CLASSIFICATION OF KNEE MRI IMAGES

RAJNEET KAUR

University Institute of Engineering & Technology, Panjab University

MR. NAVEEN AGGARWAL

Assistant Professor, *University Institute of Engineering & Technology, Panjab University*

Abstract

Classification is very important part of digital image analysis. It is a computational procedure that sort images into groups according to their similarities. MRI is latest medical imaging technology. Magnetic Resonance Imaging used for Knee scans is very useful and effective technique to detect the knee joint defects. It is a non-invasive method to take picture of knee joint and the surrounding images. Classification of knee MR Images is done for the analysis purpose. In the preprocessing steps, segmentation is done using active contour without edges by chen and vese. Region containing cartilage thickness is separated out. During feature extraction part total 46 features have been extracted out of which 19 are DICOM image header features, 13 are haralick features and others are statistical moment features. Database file is prepared for 704 Knee MRI images and 46 attribute. In pre-processing 5 features has been removed. Database file is given as input for classification. As the result of classification FT algorithms classifies all instances correctly. Classification is done on the base of parameter 'Slice Thickness'.

Key words: Classification, Segmentation, Feature extraction, Active contour, Haralick features.

1. Introduction

Over the last five years, new generations of medical data mining tools have dramatically impacted the health care industry by improving the diagnosis of medical diseases and by reducing the time pressure on physicians. Classification is a computational procedure that sorts images into groups ("classes") according to their similarities [1]. There are two different approaches to classification: supervised and unsupervised. Both make use of the similarity measure introduced above, but one (supervised) classifies a set of images according to their similarity (speak: closeness in our high-dimensional space) with certain pre-given images ("references" or "templates"), the other (unsupervised) classifies the images according to their intrinsic grouping or clustering within the set [2]. For classification of images firstly segmentation of images is done.

1.1 Segmentation

In computer vision, segmentation refers to the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze [3]. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (see edge detection). Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic.

Several general-purpose algorithms and techniques have been developed for image segmentation. The segmentation techniques used is 'Active Contour without edges' by chen and vese. The active contour model by chen and vese [4] is used to detect the objects in the images using the technique called as curve evolution which was originally proposed by mumford-shah [5] function for segmentation and for defining level sets. The basic idea of this model is similar to the mumford shahs function that is detecting the object using the curve subject to constraints from a given image. For instance starting with object to be detected the curve starts from the one point of image and moves towards the interior normal and when the boundary of the object comes , the curve stop moving further. This is very useful technique as it is used to detect the objects in images whose boundaries are not clearly defined by the gradient. This advantage of this technique is that it helps in minimizing the energy which can be seen as a particular case of minimal partition problem. In the level set formulation, this problem becomes a 'mean-curvature flow' like evolving the active contour, which will stop on the desired and properly defined boundary. However the stopping criteria used in this completely depends upon the segmentation of image not on the gradient of the image. By the functional for "minimal partition problem", and by given curve $C= \Omega\omega$, with ω =subset (Ω)

An open subset and two unkown constants C1 and C2, denoting $\Omega 1= \omega$, $\Omega 2=\Omega-\omega$. We have proposed to minimize the following energy with respect to C1, C2 and C.

$$F(C1,C2, C)= \int_{\Omega 1= \omega} [(u_0(x,y)- C1)^2 dxdy \; + \int_{\Omega 2=\Omega-\omega} [(u_0(x,y)- C2)^2 dx\, dy + \nu \,|C| \qquad (1)$$

In order to understand what's going on with this idea, let's see some figures from the original paper.

In the paper: Active contours without edges [26], the first two terms have been interpreted to two forces. The first term is the force to shrink the contour. The second term is the force to expand the contour. These two forces get balanced when the contour reaches the boundary of our interested object.

$$F1(C) +F2(C)= \int_{inside(c)} | u_{0-C1} |^2 \, dx + \int_{outside(c)} | u_{0-C2} |^2 \, d \qquad (2)$$

1.2 Feature Extraction

Large number of algorithms has been proposed for the extraction of features from knee MRI images. Texture analysis serve as a base for various feature description. Statistical, structural, spectral, filtering, histograms, transformation and many more methods are used for texture feature extraction. The global features capture the gross essence of the shapes while the local features describe the interior details of the trademarks. In the feature extraction part total 45 features have been extracted from Knee MRI images. Out of which 19 are DICOM images header features [6], 13 are haralick texture features [7] and rest are images statistical features [8]. DICOM header features are extracted out in order to check either all images have been taken under similar environmental conditions or not.

