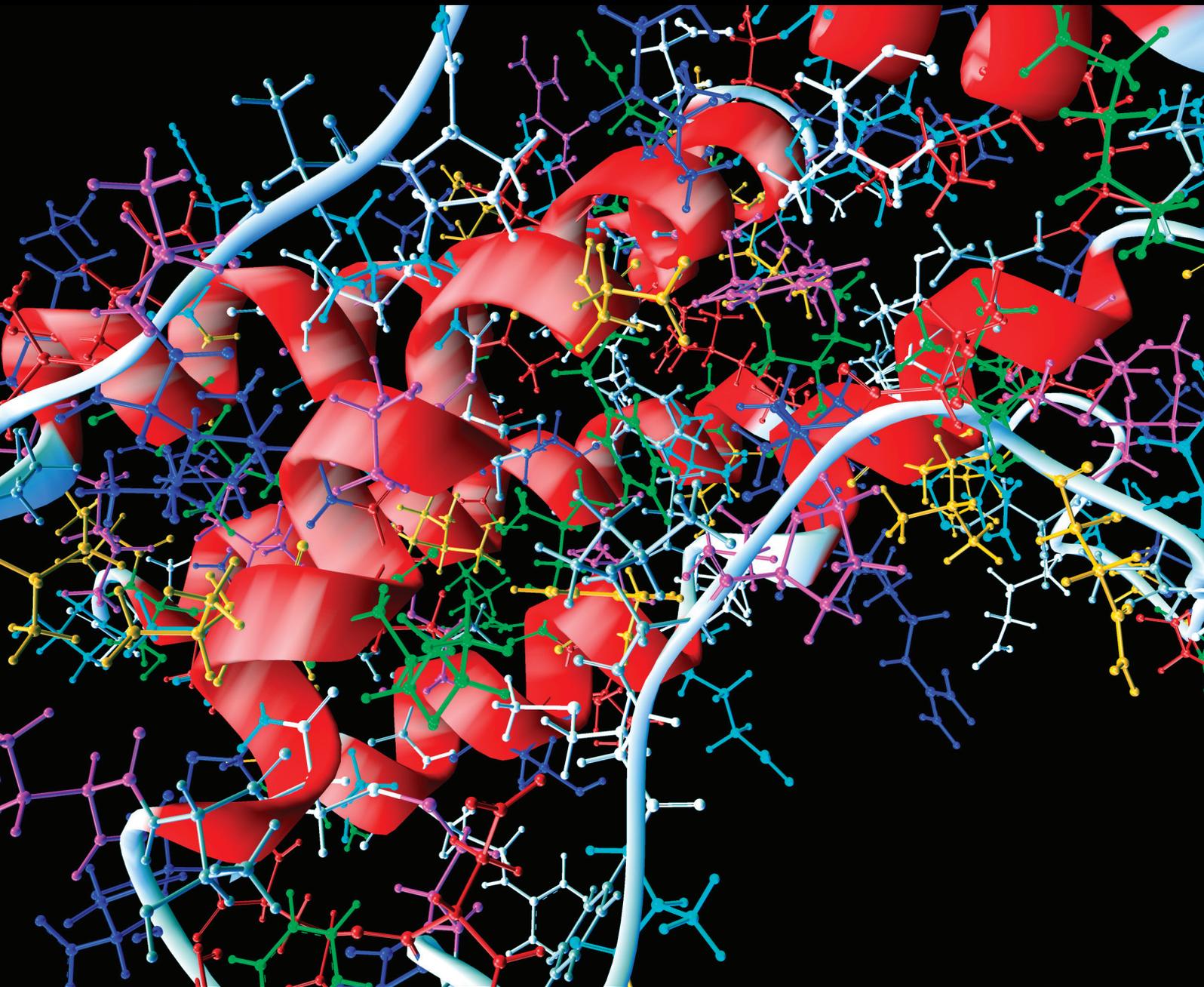


Computational and Mathematical Methods in Medicine

# Biomedical Signal and Image Processing for Clinical Decision Support Systems 2014

Guest Editors: Kayvan Najarian, Kevin R. Ward, and Shahram Shirani





---

**Biomedical Signal and Image Processing  
for Clinical Decision Support Systems 2014**

Computational and Mathematical Methods in Medicine

---

**Biomedical Signal and Image Processing  
for Clinical Decision Support Systems 2014**

Guest Editors: Kayvan Najarian, Kevin R. Ward,  
and Shahram Shirani



---

Copyright © 2015 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in “Computational and Mathematical Methods in Medicine.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Editorial Board

Emil Alexov, USA  
Georgios Archontis, Cyprus  
Dimos Baltas, Germany  
Chris Bauch, Canada  
Maxim Bazhenov, USA  
Thierry Busso, France  
Carlo Cattani, Italy  
Sheng-yong Chen, China  
William Crum, UK  
Ricardo Femat, Mexico  
Alfonso T. García-Sosa, Estonia  
Damien Hall, Australia

Volkhard Helms, Germany  
Seiya Imoto, Japan  
Lev Klebanov, Czech Republic  
Quan Long, UK  
C-M Charlie Ma, USA  
Reinoud Maex, France  
Michele Migliore, Italy  
Karol Miller, Australia  
Ernst Niebur, USA  
Kazuhisa Nishizawa, Japan  
Hugo Palmans, UK  
David James Sherman, France

Sivabal Sivaloganathan, Canada  
Nestor V. Torres, Spain  
Nelson J. Trujillo-Barreto, Cuba  
Gabriel Turinici, France  
Kutlu O. Ulgen, Turkey  
Edelmira Valero, Spain  
Guang Wu, China  
Huaguang Zhang, China  
Yuhai Zhao, China  
Xiaoqi Zheng, China  
Yunping Zhu, China

# Contents

**Biomedical Signal and Image Processing for Clinical Decision Support Systems 2014**, Kayvan Najarian, Kevin R. Ward, and Shahram Shirani  
Volume 2015, Article ID 974592, 2 pages

**Breast Cancer Detection with Reduced Feature Set**, Ahmet Mert, Niyazi Kılıç, Erdem Bilgili, and Aydin Akan  
Volume 2015, Article ID 265138, 11 pages

**A Method for Lung Boundary Correction Using Split Bregman Method and Geometric Active Contour Model**, Changli Feng, Jianxun Zhang, and Rui Liang  
Volume 2015, Article ID 789485, 14 pages

**High Order Statistics and Time-Frequency Domain to Classify Heart Sounds for Subjects under Cardiac Stress Test**, Ali Moukadem, Samuel Schmidt, and Alain Dieterlen  
Volume 2015, Article ID 157825, 15 pages

**Developing an Intelligent Automatic Appendix Extraction Method from Ultrasonography Based on Fuzzy ART and Image Processing**, Kwang Baek Kim, Hyun Jun Park, Doo Heon Song, and Sang-suk Han  
Volume 2015, Article ID 389057, 10 pages

**Segmentation and Tracking of Lymphocytes Based on Modified Active Contour Models in Phase Contrast Microscopy Images**, Yali Huang and Zhiwen Liu  
Volume 2015, Article ID 693484, 9 pages

**Fractional Diffusion Based Modelling and Prediction of Human Brain Response to External Stimuli**, Hamidreza Namazi and Vladimir V. Kulish  
Volume 2015, Article ID 148534, 11 pages

**Color Enhancement in Endoscopic Images Using Adaptive Sigmoid Function and Space Variant Color Reproduction**, Mohammad S. Imtiaz and Khan A. Wahid  
Volume 2015, Article ID 607407, 19 pages

**A Unified Framework for Brain Segmentation in MR Images**, S. Yazdani, R. Yusof, A. Karimian, A. H. Riazi, and M. Bennamoun  
Volume 2015, Article ID 829893, 17 pages

**Developing a New Computer-Aided Clinical Decision Support System for Prediction of Successful Postcardioversion Patients with Persistent Atrial Fibrillation**, Mark Sterling, David T. Huang, and Behnaz Ghoraani  
Volume 2015, Article ID 527815, 10 pages

## Editorial

# Biomedical Signal and Image Processing for Clinical Decision Support Systems 2014

**Kayvan Najarian,<sup>1</sup> Kevin R. Ward,<sup>2</sup> and Shahram Shirani<sup>3</sup>**

<sup>1</sup>*Department of Computational Medicine and Bioinformatics and Department of Emergency Medicine, Michigan Center for Integrative Research in Critical Care, University of Michigan, Ann Arbor, MI, USA*

<sup>2</sup>*Department of Emergency Medicine, Michigan Center for Integrative Research in Critical Care, University of Michigan, Ann Arbor, MI, USA*

<sup>3</sup>*Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada*

Correspondence should be addressed to Kayvan Najarian; [kayvan@umich.edu](mailto:kayvan@umich.edu)

Received 6 April 2015; Accepted 6 April 2015

Copyright © 2015 Kayvan Najarian et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abundance of data produced by a growing number of new diagnostic and monitoring devices has provided caregivers with a large number of data streams to consider at the time of clinical decision-making. One of the main challenges of today's medicine is to analyze these tremendous amounts of complex patient data in an increasing number of complex clinical decisions. Moreover, the human eye may not be able to detect and interpret complex patterns hidden in various clinical data, in particular when interpreting physiologic signals and images. Considering the need to make rapid clinical decisions, typically in stressful environments, the urgency to develop efficient quantitative approaches to analyze complex patient data can be further recognized. Specifically, there is clear need for novel signal and image processing algorithms that can create recommendations and/or predictions for healthcare providers, in a variety of clinical decision-making. The performance and capabilities of these quantitative methods are expected to match the complexity and size of the rapidly evolving imaging and measurement systems.

This special issue is the second of the series, intended as an update on the current status of, and advances in, biomedical signal and image processing methods used for clinical decision support systems. The quantitative methods presented in this issue cover a wide spectrum of algorithmic solutions designed for a variety of clinical applications.

The paper by C. Feng et al. introduces an algorithm for correction of lung boundary in X-ray computed tomography

(CT) that utilizes split Bregman method and geometric active contour model (ASM). K. B. Kim et al. apply fuzzy ART and image processing techniques to develop an automatic method to extract appendix in ultrasonography. S. Yazdani et al. present an automatic hybrid image segmentation method that integrates the modified statistical expectation maximization (EM) method and the spatial information combined with support vector machines (SVMs) and apply that to segmentation of brain MR images. Modified active contour models (ACMs) are used in the paper by Y. Huang and Z. Liu to segment and track lymphocytes in phase contrast microscopy (PCM) images. In an attempt to improve the endoscopic images, used in diagnosis of various gastrointestinal (GI) tract related diseases, M. S. Imtiaz and K. A. Wahid present a computational method that utilizes an adaptive sigmoid function and space-variant color reproduction for color enhancement. The performances of different methods of feature reduction methods, combined with a variety of classifiers, in detection of malignant tumors in breast images are compared by A. Mert et al. In a paper by M. Sterling et al. a computer-aided clinical decision support system is designed to predict the success of postcardioversion treatments among patients with persistent atrial fibrillation.

As more advanced imaging and monitoring systems are developed and more clinical measurements become available, the quantitative algorithms need to be further improved to analyze the resulting complex data. These algorithms are

expected to help not only extract new hidden knowledge from complex patient data, but also provide rapid predictive recommendations to assist healthcare providers in making better and more informed decisions. In addition, it should not be lost that these and other new computational approaches to data will actually better lead us to develop new monitoring and image systems proactively. The papers presented in this special issue outline some of the current computational methods in biomedical and signal analysis.

*Kayvan Najarian*  
*Kevin R. Ward*  
*Shahram Shirani*

## Research Article

# Breast Cancer Detection with Reduced Feature Set

Ahmet Mert,<sup>1</sup> Niyazi Kılıç,<sup>2</sup> Erdem Bilgili,<sup>1</sup> and Aydın Akan<sup>2</sup>

<sup>1</sup>Department of Electrical and Electronics, Piri Reis University, 34940 Istanbul, Turkey

<sup>2</sup>Department of Electrical and Electronics, Istanbul University, 34320 Istanbul, Turkey

Correspondence should be addressed to Ahmet Mert; [amert@pirireis.edu.tr](mailto:amert@pirireis.edu.tr)

Received 12 September 2014; Revised 14 December 2014; Accepted 25 December 2014

Academic Editor: Kevin Ward

Copyright © 2015 Ahmet Mert et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper explores feature reduction properties of independent component analysis (ICA) on breast cancer decision support system. Wisconsin diagnostic breast cancer (WDBC) dataset is reduced to one-dimensional feature vector computing an independent component (IC). The original data with 30 features and reduced one feature (IC) are used to evaluate diagnostic accuracy of the classifiers such as  $k$ -nearest neighbor ( $k$ -NN), artificial neural network (ANN), radial basis function neural network (RBFNN), and support vector machine (SVM). The comparison of the proposed classification using the IC with original feature set is also tested on different validation (5/10-fold cross-validations) and partitioning (20%–40%) methods. These classifiers are evaluated how to effectively categorize tumors as benign and malignant in terms of specificity, sensitivity, accuracy,  $F$ -score, Youden's index, discriminant power, and the receiver operating characteristic (ROC) curve with its criterion values including area under curve (AUC) and 95% confidential interval (CI). This represents an improvement in diagnostic decision support system, while reducing computational complexity.

## 1. Introduction

Breast cancer is one of the leading causes of death among all cancers for women [1]. Early detection and correct diagnosis of cancer are essential for the treatment of the disease. However, the traditional approach to cancer diagnosis depends highly on the experience of doctors and their visual inspections. Naturally, human beings can make mistakes due to their limitations. Humans can recognize patterns easily. However, they fail when probabilities have to be assigned to observations [2]. Although several tests are applied, exact diagnosis may be difficult even for an expert. That is why automatic diagnosis of breast cancer is investigated by many researchers. Computer aided diagnostic tools are intended to help physicians in order to improve the accuracy of the diagnosis [3–5].

A study was carried out to demonstrate that the machine learning may improve the accuracy of diagnosis. In Brause's work, the result shows that the most experienced physician can diagnose with 79.97% accuracy while 91.1% correct diagnosis is achieved with the help of machine learning [6].

Tumors are classified as benign and malignant. Benign tumors are not cancerous or life threatening. However these can increase the risk of getting breast cancer. Malignant

tumors are cancerous and more alarming than benign tumors. Although significant studies are performed for early detection, about 20% of all women with malignant tumors die from this disease [7].

In order to improve accuracy of breast mass classification as benign and malignant, the performance of back-propagation artificial neural network (ANN) was evaluated [8]. Moreover, the fast learning rates and generalization capabilities of radial basis function neural networks (RBFNN) have showed excellent accuracy in microcalcification detection task [9, 10]. The advantages of RBFNN are simple structure, good performance with approaching nonlinear function, and fast convergence velocity. Thus, it has been widely used in pattern recognition and system modeling [11, 12]. On the other hand, the structure of RBFNN increases when the net's input dimension increases. Moreover, the irrelevant components in the inputs will decrease the generalization performance of RBFNN [13].

Support vector machine (SVM) is an effective statistical learning method for classification [14]. SVM is based on finding optimal hyperplane to separate different classes mapping input data into higher-dimensional feature space. SVM has advantage of fast training technique, even with large number

of input data [15, 16]. Therefore it has been used for many recognition problems such as object recognition and face detection [17–19].

Principal component analysis (PCA) is a technique to reduce dimensionality using second order statistical information [20]. Independent component analysis (ICA) is a recently developed method in pattern recognition and signal processing fields [21, 22]. It involves higher order statistics to extract independent components that involve richer information than PCA. ICA can be used to reduce dimensionality before training  $k$ -NN, ANN, RBFNN, and SVM. Consequently the complexity of classifiers can be reduced; convergence velocity and performance can be increased [13, 23].

The objective of the proposed study is to analyze the effect of feature reduction using ICA on classification of the tumors as benign or malignant. Thus, the dimension of WDBC dataset is reduced into only one feature using ICA. The reduced data is subdivided into test and training data using 5/10-fold cross-validation and 20% partitioning to evaluate the performance of  $k$ -NN, ANN, RBFNN, and SVM. Performance measures including accuracy, specificity, sensitivity,  $F$ -score, Youden’s index, and discriminant power are computed and the receiver operating characteristic (ROC) curve is plotted to compare the classifiers. Section 2 summarizes background knowledge on dataset, ICA,  $k$ -NN, ANN, RBFNN, SVM, and performance measures. In Section 3, the methodology deployed in this study is described. In Sections 4 and 5 experimental results are presented and discussed. Finally, there is a conclusion part in Section 6.

## 2. Materials and Methods

**2.1. Dataset Information.** WDBC dataset includes 569 instances with class distribution of 357 benign and 212 malignant. Each sample consists of ID number, diagnosis (B = benign, M = malignant), and 30 features. Features have been computed from a digitized image of a fine needle aspirate (FNA) of a breast mass shown in Figure 1.

Ten real-valued features given in Table 1 calculated for each cell nucleus, and the mean, standard error, and “worst” or largest (mean of the three largest values) of these features were calculated for each image, resulting in 30 features [24].

**2.2. Independent Component Analysis.** The basic model of ICA is as follows. Suppose that the observed signal is the linear combination of two independently distributed sources. The observed signal can be written as follows:

$$x = As, \quad (1)$$

where  $s$  is a vector that consists of the source signals,  $A$  is an unknown mixing matrix composed of constant elements, and  $x$  is a vector of observed values. The unknown mixing matrix,  $A$ , is estimated using the ICA, and then separating matrix  $W$  is computed which is the inverse of  $A$ . The original signal can be found by

$$\hat{s} = Wx. \quad (2)$$

The computing of the independent components (ICs) begins with centering data by removing the mean values of the variable, as in principal component analysis (PCA). Whitening,

TABLE 1: Real-valued features computed for each cell nucleus.

Number	Ten real-valued features
1	Radius (mean of distances from center to points on the perimeter)
2	texture (standard deviation of grey-scale values)
3	Perimeter
4	Area
5	Smoothness (local variation in radius lengths)
6	Compactness (perimeter <sup>2</sup> /area – 1.0)
7	Concavity (severity of concave portions of the contour)
8	Concave points (number of concave portions of the contour)
9	Symmetry
10	Fractal dimension (“coastline approximation” – 1)

also known as sphering data, is the next step. Data which have been whitened are uncorrelated (as PCA). On the other hand, all variables have variances of one. PCA can be used for both these computations because it decorrelates the data and gives information on the variance of the decorrelated data in the form of the eigenvectors [25]. ICs are determined by applying a linear transformation to the uncorrelated data:

$$ic_i = b_i^T x, \quad (3)$$

where  $ic$  is the independent component and  $b$  is the vector to reconstruct  $ic$ . There are many different approaches to estimate  $b$  using an objective function that relates to variable independence. In this study, FASTICA algorithm has been used to compute ICs, due to its flexibility and interactive mode [26].

**2.3. Artificial Neural Networks.** Feedforward neural network (FFNN) is most popular ANN structure due to its simplicity in mathematical analysis and good representational capabilities [27, 28]. FFNN has been used successfully to various applications such as control, signal processing, and pattern classification. FFNN architecture is shown Figure 2.

$N$  states the number of input patterns and  $M$  states the number of neurons in hidden layer. Neurons in the hidden layer receive weighted inputs from a previous layer and transfer output to the neurons in the next layer in FFNN, and these computations can be described as

$$y_{\text{net}} = \sum_{i=1}^n x_i w_i + w_0,$$

$$y_{\text{out}} = f(y_{\text{net}}) = \frac{1}{1 + e^{-y_{\text{net}}}}, \quad (4)$$

$$E = \frac{1}{2} \sum_{i=1}^k (y_{\text{obs}} - y_{\text{out}})^2,$$

where  $w_0$  is bias,  $w_i$  is the weight of each input neuron,  $x_i$  is input neuron,  $y_{\text{net}}$  is composed of the summation of weighted inputs,  $y_{\text{out}}$  is the output of system,  $f(y_{\text{net}})$  denotes

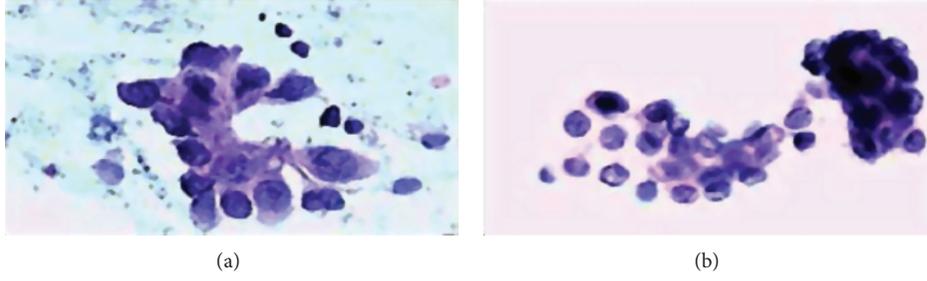


FIGURE 1: FNA biopsies of breast. Malignant (a) and benign (b) breast tumors [24].

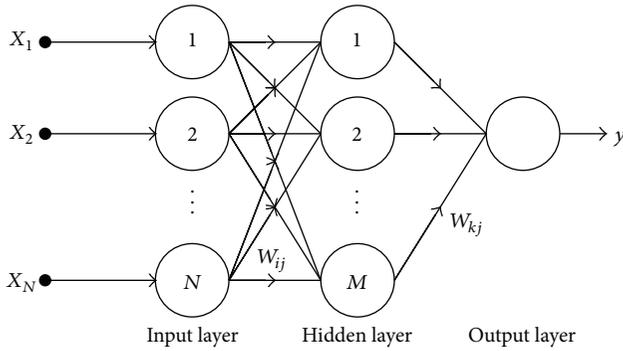


FIGURE 2: Architecture of feedforward neural network.

the nonlinear activation function,  $y_{\text{obs}}$  is the observed output value of neural network, and  $E$  is the error between output value and network result [29].

A RBFNN also consists of feedforward architecture with three layers, but the hidden layer uses Gaussian function mostly and is called radial basis layer. Each neuron consists of a radial basis function (RBF) centered on a point. The centers and spreads are computed by the training. A hidden neuron computes the Euclidean distance of input vector and the test case from the neuron's center point. Thus, it applies the RBF kernel function to the distance using the spread values.

**2.4. Support Vector Machine (SVM).** SVM is a supervised learning algorithm studied for data classification and regression. It was proposed by Boser et al. [30] and Vapnik [31]. SVM algorithm is used to find a hyperplane that separates the classes minimizing training error and maximizing the margin in order to increase generalization capability.

When the datasets are linearly separable, a linear SVM algorithm can be used to classify them. The algorithm tries to maximize the margin. Support vectors are the points lying on the margins that are shown in Figure 3.

The discriminant function of the hyperplane can be described by the following equation:

$$g(x) = w^T x + b, \quad (5)$$

where  $x$  describes data points,  $w$  is a coefficient vector, and  $b$  shows offset from the origin. In case of linear SVM  $g(x) \geq 0$  for the closest point on the one of the class,  $g(x) < 0$  for the closest point belongs to another class. Margin

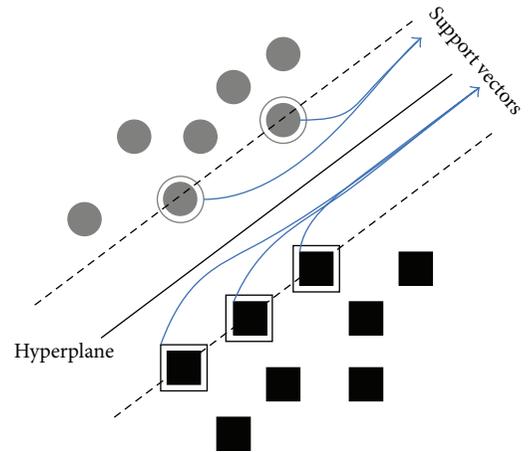


FIGURE 3: The separating hyperplane with support vectors.

$(2/\|w\|^2)$  should be maximized for better generalization ability minimizing the cost function as follows:

$$J(w) = \frac{1}{2} \|w\|^2 \quad (6)$$

$y_i(W^T x_i + b) \geq 1$   $i = 1, 2, \dots, n$  and  $y_i = \{+1, -1\}$  denotes class labels.

This is a quadratic optimization task with respect to a set of linear inequality constraints. From Karush-Kuhn-Tucker (KKT) conditions the Lagrange function is found by

$$L_p(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i \{y_i (W^T x_i + b) - 1\}, \quad (7)$$

where  $\alpha_i$  are Lagrange multipliers and  $L_p$  must be minimized to find out optimal  $w$  and  $b$ . The optimization equation can be written as

$$\text{Maximize} \left[ \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j \right]. \quad (8)$$

The other usage of SVM is that it can solve nonlinear classification problems through the trick of a kernel function. The kernel function maps data points onto a higher-dimensional space in order to construct a hyperplane separating the classes. The new discriminant function is found by

$$g(x) = W^T \Phi(X) + b, \quad (9)$$

TABLE 2: A confusion matrix for binary classification.

Actual value	Recognized value	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

where  $\Phi(X)$  represents the mapping of input vectors, onto the kernel space  $X$ . Therefore, the optimization equation can be written as:

$$\text{Maximize } \left[ \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \right], \quad (10)$$

where  $K(x_i, x_j)$  is the kernel function equals to  $\{\Phi(x_i), \Phi(x_j)\}$ . The kernel functions can be radial basis function (RBF), polynomial or any symmetric functions which satisfy the Mercer conditions [32].

**2.5. Performance Measures.** There are several ways to evaluate the performance of classifiers. Confusion matrix keeps the correct and incorrect classification results to measure the quality of the classifier. Table 2 shows a confusion matrix for binary classification, where TP, TN, FP, and FN denote true positive, true negative, false positive, and false negative counts, respectively.

The most common empirical measure to assess effectiveness is the accuracy for classifier and it is calculated by

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \quad (11)$$

Sensitivity measures the proportion of actual positives which are correctly identified and specificity measures the proportion of negatives which are correctly identified. These are formulated by

$$\begin{aligned} \text{Sensitivity} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{Specificity} &= \frac{\text{TN}}{\text{TN} + \text{FP}}. \end{aligned} \quad (12)$$

*F*-score is a measure of test accuracy. It considers both precision and the recall to compute. These are calculated by

$$\begin{aligned} \text{precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{F-Score} &= \frac{(\beta^2 + 1) \times \text{precision} \times \text{recall}}{\beta^2 \times \text{precision} + \text{recall}}, \end{aligned} \quad (13)$$

where  $\beta$  is the bias and *F*-Score is balanced when  $\beta = 1$ . It favors recall when  $\beta < 1$  and favors precision otherwise.

