Utilizing Clustering FCM And Contrast Based Enhancement to Isolate Covid-19 Infected Regions in CT Lung Images

Assel Ali Hussein[1]*, Rabab Saadoon Abdoon[2]
[1, 2]University of Babylon, College of Science, Department of Physics, Iraq
Email: Asselali9009@gmail.com

Abstract
Covid-19 disease that directly affecting lungs is an acute disease caused death of many people around word, so the early detecting of it and asses the relative ratio of the lung infection is a vital need. In this work, Histogram based contrast adjustment was implemented to enhance four lung abnormal CT scan images to highlight the abnormal regions within the experimental images. Fuzzy c-mean algorithm then was applied to segment the images in order to detect and isolate the infected regions. Besides, several morphological operations were employed to extract the refined infected Covid-19 areas effectively with accuracy of 96%.
Keywords: Clustering, covid-19, Contrast adjustment, Fuzzy c-mean, CT scan

1. Introduction

Coronavirus disease (Covid-19) On January 30, 2020, the World Health Organization announced the outbreak of the Corona virus, which is of international concern because it represents a health emergency. This virus directly affects the lungs and thus affects the physical health of millions of people around the world. Corona virus also poses a worldwide threat to mental health [1]. Many researchers from different parts of world studied Covid-19 such as: Lone, Dalila, et al; Pockhrel, Sumitra, et al; Daniel,Sir John, et al; Mercer, Tim R, et al; Del Rio, Carlos, et al; Pizarro-Ortega, Carlos Ivan, et al; Wunsch, Kathrin, et al; [2-8]. There are two basic imaging modalities available to clinicians for covid-19 infections which are: x-ray and Computerized tomography, CT scan [9]. CT scan is a quick, cheap, good and accurate test for Covid-19 [10]. Processing CT lung images helps detecting the infected regions and its position as well as their surface area. The implemented techniques varied such as region-based techniques, edge detection techniques, thresholding techniques, histogram equalization, clustering technique and other more complicated methodologies that depending on a statistical analysis, for more details see [11-14].
In this study, contrast adjusting based on images histogram was utilized to highlight the different region within CT scan lung images as first stage of the work. Fuzzy c-mean was implemented to segment the first stage resultant images to detect and isolate and extract covid-19 infected regions. Many researchers have done multiple studies on Covid-19 for example, Loey et al., proposed a model consisting of dataset preprocessing and then using transfer learning to train. This dataset consisted of 749 CT images. The test's accuracy was estimated using a confusion matrix, where transfer learning models were evaluated to diagnose COVID-19 using the original dataset and with the preprocessing procedure. Resnet50 achieved the highest accuracy of 82% with preprocessing [15]. Santa Cruz in this study proposed a multi-stage model for diagnosing COVID-19 through CT scan using six pre-trained models at one stage. Then, at another stage, the ensemble strategy was applied to collect the results of six models to predict the outcome. AUC is 90.82 %, an accuracy is 86.70 %, and an F1 score is 85.86 % [16]. Li et al., developed for COVID-19 diagnosis a tool transfer Learning 3D using a CT scan. The first part is pre-processed for 3D CT. Then the ROI (lungs) is extracted using U-Net. The second part depends on extracting the features maps by the ResNet-50 model then be max-pooled. The third and final part is feeding a fully connected layer with the feature maps and used "soft-max" as an activation function for multi-class (COVID-19, CAP, and non-pneumonic). Aurea under curve is 0.96 [17]. Previous studies trained networks for the purpose of classifying the incoming images, whether they were Covid-19 or not, without isolating and extracting the affected areas. While present study this research, the affected area was isolated and extracted by image processing methods, and this work was not done by any previous studies.

2. Fuzzy C- Mean Method

Fuzzy c mean, FCM algorithm based on grouping the input data into separate clusters. Its technique involved the probability concept. FCM is a soft clustering method that means each data item might belong to more than one cluster at the same time with a varying degree of membership according to a specific membership function. FCM works on basis of the minimization of the objective function J_m, which is represented in the following equation [9]:

$$\boldsymbol{Jm} = \sum_{i=1}^{n} \boldsymbol{u_{ij}} \sum_{i=1}^{c} \|\boldsymbol{Xi} - \boldsymbol{Cj}\|\boldsymbol{2} \text{ for } 1 \le m < \infty \quad (1)$$

Where u_{ij} is the degree of the membership, x_i the data item belong to some degree to the cluster j, c_j is the j^{th} cluster seed and $\|x_i - c_j\|$, is the distance between each designated dataset and its corresponding centroid [18].