1.3 Classification

Classification is a data mining (machine learning) technique used to predict group membership for data instances Classification is a problem of detecting the classes of data with the help of some already known classes. This is also called as supervised classification Thus the requirement is that new individual items are placed into groups based on quantitative information on one or more measurements, traits or characteristics, etc. and based on the training set in which previously decided groupings are already established. On the other hand the classifictaion in which no expert is present for prediction is called as Unsupervised Classification.
A model is defined in which with the help of set of attributes classes are defined. There are various classifiers and algorithms available. Weka is macning learing tool available which is used for the classification [9].

2. Implementation

Knee MRI scans has been collected and after image processing total 46 features have been extracted from the Knee MR images .A database file of 704 tuples and 46 attributes has been made in ASCII in CSV format, then conversion of this file to CSV file is done. CSV files are readable in Weka. The generated CSV file is opened in Weka and then different processes like data cleaning, data processing and data transformation are applied on to the input database file. These steps act as pre-processing steps for the classification of data. Along with this attribute removal is also done. Some of the attributes like patient's name, patient weight, body part examined etc are removed as they contribute nothing in classification process. chave been implemented and results have been compared on different parameters. Each of the method is tested at the Percentage split set to 70% on dataset.

2.1 Segmentation

2.1.1 Active Contour without edges

Distance map of initial mask: Chan and Vese have used phi0 which is a distance map of initial mask in his work, but we can modify it to phi1, because converting image to double and then adding to phi calculation does not make any difference. So we can emit this parameter and shorten the equation.

```
phi0 = bwdist(m)-bwdist(1-m)+im2double(m)-.5;
phi1 = bwdist(m)-bwdist(1-m)-.5;
subplot(2,2,1); plot(phi0),title('phi0'),subplot(2,2,2); plot(phi1),title('phi1');
```

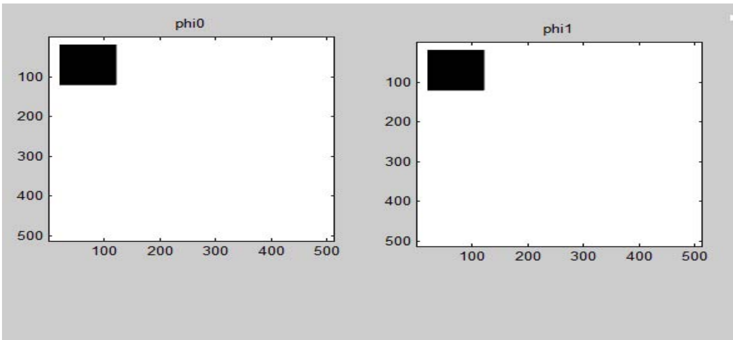


Figure 1: Distance Mask Of intial Mask

MRI images are larger in size. In order to process them it is needed to change the value of parameter μ if μ is small, then only smaller objects will be detected if μ is larger, then it will work for larger objects also or objects formed by grouping.

