Development of a Medical Image Segmentation Algorithm based on Fuzzy C-Means Clustering

EL FAHSSI KHALID[1], OUNASSER SAIDA[2], MOHAMED TAJ BENNANI[3], ABENAOU ABDENBI[2]

[1]Department Computer Science, LISAC,
Faculty of Science Dhar El Mahraz, Sidi Mohamed Ben Abdellah University of Fez,
MOROCCO

[2]Applied Mathematics and Intelligent Systems Engineering (MAISI),
Ibn Zohr University, ENSA-Agadir,
MOROCCO

[3]LPAIS Laboratory,
Computing Science Department, Faculty of Sciences Dhar El Mahraz Fez,
Sidi Mohamed Ben Abdellah University,
Fez,
MOROCCO

Abstract: - Breast mass segmentation in mammography plays a very important role in computer-aided diagnosis (CAD) systems. In this article, we propose a mammography image segmentation method based on a combined approach. The fuzzy clustering method and thresholding segmentation. Subsequently, we use the wavelet transform and the Canny filter for edge detection.

Key-Words: - Images, Processing, Breast cancer, Segmentation, Mammography, CAD.

Received: July 29, 2023. Revised: October 28, 2023. Accepted: December 15, 2023. Published: December 31, 2023.

1 Introduction

Modern methods of medical diagnosis and biomedical research largely rely on the analysis of images obtained through various technical means (optical and electron microscopes, X-ray and thermographic devices, tomography, etc.). However, effectively addressing diagnostic and scientific challenges when using visual information necessitates knowledge of specific methods for image acquisition, registration, digital processing, and analysis. This becomes particularly evident when utilizing new types of information systems that solve the problem of extracting hidden information for diagnosis (e.g., CT scanners, laser confocal microscopes, ultrasound diagnostic devices, etc.).

Magnetic Resonance Imaging (MRI) is a powerful method in medical diagnosis as it enables the extraction of valuable information about the patient's internal organs and tissues. MRI is widely used to monitor disease dynamics such as breast cancer, Alzheimer's disease, brain tumors, and more. Visual inspection of these images allows specialists to detect the emergence of certain anomalies. However, the massive number of medical images stored in the database makes visual analysis challenging, and in some cases, it may not lead to a better diagnostic outcome. Indeed, diagnostic tools and digital image processing methods significantly influence the diagnostic result.

One of the current trends in the development of medical informatics is digital image processing, including image quality enhancement, restoration of damaged images, and recognition of individual elements. Diagnosis and identification of pathological processes are among the most important tasks in the processing and analysis of medical images. Early diagnosis of the mentioned diseases can reduce mortality rates among patients.
To automate the process of MRI image analysis and improve diagnostic outcomes, it is necessary to develop new algorithms capable of addressing problems such as image classification, contour detection, and image segmentation.

2 Segmentation Methods

Numerous studies have been conducted in the field of medical image processing, [1], [2], such as the utilization of Magnetic Resonance Imaging [3], the extraction of white brain tissue regions and single-cell detection, [4] and topological visualization of human brain proliferation via MRI, [5]. However, there is no single method that would provide an acceptable analysis for any image. For the diagnosis of brain tumors, for example, many researchers employ a wide range of techniques based on MRI image segmentation. Regardless of the approach used, the problem of MRI image segmentation remains one of the fundamental challenges in digital image analysis and processing.

We can categorize segmentation methods into three main classes:

Region-based segmentation [6], edge-based segmentation [7], and pixel-based classification or thresholding segmentation, [8].

Methods within the first class [9], [10], [11], [12], search for sets of pixels that exhibit a certain degree of similarity. These techniques reduce operator involvement by automating certain aspects of low-level operations such as threshold selection, histogram analysis, classification, etc. Generally, these methods only use local information for each pixel and do not incorporate information about the shape of objects and their boundaries. Edge-based methods rely on the evolution of a curve based on internal and external forces, such as image gradient, to delineate the boundary of the object structure under analysis or pathology.

Pixel thresholding methods are more frequently employed for image binarization, [13], [14], [15]. These methods are straightforward, and require minimal computational costs, but are effective only when all objects and backgrounds in the images are clearly distinguished in terms of color or grayscale. Hence, it is necessary to control the segmentation process and adjust the results interactively.

In this context, in the article [16], a segmentation of medical images is proposed based on morphological operators in combination with threshold value selection. In [17], morphological operations were also used in combination with threshold and division-based segmentation. Furthermore, fuzzy methods are often employed for image segmentation, such as Fuzzy C-Means clustering (FCM), [18], [19].

