

SEGMENTATION AND CLASSIFICATION OF BRAIN TUMOR CT IMAGES USING SVM WITH WEIGHTED KERNEL WIDTH

Kimia Rezaei¹ and Hamed Agahi²

¹Corresponding author: Department of Electrical Engineering, Fars science and research branch, Islamic Azad University, Iran

²Associate professor, Department of Electrical Engineering, Shiraz branch, Islamic Azad University, Fars, Iran

ABSTRACT

In this article a method is proposed for segmentation and classification of benign and malignant tumor slices in brain Computed Tomography (CT) images. In this study image noises are removed using median and wiener filter and brain tumors are segmented using Support Vector Machine (SVM). Then a two-level discrete wavelet decomposition of tumor image is performed and the approximation at the second level is obtained to replace the original image to be used for texture analysis. Here, 17 features are extracted that 6 of them are selected using Student's t-test. Dominant gray level run length and gray level co-occurrence texture features are used for SVM training. Malignant and benign tumors are classified using SVM with kernel width and Weighted kernel width (WSVM) and k-Nearest Neighbors (k-NN) classifier. Classification accuracy of classifiers are evaluated using 10 fold cross validation method. The segmentation results are also compared with the experienced radiologist ground truth. The experimental results show that the proposed WSVM classifier is able to achieve high classification accuracy effectiveness as measured by sensitivity and specificity.

KEYWORDS

Brain tumor, Computed tomography, Segmentation, Classification, Support vector machine.

1. INTRODUCTION

A brain tumor or intracranial neoplasm occurs when some abnormal cells are shaped inside the brain. Two main types of tumors exist: malignant or cancerous tumors and benign tumors. Medical image processing has been developed rapidly in recent years for detecting abnormal changes in body tissues and organs. X-ray computed tomography (CT) technology uses computer-processed X-rays to produce tomographic images of a scanned object, which makes inside the object visible without cutting. CT images are most commonly used for detection of head injuries, tumors, and Skull fracture. Since various structures have similar radiodensity, there is some difficulty separating them by adjusting volume rendering parameters. The manual analysis of tumor based on visual interpretation by radiologist may lead to wrong diagnosis when the number of images increases. To avoid the human error, an automatic system is needed for

analysis and classification of medical images. Image segmentation is the process of partitioning a digital image into a set of pixels based on their characteristics and in medical images, texture contents are considered as pixels characteristics. There are various methods for segmentation. Here Support Vector Machine (SVM) with kernel function is constructed to segment the tumor region by detecting tumor and non-tumor areas. The segmentation results are obtained for the purpose of classifying benign and malignant tumors. Classification is the problem of identifying to which of a set of categories a new observation belongs, on the basis of a training set of data whose category membership had been defined. There are various algorithms for classification using a feature vector containing image texture contents. SVM, which is considered as a supervised learning system for classification, is used here.

2. LITERATURE SURVEY

There are a lot of literatures that focus on brain tumor CT images segmentation, classification and feature extraction. Padma et al. [1] proposed co-occurrence, gray level and new edge features by means of SVM classifier for segmentation of tumor from brain CT images. The method is applied on real data of 80 tumor images and it is inferred that better accuracy had been achieved compared with the fuzzy c-means clustering method. Nandpuru et al. [2] proposes SVM classification technique to recognize normal and abnormal brain Magnetic Resonance Images (MRI). First, skull masking applied for the removal of non-brain tissue like fat, eyes and neck from images. Then gray scale, symmetrical and texture features were extracted for classification training process. Rajini et al. [3] proposes a new approach for automatic diagnosis of normal and abnormal MR images which consists of two stages of feature extraction using discrete wavelet transformation and image classification by means of feed forward back propagation artificial neural network and k-nearest neighbors (k-NN) classifier. Sridhar et al. [4] proposed a method to classify 5 classes of Brain tumors MR images using Probabilistic Neural Network. Discrete Cosine Transform was applied for dimensionality reduction and extraction of 16 different features. Sundararaj et al. [5] used various intensity based and texture features such as skewness and coarseness. They constructed a linear SVM classifier for the diagnosis of normal and abnormal brain CT images. Padma et al. [6] used wavelet co-occurrence texture features by means of bidirectional associative memory type artificial neural network for the segmentation of tumor in brain CT images. They performed Genetic Algorithm for feature selection. The algorithm performance evaluation represents the outperformance of this method. Kaur, T et al. [7] proposed an automatic segmentation method on brain tumor MR images that performs multilevel image thresholding, using the spatial information encoded in the gray level co-occurrence matrix. Kaur, T et al. [8] proposed a technique which exploits intensity and edge magnitude information in brain MR image histogram and GLCM to compute the multiple thresholds. Verma, A. K. et al. [9] decomposed corrupted images using symlet wavelet then proposed a denoising algorithm utilizes the alexander fractional integral filter which works by the construction of fractional masks window computed using alexander polynomial.