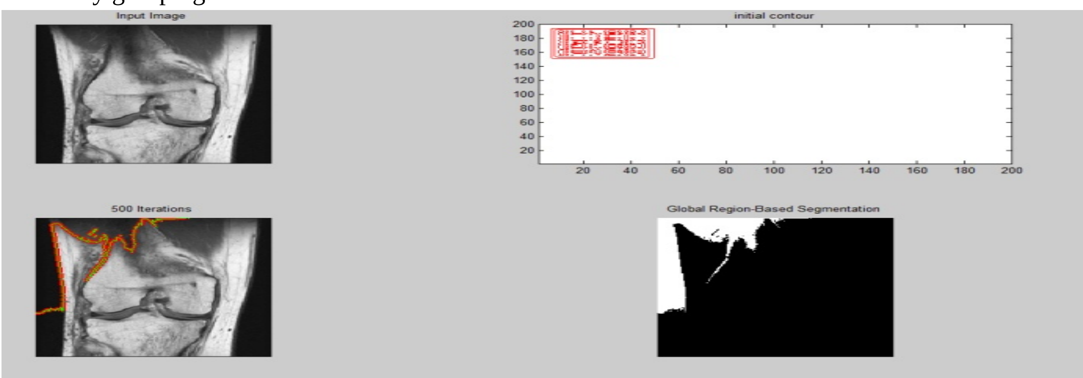


Figure 2: 'chan' or 'vector' method at dt=0.5

Increment the value of dt accordingly.
Dt=2.5

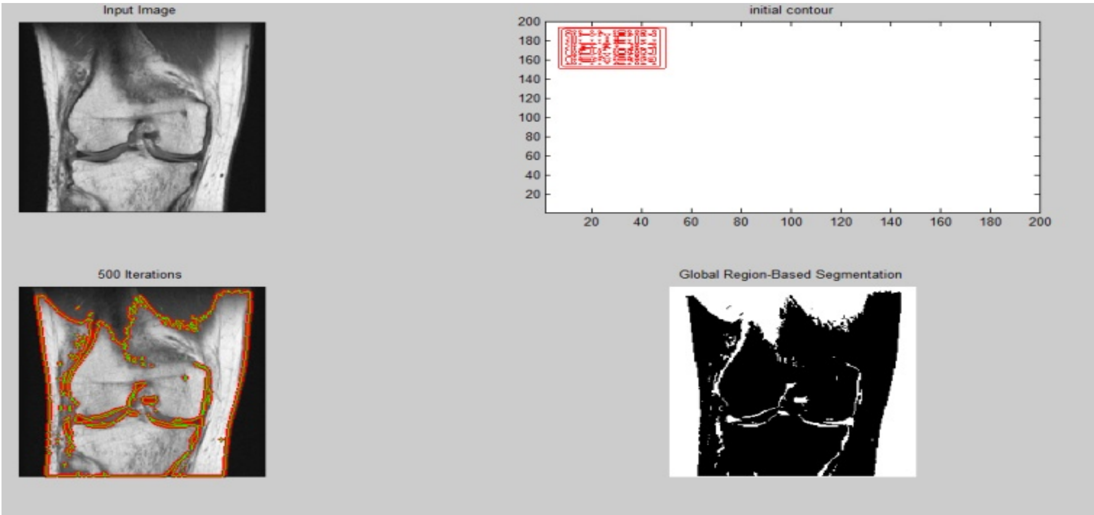


Figure 3: 'chan' or 'vector' method at dt=2.5

The output obtained by both 'chan' and 'vector' method is approximately same. But the difference in both the methods is this that if image is noisy then 'vector' method gives better result than the 'chan' method. So in the case of noisy images it is preferable to use 'vector' method.

2. 'Multiphase' method: In the case of 'multiphase method'

