Segmentation of Mammograms Using a Novel Intuitionistic Possibilistic Fuzzy C-Mean Clustering Algorithm

Chiranji Lal Chowdhary and D.P. Acharjya

Abstract There is a partitioning of a data set X into c-clusters in clustering analysis. In 1984, fuzzy c-mean clustering was proposed. Later, fuzzy c-mean was used for the segmentation of medical images. Many researchers work to improve the fuzzy c-mean models. In our paper, we proposed a novel intuitionistic possibilistic fuzzy c-mean algorithm. Possibilistic fuzzy c-mean and intuitionistic fuzzy c-mean are hybridized to overcome the problems of fuzzy c-mean. This proposed clustering approach holds the positive points of possibilistic fuzzy c-mean that will overcome the coincident cluster problem, reduces the noise and brings less sensitivity to an outlier. Another approach of intuitionistic fuzzy c-mean improves the basics of fuzzy c-mean by using intuitionistic fuzzy sets. Our proposed intuitionistic possibilistic fuzzy c-mean technique has been applied to the clustering of the mammogram images for breast cancer detector of abnormal images. The experiments result in high accuracy with clustering and breast cancer detection.

Keywords Intuitionistic fuzzy c-mean · Possibilistic c-mean · Membership degree · Non-membership degree · Hesitation degree

1 Introduction

Segmentation arises as the basic technique towards image processing analysis because this is an important building block in it and also divides the image into various parts (regions) with homogeneous features. In the segmentation process, the image is partitioned into various non-overlapping and meaningful homogeneous regions. For common images, the segmentation is based on unsupervised clustering techniques but it becomes challenging in case of medical imaging because of poor

C.L. Chowdhary (✉)
School of Information Technology and Engineering, VIT University, Vellore, India
e-mail: c.l.chowdhary@gmail.com

D.P. Acharjya
School of Computing Science and Engineering, VIT University, Vellore, India

B.K. Panigrahi et al. (eds.), *Nature Inspired Computing*, Advances in Intelligent Systems and Computing 652, https://doi.org/10.1007/978-981-10-6747-1_9

contrast and noise. A number of techniques were offered by different researchers, and we may assume thresholding technique as the easiest over the other techniques. Thresholding became more difficult in the case of distributed illuminations. It is better to use local thresholding over global in such cases, and thresholds are found for each subdivision of image to get a local variation. Many authors have proposed clustering technology in decision-making, problem-solving and image segmentation to threshold an image. Otsu [1] has used thresholding techniques for a class variance to maximize the class reparability. In the last one decade, fuzzy approach is used in medical images due to the presence of uncertainty issues present in the form of vagueness, boundaries and variations in grey levels.

The main role of clustering is to separate unlabelled data into discrete sets. Clustering is a major technique of unsupervised learning. Many clustering techniques such as conventional *k*-mean, fuzzy *c*-mean, artificial neural network and genetic algorithms were used by different researchers. It is challenging to find an optimal threshold in thresholding technique. The *k*-mean clustering approach limits each data in a particular cluster which cannot be acceptable in all applications. Fuzzy *c*-mean is mostly used in image segmentation and assigns each pixel to unlabelled fuzzy clusters to keep each pixel in all clusters with different degrees of membership. So, the membership of fuzzy *c*-mean cannot represent the degrees of belonging of the data. Fuzzy *c*-mean was successfully incorporated into the segmentation of medical images to overcome the uncertainty and unknown noise available in them. The problems in the fuzzy *c*-mean membership do not represent the degree of belongings of the data. This problem was solved by Krishnapuram and Keller [2] by proposing possibilistic *c*-means. Each element of the kth column can be any numbers between 0 and 1, as long as the least one of them is positive. Possibilistic *c*-mean was useful to identify outliers (noise points). In 1997, some researchers [3] suggested possibilistic fuzzy *c*-mean algorithm to generate membership and typicality values where clustering unlabelled data. Later in 2005, [4] came with possibilistic fuzzy *c*-mean algorithm which produces membership and possibility simultaneously with useful point models or cluster centres for each centre. Possibilistic fuzzy *c*-mean was useful for fuzzy rule-based system identification.