The above literature survey illustrates that all the above methods are considered co-occurrence texture features only and some of the methods are proposed for the purpose of classification only and some for segmentation only.

3. MATERIALS AND METHODS

First, image noises are removed using median and wiener filter. Some features must be extracted from brain tumor images for the purpose of classifier training. Hence, a two-level discrete wavelet decomposition of tumor image is performed and the approximation at the second level is obtained to replace the original image to be used for texture analysis. Here, 17 features are extracted that 6 of them are selected using Student's t-test. Dominant gray level run length and gray level co-occurrence texture features are used for SVM training. Malignant and benign tumors are classified using SVM with kernel width and weighted kernel width (WSVM) and k-Nearest Neighbors (k-NN) classifier. The proposed methodology is applied to real brain CT images datasets collected from Shiraz Chamran hospital. All images are in DICOM format with a dimension of 512×512 . The proposed algorithm is implemented in Matlab software.

3.1 Image Pre-processing and Enhancement

Medical images corrupt through imaging process due to different kinds of noise. In pre-processing stage, noise and high frequency artifact present in the images are removed. The median filter is a nonlinear digital filtering method, often used for noise reduction on an image or signal [10]. This technique is performed to improve the results of later processing. Median filter is mostly used to remove noise from medical images. Wiener filter produces an estimate of a target random process by means of linear time-invariant filter [11]. Wiener filter is also a helpful tool for the purpose of medical images noise reduction. Here images noise removing process is carried out by using median filter and wiener filter.

3.2 Tumor Segmentation using SVM Classifier

In this paper, SVM classifier is chosen for tumor identification[12]. SVM is a machine learning technique combining linear algorithms with linear or non-linear kernel functions that make it a powerful tool for medical image processing applications. To apply SVM into non-linear data distributions, the data should be transformed to a high dimensional feature space where a linear separation might become feasible. In this study, a linear function is used.

Training an SVM involves feeding studied data to the SVM along with previously studied decision values, thus constructing a finite training set. To form the SVM segmentation model, feature vectors of tumor and non-tumor area, distinguished with the help of radiologist, are extracted. 25 points covering tumor area and 25 points covering the non-tumor area are selected. These points not only cover all the tumor and no-tumor areas but also are enough as an input for training a SVM classifier due to its powerful learning even through using few numbers of training inputs. For each point (pixel), two properties of position and intensity are considered to form the feature vector or training vector. Totally 50 feature vectors are defined as input to the SVM classifier to segment the tumor shape. Accordingly, there is a 25×3 matrix of tumor area and a 25×3 matrix of non-tumor area. In segmentation phase, matrix t is given as input to the SVM for training and pixels are labeled so that their classes can be designated.

$$t_i = (x_i, y_i, I_i(x_i, y_i)) \quad i = 1, \dots, 50 \quad (1)$$

i represent the number of training vectors. (x_i, y_i) and $I_i(x_i, y_i)$ represent the position and intensity of the selected points, respectively. Pixel selection using Matlab is displayed in Figure 1.