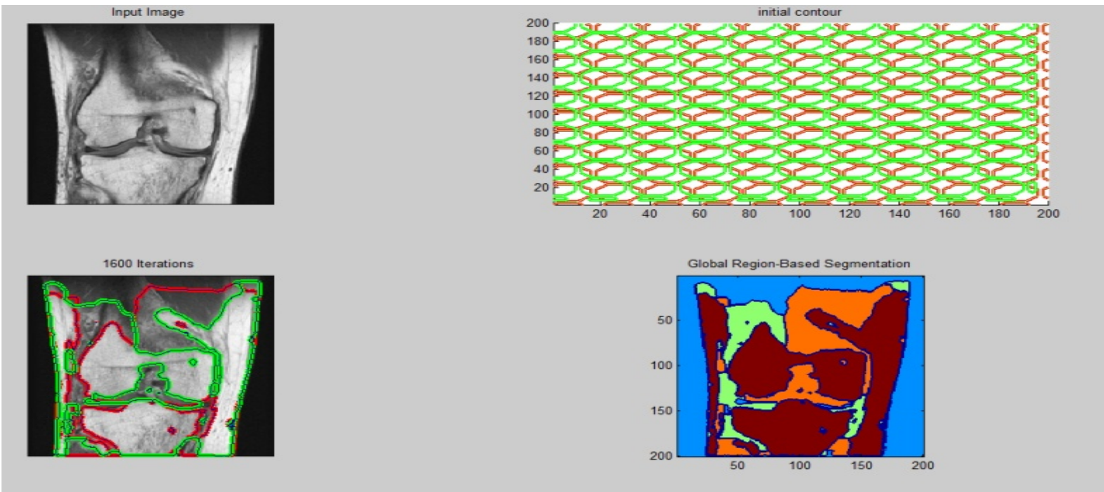
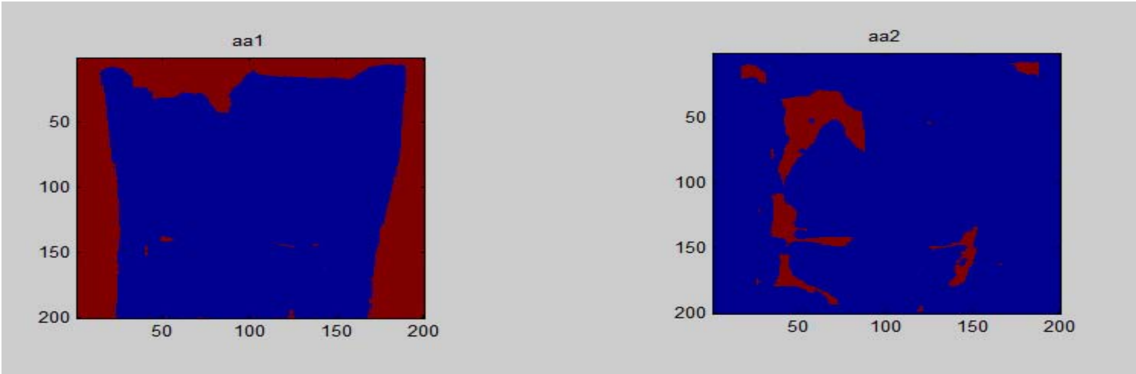
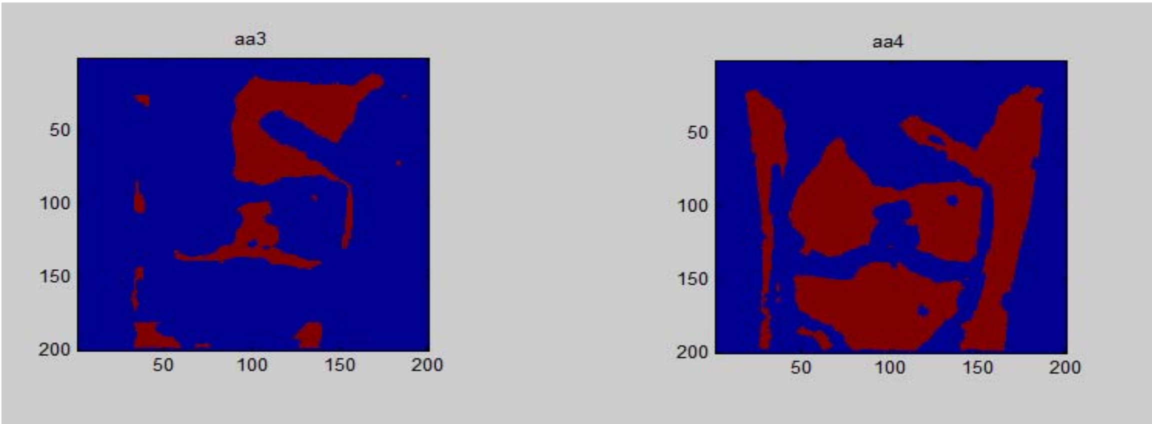


Figure 4: 'multiphase' method

Four regions aa1, aa2, aa3 and aa4 together form the complete segmented image. We can extract each region separately. By separating each region part we can get our region of interest without applying any other technique on segmented image.



(a) (b)



(c) (d)

Figure 5: Sub regions of Active Contour without edge

2.2 Feature Extraction

Following 46 features comprises feature vector:

1.File Size	17.Flip Angle	33. Area
2.Width	18.Rows	34. Mean
3.Height	19.Colums	35. Standard Deviation
4.Bit Depth	20.Angular Moment	36.Minimum
5.PatientName	21.Contrast	37.Maximum
6.Patient Birth Date	22.Correlation	38.X
7.Patient Sex	23.Entropy	39.Y
8.Patient’s Age	24.Inverse Difference Moment	40.Center of Mass X

9.Patient’s Weight	25.Sum Average	41.Center of Mass Y
10.Body part examined	26.Sum Variance	42.Perimeter
11.Slice Thickness	27.Sum Entropy	43.Int. Density
12.Image Frequency	28.Difference Average	44.Median
13.Image Nucleus	29.Differnce Variance	45.Skewness
14.Magnetic Field Strength	30.Differnce Entropy	46.Kurtosis
15.Spacing between Slices	31.Infoiormation of Correlation1	
16. Pixel Bandwidth	32.Infoiormation of Correlation2	

2.3 Classification

During the classification pre-processing of data is required to be done.

