Journal of Computer Science and Information Technology
June 2016, Vol. 4, No. 1, pp. 39-66
ISSN: 2334-2366 (Print), 2334-2374 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/jcsit.v4n1a3
URL: https://doi.org/10.15640/jcsit.v4n1a3

Entropy Guided Clustering Improvements and Statistical Rule-Based refinements for Bone Segmentation of X-Ray Images

Jamshid Tamouk[1] & Adnan Acan[2]

Abstract

During the recent decade, segmentation algorithms have been applied extensively to medical images in order to assist physician treating and diagnosing variety of diseases. Among all types of medical images, X-ray imaging is one of the frequently used imaging methods, especially for diagnosing bone diseases. Great numbers of segmentation methods have been applied on X-ray images aiming to segment bones with higher accuracy; however, it is hard to find a single method that is capable of segmenting all body parts with equal level of quality. Therefore, researchers put significant efforts on the combination different methods or improvement of existing methods using some techniques to achieve a desired segmentation outcome. In this research, improvement and refinement methods are proposed for three well-known clustering algorithms, namely the K-means, fuzzy c-means and spatial fuzzy c-means algorithms. The proposed methods include a novel entropy guided improvement strategy for clustering and a statistics based cellular rule engine for further refinement of pixel clusters. The improved algorithms are applied to X-ray images and experimental evaluations in comparison to well-known methods exhibit that significant improvements are achieved for all test cases.

Keywords: X-Ray Bone Segmentation; K-means; Fuzzy c-means; Spatial Fuzzy c-means; Entropy; Rule-based refinements.

I. Introduction

The process of extracting regions of interest from an image is known as image segmentation. It is mostly considered as the first step and a critical task of automatic image analysis.

[1] Computer Engineering Dept., Eastern Mediterranean University, Famagusta, Cyprus. jamshid.tamouk@cc.emu.edu.tr
[2] Computer Engineering Dept., Eastern Mediterranean University, Famagusta, Cyprus. adnan.acan@emu.edu.tr

In the automatic image segmentation, due to having no interaction with users, reaching highly accurate results is usually a hard problem. A great number of researchers have attempted to raise the accuracy of image segmentation and applications have been made in vast majority of areas such as analysis of satellite and medical images, face detection, and etc. In the past decade, segmentation has been applied extensively to medical images in order to assist physician in diagnosing variety of diseases, such as bone diseases, to survey the existence and exact location of fracture or other bone injuries.

X-ray images include two major regions, which are tissue (high/low density tissue) and bone (with density higher than that of tissues). In X-ray images, bone appears brighter than tissue; however bone segmentation is still a challenging task due to complexity of X-ray images, that is mainly caused by existence of noise and artifacts in these types of images and most importantly closeness of grayscale levels of some tissues and bones (lack of the contrast between regions) (Mahendran and Baboo 2011). Overlapping of the tissue and bone regions in this type of images is another problem to be considered. Due to these mentioned difficulties, most of the segmentation methods are unable to segment the bones with desired accuracy. Hence, the fundamental objective of this research is to detect/segment the bones over X-ray images as complete and accurate as possible. For this purpose, the weaknesses of three well-known methods are analyzed and an improvement method is proposed to get better segmentation results in cases of lack of contrast or overlapping between bones and tissue regions.

There are several segmentation methods applied on X-ray images for the purpose of segmenting the bones with higher accuracy, however, because of the visual complexity in different bone structures, there is no single method which can be used successfully for the whole body segmentation. That is the reason why researchers attempted to combine different methods or improve existing ones, using mathematical or experience based approaches, for the purpose of getting a good segmentation result. K-means, fuzzy c-means and spatial fuzzy c-means algorithms are well-known clustering methods applied successfully to a variety of applications. Considering the X-ray image segmentation problem, they have significant weaknesses in segmenting regions with boundary insufficiencies (i.e., missing edges), lack of texture contrast between regions of interest (bones) and background (background and soft tissues), and presence of noise.

The main reason behind the weakness of the above mentioned methods for these problematic cases is their holistic approach in clustering the image pixels. The proposed entropy-driven improvements in clustering and statistical rule-based refinement methods advance the capability of their integrated algorithm to get significantly better results in dealing with missing edges, overlapped regions, low contrast and noisy X-ray images.

Considering the state of the art in employing K-means, fuzzy c-means and spatial fuzzy c-means algorithms for X-ray image segmentation, the following research works are found particularly useful on pioneering the ideas for the proposed improvement method. In this respect, Ng (2003) modified the K-means algorithm by changing the distance equation for the purpose of texture segmentation. Instead of using Euclidian distance measure, the author added an additional spatial proximity measure to the feature space distance. This modified distance measure is defined as,

$$(X_i,X_j) = (1-\lambda) \sum_k |X_i\ (k) - X_j\ (k)| + \lambda\ (|x_i - x_j| + |y_i - y_j|),\ 0 \leq \lambda < 1$$

Where $(x_i;y_i)$ and $(x_j;y_j)$ are pixel coordinates on the image that are normalized with regard to the image dimensions. The first term in the above equation is for measuring the degree of dissimilarity between two pixels (feature vectors) in Manhattan distance. The second term is for measuring the spatial separation between the locations of the two pixels by computing the distance using geometric space, where the parameter λ is used to adjust the relative emphasis between the two distances measures. Based on this research, improvements in both segmentation quality and speed of K-means method are achieved for some average and high contrast texture images; however, results of using this approach with low contrast images like X-rays is not mentioned in this paper.

Sarmila and Sujatha(2012), proposed a hybrid segmentation method for malignancy detection using FCM and active contour model. They improved the segmentation speed considerably by this proposed combination; however, evaluation of this method in terms of segmentation accuracy is not given. Based on the published experimental results, the presented hybrid method seems to be insufficient to improve the segmentation quality in terms of extracting the objects from their background.

Venkateswaran, and Muthukumar (2010) improved generalized spatial fuzzy c-means (GSFCM) by applying a genetic algorithm and they named this approach as GAGSFCM method. Based on their research using MR images, they succeeded to get higher accuracy than FCM and GSFCM algorithms. However, computational complexity of this method is too high compared to those of the FCM and GSFCM algorithms.

Shamsi and Seyedarabi (2012) modified the spatial fuzzy c-means algorithm for clustering MR images by giving a weight to each pixel (each point of dataset) in relation to every cluster. Therefore, they managed to make this method robust under noisy environments. Consequently, the resulting algorithm had better segmentation results compared to FCM and SFCM; however, the differences are not remarkable especially in cases of images with less noise.

The rest of this paper is organized as follows: Section 2 includes a review of the relevant literature on three well-known automatic segmentation methods, namely K-means, fuzzy c-means, and spatial fuzzy c-means algorithms. Section 3 illustrates a detailed explanation of the proposed improvement approaches for segmentation and pixel-wise refinements. Section 4 shows the experimental evaluations over several X-ray images selected to illustrate the achieved improvements both visually and computationally using well-known image quality assessment measures. Finally, conclusions and future work plans are discussed in Section 5.

2. Automatic Methods for Image Segmentation

Image segmentation algorithms can be grouped in three categories as follows: manual, semi-automatic and fully automatic. The three well-known automatic segmentation methods which are the target of this research are discussed in the following subsections.

2.1. K-means algorithm:

This method is one of the simplest, easiest, and widely used unsupervised techniques that are employed for solving the clustering problems. The main process of this algorithm starts by setting out ‘K’ (the number of clusters) by a user-defined or an automatic method.

Then, the process continues by defining the cancroids of clusters randomly, one for each cluster. The next step is to find the pixel-centroid distances for each pixel and consequently assigning them to the nearest cluster until no unflustered pixels remain. After this initialization step, centroid of each cluster is recalculated to update the positions of cluster centers, which is then followed by assigning each pixel to the new closest cluster (new centroid). This procedure is repeated until no changes occur in cluster centers (Mohd, Beg, Herawan, Noraziah, and Rabbi 2011).

2.2. Fuzzy c-means algorithm:

Similar to K-means method, this is an unsupervised technique that has been successfully applied to solve clustering problems. It works based on grouping of similar points/pixels into same clusters. It attempts to iteratively minimize a cost function which is the distance of pixels to the cluster centers in the feature domain.