5doi.org/10.31838/hiv23.01.136 Received: 14.09.22, Revised: 18.10.22, Accepted: 17.12.22

Partitioning process of the dataset is accomplished by iteratively minimizing the objective function. In each round the membership values u_{ij} of each data item is updated as well as the clusters centroids cj as shown in the following equation [19].

$$u_{ij} = \frac{\sum_{i=1}^{n} u_{ij}^{m} . x_i}{\sum_{=1i}^{x} x_{ij}^{m}} \quad (2)$$

Histogram

Thresholding is the simplest and the most commonly used method of segmentation. When an image *f (x, y)* contains an object having homogeneous intensity and a background with different intensity levels, such an image can be segmented into two regions by simple thresholding, the segmented image *g (x, y)* can be expressed as [20]:

$$g(x,y) = \begin{cases} 1 \; if \; f\,(x,y) > T \\ 0 \; otherwise \end{cases} \quad (3)$$

Where represents the threshold value. The choice of threshold T can be dependent on the image histogram. If the image has a single object and a background of homogeneous density, it often has a bimodal histogram as shown in the Figure (1).

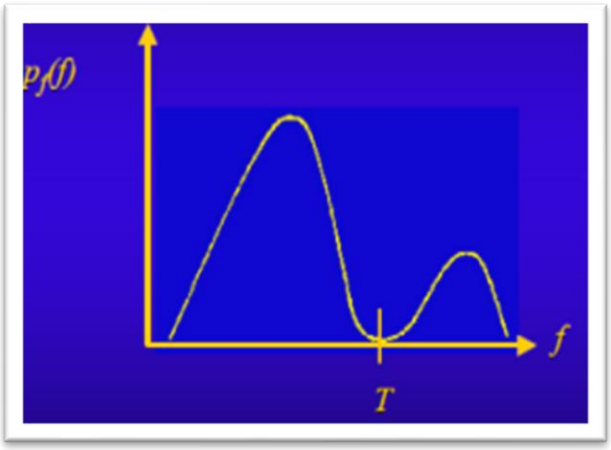

Figure (1): Bimodal histogram.

The histogram of pixel values can be denoted by: h_0; h_1; h_k ;h_N; where h_k specifies the number of pixels in an image with gray scale value k and N is the maximum pixel intensity value (typically 255). If the histogram is noisy, the calculation of the local histogram minimum is difficult, and in this case, low pass filter is utilized for smoothing to reduce the noise. If the intensity of the object or of the background is slowly varying, the histogram may not contain two clearly distinguished lobes. In this case, a spatially varying threshold can be applied.

In certain cases, the region boundary is desired. If the segmented image *g(x, y)* is available, the boundary obtained by finding the transitions from one region to another is:

$$b(x,y) = \begin{cases} 1 \; if \; \{(g(x,y) \in R_i \; and \; g(x,y-1) \in R_j \, , i \neq j \\ or \; (g(x,y) \in R_i \; and \; g(x-1,y) \in R_j \, , i \neq j \\ 0 \; otherwise \end{cases} \quad (4)$$

Image Enhancement

Image enhancement is perception strengthening of information in likeness for human viewers and guarantees the provision of improved input for many images processing approaches that are automatic. Adaptation of one or more characteristics of the digital image is done with the selection of characteristics including the enhancement process being particular to specific work because of the availability of differentiation techniques for digital image enhancement without modulation. Enhancement methods can be expressed as [21]:

$$g\,(x,y) = T\,[\,f\,(x,y)\,] \quad (5)$$

Where: f(x, y) this is input image, g(x, y) this is processed image, T: this is an operator on (f) and can be operate on the groups of input images [22]. The goal from contrast adjusting is to increase the dynamic domain of the gray levels in the digital image being treating meted by passing the value of all pixels that is between 0.1 and 0.9 to convert to 0 and 1, depending on the vision request [23].in this paper using the Contrast adjustment remaps image intensity values to the full display range of the data type. An image with good contrast has sharp differences between black and white.

Morphological Operations

Morphological operators are used in image processing due to their robust performance in preserving the shape of a signal, while suppressing the noise [24]. The basic idea in mathematical morphology is to convolve an image with a given mask (known as the structuring element) [25].