In [17], for brain tumor extraction, an approach based on morphological operations with threshold-based segmentation and watershed line (LPE) methods was utilized. Additionally, the Fuzzy C-Means method, [18], [19], is widely used in medical image segmentation. This method aims to separate objects from each other and the background in the image by extracting contours or segmenting them into homogeneous regions.

In [20], to detect tumor boundaries in MRI images for various cases of brain tumors, a hybrid approach was proposed, combining the watershed method and the Canny edge detection method. One of these methods for representing medical image contours uses color encoding, [21].

3 Thresholding Methods

This pertains to a fundamental method in image segmentation, [22]. The general principle of thresholding involves finding an appropriate threshold value and then classifying all pixels in the image based on the value of their grayscale levels compared to this threshold to separate the regions of interest from the image background. Some threshold determination methods are based on parameters other than grayscale level, such as entropy, [23] or Tsallis entropy, [24]. For example, [25], formulated the image thresholding problem as an iterative discriminant analysis problem that allows for the selection of an optimal threshold value. The criterion used for threshold selection is based on maximizing a statistical measure of separation between classes. In all cases, the threshold obtained through the methods mentioned above is ultimately used for pixel classification based on their grayscale levels.

In general, thresholding methods can be classified into two categories:

• Global thresholding methods: These methods are widely used in the segmentation of mammographic images, [26], to detect tumor areas or microcalcifications. The principle of these methods is to determine the threshold value using the overall information contained in the image. This information is often presented in the form of a histogram of grayscale levels in the image, [27], [28]. Despite their widespread use, global thresholding is not very effective in accurately identifying regions of interest. In reality, mammographic images represent the projection of a 3D scene onto a 2D observation space. This projection results in significant overlaps of regions that make up the breast tissue, [29], which limits the effectiveness of these methods.

• Local thresholding methods: These methods aim to locally refine the threshold value to better identify regions of interest. The threshold value is determined by considering only the information contained in the local neighborhood of each pixel, [29]. These methods have often demonstrated better detection efficiency compared to global thresholding methods. It's worth noting that local thresholding methods have not only been used for image segmentation but have also been utilized as a preprocessing step for other algorithms, such as those based on Markov fields, [30].

4 The Proposed Method

In [31], the approach used has led to improved segmentation compared to classical methods. However, the complexity of the method lies in the selection of the initial contour point.

The major disadvantage of the level set method is that it requires considerable thought to construct appropriate speeds to progress the level set function. Very high computational complexity and often very slow convergence.

Among the disadvantages of the approach is the lack of a reset term. In reality, the reset is a very important step for the evolution of the contour, on the one hand. On the other hand, the approach uses an algorithm with very high computational complexity and often very slow convergence, which limits its use in real-time applications. Indeed, this algorithm requires updating and calculating the function for all points in the image and not just for the zero level curve.

In this article, we propose a segmentation approach based on the enhancement of the Otsu thresholding and fuzzy C-means clustering, as illustrated in Figure 1. The proposed method has been applied for breast tumor detection.

For contour detection, various algorithms can be used, such as Roberts, Prewitt, and Sobel, and more complex ones like wavelet and Canny. In this article, we used the Canny operator for contour detection as it is considered to be a more efficient tool for contour extraction.

After segmenting the image into a series of homogeneous classes using the proposed segmentation method, we apply the Canny operator. The steps of this method can be described as follows. In the first step, a Gaussian filter is used to smooth the original images. Additionally, for each pixel in the image, we calculate the magnitude and direction of the gradient. The next step involves selecting edge pixels (extreme pixels). A pixel is considered an edge pixel if its gradient value is greater than that of its two neighbors in the direction of the gradient.