Unlike K-means method, this method allows assignment of one point/pixel to more than one clusters in which the measure of belonging to a particular cluster is called the membership value. This degree is calculated based on similarity between any point and the cluster centers. It means that membership value will be higher if pixel and the cluster center is more similar and vice versa (Sarmila, et al. 2012).

2.3. Spatial fuzzy c-means algorithm:

In general, image pixels in immediate neighborhoods most probably have either the same or close grayscale levels. In this case, these pixels will belong to the same cluster with a high probability that may cause uncertainty in determining the exact object boundaries. Spatial fuzzy c-means algorithm handles this problem through using the spatial relationships among neighbor pixels. The algorithmic description of the spatial fuzzy c-means algorithm is exactly the same as that of fuzzy c-means algorithm except the following modification: The membership values are altered after computing the spatial function, h_{ij}, within the neighborhood of each pixel, X_j, as follows:

$$h_{ij} = \sum_{k \in NB\,(X_j)} u_{ik}$$

Where NB (X_j) represents a square window centered on pixel (X_j) in the spatial domain. Just like the membership values u_{ij}, the spatial function value (h_{ij}) represents the probability that pixel (X_j) belongs to i^{th}cluster. Finally, the spatial function is used to update the membership function values as given below:

$$\acute{u}_{ij} = \frac{u_{ij}^{p} h_{ij}^{q}}{\sum_{k=1}^{c} u_{kj}^{p} h_{kj}^{q}}$$

where $u_{ik}$and $\acute{u}_{ij}$ are the current and updated membership values, p and q are parameters to control the relative importance of both functions (spatial function and membership function)(Chuang, Tzeng, Chen, Wu, Chen 2006 , and Jaffar, Ahmed, Naveed, Hussain, and Mirza 2009).

3. The Proposed Method: Entropy Guided Cellular Automata Rule Engine for Tissue / Bone Clustering

The three automatic segmentation methods explained above exhibit insufficient performance in segmenting the bones in X-ray images correctly when they are used with low contrast images, noisy images, and when bones and tissues have overlapping in terms of grayscale levels. Consequently, they may misclassify part of bones as tissue and vice versa. In this research, to deal well with these problematic issues, an entropy-guided improvement for clustering and a statistical rule-based refinement for pixel clusters are proposed for the purpose of better differentiation of tissues and bones. As it is illustrated in Figure 1, the proposed method contains three phases as follow:

1. Segmenting a given X-ray image using only one of the three automatic segmentation methods improved by the entropy rule.

2. Passing both the segmented and the original images to the statistical rule engine to find out whether the tissue-labeled pixels really belong to the tissues or should be assigned to the bone cluster.
3. Merging the results of "Segmentation" and "statistical rule engine" phases to get the final output.

These three phases are presented in detail in the following subsections:

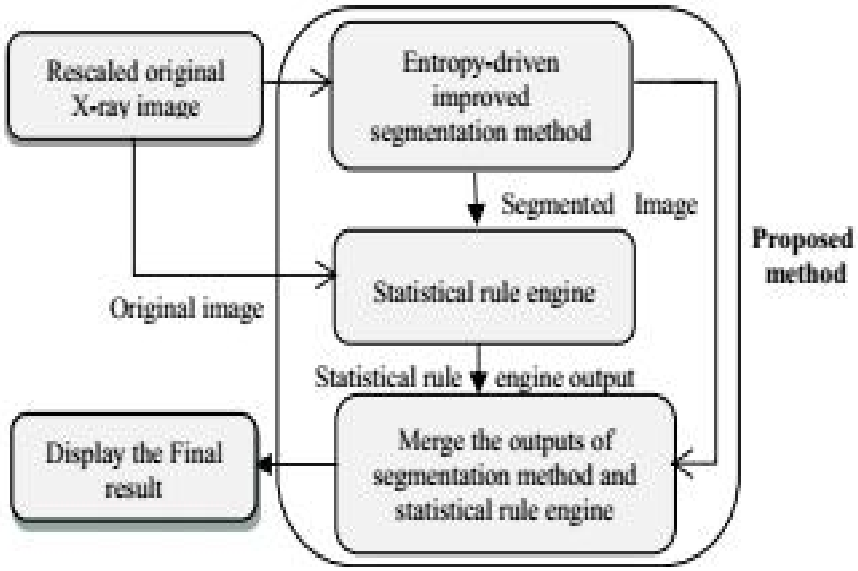


Fig. 1: The three phases of the proposed method

3.1. Automatic segmentation of X-ray images: Entropy-driven improvement for clustering algorithms

As explained above, the first phase of the proposed approach is the automatic segmentation of the input X-ray image using one of the improved entropy driven K-means, fuzzy c-means, or spatial fuzzy c-means algorithms. Application details of these algorithms are as follows: These improved clustering algorithms (i.e. in phase one) have two important characteristics. First of all, they do not need to specify any user-defined value as the number of clusters. Secondly, they have ability of eliminating clusters belonging to soft tissues and keeping only the clusters that most probably belonging to bones.

An algorithmic description of the entropy-guided improvements applied on the three well-known clustering algorithms is given below. A flowchart illustration is also provided in Figure 2 to better describe the flow of data throughout the computational steps.

1. Take the original image as input and find the number of clusters using the histogram method (Kanthan and Sujatha 2013).Number of peaks in the image histogram approximately shows the number of clusters in the image. Since the output of histogram method often indicates more clusters than the really present, some of the clusters should be removed or merged with the others to reduce the number of clusters to its actual value.
2. Cluster the input X-ray image with one of the clustering algorithms using the number of clusters found in step 1.
3. Sort the clusters based on their grayscale levels in ascending order. For example, cluster with grayscale values between 0 and 10 is labeled as the first cluster (i.e. indexed as J=1). Similarly, cluster with grayscale values close to 255 is set as the last cluster (i.e. indexed as J=K).
4. Set J=1 (Ignore the first low grayscale cluster which surely contains the background and tissue pixels). We have taken the advantage of the fact that tissues have low grayscale values compared to bones, hence the probability that the pixels in the first cluster belong to the background or soft tissues is very high.
5. Remove/Eliminate the first J clusters that belong to the background or soft tissues (if they are not removed before) simply by changing their grayscale pixel values to zero, and afterwards assume that the remaining clusters (with none zero pixel values) represent a single segment (bones).
6. Calculate entropy for both the original image and the segmented image obtained from step 5.
7. If entropy of the original image is less than α times the entropy of the segmented image (It means we need to remove more clusters with low grayscale values to reach better segmentation result) then set J=J+1 and go back to step 5, otherwise go to step 8 (for bone segmentation over x-ray images, 'α' is set to a fix number equal to 3.0).
8. Save the remaining clusters as output of the improved entropy-guided clustering algorithm.

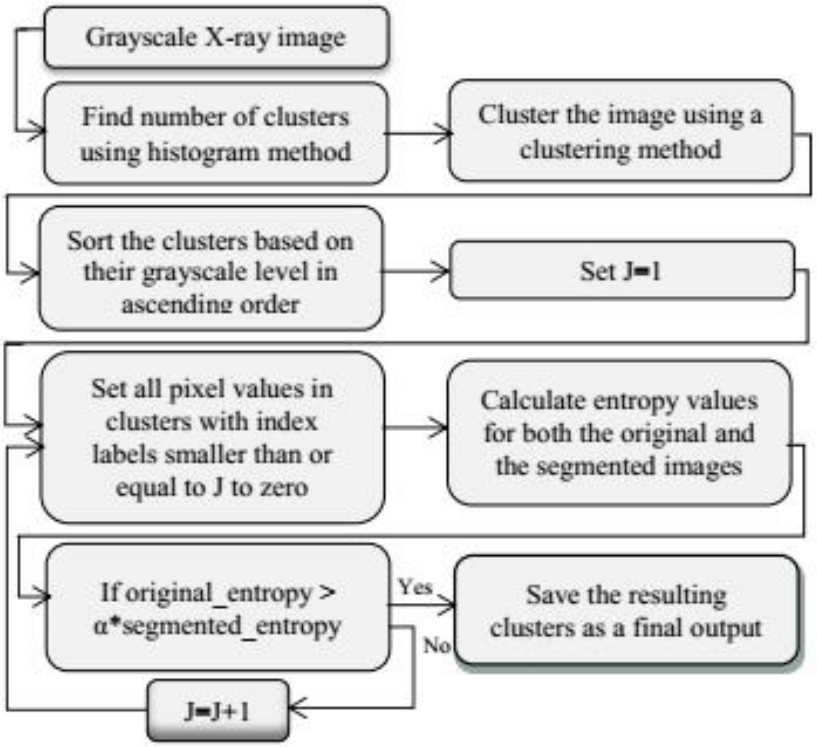


Figure 2: Flowchart description of the improved entropy-guided clustering algorithms.