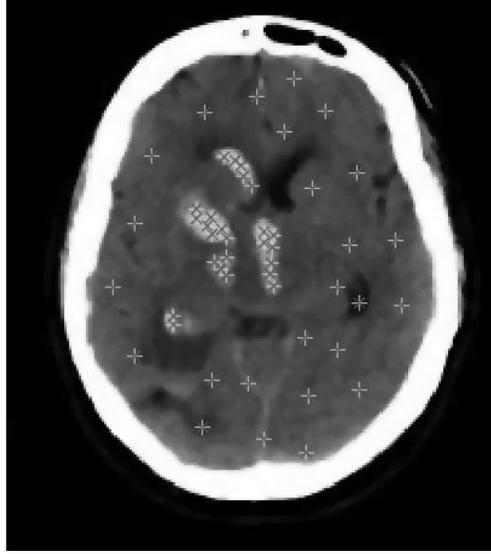


Figure 1. Pixel Selection using Matlab

3.3 Processing the Segmented Tumor Image on the Basis of 2D Discrete Wavelet Decomposition

Discrete Wavelet Decomposition is an effective mathematical tool for texture feature extraction from images. Wavelets are functions based on localization, which are scaled and shifted versions of some fixed primary wavelets. Providing localized frequency information about the function of a signal is the major advantage of wavelets.

Here a two-level discrete wavelet decomposition of tumor image is applied, which results in four sub-sets that show one approximation representing the low frequency contents image and three detailed images of horizontal, vertical and diagonal directions representing high frequency contents image [13]. 2D wavelet decomposition in second level is performed on the approximation image obtained from the first level. Second level approximation image is more homogeneous than original tumor image due to the removing of high-frequency detail information. This will consequence in a more significant texture features extraction process.

3.4 Feature Extraction

Texture is the term used to characterize the surface of an object or area. Texture analysis is a method that attempts to quantify and detect structural abnormalities in various types of tissues. Here dominant gray-level run length and gray-level co-occurrence matrix method is used for texture feature extraction.

The dominant gray-level run length matrix [14] is given as:

$$\varphi(d, \theta) = [p(i, j | d, \theta)] \quad 0 < i \leq N_g, \quad 0 < j \leq R_{\max} \quad (2)$$

Where N_g is the maximum gray level and R_{\max} is the maximum run length. The function $p(i, j|\theta)$ calculates the estimated number of runs in an image containing a run length j for a gray level i in the direction of angle θ . Dominant gray-level run length matrices corresponding to $\theta = 0^\circ, 45^\circ, 90^\circ$ and 135° are computed for approximation image derived from second level wavelet decomposition. Afterward, the average of all the features extracted from four dominant gray level run length matrices is taken.

A statistical method of analyzing texture considering the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM) [15]. The GLCM functions characterize the texture of the given image by computing how often pairs of pixel with certain values and in a specified spatial relationship occur in an image. The gray-level co-occurrence matrix is given as:

$$\varphi(d, \theta) = [p(i, j|d, \theta)] \quad 0 < i \leq N_g, \quad 0 < j \leq N_g \quad (3)$$

Where N_g is the maximum gray level. The element $p(i, j|d, \theta)$ is the probability matrix of two pixels, locating within an inter-sample distance d and direction θ that have a gray level i and gray level j . Four gray-level co-occurrence matrices, with $\theta = 0^\circ, 45^\circ, 90^\circ$ and 135° for direction and 1 and 2 for distance, are computed for approximation image obtained from second level wavelet decomposition. Then, 13 Haralick features [16] are extracted from each image's GLCM and the average of all the extracted features from four gray-level co-occurrence matrices is taken.

3.5 Feature Selection

Feature selection is a tool for transforming the existing input features into a new lower dimension feature space. In this procedure noises and redundant vectors are removed. Here, Two-sample Student's t-test is used for feature selection which considered each feature independently [17]. In this method, significant features are selected by computing the mean values for every feature in benign tumor class and malignant tumor class. Then, mean values of both classes are compared.

The T-test presumed that both classes of data are distributed normally and have identical variances. The test statistics can be calculated as follows:

$$t = x_b - x_m / \sqrt{\frac{\text{var}_b}{n_b} + \frac{\text{var}_m}{n_m}} \quad (4)$$

Where, x_b and x_m are mean values from benign and malignant classes. var_b and var_m represent variances of benign and malignant classes. n_b and n_m show the number of samples (images) in each class. This t value followed Student t-test with $(n_b + n_m - 2)$ degrees of freedom for each class.