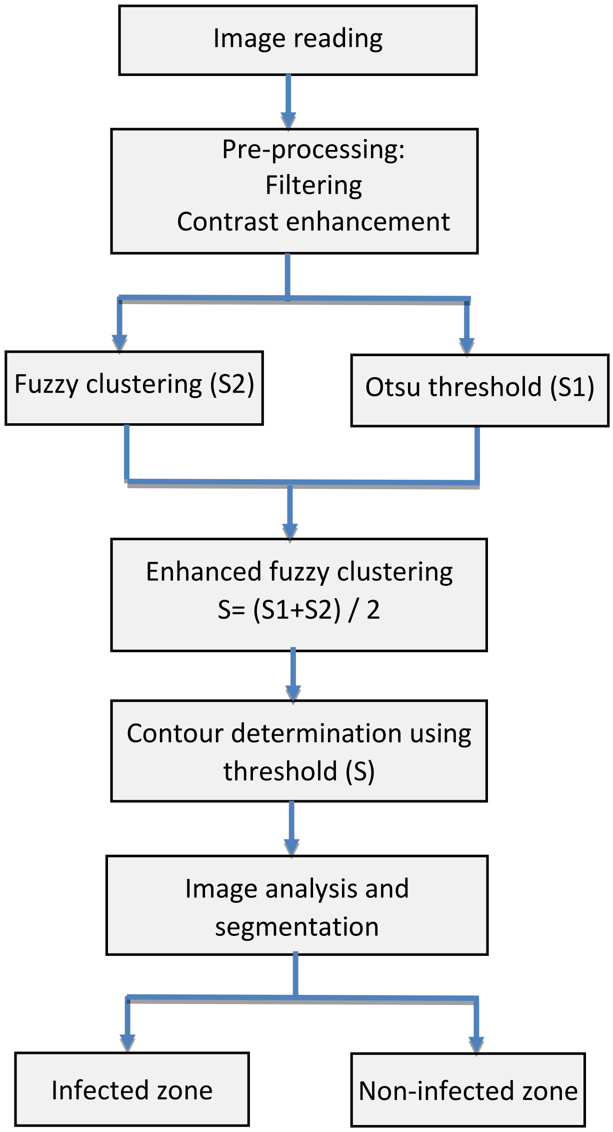

Fig. 1: Functional Diagram of the Proposed Method

However, the traditional Canny operator only extracts gradients in the x and y directions within 2x2 neighborhoods. This method loses some important contour information in the slope direction (45° and 135°). To address this issue, we use a method that calculates information about the gradient magnitude and direction to determine the gradient module of the pixel. This method takes into account both localization accuracy and noise immunity. For each pixel (i, j) in an image I, the partial derivatives with respect to x and y for 45° and 135° are as follows:

$$P_x(i,) = (i+1,) - (i-1,),$$
$$P_y(i,) = (i,+1) - (i,-1),$$
$$P_{45^\circ}(i,j) = I(i-1,j+1) - (i-1,j-1),$$
$$P_{135^\circ}(i,) = (i-1,+1) - (i+1,-1).$$

The gradient magnitude is:

$$G(i,j) = \sqrt{P_x^2(i,j) + P_y^2(i,j) + P_{45}^2(i,j) + P_{135}^2(i,j)} \quad (1)$$

Gradient direction:

$$\theta(i,j) = arctan\frac{P_x(i,j)}{P_y(i,j)} \quad (2)$$

5 Informative Features Extraction

The certainty of object classification in the segmented image is strongly linked to the method used for extracting informative features. In this article, we proposed a classification algorithm based on the application of Discrete Wavelet Transform (DWT) and the Gray-Level Co-occurrence Matrix (GLCM). The coefficients obtained through discrete wavelet transform are analyzed to extract some of the most informative statistical features, such as energy, contrast, correlation, entropy, homogeneity, and others.

Finally, the procedure for diagnosing mammography images involves using a machine learning algorithm for the classification of tumor-infected areas (benign or malignant). The segmentation method proposed in this article has yielded good results, making it easier to choose a simple and effective learning algorithm, such as the Support Vector Machine (SVM) method.

6 Results and Discussion

In this experimental study, we used the database (mini-MIAS), [32], for mammography analysis (breast cancer).In this work, we formed two classes of data with 100 images each of size 256 x 256. One class consists of healthy breast MRI images. While the second consists of MRI images with abnormalities. The two classes are used for learning the diagnostic system. An additional breast cancer dataset consisting of 50 images is used to test the system. The results of the experimental study of breast tumor segmentation and detection are shown in Figure 2.

We can also notice that the brightness and contrast of the image vary from one image to another (The upper part of the figure from left to right). The first column (from left) represents the original images, and the second column represents the image results after the preprocessing step.The third and last column illustrates the result after masking and processing (final edge images). Comparing the proposed method with the methods [28], [29] for identifying and segmenting a breast tumor presented, one can notice the developed algorithms capable of segmenting tumors and normal regions better than other methods.Indeed , the results obtained in [28], [29] include a non-tumor region in the segmented image, which represents an ambiguity for the diagnosis (without outline of the tumors). The proposed method successfully detects the tumor region (as well as the normal breast region in the original image) as shown in Figure 2.