Other two measures which are used to analyze the performance of a classifier in medical diagnosis are discriminant power (DP) and Youden's index. DP evaluates how

well a classifier distinguishes between positive and negative samples:

$$\text{DP} = \frac{\sqrt{3}}{\pi} (\log X + \log Y), \quad (14)$$

where

$$X = \frac{\text{sensitivity}}{1 - \text{sensitivity}}, \quad Y = \frac{\text{specificity}}{1 - \text{specificity}}. \quad (15)$$

The result can be summarized as follows:  $\text{DP} < 1$  then "poor discriminant,"  $\text{DP} < 2$  then "limited discriminant,"  $\text{DP} < 3$  then "fair discriminant" and other cases then "good discriminant." Youden's index evaluates a classifier's ability to avoid failure [33] and is described as

$$\gamma = \text{sensitivity} - (1 - \text{specificity}). \quad (16)$$

Youden's index is used summary measure of the receiver operating characteristic (ROC) curve. The diagnostic performance of a test or a classifier to distinguish diseased cases from normal cases is evaluated using the ROC curve analysis [34].

In this study, an attempt has been made to evaluate the performance of the classifiers computing the aforementioned measures for 5/10-fold cross-validations (CV) and 20% data partitioning. For 5-CV or 10-CV, the data are divided into 5 or 10 subsets, and each subset is sequentially deployed as test data while others are deployed as training data. Thus 5 or 10 iterative processes are evaluated to determine distinguishing capability of the classification model. Data partitioning is easier and less reliable than CV method. In our simulations, once 20% of the data is randomly selected as test data, the other samples are used for training.

### 3. Methodology

In this study, the original 30 features of WDBC data and reduced one feature using ICA are deployed to evaluate the classifiers performance on breast cancer decision. Thus, the proposed model shown in Figure 4 is applied to WDBC data that have 30 features and 569 instances (patients) were used to train and test the models.

First, the dimensionality of the data is reduced using ICA and partitioned into subsamples using 5/10-CV and 20% partitioning to evaluate the classifiers. The subsamples have been used sequentially to train and test ANN, RBFNN, SVM, and *k*-NN. The outputs of the classifiers have been evaluated to find out performance measures.

First, ICA is used to compute ICs. Since the first IC has distinctly large eigenvalue given in Figure 5, it has been selected as a feature vector.

In other words, one IC can successfully identify the thirty features with the retained 98.205% of nonzero eigenvalues. In addition, the distribution of the IC is given in Figure 6 to indicate its distinguishing capability.

The data are divided into subsets using 5/10-CV and 20% partitioning to test and train classifiers. After training process, the test data are used to evaluate diagnostic performances of the classifiers in terms of sensitivity, specificity, accuracy, *F*-score, Youden's index, DP, and ROC curve.

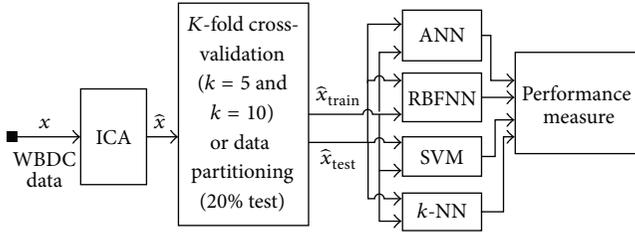


FIGURE 4: The basic model of the study.

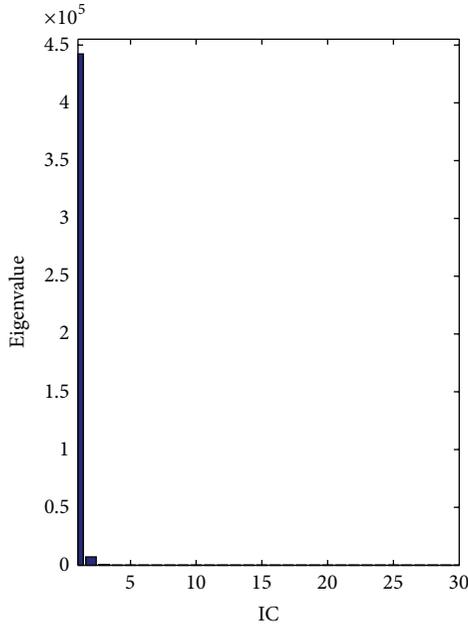


FIGURE 5: Corresponding eigenvalues of the WDBC data.

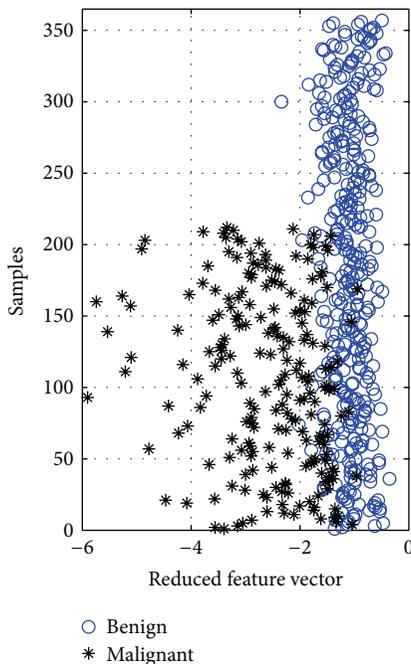


FIGURE 6: The distribution of computed IC (reduced feature vector).

For training processes,  $k$ -NN classifier, one-dimensional Euclidean distance,  $d = \sqrt{(x_{\text{test}} - x_{\text{training}})^2}$  between test and training samples [35]. The results of  $k$ -NN classifier are obtained for the  $k$  values from 1 to 25, and then the performance measures at the best  $k$  value are stored. The model of ANN is selected as feedforward neural network with one hidden layer. The total number of neurons in the hidden layer is sequentially increased to find the maximum accuracy. Moreover, the activation function of the hidden layer of the network has been chosen as log-sigmoid transfer function. In order to train the network, gradient descent with momentum and adaptive learning rate backpropagation algorithm is used. RBFNN is also evaluated varying the spread value ( $\sigma$ ). For SVM, linear, quadratic, and RBF kernels are used to explore which type of separating hyperplane is more suitable for breast cancer classification.

### 4. Results

One-dimensional feature vector of WDBC data reduced using ICA is used for training and testing the classifiers. The accuracy, sensitivity, and specificity of one dimensionality have been performed using 5/10 CV technique and 20% of data as test data. Also, the success of the breast cancer classification is generally evaluated on the basis of sensitivity value because the classifying of the malignant mass is more important than the benign mass.

The accuracy of the  $k$ -NN classifier has been computed for varying  $k$  values between 1 and 25. The comparison graph of the effect of ICA on accuracy of  $k$ -NN classifier is shown in Figure 7.

The maximum accuracy results when 20% test data with 30 features is 96.49% where  $k = 5$ . However, reduced one feature vector using ICA provides the accuracy of 92.98% where  $k = 5$  and 20% test data is selected. Moreover, the accuracy of  $k$ -NN classifier is decreased from 93.15% (30 features) to 91.04% (1 feature by ICA) when 10-CV is used to test and train.

Accuracy graph of ANN has been plotted varying neuron numbers in the hidden layer for 10/5-CV and 20% test data. The accuracy graph of ANN classifier is given in Figure 8.

ANN classifier has nearly perfect accuracy value of 99.12% (the number of neurons is four) when original 30 features and 20% test data are selected. The effect of ICA on reducing into one feature is changed accuracy value to 91.23% where neuron number is nine. In addition, the accuracy value is changed from 97.54% to 90.51% when 10-CV is used.

Spread value of RBFNN is adjusted between 0 and 60 to get maximum accuracy for 20% test data ratio and 10/5-CV. The accuracy graph of RBFNN is shown in Figure 9.

Referring to the accuracy graph of RBFNN, maximum accuracy, 95.12%, is obtained where spread value is 48 for 20% test data. This value is decreased to 90.35% when reduced one-dimensional feature vector by ICA is used. However, when 10-CV is used, the effect of ICA increases the accuracy from 87.18% (with 30 features) to 90.49% (with 1 feature reduced by ICA).

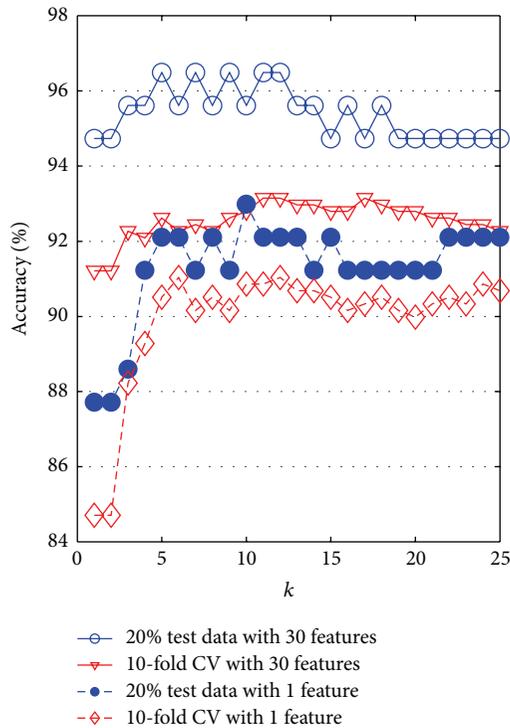
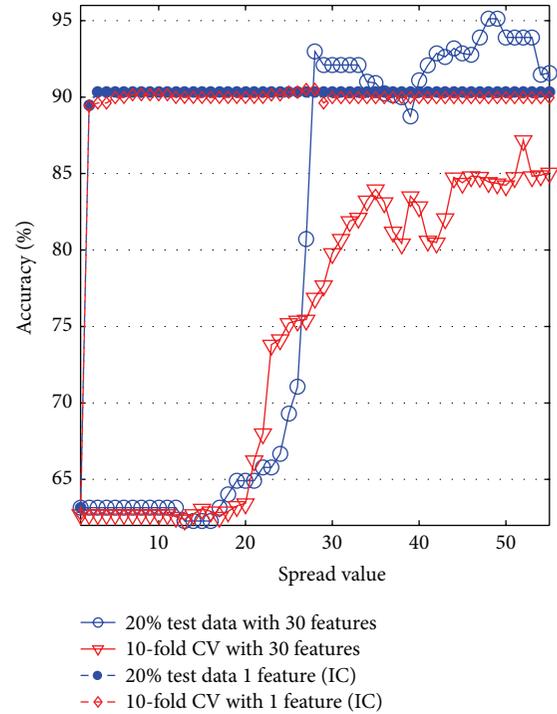
FIGURE 7: The graph of accuracy of  $k$ -NN classifier.

FIGURE 9: The accuracy graph of RBFNN.

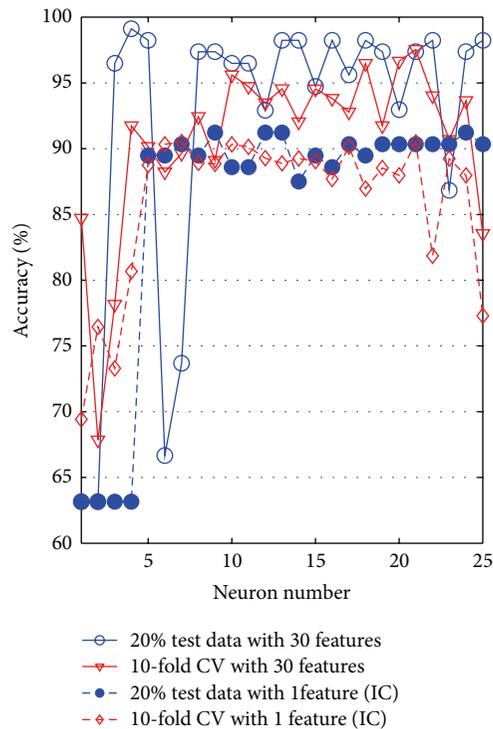


FIGURE 8: The accuracy graph of ANN.

Accuracy evaluation of SVM has been computed for kernel functions including linear, polynomial, and RBF with kernel function parameters such as RBF sigma value for RBF kernel and polynomial degree for polynomial kernel. The accuracy graph of SVM classifier is presented in Figure 10

where the axes of polynomial degree indicate linear kernel when its value equals one.

Generally, SVM classifier with linear kernel provides more accurate result than polynomial and RBF kernel. Its accuracy is 98.25% for 30 features and 90.35% for reduced 1 feature when 20% of data is used as test data. In contrast to polynomial kernel, effect of ICA increases the accuracy of SVM with RBF kernel from 89.47% (30 features) to 91.23% (1 feature). When 10-CV is used, the accuracy is decreased from 97.54% (30 features, linear kernel) and 95.25% (30 features, RBF kernel) to 90.33% and 90.86% (reduced 1 feature by ICA).

$k$ -NN, ANN, RBFNN, and SVM have been tested and trained to find out maximum accuracy adjusting their parameter. The performance measures such as accuracy, specificity, sensitivity,  $F$ -score, Youden's index, and discriminant power of the classifiers are compared to each other. The parameters of the classifiers which provide maximum accuracy are selected to be compared to the other classifiers. In addition to these performance measures, the ROC curve of three classifiers is plotted to enhance visibility of the comparison.

10-CV and one-dimensional feature vector reduced by ICA are used to compare the performances of classifiers. In input data of classifiers, the test data are compared to the original class label to find out TP, TN, FP, and FN values. These values for classifiers are given in the form of confusion matrix in Table 3.

RBFNN classification using 30 original features provides worse performance than reduced one-dimensional feature vector; refer to Table 3. The other classification used with 30 features has slightly higher true values when compared to classification with one feature reduced by ICA.

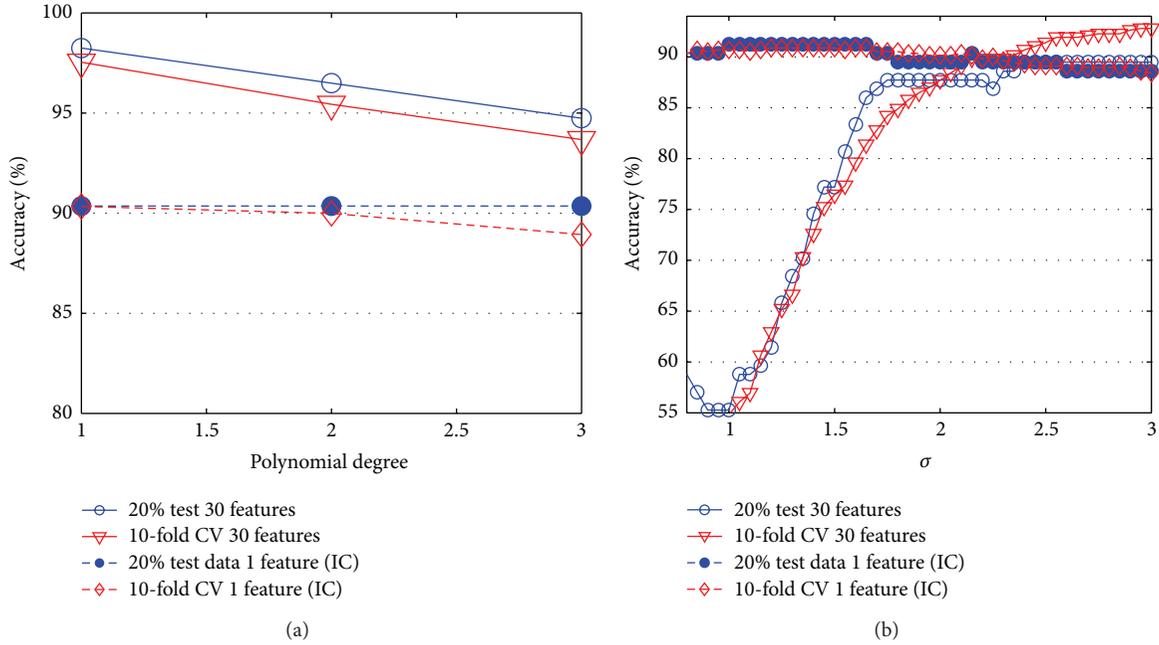


FIGURE 10: The accuracy graphs of SVM classifiers.

TABLE 3: The confusion matrices of the classifiers using reduced one dimensionality by ICA (1F denotes one feature and 30F denotes original features).

<i>k</i> -NN classifier ( <i>k</i> = 6)					ANN classifier (neuron number 7)				
Actual value	Recognized value				Actual value	Recognized value			
	Malignant		Benign			Malignant		Benign	
	1F	30F	1F	30F		1F	30F	1F	30F
Malignant	338 (TP)	346	19 (FN)	11	Malignant	346 (TP)	357	11 (FN)	0
Benign	32 (FP)	28	180 (TN)	184	Benign	43 (FP)	14	169 (TN)	198
RBFNN classifier (spread = 28)					SVM classifier ( $\sigma = 1.3$ )				
Actual value	Recognized value				Actual value	Recognized value			
	Malignant		Benign			Malignant		Benign	
	1F	30F	1F	30F		1F	30F	1F	30F
Malignant	345 (TP)	334	12 (FN)	23	Malignant	348 (TP)	343	14 (FN)	9
Benign	43 (FP)	138	169 (TN)	74	Benign	43 (FP)	13	169 (TN)	199

The performance measures of *k*-NN, ANN, RBFNN, and SVM classifiers such as sensitivity, specificity, accuracy, *F*-score, discriminant power (DP), and Youden’s index are given in Table 4 to compare the effect of ICA on the classification.

Discriminant power evaluates how well a classifier distinguishes between positive and negative samples. DP of ANN and SVM with 30 original features differs from 3 which means good discriminant. When ICA is used to reduce to one dimensionality, DP falls to 2.769 (SVM) and 2.655 (ANN). In other words, discriminants turn to fair.

A higher value of Youden’s index shows better ability to avoid failure. *k*-NN results in the highest value of Youden’s index; refer to Table 4. Youden’s index is used to plot the ROC curve of a classifier. The true positive rate (sensitivity) is plotted in function of the false positive rate (1 – Specificity)

for cut-off points in a ROC curve. The ROC curve can be used to compute area under the ROC curve (AUC) and 95% confidence interval (CI). AUC equals 1 when all test data is assigned to true class labels. Higher AUC indicates that higher accuracy 95% CI is another indicator of the ROC curve which can be used to test whether a classifier can distinguish the classes. If its value is not 0.5, it means the classifier can distinguish the classes. The ROC curves of the *k*-NN, ANN, RBFNN, and SVM classifiers using one-dimensional feature vector reduced by ICA and 30 features are presented in Figure 11.

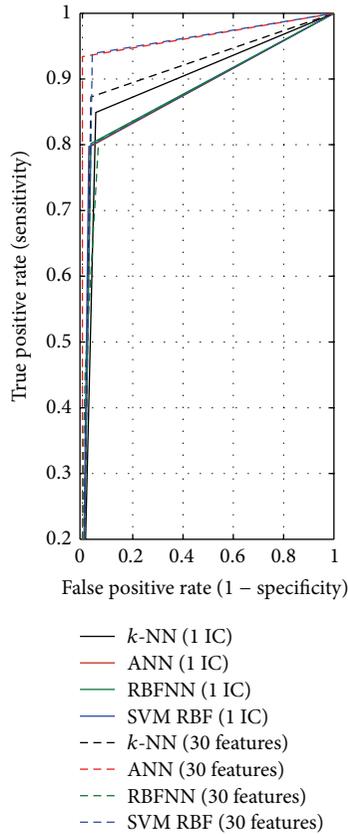
The criterion values of the ROC curves of classifiers are given in Table 5. AUC of the ANN (0.966) and SVM (0.949) results in higher value when 30 original features are used. However, when classification with 1 feature reduced by ICA

TABLE 4: The comparison of ICA algorithm's effect on the classifiers' performance measures (sensitivity, specificity, accuracy, and  $F$ -score in %).

Measures	$k$ -NN		ANN		RBFNN		SVM (RBF K.)	
	1F	30F	1F	30F	1F	30F	1F	30F
$F$ -score	92.98	94.65	92.76	98.07	92.61	80.57	93.04	96.21
DP	2.539	2.912	2.655	$Inf$	2.606	1.131	2.769	3.267
$\Upsilon$	0.795	0.839	0.766	0.934	0.763	0.284	0.772	0.899
Accuracy	91.03	93.14	90.5	97.53	90.49	87.17	90.86	95.25
Specificity	84.9	87.26	79.71	93.39	79.71	34.9	79.71	93.86
Sensitivity	94.67	96.63	96.91	100	96.63	93.55	97.47	96.07

TABLE 5: Criterion values of the ROC curves of  $k$ -NN, ANN, RBFNN, and SVM.

Criterion	$k$ -NN		ANN		RBFNN		SVM	
	1F	30F	1F	30F	1F	30F	1F	30F
AUC	0.880	0.911	0.879	0.956	0.881	0.877	0.879	0.945
95% CI	0.86–0.92	0.89–0.94	0.85–0.91	0.94–0.98	0.85–0.91	0.85–0.91	0.85–0.91	0.92–0.97

FIGURE 11: The ROC curves of  $k$ -NN, ANN, RBFNN, and SVM classifiers.

is evaluated,  $k$ -NN (0.897) and SVM (0.885) result in higher AUC. It means  $k$ -NN and SVM classifiers using reduced one feature distinguish samples more correctly.

Table 5 shows that the accuracy of the  $k$ -NN (91.03%) is better than the accuracy of ANN, RBFNN, and SVM (90.50%, 90.49%, and 90.86%). Generally, one feature reduced by ICA

TABLE 6: CPU time for classification.

Classifier	Partitioning	IC (seconds)	30 features (seconds)
$k$ -NN	20%	8.02	8.31
	10-CV	13.52	14.77
ANN	20%	11.12	13.9
	10-CV	76.72	118.21
RBFNN	20%	14.9	20.03
	10-CV	90.49	129.84
SVM (poly)	20%	7.17	7.28
	10-CV	7.47	9.13
SVM (RBF)	20%	9.02	43.30
	10-CV	10.72	19.05

decreases the accuracy of  $k$ -NN, ANN, and SVM. However, it increases the accuracy of RBFNN.

The aforementioned classification methods are analyzed in terms of computing time given in Table 6 to compare the computational complexities to the classifications with the original 30 features.

The proposed methods have lower computing time when compared to classification of the original dataset. In case of neural network classifications with 30 features, network constructions consume highly more time than classification with one IC. The measured durations of 13.9 and 20.03 seconds are decreased to 11.12 and 14.9 seconds when ANN and RBFNN with 20% partitioning are deployed. Particularly, the effect of using IC as feature on complexity is existed when 10-CV is selected. The consumed time of the ANN and RBFNN is decreased from 118.21 and 129.84 seconds to 76.72 and 90.49 seconds, respectively. In addition, ICA decreases computational times of the SVM and  $k$ -NN classifications, but the rates are less than the neural networks.

## 5. Discussions

Sensitivity/specificity indicates the proportion of actual positives/negatives which are correctly identified. While use of

TABLE 7: Comparison of the methods and accuracy of previous studies and this study.

Author	Method	Feature number	Accuracy	Sensitivity
Krishnan et al. [36]	40% test data, SVM (poly.)	30	92.62%	92.69%
	40% test data, SVM (RBF)		93.72%	94.50%
Bagui et al. [37]	64% test data, $k$ -RNN	30	96.00%	95.09%
	64% test data, $k$ -RNN	Best 3	98.10%	98.05%
Sweilam et al. [38]	PSO + SVM	30	93.52%	91.52%
	QPSO + SVM		93.06%	90.00%
Mangasarian et al. [39]	10-CV, MSM-T	Best 3	97.50%	—
Mert et al. [40]	10-CV, PNN	3 (2IC + DWT)	96.31%	98.88%
	LOO, PNN		97.01%	97.78%
Zheng et al. [41]	$K$ -SVM	6	97.38%	—
This study	10-CV, $k$ -NN	1 feature reduced by ICA	91.03%	94.67%
	40% test, $k$ -NN		92.56%	94.02%
	10-CV, ANN		90.50%	96.91%
	40% test, ANN		90.89%	97.00%
	10-CV, RBFNN		90.49%	96.63%
	40% test, RBFNN		89.98%	96.01%
	10-CV, SVM (linear)		90.33%	96.35%
	40% test, SVM (linear)		90.01%	95.00%
	10-CV, SVM (quadratic)		89.98%	95.24%
	40% test, SVM (quadratic)		91.01%	96.42%
10-CV, SVM (RBF)	90.86%	97.47%		
40% test, SVM (RBF)	91.03%	97.56%		

one-dimensional feature vector reduced by ICA decreases accuracy slightly, it increases sensitivity values of SVM and RBFNN classifiers. The maximum sensitivity measure belongs to SVM with RBF kernel when one feature is used. The graph of the effect of ICA on sensitivity measures of classifiers is shown in Figure 12.

Sensitivity refers successfully identified malignant samples in cancer classification. Thus, higher sensitivity means higher diagnostic capability of malignant tumors and it can be used to help physicians to diagnose cancerous mass more correctly. The accuracy and sensitivity measures of previous classification studies and this study on WDBC dataset are given in Table 7 to compare the effect of feature reduction using ICA. It should be noted that the studies on WDBC differ from studies on WBC dataset which consists of 699 instances with 10 attributes.

Higher number of features used to classify breast cancer as benign and malignant results in slightly higher accuracy. Feature reduction into one using ICA decreases the accuracy of  $k$ -NN, ANN, and SVM slightly. However, it increases the accuracy of RBFNN and the sensitivity values of SVM and RBFNN.

Referring to Table 7, the sensitivity measures of the classifiers used with one-dimensional feature vector reduced by ICA in this study perform better than the other studies. However, accuracy rates of the proposed classifications ( $90.53\% \pm 0.34$ ) are lower than the previous methods ( $94.93\% \pm 2.07$ ). The study of WDBC data creators [39] set has the highest accuracy (97.50%) using multisurface method tree (MSM-T) with 3 selected features. Similarly, hybrid methods are

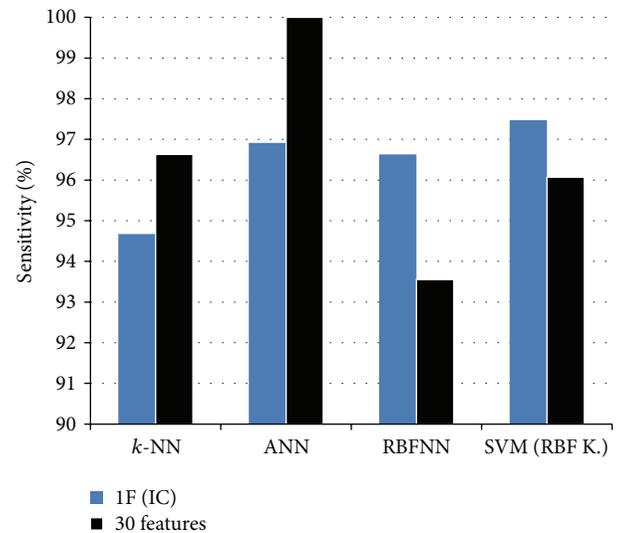


FIGURE 12: Sensitivity measures of the classifiers.

more successful than the others. Breast cancer classifications using probabilistic neural network (PNN) with hybrid feature reduction using discrete wavelet transform (DWT) and ICA [40] or classification using SVM with 6-dimensional feature space obtained by  $K$ -means algorithm [41] have accuracy rates of 96.31% and 97.38% for 10-CV. Particularly, SVM based studies [36, 38] with 30 features have near scores to our one-dimensional results.