Carvalho [5] proposed an adaptive and non-adaptive fuzzy *c*-mean for symbolic interval data. This does a fuzzy partition and model for every cluster by optimizing an adequacy criterion on proper Euclidian distance between vectors of intervals. Ji et al. [6] worked with an adaptive method to find the weights of local spatial factors in the objective function for local spatial continuity. They applied its magnetic resonance imaging images and found significant success with the proposed possibilistic fuzzy *c*-mean method. This method was found more robust and efficient for many levels of noise and was able to achieve higher accuracy in the context of brain magnetic resonance imaging image segmentation. Fuzzy *c*-mean cannot achieve accurate results due to noisy conditions. So, [7] proposed exponential fuzzy *c*-mean to enhance membership issues and result in more meaningful membership degree over fuzzy *c*-mean. Wahid et al. [8] proposed a genetic algorithm-based clustering technique which exploits multiple views to generate different clustering solutions and then selects a combination of clusters to form a final clustering solution. With the use of research work on theory and application of intuitionistic fuzzy sets by Attansov [9], Chaira [10] proposed intuitionistic fuzzy *c*-mean technique. This method is useful in clustering different regions of the medical images and helps to find abnormalities in image. In intuitionistic fuzzy sets, there are membership degree, non-membership degree and hesitation degree. Hesitation is an another uncertainty parameter.

In our proposed algorithm, possibilistic FCM was integrated with intuitionistic fuzzy *c*-mean. Traditional clustering methods are unable to handle the cases of noisy data and outliers. So we have used possibilistic approach to overcome this condition for improving membership assignments. To improve the possibilistic *c*-mean algorithm, we have embedded intuitionistic fuzzy *c*-mean. Intuitionistic fuzzy *c*-mean algorithm on uncertainty issue and there hesitation degree is discussed while defining the membership function. The remaining part of the proposed paper is organized in the following sections: Sect. 2 describes the preliminaries carried out on the clustering approaches; Sect. 3 covers the construction of intuitionistic possibilistic fuzzy *c*-mean; Sect. 4 describes the result analysis; and finally, Sect. 5 concludes the paper.

2 Preliminaries of Clustering

Clustering approach is in concern with image segmentation issues and that will decide about that how the pixels of an image concern with each other in a proper way. A clustering algorithm will divide the set of pixels into clusters such that pixels contained by a given cluster claim high degree of similarity over the pixels belonged to different clusters. An extensive literature provides information that different clustering methods are applied in many research areas such as pattern recognition, data mining, big data, taxonomy, image processing and information retrieval. Clustering approaches found to be categorized into two methods as hierarchical and partitioning methods. Hierarchical methods are nested in sequence of partitions of the input data, and partitioning methods obtain a single partition of the input data in a fixed number of clusters. According to the traditional *k*-means clustering method [11], every data belong just to one cluster, but this cannot be convincing in few applications. Fuzzy *c*-mean (FCM) clustering is a fuzzy type of *k*-mean clustering [7, 12]. In FCM, there is a combination of fuzzy approach that is used to allow data to be added in all clusters with diverse degrees of membership.

The fuzzy clustering approach [7, 12] is partitioned into a set B in k clusters, and the set is having N members as $B = \{b_1, b_2, b_3, \ldots b_N\}$. There may be a situation of uncertainty that the data b_l were consigned to many clusters with diverse degrees of membership u_{lm}. The belongingness of a data with a cluster is decided by paralleling its distance or dissimilarity d_{lm}^2 from the cluster centroid v_m. The measurements of distances are calculated with the help of Euclidean formula:

$$\mathrm{FCM} = \sum_{m=1}^{k} \sum_{l=1}^{N} u_{ij}^{p} d_{lm}^{2},\; p \in (1, \infty), \sum_{m=1}^{k} u_{lm} = 1. \tag{1}$$