There are two basic morphological operators: erosion and dilation, opening and closing are two derived operations in terms of erosion and dilation. Here a brief introduction to the morphological operations [26]:

The primitive operations of mathematical morphology, dilation, and erosion are based on set theory. A binary image is considered as a set of points (white region) containing in some universal set (background, or black region). The morphological structuring element is a second set of points required for the transformation. The shape of this structuring element defines the results produced by the morphological filter.

3. Materials and Methods

- The method of work consists of a number of steps, which are as follows:
- ● Inserting images, which are four CT scan images of the lung.
- ● Pre-processing of the entered images by taking a histogram of images and then performing a process contrast adjustment Based on the histogram, CT scan images of the lung for areas of infection with the Covid-19 virus.
- ● Applying one of the grouping methods, the Fuzzy c-mean method, to images after pre-processing.
- ● Calculate the surfaces area of the extracted injury areas.
- The steps of this work can be summarized in the following block diagram:

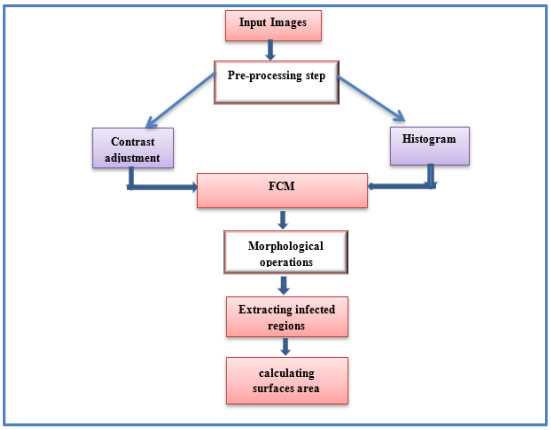

Figure (2): Block diagram of the work.

The experimental dataset are four lung CT scan images, acquired from Dr.Zaid Al-Khafaji Center, Ashur center and Fadak center with sizes of (252x331), (252x331), (252x331) and (252x331) pixels respectively. Figure (3) illustrates these input images.

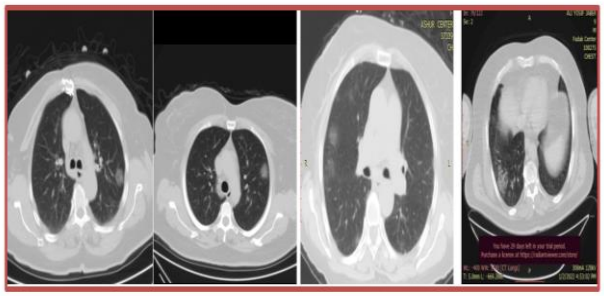

Figure (3): Input CT scan lung images.

4. Results

The results are presented as follows:

I. pre-processing

This stage of the work involved two steps which are:

Images Histogram plotting

The histogram of the hole lung images and the histogram of the corresponding extracting samples from the infected regions where plotted as shown in Figure (4) in order to estimate the gray level range of the infected region that enable as to determine the proper values utilized in contrast adjustment step.

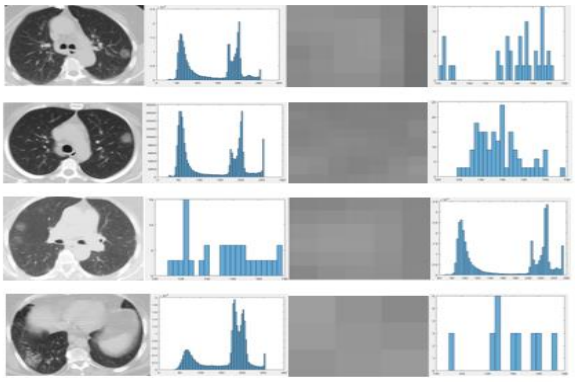

Figure (4): Histogram of the input whole lung images in (a) and the histogram of the corresponding samples of infected regions in(b).

Figure (4) shows the histogram plot of the whole lung and of samples from the infected areas within the lung. The first and second columns are the images and their histogram, while the third and fourth columns are the samples of the infected regions and their histograms.

Contrast adjustment

Many different interest ranges tested to highlight the infected regions contrast, adjusting was applied depending on the histogram information and Figure (5) Displays the results of this step.