1. Discretization of data- when too many values exist for a attribute, then this requires performing discretization on numeric or continuous attributes. Perform discretization of all non-class attributes in the data set into 10 equal-width bins as follows: under “Filter” in the “Preprocess” window of the Explorer, select ‘filters’->’unsupervised’->’attribute’->’Discretize’. Use default parameters for the ’Discretize’ filter. After you make sure that all non-class attributes are nominal, perform classification on this set using the four classification methods and each of the evaluation strategies indicated above.

2.Attribute Removal- Many attributes does not contribute to classify the data, those attributes need to be removed before the classification algorithm because if they will be included then results may not be accurate and it takes a lot of time to produce the output.

Information gain algorithm is applied and attributes are arranged in the order of significance. Slice thickness attributes shows the highest priority. Classification is done on the base of attribute ‘slice thicknesses. Learning rate of four algorithms has been obtained by starting the training from 1% till 99%. It has been observed that all algorithms give almost constant value after 50% of percentage slit in training. Classification algorithms have been implemented on the 704 Knee MRI images dataset using the percentage split of 50% during training. Algorithms which are compared against different parameters are ZeroR, Naïve Bayes, Id3, FT and J48 algorithm. Values of different parameters have been compared in order to evaluate which algorithm gives better results in case of Knee MR Images.

2.3.1 Comparison of correctly and incorrectly classified instances of different algorithms

In order to find out which algorithm correctly classifies maximum number of instances. 50% of attributes have been taken in training phase and rest in the testing phase. Obtained values have been plotted below.

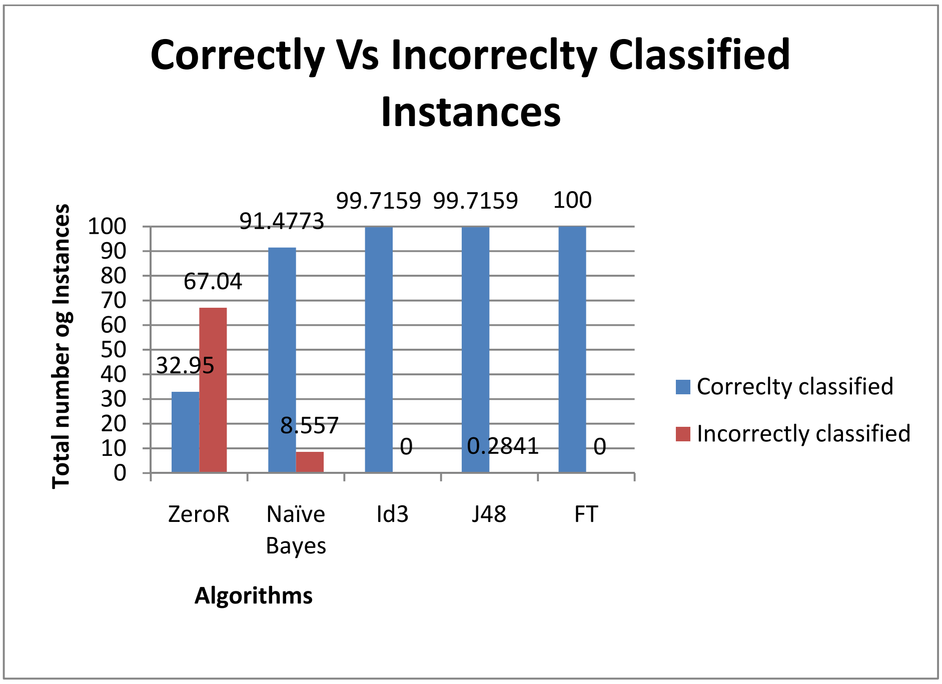


Figure 6: Correctly Vs Incorrectly Classified Instances

It has been observed from the plotted vales that FT correctly classified 100% of instances whereas ZeroR classifies 32.95%, Naïve Bayes classifies 91.4773%, Id3 classifies 99.7159% and J48 classifies 0.2841% of instances. FT gives maximum number of correctly classified instances and ZeroR gives minimum number of correctly classified instances.

2.3.2 Comparison of Root Mean Square Error

Root means square error of different algorithms has also been compared. Calculated values have been plotted below.