The role of fuzzifier parameter (p) is about mechanizing the influence of the membership degree over the objective function. The value of membership degree and centroid is expressed in (2) and (3) correspondingly:

$$u_{lm} = \frac{1}{\sum_{q=1}^{k} \left(\frac{d_{lm}^{2}}{d_{lq}^{2}} \right)^{\frac{1}{p-1}}} \tag{2}$$

$$v_{m} = \frac{\sum_{l=1}^{N} u_{lm}^{p} x_{l}}{\sum_{l=1}^{N} u_{lm}^{p}} \tag{3}$$

FCM clustering results are better over k-mean results but more sensitive to noise. One limitation is that the cluster-wise addition of all the membership degrees for every data to one leads the abnormal points to be members of clusters. By integrating the possibilistic approach with the fuzzy c-mean, the drawback of FCM can be resolved and that approach can be given the name of possibilistic fuzzy clustering (PFCM). The main equation for possibilistic fuzzy c-mean approach can be written as:

$$\mathrm{PFCM} = \sum_{m=1}^{k} \sum_{l=1}^{N} u_{lm}^{p} d_{lm}^{p} + \sum_{m=1}^{k} \lambda_{m} \left(\sum_{l=1}^{N} 1 - u_{lm} \right). \tag{4}$$

The membership degree function and positive number are as follows in (5) and (6):

$$u_{lm} = \frac{1}{1 + \left(\frac{d_{lm}^{2}}{\lambda_{m}} \right)^{\frac{1}{p-1}}}, \tag{5}$$

$$\lambda_{m} = W \frac{\sum_{l=1}^{N} u_{lm}^{p} d_{lm}^{2}}{\sum_{l=1}^{N} u_{lm}^{p}}, \tag{6}$$

In (6), W is an amendable weight which is typically set to one.

Equation (3) of FCM is used to achieve an optimum solution to update in the centroid. In case if all clusters are coincident clusters, Eq. (4) may be really minimized. The membership degree from Eq. (5) relies heavily on the gap between the data and specific cluster without any consideration of other clusters.

Another improved fuzzy clustering is based on intuitionistic fuzzy clustering algorithm. The conventional fuzzy c-mean function is modified by the use of intuitionistic fuzzy sets. The cluster centres are modified so that we can integrate intuitionistic properties with fuzzy c-mean method. Atanassov [9] proposed intuitionistic fuzzy sets and said about the existence of hesitation degree. According to him, it cannot be always true that summation of degree of membership and degree of non-membership is being 1. There may be a possibility of hesitation degree, and a hesitation degree is defined as 1 minus the sum of membership and non-membership degrees. The hesitation degree is given as follows:

$$\pi_A = \text{hesitation\_degree} = 1 - (\text{membership\_degree} + \text{non\_membership\_degree}) \tag{7}$$

Hesitation degree is initially calculated using Eq. (7), and the intuitionistic fuzzy membership values are obtained as follows:

$$u_{lm}^{*} = u_{lm} + \pi_{lm} \tag{8}$$

where $u_{lm}^{*}(u_{lm})$ denotes the intuitionistic fuzzy membership of the mth data in lth class. Replacing Eq. (8) in (3), the modified cluster centre will be:

$$v_m^{*} = \frac{\sum_{l=1}^{N} u_{lm}^{*,p} x_l}{\sum_{l=1}^{N} u_{lm}^{*,p}} \tag{9}$$

Using Eq. (9), the cluster centre is updated and simultaneously the membership matrix is updated. At each iteration, the cluster centre and the membership matrix are updated and the algorithm stops when the updated membership matrix and the previous matrix. Thus, the criterion function in conventional FCM is modified using intuitionistic fuzzy sets.