## 6. Conclusions

In this study, the effect of dimensionality reduction using independent component analysis (ICA) on breast cancer decision support systems with several classifiers such as artificial neural network (ANN),  $k$ -nearest neighbor ( $k$ -NN), radial basis function neural network (RBFNN), and support vector machine (SVM) is investigated. The results of the applied original thirty features of Wisconsin diagnostic breast cancer (WDBC) are compared with the reduced one dimension by ICA. The accuracy rates of the classifications with thirty original features except RBFNN have slightly decreased from 97.53%, 91.03%, and 95.25% to 90.5%, 91.03%, and 90.86%, respectively. However, the one-dimensional feature vector causes RBFNN classifier to be more distinguishing with the increased accuracy from 87.17% to 90.49%. Furthermore, the sensitivity rates which define the successfully recognized malignant samples are increased from 93.5% to 96.63% for RBFNN and from 96.07% to 97.47% for SVM, while the others have slight decrease at the rate between 0.96% and 3.09%. If the objective is to increase the rate of the successfully identified malignant breast cancer using RBFNN or decrease computational complexity without loss of the high accuracy rate, feature reduction applying ICA can be a high performance solution.

## Conflict of Interests

The authors declare that they have no conflict of interests.

## Acknowledgment

This work was supported by the Istanbul University Scientific Research Projects, Project numbers YADOP-6987, 36196, 38262, 42330, and 35830.

## References

- [1] I. Christoyianni, E. Dermatas, and G. Kokkinakis, "Fast detection of masses in computer-aided mammography," *IEEE Signal Processing Magazine*, vol. 17, no. 1, pp. 54–64, 2000.
- [2] N. Salim, *Medical Diagnosis Using Neural Network*, Faculty of Information Technology University, 2013, <http://www.generation5.org/content/2004/MedicalDiagnosis.asp>.
- [3] A. Tartar, N. Kilic, and A. Akan, "Classification of pulmonary nodules by using hybrid features," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 148363, 11 pages, 2013.
- [4] N. Kilic, O. N. Ucan, and O. Osman, "Colonic polyp detection in CT colonography with fuzzy rule based 3D template matching," *Journal of Medical Systems*, vol. 33, no. 1, pp. 9–18, 2009.
- [5] A. Mert, N. Kiliç, and A. Akan, "Evaluation of bagging ensemble method with time-domain feature extraction for diagnosing of arrhythmia beats," *Neural Computing and Applications*, vol. 24, no. 2, pp. 317–326, 2014.
- [6] R. W. Brause, "Medical analysis and diagnosis by neural networks," in *Proceedings of the 2nd International Symposium on Medical Data Analysis (ISMDA '01)*, pp. 1–13, Madrid, Spain, October 2001.
- [7] T. S. Subashini, V. Ramalingam, and S. Palanivel, "Breast mass classification based on cytological patterns using RBFNN and SVM," *Expert Systems with Applications*, vol. 36, no. 3, pp. 5284–5290, 2009.
- [8] M. N. Gurcan, H.-P. Chan, B. Sahiner, L. Hadjiiski, N. Petrick, and M. A. Helvie, "Optimal neural network architecture selection: improvement in computerized detection of microcalcifications," *Academic Radiology*, vol. 9, no. 4, pp. 420–429, 2002.
- [9] A. P. Dhawan, Y. Chitre, C. Bonasso, and K. Wheeler, "Radial-basis-function based classification of mammographic microcalcifications using texture features," in *Proceedings of the 17th IEEE Engineering in Medicine and Biology Annual Conference*, pp. 535–536, September 1995.
- [10] A. T. Azar and S. A. El-Said, "Superior neuro-fuzzy classification systems," *Neural Computing and Applications*, vol. 23, no. 1, supplement, pp. 55–72, 2012.
- [11] M. Jia, C. Zhao, F. Wang, and D. Niu, "A new method for decision on the structure of RBF neural network," in *Proceedings of the 2006 International Conference on Computational Intelligence and Security*, pp. 147–150, November 2006.
- [12] J. K. Sing, S. Thakur, D. K. Basu, M. Nasipuri, and M. Kundu, "High-speed face recognition using self-adaptive radial basis function neural networks," *Neural Computing & Applications*, vol. 18, no. 8, pp. 979–990, 2009.
- [13] R. Huang, L. Law, and Y. Cheung, "An experimental study: on reducing RBF input dimension by ICA and PCA," in *Proceedings of the 2002 International Conference on Machine Learning and Cybernetics*, vol. 4, pp. 1941–1945, November 2002.
- [14] V. N. Vapnik, *Statistical Learning Theory*, John Wiley & Sons, New York, NY, USA, 1989.
- [15] M. F. Akay, "Support vector machines combined with feature selection for breast cancer diagnosis," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3240–3247, 2009.
- [16] B. Wang, H. Huang, and X. Wang, "A support vector machine based MSM model for financial short-term volatility forecasting," *Neural Computing and Applications*, vol. 22, no. 1, pp. 21–28, 2013.
- [17] M. Pontil and A. Verri, "Support vector machines for 3D object recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 6, pp. 637–646, 1998.
- [18] J. Zhou, G. Su, C. Jiang, Y. Deng, and C. Li, "A face and fingerprint identity authentication system based on multi-route detection," *Neurocomputing*, vol. 70, no. 4–6, pp. 922–931, 2007.
- [19] E. Gumus, N. Kilic, A. Sertbas, and O. N. Ucan, "Evaluation of face recognition techniques using PCA, wavelets and SVM," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6404–6408, 2010.
- [20] S. Kara, A. Güven, and S. İçer, "Classification of macular and optic nerve disease by principal component analysis," *Computers in Biology and Medicine*, vol. 37, no. 6, pp. 836–841, 2007.
- [21] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Networks*, vol. 13, no. 4–5, pp. 411–430, 2000.
- [22] M. P. S. Chawla, "A comparative analysis of principal component and independent component techniques for electrocardiograms," *Neural Computing and Applications*, vol. 18, no. 6, pp. 539–556, 2009.
- [23] S. D. Villalba and P. Cunningham, "An evaluation of dimension reduction techniques for one-class classification," *Artificial Intelligence Review*, vol. 27, no. 4, pp. 273–294, 2007.

- [24] W. H. Wolberg, W. N. Street, and O. L. Mangasarian, "Machine learning techniques to diagnose breast cancer from image-processed nuclear features of fine needle aspirates," *Cancer Letters*, vol. 77, no. 2-3, pp. 163-171, 1994.
- [25] K. H. Liu, B. Li, Q. Q. Wu, J. Zhang, J. X. Du, and G. Y. Liu, "Microarray data classification based on ensemble independent component selection," *Computers in Biology and Medicine*, vol. 39, no. 11, pp. 953-960, 2009.
- [26] 2013, <http://research.ics.tkk.fi/ica/fastica/>.
- [27] J. Bilski, "The UD RLS algorithm for training feedforward neural networks," *International Journal of Applied Mathematics and Computer Science*, vol. 15, pp. 115-123, 2005.
- [28] N. Sivri, N. Kilic, and O. N. Ucan, "Estimation of stream temperature in Firtina Creek (Rize-Turkiye) using artificial neural network model," *Journal of Environmental Biology*, vol. 28, no. 1, pp. 67-72, 2007.
- [29] O. A. Abdalla, M. H. Zakaria, S. Sulaiman, and W. F. W. Ahmad, "A comparison of feed-forward back-propagation and radial basis artificial neural networks: A Monte Carlo study," in *Proceedings of the International Symposium in Information Technology (ITSim '10)*, vol. 2, pp. 994-998, Kuala Lumpur, Malaysia, June 2010.
- [30] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pp. 144-152, July 1992.
- [31] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.
- [32] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Wiley, New York, NY, USA, 1953.
- [33] W. J. Youden, "Index for rating diagnostic tests," *Cancer*, vol. 3, no. 1, pp. 32-35, 1950.
- [34] L. L. Pesce and C. E. Metz, "Reliable and computationally efficient maximum-likelihood estimation of proper binormal ROC curves," *Academic Radiology*, vol. 14, no. 7, pp. 814-829, 2007.
- [35] J. Hamidzadeh, R. Monsefi, and H. S. Yazdi, "DDC: distance-based decision classifier," *Neural Computing & Applications*, vol. 21, no. 7, pp. 1697-1707, 2012.
- [36] A. M. Krishnan, S. Banerjee, C. Chakraborty, and A. K. Ray, "Statistical analysis of mammographic features and its classification using support vector machine," *Expert Systems with Applications*, vol. 37, no. 1, pp. 470-478, 2010.
- [37] S. C. Bagui, S. Bagui, K. Pal, and N. R. Pal, "Breast cancer detection using rank nearest neighbor classification rules," *Pattern Recognition*, vol. 36, no. 1, pp. 25-34, 2003.
- [38] N. H. Sweilam, A. A. Tharwat, and N. K. Abdel Moniem, "Support vector machine for diagnosis cancer disease: a comparative study," *Egyptian Informatics Journal*, vol. 11, no. 2, pp. 81-92, 2010.
- [39] O. L. Mangasarian, W. N. Street, and W. H. Wolberg, "Breast cancer diagnosis and prognosis via linear programming," *Operations Research*, vol. 43, no. 4, pp. 570-577, 1995.
- [40] A. Mert, N. Kılıç, and A. Akan, "An improved hybrid feature reduction for increased breast cancer diagnostic performance," *Biomedical Engineering Letters*, vol. 4, no. 3, pp. 285-291, 2014.
- [41] B. Zheng, S. W. Yoon, and S. S. Lam, "Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1476-1482, 2014.

## Research Article

# A Method for Lung Boundary Correction Using Split Bregman Method and Geometric Active Contour Model

Changli Feng,<sup>1,2</sup> Jianxun Zhang,<sup>2</sup> and Rui Liang<sup>2</sup>

<sup>1</sup> Department of Information Science and Technology, Taishan University, Taian 271021, China

<sup>2</sup> Tianjin Key Laboratory of Intelligent Robotics, Institute of Robotics and Automatic Information System, College of Computer and Control Engineering, Nankai University, No. 94 Weijin Road, Tianjin 300071, China

Correspondence should be addressed to Jianxun Zhang; zhjx@robot.nankai.edu.cn

Received 11 June 2014; Revised 18 September 2014; Accepted 18 September 2014

Academic Editor: Kayvan Najarian

Copyright © 2015 Changli Feng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to get the extracted lung region from CT images more accurately, a model that contains lung region extraction and edge boundary correction is proposed. Firstly, a new edge detection function is presented with the help of the classic structure tensor theory. Secondly, the initial lung mask is automatically extracted by an improved active contour model which combines the global intensity information, local intensity information, the new edge information, and an adaptive weight. It is worth noting that the objective function of the improved model is converted to a convex model, which makes the proposed model get the global minimum. Then, the central airway was excluded according to the spatial context messages and the position relationship between every segmented region and the rib. Thirdly, a mesh and the fractal theory are used to detect the boundary that surrounds the juxtaleural nodule. Finally, the geometric active contour model is employed to correct the detected boundary and reinclude juxtaleural nodules. We also evaluated the performance of the proposed segmentation and correction model by comparing with their popular counterparts. Efficient computing capability and robustness property prove that our model can correct the lung boundary reliably and reproducibly.

## 1. Introduction

X-ray computed tomography (CT) is the most sensitive way for lung nodules detection [1], and detecting lung nodules using CT has become an increasingly important issue [2]. In this task, the lung region segmentation is a critical first step which can minimize the analytical region and system computation. Furthermore, the automated lung segmentation method is also needed due to the large number of axial slices that are produced by the multidetector-row CT scanner [3]. In addition, because juxtaleural nodules are contiguous with the chest wall and they have similar density to lung region, those nodules are always incorrectly excluded from the lung region. Thus the method, which can extract the lung region without the loss of any juxtaleural nodules from CT slices, should not only include the lung segmentation process, but also correct the defective boundary that surrounds the juxtaleural nodule.

Several methods have been proposed for segmenting lung region from CT slice images. The most used techniques are the intensity-based methods [4–7], which are based on the intensity difference between the lung tissue and other tissues. In this kind of models, to separate the whole region into two parts, a fixed threshold value is selected. Although this method is simple and fast to implement, it has an inherent limit that there is not an ideal fixed threshold that works well for all the images. Thus, the optimal threshold [8] is proposed, which gives an adaptive threshold for every slice image. But this model is also based on the theory of the fixed threshold method which makes the model lack the ability of segmenting the CT images that contain inhomogeneous intensity.

As another type of segmentation technique, the active contour model is also introduced to obtain a more accurate segmented result in the chest CT images. In [9–12], the proposed active contour models combine the gradient information or the curvature information which makes

the level set function stop on the boundary with high gradient or curvature values. However this kind of model does not easily converge to boundary concavity, which leads to a rough boundary [13]. As another type of information, the regional information is also used in active contour models [14–19]. Chan-Vese [14] proposes the well-known CV model which can segment the intensity homogeneous images quickly and robustly. Li [15] presents a local binary fitting (LBF) model which can segment the intensity inhomogeneous images successfully. Yang [20] combines the CV model, the LBF model, the global convex segmentation method, and the split Bregman method together to construct a new active model that is a global convex and can get the global minimum no matter where the initial contour lies.

As for the edge correction methods, the existing methods can be subdivided into two families: the morphology-based methods [5, 7, 21–24] and the geometry-feature-based models [25, 26]. The commonly used rolling-ball methods [7, 21, 22], which drive a ball along the edge contour to rein-clude juxtaleural nodules, belong to the morphology-based methods. In [5, 23, 24], sequences of erosion and dilation operations are combined to smooth the segmented lung boundary; they belong to the morphology-based methods too. Those morphology-based methods are easy and fast to implement, but the correction accuracy hugely depends on the radii of the morphology template. What is more, there are no proper radii that work well all the time. On the other side, the curvature information of the lung contour is used to identify the juxtaleural nodules [25, 26], but the curvature information is not a robust criterion for its sensitive properties to the rapid curvature changes and the noise. Yim and Hong [27] propose a contour tracing method to smooth the 2D contour; Pu et al. [28] present an adaptive border marching (ABM) method to rein-clude all the juxtaleural nodules; Yim et al. [2] correct the boundary by evolving the initial identified boundary with a defined scope and finding the optimal boundary. However, all those methods depend on the initial identified boundary greatly. If the initial boundary is recognized incorrectly, the correction method effect may be worse.

In this paper, a new lung extract method without the loss of the juxtaleural nodules is proposed by means of the CV model, the LBF model, the innovative edge detection function, the split Bregman method, and the geometric active contour model. First of all, a new edge detection function is defined with the help of the classical structure tensor theory. Then a new active contour model, which combines the global information, the local information, and the new edge information, is proposed. Besides, the proposed active contour model is converted to a convex segmentation model and is solved by the split Bregman method. After that, the central airway is excluded automatically. In the next step, the fractal theory is applied to detect the defective edge that is caused by juxtaleural nodules and the blood vessel and the airway. Finally, the geometric active contour model is introduced to correct that detected defective boundary.

The organization of the paper is as follows. The new edge detection function, the segmentation active contour model,

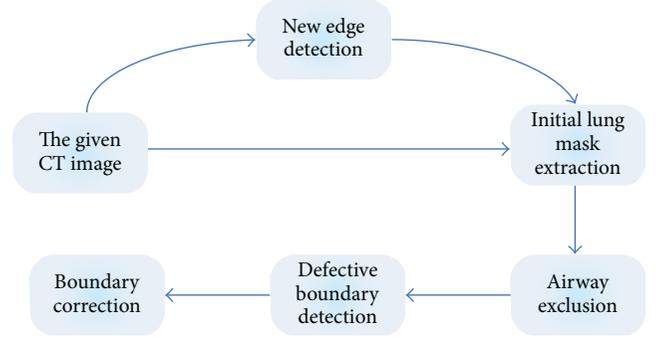


FIGURE 1: The schemes of the proposed lung correction method.

the correction geometric method, and other processes are discussed in Section 2. Numeric experiments are given in Section 3. The conclusion of the paper is offered in Section 4.

## 2. Methods

As shown in Figure 1, the proposed model consists of three major steps. Firstly, the lung initial mask is extracted by an active contour model. Secondly, according to the airway position in the preslice CT image and the location relationship between human organs and the rib tissue, the central airway in every slice of the CT scan is excluded automatically. Then, the lung region is divided into several smaller blocks using a grid line. Thirdly the fractal theory is used to detect the lung boundary that is contiguous to juxtaleural nodules. Finally, the geometric active contour model is introduced to correct the defective lung boundary that is caused by juxtaleural nodules.

*2.1. The New Edge Detection Function.* As a traditional edge detection function,  $g$  is defined as follows:

$$g(\mathbf{I}) = \frac{1}{1 + |\nabla(G_\sigma * \mathbf{I})|^2}, \quad (1)$$

where  $G_\sigma$  is a Gaussian kernel function with the standard deviation of  $\sigma$ . From (1), it is easy to find that  $g$  is sensitive to different noise sources and weak boundaries. So when  $g$  is used, it may make the level set evolution unstable if the detected edge is affected by different noise sources or weak boundaries.

In order to overcome the disadvantage, the classical structure tensor theory [29] is used to construct a new edge detection function. For a scalar image  $\mathbf{I}$ , the classical structure tensor  $J_\sigma$  is defined as follows:

$$J_\sigma = \begin{pmatrix} K_\sigma * I_x^2 & K_\sigma * I_x I_y \\ K_\sigma * I_x I_y & K_\sigma * I_y^2 \end{pmatrix}, \quad (2)$$

where  $K_\sigma$  is a Gaussian kernel function of size  $3 \times 3$  with the standard deviation  $\sigma$  and  $I_x$  and  $I_y$  denote the partial derivatives of the given image.

To be expressed simply,  $J_\sigma$  is written as  $\begin{pmatrix} J_{xx} & J_{xy} \\ J_{xy} & J_{yy} \end{pmatrix}$ . Note that the matrix  $J_\sigma$  is positive semidefinite; its eigenvalues are

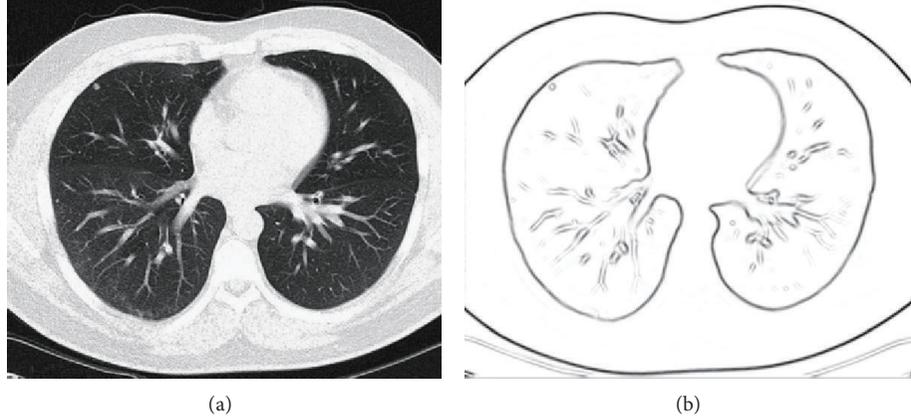


FIGURE 2: The edge detection result of  $g^*$ . (a) The original image; (b) the detected edge.

labeled as  $\rho_1, \rho_2, \rho_1 \geq \rho_2$ . What is more, it is easy to find that  $\rho_1 \approx \rho_2 \approx 0$  in the flat region,  $\rho_1 \gg \rho_2 \approx 0$  when the point is on the straight line edge, and  $\rho_1 \gg \rho_2 \gg 0$  when it is a corner point. That is, the difference  $\Delta\rho$  between the two eigenvalues gets a small value in the flat region and a big value on the edge of the image, respectively.

Thus we can select a threshold to distinguish those two regions from each other; let  $S_T = \{x \mid x \in T\}$ ,  $T = \rho_1 - \rho_2$ ; then the threshold is defined as follows:

$$T_c = \bar{T} + \kappa\sigma_T, \quad (3)$$

where  $\bar{T}$  and  $\sigma_T$  are the mean and the standard deviation of  $S_T$ , respectively, and  $\kappa$  is a constant-weight that controls the difference between  $T_c$  and  $\bar{T}$ .

By virtue of  $T_c$ , a new edge detection function is calculated as follows:

$$g^*(x, y) = \begin{cases} 0, & T(x, y) \leq T_c \\ 1, & T(x, y) > T_c. \end{cases} \quad (4)$$

In Figure 2(b), there is the detected edge of  $g^*$ , from which it is easy to find that  $g^*$  gives a perfect edge of the image in Figure 2(a). Furthermore, this result is very similar to the observation of human. What is important is that, except for offering of the edge information to the proposed active contour model,  $g^*$  can also serve as an adaptive weight of the proposed model, which is shown in the following subsection.

**2.2. Initial Lung Mask Extraction.** Due to the partial volume effect (PVE) or the similar density between the juxtaleural nodules and the chest wall, CT images often have inhomogeneous intensity around the lung boundary. However, the most of existing methods cannot segment those CT images accurately. So, we propose a new combined model to segment this kind of CT images. The new model is a combination of the global-region-information and local-region-information with an adaptive weight. Besides, we convert the new active contour model to a convex minimization model which can ensure the robustness and efficiency of segmentation result.

As mentioned above, CV model and LBF model are active contour models (ACM) that are based on global intensity information and local region intensity information correspondingly; thus those two models are integrated into the new model dynamically to inherit their advantages. Its energy function is as follows:

$$E^{GLg} = \alpha E^{CV} + \beta E^{LBF} + \gamma E_{g^*}^L + \nu E^R, \quad (5)$$

where  $\alpha, \beta, \gamma$ , and  $\nu$  are four positive constants which control the energy ratio of the CV model, the LBF model, the length term, and the regularization term correspondingly in the proposed energy function;  $E_{g^*}^L$  and  $E^R$  are the length term and the regularization term correspondingly;  $E^{CV}$  and  $E^{LBF}$  are the energy function of CV model and LBF model, which can be calculated by

$$\begin{aligned} E^{CV} &= \sum_{i=1}^2 \eta_i M_i(\phi(x)) |I(x) - c_i|^2 dx, \\ E^{LBF} &= \sum_{i=1}^2 \xi_i \int \int K_\omega(x-y) |I(y) - f_i(x)|^2 M_i(\phi(y)) dy dx \\ E_{g^*}^L &= \int_{\Omega} g^*(x) \delta_\epsilon(\phi(x)) |\nabla\phi(x)| dx \\ E^R &= \int_{\Omega} |\nabla\phi(x) - 1|^2 dx, \end{aligned} \quad (6)$$

where  $\eta_1, \eta_2, \xi_1$ , and  $\xi_2$  are four positive constants that balance the intensity-based energy in the inner and outer of the curve  $C$  and  $K_\omega$  denotes a Gaussian kernel function with standard

deviation of  $\omega$  and window size of  $(4\omega + 1) \times (4\omega + 1)$ . The rest of the terms are defined as follows:

$$\begin{aligned} H_\varepsilon(x) &= \frac{1}{2} \left[ 1 + \frac{\pi}{2} \arctan\left(\frac{x}{\varepsilon}\right) \right] \\ \delta_\varepsilon(x) &= \frac{1}{\pi} \frac{\varepsilon}{\varepsilon^2 + x^2} \\ M_1 &= H_\varepsilon(x), \quad M_2 = 1 - M_1 \\ c_i &= \frac{\int_\Omega M_i(\phi(x)) I(x) dx}{\int_\Omega M_i(\phi(x)) dx}, \quad i = 1, 2 \\ f_i &= \frac{K_\omega * (M_i(\phi(x)) I(x))}{K_\omega * M_i(\phi(x))}, \quad i = 1, 2 \\ \alpha &= \alpha' \cdot g^*, \quad \beta = \beta' (1 - g^*), \end{aligned} \quad (7)$$

where  $\alpha'$  and  $\beta'$  are two positive constants that denote the adaptive weights of the first two terms in (5).

It is worth noting that the proposed  $g^*$  has two roles in (5). One is the factor of the length term which can drive the evolving level set function toward the detected  $g^*$ . By this term the computational efficiency will be improved. The other role is the adaptive weight, which can assign appropriate combined energy to every point. With the help of this term, the edge point will be given more local information which helps to determine the precise boundary. It is the major improvement of our proposed method which makes our proposed model outperform their  $g$ -based counterparts [20].

The energy function (5) can be minimized by solving the following gradient flow:

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= \alpha \delta_\varepsilon(\phi) [-\eta_1 |I - c_1|^2 + \eta_2 |I - c_2|^2] \\ &+ \beta \delta_\varepsilon(\phi) [-\xi_1 e_1 + \xi_2 e_2] + \nu \left( \nabla^2 \phi - \operatorname{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) \right) \\ &+ \gamma \delta_\varepsilon(\phi) \operatorname{div} \left( g^* \frac{\nabla \phi}{|\nabla \phi|} \right), \end{aligned} \quad (8)$$

where  $e_i(x) = \int K_\omega(y - x) |I(x) - f_i(y)|^2 dy$ .

