

# Fuzzy-based Medical X-Ray Image Classification

Fatemeh Ghofrani, Mohammad Sadegh Helfroush, Mahmoud Rashidpour, Kamran Kazemi

Department of Electrical and Electronic Engineering, Shiraz University of Technology, Shiraz, Iran

Submission: 28-01-2012 Accepted: 31-03-2012

## ABSTRACT

In this paper, a novel fuzzy scheme for medical x-ray image classification is presented. In this method, each image is partitioned into 25 overlapping subimages and then, we extracted the shape-texture features from shape and directional information of each subimage. In the classification step, we apply a fuzzy membership to each subimage considering the Euclidean distance between feature vector of each subimage and average of feature vectors of training subimages. At last, a hard classification of the test image can be obtained by performing a max operation on the summation of fuzzy memberships. The proposed method is evaluated for image classification on 2655 radiographic images from IRMA dataset with 300 training samples and 2355 test samples. Classification accuracy rates obtained by fuzzy classifier are higher than that of obtained by multilayer perceptron or even SVM classifier.

**Key words:** Fuzzy classifier, medical x-ray images, multilayer perceptron, shape-texture features, support vector machines

## INTRODUCTION

By increasing uses of medical X-ray images for diagnosing, efficient medical images retrieval becomes a necessity. Classification is one of the most important stages of image retrieval. Thus a certain classification plays a fundamental role in development of image retrieval.

Several medical X-ray image classification schemes are presented in the literature.<sup>[1-5]</sup> In 2003, Keysers *et al.*<sup>[6]</sup> developed a new content-based medical image retrieval scheme. In this method, images have been classified based on image modality, body orientation, anatomic region, and biological system. Performance of this classifier has been evaluated on 1617 training images and 332 test images from six classes of IRMA dataset and an error rate of 8% was obtained. In 2007, a new image classification method has been proposed by Mueen *et al.*<sup>[7]</sup> In this method, multilevel features have been extracted from 9000 training images and 1000 test images. Then Support Vector Machines (SVM) have been utilized in the classification stage. It should be noted that a support vector machine learns the decision surface from two distinct classes of the input points.<sup>[8]</sup> The accuracy rate of this method for 57 classes was 89%. In 2009, Iakovidis *et al.*<sup>[5]</sup> presented a content-based medical image retrieval scheme. This method that utilizes similarity measures, defined over higher level patterns that were associated with clusters of low level image feature spaces. Performance of

this scheme has been evaluated on 9000 training images and 1000 test images from 116 classes of IRMA dataset. The accuracy rate of this scheme on a subset of available data, generated according to the guidelines provided in<sup>[1]</sup> was 78%. In this method, training images were registered in the database. So some additional computation for registration of training images was necessary and it may not be so attractive. Also, a large number of training images was used in this method. With respect to this point that labeling of a large number of samples is expensive and time consuming, and majority of proposed methods for x-ray image classification are based on a large number of training samples, the basic motivation for medical x-ray image classification using a small number of training samples still remains as a challenge.

Recently, fuzzy set theory is used in many applications<sup>[9]</sup> and particularly in the field of pattern clustering and classification tasks.<sup>[10-12]</sup> Recent work in fuzzy-based medical image classification schemes has focused on extracting the fuzzy features. For example, Iakovidis *et al.*<sup>[13]</sup> developed a novel approach for thyroid ultrasound pattern representation. In this approach, a feature extraction scheme based on fusion of a fuzzy distribution of local binary patterns and ultrasound echogenicity represented by the fuzzy gray-level histogram is presented. In Ref. [14] a new method of content-based radiology medical image retrieval is presented. The description of images, in their work, relies on a fuzzy rule based Compact Composite

### Address for correspondence:

Corresponding author at :Department of Electrical and Electronic Engineering, Shiraz University of Technology, Shiraz, Iran.  
E-mail: f.ghofrani@sutech.ac.ir

Descriptor (CCD), which includes global image features capturing both brightness and texture characteristics in a 1D Histogram. These approaches do not take advantages of fuzzy set theory in the classification stage. In this paper, in order to obtain a good performance, we propose a classifier based on fuzzy set theory.

The aim of this paper is to introduce a new efficient scheme for medical x-ray image classification. Unfortunately, a typical characteristic of medical x-ray images is their large variations *within* a class, and also strong visual similarities across *different* classes.<sup>[1]</sup> Hence, the determination of a category is, in general, a difficult task. It is fortunate that in this case, fuzzy set theory allows us to easily determine the degree of membership of each image to different categories. Considering these points, in this paper a fuzzy-based scheme for medical x-ray image categorization is proposed. The results show that a fuzzy classifier performs more efficient than well-known classifiers such as multilayer perceptron (MLP) and even support vector machines (SVM).

The paper is organized as follows. Section II presented the procedure of feature extraction. In Section III, we described our fuzzy classification scheme. Experimental results are shown in section IV. Finally, conclusions are summarized in section V.