In order to demonstrate the effect of factor α on the performance of the underlying clustering algorithms, we took the K-means algorithm into consideration and clustering outputs with two different values of this parameter are generated. Remembering that our fundamental objective is to cluster pixels that most probably belong to the bones, we want to keep only the pixels that certainly belong to the bonesto build the desired segmentation output. Rest of the pixels will be eliminated (by changing the corresponding pixel values to zero).

As described above, this is achieved using the entropy information over the original and the clustered images. If the entropy over the original image is less than "α times" entropy of the segmented image, then pixels corresponding to bone segment are not completely identified yet, hence we need to remove more clusters from the segmented image, otherwise all pixels within the bone segment are extracted and the procedure can be terminated.

For instance, Figure 3 illustrates the segmentation results using the improved K-means algorithm proposed in this research with two different entropy factors, α. When α is set equal to 3 (number of main entities, background, tissue, and bone), tissue labeled pixels are better differentiated from the ones identified as bone. However, by setting α=2, the improved K-means method works as a binary segmentation method (i.e. there are only two clusters, background and object).

a b c

Fig. 3: Improved K-means algorithm: effect of factor α on the performance of algorithm. a: original image, b:clustered image using α=3, c:clustered image using α=2.

3.2. Statistical rule engine for tissue/bone refinement

The fundamental objective of the statistical rule engine for tissue/bone refinement is to find out if the tissue-labeled pixels resulting from one of the three improved clustering algorithms (mentioned above) are really belonging to tissues or should be assigned to bone segment.

This engine is based on statistical analysis of pixels within a square window (mask) to further refine the labels of tissue and background clustered pixels. That is, the statistical rule engine processes only the pixels which are already labeled as tissue or background. The reason why rule engine should only be applied on pixels with the label of tissue or background is that, in the phase one of our proposed method, we have already found the pixels that certainly belong to the bone.

Therefore, there is no need to further investigate the possibility of their membership to the bone segment. The statistical rule engine works pixel-wise as follows: for every tissue- or background-labeled pixel on the segmented image, calculate the mean and standard deviation of the pixels within a predefined square mask in the original image.

If the grayscale level of the pixel under consideration, (i.e. pixel in center of mask) is greater than mask's mean plus γ times the mask's standard deviation. Then this pixel certainly belongs to the bone (should be assigned to the bone segment), otherwise it belongs to soft tissue or background. Parameter γ is set to 2.0 based on the Gaussian distribution. The fundamental algorithmic steps of phase 2 are given below:

1. Set the mask/window size=$\beta \times \beta$(β is a fix number for all images, e.g. for this paper it is been set equal to 15) for both original and segmented images, and set the masks' centers on the first (top-left) pixel of images.
2. Move the masks one pixel ahead on both images.
3. If the pixel value of mask center in the segmented image is equal to zero then calculate mean and standard deviation of pixels under the mask in original image, otherwise go back to step 2.
4. If pixel value of mask center in the original image is greater than T=mean + γ*std, then change the center pixel value in the segmented image from zero to one, otherwise without changing the pixel value go to step 5.
5. If all pixels in the segmented image are considered as mask centers and processed as above, then go to step 6, otherwise go to step 2.
6. Remove the salt (noise) present on the segmented image using Otsu's threes holding method (Otsu 1979) (if the grayscale value of the segmented image is smaller than Otsu's threes holding of original image, change the pixel value in segmented image to zero)
7. Find and keep only the largest segment in the image resulting from step 6 (by finding connected components and selecting the largest component/segment).
8. Add the output of this process (step 7) as an input to the merging step of the proposed method (see Figure 1 and 4).

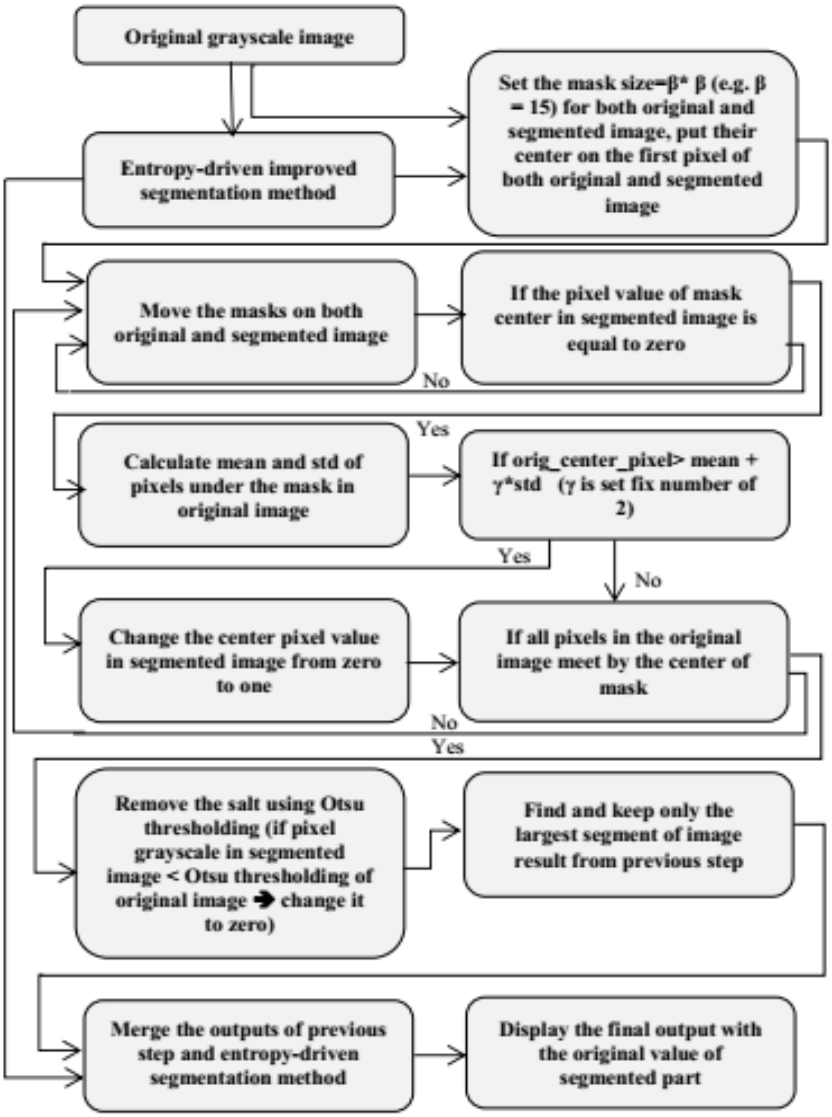
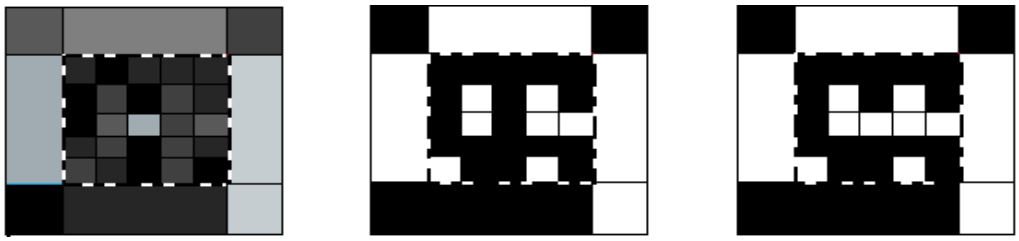


Figure 4: Flowchart description of the statistical rule engine.

