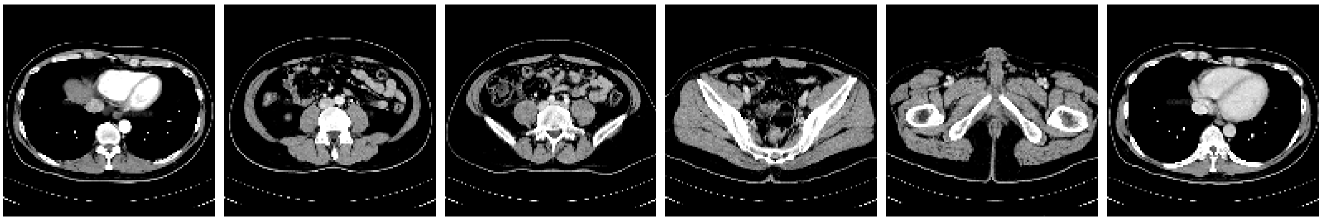
REFERENCES

[1] N. Sharma and L. M. Aggarwal, “Automated medical image segmentation techniques,” *Journal of Medical Physics*, vol. 35, pp. 3–14, 2010.

[2] H. Ng, S. Ong, K. Foong, P. Goh, and W. Nowinski, “Medical image segmentation using k-means clustering and improved watershed algorithm,” in *Proceedings of 7th IEEE Southwest Symposium on Image Analysis and Interpretation*, 2006, pp. 61–65.

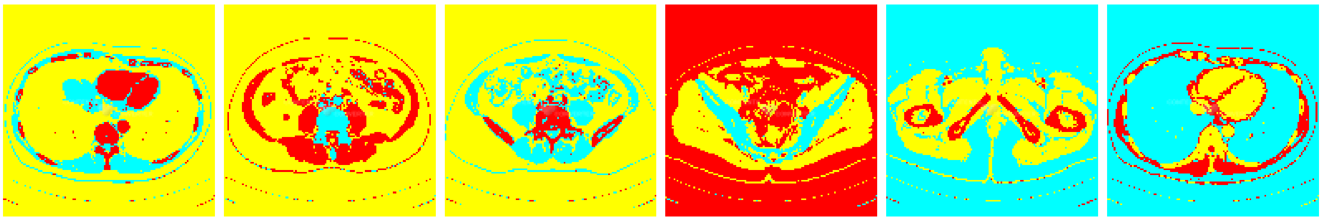
[3] C. Zhang and P. Wang, “A new method of color image segmentation based on intensity and hue clustering,” in *In Proceedings of 15th International Conference on Pattern Recognition*, vol. 3, 2000, pp. 613–616.

ISBN: 978-988-19253-2-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)



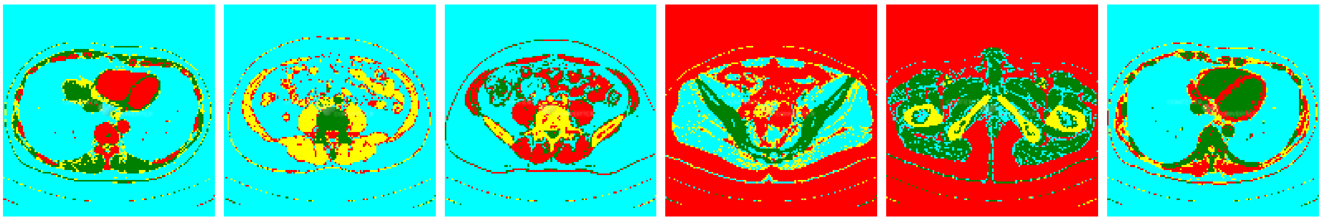
(a) CT52 (b) CT131 (c) CT139 (d) CT155 (e) CT176 (f) CT181

Fig. 3. Original CT Images



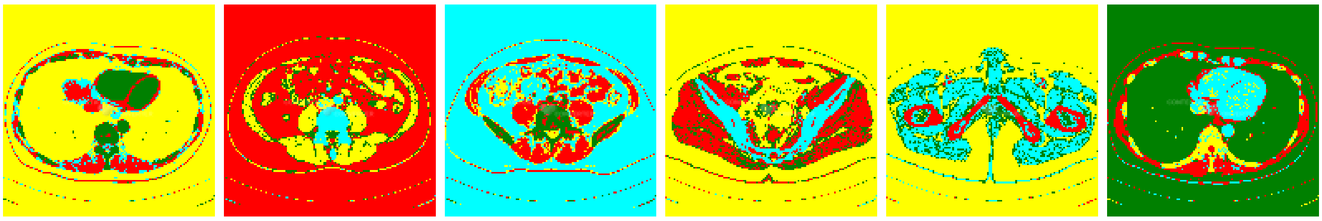
(a) CT52 (b) CT131 (c) CT139 (d) CT155 (e) CT176 (f) CT181

