ORIGINAL ARTICLE

Received: October 2017
Accepted: February 2018

Analysis of MRI Images of the Liver, using a Combination of Wavelet and Principle Component Analysis (Pca) and Support Vector Machine (SVM) for the Diagnosis and Classification of Benign and Malignant Tumors

Bahman Cheraghi Gharakhanloo[1],Bashir Bagheri Nakhjavanlo[2],
Ali Mohammad Mohammadi[3]

ABSTRACT

Diagnosis of the tumors' tissues in the liver and distinguishing the malignant tumors from benign is a critical issue in medicine. In this regard, so many methods have been proposed to make the accurate tumor detection and classification algorithms using Machine Learning and Computer Vision techniques. In this study, first we analyzed the liver's MR images using Discrete Wavelet Transform techniques for dimensionality reduction and feature extraction, and then Principal Component Analysis technique has been employed to select the essential features for classification, and finally the selected features were used to train Support Vector Machine algorithm. In classification, we used the different kernels for SVM and the result of each classifier was compared. The outcome of the algorithm indicates the high performance of our method when there are few training data available.

Keywords: Liver, Tumor, Malignant, Benign, Discrete Wavelet Transform, Support Vector Machine, Principal Component Analysis.

1. Department of computer engineering, Karaj Branch, Islamic Azad University, Karaj, Iran.
2. Department of Mathematics and Computer, Firoozkooh Branch, Islamic Azad University, Firoozkooh, Iran.
3. Department of Computer Engineering, West Tehran Branch, Islamic Azad University, Tehran, Iran.

***Corresponding Author:**
Ali Mohammad Mohammadi
Department of Computer Engineering, West Tehran Branch, Islamic Azad University, Tehran, Iran.
Tel: (+98)9126474391
Email address:
mohammadi81ali@gmail.com

BCCR
2018; 10(1): 34-41
www.bccrjournal.com

INTRODUCTION:

Liver cancer is one of the most dangerous diseases, occurring due to the irregular and uncontrolled growth of a number of cells within the liver that can attack adjacent healthy tissues[1]. The existence of a mass in the body does not necessarily mean a malignant tumor; some types of these masses are benign and have no harm. Benign tumors are those that grow alongside the tissue and are uniform, known as non-cancerous tumors, and the ones that stick to the body tissues and spread to its other parts are called malignant or cancerous tumors. The proper diagnosis is the base for the development of an optimal plan for liver cancer treatment[2,3]. One of the most popular medical imaging methods is Magnetic Resonance imaging (MRI), used to diagnose the disease and design a treatment plan[4].

For the analysis of liver images, first features extraction from the existing images must be done, where wavelet is a very efficient tool with multi-resolution property, i.e. it can be used to analyze an image at different levels in terms of resolution. However, this technique requires a lot of processing, so we used dimensional reduction technique to reduce the feature vector[5].

In dimensional reduction, PCA in the mathematical definition is an orthogonal linear transformation that takes data to a new coordinate system, so that the largest data variance is placed on the first coordinate axis, the second largest variance on the second coordinate axis and it goes on like this for the rest. PCA can be used to reduce the dimension of the data, as a result of which the components of the dataset extracted from the image with the greatest effect on the variance are maintained and, accordingly, the most appropriate features are extracted[6].

According to the features obtained for differentiation of the tumor from the main tissue and classification of its types, among many methods presented by the researchers, such as logistic regression[7], artificial neural network[8], and especially the deep neural network that needs a lot of initial data for training, we used SVM technique, which offers high flexibility and precision in case of few training data for supervised classification. We used four different kernels for implementation in SVM. K-fold cross validation technique was used to prevent over-fit of the learner algorithm in the training phase[9]. By default, SVM was implemented with a linear kernel, and then three other kernels (HPOL, IPOL, GRB) were selected and implemented. Finally the results of the system with four different kernels were compared.