To illustrate how the statistical rule engine works, an example using a 5×5 mask is presented in Figure 5 where the center pixel has bright color on the original image (Figure 5.a). This pixel is classified as tissue in the segmented image (Figure 5.b).

Hence, to further refine the cluster of this pixel, for being sure about its assigned label, the mean and standard deviation under the mask in original image are calculated and if the pixel value of the original image at this location is greater than T=mean + 2* standard_deviation, this pixel will be assigned to bone segment.

Figure 5-c shows that the statistical rule engine assigns the center pixel to the bone segment by changing its color from black (zero) to white (one).



a b c

Fig. 5: a: mask over the original image, b: mask over the segmented image, c: output of the statistical rule engine.

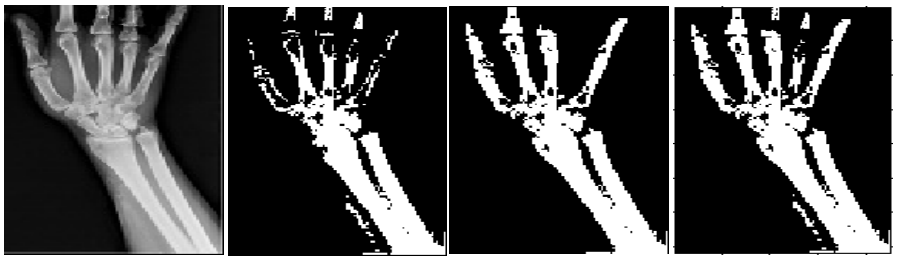
As a matter of fact, the proposed statistical rule engine for refining pixel clusters usually creates a lot of salt noises and outliers as shown in Figure 6.c. Then, we need to filter out the noise and, for this purpose; we used the Otsu's thresholding method. Figure 6.d presents the result of this noise removal procedure. Furthermore, as a result of operations performed so far within the framework of the proposed approach, bones constitute the largest segments that further helps us to decide which segments to keep as bones and which are to be removed totally.

Figure 6.e demonstrates the idea behind keeping the largest segment as bones while removing the ones which originate from the clones of tissue pixels to bone pixels. The final step of the proposed method includes the merging of the two outputs originating from the improved K-means algorithm and the statistical rule engine refinement proposed above. Output of this merging step is presented in figure 6.f. The merging operator is a simple pixel wise 'or' operator. Accordingly, if one of the two outputs identifies a pixel as bone, it belongs to bone segment, otherwise it is classified as tissue or background.

a b c d e f

Fig. 6: X-ray image of foot. a: original image, b: segmented image using improved K-means, c: output image of the statistical rule refinement before denoising, d: output image after salt removing, e: output image after removing all segment except the largest one, f: merging the Fig. 6.b and Fig. 6.e as a final segmentation output

Starting from an original X-ray image, Figure 7is another illustration of the outputs of all steps of the proposed methods. Output of the improved K-means algorithm is shown in Figure 7.b. It can easily be seen that parts of finger bones are segmented as tissue/background. The proposed statistical refinement rule engine followed the Otsu's denoising and the largest segment methods and generated the image in Figure 7.c. This image is better than the one in Figure 7.b in terms of separating bones from tissues, however when applied alone, it may misclassify some bones as tissue. After merging the outputs of improved K-means and statistical rule engine, simply by an 'or' operation, precise segmentation of bones is obtained in Figure 7.d. Comparing Figure 7.b and Figure 7.d clearly indicates the improvement obtained by proposed methods.



a b c d

Fig. 7: X-ray image of hand. a: original image, b: segmented image using improved K-means, c: output of statistical rule engine followed by a denoising and largest segment method, d: final output of proposed method obtained from merging b and c.

4. Experimental Evaluation

In this research work, MATLAB R2013b platform is used for implementing the proposed methods. X-ray images used in experimental work are collected from different web sites (www.commons.wikimedia.organd www.fineartamerica.com).The selection of these images is carried out carefully to exhibit the capabilities of the proposed methods under different conditions such as low contrast and/ or grayscale overlapping between bones and tissues. Basically, the X-ray images used in this research belongs to the following body parts: pelvic (hip bone, sacrum, and coccyx), hand (carpals, metacarpals, and phalanges), arm (ulna, radius, and humerus) and foot (femur, tibia, and fibula).

There are three algorithmic parameters to be considered in experimental work, namely entropy-related factor α used in improved K-means, Fuzzy and Spatial fuzzy c-means algorithms, mask size β, and γ used in statistical rule engine. For all algorithms and trials, entropy-related factor α is set to a fixed value equal to 3.0 for X-ray images and 2.0 for binary images. Mask size β and statistical rule engine parameter γ are set to fixed values 15 and 2, respectively, for all images and algorithms. It can easily be seen that only the parameter α needs to be tuned by the user depending on the image type under consideration.

To compare the performance of the proposed methods with the performances of the three underlying automatic image segmentation methods, namely the K-means, fuzzy c-means, and the spatial fuzzy c-means algorithms, nine well-known quality assessment measures are taken into account. These assessment measures are namely, Accuracy, F-measure, PSNR, Jaccard index, Structural content, Normalized cross-correlation, Average difference, Normalized absolute error, and Kolmogorov-Smirnov test [Jaccard 1912,Massey 1951,and Varnan, Jagan, Kaur, Jyoti and Rao 2011].Each of these assessment measures and their interpretation in terms of segmentation quality are briefly described below in Table 1. Abbreviations used in tables below are as follows: ACC = Accuracy, FM = F-measure, JI = Jaccard Index, SC=Structural Content, NCC =Normalized Cross Correlation, AD = Average Difference, KS = Kolmogorov Smirnov, NAE = Normalized Absolute Error, KM = K-means, KMN= Improved K-means, FCM = Fuzzy c-means, FCMN= Improved Fuzzy c-means, SFCM = Spatial fuzzy c-means, SFCMN= Improved Spatial fuzzy c-means, and Imp % = Percent improvement.

Table 1: Methods for quality assessment used in this research

ssesm. Method	Definition	Interpretation
ACC	Accuracy is used to find the percentage of correctly detected pixels (both true positive and true negative) to all pixels of input image. That is, *ACC=(TP+TN)/N*, where TP is the true positive, TN is the true negative and N is the number of pixels in the given image.	The higher the accuracy better is the quality of segmentation.
FM	F-measure is an indication of a test's accuracy. It is defined based on two measurements called precision and recall. Precision is the number of correct results divided by the number of all returned results, whereas recall is the number of correct results divided by the number of results that should have been returned. Accordingly, F-measure is defined as, *F=2*(precision*recall)/(precision+ recall)*	F-measure score reaches its best value at 1 and worst score at 0.
PSNR	This is the peak signal-to-noise ratio (PSNR) between ground truth and segmentation result, and defined as *PSNR =10*log10((255^2)/MS), MS =* $\sum \sum (g-s)^2 / N$ where, g is ground truth, s is segmented image and N is total number of pixels.	In PSNR, greater values indicate greater image similarity.
JI	Jaccard index computes the percentage of all correctly detected bone pixels to the number of pixels in the image without the pixels that are correctly detected as background and tissue. Considering that A is the ground truth and B is the segmented image, this quality measure is defined as, J=M11/ (M01+M10+M11) , where M11 = number of pixels where A and B both have a value of 1, M01 = number of pixels where the pixel of A is 0 and the pixel of B is 1, M10 = number of pixels where the pixel of A is 1 and the pixel of B is 0.	Jaccard index is a similarity measurement method increasing values of which indicate better quality in the segmentation result.