Fig. 4. Clustering results of KM with $K = 3$



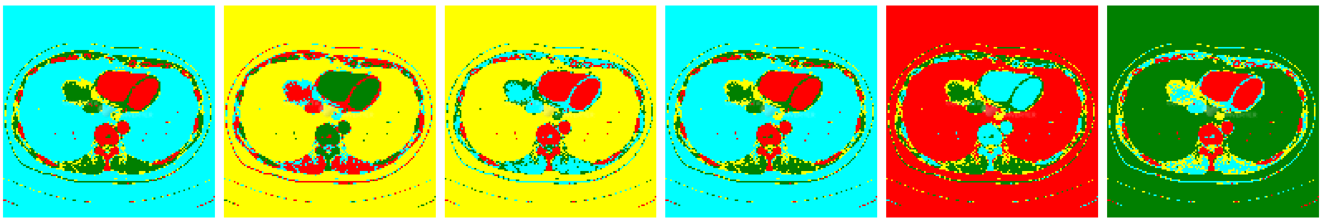
(a) CT52 (b) CT131 (c) CT139 (d) CT155 (e) CT176 (f) CT181

Fig. 5. Clustering results of KM with $K = 4$



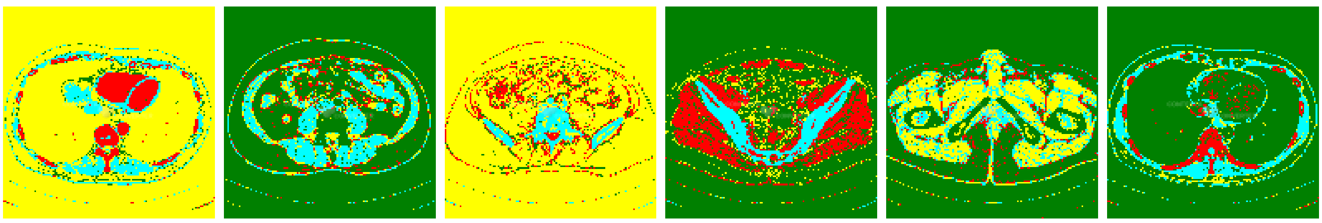
(a) CT52 (b) CT131 (c) CT139 (d) CT155 (e) CT176 (f) CT181

Fig. 6. Clustering results of FCM with $K = 4$ and $m = 3$



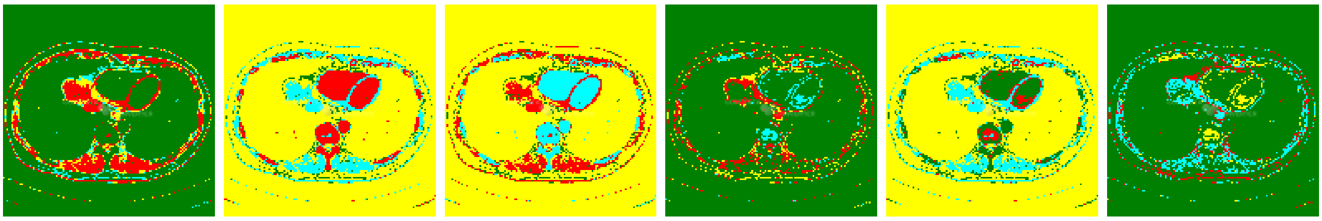
(a) $m = 2$ (b) $m = 3$ (c) $m = 4$ (d) $m = 5$ (e) $m = 6$ (f) $m = 7$

Fig. 7. Clustering results of CT52 using FCM with $K = 4$



(a) CT52 (b) CT131 (c) CT139 (d) CT155 (e) CT176 (f) CT181

Fig. 8. Clustering results of FCMm with $K = 4$ and $m = 3$



(a) $m = 2$ (b) $m = 3$ (c) $m = 4$ (d) $m = 5$ (e) $m = 6$ (f) $m = 7$

Fig. 9. Clustering results of CT52 using FCMm with $K = 4$

[4] A. Suruliandi and K. Ramar, “Local texture patterns - a univariate texture model for classification of images,” in *Proceedings of the International Conference on Advanced Computing and Communications*, 2008, pp. 32–39.

[5] L. Zhou and H. Wang, “Open/closed eye recognition by local binary increasing intensity patterns,” in *Proceedings of the IEEE Conference on Robotics, Automation and Mechatronics*, 2011, pp. 7–11.

[6] A. R. Rivera, R. Castillo, and O. Chae, “Local directional number pattern for face analysis: Face and expression recognition,” *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1740–1752, May 2013.

[7] T. Ojala, M. Pietikainen, and D. Harwood, “A comparative study of texture measures with classification based on featured distributions.” *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[8] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[9] C.-H. Wu, L.-W. Lu, , and Y.-Y. Li, “A study on pattern encoding of local binary patterns for texture-based image segmentation,” in *In Proceedings of the 6th International Conference on Machine Learning and Cybernetics (ICMLC)*, 2014, pp. 13–16.

[10] D.-Q. Zhang and S.-C. Chen, “A novel kernelized fuzzy c-means algorithm with application in medical image segmentation,” *Artificial Intelligence in Medicine*, vol. 32, pp. 37 – 50, 1 2004.

[11] J. B. MacQueen, “Some methods for classification and analysis of multivariate observations,” in *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, 1, Ed. Berkeley, University of California Press, 1967.

[12] J. C. Dunn, “A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters,” *Journal of Cybernetics*, vol. 3, pp. 32 – 57, 1974.

[13] Y. Dodge, “An introduction to l1-norm based statistical data analysis,” *Computational Statistics & Data Analysis*, vol. 5, no. 4, pp. 239 – 253, 1987.

ISBN: 978-988-19253-2-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)