3 Construction of Proposed Intuitionistic Possibilistic Fuzzy Clustering

Several medical image segmentation systems suggested by different authors have used the conventional k-mean clustering approach for tumour detection. But it has the limitations such as inadequate detection of tumour, predominantly in case of malignant cases. Some other researchers have used fuzzy c-mean approach which can detect malignant tumour mass more precisely over k-mean. These clustering methods are unable to handle the cases of noisy data and outliers. Possibilistic approach was used to overcome such condition for improving membership assignments. We have proposed an integrated intuitionistic fuzzy c-mean system to improve the possibilistic c-mean algorithm. Such medical image segmentation system will be called as intuitionistic possibilistic fuzzy c-mean (IPFCM) clustering system, and this system is taking the advantages of both approaches. Construction of the proposed system is in four paces: initiative pre-processing, main

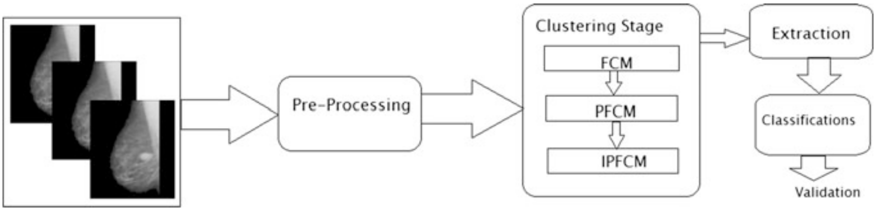

Fig. 1 Proposed intuitionistic possibilistic fuzzy clustering system

segmentation, i.e. clustering, statistical texture feature extraction and classification. The leading idea of our proposed work is based on segmentation part which does integration of possibilistic fuzzy *c*-mean with intuitionistic fuzzy *c*-mean and reduces the number of iterations to help in minimizing the execution time. The important portions of our IPFCM clustering method will be explained in the next paragraphs (Fig. 1).

Medical images are difficult to interpret, so the image is converted into the accessible form. This step involves removing unwanted parts which present in the background of the medical images. The main objective of this step is to improve the image quality by removing the unwanted areas. Disparate possibilistic clustering methods that try to minimize the membership degree of noisy data, the intuitionistic possibilistic fuzzy clustering (IPFCM) approach, assigns hesitation degree to them with membership and non-membership degree. Intuitionistic possibilistic fuzzy *c*-mean algorithm is applied to improving the detection duties of breast cancer detection system to assist a radiologist to some extent. After, medical images were enhanced using intuitionistic possibilistic fuzzy *c*-mean algorithm, choose intuitionistic fuzzy *c*-mean algorithm to separate between each cluster of pixels.

3.1 IPFCM Algorithm 1

Step 1. Initialization: Initialize B, k, d, u, v and the parameters used in the proposed algorithm.

Step 2. Evaluate the Eq. $\mathrm{PFCM} = \sum_{m=1}^{k} \sum_{l=1}^{k} u_{lm}^{P} d_{lm}^{P} + \sum_{m=1}^{k} \lambda_m \left(\sum_{l=1}^{N} 1 - u_{lm} \right)$ by Eq. (4).

Sub-Step 2 (a). Find $u_{lm} = \frac{1}{1 + (\frac{d_{lm}^2}{\lambda_m})^{\frac{1}{P-1}}}$ by Eq. (5).

Sub-Step 2(b). Find $\lambda_m = W \frac{\sum_{l=1}^{N} u_{lm}^{P} d_{lm}^{2}}{\sum_{l=1}^{N} u_{lm}^{P}}$ by Eq. (6).

Step 3. Hesitation degree is initially calculated using Eq. (7)

Step 4. Intuitionistic fuzzy membership value is obtained as: $u_{lm}^{*} = u_{lm} + \pi_{lm}$, where $u_{lm}^{*}(u_{lm})$ denotes the intuitionistic fuzzy membership of the mth data in lth class.