Table 1 continues. Methods for quality assessment used in this research

Assesm. Method	Definition	Interpretation
SC	It is a correlation based measure that is used to measure the similarity between two images (ground truth and segmented image), that is defined as $SC=\sum \Sigma \quad s^2 / \sum \Sigma \quad g^2$ where, g is ground truth and s is the segmented image.	Higher scores of structural content mean more similarity between ground truth and the segmented image.
NCC	Normalized cross-correlation (NK) measures the similarity between ground truth and the segmented image. $NK=\left(\sum \Sigma \quad g*s\right) / \sum \Sigma \quad g^2$ where,g is ground truth and s is segmented image.	A higher value of NK means better segmentation quality. The range of this measure is [-1, 1] and, a score of 1 means complete similarity between two images.
AD	It is a dissimilarity measure based of the average of difference between the ground truth and segmented image. $AD=\left(\sum \Sigma \quad g-s\right) /N$ where g is ground truth and s is the segmented image and N is total number of pixels.	Lower values of this assessment measure means higher similarity between ground truth and the segmented image.
NAE	This assessment measure is based on the calculation of dissimilarity between ground truth and the segmented image. It is simply the average of absolute differences between pixels of ground truth and the segmented image:$NAE=\sum \Sigma \quad \lvert g-s\rvert / \sum \Sigma \quad g$ where,g is ground truth and s is the segmented image.	Lower scores of NAE mean better segmentation quality.
KS	The Kolmogorov–Smirnov test score is a measure of distance between the distribution function of the segmented image and the distribution function of the ground truth.	Lower values of Kolmogorov-Smirnov test score means higher similarity of two images.

The following figures are presented to demonstrate the success of the proposed approach under various well-known image quality measures. Each image is also provided with a ground truth that is prepared manually from the original image, to be used in computation of some assessment scores. Comparison with the well-known algorithms with respect to each assessment score are also provided under each figure.

Additionally, the improved versions of well-known algorithms, implemented through using of the proposed entropy guided clustering and statistical rule engine are named by appending letter 'N' to their short names, e.g. KM and KMN that represent the K-means and its improved implementation, respectively.

Figure 8 is an X-ray image of pelvic including hip bone, sacrum, coccyx and parts of femur and backbone (lumbar lordosis). This image is significant to exhibit the segmentation of a complex bone structure consisting of several joints, different thicknesses, and overlapping pixel values within tissue and bone regions. In this figure, segmentation results of three well-known methods and their improved implementations are put side-by-side to make the visual and computational comparisons easier. Table 2 illustrates the scores achieved with respect to nine assessment measures.

For all the three automatic segmentation algorithms, their corresponding improved implementations performed significantly better. The improvements achieved for the KM and SFCM methods are around 20% for most of the assessment scores, whereas for the FCM algorithm the proposed improvement and refinement methods achieved around 10% improvements in segmentation quality for most of the quality measures. By visual inspection, we can say that the proposed methods are quite powerful in identifying the boundaries of bone segments; however, some parts of bones having very similar pixel values to tissues are classified as belonging to tissue/background segment. This fact is particularly seen on the flat ends of the hip bone and femur.

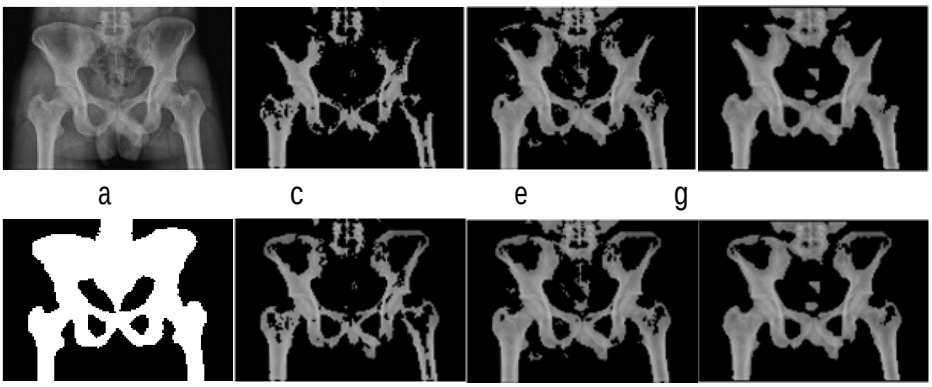


Fig. 8: X-ray image of pelvic (hip bone, sacrum, and coccyx), parts of femur and backbone (lumbar lordosis) a: original image, b: ground truth, c: KM segmentation, d: KMN segmentation, e: FCM segmentation, f: FCMN segmentation, g: SFCM segmentation, h: SFCMN segmentation.

Table 2: Segmentation quality assessment scores of KM, KMN, FCM, FCMN, SFCM, SFCMN methods for Figure 8.

Method Alg.	KM	KMN	Imp.%	FCM	FCMN	Imp.%	SFCM	SFCMN	Imp.%
ACC	0.751	**0.801**	6.631	0.801	**0.810**	1.161	0.810	**0.855**	5.505
FM	0.594	**0.702**	18.29	0.702	**0.723**	2.934	0.723	**0.809**	11.97
PSNR	30.18	**31.24**	3.522	31.24	**31.54**	0.948	31.54	**33.345**	5.722
JI	0.422	**0.541**	28.19	0.541	**0.566**	4.602	0.566	**0.680**	20.09
SC	0.482	**0.571**	18.30	0.571	**0.609**	6.641	0.609	**0.718**	18.04
NCC	0.473	**0.555**	17.52	0.555	**0.587**	5.599	0.587	**0.674**	14.95
AD	29.61	**23.73**	19.85	23.73	**21.92**	7.632	21.92	**15.42**	29.65
NAE	0.542	**0.444**	18.22	0.444	**0.419**	5.614	0.419	**0.332**	20.64
KS	0.240	**0.184**	23.48	0.184	**0.168**	8.651	0.168	**0.091**	45.92

Figure 9 is another X-ray image of pelvic. This image includes large part of hip bone, and small parts of femur and backbone (lumbar lordosis). Difficulties associated with the segmentation of this image are due to huge overlapping between pixel values of tissues and bone regions. Considering the nine quality assessment methods, Table 3illustrates the segmentation quality scores achieved by all algorithms in the experimental suit. Again, for all the three automatic segmentation algorithms, their corresponding improved implementations performed better.

The improvements achieved for the KM and SFCM methods are approximately10% for most of the assessment scores, whereas for the FCM algorithm the proposed methods achieved around 5% improvements. Visual inspection of the segmentation results also exhibits that segmentations associated with the improved versions of algorithms are better in detecting boundaries and complete shapes of bones. However still there are some parts of bones with very close pixel values to those of tissues. Therefore, these parts are classified wrongly as tissue/background segment, especially on the flat ends of the hip bone and part of a backbone which called lumbar lordosis.

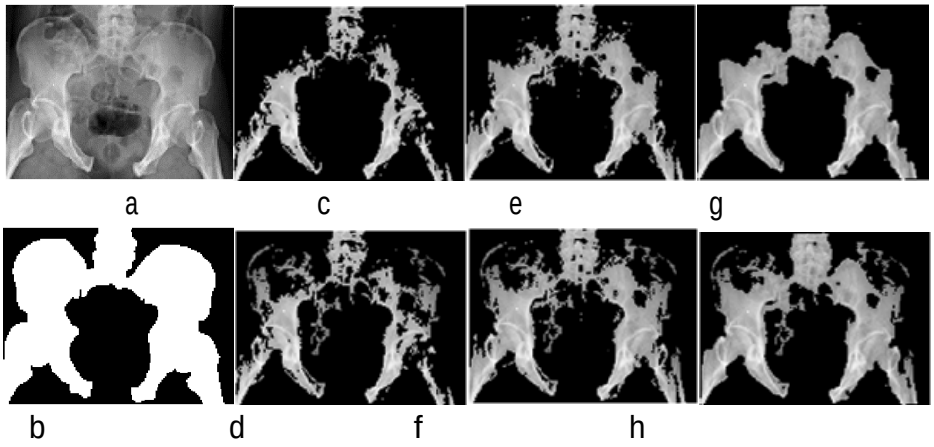


Fig. 9: X-ray image of pelvic. a: original image, b: ground truth, c: KM segmentation, d: KMN segmentation, e: FCM segmentation, f: FCMN segmentation, g: SFCM segmentation, h: SFCMN segmentation.

Table 3: Segmentation quality assessment scores of KM, KMN, FCM, FCMN, SFCM, SFCMN methods for Figure 9.