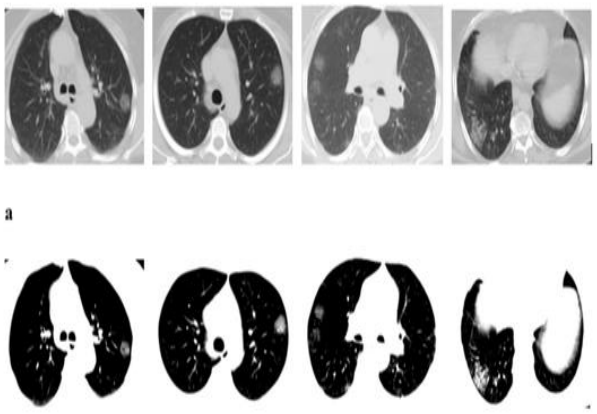

Figure (5): Results of contrast adjustment for CT-scan images input image. (a) input images,

(b) the corresponding contrast adjusted images

The next step is to extract the only whole lung region by applying k-means algorithm with two clusters to determine the surface area of the lung. Figure (6) presents the results of this stage.

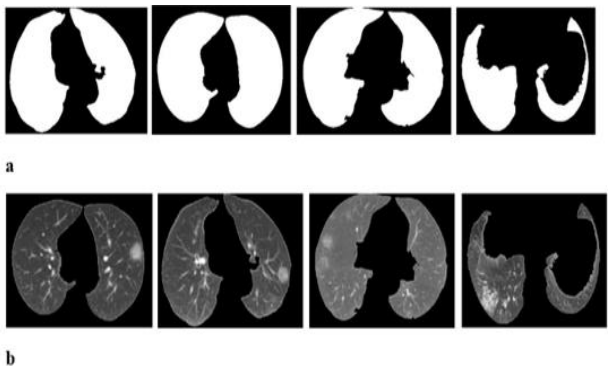

Figure (6): results of implementing k-means algorithm to extract the only whole lungs. (a) the class of lung regions,

(b)the corresponding gray level images

II. Detecting covid-19 lung infected regions

In this stage of the work, FCM algorithm was implemented to isolate and extract the lung infected regions.

FCM clustering algorithm

Clustering FCM was applied with different number of clusters (3, 4 and 5) to segment the adopted images. Figures (7-9) present the results of this stage of the work.

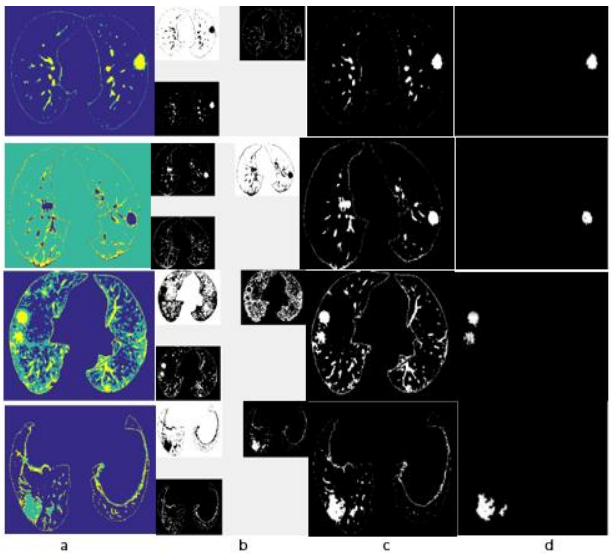

Figure (7): Results of implementing FCM Clustering with three clusters. (a) segmented image, (b) the discret clusters, (c) the cluster of the infected region and (d) the final extracted infected region.

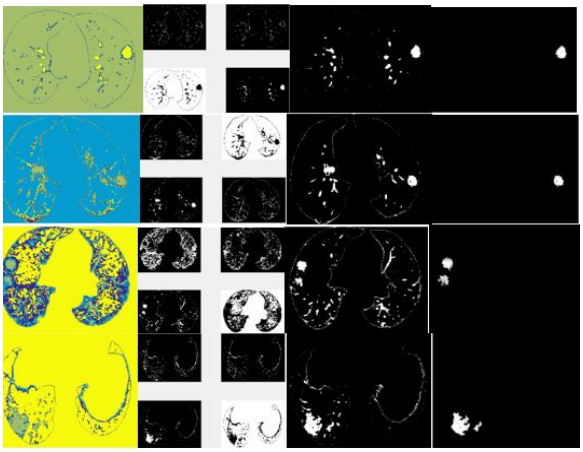

Figure (8): Results of implementing FCM Clustering with four clusters. (a) segmented image, (b) the discrete clusters, (c) the cluster of the infected region and (d) the final extracted infected region.