## Feature Extraction

Texture, shape, color, and spatial location features are often used in pattern classification. Since x-ray images are gray level and their texture characteristic is very similar, so color and texture features may not be suitable for medical x-ray images classification. In this paper, a combination of shape and texture features is used. Feature extraction algorithm diagram is shown in Figure 1. The features extraction procedure is described as follows:

### Histogram Adjustment

In this stage, the contrast of input images is improved by mapping the intensity values of each image to new values, such that 1% of data are saturated at low and high intensities of original image.

### Image Denoising

In images in which the brightness gradient generated by the noise is greater than that of the edges, and the level of the noise varies significantly across the image, a global noise estimate does not provide an accurate local estimate and the local value of the gradient provides too partial a piece of information for distinguishing noise-related and edge-related gradients.<sup>[15]</sup> Hence, in order to remove noise from images, the anisotropic diffusion filter that was proposed by Perona *et al.*,<sup>[16]</sup> is applied. They put forward an anisotropic diffusion (AD) equation to smooth a noisy image that is given by the expression

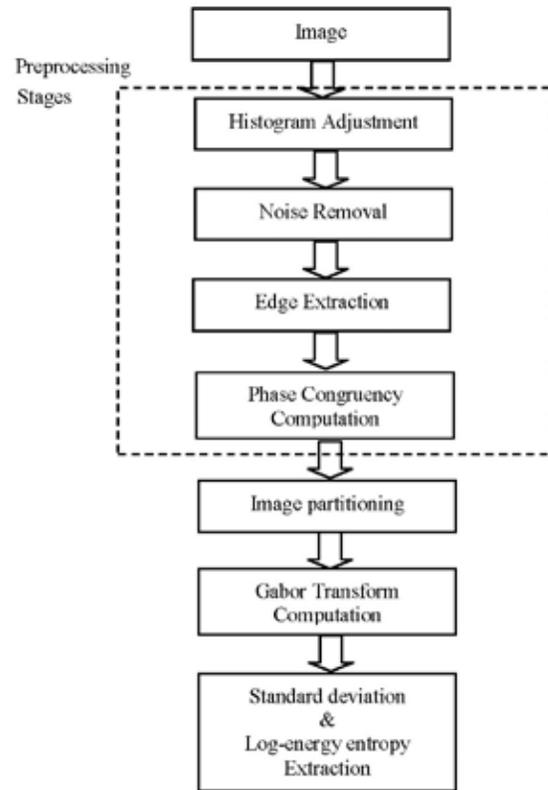


Figure 1: Feature extraction algorithm diagram

$$\frac{\partial u(x, y, t)}{\partial t} = \text{div}(g(|\nabla u(x, y, t)|)\nabla u(x, y, t)), \quad (1)$$

where  $u(x, y, t) : \Omega \times [0, +\infty) \rightarrow R$  is a scale image obtained by convolving the original image  $u_0(x, y)$  with a Gaussian kernel  $G(x, y; t)$  of variance  $t$ ,  $g(|\nabla u|)$  is a decreasing function of gradient.

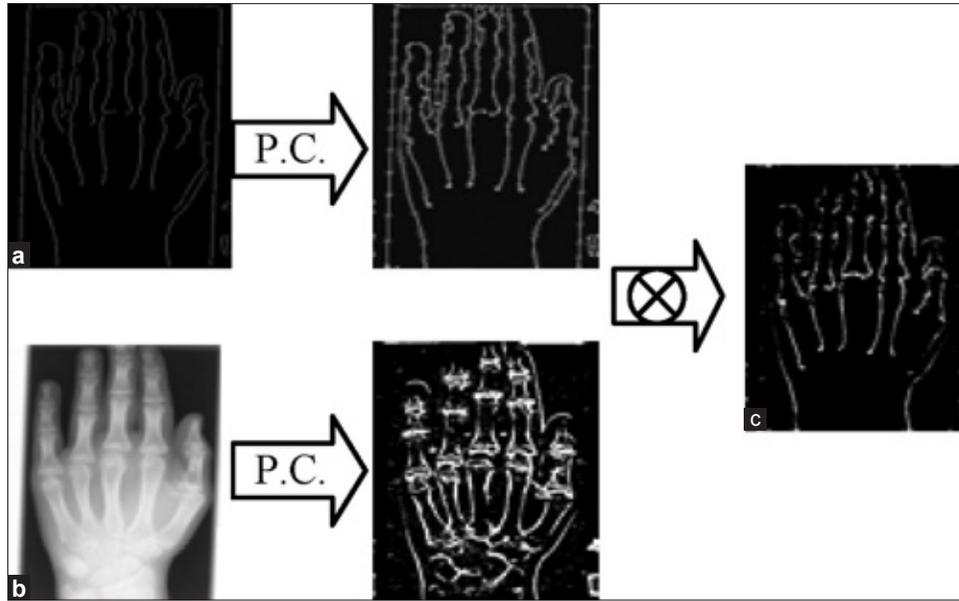
### Detecting Edge Using Canny Edge Detection

The Canny method<sup>[17]</sup> finds edges by looking for local maxima of the gradient of each image. The gradient is obtained using the derivative of a Gaussian filter. The method uses two thresholds to detect strong and weak edges, and includes the weak edges in the output only if they are connected to strong edges. This method is therefore less likely than the others to be fooled by noise, and more likely to detect true weak edges.