In statistics, the p -value is a function of the observed sample results, used to test a statistical hypothesis and figuring out that the hypothesis under consideration is true or false. Here, the p -value is calculated based on test statistics and degrees of freedom [18]. Then, the optimal features are selected on the basis of the condition $P < 0.001$.

3.6 Classification Using k-NN Classifier

In this paper, the main objective of classification is the identification of benign and malignant tumors in brain computed tomography images. The k -nearest neighbor classifier is a nonparametric supervised classifier that performs propitious for optimal values of k . k -NN algorithm consists of two stages of training and testing. In training stage, data points are given in n -dimensional space [19]. These training data are labeled so that their classes can be specified. In the testing stage, unlabeled data are given as input and the classifier generates the list of the k nearest data points (labeled) to the testing point. Then the class of the majority of that list is identified through the algorithm.

k -NN algorithm:

1. Define a suitable distance metric.
2. In training step, all the training data set P are put in pairs.

$$P = \{(y_i, C_i), i = 1, \dots, n\} \quad (5)$$

Where y_i is a training pattern in the training data set, C_i is its class and n is the number of training patterns.

3. In testing step, the distances between testing feature vector and training data are computed.
4. The k -nearest neighbors are chosen and the class of the testing example is specified.

The result of classification in testing stage is used to evaluate the precision of the algorithm. If it was not satisfactory, the k value can be changed till achieving the desirable result.

3.7 Classification using SVM classifier

Support vector machine algorithm depends on the structural risk minimization principle. Compared with artificial neural networks, SVM is less computationally complex and function well in high-dimensional spaces. SVM does not suffer from the small size of training dataset and obtains optimum outcome for practical problem since its decision surface is specified by the inner product of training data which enables the transformation of data to a high dimensional feature space. The feature space can be defined by kernel function $K(x, y)$. The most popular kernel is the (Gaussian) radial basis function kernel, which is used here and is defined as follows:

$$k(x, y) = \exp(-\|x - y\|^2 / 2\sigma^2) \quad (6)$$

Where σ is the kernel width and chosen by the user. For the purpose of diminishing the coexisting over-fitting and under-fitting loss in support vector classification using Gaussian RBF kernel, the kernel width is needed to be adjusted, to some extent, the feature space distribution. The scaling rule is that in dense regions the width will be narrowed (through some weights less than 1) and in sparse regions the width will be expanded (through some weights more than 1) [20]. The Weighted Gaussian RBF kernel is as follows:

$$k(x, y) = \exp(-\lambda_weight(x) \times \lambda_weight(y) \times \lambda \times \|x - y\|^2) \quad (7)$$

Where λ_weight is a variable changing in a very small range around 1.

3.8 Sementation Performance Evaluation

The performance result of segmentation is evaluated by computation of segmentation accuracy. The segmentation accuracy is calculated as the direct ratio of the number of tumor pixels common for ground truth and the output of the segmented tumor to the total ground truth tumor pixels. The ground truth is indicated from the boundary drawings of the radiologist. segmentation accuracy and segmentation error are as fallows:

$$\text{Segmentation accuracy} = (\text{no. of pixels matche} / \text{total no. of tumor pixels in ground truth}) \times 100 \quad (8)$$

3.9 Classification Performance Evaluation

In order to compare the classification results, classifiers performances were evaluated using round robin (10-fold cross-validation) method [21]. In this method, the total number of data is divided into 10 subsets. In each step one subset is left out and the classifier is trained using the remainders. Next, the classifier is applied to the left out subset to validate the analysis. This process is iterated until each subset is left out once. For instance, in the n-sample images, the round robin method trains the classifier using n – 1 samples and then applies the one remaining sample as a test sample. Classification is iterated until all n samples have been applied once as a test sample. The classifier's accuracy is evaluated on the basis of error rate. This error rate is defined by the terms true and false positive and true and false negative as follows:

$$\text{sensetivity} = TP / (TP + FN) \times 100 \quad (9)$$

$$\text{specificity} = TN / (TN + FP) \times 100 \quad (10)$$

$$\text{accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (11)$$

Where TN is the number of benign tumors truly identified as negative, TP is the number of malignant tumors truly identified as positive, FN, malignant tumors falsely identified as negative and FP, benign tumors falsely identified as positive. Sensitivity is the ability of the method to recognize malignant tumors. Specificity is the ability of the method to recognize benign tumors. Accuracy is the proportion of correctly identified tumors from the total number of tumors.

4. RESULTS AND DISCUSSION

The proposed method is applied to real brain CT images. The data set consists of volume CT data of 20 patients (10-benign, 10- malignant). Number of slices varies across the patient's 4-6 benign slices and 4-6 malignant slices. In total there were 100 slices (50-benign, 50- malignant). All images were with a dimension of 512×512 , gray scale and in DICOM format. The proposed algorithm is implemented in Matlab 2014b.

Result of an input real CT image and the segmented tumor image using SVM classifier is represented in Figure 2.

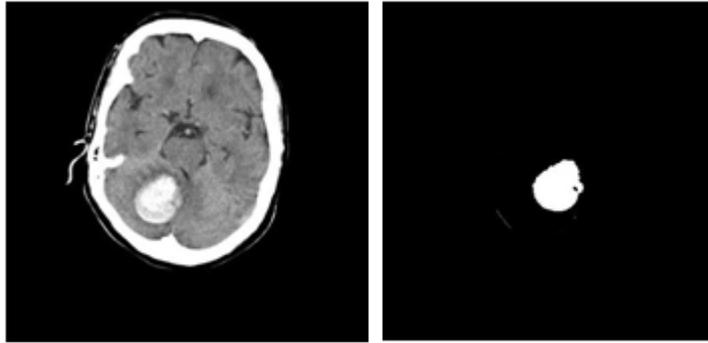


Figure 2. Input CT image and segmented tumor using SVM classifier

The quantitative results in terms of performance measures such as segmentation accuracy and segmentation error for real data of 50 benign slices of 10 patients (5 slices for each patient) and 50 malignant slices of 10 patients(5 slices for each patient), are calculated and tabulated in Table 1.

Table 1. Segmentation Accuracy of 10 Patients with 100 Slices

Patients	Malignant slices (50)	Benign slices (50)
Segmentation accuracy		
1	88.85	89.83
2	89.09	88.98
3	87.89	88.72
4	88.87	87.98
5	89.79	89.65
6	89.86	87.98
7	88.98	87.86

8	88.79	89.65
9	89.78	88.94
10	89.92	88.57
average accuracy	89.182	88.816

Tumor image wavelet approximation and its details in horizontal, vertical and diagonal directions at second level of wavelet decomposition can be observed in Figure 3.

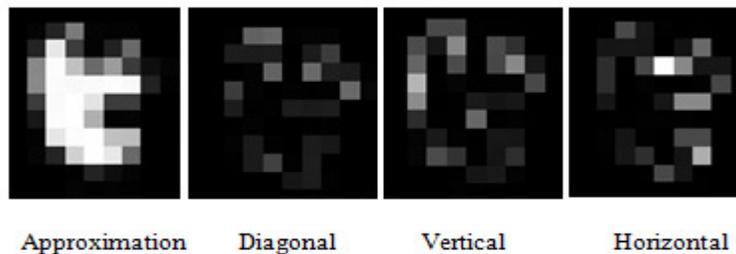


Figure 3. Wavelet Decomposition Images of Segmented Tumor

17 features are extracted from the wavelet approximation tumor image of each slice that 6 of them are selected by means of Student's t-test. The best textural features selected are long-run low-gray-level emphasis (LLGE), long-run high-gray-level emphasis (LHGE), energy, contrast, variance and inverse difference moment (IDM). The feature selection outcome is consistent with the knowledge of radiologist. For instance, feature variance computes the heterogeneity of a CT slice and LHGE captures the heterogeneous nature of the texture feature. According to radiologist, it can be inferred from the presence of heterogeneity that an abnormal slice is malignant. Feature IDM measures the homogeneity of a slice and feature LLGE demonstrate the homogeneous nature of the texture feature. Conforming to radiologist, the existence of homogeneity indicates that an abnormal slice is benign. These six features are given as inputs to the K-NN, SVM and WSM classifiers.

The performance of classifiers is evaluated using 10-fold cross-validation method and tabulated in Table 2.

Table 2. Classifier Performances Comparison

classifier	K-NN	SVM	WSVM
sensitivity	76%	77%	78%
specificity	72%	75%	76%
accuracy	74%	76%	77%

Classification accuracy effectiveness is measured by sensitivity and specificity. Compared to K-NN and SVM with the classification accuracy of 74% and 76% respectively, WSVM performed better with the accuracy of 77%. The stated results of the comparison of three classifier performances are represented using a bar graph as shown in Figure 4.

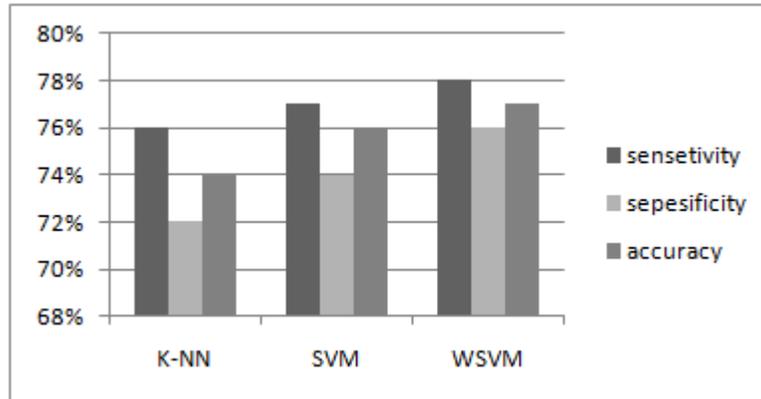


Figure 4. Classifiers Performance Evaluation

5. CONCLUSION

The work in this research involved using SVM with kernel function to classify Brain tumor CT images into benign and malignant. From the experimental results, it is inferred that the best classification performance is achieved using the WSVM. Furthermore, these results show that the proposed method is effective and efficient in predicting malignant and benign tumors from brain CT images. For future work, the proposed method can be applied to other types of imaging such as MRI and even can be used for segmentation and classification of tumors in other parts of body.

ACKNOWLEDGEMENT

The authors wish to acknowledge Shiraz Chamran hospital for providing the original brain tumor images in DICOM format.

REFERENCES

- [1] Padma Nanthagopal, A., Sukanesh Rajamony, R., 2012, A region-based segmentation of tumour from brain CT images using Nonlinear Support Vector Machine classifier, *J. Med. Eng. Technol.*, 36, (5), 271–277
- [2] Nandpuru, H.B., Salankar, S.S., Bora, V.R., 2014, MRI Brain Cancer Classification Using Support Vector Machine, *IEEE. Conf. Electrical Electronics and Computer Science*, 1–6
- [3] Hema Rajini N., Bhavani R., 2011, Classification of MRI Brain Images using k-Nearest Neighbor and Artificial Neural Network., *IEEE-International Conference on Recent Trends in Information Technology*, ICRTIT Chennai, Tamil Nadu., 563 – 568
- [4] Sridharn. D., Murali Krishna.IV., 2013, Brain Tumor Classification Using Discrete Cosine Transform and Probabilistic Neural Network, *2013 International Conference on Signal Processing Image Processing & Pattern Recognition (ICSIPR)*, Coimbatore, 92 – 96