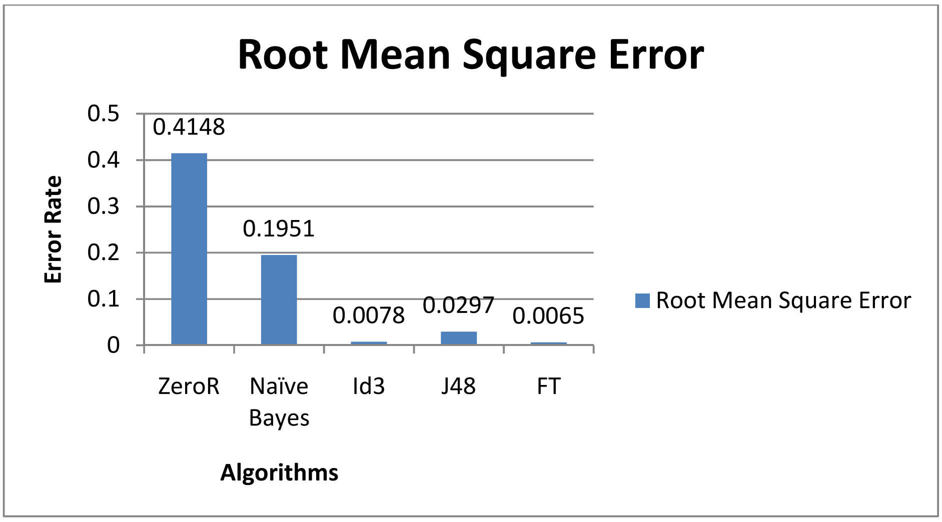


Figure 7: Comparison of Root Mean Square Error

It has been observed from the plotted values that FT gives minimum value and ZeroR gives maximum value of Root Mean Square Error.

3. Conclusion

Magnetic Resonance Imaging used for Knee scans is very useful and effective technique to detect the knee joint defects. It is a non-invasive method to take picture of knee joint and the surrounding images. Real Knee MRI data has been collected from MRI centers. It has been observed that Active Contour without Edges gives better segmentation results as different regions can be easily extracted out. In the next phase, total 46 features have been calculated and in the pre-processing 5 features which give the detail of patient's personal data have been removed. A database file consisting of 704 images with 41 lists of attributes is prepared and it used for classification process in next phase. Classification is implemented using ZeroR, Id3, J48, FT and Naïve Bayes. It can be concluded that FT gives maximum number of correctly classified instances and minimum values of Root Mean Square Error as compared to the other algorithms.

4. Future work

The thesis work is conducted only on 704 knee MR Images. Database can be extended and same methodology can be applied to the database containing images in thousands and many more. Only MRI Knee data has been used, the same approach can be extended to different medical imaging technologies like CT scan etc. Different segmentation algorithms and classification algorithms can be used for segmentation.

5. References

[1] Wu, X. Kumar, V. Ross Quinlan, J. Ghosh, J. Yang, Q. Motoda, H. McLachlan, G. J. Ng, A. Liu, B. Yu, P. S., "Top 10 algorithms in data mining", Knowledge and Information Systems, vol 14; number 1, pages 1-37, ISBN:0219-1377, Springer, 2008.

[2] Fayyad, U. M., G. P. Shapiro, P. Smyth.,"From Data Mining to Knowledge Discovery in Databases", AI Magazine, vol 17, number 3, pages 37-54, ISBN: 0738-4602, American Association of Artificial,1996.

[3] Qi Luo,Wuhan,"Advancing Knowledge Discovery and Data Mining", Procedings of Knowledge Discovery and Data Mining,IEEE WKDD, pages 3-5, ISBN: 978-0-7695-3090-1, 2008

[4] T. Chan and L. Vese, "Active contours without edges," in IEEE Trans on Image Processing Vol 10,pages 266-278, ISBN: 1057-7149, 2001.

[5] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems", Commun. Pure Appl. Math, vol. 42, pages 577–685, 1989.

[6] Rosset, A. Spadola, L. Rati., "OsiriX: An Open-Source Software for Navigating in Multidimensional DICOM Images", JOURNAL OF DIGITAL IMAGING, vol 17, part 3, pages 205-216, ISBN: 0897-1889 , W B SAUNDERS CO, 2004.

[7] R. M. Haralick and K. Shanmugam, "Computer Classification of Reservoir Sandstones," IEEE Transactions on Geoscience Electronics, vol. 11, pages. 171-177, 1973.

[8] Brandt, S. Laaksonen, J. Oja, E., "Statistical Shape Features in Content-based Image Retrieval", International Conference On Pattern Recognition, vol 15; vol 2, pages 1062-1065, ISBN: 1051-4651,2000.

[9] Holmes, G.; Donkin, A.; Witten, I.H., "WEKA: a machine learning workbench" , Intelligent Information system, proceedings of second Australian and new Zealand conference, pages 357-361, ISBN: 0-7803-2404-8 , dec 1994.