The main components of the proposed algorithm

Generally, our method consists of three main steps: the first stage involves the preprocessing, which is the stage of extracting the features and reducing the dimension, the second involves the training of a kernel-based SVM, and the third one involves the addition of images from the liver to the trained algorithm and receiving the relevant results, with the whole process shown in **Figure 1.**

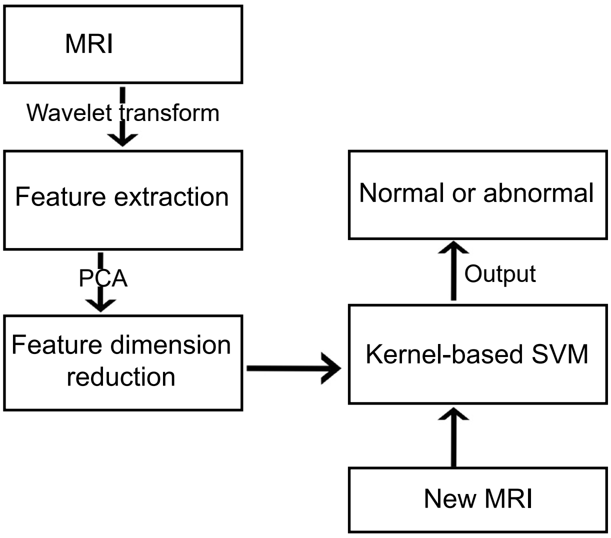

Figure 1. The main components of the proposed hybrid algorithm

Dimensionality Reduction and Extracting features with discrete wavelet

At this stage, which is preprocessing, to analyze the existing images, we extracted the features of the image, using discrete wavelet transform method. Wavelet analysis is one of the relatively new and exciting achievements of pure mathematics, which is based on several decades of research in frequency and wave analysis. Today, wavelet analysis is applied to many field of science and engineering and there are new possibilities for understanding its mathematical aspects as well as increasing its applications. The wavelet is a given function with a mean of zero in terms of whose transforms and dilations the expansion is performed. It is possible to formulate any applicationwhich is based on fast Fourier transform (FFT) by means of wavelets and obtain more spatial (or temporal) information. Overall, this affects the processing of signal, image and fast numerical algorithms for calculation of integral operators. This analysis, as a numerical tool, can greatly reduce the computational complexity and make it easy and fast for the computation to be done. Today, wavelet analysis has different and numerous applications, including its use in medical imaging (MRI) and CT scan.

Two-dimensional wavelet (2D)

Wavelet transformations in two-dimensional mode require a two-dimensional scalar function such as φ(x, y) and three two-dimensional wavelets: $\psi^{H}(x,y)$, ψ^{V} (x,y), and ψ^D (x,y). Each of them is the product of two one-dimensional functions. It is easy to scale up a two-dimensional to one-dimensional DWT with respect to wavelet functions and two-dimensional scaling[10].

Dimensional reduction through PCA

After acquiring the features which are extracted from MR images using DWT, we applied Principal Component Analysis to select the best feature to use in classification stage. The PCA is the most widely used algorithm, and in which we act based on the identification of the principle components, and obtain the vector of the features of the data covariance matrix, i.e. the main components. Concerning vector x one can assume that its components have a mean zero and, if not, we can easily calculate the mean and subtract from them. Now, if we assume that Y is the covariance matrix x and A is the eigenvector of matrices, and B is the eigenvalue matrix calculated from covariance matrix Y, and represent each member of A with Ø and each member of B with λ, then the principle components matrix P containing corresponding eigenvectors will be larger eigenvalue[11].

$\lambda_1 > \lambda_2 > \ldots > \lambda_m > \lambda_1 > \ldots \lambda_n$

P=[Ø_1,Ø_2,…,Ø_m]

New feature vectors are obtained using the following equation:

C = P.X

Classification of SVM

At this stage, the features selected by PCA are given to an algorithm called support vector machine (SVM). Then, based on the features the algorithm is trained in order to be able to accurately handle the classification after receiving the new data and to determine the new location of the tumor and its type according to the new liver images. SVM is actually a binary classifier that separates two classes, using a linear boundary. In this method, the samples forming the boundaries of the classes are obtained, using all the bands and an optimization algorithm. These samples are called support vectors. A number of training points with the shortest distance to the decision boundary can be considered as a subset for definition of the decision boundaries and as support vector[12].

Linear kernel

The kernel-free and default mode of this algorithm is known as the linear kernel. Assume that the data is composed of two classes, and that the classes in total have xi , i =1,...,L training points, where xi is a vector. These two classes are labeled with y i = ± 1. To calculate the decision boundary of the two completely separate classes, optimal margin method is used [180]. In general, a linear decision boundary can be written as follows:

(1): $w.x + b = 0$

x is a point on the decision boundary, and w is an n-dimensional vector perpendicular to the decision boundary. b/w is the origin distance to the decision boundary and w.x represents the inner multiplication of the two vectors x and w. Since, by multiplying a constant on both sides (1), equality will still be established, for the definition of the unity of the values w and b, the following conditions are applied to them.

(2): $Y_i(w.\, x_i + 1) > 1$ (x_i is a support vector)

$Y_i(w.\, x_i + 1) = 1$ (xi is not a support vector)

The first step in calculating the optimal decision boundary is to find the closest samples of the two training classes. In the next step, the distance between those points is calculated in a straight line perpendicular to the boundaries separating the two classes completely. The optimal decision boundary is calculated by solving the following optimization problem:

(3): $\max_{w,b} \min_{i=1,\ldots,L} [y_i \frac{(w.x_i+b)}{|w|}]$

Based on equation 2 the above equation, through performing a series of mathematical operations, changes to the one below.

(4):

$$\min_{w,b} \frac{1}{2} |w|^2 ,\ y_i (w.x_i + b) - 1 \geq 0 \quad i = 1, \ldots, L$$

Solving the problem of optimization (4) is a difficult task. To simplify it by using the Lagrangian indeterminate coefficients, this optimization problem can be converted to the following form where λ_is are Lagrange coefficients.

(5):

$$\min_{i=1,\ldots,L} [-\tfrac{1}{2}\textstyle\sum_{i=1}^{L} \sum_{j=1}^{L} \lambda_i y_i (x_i x_j) \lambda_j y_j + \sum_{i=1}^{L} \lambda_i]$$

$$\lambda_i \geq 0 \quad i = 1, \ldots, L$$

$$\sum_{i=1}^{L} \lambda_i y_i = 0$$

After solving the above optimization problem and finding Lagrange coefficients, w is calculated by means of the following equation.

(6): $W = \sum_{i=1}^{L} \lambda_i y_i x_i$

λ_i of support vectors will be greater than zero and λ_i of the other points will be zero. Hence, considering (6) and zero λ_i for xies that are not support vectors, to obtain a decision boundary, we only need a limited number of points of instruction which are the support vectors; the image classification by means of SVM requires just a limited number of training points. After finding w using the following equation, the value of b is calculated through the following equation for different support vectors and the final bs are obtained by averaging the resulted b.

(7): $\lambda_i [y_i (w.\, x_i + b) - 1] = 0 \qquad i = 1, \ldots, L$

The final classifier is obtained from (8).

(8): $f(X, W, b) = \mathrm{sgn}(W.X + b)$

The traditional SVM has built a cloud page for data classification, so it cannot properly handle the conditions in which different data from different types are located on different regions of a surface cloud. In such cases, the isolation method was implemented based on the kernel[13]. We focused on increasing accuracy of the algorithm for which we tested three polynomials: homogeneous polynomials, heterogeneous polynomials, and Gaussian kernels[14], whose equations have been shown in **Table 1**.

Table 1. Different kernels for SVM

Parameter	Equations	Kernels
d	$K(x_i,x_j)=(x_i.x_j)^d$	Heterogeneous polynomials
d	$K(x_i,x_j)=(x_i.x_j+1)^d$	Heterogeneous polynomials
γ	$K(x_i,x_j)=\exp(-\gamma\|x_i-x_j\|^2)$	Gaussian

Cross-validation K-Fold

Since the classifier is trained according to the existing training data, our classifier may have a high accuracy just for these data, and in actual circumstances in the face of other data, it has proper accuracy. Therefore, the algorithm should be suitably generalized to prevent over-fit, and to make algorithm more reliable in the face of other datasets for this purpose, a stage called cross-validation is used. There are different methods for cross-validation, among which K-Fold is the one we used in our study. In this method, we divide all the training data into a subset of k, so that in each time of run, the i-th set is tested and build it with the rest of the model. K-fold is completely randomly divided, but there is another method called Stratified k-fold, where each fold separated is close to the distribution of the same class[15]. Another important issue is determining the appropriate value for k. If the value is too large, the estimator bias of the real error rate will be small, but the estimator variance will be large, as a result of which the algorithm's calculation time will increase. By contrast, if K is too small, the calculation time will be low, but the estimator variance will be small and the estimator bias will be large[16]. The value of k can be considered to be a number between 1 and 10. We considered number five appropriate, which is usually obtained by error and trial.

Tests and comparison of methods

The implementation of the algorithm and the corresponding tests has been done by MATLAB 2015b. In fact, we use the tools provided in MATLAB to implement the wavelet and support vector machine algorithms. Also MRI images were downloaded from the reference (http://www.mr-tip.com), and we used T-2 images due to their high resolution (256×256 pixels). We used liver images randomly for various diseases, among which 120 were abnormal, indicating a defective liver, and 130 were normal, indicating a healthy one. **Table 2**, indicates the number of normal and abnormal liver samples that we use to train and test our algorithm.

Table 2. The number of the samples used to train and test the proposed algorithm.

Total number of images	Training images 200		Testing images 50	
	Normal	abnormal	Normal	abnormal
	100	100	30	20

Feature extraction, using wavelet

Using the wavelet, we processed the image at three levels, and by reducing the wavelet approximation and removing the rest of the parts the original image was reduced to a large amount in terms of volume and converted to an image of 32 × 32, where a specific section was visible with the box in the upper left corner and above the image on the right in **Figure 2**.

Dimensional reduction:

The original image, whose size has significantly been reduced in the feature extraction step, needs to be reduced again in dimensions by PCA method, with k=1 computational variance of 42% and k = 20 computational variance of 96% that show minimum and maximum values of variance respectively.

Accuracy and the time of hybrid algorithm

The hybrid algorithm was run with all four kernels and the results of the classification were shown in each step, and the results were compared with those obtained by using the previous methods as shown in **Table 3**.

The calculation time of the algorithm for each 256× 256 image was 0.0489. In this process, the feature extraction step had the maximum time consumed, i.e. 0.031, and according to **Table 3**, the hybrid algorithm with Gaussian kernel had an accuracy of 98.82%.

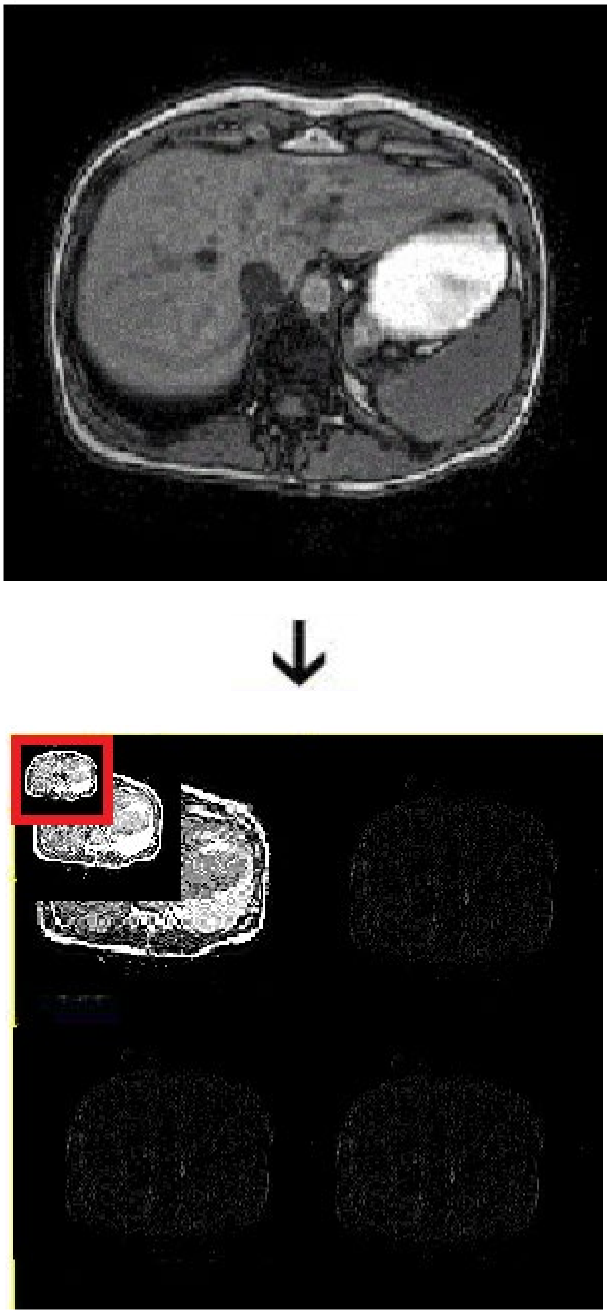

Figure 2. Liver image and its wavelet output

CONCLUSION:

In this paper, we introduced a hybrid algorithm for identification and differentiation of the liver tumor from MRI images. The wavelet, due to its nature, causes a minimal loss of the valuble features of the image. So in pre-processing stage with the wavelet, valuable information was obtained from the image. PCA reduced the search space from 1024 to 20, which had a great effect on speed and accuracy of the final result. In the classification stage, we used LIN, HPOL, IPOL, and GRB kernels among which GRB kernel, acting as an exponential function, was able to classify more distance as the decision line margin, and produced the best result with accuracy of 98.82% even among the trained kernels. Moreover, one of the benefits of this classification algorithm is obtaining good results in case of low training data and, accordingly, our proposed method, DWT + PCA + SVM, is based on GRB kernel, which can be used as an appropriate tool for diagnosis and treatment of liver disease.

This algorithm is not only useful in the analysis of the liver image but also it can be used in cases which need classification as a useful tool for medical diagnoses.

In future’s studies the effectiveness of this algorithm in the analysis of other medical images can be investigated, and through an alternative method, which is more developed than the wavelet and an alternative method of dimension decrease, such as ICA, as well as comparing the other kernells of SVM, a more suitable method can be offered.

Table 3. The results of the algorithms available in the research literature and those of our proposed method

Results from research literature	
[12] DWT+SOM	%94
[12] DWT+SVM with linear kernel	96%
[12] DWT+SVM with RBF based kernel	98%
[41] DWT+PCA+ANN	97%
[41] DWT+PCA+kNN	98%
[25] [41] DWT+PCA+ACPSO+FNN	98.75%
[17] Cascaded Fully Convolutional Neural Networks	94%
[19] Moth-flame optimization	95.66%
[20] WPT+ SVM	88.8%
[21] WPT+Multi-class SVM	95.4
Results obtained from our proposed algorithm with four kernels	
DWT+PCA+KSVM (LIN)	92.12%
DWT+PCA+KSVM (HPOL)	94.21%
DWT+PCA+KSVM (IPOL)	95.11%
DWT+PCA+KSVM (GRB)	98.82%

REFERENCES:

1. Chen EL, Chung PC, Chen CL, Tsai HM, Chang CI. An automatic diagnostic system for CT liver image classification. IEEE Trans Biomed Eng. 1998 Jun;45(6):783-94.
2. Pavlopoulos S, Kyriacou E, Koutsouris D, Blekas K, Stafylopatis A, Zoumpoulis P. Fuzzy neural network-based texture analysis of ultrasonic images. IEEE Eng Med Biol Mag. 2000 Jan-Feb;19(1):39-47.
3. Lee CC, Chung PC, Tsai HM. Identifying multiple abdominal organs from CT image series using a multimodule contextual neural network and spatial fuzzy rules. IEEE Trans Inf Technol Biomed. 2003 Sep;7(3):208-17.
4. Huang YL, Chen JH, Shen WC. Diagnosis of hepatic tumors with texture analysis in nonenhanced computed tomography images. Acad Radiol. 2006 Jun;13(6):713-20.
5. Qiu H, Lee J, Lin J, Yu G. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. J Sound Vib. 2006 Feb 7;289(4-5):1066-90.
6. Zhang YD, Wu LN, Wei G. A New Classifier for Polarimetric Sar Images. Prog Electromagn Res. 2009;94:83-104.
7. Hosmer DW, Taber S, Lemeshow S. The Importance of Assessing the Fit of Logistic-Regression Models - a Case-Study. Am J Public Health. 1991 Dec;81(12):1630-5.
8. Cagnoni S, Coppini G, Rucci M, Caramella D, Valli G. Neural-Network Segmentation of Magnetic-Resonance Spin-Echo Images of the Brain. J Biomed Eng. 1993 Sep;15(5):355-62.
9. Genov R, Chakrabartty S, Cauwenberghs G. Silicon support vector machine with on-line learning. Int J Pattern Recogn. 2003 May;17(3):385-404.
10. Garge DM, Bapat VN. A low cost wavelet based mammogram image processing for early detection of breast cancer. Indian Journal of Science and Technology. 2009 Sep 1;2(9):63-5.
11. Smith LI. A tutorial on principal components analysis2002 Feb 26.
12. Wang L, Chan KL, Zhang Z. Bootstrapping SVM active learning by incorporating unlabelled images for image retrieval:629-34. 2003.
13. Acevedo-Rodríguez J, Maldonado-Bascón S, Lafuente-Arroyo S, Siegmann P, López-Ferreras F. Computational load reduction in decision functions using support vector machines. Signal Processing. 2009 Oct 1;89(10):2066-71.
14. Deris AM, Zain AM, Sallehuddin R. Overview of support vector machine in modeling machining performances. Procedia Engineering. 2011 Jan 1;24:308-12.
15. May RJ, Maier HR, Dandy GC. Data splitting for artificial neural networks using SOM-based stratified sampling. Neural Networks. 2010 Mar 1;23(2):283-94.
16. Armand S, Watelain E, Roux E, Mercier M, Lepoutre FX. Linking clinical measurements and kinematic gait patterns of toe-walking using fuzzy decision trees. Gait & posture. 2007 Mar 1;25(3):475-84.
17. Christ PF, Ettlinger F, Grün F, Elshaera ME, Lipkova J, Schlecht S, Ahmaddy F, Tatavarty S, Bickel M, Bilic P, Rempfler M. Automatic liver and tumor segmentation of ct and mri volumes using cascaded fully convolutional neural networks. arXiv preprint arXiv:1702.05970. 2017 Feb 20.
18. Chlebus G, Meine H, Moltz JH, Schenk A. Neural network-based automatic liver tumor segmentation with random forest-based candidate filtering. arXiv preprint arX-

iv:1706.00842. 2017 Jun 2.

19. Said S, Mostafa A, Houssein EH, Hassanien AE, Hefny H, editors. Moth-flame Optimization Based Segmentation for MRI Liver Images. International Conference on Advanced Intelligent Systems and Informatics; 2017: Springer.
20. Cao G, Shi P, Hu B. Liver fibrosis identification based on ultrasound images. InEngineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the 2006 Jan 17 (pp. 6317-6320). IEEE.
21. Virmani J, Kumar V, Kalra N, Khandelwal N. SVM-based
22. characterization of liver ultrasound images using wavelet packet texture descriptors. Journal of digital imaging. 2013 Jun 1;26(3):530-43.

Rouhollahi MR, Mohagheghi MA, Mohammadrezai N, Motlagh AG, Harirchi I. Situation analysis of the National Comprehensive Cancer Control Program in the IR of Iran; assessment and recommendations based on the IAEA imPACT mission. Archives of Iranian medicine. 2014 Apr 1;17(4):222.