Step 5. Replace Eq. in Step (4) in Sub-Section Eq. 2(b), the modified cluster center is will be: $\lambda_m = W \frac{\sum_{l=1}^{N} u_{lm}^{*,P} d_{lm}^{2}}{\sum_{l=1}^{N} u_{lm}^{*,P}}$ the cluster center is updated and simultaneously the membership matrix is updated.

Step 6. Determine the termination of iteration. Check convergence criterion. If convergence has been reached, stop the iteration, otherwise, go to Step 2.

4 Result Analysis

We have proposed algorithm on mammogram images. On common conditions and breast cancer images, we have compared different clustering algorithms. These clustering algorithms such as fuzzy *c*-mean, possibilistic fuzzy *c*-mean and intuitionistic fuzzy c-mean algorithms are compared with intuitionistic possibilistic fuzzy *c*-mean algorithm.

The importance of clustering is its ability to detect the tumour regions. Our proposed algorithm is having good results of detecting tumours over the other existing clustering algorithms (Fig. 2).

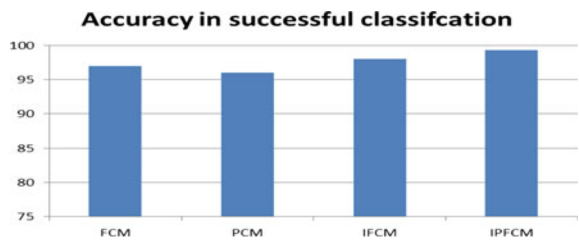

Fig. 2 Comparison between different clustering algorithms

5 Conclusions

Our proposed novel algorithm on intuitionistic possibilistic fuzzy clustering is hybridization of possibilistic fuzzy c-mean with intuitionistic fuzzy c-mean algorithm. We have tested our algorithm on breast cancer images (mammograms from MIAS dataset), and we have observed that efficiency of our approach is better over others.

References

1. Otsu, N.: A threshold selection method from gray level histograms. IEEE Trans. Syst. Man Cyber. **9**, 62–66 (1979)
2. Krishnapuram, R., Keller, J.M.: A possibilistic approach of clustering model. IEEE Trans. Fuzzy Syst. **1**, 98–110 (1993)
3. Pal, N.R., Pal, K., Bezdek, J. C.: A mixed c-means clustering model. Interrnational Conference Fuzzy Systems, Spain, pp. 11–21 (1997)
4. Pal, N.R., Pal, K., Keller, J. M., Bezdek, J.C.: A possibilistic fuzzy c-mean clustering algorithm. **13** (2005)
5. de Carvalho, F.D.A.: Fuzzy c-mean clustering for symbolic interval data. Pattern Recogn. Lett. (2006)
6. Ji, Z.X., Sun, Q.S., Xia, D.S.: A modified possibilistic fuzzy c-means clustering algorithm for bias MR images. Comput. Med. Imaging Graph. **35**, 383–397 (2011)
7. Treerattanapitak, K., Juruskulchai, C.: Outlier Detection with possibilistic exponential fuzzy clustering. 8th FSKD proceeding, pp. 453–457 (2011)
8. Wahid, A., Gao, X., Andreae, P.: Multi-view clustering of web documents using multi-objective genetic algorithm. IEEE Congress on Evolutionary Computation (2014)
9. Attanassov, K.T.: Intuitionistic fuzzy sets. Fuzzy Sets Syst. **20**, 87–96 (1986)
10. Chaira, T.: A novel intuitionistic fuzzy c-means clustering algorithm and its application to medical images. Appl. Soft. Comput. **11**, 1711–1717 (2011)
11. MacQueen, J.B.: Some methods for classification and analysis of multivariate observations. In: Proceedings of 5th Berkeley Symposium Mathematical Statistics and Probability. vol. 1, pp. 281–297 (1966)
12. Treerattanapitak, K., Juruskulchai, C.: Possibilistic exponential fuzzy clustering. J. Comput. Sci. Technol. **28**, 311–322 (2013)