Method Alg.	KM	KMN	Imp.%	FCM	FCMN	Imp.%	SFCM	SFCMN	Imp.%
ACC	0.694	**0.736**	6.037	0.772	**0.805**	4.286	0.797	**0.830**	4.153
FM	0.587	**0.670**	14.21	0.724	**0.778**	7.501	0.767	**0.816**	6.296
PSNR	29.22	**30.02**	2.729	29.58	**29.82**	0.827	31.33	**32.39**	3.399
JI	0.415	**0.504**	21.39	0.567	**0.607**	6.998	0.622	**0.689**	10.62
SC	0.475	**0.546**	15.00	0.515	**0.535**	3.903	0.655	**0.706**	7.818
NCC	0.473	**0.534**	12.76	0.512	**0.531**	3.653	0.635	**0.678**	6.806
AD	48.36	**41.40**	14.39	44.51	**42.39**	4.773	31.03	**25.77**	16.95
NAE	0.547	**0.483**	11.75	0.506	**0.484**	4.409	0.375	**0.327**	12.87
KS	0.305	**0.244**	19.87	0.279	**0.263**	5.874	0.173	**0.124**	28.10

The X-ray image of pelvic (hip bone, sacrum, and coccyx), femur and very small part of backbone (lumbar lordosis) is illustrated in Figure 10. This image is particularly interesting due to the complete embedding of bones within the whole body tissue. Low contrast and overlapping of regions between issues and bones makes segmentation of this image very difficult. Similarly to the previous two figures, the segmentation results of the three segmentation methods and their improved implementations are represented below. Significant improvements in segmentation quality are both visually and computationally observable.

As seen in Table 4, improvements achieved in the three algorithms are very similar except the KS assessment score. While KMN algorithm is 25% better with respect to KM, FCMN exhibit the same score with FCM. This is the only X-ray image for which SFCMN is worse than SFCM with respect to the KS assessment score.

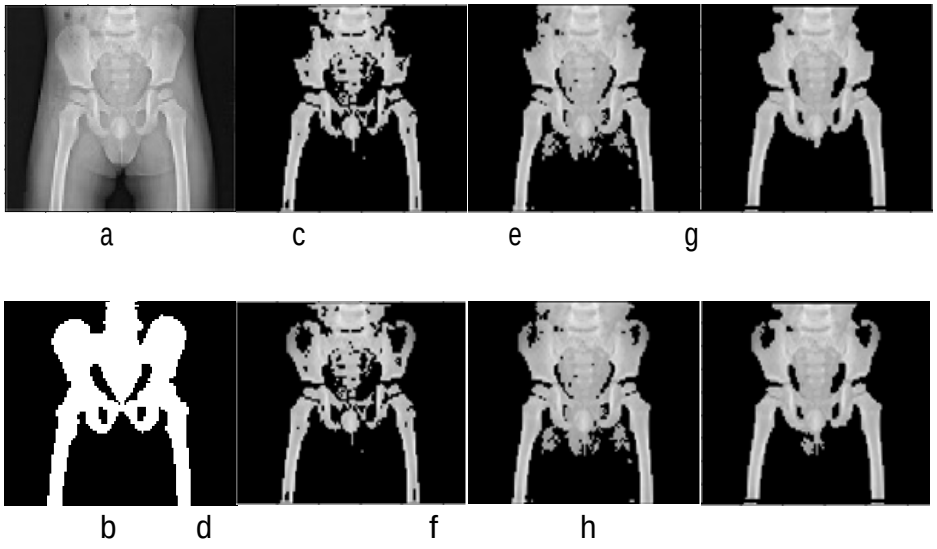


Fig. 10: X-ray image of pelvic (hip bone, sacrum, and coccyx), femur and part of backbone (lumbar lordosis).a: original image, b: ground truth, c: KM segmentation, d: KMN segmentation, e: FCM segmentation, f: FCMN segmentation, g: SFCM segmentation, h: SFCMN segmentation.

Table 4: Segmentation quality assessment scores of KM, KMN, FCM, FCMN, SFCM, SFCMN methods for Figure 10.

Method Alg.	KM	KMN	Imp.%	FCM	FCMN	Imp.%	SFCM	SFCMN	Imp.%
ACC	0.883	**0.899**	1.813	0.900	**0.911**	1.210	0.915	**0.925**	1.071
FM	0.776	**0.812**	4.690	0.821	**0.839**	2.119	0.856	**0.876**	2.383
PSNR	34.51	**35.37**	2.473	35.99	**36.62**	1.753	37.96	**39.83**	4.905
JI	0.634	**0.684**	7.915	0.697	**0.722**	3.646	0.748	**0.780**	4.236
SC	0.865	**0.902**	4.278	0.969	**0.988**	2.044	1.046	**1.089**	4.112
NCC	0.769	**0.806**	4.800	0.847	**0.865**	2.209	0.904	**0.935**	3.452
AD	14.39	**11.96**	16.85	9.881	**8.613**	12.83	6.214	**4.119**	33.70
NAE	0.362	**0.317**	12.45	0.306	**0.284**	7.288	0.262	**0.237**	9.729
KS	0.063	**0.047**	25.56	**0.036**	**0.036**	0	**0.032**	0.035	-9.52

To illustrate and compare the success of algorithms for thin and multiply joined bone structures, Figure 11 shows an X-ray image of hand including carpals, metacarpals, and phalanges. There are overlapping regions between tissues and bones in regions around the metacarpals, phalanges and regions between ulna and radius which are connected to the end of carpals.

Table 5includes scores computed over the nine assessment measures. Surprisingly we get the same results for both K-means and fuzzy c-means algorithms and consequently, the same degree of improvement in their improved versions. Here again the improved methods performed significantly better than their original ones. For instance, the KM and FCM algorithms are identically improved in a diverse range of approximately from 5%to62%, whereas SFCM is improved in a range from10% to 80% for the different quality measures. Based on the achieved results for this image, it is crystal clear that proposed method out performed its conventional counterparts by detecting almost all details of the thin bones and joints.

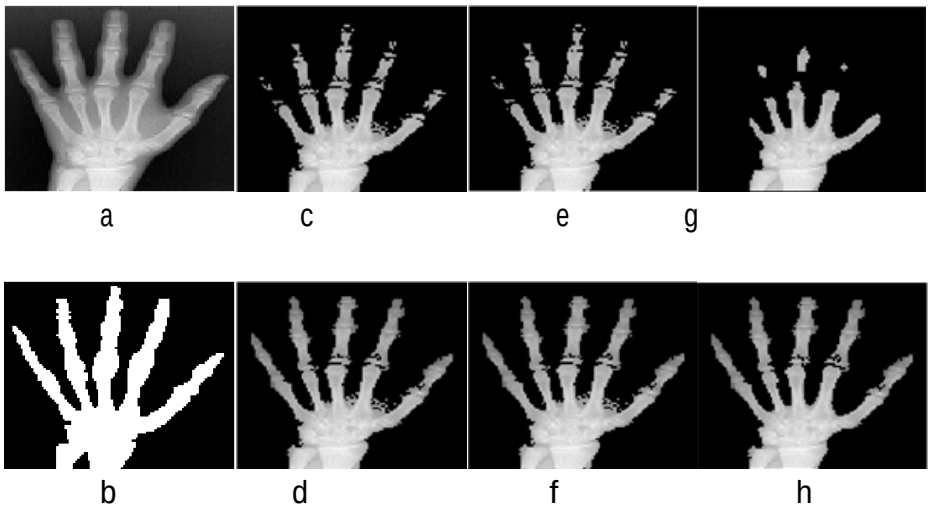


Fig. 11: X-ray image of hand includes: carpals, metacarpals, and phalanges. a: original image, b: ground truth, c: KM segmentation, d: KMN segmentation, e: FCM segmentation, f: FCMN segmentation, g: SFCM segmentation, h: SFCMN segmentation.

Table 5: Segmentation quality assessment scores of KM, KMN, FCM, FCMN, SFCM, SFCMN methods for Figure 11.