### Calculating Phase Congruency

As shown in Figure 2, to remove unnecessary edges and boundaries and also to increase the edge features accuracy, the phase congruency of shape image and original image are computed.<sup>[19]</sup> The phase congruency of original image and the resulting image of previous step is calculated. Then these two resulting images are multiplied. The measure of phase congruency proposed by Morrone *et al.*<sup>[20]</sup> is

$$PC_1(x) = \frac{|E(x)|}{\sum_n A_n(x)}, \quad (2)$$



**Figure 2:** Phase congruency computation results; (a) Computing phase congruency of edge images; (b) Computing phase congruency of original images; (c) Multiplication results of edge images<sup>[18]</sup>

where  $|E(x)|$  (local energy), is the magnitude of the vector from the origin to the end point and  $A_n(x)$  is the amplitude of local, complex valued, Fourier components at a location  $x$  in the signal.

### Partitioning Image into 25 Subimages

In this stage, each image is partitioned into 25 subimages. In order to keep the information of boundaries of subimages, image partitioning into overlapping subimages is used. Note that by using the overlapping subimages scheme, each image can be partitioned into  $(2n+1)^2$  subimages for  $n=1,2,3,\dots$  and in this paper, partitioning into 25 subimages ( $n=2$ ) is proposed, because 9 subimages ( $n=1$ ) cannot illustrate details clearly and by using the 49 subimages ( $n=3$ ), the number of extracted features will be increased.

### Computing Discrete Gabor Transform

After preprocessing stages, we compute Gabor transform of each subimage. Gabor filters are a group of wavelets, with each wavelet capturing energy at a specific frequency and a specific direction.<sup>[18]</sup> Frequencies and directions utilized in this paper are shown in Figure 3. For a given image  $I(x,y)$  with size  $P \times Q$ , discrete Gabor transform<sup>[18]</sup> is obtained using the following expressions:

$$G_{m,n}(x,y) = \sum_s \sum_t I(x-s, y-t) \varphi_{mn}^*(s,t), \quad (3)$$

where  $s$  and  $t$  are the filter mask size variables, and  $\varphi_{mn}^*$  is the complex conjugate of  $\varphi_{mn}$  which is a class of self-similar functions generated from expansion and rotation of the following mother wavelet:

$$\varphi_{mn} = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \cdot \exp(j2\pi\omega x), \quad (4)$$

where  $\omega$  is the modulation frequency. The self-similar Gabor wavelets are obtained through the generating function:

$$\varphi_{mn}(x,y) = a^{-m} \varphi(\tilde{x}, \tilde{y}), \quad (5)$$

where  $m$  and  $n$  indicate the scale and direction of the wavelet respectively, with  $m=0,1,\dots,M-1$ ,  $n=0,1,\dots,N-1$ , and

$$\begin{aligned} \tilde{x} &= a^{-m}(x \cos \theta + y \sin \theta), \\ \tilde{y} &= a^{-m}(-x \sin \theta + y \cos \theta), \end{aligned} \quad (6)$$

where  $a > 1$  and  $\theta = \frac{n\pi}{N}$ . The variables in the previous equations are defined as follows:

$$\begin{aligned} a &= \left(\frac{U_h}{U_l}\right)^{\frac{1}{M-1}}, \\ W_{m,n} &= a^m U_l, \\ \sigma_{x,m,n} &= \frac{(a+1)\sqrt{2 \ln 2}}{2\pi a^m (a-1)U_l}, \\ \sigma_{y,m,n} &= \frac{1}{2\pi \tan\left(\frac{\pi}{2N}\right)} \sqrt{\frac{U_h^2}{2 \ln 2} - \left(\frac{1}{2\pi\sigma_{x,m,n}}\right)^2}. \end{aligned} \quad (7)$$

In our implementation, the following values are used for mentioned parameters as commonly used in the literature:  $U_l = 0.05$ ,  $U_h = 0.4$ ,  $M=3$ ,  $N=6$ ,  $s$  and  $t$  range from 0 to 33, i.e., filter mask size is  $33 \times 33$ .

### Extracting Shape- texture features

Finally, in the last stage of feature extraction, two features are extracted from any filtered subimage. These features are calculated as follows:<sup>[19]</sup>

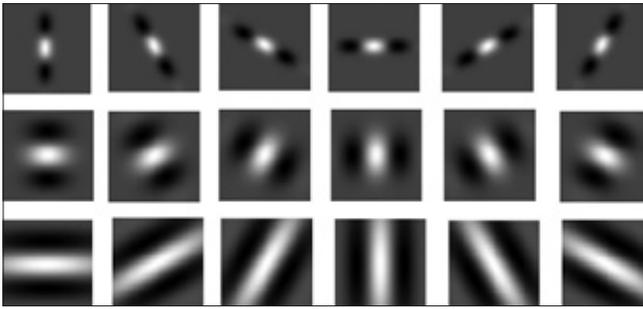


Figure 3: Frequencies and directions of Gabor filters used in this work

$$std = \sqrt{\frac{1}{N} \sum_{i=0}^{P-1} \sum_{j=0}^{Q-1} (I(i, j) - E)^2},$$

$$E = \frac{1}{N} \sum_{i=0}^{P-1} \sum_{j=0}^{Q-1} I(i, j), \tag{8}$$

where  $I(i, j)$  is the result of applying Gabor transform to each subimage with size  $P \times Q$  :

$$Log\text{-energy entropy} = -\sum_{i=0}^{P-1} \sum_{j=0}^{Q-1} (A(i, j))^2. \tag{9}$$

Discrete wavelet transform is applied to  $I(i, j)$  and finally  $A(i, j)$  are coefficients of the low-frequency subband in an orthonormal basis.