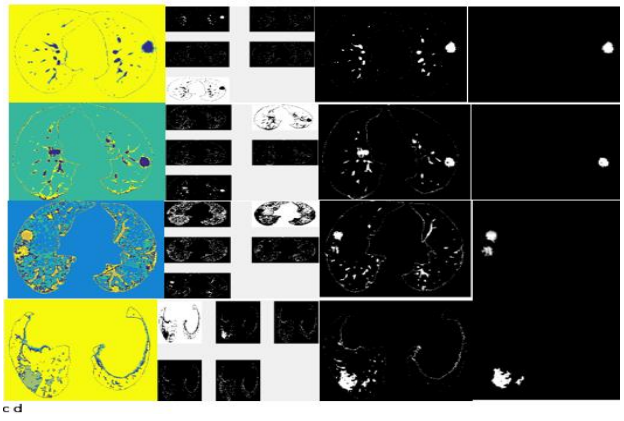

Figure (9): Results of implementing FCM Clustering with five clusters. (a) segmented image, (b) the discrete clusters, (c) the cluster of the infected region and (d) the final extracted infected region.

In Figures (7- 9), the first colume represents the clustered images, the second one shows the discrete clusters, third one shows the segment that infected region belong to, and last column presents the refined extracted infected region after applying many morphological operations. The results showed adquate extraction of the infected regions according to the radiologist consultation, and according to her, the number three of clusters is the more approporate cluster number that adqaue for extracting covid-19 infected regions.

The extracted surface areas of covid-19 infected regions were calculated and presented in Table (1).

Table 1: The values of the calculated extracted surface areas infected regions for FCM.

Surface Area (pixels)			
Images	Number of clusters		
	three	four	five
Image1	901	838	835
Image2	832	748	785
Image3	1445	1083	981
Image4	2525	2343	2318

Table 2: Precent relative extracted infected region with respect to the whole infected lungs.

Precent Relative of the infected regions %			
Images	Number of clusters		
	three	four	five
Image1	0.015	0.014	0.014
Image2	0.015	0.0137	0.016
Image3	0.018	0.0135	0.012
Image4	0.056	0.052	0.051

Radiologists Delineation

The manual delineation of the abnormal regions by the radiologist is illustrated in Figure (10) for lung CT scan

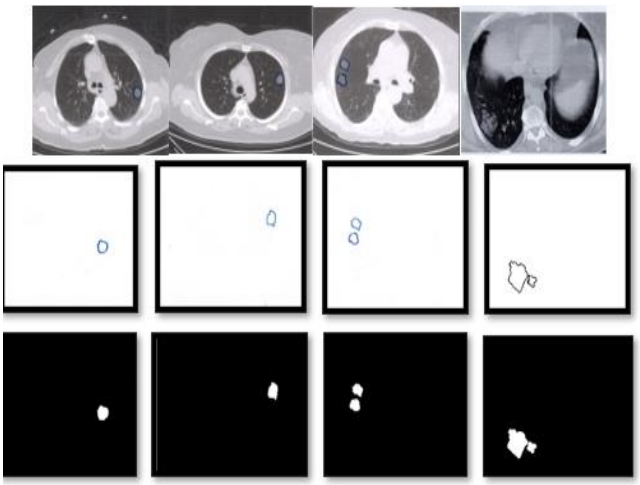

Figure (10): Identification of abnormal areas by a radiologist on computed tomography images of the lung.

9- Accuracy

The accuracy of the implemented method was calculated depending on the number of images that processed correctly and the whole number of the input processed image, and it found equals to 96% for FCM.

Accuracy was determined based on the number of images correctly segmented into the total number of processed images utilizing the method used and the result multiplied by 100.

5. 10- Conclusions

In this work, Fuzzy c-mean method was implemented with different number of clusters to detect and isolate affected areas in CT scan images of patients infected with COVID-19, with the help of contrast control and morphological processes such as widening and opening. The appropriate number of clusters to isolate and extract the affected area was (3) by presenting the results to the radiologist. The results confirmed the quality performance of the implemented method for isolating and extracting areas affected by Covid-19, with a high accuracy of 96%.

6. Acknowledgement

The authors would like to express their thanks to Dr. Ruaa Falah al- Jubawi worked in Zaid Al-Khafaji Clinic Center for her consultation and reviewing the quality of the extracted infected regions. As well as, authors are grateful to Dr. Zaid Hadi Al-Khafaji Center, Ashur Clinic and Fadak Clinic for providing the Covid-19 images.

References

Talevi, Dalila, et al. "Mental health outcomes of the CoViD-19 pandemic." Rivista di psichiatria 55.3 (2020).

Lone, Shabir Ahmad, and Aijaz Ahmad. "COVID-19 pandemic–an African perspective." Emerging microbes & infections 9.1 (2020): 1300-1308.

Pokhrel, Sumitra, and Roshan Chhetri. "A literature review on impact of COVID-19 pandemic on teaching and learning." Higher Education for the Future 8.1 (2021).

Daniel, Sir John. "Education and the COVID-19 pandemic." Prospects 49.1 (2020)

Mercer, Tim R., and Marc Salit. "Testing at scale during the COVID-19 pandemic." Nature Reviews Genetics 22.7 (2021): 415-426.

Del Rio, Carlos, Saad B. Omer, and Preeti N. Malani. "Winter of Omicron—the evolving COVID-19 pandemic." Jama 327.4 (2022): 319-320.

Pizarro-Ortega, Carlos Ivan, et al. "Degradation of plastics associated with the COVID-19 pandemic." Marine Pollution Bulletin (2022): 113474.

Wunsch, Kathrin, Korbinian Kienberger, and Claudia Niessner. "Changes in Physical Activity Patterns Due to the COVID-19 Pandemic: A Systematic Review and Meta-Analysis." International journal of environmental research and public health 19.4 (2022): 2250.

Bernheim, A., Mei, X., Huang, M., Yang, Y., Fayad, Z. A., Zhang, N., ... & Chung, M. Chest CT findings in coronavirus disease-19 (COVID-19 (2020).

Yang, X., He, X., Zhao, J., Zhang, Y., Zhang, S., & Xie, P. (2020). COVID-CT-dataset: a CT scan dataset about COVID-19 (2020).

Nashid Alam and Mohammed J. Islam, " Pectoral Muscle Elimination on Mammogram Using KMeans Clustering Approach", Department of Computing Science and Engineering Shahjalal University of Science and Technology, Sylhet, Bangladesh, Vol. 4, No. 1, PP. 11-21, (2014).

M.Punitha and K.Perumal, "Hybrid Segmentation with Canny Edge and K Means Clustering To Extract the Mammogram Tumor", International Journal of Computer Science Trends and Technology (IJCST), Vol. 6, Issue 4, PP.11-16, (2018). 12- S.P.

Meharunnisa, M. Ravishanakr and K. Suresh, "Improvement in Detection Accuracy of Digital Mammogram Using Point Transform and Data Mining Technique", Department of Electronics and Instrumentation Engineering, Visvesvaraya Technological University, India, Vol. 9, Issue 1, PP.1838-1843, (2018).

Rabab Saadoon Abdoon and Ola Saad Khudair, "Employing Histogram Equalization Enhancement Technique to Segment Different Medical Images of Three Organs." Journal of Advanced Microscopy Research, Vol. 13, No. 4, PP. 494-502, (2018).

M. Loey, G. Manogaran, and N. E. M. Khalifa, —A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images, ‖ Neural Comput. Appl., pp. 1–13, (2020).

J. F. Hernández Santa Cruz, —An ensemble approach for multi-stage transfer learning models for COVID-19 detection from chest CT scans, ‖ Intell. Med., vol. 5, no. June 2020, p. 100027, (2021).

L. Li et al., —Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT., ‖ Radiology, p. 200905, 2020, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/32191588.

Dunn JC. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. Journal of Cybernetics (1973).

Bezdek JC. Pattern recognition with fuzzy objective function algorithms. Springer Science & Business Media (2013).

Gonzalez R. C. and Woods R. E., "Digital Image Processing", Prentice Hall, (2002).

B. Fittes, "Gray-level transformations for interactive image enhancement. MS Thesis. Final Technical Report," (1975).

R. C. Gonzalez and R. E. Woods, "Digital image processing," ed: Prentice Hall New Jersey, (2002).

Bezdek JC. Pattern recognition with fuzzy objective function algorithms. Springer Science & Business Media (2013).

Kapur T, Grimson WE, Wells III WM, Kikinis R. Segmentation of brain tissue from magnetic resonance images. Medical image analysis. 1996 Jun 1;1(2):109-27.

Bezdek JC, Hall LO, Clarke L. Review of MR image segmentation techniques using pattern recognition. Medical physics. 1993 Jul;20(4):1033-48.

Klingler JW, Vaughan CL, Fraker TD, Andrews LT. Segmentation of echocardiographic images using mathematical morphology. IEEE Transactions on Biomedical Engineering. 1988 Nov;35(11):925-34.