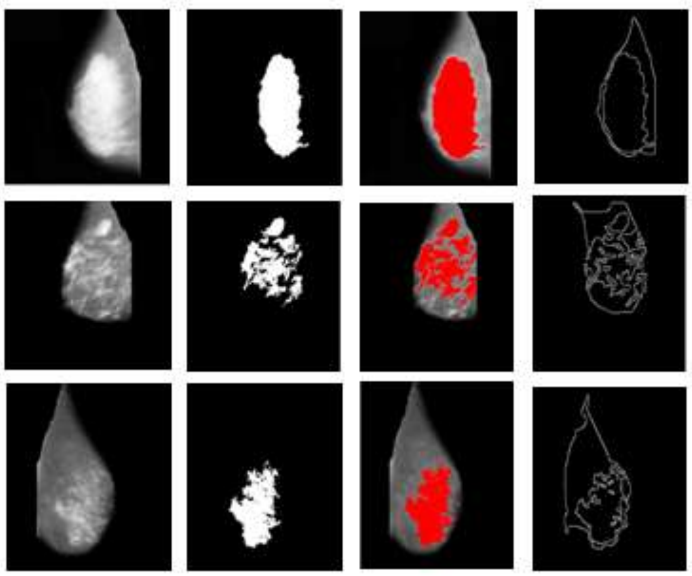

Fig. 2: Experimental result of the proposed method for breast tumor detection

7 Conclusion

The experimental results obtained show that the effectiveness of the proposed method of segmentation by groups. The effectiveness of the approach lies in the combination of two approaches, the Otsu thresholding method and fuzzy clustering.The application of the discrete wavelet transform (DWT) and the GLCM co-occurrence matrix ensured an improvement in the analysis quality during the extraction of image features.

In comparison with other segmentation methods, the new threshold used in the proposed approach significantly increased the precision and quality of medical image segmentation. Which will subsequently contribute to improving the diagnostic results of MRI images while detecting breast tumors.

References:

[1] Lo SCB, Chan H-P, Lin JS, Li H, Freedman M, Mun SK. Artificial convolution neural network for medical image pattern recognition, *Neural networks,* 1995. T. 8., №. 7-8, C. 1201-1214.

[2] Guotai Wang, Wenqi Li, Sebastien Ourselin, Tom Vercauteren. Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks, *International MICCAI brain lesion workshop*, Springer, Cham, 2017, C. 178-190.

[3] Xiaojian Kang, Timothy J Herron, And U Turken, David L Woods. Diffusion properties of cortical and pericortical tissue: regional variations, reliability, and methodological issues, *Magnetic Resonance Imaging*, 2012. – T. 30. – №. 8. – C. 1111-1122.

[4] Mariana Leite, David Gobbi, Marina Salluzi, Richard Frayne, Roberto Lotufo, Letícia Rittner. 3D texture-based classification applied on brain white matter lesions on MR images, Medical Imaging 2016: Computer-Aided Diagnosis. – *International Society for Optics and Photonics*, 2016. – T. 9785. – C. 97852N.

[5] Schultz T., Theisel H., Seidel H. P. Topological visualization of brain diffusion MRI data, *IEEE Transactions on Visualization and Computer Graphics,* 2007. T. 13., №. 6. – C. 1496-1503.

[6] J. Singh, H Kaur, Plant Disease Detection Based on Region-Based Segmentation and KNN Classifier, *International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB).*

[7] Bahadir Karasulu, Serdar Korukoglu, A simulated annealing-based optimal threshold determining method in edge-based segmentation of grayscale images, *Applied Soft Computing Journal*, Vol. 11, Issue 2, p. 2246-2259, 2011

[8] Natthan Singh, Shivani Goyal, Determination and Segmentation of Brain Tumor Using Threshold Segmentation with Morphological Operations, *Soft Computing: Theories and Applications*, 2018, Vol. 584.

[9] Weibei, Dou, Su, Ruan, Yanping, Chen, Daniel, Bloyet, and Jean-Marc, Constans. (2007) “A framework of fuzzy information fusion for segmentation of brain tumor tissues on MR images.” *Image and Vision Computing* 25: 164-171.

[10] Gul, Moonis, Jianguo, Liu, Jayaram K., Udupa, and David B., Hackney. (2002) “Estimation of tumor volume with fuzzy-connectedness segmentation of MR images”. *American Journal of Neuroradiology*, 23: 352-363.

[11] Meritxell, Bach Cuadra, Claudio, Pollo, Anton, Bardera, Olivier, Cuisenaire Jean-Guy Villemure, and Jean-Philippe, Thiran (2004) “Atlas-based segmentation of pathological MR brain images using a model of lesion growth.” *IEEE Transactions on Medical Imaging* 23 (10): 1301-1313.

[12] Jason, J. Corso, Eitan, Sharon, and Alan, Yuille. (2006) “Multilevel segmentation and integrated Bayesian model classification with an application to brain tumor segmentation”, in *MICCAI2006, Copenhagen, Denmark, Lecture Notes in Computer Science*, 4191, Springer, Berlin: 790-798.