As most active contour models, new model (5) is also prone to the local minimum. As mentioned above, LCV model is prone to the local minimum. To overcome this defect, (5) is converted to a convex minimization problem; then it is solved by the split Bregman method. First of all, in general, let  $\xi_1 = \xi_2 = \eta_1 = \eta_2 = 1$ ; then (8) is written as follows:

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= \gamma \delta_\varepsilon(\phi) \operatorname{div} \left( g^* \frac{\nabla \phi}{|\nabla \phi|} \right) + \mu \delta_\varepsilon(\phi) (F_1 + F_2) \\ &+ \nu \left( \nabla^2 \phi - \operatorname{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) \right), \end{aligned} \quad (9)$$

where

$$\begin{aligned} F_1 &= \alpha' g^* \left( -(\mathbf{I} - c_1)^2 + (\mathbf{I} - c_2)^2 \right) \\ F_2 &= \beta' (1 - g^*) (-e_1 + e_2). \end{aligned} \quad (10)$$

Note that (9) is nonconvex too; however, it can be transformed into a convex one by the globally convex segmentation (GCS) method [30]. To apply this method, the third term in (9) should be dropped. In fact, according to [20], the regularization term is used to eliminate the reinitialization process and maintain the level set function as an approximate signed distance function near the zero level set. At the same time, this term is not contained in the classical region-based models, such as the famous CV models. With only the data fitting term and the arc length term, these models work well. Actually, the reinitialization process is not encouraged for most experiments since it may cause some subtle side effects, such as preventing the detection of interior boundaries within an object, as pointed out in CV model [14]. Furthermore, in split Bregman method, we restrict  $\phi$  to the interval  $[0, 1]$ . In this way, the level set function  $\phi$  will not blow up to very large values on both sides of the zero level set and will not cause inaccurate computation or erroneous segmentation results. Thus dropping this term is reasonable and will not affect the segmentation results of the model.

Let  $\gamma = 1$ ; then we get the new gradient flow as follows:

$$\frac{\partial \phi}{\partial t} = \delta_\varepsilon(\phi) \left( \operatorname{div} \left( g^* \frac{\nabla \phi}{|\nabla \phi|} \right) + \mu (F_1 + F_2) \right). \quad (11)$$

With the help of the GCS method, we drop  $\delta_\varepsilon(\phi)$  to get a simplified gradient flow. It is worth noting that, through  $\delta_\varepsilon(\phi)$  which is dropped, the optimality solution of the simplified gradient flow is also equivalent to its original form [20]. The simplified form is defined as follows:

$$\frac{\partial \phi}{\partial t} = \left[ \operatorname{div} \left( g^* \frac{\nabla \phi}{|\nabla \phi|} \right) + \mu (F_1 + F_2) \right]. \quad (12)$$

Based on (12), a new energy function can be constructed as follows:

$$E^{\text{GLg}}(\phi) = \int g^* |\nabla \phi(\mathbf{x})| d\mathbf{x} + \int \mu \phi(\mathbf{x}) s(\mathbf{x}) d\mathbf{x}, \quad (13)$$

where  $s(\mathbf{x}) = -(F_1(\mathbf{x}) + F_2(\mathbf{x}))$ .

It is interesting to find that (13) has the same solution as (11). Thus the purpose becomes finding the minimum of (13).

To guarantee model (13) gets the unique global minimum, the range of  $\phi$  is restricted within a bounded interval. In this paper, the interval is  $[0, 1]$ . Thus (13) can be written as follows:

$$\begin{aligned} \min_{\phi \in [0,1]} E^{\text{GLg}}(\phi) &= \min_{\phi \in [0,1]} \left[ \int g^*(\mathbf{x}) |\nabla \phi(\mathbf{x})| d\mathbf{x} \right. \\ &\quad \left. + \int \mu \phi(\mathbf{x}) s(\mathbf{x}) d\mathbf{x} \right]. \end{aligned} \quad (14)$$

Generally, (14) can be expressed as

$$\min_{\phi \in [0,1]} E^{\text{GLg}}(\phi) = \min_{\phi \in [0,1]} \left[ |\nabla \phi|_{g^*} + \mu \langle \phi, s \rangle \right], \quad (15)$$

where

$$\begin{aligned} |\nabla\phi|_{g^*} &= \int g^*(\mathbf{I}(\mathbf{x})) |\nabla\phi(\mathbf{x})| d\mathbf{x}, \\ \langle\phi, \mathbf{s}\rangle &= \int \phi(\mathbf{x}) s(\mathbf{x}) d\mathbf{x}. \end{aligned} \quad (16)$$

In order to apply the split Bregman method, we present a new term  $\vec{\mathbf{d}}$  and a penalty term that can ensure  $\vec{\mathbf{d}}$  approximate to  $|\nabla\phi(\mathbf{x})|$  in the iteration. By those two terms, (14) is reformatted into the following form:

$$\min_{\phi \in [0,1]} E^{\text{GLg}}(\phi) = \min_{\phi \in [0,1]} \left[ |\vec{\mathbf{d}}|_{g^*} + \mu \langle\phi, \mathbf{s}\rangle + \frac{\lambda}{2} \|\vec{\mathbf{d}} - \nabla\phi\|^2 \right], \quad (17)$$

where  $\lambda$  is a positive constant parameter.

Then the Bregman iteration is used to meet the condition  $\vec{\mathbf{d}} = |\nabla\phi(\mathbf{x})|$  [31], so we get the following optimization problem:

$$\begin{aligned} (\phi^{k+1}, \vec{\mathbf{d}}^{k+1}) &= \arg \min_{\phi \in [0,1], \vec{\mathbf{d}}} \left( |\vec{\mathbf{d}}|_{g^*} + \mu \langle\phi, \mathbf{s}\rangle \right. \\ &\quad \left. + \frac{\lambda}{2} \|\vec{\mathbf{d}} - \nabla\phi - \vec{\mathbf{d}}^k\|^2 \right), \end{aligned} \quad (18)$$

where  $\vec{\mathbf{b}}^{k+1}$  is got by the Bregman iteration:

$$\vec{\mathbf{b}}^{k+1} = \vec{\mathbf{b}}^k + (\nabla\phi^{k+1} - \vec{\mathbf{d}}^{k+1}). \quad (19)$$

The optimization solution of  $\phi^{k+1}$  is obtained by the optimization condition:

$$\Delta\phi = \frac{\mu \cdot s}{\lambda} + \nabla(\vec{\mathbf{d}} - \vec{\mathbf{b}}), \quad \phi \in [0, 1]. \quad (20)$$

For (20), we get the approximated solution of  $\phi$  by means of the Gauss-Seidel method [31], which is computed by

$$\begin{aligned} \alpha_{i,j} &= d_{i-1,j}^{x,k} - d_{i,j}^{x,k} + d_{i,j-1}^{y,k} - d_{i,j}^{y,k} \\ &\quad - b_{i-1,j}^{x,k} + b_{i,j}^{x,k} - b_{i,j-1}^{y,k} + b_{i,j}^{y,k}, \\ \beta_{i,j} &= \frac{1}{4} \left( \phi_{i-1,j} + \phi_{i+1,j} + \phi_{i,j-1} \right. \\ &\quad \left. + \phi_{i,j+1} - \frac{\lambda \cdot s}{\mu} + \alpha_{i,j} \right), \\ \phi_{i,j} &= \max \left\{ \min \left\{ \beta_{i,j}, 1 \right\}, 0 \right\}. \end{aligned} \quad (21)$$

Finally, the solution of  $\vec{\mathbf{d}}^{k+1}$  is acquired by the Shrink process [31]:

$$\begin{aligned} \vec{\mathbf{d}}^{k+1} &= \text{shrink}_{g^*} \left( \vec{\mathbf{b}}^k + \nabla\phi^{k+1}, \frac{1}{\lambda} \right) \\ &= \text{shrink} \left( \vec{\mathbf{b}}^k + \nabla\phi^{k+1}, \frac{g^*}{\lambda} \right), \end{aligned} \quad (22)$$

where the Shrink process is defined as follows:

$$\text{shrink}(x, \theta) = \begin{cases} \frac{x}{|x|} \max(|x| - \theta, 0), & x \neq 0 \\ 0, & x = 0. \end{cases} \quad (23)$$

As shown in Figure 3, for the given images  $\mathbf{I}$ , update  $s$  in (21) by calculating the mean intensity in the regions  $\phi > 0$  and  $\phi < 0$ . Then update  $\phi$  according to (21) until its convergence is achieved. By this way, the final segmentation result is obtained by the boundary of the following set:

$$\{x \in \Omega \mid \phi^{\text{final}}(x) > 0.5\}. \quad (24)$$

The segmentation result of the given image in Figure 3(a) is shown in Figure 3(b), and the binary image of the segmentation result is displayed in Figure 3(c).

**2.3. Airway Exclusion.** Because the inner of airway has the approximate intensity with the lung region, it is often included in the final segmentation result of the intensity-information-based image segmentation model. However, this part of tissue, especially the central airway, is useless for lung nodules detection; thus it needs to be excluded from the above segmentation result.

In fact, it is easy to find that the trachea and bronchi out of the lung region are faraway from the ribs and there is less osseous tissue in its neighborhood, while the outer boundary of the lung region is contiguous with the rib. Thus, for every subregion of the segmentation result, a morphology dilation process is applied to detect whether there is an osseous tissue in the dilated region. If none of the osseous tissues exists, this subregion is excluded from the lung region of the current CT slice. Until all the slices in the CT scan are processed, the majority of those regions are removed already.

However, the performance of the exclusion process is not always perfect all the time. For some slices of some patients, the tracheal cartilages around the trachea and bronchi out of the lung region are very obvious; the dilated neighborhood region contains some tracheal cartilage that is regarded as the osseous tissue, so the inner parts of them are preserved.

In order to tackle this problem, a further process is introduced. In the CT slice, it is seen that the tracheal cartilage is obvious in some slice, but it does not always exist in the previous or successive slices for its size limit in the  $z$ -axis. What is more, it is interesting to observe that the inner parts of the trachea and bronchi are excluded in those tracheal cartilage vanished slices. Besides, the locations of organs are approximately similar to each other among the adjacent slices. Therefore we can use these two properties to exclude the inner region of the central airway. In this method, from the second slice of the CT scan, each subregion in the current slice is checked for whether the equivalent location in the previous slice is excluded. If the same subregion is removed, the current region will be excluded from lung region too. Until all slices are handled, the major central airway-surrounded regions have been excluded already. In addition, the consequence of subsection is also given in Figure 4.

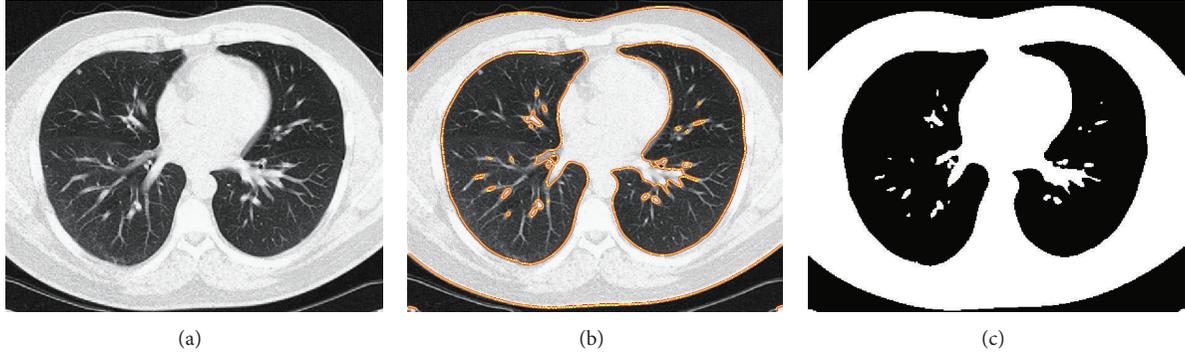


FIGURE 3: The segmentation result and the obtained initial mask. (a) The original image; (b) the segmentation result; (c) the obtained initial lung mask.

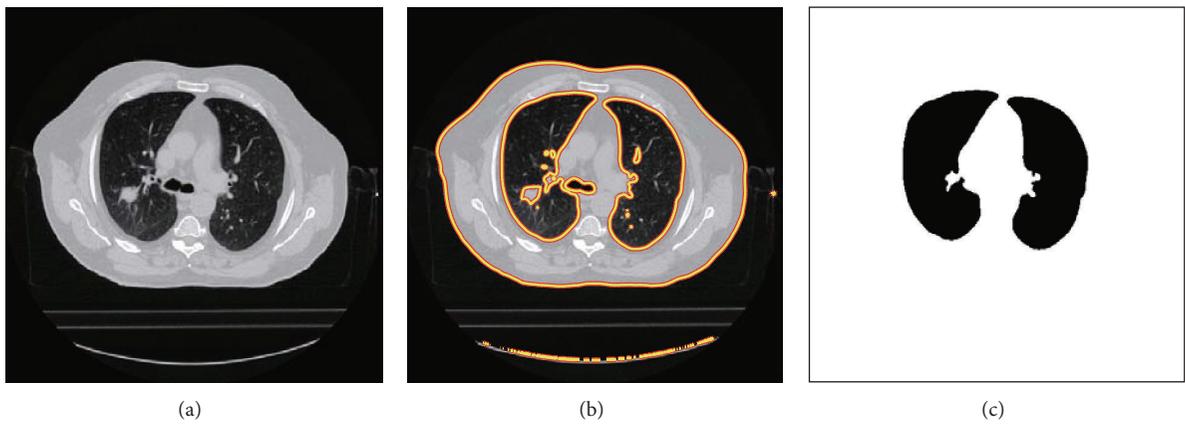


FIGURE 4: The airway-free initial lung mask. (a) The original image; (b) the segmentation result; (c) the result after extracting central airway.

**2.4. Defective Boundary Detection.** For juxtapleural nodules, they have the approximate intensity with the surrounding area, so they are often incorrectly excluded from the lung region. But those nodules hold a higher rate of being malignant tumors than other inner nodules, so the performance of the CAD system will be affected if those nodules are not reincluded into the lung region.

Hence, according to the fact that the contour of the lung region is smooth and the smoothness of the boundary on the segmented image is damaged by the juxtapleural nodule, a new boundary detection method is proposed to detect the defective boundary by means of the fractal theory.

First of all, the minimum enclosing rectangle (MER) of the lung region on 2D slice is acquired; then it is uniformly divided into  $10 \times 10$  blocks by a grid line. The number of the blocks is chosen according to the actual size of the lung region and the size of lung nodules. Then the blocks that contain the boundary of the segmentation result, which is called the boundary block, are detected. In order to visualize the detection process, the result of this subsection is shown in Figure 6. Note that when there is only one lung subregion, the blocks should be set to  $2 \times 2$  to avoid the phenomenon that the size of the block is too small to carry the operation on it.

For every boundary block, the fractal theory is adopted to account the fractal dimension of the inner boundary

line. Furthermore, for easy and automatic implementation, the box-counting method is selected to compute the fractal dimension among so many techniques [32]. By this method, every boundary block is covered by a series of grids, whose sizes have a progressive decrease. For each of the grids, the following two values are recorded: the number of square boxes intersected by the image,  $N(s)$ , and the side length of the squares,  $s$ . The regression slope  $D$  of the straight line formed by plotting  $\log(N(s))$  against  $\log(1/s)$  is the fractal dimension of the current block which can be evaluated as follows [32]:

$$\log(N(s)) = \log(K) + D \log\left(\frac{1}{s}\right), \quad (25)$$

where  $K$  is a constant and  $N(s)$  is proportional to  $(1/s)^{-D}$ .

As mentioned above, the normal boundary of the lung region on 2D slice is smooth, so the fractal dimension of the block that contains those kinds of boundary is smaller than that of the boundary block that contains the defective boundary. Until all the boundaries are calculated by the box-counting method, all the boundary blocks can be categorized into two classes: the blocks with a small fractal dimension that contain the normal smooth boundary and the blocks with a big value that contain the defective boundary that is caused by the juxtapleural nodule, which can be seen clearly from

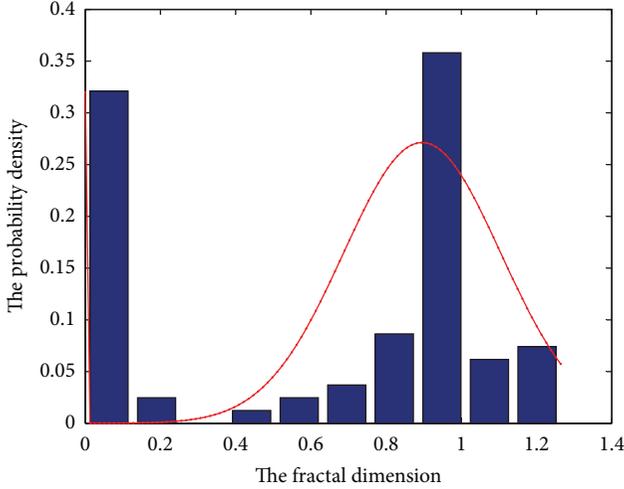


FIGURE 5: The histogram of all the fractal dimension values in all edge blocks.

the histogram in Figure 5. Thus we can select an appropriate threshold to distinguish those blocks into classes.

Let  $d_i$  denote the fractal dimension of the  $i$ th boundary block;  $N_b$  indicates the overall number of the boundary block,  $S_f = \{d_i \mid i = 1, 2, \dots, N_b\}$ ; the threshold  $T_f$  can be defined as follows:

$$T_f = \bar{d} + s_d, \quad (26)$$

where  $\bar{d}$  and  $s_d$  are the mean value and the standard deviation of  $S_f$ , respectively. Furthermore, the threshold in (26) is given by the strategy of trial-and-error empirically. In our experiments, both the mean and the valley value of the histogram are not suitable; thus we select an optimal value after many experiments.

With the help of  $T_f$ , the block is regarded as the block that contains the defective boundary if its fractal dimension is bigger than  $T_f$ . Then those detected blocks will be corrected in the following subsection. As shown in Figure 6(c), the detected defective boundary by  $\bar{T}_f$  is very accurate, and it is nearly similar to the observation of human. Note that it is the first time that the fractal theory is used to identify the defective lung boundary that surrounds juxtapleural nodules in CT slices. Besides, the lung-inner blood vessel is also excluded by the segmentation method for their higher density. In fact, those regions should reinclude the lung region to get an accurate nodule-detection result. It is interesting that, except the defective boundary that is caused by juxtapleural nodules, our model can detect those boundaries that are caused by those blood vessels too and regard them as the defective boundary. Furthermore, previous methods only take local properties of the boundary into account, which makes the model sensitive to the robustness of those local properties. But the proposed model integrated the global boundary properties with the local boundary properties together by the statistical method. Thus the detected result will be more robust to different shapes of the lung boundary.

**2.5. Boundary Correction Based on Geometric Deformable Models.** As shown in Figure 7(a), the whole image is divided into four parts:  $\Omega_1$ ,  $\Omega_2$ ,  $\Omega_3$ , and  $\Omega_4$ .  $\Omega_1$  is the outside of the lung region except for those points in the defective boundary blocks,  $\Omega_2$  denotes inner parts of the lung region,  $\Omega_3$  indicates the set of boundary points in the correct boundary block, and the detected defective boundary block region is labeled as  $\Omega_4$ .

Besides, according to the purpose of the correction process and the boundary information, we also give the different curve evolving strategies for all four kinds of regions. In  $\Omega_1$ , if the evolving curve passes through the real boundary, the curve should be shrunk. Conversely, in  $\Omega_2$ , the curve will be expanded if the curve converges into the inner of the lung. For  $\Omega_3$ , due to those edge points on the accurate boundary already, there is no need to correct them. Therefore, if the curve is on those points, it should be kept still to avoid useless computation. Finally, for  $\Omega_4$ , because they contain the defective boundary, the hole should be filled while keeping the boundary smooth. Thus the curve evolution is proposed to expand the defective boundary in the hole by a balloon-force-like technique. As shown in Figure 7(b), by this way, the defective inner boundary will be expanded to the ideal position all the time.

From Figure 7, what needs to be done is expanding the whole lung region from the inner part of the lung region by a balloon-force-like power until we get an optimal smooth boundary while keeping the correct boundary still in current position. As a result, the defective boundary is corrected; the other boundary performs the same state as before.

First of all, we propose the simplified geometric active contour which is written as follows:

$$\frac{\partial \phi}{\partial t} = \mathbf{k} |\nabla \phi|, \quad \phi(x, 0) = \phi_0, \quad (27)$$

where  $\mathbf{k}$  is a speed parameter which controls the speed and orientation of the evolution curve. In order to make the evolution stop on the correct boundary,  $\mathbf{k}$  is defined by

$$\mathbf{k}(x) = \begin{cases} 1, & x \in \Omega_1 \\ 0, & x \in \Omega_3 \\ -1, & x \in \{\Omega_2, \Omega_4\}. \end{cases} \quad (28)$$

The motivation of (28) is that  $k = -1$  can make the balloon force outward to expand the boundary curve, while  $k = 1$  can let the force inward to shrink the boundary curve. Furthermore,  $k = 0$  can keep the right boundary still and protect those points from impact and ensure the final correction result is precise.

However, without any restrictive conditions, the curve will keep expanding itself on the defective boundary all the time. And it may overpass the real lung boundary and arrives at an inappropriate position at last. Thus the curve length

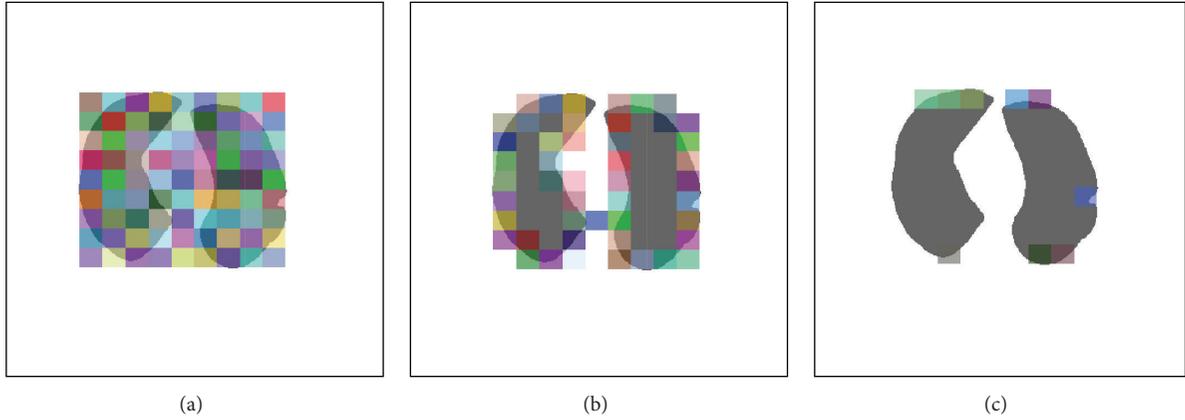


FIGURE 6: The process of detecting defective boundary block. (a) The whole block setting; (b) the boundary block; (c) the detected defective boundary block.

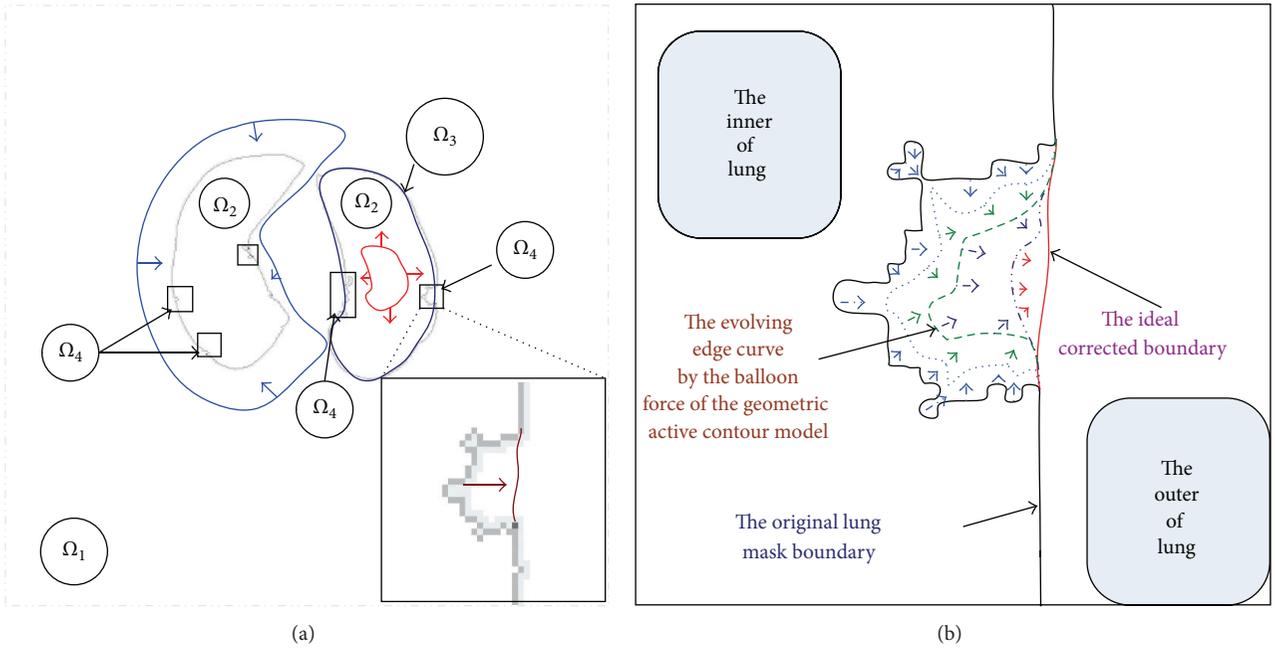


FIGURE 7: The figure of the geometric active contour model and the purpose of the proposed correction method. (a) The general description of the geometric active contour model; (b) the purpose of our method.

term and the area term are introduced by the assumption that the final evolution curve should be as smooth as possible. The introduced terms are defined by

$$\begin{aligned}
 E^L &= \int_{\Omega} \delta_{\varepsilon_1}(\phi) |\nabla \phi| dx, \quad x \in \Omega. \\
 E^A &= \int_{\Omega} H_{\varepsilon_1}(\phi(x)) dx, \quad x \in \Omega,
 \end{aligned}
 \tag{29}$$

where  $\varepsilon$  is a parameter that controls the nonzero interval in the delta Dirac function  $\delta_{\varepsilon_1}(x)$ . Their gradient flows are calculated by

$$\begin{aligned}
 \frac{\partial \phi}{\partial t_L} &= \delta_{\varepsilon_1}(\phi) \operatorname{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right), \\
 \frac{\partial \phi}{\partial t_A} &= \delta_{\varepsilon_1}(\phi).
 \end{aligned}
 \tag{30}$$

To maintain the smoothness of the evolution curve, the curve length term is embedded into the geometric active

contour model; thus the new geometric active contour model is defined by

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= \mathbf{k} |\nabla \phi| + \kappa_1 \frac{\partial \phi}{\partial t_L} + \kappa_2 \frac{\partial \phi}{\partial t_A} \\ \phi(x, 0) &= \phi_0, \end{aligned} \quad (31)$$

where  $\phi_0$  denotes the final lung mask of Section 2.3 and  $\kappa_1$  and  $\kappa_2$  denote the gradient flow distribution in the proposed geometric active contour model.

In order to maintain the evolution of the level set function stable, the Gaussian convolution method which is proposed by Zhang et al. [17] is introduced into this model. The level set update function by the Gaussian kernel function is as follows:

$$\phi'_n = \phi_{n-1} + \Delta t \Delta \phi, \quad \phi_n = G_{\sigma_1} * \phi'_n, \quad (32)$$

where  $G_{\sigma_1}$  is a Gaussian kernel function with the standard deviation of  $\sigma_1$ ,  $\Delta t$  is the time step, and  $\Delta \phi$  denote the result which is obtained by the first equation of (31). Besides,  $\phi_n$  serves as the initial contour for the next iteration. To the best of our knowledge, it is the first time that the geometric active model is proposed to correct the lung edge in computer-aided diagnosis technique.

In the correction process, the initial contour for the geometric active contour model is set as the result of the initial lung region mask without the airway and background. Then update the level set function  $\phi$  according to (32) until it arrives at a stable state. Finally, the corrected boundary, which is shown in Figure 8, is given as the zero level set of  $\phi$ . In the figure, both the result of the correction method and the evolving active contour are provided. Particularly in Figure 8(c), there are all the curves in the entire correction process, which also implies the perfect effect of the proposed correction method.

### 3. Experiments

In this section, the experiments are performed on the proposed model. The experiments focus on the following aspects: (1) the superior effect of  $g^*$  over  $g$ ; (2) the global minimum detecting ability of the proposed model; (3) the advantages of GLg over the state-of-the-art active contour model; (4) the correction effect analysis. The experiments are finished on a notebook computer with Intel 2.10 GHz CPU and Matlab 8.0. The parameters in the proposed model are set as follows:  $\alpha' = 0.9$ ,  $\beta' = 0.1$ ,  $\lambda = 1 \times 10^3$ ,  $\mu = 3\lambda$ ,  $\sigma = 3$ ,  $\kappa = -0.05$ ,  $\omega = 1$ ,  $\varepsilon = 1$ ,  $\varepsilon_1 = 0.5$ ,  $\kappa_1 = 5n$ ,  $\kappa_2 = 0.1$ , and  $n$  denotes the iteration times. The used 3 CT data sets are obtained from the hospital in Guangzhou, China; the rest of 25 sets are selected from the Lung Image Database Consortium (LIDC).

First of all, we test the boundary detecting capability of  $g^*$  in the given image. In order to display the advantages of  $g^*$ , the traditional edge detection function  $g$  is introduced too. The detected edges are shown in Figure 9, from which it is easy to find that  $g$  is affected by noise and additional background. However, due to the usage of the statistical information,  $g^*$  is insensitive to the noise and background and gets a clear edge-information-contained image. It is

worth noting that  $g^*$  gets all the edges of our interesting structures except for some ribs and the spine.

In order to demonstrate the global minimum detection capability of the proposed model, we select two different rectangle initial contours which are shown in Figures 10(a) and 10(c) correspondingly. The segmentation results are shown in Figures 10(b) and 10(d). From this experiment, it is easy to find that the proposed model gets the uniform segmentation results under two different initial contours, from which we can know that the proposed model can get the global minimum no matter what kinds of initial contour we select.

In the following experiment, we verify the advantages of GLg model over the popular optimal threshold method. The segmentation results are given in Figure 11, in which Figure 11(a) shows the segmentation result of the optimal threshold method and Figure 11(b) displays the segmentation result of the proposed model. From the results, it is clearly seen that the optimal threshold method fails to segment this image for there is intensity inhomogeneity in the box region. However, with the help of  $g^*$ , GLg model detects the true boundary in the box region. Although a circular region in the box region is identified as the outer part of the lung, this region can be corrected by filling the inner hole by the morphology method; thus this part of region is not affected by the final segmentation result. Note that the hole-filling method is a common process in the previous models, and it is not our special process to deal with the segmentation result.

To verify the segmentation ability of the proposed model, we compare the proposed model with some other models. The segmentation results are listed in Figure 12. Figure 12(a) gives the initial contour and the original image, and Figure 12(b) shows the segmentation result by an expert. Figures 12(c)–12(f) provide the segmentation results of LBF model [15], LIF model [17], GCLGIF [20] model, and the proposed model correspondingly. From those figures, it is easy to find that LBF model and LIF model are sensitive to the initial contour. As combined models, GCLGIF model and the proposed model can achieve a satisfactory segmentation result. The segmentation results of the last two models approximate the segmentation result of the expert. Particularly, it is seen that the proposed model overcomes GCLGIF model in the following experiment.

For evaluating the proposed model quantitatively, another globally convex local and global intensity fitting energy (GCLGIF) model, which is also a global-and-local-combined active contour model, is introduced in this experiment. The major difference between those two models is the different edge information and different adaptive weight function. In GCLGIF model, the edge information is given by  $g$ ; the weight function is  $\omega$ . In GLg method, the edge information and the weight are obtained from  $g^*$  simultaneously which can give more useful edge information for evolving the level set function. The experiment is performed on a CT set which contains 141 CT slices; the initial active contour for both kinds of models is obtained by the threshold of  $-500$  Hounsfield units (HUs) [5]. It is worth noting that although the proposed segmentation model can segment images accurately in every kind of initial contour

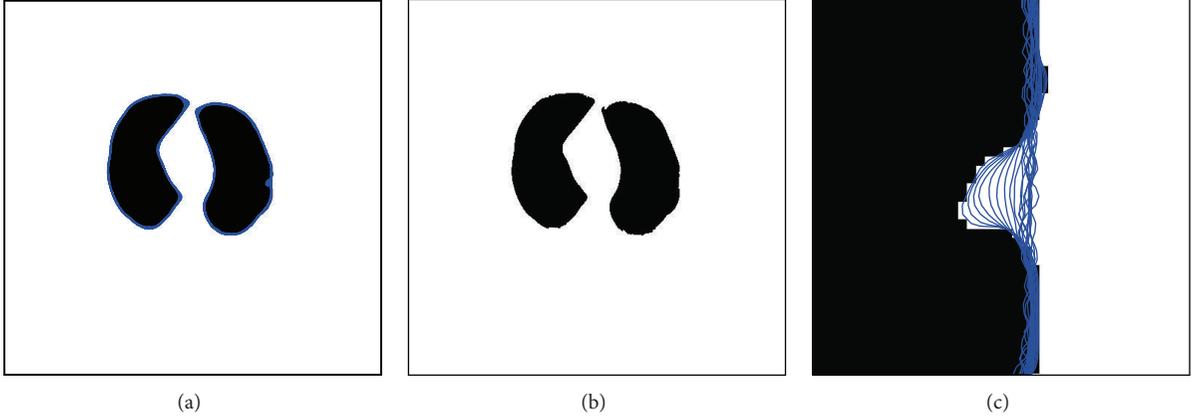


FIGURE 8: The corrected results. (a) The evolved active contour; (b) the final correction result; (c) the corrected part by the proposed method.

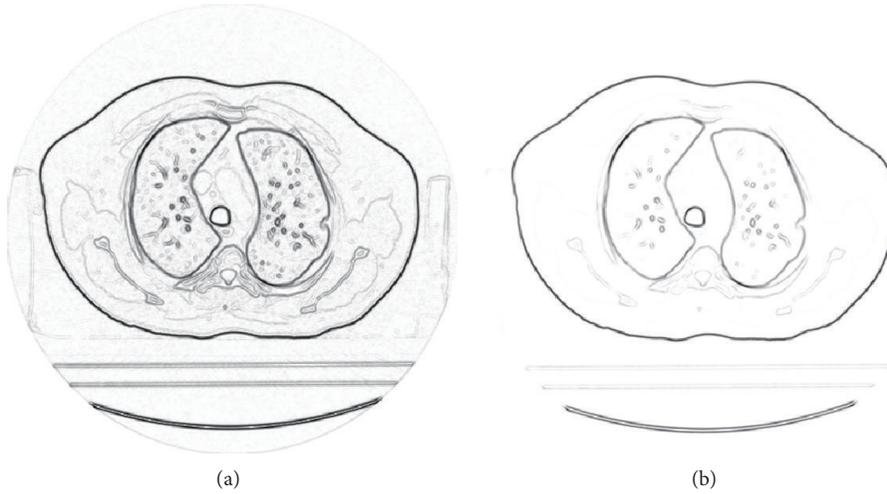


FIGURE 9: The edge detection capability of  $g$  and  $g^*$ . (a) The detected edge by  $g$ ; (b) the detected edge by  $g^*$ .

which can be found in Figure 10, the above initial contour is used for reducing computational time; it does not give a better initial contour to the segmentation model. In the experiment, if we take the GCLGIF segmentation result as the golden standard, the mean segmentation accuracy of GLg method in those 141 slices achieves 99.97%. Additionally, due to the different edge information and weight function, GLg needs less time to segment all slices than GCLGIF model. Besides, the standard deviation and the coefficient of variation of all the 141 time data are also compared which are shown in Table 1. Besides, we also offer the time improved rate (TIR) in Table 1, which is defined by

$$\text{TIR} = \frac{t_1 - t_2}{t_1} \times \%, \quad (33)$$

where  $t_1$  denotes the time of the GCLGIF model and  $t_2$  denotes the time of the GLg model.

It should be noted that not only the mean time but also the standard deviation and the coefficient of variation are smaller than those of the GCLGIF model. Besides, from

TABLE 1: The computational effect comparison between  $g$ - and  $g^*$ -based models.

Model	$\bar{t}$	Std	CoV	TIR
GCLGIF	0.37799	1.1139	2.9468	—
GLg	0.3543	0.80846	2.2818	6%

the TIR value, it is known that, under the nearly same accuracy, GLg model is more efficient than GCLGIF model.

In order to display the correction effect intuitively, the 3D model of the segmented lung and the edge corrected lung is reconstructed in Figure 13. From Figure 13(a), it is easy to find that there is a hole that surrounds the juxtaleural nodule in the signed circle region. After edge correction, the hole is almost filled by the geometric active contour model, which can be seen in Figure 13(b). It is interesting to observe that below the circular region of Figure 13(b) the rib profile is more obvious after edge correction. Unfortunately, the hole is not filled perfectly; there is a tiny defect in the circular region of Figure 13(b). The leading reason is that the evolving level set function should be regularized as a signed distance

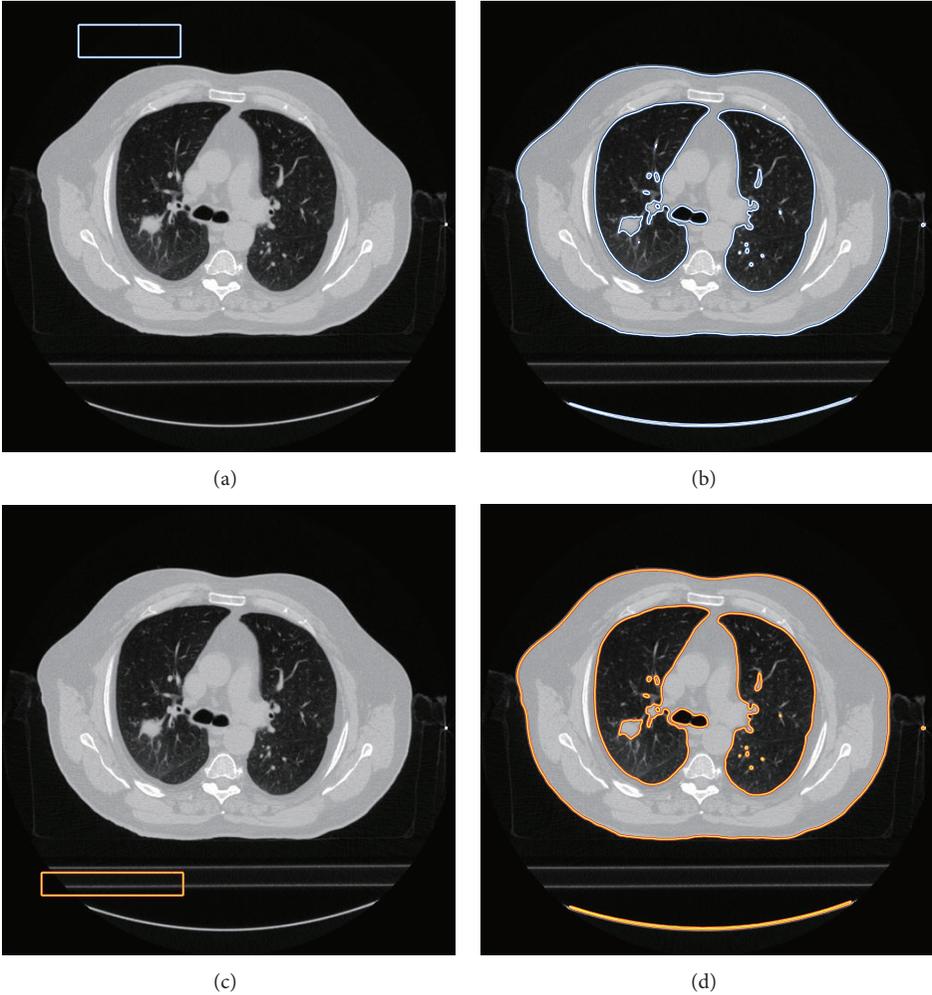


FIGURE 10: The segmentation results under two different initial contours. (a) and (c) are two initial contours; (b) and (d) are the segmentation results.

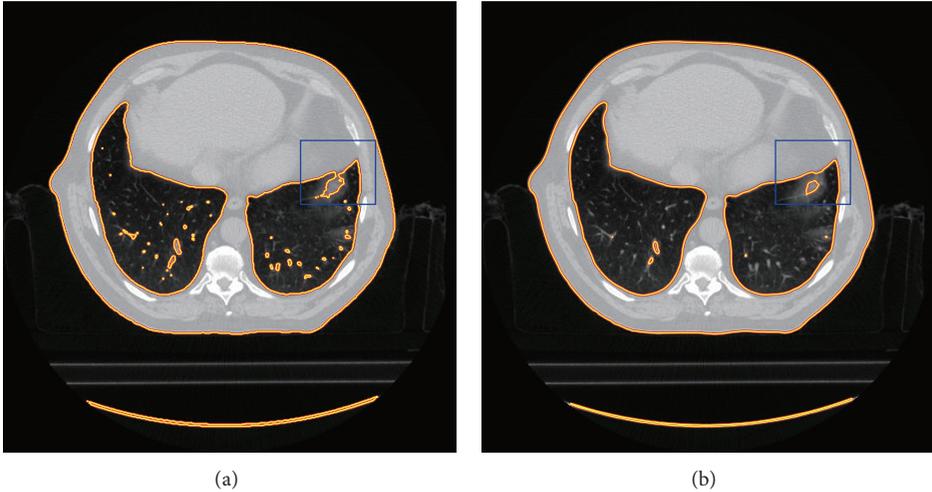


FIGURE 11: The comparison between the optimal threshold method and GLg model. (a) The segmentation result of the optimal threshold method. (b) The segmentation result of GLg model.

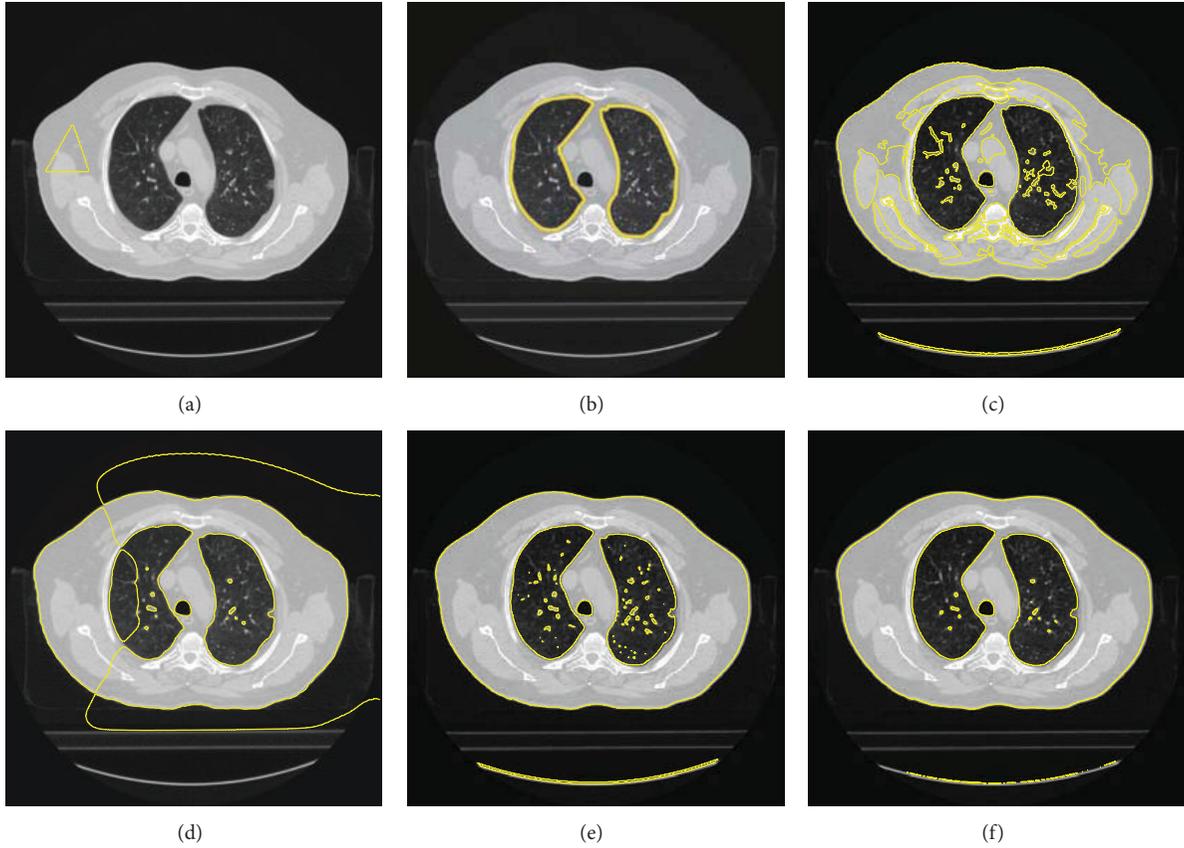


FIGURE 12: The comparison between the proposed method and other models. (a) The initial contour and the original image. (b) The segmentation result by expert. (c) The segmentation of LBF model. (d) The segmentation of LIF model. (e) The segmentation of GCLGIF model. (f) The segmentation of the proposed model.

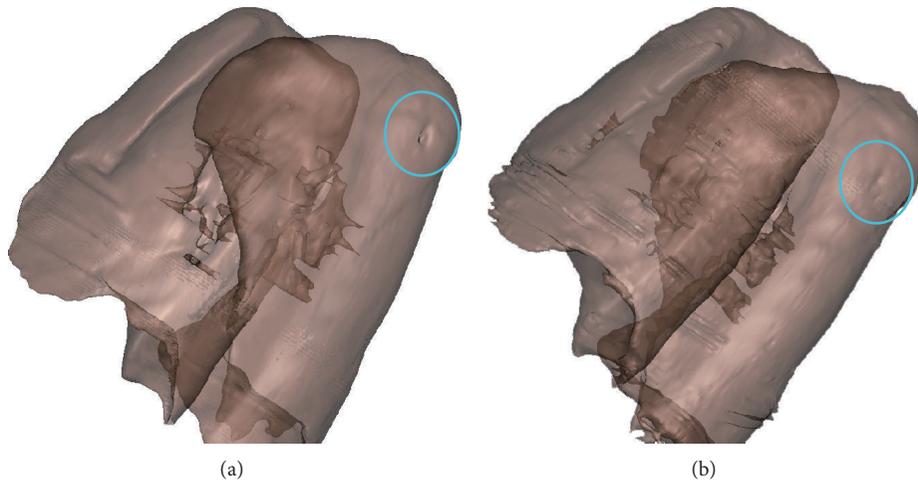


FIGURE 13: The 3D model of the corrected lung. (a) The 3D lung model before correction; (b) the 3D model after correction.

function by the Gaussian convolution, which makes the zeros level set not match the real edge absolutely which can be seen in Figure 8(c). However, if we observe the tiny sag region between two ribs in Figure 13(b), the size of the unfilled hole is very small too; that is, the loss of the imperfect filling process is very slight which can be verified by Figure 8(c) too.

Finally, in order to examine the correction accuracy of the proposed model, the correction results of GLg model are compared with the correction results by hand which are obtained under the guidance of the expert. We select 28 CT slice sets from the LIDC Database; the correction results by hand are selected as the golden standard; the mean

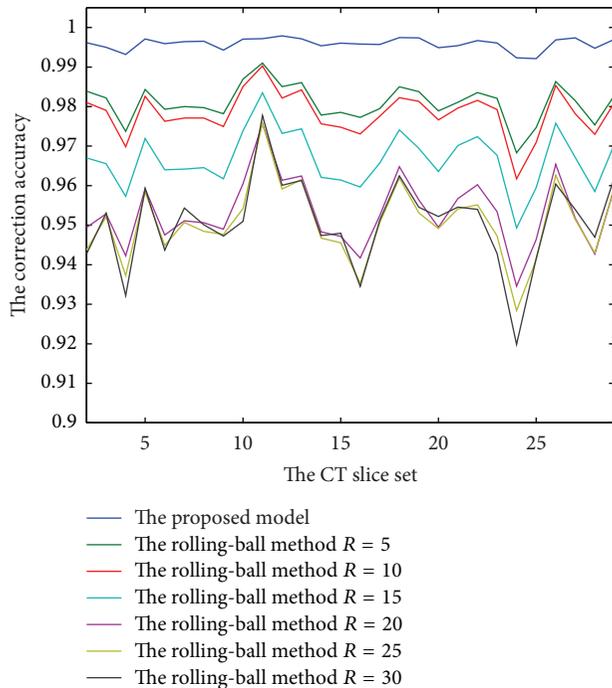


FIGURE 14: The mean correction accuracy of our method and the morphology method under 28 different CT slice sets.

correction accuracy of every set is shown in Figure 14. From those accuracy data, it is easy to find that the minimum of the proposed correction accuracy is above 99%; that is, they outperform their morphology-based counterparts, which implies the proposed model is a feasible model.

#### 4. Conclusion

A new automated lung extraction and the edge correction method combines CV model, LBF model, the globally convex segmentation method, the split Bregman method, the fractal theory, and the geometric active contour model. The detection accuracy of the proposed new edge detection function is more accurate than the traditional edge detection function. With the help of the split Bregman method, the useful information of the new adaptive weight function, and the new edge information, the proposed segmentation model that combines the CV model and LBF model can segment the given image quickly and get the global minimum no matter where the initial contour lies. Furthermore, with the help of the fractal theory and the statistical threshold, the proposed method can detect the defective edge successfully. In addition, the geometric active contour model can fill the hole that surrounds juxtapleural nodules automatically. Several experiments demonstrate the accuracy and efficiency of our correction model.

#### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

#### Acknowledgment

This work is supported by the National Science & Technology Pillar Program of China (Grant no. 2012BAI14B00).

#### References

- [1] K. T. Bae, J.-S. Kim, Y.-H. Na, K. G. Kim, and J.-H. Kim, "Pulmonary nodules: automated detection on ct images with morphologic matching algorithm—preliminary results," *Radiology*, vol. 236, no. 1, pp. 286–293, 2005.
- [2] Y. Yim, H. Hong, J. Beom Seo, N. Kim, E. Jin Chae, and Y. Gil Shin, "Correction of lung boundary using the gradient and intensity distribution," *Computers in Biology and Medicine*, vol. 39, no. 3, pp. 239–250, 2009.
- [3] P. T. Johnson, H. K. Hahn, D. G. Heath, and E. K. Fishman, "Automated multidetector row CT dataset segmentation with an interactive watershed transform (IWT) algorithm: part 2—body CT angiographic and orthopedic applications," *Journal of Digital Imaging*, vol. 21, no. 4, pp. 413–421, 2008.
- [4] Q. Li, S. Katsuragawa, and K. Doi, "Computer-aided diagnostic scheme for lung nodule detection in digital chest radiographs by use of a multiple-template matching technique," *Medical Physics*, vol. 28, no. 10, pp. 2070–2076, 2001.
- [5] S. Hu, E. A. Hoffman, and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images," *IEEE Transactions on Medical Imaging*, vol. 20, no. 6, pp. 490–498, 2001.
- [6] N. A. Avila, J. A. Kelly, A. J. Dwyer, D. L. Johnson, E. C. Jones, and J. Moss, "Lymphangioliomyomatosis: correlation of qualitative and quantitative thin-section CT with pulmonary function tests and assessment of dependence on pleurodesis," *Radiology*, vol. 223, no. 1, pp. 189–197, 2002.
- [7] S. G. Armato III and W. F. Sensakovic, "Automated lung segmentation for thoracic CT: impact on computer-aided diagnosis," *Academic Radiology*, vol. 11, no. 9, pp. 1011–1021, 2004.
- [8] M. Sonka, V. Hlavac, R. Boyle et al., *Image Processing, Analysis, and Machine Vision*, PWS Publishing Company, Pacific Grove, Calif, USA, 1999.
- [9] Y. Itai, H. Kim, S. Ishikawa et al., "Automatic segmentation of lung areas based on snakes and extraction of abnormal areas," in *Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI '05)*, pp. 377–381, IEEE, November 2005.
- [10] M. Silveira and J. Marques, "Automatic segmentation of the lungs using multiple active contours and outlier model," in *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '06)*, pp. 3122–3125, New York, NY, USA, September 2006.
- [11] M. Lee, S. Park, W. Cho, and S. Kim, "Medical image segmentation using a geometric active contour model based on level set method," in *Proceedings of the IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM '07)*, pp. 577–580, August 2007.
- [12] S. Liu and J. Li, "Automatic medical image segmentation using gradient and intensity combined level set method," in *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '06)*, pp. 3118–3121, IEEE, September 2006.
- [13] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359–369, 1998.

- [14] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
- [15] C. Li, C.-Y. Kao, J. C. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1940–1949, 2008.
- [16] L. Wang, L. He, A. Mishra, and C. Li, "Active contours driven by local Gaussian distribution fitting energy," *Signal Processing*, vol. 89, no. 12, pp. 2435–2447, 2009.
- [17] K. Zhang, H. Song, and L. Zhang, "Active contours driven by local image fitting energy," *Pattern Recognition*, vol. 43, no. 4, pp. 1199–1206, 2010.
- [18] L. Wang, C. Li, Q. Sun, D. Xia, and C.-Y. Kao, "Active contours driven by local and global intensity fitting energy with application to brain MR image segmentation," *Computerized Medical Imaging and Graphics*, vol. 33, no. 7, pp. 520–531, 2009.
- [19] X.-F. Wang, D.-S. Huang, and H. Xu, "An efficient local Chan-Vese model for image segmentation," *Pattern Recognition*, vol. 43, no. 3, pp. 603–618, 2010.
- [20] Y. Yang and B. Wu, "Split Bregman method for minimization of improved active contour model combining local and global information dynamically," *Journal of Mathematical Analysis and Applications*, vol. 389, no. 1, pp. 351–366, 2012.
- [21] S. G. Armato III, F. Li, M. L. Giger, H. MacMahon, S. Sone, and K. Doi, "Lung cancer: performance of automated lung nodule detection applied to cancers missed in a CT screening program," *Radiology*, vol. 225, no. 3, pp. 685–692, 2002.
- [22] S. G. Armato III, M. L. Giger, C. J. Moran, J. T. Blackburn, K. Doi, and H. MacMahon, "Computerized detection of pulmonary nodules on CT scans," *Radiographics*, vol. 19, no. 5, pp. 1303–1311, 1999.
- [23] M. Dajnowiec and J. Alirezaie, "Computer simulation for segmentation of lung nodules in CT images," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '04)*, vol. 5, pp. 4491–4496, IEEE, October 2004.
- [24] X. Zhang, G. McLennan, E. A. Hoffman, and M. Sonka, "Automated detection of small-size pulmonary nodules based on helical CT images," in *Information Processing in Medical Imaging*, vol. 3565 of *Lecture Notes in Computer Science*, pp. 664–676, Springer, 2005.
- [25] L. Pastor, A. Pousse, P. Manzoni, and B. Kastler, "A pulmonary nodule modeling tool as a diagnostic aid for lung HRCT images," *Computerized Medical Imaging and Graphics*, vol. 29, no. 8, pp. 631–637, 2005.
- [26] J. P. Ko and M. Betke, "Chest CT: automated nodule detection and assessment of change over time—preliminary experience," *Radiology*, vol. 218, no. 1, pp. 267–273, 2001.
- [27] Y. Yim and H. Hong, "Correction of segmented lung boundary for inclusion of pleural nodules and pulmonary vessels in chest CT images," *Computers in Biology and Medicine*, vol. 38, no. 8, pp. 845–857, 2008.
- [28] J. Pu, J. Roos, C. A. Yi, S. Napel, G. D. Rubin, and D. S. Paik, "Adaptive border marching algorithm: automatic lung segmentation on chest CT images," *Computerized Medical Imaging and Graphics*, vol. 32, no. 6, pp. 452–462, 2008.
- [29] J. Weickert, *Anisotropic Diffusion in Image Processing*, vol. 1, Teubner, Stuttgart, Germany, 1998.
- [30] T. F. Chan, S. Esedoglu, and M. Nikolova, "Algorithms for finding global minimizers of image segmentation and denoising models," *SIAM Journal on Applied Mathematics*, vol. 66, no. 5, pp. 1632–1648, 2006.
- [31] T. Goldstein, X. Bresson, and S. Osher, "Geometric applications of the split Bregman method: segmentation and surface reconstruction," *Journal of Scientific Computing*, vol. 45, no. 1–3, pp. 272–293, 2010.
- [32] K. Foroutan-pour, P. Dutilleul, and D. L. Smith, "Advances in the implementation of the box-counting method of fractal dimension estimation," *Applied Mathematics and Computation*, vol. 105, no. 2-3, pp. 195–210, 1999.

## Research Article

# High Order Statistics and Time-Frequency Domain to Classify Heart Sounds for Subjects under Cardiac Stress Test

Ali Moukadem,<sup>1</sup> Samuel Schmidt,<sup>2</sup> and Alain Dieterlen<sup>1</sup>

<sup>1</sup>MIPS Laboratory, University of Haute Alsace, 68093 Mulhouse, France

<sup>2</sup>Department of Health Science and Technology, Aalborg University, 9220 Aalborg, Denmark

Correspondence should be addressed to Ali Moukadem; [ali.moukadem@uha.fr](mailto:ali.moukadem@uha.fr)

Received 27 October 2014; Revised 2 December 2014; Accepted 1 January 2015

Academic Editor: Kayvan Najarian

Copyright © 2015 Ali Moukadem et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper considers the problem of classification of the first and the second heart sounds (S1 and S2) under cardiac stress test. The main objective is to classify these sounds without electrocardiogram (ECG) reference and without taking into consideration the systolic and the diastolic time intervals criterion which can become problematic and useless in several real life settings as severe tachycardia and tachyarrhythmia or in the case of subjects being under cardiac stress activity. First, the heart sounds are segmented by using a modified time-frequency based envelope. Then, to distinguish between the first and the second heart sounds, new features, named  $\alpha_{opt}$ ,  $\beta$ , and  $\gamma$ , based on high order statistics and energy concentration measures of the Stockwell transform (S-transform) are proposed in this study. A study of the variation of the high frequency content of S1 and S2 over the HR (heart rate) is also discussed. The proposed features are validated on a database that contains 2636 S1 and S2 sounds corresponding to 62 heart signals and 8 subjects under cardiac stress test collected from healthy subjects. Results and comparisons with existing methods in the literature show a large superiority for our proposed features.

## 1. Introduction

Cardiac auscultation is the basis for heart examination. It provides a wealth of information about structural and functional cardiac defects, using a simple, efficient, and costless medical device: the stethoscope. Invented in the nineteenth century, this acoustic instrument has proved since then to be of paramount importance to the physical examination and diagnosis of cardiac pathologies. Over the course of the past two centuries, the stethoscope underwent numerous improvements to reach the development of the electronic stethoscope capable of registering and optimizing the quality of the acoustic signal, completed by the Phonocardiographic (PCG) representation of the auscultation signal. However, the analysis of the cardiac sounds, solely based on the human ear, is limited by the experience of the clinician for a reliable diagnosis of cardiac pathologies and to obtain all the qualitative and quantitative information about cardiac activity [1]. Information, such as the temporal localization of the heart sounds, the number of their internal components,

their frequential content, and the significance of diastolic and systolic murmurs, can also be studied directly on the PCG signal [2]. In order to recognize and classify cardiovascular pathologies, advanced methods and techniques of signal processing will be used.

For that, two approaches could be considered to improve electronic stethoscopes:

- (i) stethoscope with embedded autonomous analysis, simple for home use by patients and paramedics, for the purpose of autodiagnosis and follow-up,
- (ii) stethoscope coupled with a hosting device or a server for sophisticated analysis (coupled to a PC with a Bluetooth link) for the use of professionals in order to improve performance of clinical medical diagnosis.

Whatever the approach, one of the first phases in the analysis of heart sounds, is the segmentation [3–5]. Heart sound segmentation divides the PCG signal into four parts: S1 (first heart sound), extant systole, S2 (second heart sound), and

extant diastole. First, S1 and S2 are located; then, extant systole is represented by the interval S1 to S2 and extant diastole by the interval S2 to S1.

Identification of the two phases of the cardiac cycle and of the heart sounds with robust differentiation between S1 and S2 even in the presence of additional heart sounds and/or murmurs is a first step in this challenge. Then, there is a need to measure accurately S1 and S2 [6, 7] allowing the progression to automatic diagnosis of heart murmurs with the distinction of ejection and regurgitation murmurs.

Most of the existing methods, for the direct segmentation of heart sounds, without the use of the help of ECG (see Figure 1), use the feature of systole and diastole duration to classify the first heart sound (S1) and the second heart sound (S2) [3–10]. These time intervals can become problematic and useless in several real life settings which are particularly represented by severe tachycardia, tachyarrhythmia, or subjects under cardiac stress activity.

In our earlier work on the segmentation of heart sounds [2], we applied the singular value decomposition (SVD) technique and the KNN classifier to distinguish between S1 and S2. The SVD extracts vector of 20 features is issued from the Stockwell transform [11], and then the feature vector was followed by a trained KNN classifier. This proposed method was validated on a general database (without stress data) collected from Hospital University of Strasbourg (HUS) and Mars500 project.

In this study, we use a new database of stressed subject collected in Aalborg University. This database is particularly of interest to classify S1 and S2 because it contains specific conditions where the systolic and diastolic intervals features are useless. Furthermore, our original approach adopted in this paper consists of studying qualitative features (instead of extracting blindly feature vector as done in [2]) in order to select the most appropriate single feature. This new approach makes the training phase unnecessary since it needs only a simple threshold and makes the segmentation phase less time consuming and reduces its complexity. Three original features based on time-frequency domain and high order statistics are proposed in this paper and their performances is discussed and analyzed.

The main contributions of this paper can be summarized as follows.

- (i) A first modification on the segmentation method proposed in [2] is performed in order to enhance the detection of low intensities sounds buried in noise.
- (ii) The main contribution of this study is the investigation of 3 new qualitative features to discriminate between S1 and S2 ( $\alpha_{\text{opt}}$ ,  $\beta$ , and  $\gamma$ ):
  - (a) the  $\alpha_{\text{opt}}$  feature was used in [2] to optimize the energy concentration of the Stockwell transform. However, in this study  $\alpha_{\text{opt}}$  is proposed as a feature to discriminate between S1 and S2 which is totally different and can be considered as a new approach;
  - (b) the second feature, namely,  $\beta$ , is the integration over time of the envelope obtained by a modified

measure of the instantaneous frequency of the signal. This feature aims to describe accurately the frequency content of S1 and S2 over time;

- (c) the third new feature, namely,  $\gamma$ , calculates the kurtosis of the time-frequency envelope. This feature is based on the spectrogram of the Stockwell transform (ST-spectrogram) and an analogy between the time-frequency coefficients and the probability density function is made in order to apply the kurtosis measure.
- (iii) Experimental validation based on specific database from Aalborg University of stress test subjects is performed. We note here that many studies in the literature suggest that an involvement of the some features extracted from the heart sounds (e.g., S/D-ratio) can increase the diagnostic value of the exercise test [12–14]. However, to our knowledge, our study is the first one with interests in distinguishing the heart sounds (S1 and S2) for subject under stress conditions.
- (iv) An experimental study to show the high frequency content ration (S2/S1) variation over heart rate is performed and discussed.

This paper is organized as follows. Section 2 describes the data collection process and methods proposed in this paper. The results and discussion are presented in Sections 3 and 4 giving the conclusion and the future work.

## 2. Methods

*2.1. Data Collection.* The collected database used in this study corresponds to healthy subjects under cardiac stress test from the Department of Health Science and Technology, Aalborg University [15].

Nine healthy subjects were enrolled in the study ( $M = 5$ ,  $F = 4$ ) with a median age of 32 (24–36). Informed consent was retrieved from all subjects prior to the exercise test. A Panasonic microphone was incorporated in a coupler, specially designed by the Department of Acoustics at Aalborg University, Denmark. The microphone detects the mechanical pressure differences in the coupler, caused by alterations of the sound pressure. The microphone records with a sampling frequency of 48000 Hz. The heart sound recordings are synchronized with a 3-lead ECG (see Figure 1). The microphone was fitted to the 3rd left intercostal space with a specially designed double adhesive plaster. Subsequently the subject cycled on a Monark Ergometric 894E ergometer bicycle. The workload was increased by 25 watt every two minutes with an initial workload of 25 watt. The subject cycled until subjective maximum endurance was reached. Afterward subjects that did not reach 80% of maximum heart rate defined as  $(220 \text{ bpm} - \text{age}) \pm 12$  were excluded from the study. The study was conducted according to the Danish ECG stress test guidelines. The “220 bpm – age” is a common criterion to ensure that the patient reaches their full capacity [16]. One subject did not reach this rate and was therefore excluded from the study. Recordings of heart sounds were made for 10 seconds at the end of each workload level. Acarix

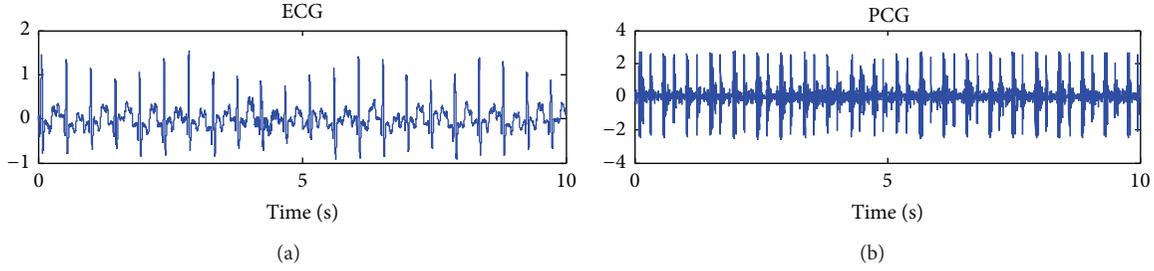


FIGURE 1: Synchronized ECG and PCG signals for a subject under cardiac stress test.

Data Acquisition System was used for recording the heart sounds and ECG [17]. The subject's heart rates before starting the experiment correspond to the first workload level. There were not any special restrictions required from the subjects before starting the experiment.

**2.2. Gold Standard.** The gold standard is generated based on the synchronized ECG signal. The ECG can provide information to classify the first and the second heart sounds, since S1 occurs subsequent to the QRS complex and S2 occurs after the T-wave [5].

The sounds are automatically classified based on the proposed features which are validated based on the corresponding synchronized ECG signals.

**2.3. The High Frequency Signature (HFS) Feature to Classify S1 and S2.** The only study in the literature that aims to classify S1 and S2 by taking into consideration another feature compared to the systolic and diastolic criteria is the study proposed by Kumar et al. in [18, 19]. The methods aim to extract the high frequency envelopes in sound segments, by applying the Shannon energy operator on the detail coefficients issued from the wavelet transform (Daubechies 6) [18]. In order to detect the heart cycles, an adaptive threshold is defined for this envelope. The algorithm aims to detect the high frequency signatures (HFS) and the low frequency signatures (LFS) [18].

Kumar et al. consider that usually S2 sounds contain higher frequency with respect to S1 sound (HFS correspond to S2 and LFS correspond to S1) excluding some rare exceptions.

The problem with the Kumar et al.'s paper can be summarized as follows.

- (i) Authors consider that S1 can contain higher frequency content compared to S2 only in rare cases like prosthetic valves, for example. This hypothesis ignores the complexity of real clinical sounds on which S1 in normal sounds can have higher signature (see Figure 2, for example) and ignore that the frequency content of S1 and S2 is related to the heart rate (see Figure 15) and the auscultation position [20, 21].
- (ii) To make the proposed method automatic and free from prior knowledge, the type of HFS signature is not identified as S2 automatically (because exceptions can occur) but it is identified by using the systolic time

interval criteria which is exactly what we aim to avoid in this study since the systolic time interval estimation is not a reliable feature in stress test data (when HR is high) or in pathological cases as severe tachycardia or tachyarrhythmia.

- (iii) Authors consider that all detected HFS exhibit one class of sound (S1 or S2) which is not a reliable hypothesis since the frequency contents of S1 and S2 can vary in the same registration (see Figure 3) due to changes in respiratory conditions [20, 21].

**2.4. Stockwell Transform.** The S-transform originates from two advanced signal processing tools, the short time Fourier transform (STFT) and the wavelet transform (WT). It can be viewed as a frequency dependent STFT or a phase corrected WT. The S-transform is becoming a valuable tool applied on many signals and domains as cardiovascular [2], EEG signals [22], geophysics [23], power system engineering [24], and so forth. The S-transform of a time varying signal  $x(t)$  is defined by [11]

$$S_x(t, f) = \int_{-\infty}^{+\infty} x(\tau) w(\tau - t, f) e^{-2\pi j f \tau} d\tau, \quad (1)$$

where the window function  $w(\tau - t, f)$  is chosen as

$$w(t, f) = \frac{1}{\sigma(f) \sqrt{2\pi}} e^{-t/2\sigma^2(f)}. \quad (2)$$

And  $\sigma(f)$  is a function of frequency as

$$\sigma(f) = \frac{1}{|f|}. \quad (3)$$

The window is normalized as

$$\int_{-\infty}^{+\infty} w(t, f) dt = 1. \quad (4)$$

This gives the direct relation between the S-transform and the Fourier spectrum by averaging the local spectrum over time:

$$\int_{-\infty}^{+\infty} S_x(t, f) dt = X(f), \quad (5)$$

where  $X(f)$  is the Fourier transform of  $x(t)$ .

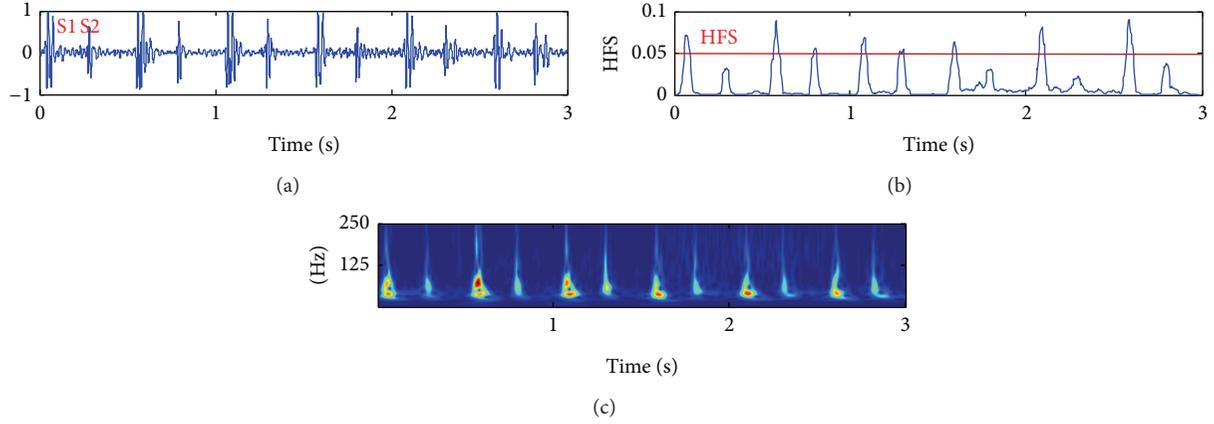


FIGURE 2: Normal heart sound with the HFS envelope and the magnitude of the S-Transform showing the higher frequency content in S1 compared to S2.

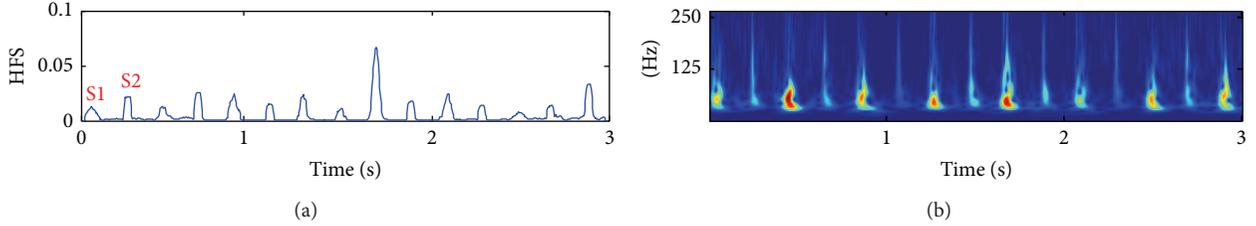


FIGURE 3: Segmented sound with the HFS envelope and the magnitude of the S-transform of the corresponding heart sound showing that HFS signature does not correspond necessarily to just one class.

**2.5. Segmentation: The Modified SSE Method.** The localization of heart sounds is established by using the SSE method (see (6)). The proposed SSE method extracts the envelope of the signal by calculating the Shannon energy of each column of the extracted S-matrix (local spectrum). Then, the extracted envelope is smoothed by applying an average filter. The SSE envelope applied on the time-frequency matrix  $S(\tau, f)$  is calculated as:

$$\text{SSE}(S(\tau, f)) = - \int_{-\infty}^{+\infty} |S(\tau, f)|^n \log(|S(\tau, f)|^n) df. \quad (6)$$

The parameter  $n$  is usually fixed to 2 [2] which is the standard coefficient of the Shannon energy measure. In this study,  $n$  is fixed to 1.5 to enhance the detection of low intensities sounds buried in noise. This occurs in heart sounds more often with S2 when the cardiac frequency is high. Figure 4 shows the compromise of attenuation of low and high intensities, as a function of the value of  $n$ . We note here that, for the SSE method, the intensities are the local spectrum coefficients of the S-transform and not the time sample intensities of the signal. Figure 5 shows the influence of the values of  $n$  in the detection of very low intensities heart sounds (S2 in this case).

## 2.6. New Proposed Features

**Feature 1** (the Gaussian parameter ( $\alpha_{\text{opt}}$ )). In another study, we have introduced a parameter  $\alpha$  to the Gaussian window in order to optimize the energy concentration of the Stockwell

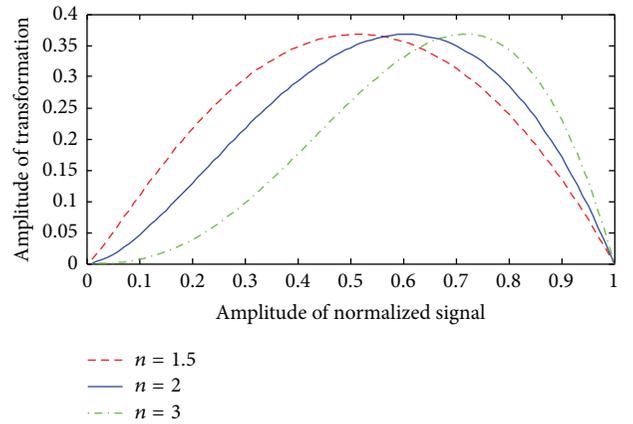


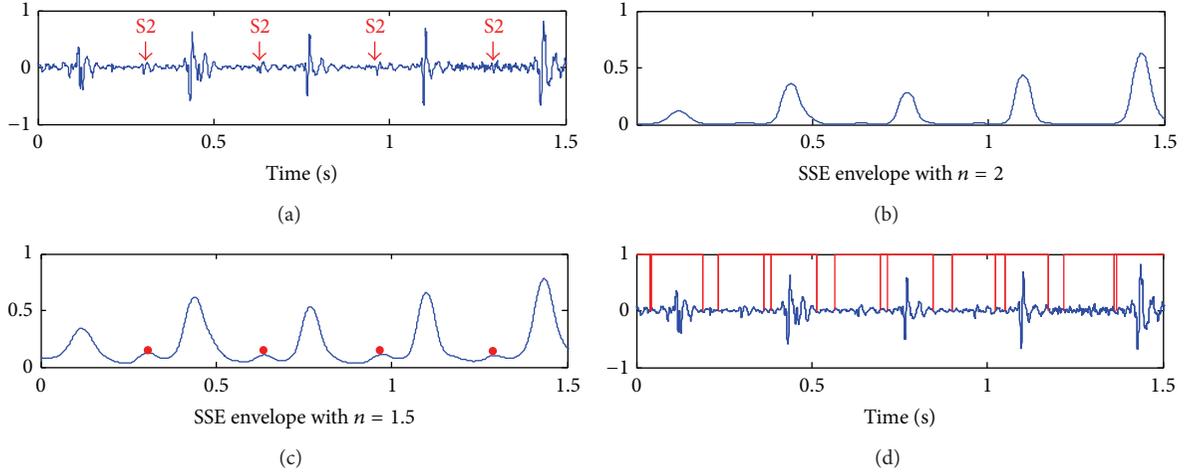
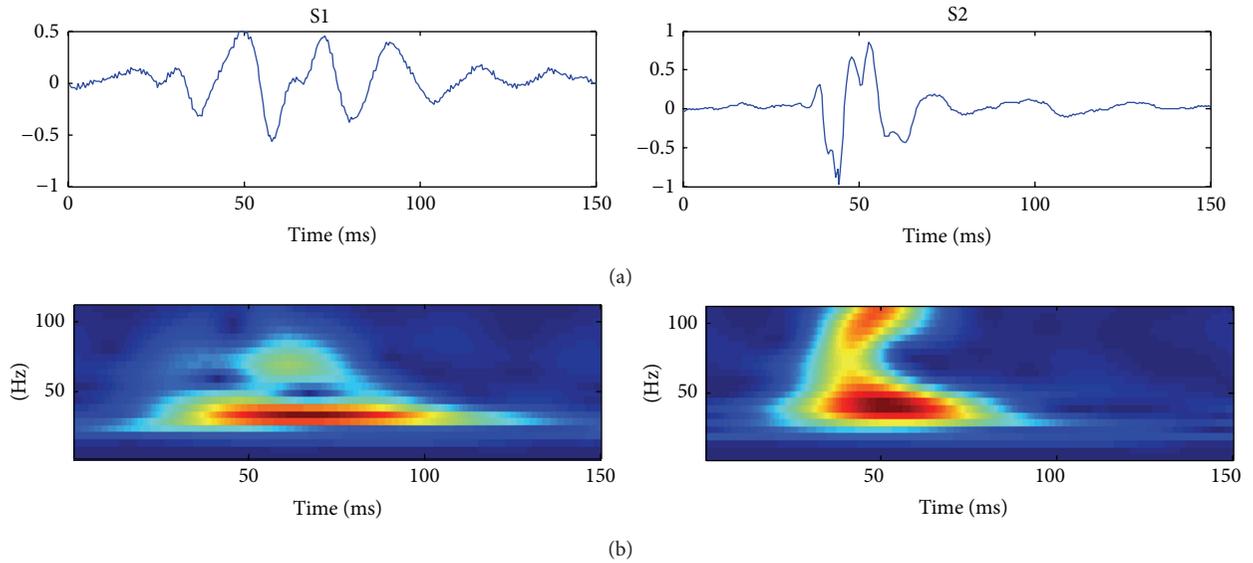
FIGURE 4: The envelope of normalized signal for values of  $n = 1.5$ , 2, and 3.

transform [2]. The parameter  $\alpha$  introduced to the Gaussian equation (2) is introduced as follows:

$$\sigma(f) = \frac{\alpha}{|f|}. \quad (7)$$

The value of  $\alpha$  which maximizes the energy concentration is considered as the optimal value. The energy concentration measure is given as

$$\text{CM}(\alpha) = \frac{1}{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |S_x^\alpha(t, f)| dt df}. \quad (8)$$


 FIGURE 5: The influence of the values of  $n$  in the SSE envelope for the detection of S2 sounds with very low intensities.

 FIGURE 6: S1 and S2 signals (a) and optimized S-transform obtained with  $\alpha = 0.8$  for S1 and  $\alpha = 0.5$  for S2 (b).

This measure has some favorable performance in comparison to other concentration measures [25].

With  $S_x^\alpha(t, f)$  being the normalized energy of the S-transform for each  $\alpha$ , it is given by [26]

$$\overline{S_x^\alpha(t, f)} = \frac{S_x^\alpha(t, f)}{\sqrt{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |S_x^\alpha(t, f)|^2 dt df}}. \quad (9)$$

The values of  $\alpha$  are chosen from a set,  $0.5 < \alpha < 2$ , with a step of 0.1. The optimal solution is reached when  $CM(\alpha)$  is maximized:

$$\alpha_{\text{opt}} = \arg \max_{\alpha} (CM(\alpha)). \quad (10)$$

In this study, we propose to test the ability of the  $\alpha_{\text{opt}}$  to discriminate between the first and the second heart sound. Since it has been used to optimize the energy concentration in the time-frequency plane, it may be interesting to test

it as a discriminator feature. From a signal theory point of view, the complexity concept of signals is intuitively related to the number of their elementary components [27] and since S1 generally contains more components than S2 [21], hence, it can be considered as a more complex signal than S2. These physiological differences will necessarily lead to different time-frequency content behavior which we will aim to reveal with  $\alpha_{\text{opt}}$  parameter.

Figure 6 shows S1 and S2 signals examples with the corresponding optimized S-transform obtained with  $\alpha = 0.8$  and 0.5, respectively.

*Feature 2* (the SSE envelope feature ( $\beta$ )). It is another new feature that we investigate in this study, namely,  $\beta$ ; it aims to integrate the normalized SSE envelope over time; it can be given as

$$\beta = \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} |S_x(t, f)|^2 \log_2(|S_x(t, f)|^2) df \right\} dt. \quad (11)$$

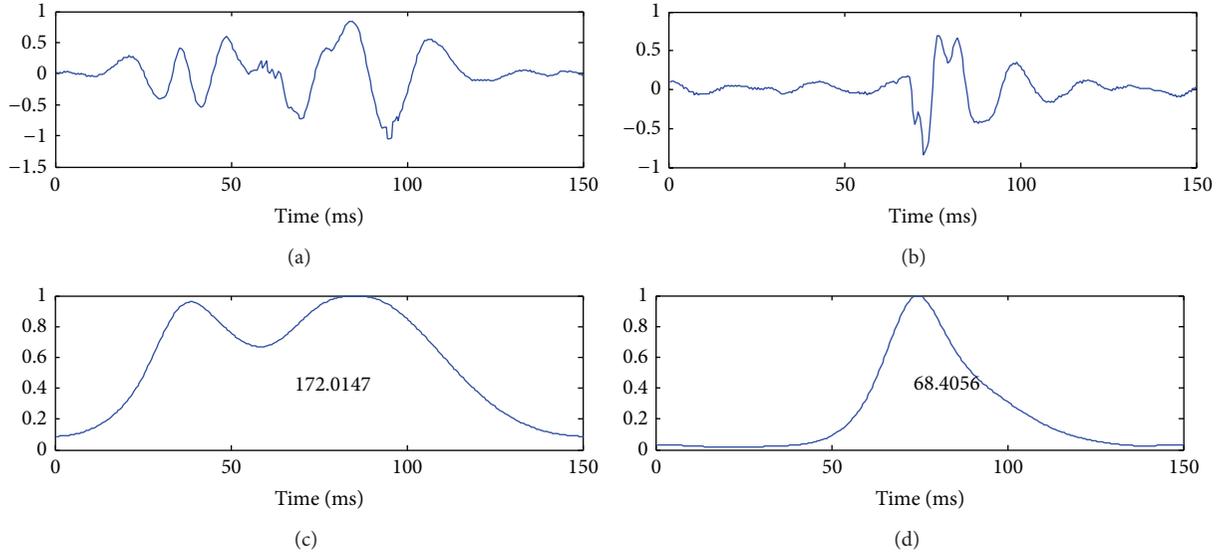


FIGURE 7: S1 (left) and S2 (right) signals and their normalized SSE envelopes with the values of  $\beta$  (bottom).

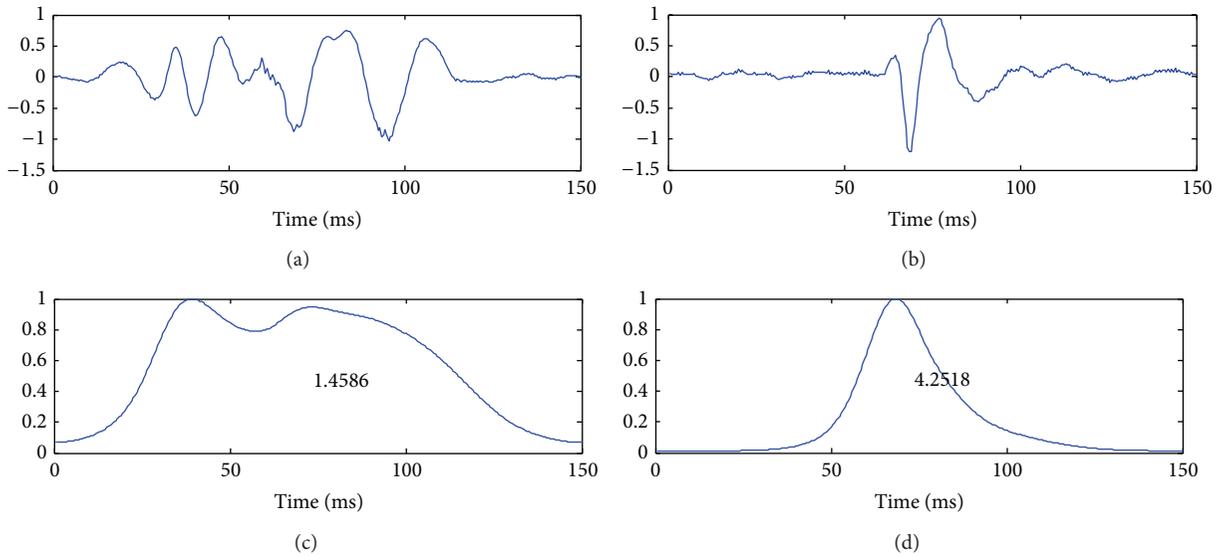


FIGURE 8: S1 (left) and S2 (right) signals and their normalized SSE envelopes with the values of  $\gamma$  (bottom).

The SSE envelope estimates the frequency energy at the local spectrum of the signal. It can be considered as a modified instantaneous frequency measure. The  $\beta$  feature aims to reveal the frequency contribution of each sound over time. Mathematically, it can be viewed as the integration over time of a modified instantaneous frequency measure. The measure is computed from the normalized SSE envelope to avoid the influence of the amplitude variations. Figure 7 shows an example of the  $\beta$  feature calculated on S1 and S2 sounds from their normalized SSE envelopes.

*Feature 3* (high order statistic feature ( $\gamma$ )). It is the third feature proposed in this paper on higher order statistic measure (kurtosis) applied on time-frequency coefficients. The

kurtosis measure is normally applied on a probability distribution to describe its shape. A normal transition between the Stockwell transform and the corresponding time-frequency energy distribution is the square of magnitude of the S-matrix, namely, in this paper, the ST-spectrogram. In this case, the time-frequency representation plays an analogous role to a 2D probability density function (PDF) [27]. Then, the kurtosis can be applied directly on this estimated PDF via time-frequency plane. However, the kurtosis may be very sensitive to noise [28]. To deal with this problem, we calculate the SSE envelope applied on the ST-spectrogram before applying the kurtosis. In this case, the values of the extracted envelope are considered as the estimated probability distribution via time-frequency plane.

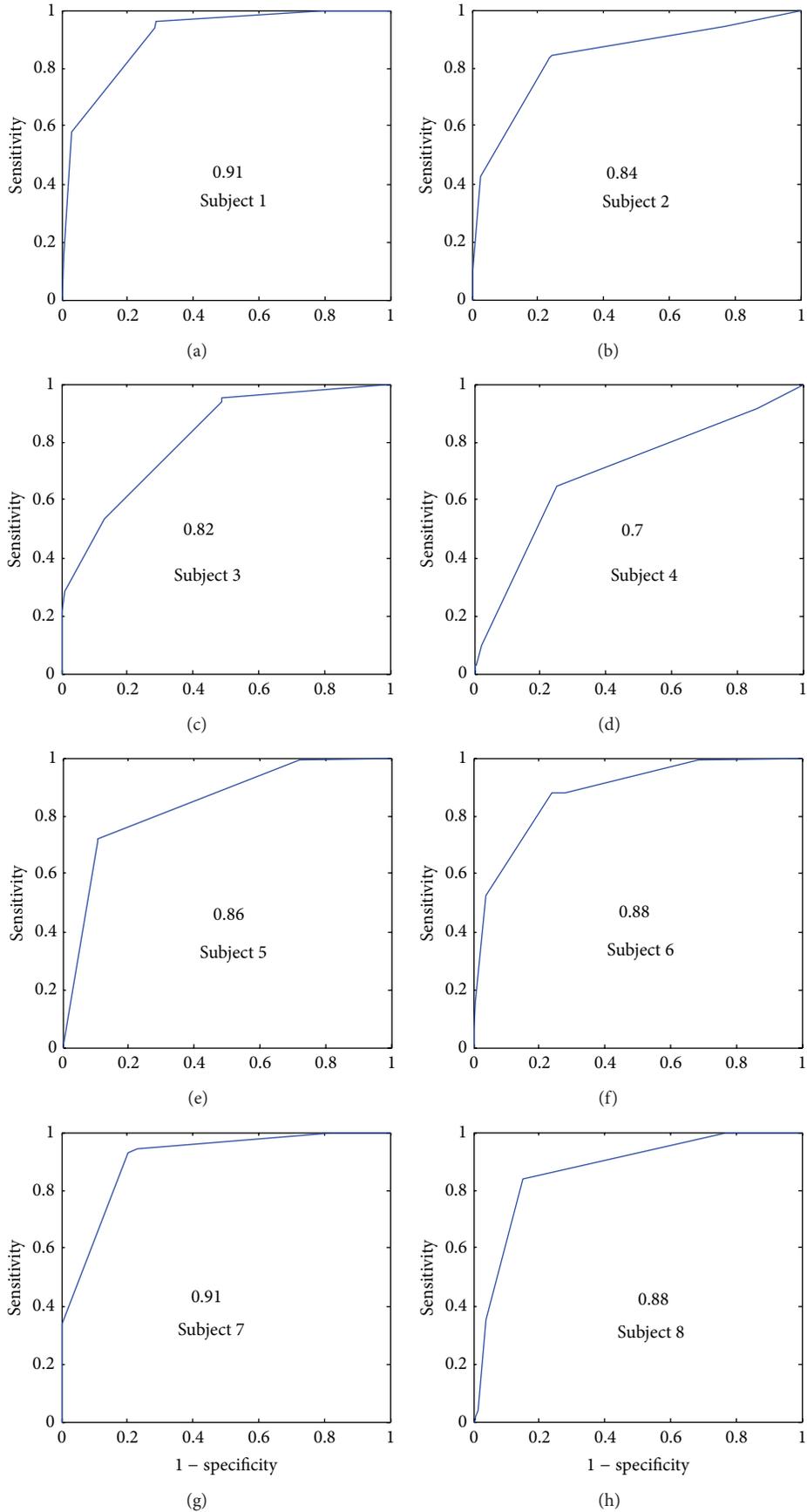


FIGURE 9: Receiver operation characteristic curves for feature  $\alpha$  and for all subjects.

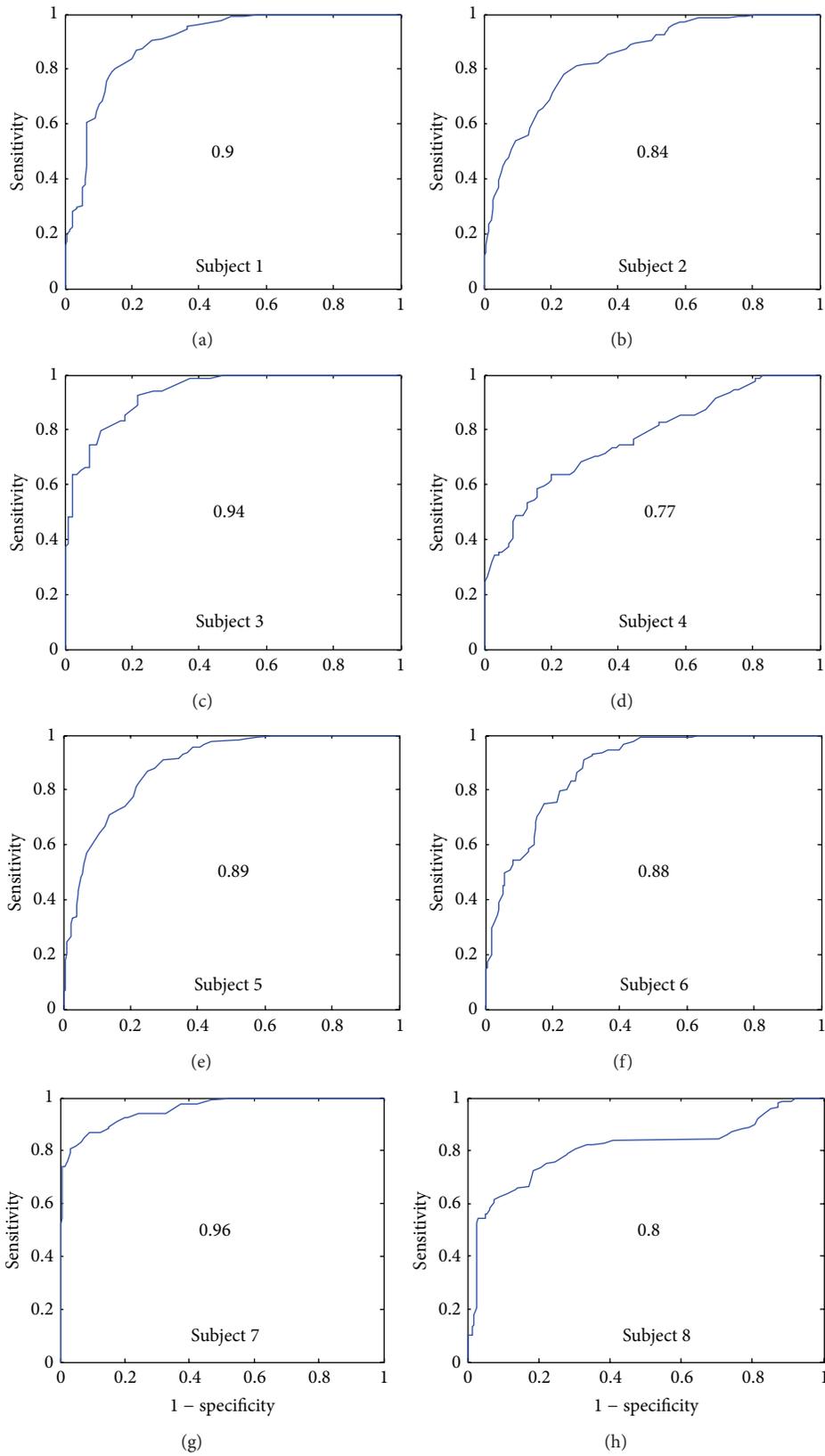


FIGURE 10: Receiver operation characteristic curves for feature  $\beta$  and for all subjects.

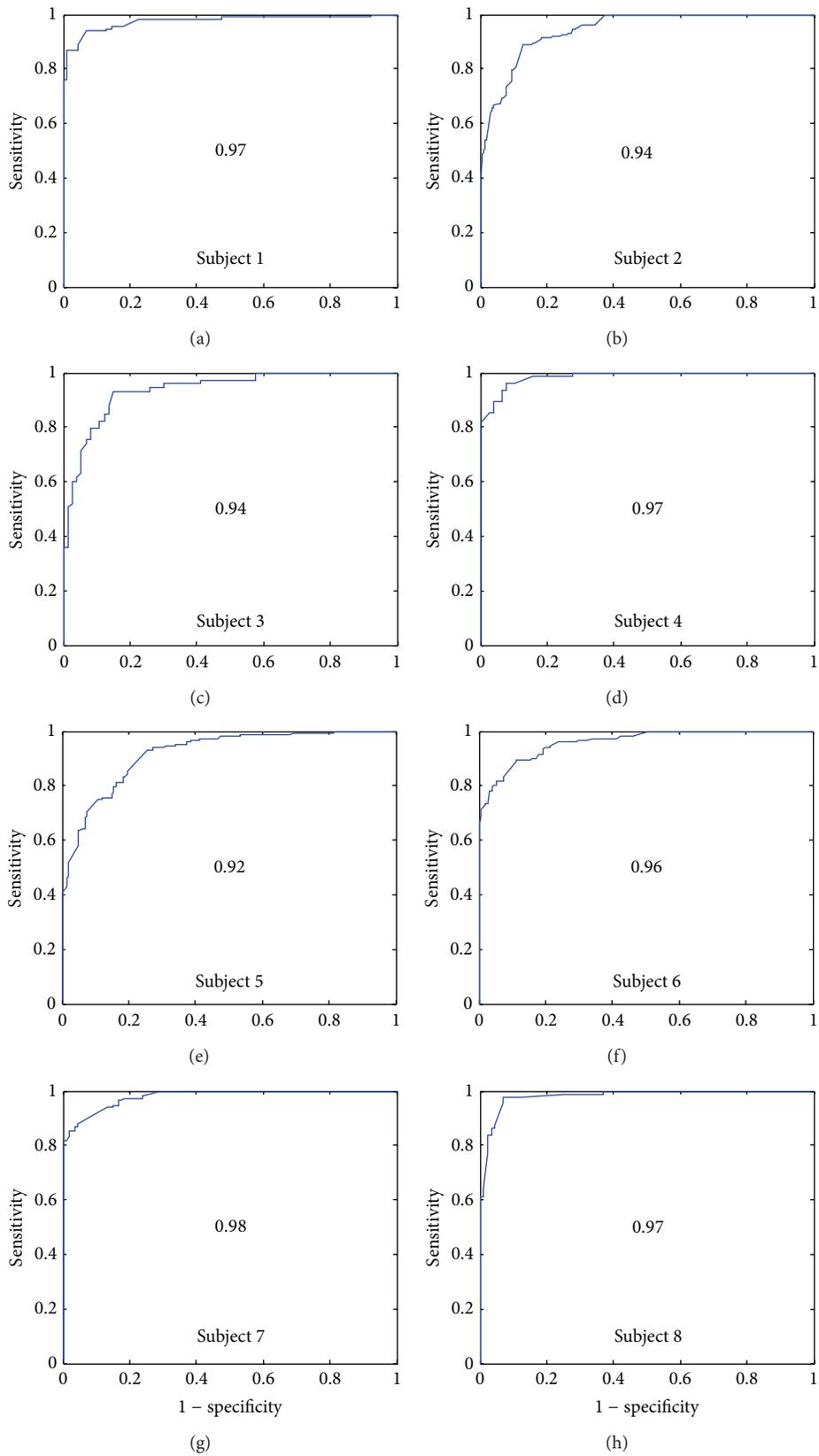


FIGURE 11: Receiver operation characteristic curves for feature  $\gamma$  and for all subjects.

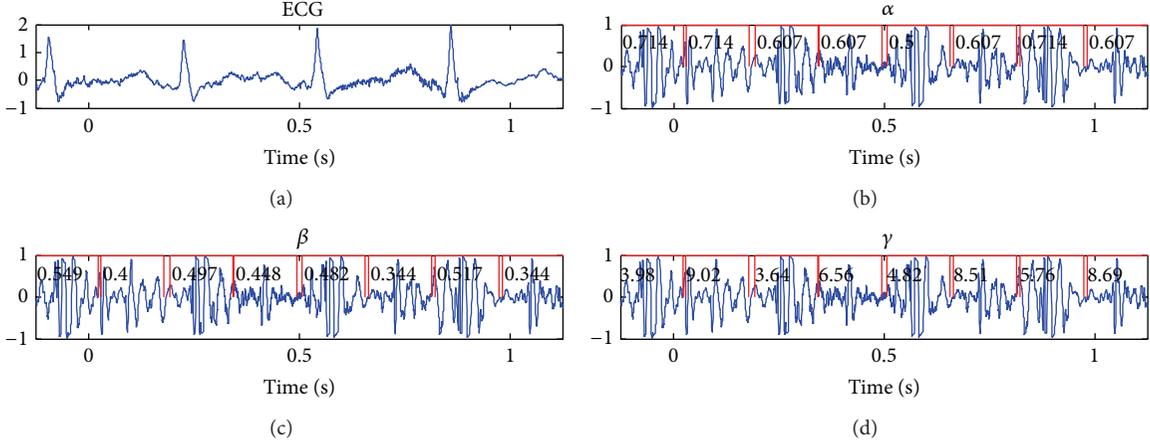


FIGURE 12: Example of a segmented stress test heart sound for subject 4 and workload level 6 (HR = 181 bpm) with the values of  $\alpha$ ,  $\beta$ , and  $\gamma$  calculated for each located sound (S1 and S2).

TABLE 1: The area under curve (AUC), mean values and standard deviations of each feature (S1( $\alpha$ ), S2( $\alpha$ ), S1( $\beta$ ), S2( $\beta$ ), S1( $\gamma$ ), and S2( $\gamma$ )), and the maximum heart rate (HR<sub>max</sub>) reached for each subject.

Subject	1	2	3	4	5	6	7	8	Mean
AUC( $\alpha$ )	0.91	0.84	0.82	0.7	0.86	0.88	0.91	0.88	<b>0.85</b>
AUC( $\beta$ )	0.9	0.84	0.94	0.77	0.89	0.88	0.96	0.8	<b>0.87</b>
AUC( $\gamma$ )	0.97	0.94	0.94	0.97	0.92	0.96	0.98	0.97	<b>0.96</b>
S1( $\alpha$ )	0.79 ± 0.05	1.02 ± 0.76	1.03 ± 0.35	0.71 ± 0.07	0.7 ± 0.03	0.8 ± 0.06	0.88 ± 0.37	0.74 ± 0.18	<b>0.83 ± 0.13</b>
S2( $\alpha$ )	0.61 ± 0.01	0.6 ± 0.03	0.66 ± 0.03	0.61 ± 0.01	0.5 ± 0.02	0.6 ± 0.04	0.61 ± 0.03	0.6 ± 0.03	<b>0.59 ± 0.04</b>
S1( $\beta$ )	0.57 ± 0.15	0.63 ± 0.1	0.6 ± 0.1	0.65 ± 0.21	0.65 ± 0.08	0.63 ± 0.11	0.67 ± 0.11	0.56 ± 0.17	<b>0.62 ± 0.03</b>
S2( $\beta$ )	0.29 ± 0.07	0.39 ± 0.12	0.26 ± 0.06	0.39 ± 0.14	0.4 ± 0.06	0.32 ± 0.13	0.32 ± 0.05	0.37 ± 0.11	<b>0.34 ± 0.05</b>
S1( $\gamma$ )	2.44 ± 0.82	2.7 ± 0.96	2.74 ± 1.33	2.91 ± 0.85	3.33 ± 1.18	3.18 ± 1.2	2.56 ± 0.87	3.12 ± 1.05	<b>2.89 ± 1.07</b>
S2( $\gamma$ )	7.47 ± 2.13	6.2 ± 1.92	6.74 ± 2.32	6.63 ± 1.6	7.02 ± 2.25	8.24 ± 2.35	7.8 ± 2.6	6.83 ± 1.49	<b>7 ± 2.5</b>
HR (bpm)	162	180	170	194	186	192	198	180	

If we consider the squared modulus of the S-transform or the ST-spectrogram, we obtain an energy distribution of the signal in time-frequency plane. The ST-spectrogram is given as

$$|S_x(t, f)|^2 = \left| \int_{-\infty}^{+\infty} x(\tau) w(\tau - t) e^{-2\pi j f \tau} d\tau \right|^2. \quad (12)$$

The ST-spectrogram is normalized as follows:

$$|S_x^{\text{norm}}(t, f)|^2 = \frac{|S_x^{\text{norm}}(t, f)|^2}{\iint |S_x^{\text{norm}}(t, f)|^2 dt df}. \quad (13)$$

The proposed feature based on the kurtosis can be given as

$$\gamma = \frac{E \left[ \left( \text{SSE}(|S_x^{\text{norm}}(t, f)|^2) - \mu_x \right)^4 \right]}{\left\{ E \left[ \left( \text{SSE}(|S_x^{\text{norm}}(t, f)|^2) - \mu_x \right)^2 \right] \right\}^2}, \quad (14)$$

where  $\text{SSE}(|S_x^{\text{norm}}(t, f)|^2)$  are the values of the SSE envelope applied on the normalized ST-spectrogram  $|S_x^{\text{norm}}(t, f)|^2$  and  $\mu_x$  is the corresponding mean value. The kurtosis measures the peakedness of the distribution. This feature will try to

describe the shape of the estimated time-frequency distribution for the segmented sounds (S1 or S2).

Figure 8 shows an example of the  $\gamma$  feature calculated on S1 and S2 sounds from their SSE envelopes based on ST-Spectrogram.

### 3. Results and Discussions

The segmentation of heart sounds is established by using the modified SSE method proposed in Section 2.5. The different proposed features are tested separately and a comparison study with the HFS feature proposed in the literature is performed. The proposed features  $\alpha_{\text{opt}}$ ,  $\beta$ , and  $\gamma$  are calculated for each segmented sound and the results are summarized in Table 1. The total number of S1 and S2 in the database is 2636 (1318 S1 and 1318 S2) sounds that correspond to 62 heart signals and 8 subjects.

**3.1. Results for Feature  $\alpha_{\text{opt}}$ .** Results show that the mean value of  $\alpha_{\text{opt}}$  is greater for S1 than S2 ( $0.83 \pm 0.13$  and  $0.59 \pm 0.04$ , resp.), which means that the width of the Gaussian window (see Figure 5) obtained with the optimization of the energy concentration is wider for S1 than S2. In other words, the

TABLE 2: Significance values (Mann-Whitney  $U$  test), range (min and max), and the area under curve (AUC) results obtained for all subjects and for each proposed feature.