### PROPOSED CLASSIFICATION METHOD

A new scheme for the fuzzy classification of medical x-ray images is proposed in the present paper. The proposed fuzzy classification scheme extracts information of different parts of input pattern relating to each class. Since all parts of an image are not equally important in discriminating all classes, the partwise membership is expected to help in the classification performance. An illustration of the proposed classification method is shown in Figure 4.

The proposed scheme works in two steps. In the first step, the system takes a feature vector of input pattern and fuzzifies its different parts using a membership function, and provides the membership of individual parts to different classes. A membership matrix thus obtained contains number of rows and columns equal to the number of subimages or parts and classes, respectively. In the present study, we have used a membership function [Figure 5] to fuzzify an input pattern. Thus, the first step of the proposed fuzzy classification scheme extracts the hidden information of different subimages to all classes that may be helpful for achieving better classification results. The advantage of using this type of membership function is that it has two fuzzification parameters, which can be tuned to produce the best classification results.

The second step of the proposed fuzzy classifier is a hard classification by applying a max operation to defuzzify the

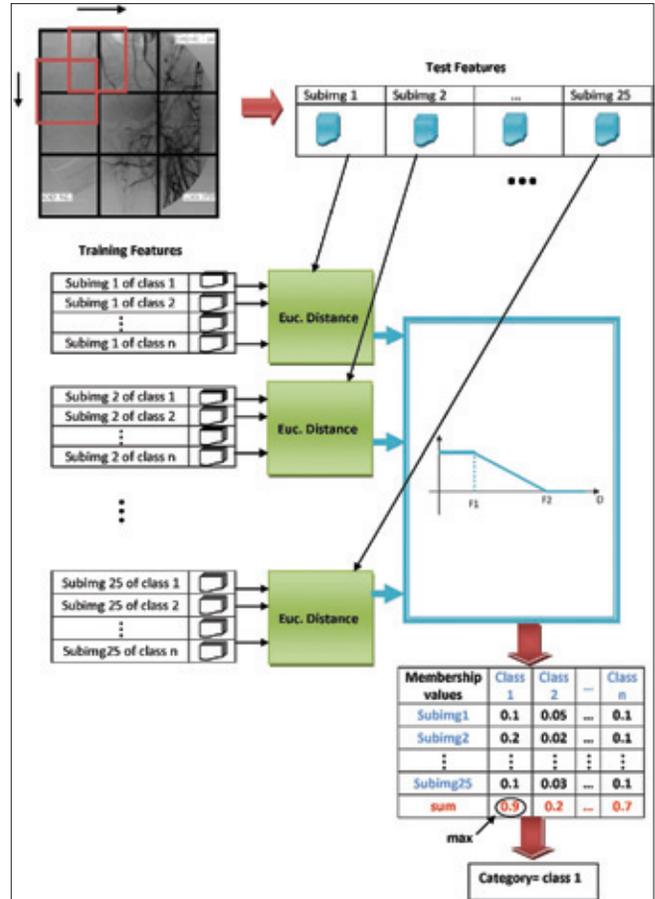


Figure 4: Proposed classification method

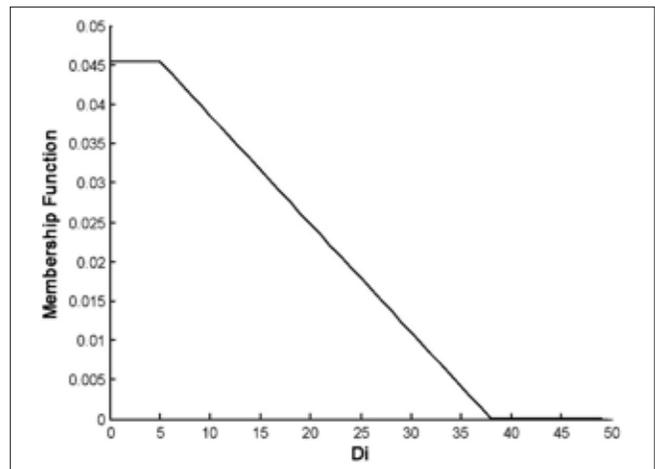


Figure 5: Membership function as a function of  $D_i$

output results. An image is assigned to the class which has the highest membership value.

### Fuzzification

The membership function generates a partwise degree of similarity of an image to different classes by fuzzification. Here we have utilized a membership function to model a class according to a Euclidean distance. The distance

is between feature vector of any subimage and average of feature vectors of training subimages for fuzzification process. By varying the values of the fuzzification parameters  $(F_1, F_2)$ , the steepness of the membership function can be controlled. This function is defined as shown in Figure 5

$$m(k,i) = \begin{cases} 1 & \text{if } D_{k,i} \leq F_1 \\ \frac{F_2 - D_{k,i}}{F_2 - F_1} & \text{if } F_1 < D_{k,i} < F_2 \\ \varepsilon & \text{if } D_{k,i} \geq F_2 \end{cases} \quad (10)$$

where  $\varepsilon$  is a very small positive value. To calculate  $D_{i,k}$ , suppose  $x_1$  and  $x_2$  are feature vectors of subimage  $i$  and the average of feature vectors of training subimages in class  $k$  with fixed size  $N$ , respectively. Then Euclidean distance between  $x_1$  and  $x_2$  or  $D_{i,k}$  is defined as follows:

$$D_{k,i}(x_1, x_2) = \sum_{l=1}^N (x_1^l - x_2^l)^2 = (x_1 - x_2)^T (x_1 - x_2). \quad (11)$$

Finally, the normalized membership function of any subimage can be defined as

$$\mu(k,i) = \frac{1}{\sum_k m(k,i)} \cdot m(k,i) \quad (12)$$

In the present study, it should be noted that the fuzzification parameters are tuned to achieve the best performance. We have selected the fuzzification parameters  $F_1$  and  $F_2$  as 15 and 850 respectively, and  $\varepsilon = 10^{-4}$  is considered.

**Defuzzification**

The last step of the proposed fuzzy classification model is a hard classification by performing a max operation to defuzzify the

output results. The image is assigned to class  $k$  corresponding to the highest membership value. Mathematically, the class number of a test image is determined as follows:

$$p = \arg \max_k (\sum_i \mu(k,i)). \quad (13)$$

**EXPERIMENTAL RESULTS**

In this section, the result of implementation of proposed method and comparison with some other classification methods are presented.

The images utilized for implementation include 2655 radiographic images with different sizes (example images are represented in Figure 6). These images are from 20 classes of IRMA database. We have taken 300 and 200 images as training samples, e.g., 15 and 10 samples for each class, and 2355 test samples. These classes are listed in Table 1. Using a small number of training samples is the most important point of this work.

Here, we present our results at a global accuracy, meaning the performance on the complete dataset, but also at a class-specific level, where we average the classification rates of all classes evaluated separately. We follow this procedure, as we encounter an unbalance database.<sup>[2]</sup>

After some preprocessing, shape-texture features are extracted using Gabor filters. Gabor filters extract the features from midfrequency or higher bands. Hence, the features that are extracted from the filtered x-ray images could not be features with rich information content. On the other hand, the spectrum of extracted edges for an image is more spread than that of original image. Thus, the features achieved from filtered extracted edges will be more efficient. First, just log-energy entropy features were extracted and



Figure 6: (a) Some images with different sizes from the images archives. (b) Some images with bad conditions from the images archives

applied for classification method. Then, we extracted another feature, standard deviation, to improve the performance of this work. Tables 2 and 3 depict the classification results using log-energy entropy feature and its combination with standard deviation feature for 300 training data, respectively. Also, the performance comparison results for 200 and

100 training samples using combination of features is shown in Tables 4 and 5. In each case, in order to reduce the feature space dimensionality and computational complexity and also to increase their capability, the Principal Component Analysis (PCA)<sup>[21]</sup> algorithm is utilized. Also, to achieve the best classification results, the optimum length of the feature vector is selected by several iterations.

Table 1: X-ray image classes

Class number	Anatomic part	Direction
1	Pelvis (cardiovascular system)	Coronal
2	Left breast	Axial
3	Right breast	Axial
4	Knee	Axial
5	Elbow	Coronal
6	Elbow	Sagittal
7	Ankle joint	Coronal
8	Knee	Coronal
9	Pelvis (musculoskeletal system)	Coronal
10	Cervical spine	Coronal
11	Lumber spine	Coronal
12	Chest	Sagittal
13	Neuro cranium	Sagittal
14	Cranium	Coronal
15	Hand	Coronal
16	Chest	Coronal
17	Foot	Coronal
18	Thoracic spine	Sagittal
19	Ankle joint	Sagittal
20	Radio carpal joint	Sagittal

To evaluate proposed classification method, three different classifiers consist of polynomial and Gaussian kernel support vector machines and MLP are utilized to compare with proposed classifier. As SVM is basically bi-classifier, we can use it to classify data into multi-class by different ways such as one-against-all, one-against-one and hierarchical. In this work, one-against-one method<sup>[22]</sup> is used. Furthermore, in the SVM classifier, two parameters exist for polynomial and Gaussian kernels:  $C$  and  $\gamma$ . For a given problem, It is not known beforehand which  $C$  and  $\gamma$  are the best, so a good selection of these parameters must be done. A common strategy is to separate the data set into two parts, of which one is considered unknown. The prediction accuracy obtained from the unknown set more precisely shows the performance on classifying an independent data set. An improved version of this procedure is known as cross-validation. In v-fold cross-validation, we first divide the training set into v subsets of equal size. Sequentially one subset is tested using the classifier trained on the remaining v-1 subsets. Thus, each instance of the whole training

Table 2: Classification accuracy rates obtained by four different classifiers using log-energy entropy feature for 300 training images

Class number	MLP	SVM		Fuzzy classifier
		Polynomial kernel	Gaussian kernel	
1	25	100	100	95
2	89.09	94.54	94.54	96.36
3	88	86	86	88
4	48.75	60	60	70
5	65	65	65	68.33
6	48	60	60	49.33
7	30.83	73.33	73.33	79.16
8	68.23	71.76	71.76	62.35
9	55	58.5	59	81
10	51.5	76	76	80
11	74.76	76.2	76.2	87.62
12	58.6	56.4	56.8	65
13	40.77	75	74.61	75
14	98	96	96	70
15	20	80	80	74
16	66	94	94	76
17	54	54	54	56
18	50	80	80	66
19	62	86	86	68
20	28.57	65.71	68.57	77.14
Classwise accuracy rate	56.1	75.42	75.6	78.61
Global accuracy rate	56.22	70.10	70.23	80.38

Figures are in percentage

Table 3: Classification accuracy rates obtained by four different classifiers using combination of log-energy entropy and standard deviation features for 300 training images

Class number	MLP	SVM		Fuzzy classifier
		Polynomial kernel	Gaussian kernel	
1	90	100	100	95
2	72.72	96.36	96.36	100
3	72	96	96	100
4	33.75	61.25	61.25	75
5	10	70	70	70
6	48	50.66	50.66	50.66
7	53.33	78.33	78.33	86.66
8	72.94	82.35	82.35	72.94
9	39.5	77	77	83.5
10	71.5	86	86	92.5
11	69.52	87.14	87.14	91.42
12	62	82.4	82.4	85.8
13	61.53	86.53	86.53	90.76
14	60	96	96	88
15	30	72	72	80
16	74	96	96	76
17	20	74	74	70
18	36	84	86	82
19	28	90	90	82
20	40	77.14	74.28	74.28
Classwise accuracy rate	52.24	82.15	82.11	82.32
Global accuracy rate	57.11	82.08	82.08	84.3

Figures are in percentage

set is predicted once so the cross-validation accuracy is the percentage of data which are correctly classified. We applied a grid-search on  $C$  and  $\gamma$  using cross-validation. Various pairs of  $(C, \gamma)$  values are tried and the one with the best cross-validation accuracy is picked. It should be noted that trying exponentially growing sequences of  $C$  and is a practical method to identify good parameters (for example,  $C = 2^{-5}, 2^{-3}, \dots, 2^{-15}, \gamma = 2^{-15}, 2^{-13}, \dots, 2^3$  ).<sup>[23]</sup> Hence, in this paper, we use a grid search and find the best pairs of  $(C, \gamma)$  values with five-fold to obtain the best results for SVM classifier. These values are listed in Table 6.

Table 2 shows that for 300 training images and by using the log-entropy features, MLP provides a classwise accuracy 56.1% and the global accuracy as 56.22%. These values are 75.42% and 70.1% with polynomial kernel SVM and 75.6% and 70.23% with Gaussian kernel. The proposed fuzzy classifier provides an improved classwise accuracy and global accuracy values and they are 78.61% and 80.38%, respectively. Thus there is an increase of nearly 22% of

classwise accuracy by the proposed method, and the corresponding global accuracy value is increased by 24% compared to MLP. Similarly, an increase of nearly 3% and 10% in classwise and global accuracy are obtained compared with the SVM classifier, respectively. Also, an improvement with the proposed fuzzy classifier can be observed with combination of log-entropy and standard deviation features for 300 training data, i.e., with the proposed classifier, polynomial kernel SVM, Gaussian kernel SVM and MLP methods classwise accuracies are 82.32%, 82.15%, 82.08%, and 52.24%, respectively. Also there is an increase in global accuracy of 3% with the proposed method compared with SVM and of 27% with MLP. From these values, it is obvious that the proposed classifier provides a better global and classwise accuracy compared with the SVM and MLP [Table 3].

The superiority of the proposed fuzzy classifier is also validated with 200 and 100 training images as shown in Tables 4 and 5, respectively. A comparison is made among

**Table 4: Classification accuracy rates obtained by four different classifiers using combination of log-energy entropy and standard deviation featears for 200 training images**

Class number	MLP	SVM		Fuzzy classifier
		Polynomial kernel	Gaussian kernel	
1	70	95	95	97.5
2	72.2	90.9	90.9	100
3	72	90	90	100
4	25	53.75	53.75	81.25
5	11.66	63.33	63.33	56.66
6	12	38.66	38.66	25.33
7	20	70.83	70.83	79.16
8	47.64	81.17	81.17	68.82
9	28.5	81	82.5	89
10	30	83.5	84	87.5
11	30.47	83.8	83.8	84.3
12	29	68.6	68.8	67.2
13	13.1	75	75	88.46
14	52	94	94	84
15	26	82	82	76
16	76	96	96	86
17	64	56	56	62
18	32	84	84	64
19	36	82	82	72
20	5.7	48.57	48.57	62.85
Classwise accuracy rate	37.7	75.9	76.01	76.6
Global accuracy rate	31.84	75.28	75.49	77.02

Figures are in percentage

**Table 5: Classification accuracy rates obtained by four different classifiers using combination of log-energy entropy and standard deviation featears for 100 training images**

Class number	SVM		Fuzzy classifier
	Polynomial kernel	Gaussian kernel	
1	95	95	95
2	89.09	89.1	94.54
3	84	84	98
4	58.75	60	67.5
5	53.33	53.33	58.33
6	34.66	33.33	52
7	47.5	47.5	51.66
8	52.94	53.52	28.82
9	84	84.5	91
10	73	73	87
11	78.1	78.57	76.66
12	63.4	63.8	92.4
13	42.69	42.7	78.46
14	92	92	84
15	86	86	88
16	94	96	80
17	40	40	26
18	92	92	78
19	80	80	64
20	34.28	34.28	60
Classwise accuracy rate	68.73	68.93	72.57
Global accuracy rate	65.43	65.69	76.1

Figures are in percentage

**Table 6: The best selected parameters for SVM classifier using the grid search**

Feature name	No. of training images	C value	$\gamma$ value	Cross-validation (%)
Log-entropy	300	$2^6$	$2^{-9}$	74.66
Log-entropy and standard deviation	300	$2^5$	$2^{-7}$	82.33
Log-entropy and standard deviation	200	$2^5$	$2^{-9}$	79.5
Log-entropy and standard deviation	100	$2^5$	$2^{-7}$	53.75

**Table 7: Summarized results obtained by four different classifiers for 300 training images**

Classification method	Feature name		Accuracy rate	
	Log-energy entropy	Standard deviation	Classwise	Global
MLP	*		56.1	56.22
Polynomial kernel SVM	*		75.42	70.1
Gaussian kernel SVM	*		75.6	70.23
Proposed method	*		78.61	80.38
MLP	*	*	52.24	57.11
Polynomial kernel SVM	*	*	82.15	82.08
Gaussian kernel SVM	*	*	82.11	82.08
Proposed method	*	*	82.32	84.3

Figures are in percentage

the four classifier mentioned earlier. Table 4 shows that the classwise accuracy with the proposed classifier using combination of features and 200 training images is 76.6%, which is more than 76.01%, 75.9%, and 37.7% obtained using Gaussian kernel SVM, polynomial kernel SVM, and MLP, respectively. Also, an improvement of global accuracy value is obtained by the proposed classifier. For example, the global accuracy values are 77.02%, 75.49%, 75.28%, and 31.54%, respectively. In addition, Tables 4 and 5 also reveal that the global accuracy rate obtained with the proposed classifier for 100 training data is nearly the same as with the SVM at 200 training data. This is particularly important when there is a scarcity of training images.

From the classification results of four different classifiers, it is observed that for all cases the proposed fuzzy classifier performed better than the MLP and even SVM with optimum parameters. For 300 training images and by using combination of features, the summarized results are shown in Table 7.

## CONCLUSION

A novel fuzzy scheme for medical x-ray image classification has been presented. In order to improve performance of medical x-ray image classification, an effective fuzzy classifier was introduced. The experimental results demonstrated the efficiency of proposed scheme. Future works involved evaluating performance of fuzzy SVM and fuzzy Neuron Network classifiers for medical x-ray image classification.

## ACKNOWLEDGMENT

The authors would like to thank TM Lehmann, Dept. of Medical Informatics, RWTH Aachen, Germany, for making the database available for the experiments.

## REFERENCES

- Greenspan H, Pinhas AT. Medical image categorization and retrieval for PACS using the GMM-KL framework. *IEEE Trans Inf Technol Biomed* 2007;11:190-202.