[13] Fatma Taher, Naoufel Werghi and Hussain Al-Ahmad, “A thresholding approach for detection of sputum cell for lung cancer early diagnosis” Vol. 2012, Issue 600 *CP2012 IET Conference on Image Processing*, July 2012.

[14] Gul, Moonis, Jianguo, Liu, Jayaram K., Udupa, and David B., Hackney. (2002) “Simple and efficient method of low-contrast grayscale image binarization”. Vol. 9972 *LNCS, p 142 – 150, International Conference on Computer Vision and Graphics, ICCVG* Warsaw, September 2016.

[15] An Algorithm for fast adaptive image binarization with applications in radiotherapy imaging, *IEEE Transactions on Medical*

Imaging, Vol. 22, Issue 1, pp.22-28, January 2003.

[16] Priya and Vivek Singh Verma, New Morphological Technique for Medical Image Segmentation, *3rd IEEE International Conference on "Computational Intelligence and Communication Technology" (IEEE-CICT 2017).*

[17] DiyaChudasama, Tanvi Patel and Shubham Joshi "Image Segmentation using Morphological Operations", *International Journal of Computer Applications* (0975 – 8887), Vol. 117 – No. 18, May 2015.

[18] Chuang, K.-S., Tzeng, H.-L., Chen, S., Wu, J., & Chen, T.-J. Fuzzy c-means clustering with spatial information for image segmentation. *Computerized Medical Imaging and Graphics*, 30(1), 9–15. 2006.

[19] Huiyu Zhou, Gerald Schaefer and Chunmei Shi, Fuzzy C-Means Techniques for Medical Image Segmentation. Fuzzy Systems in *Bioinformatics and Computational Biology*, pp. 257–271, 2009.

[20] Norhafiza Hamzah, Mohd Sapaweh Asli and Rechard Lee, Skin Cancer Image Detection using Watershed Marker-Controlled and Canny Edge Detection Techniques. *Transactions on Science and Technology*, Vol. 5, No. 1, 1 - 4, 2018.

[21] Xianhua Zeng, Shiyue Tong, Yuzhe Lu, Liming Xu and Zhiwei Huang, Adaptive Medical Image Deep Color Perception Algorithm, *IEEE Access*, Vol. 8, 2020.

[22] R. C. Gonzalez and R. E. Woods: *Digital Image Processing.* 2ed, Prentice Hall, 2002.

[23] Lo S. C. B. et al. Artificial convolution neural network for medical image pattern recognition, *Neural networks*, 1995 Т. 8 №. 7-8 С. 1201-1214.

[24] J. Mohanalin, P. K. Kalra and N. Kumar: Tsallis Entropy Based Contrast Enhancement of Micro calcifications, *International Conference on Signal Acquisition and Processing, ICSAP*, pp.3-7, 2009.

[25] M. A. Duarte, A. V. Alvarenga, C. M. Azevedo, A. F. Infantosi and W. C. Pereira: Automatic microcalcifications segmentation procedure based on Otsu's method and morphological filters, *Health Care Exchanges (PAHCE)*, pp. 102-106, 2011.

[26] H. Cheng, X. Shi, R. Min, L. Hu, X. Cai, and H. Du: Approaches for automated detection and classification of masses in mammograms, *Elsevier*, vol. 39, no.4, pp. 646-668, 2006.

[27] G. M. Victor, M.G. Daniel, and R. Juan: Iterative method for automatic detection of masses in digital mammograms for computer-aided diagnosis, *Proc. SPIE*, vol. 3661, pp. 1086-1093, 1999.

[28] T. Matsubara, T. Fujita, H. Kasai, S. Goto, M. Tani, Y. Hara, and T.T. Endo: Development of new schemes for detection and analysis of mammographic masses, *IEEE. Comput. Soc*, pp. 63-66, 1997.

[29] L. Li, W. Qian, L. P. Clarke, R. A. Clark, and J. A. Thomas: Improving mass detection by adaptive and multiscale processing in digitized mammograms, *Proc. SPIE*, vol. 3661, pp. 490-498, 1999.

[30] H. D. Li, M. Kallergi, L. P. Clarke, V. K. Jain, and R. A. Clark: Markov random field for tumor detection in digital mammography, *IEEE. Transactions on Medical Imaging*, vol. 14, no. 3, pp. 565-576, 1995.

[31] Liu, J., Wei, X. & Li, L. MR image segmentation based on level set method. *Multimed. Tools Appl*. 79, 11487–11502, 2020.

[32] The mini-MIAS database of mammograms. http://peipa.essex.ac.uk/info/mias.html (Access Date: September 7, 2023).

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare.