



RESEARCH ARTICLE

The Performance of Novel Wavelet for Lung Image Analysis

Donia Ben Hassen^{1*} and Sihem Ben Zakour²

¹Department of Informatics, University of Manouba, Tunisia

²Department of Basic Science, University of Dammam, KSA, Saudi Arabia

*Corresponding author: Dr. Donia Ben Hassen, Department of Informatics, University of Manouba, Tunisia



Abstract

Even the advancement in the medical image, chest radiography still the most common used and cheapest method in detecting lung diseases. The generated image allows physicians to screen the normal, benign and malignant tissues in the lung. In order to better classify and reduce noise in the image's information, the employ of preprocessing methods is unavoidable. A novel scheme for lesions classification in chest radiographs is presented in this paper by combining discrete wavelet analysis (WA), FCM algorithm and Support Vector Machine (SVM). The wavelet is engaged to reduce the redundant information, FCM to pre-classify then SVM to the final classification of the lung diseases. A comparative study is presented, and the results suggest that our approach help radiologists and experts to better diagnosis the disease.

Keywords

Wavelet analysis, Computer aided diagnosis, Chest radiography, Classification, FCM, SVM

Introduction

Despite the wide variety of high-tech imaging techniques available, the traditional two-dimensional chest radiograph remains ubiquitous in clinical practice. Chest radiography has defended its position in the diagnostic workflow although advances in the field of computed tomography (CT). The main arguments in favor of chest radiography are broad availability, cost effectiveness as well as the relatively low-dose exposure. However, compared to CT imaging, sensitivity and specificity for the detection of lung lesions are low [1]. One of the most important and difficult tasks is the diagnosis of lung lesions from chest radiographies is interpreting a chest radiograph because it is extremely challenging. Superimposed anatomical structures make the image

complicated. Even experienced radiologists face trouble detecting and specifying the kind of lesions. This limitation of human expert-based diagnosis has provided a strong motivation for the use of computer technology to improve the speed and accuracy of the diagnosis process. Recently, Computer-Aided Diagnosis (CAD) has become one of the major research subjects in medical imaging and diagnostic radiology.

With CAD, the performance by computers does not have to be comparable to or better than that by physicians, but needs to be complementary to that by physicians. In fact, CAD systems are currently used to help in the early diagnostic process and follow-up. This task is especially useful for lung diseases which need early diagnosis, adequate follow-up, and timely monitoring of disease indicators.

Though, uncertainty is characteristic of information derived from chest radiography. It can be inconsistent, incomplete and uncertain. In addition, this vagueness affects the diagnosis and decision making. Therefore physicians need a system which is significantly analogous to a human judgment. Despite lesion classification systems provide the foundation for lesions diagnosis and patient cure, the studies reported in developing a CAD application was limited to the distinction between the cancerous lesions and the non-cancerous. The design of a CAD which can give an idea about the nature of the lesion, for example the lesion is of 50% an infection, 10% a cancer and 30% a tuberculosis etc... may be helpful for handling this vagueness and as a support to diagnosis in this field.

To reach this goal, we propose in this paper to model and tackle imperfection and vagueness of the chest

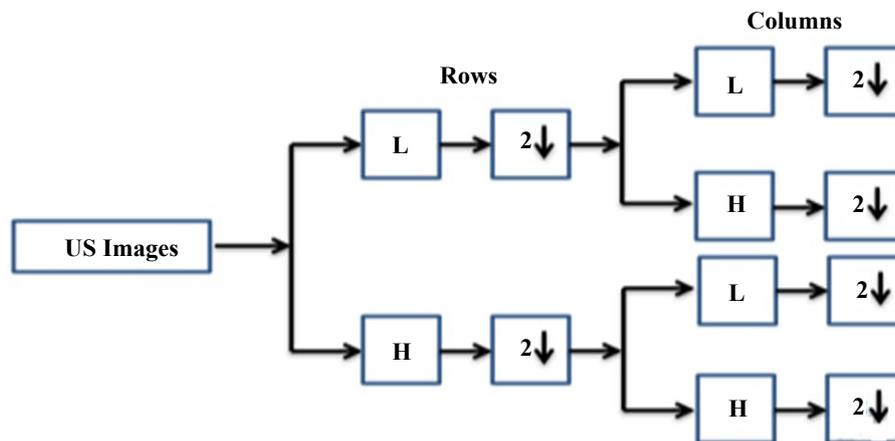


Figure 1: Wavelet analysis for image processing.

radiography and ensure the performance of our CAD system and its accuracy. Indeed, we propose a novel combined approach for classification of chest lesions. This paper is organized as follows: Section 1 describes the related works divided on the three parts Wavelet analysis, FCM and SVM, respectively. Then, the combined Wavelet-Fuzzy-SVM method is described in section 2. Finally, section 3 presents the results obtained and a comparison between the overall approaches.

Proposed Method

Wavelet analysis

A common form of multivariate analysis is Wavelet analysis. The Wavelet theory, discovered by [2] and has been employed in different scientific fields, such as physics, engineering and mathematic, data compression, speech analysis, ... Wavelet analysis [3] decomposes a function into frequency components that represent different degrees of function smoothness, with High frequency components capture the least smooth function behavior while low frequency components capture smooth, making easy the extraction of information exclusively in the time-frequency domain, as shown in Figure 1. A wavelet is a waveform, with limited duration and having an average value of zero, with irregular and asymmetric properties. It is a mathematical function which decomposes a signal (1D) or an image (2D) to details and approximation. Wavelets are ordinarily utilized to decompose and compress picture. Wavelets are computed based on different levels of decomposition. The multi-determination level allows the investigation of dim level pixels from different areas and sizes. These properties prompt to the possibility that wavelets could manage specialists to better characterize the lung maladias.

Wavelet families are functions generated from base function $w(t)$, called mother wavelet, by means of scaling and translating operations:

$$\psi_{a,\tau}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-\tau}{a}\right) \quad a > 0$$

Where $\psi(t)$ satisfy $\int_R \psi(t) dt = 0$. A is the scale parameter, τ is the translation parameter.

The scaling operation is based on “stretching” and “compressing” operations on the mother wavelet. This leads to use different frequency information of the function. The most used scaled is the binary one (shrinking by factors of two) and dyadic translations of the basic wavelet, $w(t)$. Therefore, discrete wavelet transform (DWT) can be identified by $a=2^j$ and $T=2^k$, where j and k are integers. Then, the wavelet function is expressed as follows:

$$\psi_{jk}(t) = 2^{-j} \psi(2^{-j}t - k) \quad j, k \in Z$$

The Wavelet analysis has a low pass and high pass filters. The Haar wavelet is the perfect choice in studying the time domain (compactly supported, small support, only 2 taps) but not in frequency domain. In addition, the Haar wavelet has an efficient memory exactly reversible (easy reconstruction) and computationally is the cheapest one.

Wavelets are essential and generally utilized component descriptors for surface examination, because of their adequacy in catching restricted spatial and recurrence data and multi-resolution attributes. Hence Many old works are dedicated to the application of wavelet analysis in image processing such as [4]. In this paper, the ROIs are deteriorated to 4 levels, LL, LH, HL, and HH, by utilizing 2D symlets wavelet in light of the fact that the symlets wavelet has preferable symmetry over Daubechies wavelet and accordingly more reasonable for picture preparing [5]. We remove level, vertical and corner to corner directional detail coefficients from the wavelet deterioration structure. At last, we get the wavelet includes by ascertaining the mean and the fluctuation of these wavelet coefficients. And also recently, the works are more sophisticated by combining as examples [6-9].

Once the lung images were prepared, wavelets were used for noise reduction. As the Haar wavelet filter is the simplest one, we have applied it using different lev-

els of resolution. The wavelets were then preprocessed.

The combined Wavelet-Fuzzy-SVM method

Here, we only describe the use of wavelet in our CAD system which is a continuation of our already published work in [10-13]. The other stages are briefly presented in this section.

The importance of the reduction in the feature space in the context of the classification of medical radiographic images is felt not only because the high number of features may deteriorate the quality of the classification in addition to the risk of correlated features but also the high number of extracted features for the classification may be a time consuming.

The number of reduction methods is very important. We extracted in our previous work [10] an optimal set of features according to the description of specialists which are Size, Age, Gender, The X-fraction and Y-fraction, the circularity, Skewness, Kurtosis, Entropy, Correlation, Homogeneity, Lacunarity and finally the Localisation. These features can represent the characteristics of different kinds of lesions in chest radiograph. Then, we needed to eliminate redundant variables from the subset extracted because they affect the performance of the classification. We used Stepwise Forward Selection and Principal Components Analysis (PCA). We obtained two subsets of features. We finally experimented the Stepwise/FCM/SVM classification and the PCA/FCM/SVM one. The results suggest that the second approach may be helpful to radiologists for reading chest images.

In this paper we propose to compare these results already published in [10] with those obtained using the wavelet analysis as preprocessing. The wavelet method allows us to eliminate firstly the insignificant or irrelevant features. It thus seems interesting to compare the relevance of attributes resulting from the wavelet method with those from the PCA method. We propose to realize the classification phase with two sets of features in order to know the most pertinent features. The Figure 2 below presents our followed procedure in classifying the different lung.

The Fuzzy C Means (FCM) Algorithm uses a fuzzy clustering, in which the input vector x is pre-classified with different membership values. The intelligibility is the motive force behind the use of FCM algorithm for this problem. However, a compromise between interpretability and accuracy is met. On the other hand, we focused on a more accurate solution by using SVM. Then we risk losing the linguistic sense defining the fuzzy models. Indeed, we have experimented also the possibility to increase the interpretability of SVM classifier by the hybridization with the clustering method Fuzzy C means.

Experimental Study and Results

Almost all features presented in the literature are classical features. Some of them describe the shape of the candidate region considered as the circularity, the perimeter; some features are used to study the texture such as the intensity. In this stage, we need new features that should be able to represent the characteristics of the lesions and specify the nature of the detected lesion. To this aim, we computed several features based on characteristics given by service of medical imaging of CHU Charles Nicolle described in our previous work [11].

We discussed with our collaborators in the service of medical imaging of the university hospital Charles Nicolle and we detected the five important diseases which are: Lung cancer, Metastasis, Tuberculosis, Infection, Benign tumors and we computed the correspond features.

The chest radiographs are taken from the JSRT database [14]. The database contains 247 PA chest radiographs collected from 13 institutions in Japan and one in the United States. The images were taken from films to a size of 2048×2048 pixels, a spatial resolution of 0.175 mm/pixel and 12 bit gray levels. In order to test our approach, the chest radiographs require separation of lung fields from background.

We judge the performance of our classification approach using several evaluation criteria often used in the literature. For a five class clustering problem, one

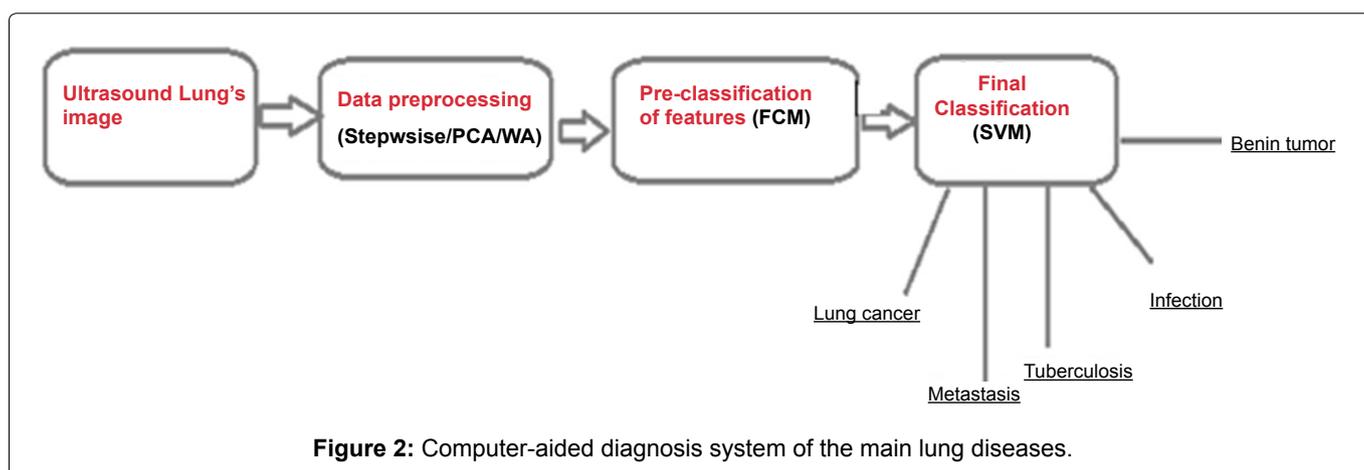


Figure 2: Computer-aided diagnosis system of the main lung diseases.

Table 1: Performances the different employed methods.

Classes	Stepwise/FCM/SVM (published in [1])			PCA/FCM/SVM (published in [1])			WA/FCM/SVM		
	Accuracy (%)	SE (%)	SP (%)	Accuracy (%)	SE (%)	SP (%)	Accuracy (%)	SE (%)	SP (%)
Cancer	70.33	73.00	71.33	74.19	80.00	60.00	90.2	94.05	92.05
Metastasis	60.25	63.88	61.56	50.94	66.66	50.00	90.01	95.03	92.42
Infection	55.66	52.76	53.00	51.72	50.00	52.00	60.72	59.73	60.27
Tuberculosis	60.23	55.85	57.00	56.56	53.33	51.88	60.79	57.74	59.84
Benign tumors	53.33	54.67	53.89	52.00	51.45	50.00	70.03	72.02	71.03

can distinguish true positive (TP) (refers to the sample correctly classified), false positive (FP) (refers to the false sample classified as true sample), False Negative (FN), it is considered as false sample classified or also false sample, and true negative (TN), it is considered as false sample classified or also true sample. From these values, measures such as accuracy, sensitivity (SE) and specificity (SP) can be computed given by the following equations.

$$SE = \frac{TP}{TP + FN}$$

$$SP = \frac{TN}{TN + FP}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Table 1 shows the results of our proposed combined method WA/FCM/SVM with these ones (Stepwise/FCM/SVM and PCA/FCM/SVM) already published and detailed in our previous work [10]. The accuracy, SE and SP are the highest one for Cancer and Metastasis by recording 90.2; 94.05; 92.05 and 90.01; 95.03; 92.42 respectively. It should be noticed that there is also a significant improvement for Infection, Tuberculosis and Benign tumors in terms of Accuracy, Sensitivity and Specificity.

We can conclude that the WA/FCM/SVM can achieve better accuracy in our classification problem.

Conclusion

The medical image plays an imperative assignment in clinical analysis. The wavelet transform is extensively employed in the medical specialization. In this work, we administered a new Computer Aid Diagnosis system for lung diseases. We have exposed a novel combined method to handle with some unclearness of the chest radiography. We have presented a combination of wavelet features with fuzzy and SVM for medical diagnosis. As the Medical data are expected to be noisy with huge size, it is a difficult practice to select the most appropriate features for an automated medical diagnosis system. The use of wavelet reduces the noise and the high dimensionality of data while the fuzzy system

also decreases the complexity of the medical data. Therefore, we proceed by the elimination of redundant features by employing Wavelet analysis. And then we have classified the lung diseases using Fuzzy and SVM methods. We have compared the performance of our proposed method with Stepwise/FCM/SVM and PCA/FCM/SVM classification. The results suggest that WA/FCM/SVM approach improves the identification of the lung diseases and help the radiologists to better read chest images. The Wavelet Technique exhibits a more efficient noise reduction and feature extraction method compared to the PCA and stepwise methods.

References

- Rühl R, Wozniak MM, Werk M, Laurent F, Mager G, et al. (2008) Csl-detector-based dual-exposure dual energy in chest radiography for lung nodule detection: results of an international multicenter trial. *European Radiology* 18: 1831-1839.
- Morley CJ, Banham AD, Johnson P, Thorburn GD, Jenkin G (1978) Physical and physiological properties of dry lung surfactant. *Nature* 271: 162-163.
- Mallat SG (1989) A theory for multiresolution signal decomposition: The wavelet representation. *J. IEEE Transactions on Pattern Analysis and Machine Intelligence* 11: 674-693.
- Scheunders P, Livens S, Van de Wouwer G, Vautrot P, Van Dyck D (1998) Wavelet-based texture analysis. *International Journal on Computer Science and Information Management* 1: 22-34.
- Wang A, Sun H, Guan Y (2006) The application of wavelet transform to Multi-modality medical image fusion. *Proceedings of the 2006 IEEE International Conference on Networking, Sensing and Control*, 270-274.
- Heba M. El-Hoseny, El-Sayed M. El-Rabaie, Wael Abd Elrahman, Fathi E Abd El-Samie (2017) Medical image fusion techniques based on combined discrete transform domains. *34th National Radio Science Conference (NRSC)*.
- Jagadesh T, Jhansi Rani R (2016) A novel speckle noise reduction in biomedical images using PCA and wavelet transform. *International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*.
- Xiao Z, Zheng C (2009) Medical Image Fusion Based on An Improved Wavelet Coefficient Contrast. *3rd International Conference on Bioinformatics and Biomedical Engineering*, 1-4.
- Xiaojun X, Youren Wang, Shuai Chen (2016) Medical

- image fusion using discrete fractional wavelet transform. *Biomedical Signal Processing and Control* 27: 103-111.
10. Ben Hassen D, Ben Zakour S, Taleb H (2016) A Hybrid Fuzzy-SVM classifier for automated lung diseases diagnosis. *Polish Journal of Medical Physics and Engineering* 22: 97-103.
 11. Ben Hassen D, Taleb H, Yaacoub IB, Najla Mnif (2014) Classification of chest lesions with using fuzzy c-means algorithm and support vector machines. *International Joint Conference SOCO'13-CISIS'13-ICEUTE'13*, 319-328.
 12. Ben Hassen D, Taleb H (2013) Automatic detection of lesions in lung regions that are segmented using spatial relations. *Clinical Imaging* 37: 498-503.
 13. Ben Hassen D, Taleb H, Ben yaakoub I, Najla Mnif (2011) A fuzzy approach to chest radiography segmentation involving spatial relations. *IJCA Special Issue on Novel aspects of digital imaging applications*, 40-47.
 14. Van Ginneken B, Stegmann MB, Loog M (2006) Segmentation of anatomical structures in chest radiographs using supervised methods: A comparative study on a public database. *Medical Image Analysis* 10: 19-40.