Method	Alg. KM	KMN	Imp.%	FCM	FCMN	Imp.%	SFCM	SFCMN	Imp.%
ACC	0.904	**0.950**	5.032	0.904	**0.950**	5.032	0.861	**0.951**	10.47
FM	0.811	**0.911**	12.4	0.811	**0.911**	12.40	0.690	**0.911**	32.06
PSNR	34.94	**36.79**	5.283	34.94	**36.79**	5.283	32.87	**38.62**	17.49
JI	0.681	**0.827**	21.31	0.681	**0.827**	21.31	0.527	**0.837**	58.91
SC	0.866	**0.925**	6.823	0.866	**0.925**	6.823	0.694	**0.940**	35.51
NCC	0.821	**0.927**	12.82	0.821	**0.927**	12.82	0.668	**0.935**	40.09
AD	11.41	**4.317**	62.18	11.41	**4.317**	62.18	19.81	**4.231**	78.64
NAE	0.273	**0.147**	46.15	0.273	**0.147**	46.15	0.418	**0.146**	65.09
KS	0.068	**0.039**	42.39	0.068	**0.039**	42.39	0.124	**0.023**	81.88

In figure 12, X-ray image of right hand is shown. In this image there is intensive overlapping between tissue and bone regions around the metacarpals and more considerably in phalanges in terms of grayscale values. As it shown in Figure 12.c, .e and .g, all three well-known methods failed to segment the bones in the phalanges region. However, the results obtained by the improved versions of these methods exhibit highly significant improvements in the bone segmentation. As it is shown in Table 6, assessment scores obtained for K-means and fuzzy c-means algorithms are around 30%, whereas the modified version of SFCM exhibits improvements around 50%. This is another example where the proposed rule-based improvement technique showed noticeable success that is significantly better than KM, FCM and SFCM algorithms in segmenting the bones.

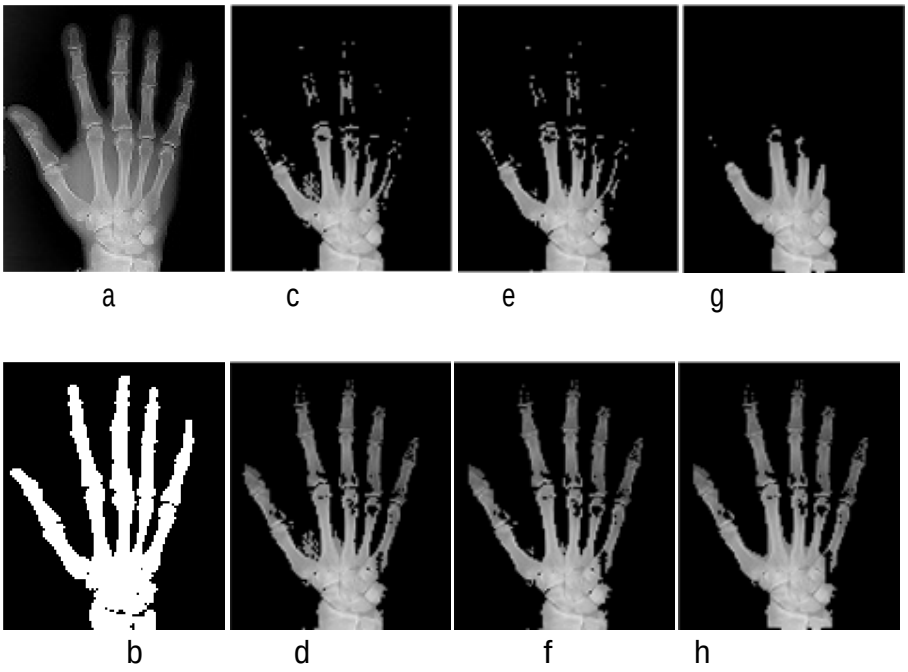


Fig. 12: X-ray image of hand includes: carpals, metacarpals, and phalanges. a: original image, b: ground truth, c: KM segmentation, d: KMN segmentation, e: FCM segmentation, f: FCMN segmentation, g: SFCM segmentation, h: SFCMN segmentation.

Table 6: Segmentation quality assessment scores of KM, KMN, FCM, FCMN, SFCM, SFCMN methods for Figure 12.

Method Alg.	KM	KMN	Imp.%	FCM	FCMN	Imp.%	SFCM	SFCMN	Imp.%
ACC	0.868	**0.946**	8.97	0.866	**0.946**	9.32	0.844	**0.947**	12.12
FM	0.688	**0.893**	29.88	0.671	**0.894**	33.16	0.607	**0.894**	47.37
PSNR	33.20	**36.84**	10.97	32.19	**37.61**	16.84	31.39	**37.78**	20.35
JI	0.524	**0.807**	53.99	0.545	**0.808**	48.19	0.435	**0.808**	85.60
SC	0.662	**0.808**	21.99	0.667	**0.808**	21.08	0.574	**0.808**	40.79
NCC	0.632	**0.765**	21.01	0.610	**0.768**	25.84	0.549	**0.771**	40.24
AD	15.73	**7.690**	51.11	14.81	**7.540**	49.10	19.40	**7.333**	62.20
NAE	0.428	**0.219**	48.80	0.436	**0.210**	51.93	0.514	**0.208**	59.60
KS	0.112	**0.032**	71.48	0.126	**0.032**	74.76	0.139	**0.032**	77.11

Another complicated X-ray image to illustrate the comparative performance improvements brought by the proposed method is shown in Figure 13. This image covers hand (carpals, metacarpals, and phalanges) and arm (ulna and radius). Its complexity is due to the number and different types of bones existing on it and too much overlapping between grayscale values of tissues and bones. Table 7 illustrates the amount of improvements achieved by each improved version of KM, FCM and SFCM algorithms using nine quality measures. The improvements achieved for the KM algorithm is around 10%,5% for FCM and 30% for SFCM. Visual inspections of the segmented images also demonstrate the clear enhancements obtained by the proposed method.

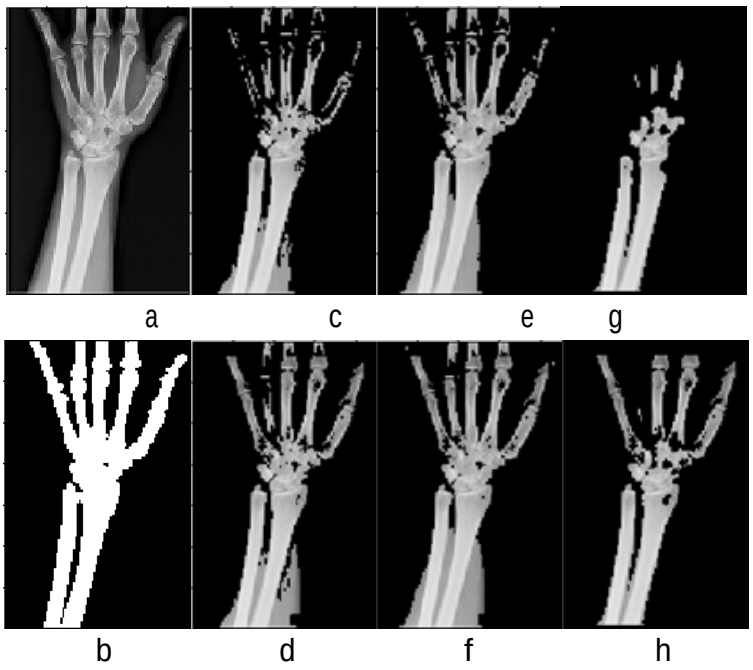


Fig. 13: X-ray image of hand and arm contains: carpals, metacarpals, and phalanges, ulna, and radius. a: original image, b: ground truth, c: KM segmentation, d: KMN segmentation, e: FCM segmentation, f: FCMN segmentation, g: SFCM segmentation, h: SFCMN segmentation.

Table 7: Segmentation quality assessment scores of KM, KMN, FCM, FCMN, SFCM, SFCMN methods for Figure 13.