- Jeanne V, Unay D, Jacquet V. Automatic Detection of Body Parts in XRay Images. Miami, FL, USA: Proc. of IEEE Computer Vision and Pattern Recognition (CVPR) Workshops; 2009. p. 25-30.
- Bertalya, Prihandoko, Kerami D, Kusuma TM. Classification of Xray Images Using Grid Approach. *Proc. Int. Conf. on Signal Image Technology and Internet Based Systems*. Washington DC, USA: SITIS; 2008. p. 314-9.
- Zhang Q, Tai X. The Xray Chest Image Retrieval Based on Feature Fusion. *Proc.Int. Conf. on Audio, Language and image processing*. Shanghai, China: ICALI; 2008. p. 899-903.
- Iakovidis DK, Pelekis N, Kotsifakos EE, Kopanakis I, Karanikas H, Theodoridis Y. A pattern similarity scheme for medical image retrieval. *IEEE Trans Inf Technol Biomed* 2009;13:442-50.
- Keyser D, Dahmen J, Ney H, Wein BB, Lehmann TM. Statistical framework for model-based image retrieval in medical applications. *J Electron Image* 2003;12:59-68.
- Mueen A, Baba MS, Zainuddin R. Multilevel feature extraction and x-ray image classification. *J Appl Sci* 2007;7:1224-9.
- Lin CF, Wang SD. Fuzzy support vector machines. *IEEE Trans Neural Netw* 2002;13:464-71.
- Morillas S, Gregori V, Hervas A. Fuzzy peer groups for reducing mixed gaussian-impulse noise from color images. *IEEE Trans Image Process* 2009;18:1452-66.
- Min R, Cheng HD. Effective image retrieval using dominant color descriptor and fuzzy support vector machine. *Int J Pattern Recognit* 2009;42:147-57.
- Iakovidis DK, Keramidas EG, Maroulis D. Fusion of fuzzy statistical distributions for classification of thyroid ultrasound patterns. *Artif Intell Med* 2010;50:33-41.
- Lin GC, Wang CM, Wang WJ, Sun SY. Automated classification of multispectral MR images using unsupervised constrained energy minimization based on fuzzy logic. *Magn Reson Imaging* 2010;28:721-38.
- Iakovidis DK, Keramidas EG, Maroulis D. Fusion of fuzzy statistical distributions for classification of thyroid ultrasound patterns. *Artif Intell Med* 2010;50:33-41.
- Chatzichristofis SA, Boutalis YS. "Content based radiology image retrieval using a fuzzy rule based scalable composite descriptor. *Springer. Multimed Tools Appl* 2010;46:493-519.
- Perona P, Malik J. Scale-Space and Edge Detection Using Anisotropic Diffusion. *IEEE Trans Pattern Anal Mach Intell* 1990;12:629-39.
- Fu S, Ruan Q, Wang W, Li Y. Adaptive Anisotropic Diffusion for Ultrasonic Image Denoising and Edge Enhancement. *Int J Inf Technol* 2006;2:284-92.
- Canny J. A computational approach to edge detection. *IEEE Trans Pattern Anal Mach Intell* 1986;8:679-98.
- Zhang D, Wong A, Indrawan M, Lu G. Content-based image retrieval using gabor texture features. Sydney, Australia: Proceedings of the First *IEEE Pacific-Rim Conference on Multimedia*; 2000. p. 392-95.
- Mohammadi SM, Helfroush MS, Kazemi K. Novel shape texture feature extraction for medical Xray image classification. *Int J Innov Comput Inf Control* 2012;8:659-76.
- Morrone MC, Ross JR, Burr DC, Owens RA. Mach bands are phase dependent. *Nature* 1986;324:250-3.
- Jain AK, Bhandrasekaran B. Dimensionality and sample size considerations in pattern recognition practice. In: Krishnaiah PR, Kanal LN, editors. vol. 2, chap 39. North Holland; 1982. p. 835-55.
- Kressel U. Pairwise classification and support vector machines. In: Scholkopf B, Burges CJ, Smola AJ, editors. *Advances in Kernel Methods: Support Vector Learning*. Cambridge, MA: MIT Press; 1999; p. 255-68.
- Hsu CW, Chang CC, Lin CJ. *A Practical Guide to Support Vector Classification*. Taiwan Department of Computer Science National Taiwan University, Taipei 106; 2010.

**How to cite this article: \*\*\***

**Source of Support: Nil, Conflict of Interest: None declared**

## BIOGRAPHIES



**Fatemeh Ghofrani** received the B.S. degree in Electrical engineering in 2009 from Shiraz University, Iran and the M.S. degree in Communication engineering from Shiraz University of Technology, Iran, in 2011 with First Class Honour. Her research interests include image processing, pattern recognition and machine learning.

**E-mail:** f.ghofrani@sutech.ac.ir



**Mohammad Sadegh Helfrroush** received the B.S. and M.S. degrees in Electrical engineering from Shiraz University of Shiraz and Sharif University of Technology, Tehran in 1993 and 1995, respectively. He performed his Ph.D. degree in Electrical Engineering from Tarbiat Modarres University, Tehran, Iran. Actually he is working as associate professor in department of Electrical and Electronics Engineering, Shiraz University of Technology, Shiraz, Iran.

**E-mail:** ms\_helfrroush@sutech.ac.ir



**Kamran Kazemi** received the B.S. and M.S. degrees in Electrical engineering from Ferdowsi University of Mashhad and K. N. Toosi University of technology, Tehran in 2000 and 2002, respectively. He performed his Ph.D. degree in Electrical Engineering

and Biomedical Engineering as cooperation between K. N. Toosi University of Technology, Tehran, Iran and University of Picardie Jules Verne (UPJV) Amiens, France. Actually he is working as assistant professor in department of Electrical and Electronics Engineering, Shiraz University of Technology, Shiraz, Iran and also as invited researcher in the GRAMFC (Groupe de Recherche sur l'Analyse Multimodale de la Fonction Cérébrale) lab at the University of Picardie Jules Verne (UPJV), Amiens, France.

**E-mail:** kazemi@sutech.ac.ir



**Mahmoud Rashidpour** received the B.S. degree in Electrical engineering and M.S. degree in Communication Engineering from Shiraz University, Iran. He performed his Ph.D. degree in Communication Engineering from University of Tehran, Iran in 2004. Actually he is working as assistant professor in department of Electrical and Electronics Engineering, Shiraz University of Technology, Shiraz, Iran.

**E-mail:** rashidpour@sutech.ac.ir