- [5] Sundararaj, G.K.; Balamurugan, V., 2014, Robust Classification of Primary Brain Tumor in Computer Tomography Images Using K-NN and Linear SVM, 2014 International Conference on Contemporary Computing and Informatics (IC3I), Mysore, 1315 – 1319
- [6] Padma, A., Sukanesh, R., 2011, A wavelet based automatic segmentation of brain tumor in CT images using optimal statistical texture features, *Int. J. Image Process.*, 5, (5), 552–563
- [7] Kaur, T., Saini, B. S., & Gupta, S. (2016). Optimized Multi Threshold Brain Tumor Image Segmentation Using Two Dimensional Minimum Cross Entropy Based on Co-occurrence Matrix. In *Medical Imaging in Clinical Applications* (pp. 461-486). Springer International Publishing.
- [8] Kaur, T., Saini, B. S., & Gupta, S. (2016) A joint intensity and edge magnitude-based multilevel thresholding algorithm for the automatic segmentation of pathological MR brain images. *Neural Computing and Applications*, 1-24.
- [9] Verma, A. K., & Saini, B. S. ALEXANDER FRACTIONAL INTEGRAL FILTERING OF WAVELET COEFFICIENTS FOR IMAGE DENOISING. *Signal & Image Processing : An International Journal (SIPIJ)* Vol.6, No.3, June 2015
- [10] Chun-yu ,N., 2009, Research on removing noise in medical image based on median filter method, *IT in Medicine & Education, ITIME '09. IEEE International Symposium*, Jinan, 384 – 388
- [11] Benesty, J.; Jingdong Chen; Huang, Y.A., 2010, Study of the widely linear Wiener filter for noise reduction, *Acoustics Speech and Signal Processing (ICASSP)*, 2010 IEEE International Conference, Dallas, TX, 205- 208
- [12] El-Naqa, I., Yang, Y., Wernick, M.N., Galatsanos, N.P., Nishikawa, R.M, 2002, A support vector machine approach for detection of microcalcifications, *IEEE Trans. Med. Imag.*, 21, (12), 1552–1563
- [13] Zhengliang Huan; Yingkun Hou, 2008, An Segmentation Algorithm of Texture Image Based on DWT, *Natural Computation*, 2008. *ICNC '08. Fourth International Conference*, Jinan, 5, 433- 436
- [14] Tang, X., 1998, Texture information in run length matrices, *IEEE Trans. Image Process.*, 7, (11), 234–243
- [15] Khuzi, M., Besar, R., Zaki WMD, W., Ahmad, N.N., 2009, Identification of masses in digital mammogram using gray level co-occurrence matrices, *Biomed. Imag. Interv. J.*, 5, (3), 109–119
- [16] Haralick, R.M., Shanmugam, K., Dinstein, I., 1973, Texture features for Image classification, *IEEE Trans. Syst. Man Cybern.* 3, (6), 610–621
- [17] Levner, I., Bulitko, V., Lin, G., 2006, Feature extraction for classification of proteomic mass spectra: a comparative study, *Springer-Verlag Berlin Heidelberg, Stud Fuzz*, 207, 607–624
- [18] Soper D.S.: ‘P-value calculator for a student t-test (OnlineSoftware)’, 2011, <http://www.danielsoper.com/statcalc3>
- [19] F. Latifoglu, K. Polat, S. Kara, S. Gunes, 2008, Medical diagnosis of atherosclerosis from carotid artery Doppler signals using principal component analysis (PCA), k-NN based weighting pre-processing and Artificial Immune Recognition System (AIRS), *J. Biomed. Inform.* 41, 15–23.
- [20] Yuvaraj N., Vivekanandan P., 2013, An Efficient SVM based Tumor Classification with Symmetry Non-Negative Matrix Factorization Using Gene Expression Data , *Information Communication and Embedded Systems (ICICES)*, 2013 International Conference, Chennai, 761– 768

- [21] Liao Y.-Y., Tsui, P.-H., Yeh, C.-K., 2009, Classification of benign and malignant breast tumors by ultrasound B-scan and nakagami-based images, J. Med. Biol. Eng. 30, (5), 307–312

AUTHORS

Dr. Hamed Agahi, has obtained his doctoral degree from Tehran University, Iran. He has 4 years of teaching experience. He is currently working as an Assistant Professor and the head of researchers and elite club in Shiraz Azad University, Iran. He has published many papers in scientific journals and conference proceedings. His research interests include pattern recognition, image processing, signal processing and machine vision and applications.



Kimia Rezaei received her Bachelor degree from Fasa Azad University, Iran, and the Master degree in telecommunications engineering from Shiraz Azad University, Iran. She has published one paper in national conference in Iran. She is currently working as Telecommunications Engineer in Sahand Telecommunication company in Iran. Her research interest is focused on pattern recognition and Image processing related research programs targeted for practical applications.