Method Alg.	KM	KMN	Imp.%	FCM	FCMN	Imp.%	SFCM	SFCMN	Imp.%
ACC	0.877	**0.910**	3.658	0.893	**0.913**	2.229	0.840	**0.917**	9.256
FM	0.743	**0.829**	11.60	0.798	**0.824**	3.308	0.604	**0.828**	37.24
PSNR	34.13	**36.56**	7.14	33.79	**34.31**	1.535	32.21	**35.27**	9.52
JI	0.591	**0.708**	19.81	0.644	**0.721**	11.93	0.432	**0.707**	63.57
SC	0.830	**0.941**	13.35	0.788	**0.833**	5.815	0.585	**0.800**	36.90
NCC	0.758	**0.847**	11.82	0.806	**0.840**	4.281	0.563	**0.779**	38.32
AD	13.86	**8.186**	40.94	15.71	**13.62**	13.27	23.39	**11.54**	50.66
NAE	0.372	**0.281**	24.36	0.376	**0.354**	5.907	0.515	**0.266**	48.30
KS	0.075	**0.072**	3.605	0.076	**0.075**	1.189	0.146	**0.069**	53.11

Figure 14 is an X-ray image of arm containing hummers, ulna, and radius. It is an image with very high overlapping with the tissue in the ulna, and radius regions and some overlapping in the joint area. After applying the proposed method to the three well-known methods used in this research, the achieved improvements in segmenting bones are shown visually and computationally in Figure 14.d, .f, and .h. The corresponding improvements with respect to the listed assessment measures are shown in Table 8.

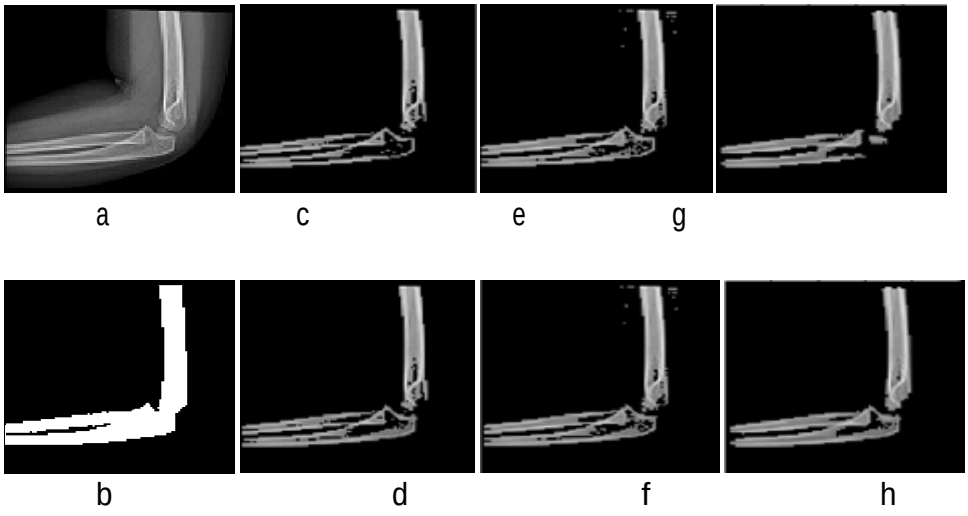


Fig. 14: X-ray image of arm contains hummers, ulna, and radius. a: original image, b: ground truth, c: KM segmentation, d: KMN segmentation, e: FCM segmentation, f: FCMN segmentation, g: SFCM segmentation, h: SFCMN segmentation.

Table 8: Segmentation quality assessment scores of KM, KMN, FCM, FCMN, SFCM, SFCMN methods for Figure 14.

Method Alg.	KM	KMN	Imp.%	FCM	FCMN	Imp.%	SFCM	SFCMN	Imp.%
ACC	0.949	**0.960**	1.117	0.964	**0.966**	0.228	0.960	**0.973**	1.386
FM	0.812	**0.857**	5.554	0.876	**0.884**	0.982	0.861	**0.911**	5.892
PSNR	37.01	**38.07**	2.872	38.74	**39.06**	0.844	38.35	**40.35**	5.236
JI	0.684	**0.750**	9.729	0.779	**0.792**	1.760	0.755	**0.837**	10.82
SC	0.812	**0.855**	5.322	0.888	**0.898**	1.058	0.844	**0.913**	8.18
NCC	0.811	**0.853**	5.168	0.880	**0.889**	0.966	0.837	**0.904**	8.030
AD	5.746	**4.464**	22.313	3.711	**3.438**	7.359	4.527	**2.725**	39.80
NAE	0.247	**0.194**	21.484	0.169	**0.158**	6.106	0.204	**0.130**	36.45
KS	0.051	**0.039**	22.53	0.032	**0.029**	8.696	0.035	**0.020**	41.45

5. Conclusion and Future Work

Considering X-ray images and the bone segmentation problem, a novel Entropy-driven improvement for clustering and a statistical rule-based refinement are presented to improve segmentation quality of three well-known algorithms. To evaluate the comparative performance of the proposed improvement methods computationally nine well-known image quality assessment methods that measure the resulting segmentations from different points of quality measures are used.

Experimental evaluations based on visual inspections and computational results showed that objectives of the proposed improvement method are reached at satisfactory levels. Bone segmentation of X-ray images with varying degrees of complexity is carried out with better quality scores compared to the achievements of the underlying fundamental algorithms. The level of improvements goes up to 80% for some severely complex X-ray images. As a future work, it could be very effective to use interval type-2 fuzzy sets to ensemble the results earned from different integration of the proposed method. This may help to handle uncertainties in medical images more efficiently.

References

CHUANG,K. S., TZENG, H. L.,CHEN,S.,WU,J., CHEN,T. J. (2006), "Fuzzy c-means clustering with spatial information for image segmentation" , Elsevier, Computerized medical imaging and graphics, vol. 30.

JACCARD, P. (1912), "The distribution of the flora in the alpine zone", New Phytol. 11 (2), 37–50.

JAFFAR, M. A., AHMED, B.,NAVEED, N.,HUSSAIN, A., and MIRZA, A. M. (2009), "Color video segmentation using fuzzy c-mean clustering with spatial information", Wse as Transactions on Signal Processing, issue 5, volume 5.

KANTHAN, M. R., and SUJATHA,S.N.N.(2013), "Automatic Grayscale Classification using Histogram Clustering for Active Contour Models", International Journal of Current Engineering and Technology, ISSN 2277 – 4106.

MAHENDRAN,S. K., and BABOO, S.S. (2011), "Enhanced automatic X-ray bone image segmentation using wavelets and morphological operators", International conference on information and electronics engineering, IPCSIT vol.6, IACSIT Press, Singapore.

MASSEY, F. J. (1951), "The Kolmogorov-Smirnov test for goodness of fit", Journal of the American statistical association. vol. 46, no. 253, pp. 68–78.

MOHD,W. M. B. W.,BEG,A. H.,HERAWAN,T.,NORAZIAH,A., and RABBI,K. F.(2011), "Improved parameter less K-means: auto-generation cancroids and distance data point clusters", IJIRR.

NG, B. W. (2003), "Wavelet based image texture segmentation using a modified K-means algorithm", Ph.D. Thesis, University of Adelaide, Australia

OTSU, N. (1979), "A threshold selection method from gray-level histograms", IEEE Trans. Syst. Man Cybern, 9: 62-66.

SARMILA, S. S., SUJATHA, S.N.(2012), "Hybrid segmentation method for malignancy detection using fuzzy-C-means and active contour model", International journal of engineering inventions, ISSN: 2278-7461, ISBN: 2319-6491, vol. 1, issue 11.

SHAMSI,H. and SEYEDARABI, H. (2012), "A modified fuzzy c-means clustering with spatial information for image segmentation", International journal of computer theory and engineering, vol. 4, no. 5.

VARNAN, C. S., JAGAN, A., KAUR, J., JYOTI, D.,RAO, D. S. (2011), "Image quality assessment techniques pn spatial domain", IJCST, vol. 2, issue 3.

VENKATESWARAN, R., MUTHUKUMAR, S. (2010), "Genetic approach on medical image segmentation by generalized spatial fuzzy c- means algorithm", IEEE International conference on computational intelligence and computing research. www.commons.wikimedia.org/w/index.php?search=medical+x-xay+image&title=Special%3A Search & uselang=en www.fineartamerica.com/art/all/Medical+images/all.