Feature	$P$ value	Range (S1)	Range (S2)	AUC
$\alpha$	<0.0001	0.5–2	0.5–0.92	0.85
$\beta$	<0.0001	0.3–1	0.13–0.88	0.87
$\gamma$	<0.0001	1–7.8	2–19.2	<b>0.96</b>

algorithm needs a higher frequency resolution for S1 than S2. The first heart sound has a booming quality and is lower pitched, duller, and longer than the second heart sound [19]; the S1 can be considered more complex (containing more components) than S2 from a physiological point of view and in term of frequency components which explain the need of higher frequency resolution or larger analysis window for S1 compared to S2.

Figure 9 shows the AUC for the  $\alpha_{\text{opt}}$  feature and for each subject. The lowest AUC corresponds to the subject 4 (0.7). The highest AUC is 0.91 and the total average of AUCs is 0.85.

The probability that the two groups (S1( $\alpha$ ) and S2( $\alpha$ )) come from distributions with different medians is calculated by the Mann-Whitney  $U$  test ( $P < 0.0001$ ) (Table 2). Significant differences between the two groups, with 95% confidence, are found. The classification results are promising. This is very interesting since this parameter  $\alpha$  was also used to refine the boundaries detection of S1 and S2 in the segmentation process.

**3.2. Results for Feature  $\beta$ .** Results for  $\beta$  feature show that the mean value of  $\beta$  is greater for S1 than S2 ( $0.62 \pm 0.03$  and  $0.34 \pm 0.05$ , resp.). This feature is the result of the integration over time of the SSE. The SSE envelope resumes the frequency content over time; it can be viewed as an instantaneous frequency measure followed by a nonlinear filter to attenuate the low and the high frequency intensities. Hence, the  $\beta$  feature can be considered as an integration of the modified instantaneous frequency measure which will be higher for physiologically richer signals (S1 in this case).

The probability that the two groups (S1( $\beta$ ) and S2( $\beta$ )) come from distributions with different medians is calculated by the Mann-Whitney  $U$  test ( $P < 0.0001$ ) (Table 2). Significant differences between the two groups, with 95% confidence, are found.

Figure 10 shows the AUC for the  $\beta$  feature and for each subject. The lowest AUC corresponds to the subject 4 (0.77). The highest AUC is 0.96 and the total average is 0.87 which is higher than the  $\alpha_{\text{opt}}$  feature. The low AUC results for subject 4 obtained with  $\alpha_{\text{opt}}$  and  $\beta$  can be explained by the high noise level in the acquired signal due to the acquisition process.

**3.3. Results for the Feature  $\gamma$ .** Results for  $\gamma$  feature show that the mean value of  $\gamma$  is greater for S2 than S1 ( $7 \pm 2.5$  and  $2.89 \pm 1.07$ , resp.). The  $\gamma$  feature operates on the distribution extracted via the spectrogram of the Stockwell transform (ST-spectrogram) which can be viewed as probability density function. The objective is to find a robust statistical description allowing us to discriminate accurately between the first and the second heart sounds.

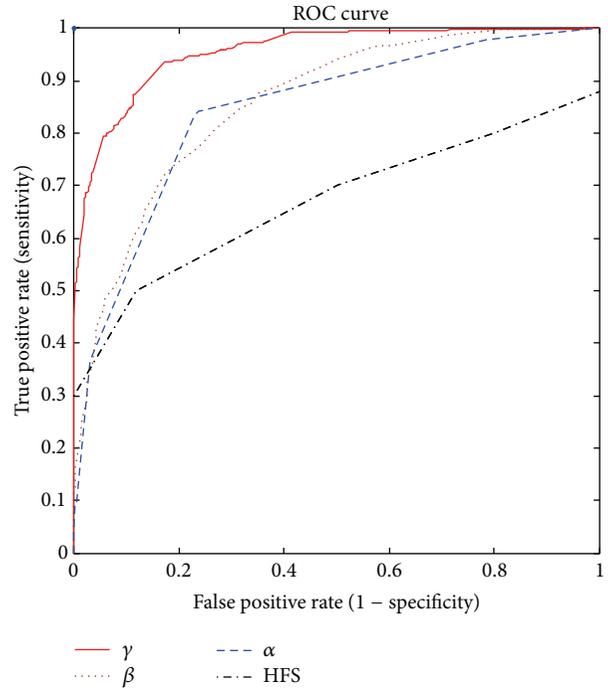


FIGURE 13: Global receiver operation characteristic curves for  $\alpha$  (Feature 1, AUC = 0.85),  $\beta$  (Feature 2, AUC = 0.87),  $\gamma$  (Feature 3, AUC = 0.96), and HFS (AUC = 0.6) features.

The S2 distribution is a heavier tail and a higher peak than the S1 distributions. This can be explained by the fact that S1 sounds are generally longer than S2 in time and they have lower frequency signature. This will lead to higher kurtosis estimation for S2. The results for the feature showed very good performance for all subjects (Figure 11) with 0.96 of total average (AUC). The probability that the two groups (S1( $\gamma$ ) and S2( $\gamma$ )) come from distributions with different medians is calculated by the Mann-Whitney  $U$  test ( $P < 0.0001$ ) (Table 2). Significant differences between the two groups, with 95% confidence, are found.

Figure 12 shows the results of segmented sound corresponding to subject 4 at the workload number 6 with a HR = 181 bmp and with different  $\alpha$ ,  $\beta$ , and  $\gamma$  features.

**3.4. Comparison with the HFS Feature.** The HFS feature shows lower results with 0.6 AUC (Figure 13). This is not surprising because the HFS method is based on several imprecise hypotheses. First, as we mentioned it above, not all HFS signatures correspond necessarily to class (S1 or S2) as authors propose in [18]. Figure 3 presents a normal heart sound with the corresponding HFS which shows clearly that the HFS does not correspond necessarily to one class. Moreover, the HFS method still needs the systolic duration to classify S1 and S2 which is not reliable when HR is very high (stress test for example). Finally, the HFS method explores only the frequency content of the sounds without any information on time; this becomes problematic for the nonstationary signals as the case in this study with S1 and S2 sounds (where the frequency of the signal varies over time).

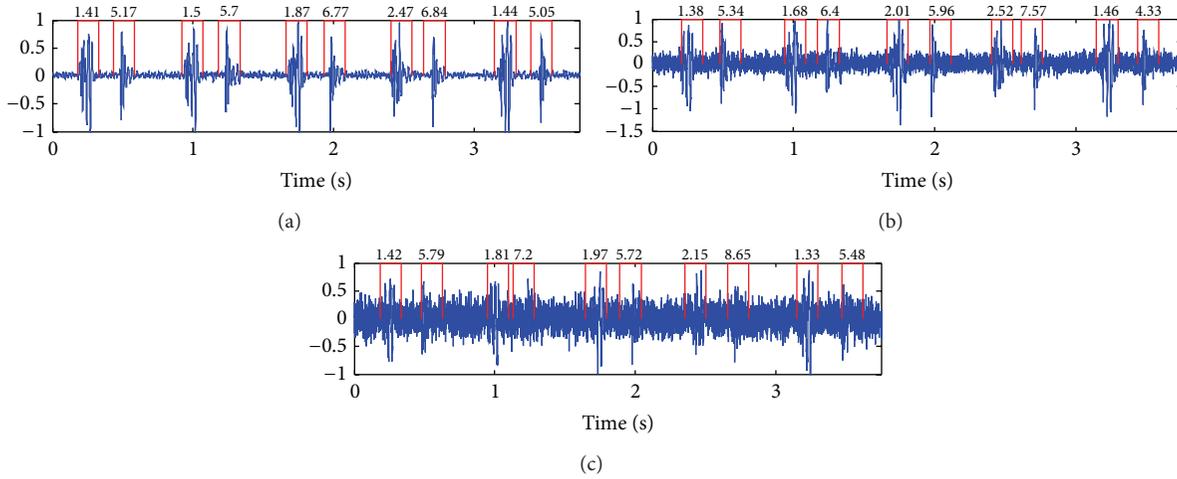


FIGURE 14: Example of a segmented stress test heart sounds with three different SNR ratios (12, 5, and 0 dB) with the values of  $\gamma$  calculated for each located sound (S1 and S2).

Figure 13 shows that the highest performance is reached by the proposed  $\gamma$  feature with AUC = 0.96.

**3.5. Robustness of the Proposed Features against Noise.** In this subsection, we study the robustness of the proposed features against noise. The sounds collected in the database were already contaminated with both physiological noise and background noise. Here, we will study clearly the robustness against noise of the proposed features by estimating the SNR ratio on selected heart sounds from the database and we add additive white Gaussian noise with three different levels. The average of the SNR for all sounds in the database is estimated to be 10 dB. To test the robustness against noise, two other levels of noise are added and the results are showed in Table 3.

Results in Table 3 show clearly the high robustness of the proposed features against noise. This is not surprising since these features are based on time-frequency domain [2, 29]. The HFS feature does not show reliable results on the sounds used in this study. Figure 14 shows a selected sound (from subject 1) with three different levels of noise and the corresponding  $\gamma$  features being calculated on the segmented S1 and S2 sounds.

**3.6. High Frequency Content Ratio S2/S1 Variation over Heart Rate.** This section aims to analyze the frequency content of S1 and S2 over the HR. As we mentioned it before, the fact that the frequency content of S2 exhibits higher frequency content than S1 cannot be generalized and adopted as robust feature to characterize the physiology related to the heart sounds. It is known, for example, that the intensity and frequency of S1 are affected by the velocity of the forces responsible for the acceleration and deceleration of the blood masses, which on the other hand are directly related to the HR [20].

To show that experimentally, the high frequency content of S1 and S2 for the 8 subjects of the database is estimated. Then, the mean and the standard deviation of the ratio of the frequency content of S2 over the frequency content of S1 ( $S2/S1$ ) are calculated in relation with the heart rate. For each

TABLE 3: AUC results for 3 levels of noise added on different sounds in the database.

Feature	10 dB	5 dB	0 dB
$\alpha$	0.85	0.82	0.8
$\beta$	0.87	0.85	0.83
$\gamma$	0.96	0.93	0.91
HFS	0.6	0.58	0.57

subject, 4 sounds are considered at 4 different stress levels (workload levels).

The results in Figure 15 show clearly the direct relation of the frequency content of the heart sounds and the heart rate. Normally, S2 has a higher frequency content than S1 ( $S2/S1 > 1$ ) except for subject 5. The red line in Figure 15 indicates when the frequency content of S1 exceeds the frequency content of S2. When the heart rate increases, the high frequency content of S2 decreases which decrease the ratio  $S2/S1$  (see Figures 15 and 16). This confirms our motivation to propose other features than the high frequency content signature to discriminate between S1 and S2.

## 4. Conclusions and Future Work

The main objective of this paper is to study the ability of new features to segment and discriminate S1 and S2 in stress test conditions data. First, we have proposed a modified version of the SSE segmentation method to take into consideration the very low intensities sounds which can occurs more often with S2 when the cardiac frequency is high. Then, we have investigated three new time-frequency features computed from the S-transform which can be considered as a hybrid method between the STFT and the wavelets. The proposed features are validated on a database that contains 2636 S1 and S2 sounds (1318 S1 and 1318 S2) that correspond to 62 heart signals and 8 subjects under cardiac stress test collected from healthy subjects.

Classifying S1 and S2 based on the PCG signal without any other reference is a hard task since they are very sensitive

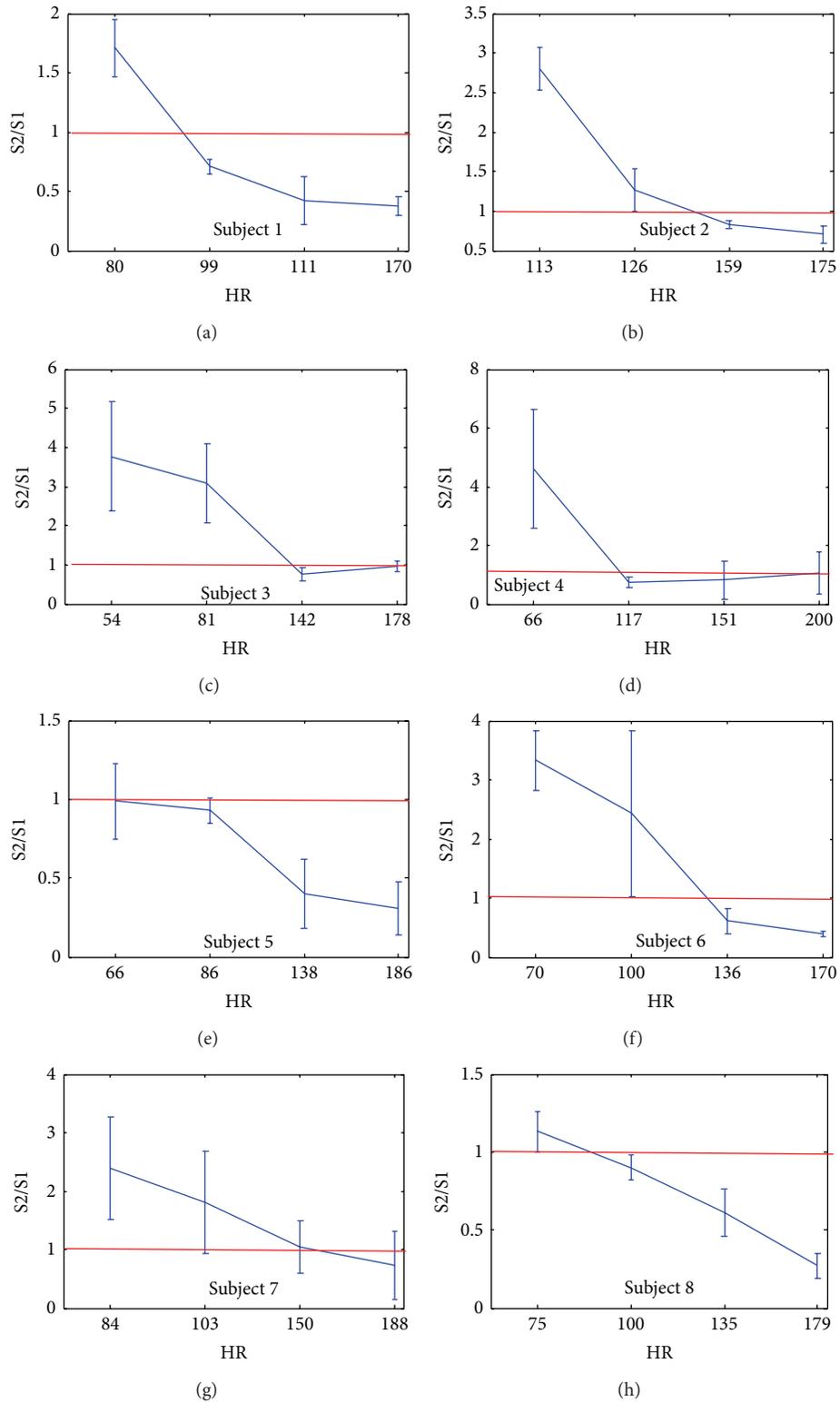


FIGURE 15: The variation of the high frequency content ratio ( $S2/S1$ ) over the HR for all subjects. The red lines indicate the high frequency content of S2 becoming lower than that of S1.

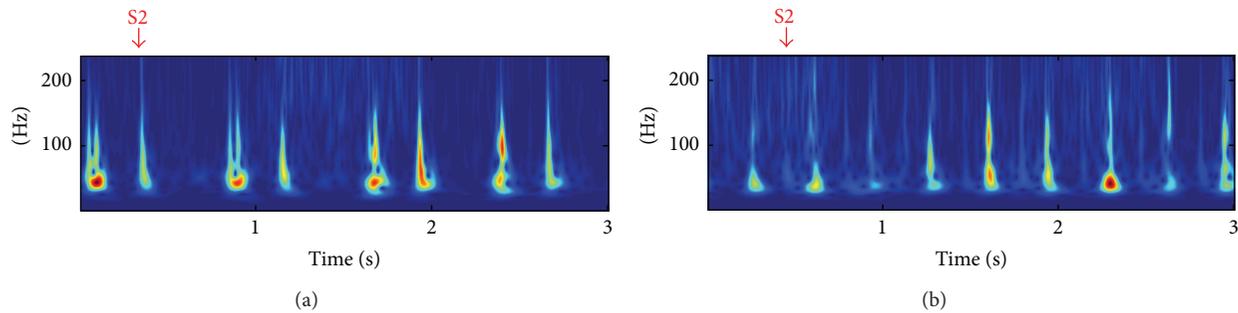


FIGURE 16: The S-transform of two sounds corresponding to the same subject and two different heart rates ((a) HR = 80 bpm, (b) 142 bpm)) showing the high frequency content of S2 decreasing when the heart rate is higher.

to several parameters like breathing, cardiac frequency, and other biomedical and environmental conditions.

The proposed features aim to describe the time-frequency behavior of each sound. The  $\alpha_{\text{opt}}$  feature corresponds to the optimal width of the Gaussian window that maximizes the energy concentration of the signal. The  $\beta$  feature is the result of integration over time of the SSE envelope (which can be viewed as a modified measure of instantaneous frequency), while the  $\gamma$  feature is based on the kurtosis of the extracted SSE envelope via the normalized ST-spectrogram. The  $\gamma$  shows the highest performance with AUC = 0.96. This is very interesting since it shows the ability to discriminate accurately between clinical S1 and S2 sounds by using a single feature which simplifies the segmentation module.

The comparison with the existing method, namely, HFS (high frequency signature), in the literature [18, 19] shows a large superiority for our proposed features, most notably the  $\gamma$  feature.

Moreover, an experimental validation is performed in this study to show the high frequency content ratio (S2/S1) variation over heart rate.

Finally, the proposed methods might have high potential to study changes in the shape of the hearts sound due to splits between M/T and A/P components in the respiration cycle or to classify other biomedical and nonstationary signals.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

The authors would like to thank the financial support from the French Telemedicine Project, E-Care. Thanks again are due to the reviewers for their constructive comments and suggestions, which have improved the paper substantially.

## References

- [1] S. M. Debbal and F. Berekci Reguig, "Choix de l'ondelette analysante et classification des signaux phonocardiogrammes en fonction des souffles surajoutés," *Afrique Science*, vol. 1, no. 1, pp. 1–13, 2004.
- [2] A. Moukadem, A. Dieterlen, N. Hueber, and C. Brandt, "A robust heart sounds segmentation module based on S-transform," *Biomedical Signal Processing and Control*, vol. 8, no. 3, pp. 273–281, 2013.
- [3] Z. Dokur and T. Ölmez, "Heart sound classification using wavelet transform and incremental self-organizing map," *Digital Signal Processing: A Review Journal*, vol. 18, no. 6, pp. 951–959, 2008.
- [4] C. N. Gupta, R. Palaniappan, S. Swaminathan, and S. M. Krishnan, "Neural network classification of homomorphic segmented heart sounds," *Applied Soft Computing Journal*, vol. 7, no. 1, pp. 286–297, 2007.
- [5] A. A. Sepehri, A. Gharehbaghi, T. Dutoit, A. Kocharian, and A. Kiani, "A novel method for pediatric heart sound segmentation without using the ECG," *Computer Methods and Programs in Biomedicine*, vol. 99, no. 1, pp. 43–48, 2010.
- [6] V. Nigam and R. Priemer, "Assessing heart dynamics to estimate durations of heart sounds," *Physiological Measurement*, vol. 26, no. 6, pp. 1005–1018, 2005.
- [7] H. Naseri and M. R. Homaeinezhad, "Detection and boundary identification of phonocardiogram sounds using an expert frequency-energy based metric," *Annals of Biomedical Engineering*, vol. 41, no. 2, pp. 279–292, 2013.
- [8] Z. Dokur and T. Ölmez, "Feature determination for heart sounds based on divergence analysis," *Digital Signal Processing*, vol. 19, no. 3, pp. 521–531, 2009.
- [9] Z. Yan, Z. Jiang, A. Miyamoto, and Y. Wei, "The moment segmentation analysis of heart sound pattern," *Computer Methods and Programs in Biomedicine*, vol. 98, no. 2, pp. 140–150, 2010.
- [10] S. E. Schmidt, C. Holst-Hansen, C. Graff, E. Toft, and J. J. Struijk, "Segmentation of heart sound recordings by a duration-dependent hidden Markov model," *Physiological Measurement*, vol. 31, no. 4, pp. 513–529, 2010.
- [11] R. G. Stockwell, L. Mansinha, and R. P. Lowe, "Localization of the complex spectrum: the S transform," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 998–1001, 1996.
- [12] G. Ferro, A. Guinta, S. Maione, G. Carella, A. Genovese, and M. Chiariello, "Diastolic time during exercise in normal subjects and in patients with coronary artery disease. A plethysmographic study," *Cardiology*, vol. 71, no. 5, pp. 266–272, 1984.
- [13] G. Ferro, C. Duilio, L. Spinelli, G. A. Liucci, F. Mazza, and C. Indolfi, "Relation between diastolic perfusion time and coronary artery stenosis during stress-induced myocardial ischemia," *Circulation*, vol. 92, no. 3, pp. 342–347, 1995.
- [14] V. Gemignani, E. Bianchini, F. Faista et al., "Assessment of cardiologic systole and diastole duration in exercise stress

- test with a transcutaneous accelerometer sensor,” *Computers in Cardiology*, vol. 35, pp. 153–156, 2008.
- [15] S. M. M. Rønved, I. Gjerløv, A. Brokjær, and S. E. Schmidt, “Phonocardiographic recordings of first and second heart sound in determining the systole/diastole-ratio during exercise test,” in *Proceedings of the 15th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics (NBC '11)*, vol. 34, pp. 85–88, Aalborg, Denmark, June 2011.
- [16] K. Saunamäki, K. Egstrup, L. Krusell et al., *Vejledende retningslinjer for klinisk arbejdstest i relation til iskæmisk hjertesygdom*, Dansk Cardiologisk Selskab, 2001.
- [17] J. Hansen, N. H. Zimmermann, S. Schmidt, D. Hammershoi, and J. Struijk, “System for acquisition of weak murmurs related to coronary artery diseases,” in *Proceedings of the Computing in Cardiology*, pp. 213–216, Hangzhou, China, September 2011.
- [18] D. Kumar, P. Carvalho, M. Antunes et al., “Detection of S1 and S2 heart sounds by high frequency signatures,” in *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '06)*, pp. 1410–1416, IEEE, New York, NY, USA, September 2006.
- [19] D. Kumar, P. Carvalho, M. Antunes, R. P. Paiva, and J. Henriques, “An adaptive approach to abnormal heart sound segmentation,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '11)*, pp. 661–664, 2011.
- [20] J. M. Felner, “The first heart sound,” in *Clinical Methods: The History, Physical, and Laboratory Examinations*, H. K. Walker, W. D. Hall, and J. W. Hurst, Eds., chapter 22, Butterworths, Boston, Mass, USA, 1990.
- [21] J. M. Felner, “The second heart sound,” in *Clinical Methods: The History, Physical, and Laboratory Examinations*, H. K. Walker, W. D. Hall, and J. W. Hurst, Eds., chapter 23, Butterworths, Boston, Mass, USA, 3rd edition, 1990.
- [22] S. Assous and B. Boashash, “Evaluation of the modified S-transform for time-frequency synchrony analysis and source localisation,” *Eurasip Journal on Advances in Signal Processing*, vol. 2012, no. 1, article 49, 2012.
- [23] C. R. Pinnegar and L. Mansinha, “The S-transform with windows of arbitrary and varying shape,” *Geophysics*, vol. 68, no. 1, pp. 381–385, 2003.
- [24] M. Biswal and P. K. Dash, “Detection and characterization of multiple power quality disturbances with a fast S-transform and decision tree based classifier,” *Digital Signal Processing*, vol. 23, no. 4, pp. 1071–1083, 2013.
- [25] L. Stanković, “Measure of some time-frequency distributions concentration,” *Signal Processing*, vol. 81, no. 3, pp. 621–631, 2001.
- [26] E. Sejdić, I. Djurović, and J. Jiang, “A window width optimized S-transform,” *Eurasip Journal on Advances in Signal Processing*, vol. 2008, Article ID 672941, 2008.
- [27] R. G. Baraniuk, P. Flandrin, A. J. Janssen, and O. J. Michel, “Measuring time-frequency information content using the Rényi entropies,” *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1391–1409, 2001.
- [28] R. Ferzli, L. J. Karam, and J. Caviedes, “A robust image sharpness metric based on kurtosis measurement of wavelet coefficients,” in *Proceedings of the 1st International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, 2005.
- [29] E. Sejdić, I. Djurović, and J. Jiang, “Time-frequency feature representation using energy concentration: an overview of recent advances,” *Digital Signal Processing*, vol. 19, no. 1, pp. 153–183, 2009.

## Research Article

# Developing an Intelligent Automatic Appendix Extraction Method from Ultrasonography Based on Fuzzy ART and Image Processing

Kwang Baek Kim,<sup>1</sup> Hyun Jun Park,<sup>2</sup> Doo Heon Song,<sup>3</sup> and Sang-suk Han<sup>4</sup>

<sup>1</sup> Department of Computer Engineering, Silla University, Busan 617-736, Republic of Korea

<sup>2</sup> Department of Computer Engineering, Pusan National University, Busan 609-735, Republic of Korea

<sup>3</sup> Department of Computer Games, Yong-In Songdam College, Yongin 449-040, Republic of Korea

<sup>4</sup> Department of Radiology, Inje University Busan Paik Hospital, Busan 614-735, Republic of Korea

Correspondence should be addressed to Kwang Baek Kim; [gbkim@silla.ac.kr](mailto:gbkim@silla.ac.kr) and Sang-suk Han; [hansono@hanmail.net](mailto:hansono@hanmail.net)

Received 15 July 2014; Revised 19 September 2014; Accepted 26 September 2014

Academic Editor: Kevin Ward

Copyright © 2015 Kwang Baek Kim et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ultrasound examination (US) does a key role in the diagnosis and management of the patients with clinically suspected appendicitis which is the most common abdominal surgical emergency. Among the various sonographic findings of appendicitis, outer diameter of the appendix is most important. Therefore, clear delineation of the appendix on US images is essential. In this paper, we propose a new intelligent method to extract appendix automatically from abdominal sonographic images as a basic building block of developing such an intelligent tool for medical practitioners. Knowing that the appendix is located at the lower organ area below the bottom fascia line, we conduct a series of image processing techniques to find the fascia line correctly. And then we apply fuzzy ART learning algorithm to the organ area in order to extract appendix accurately. The experiment verifies that the proposed method is highly accurate (successful in 38 out of 40 cases) in extracting appendix.

## 1. Introduction

The appendix vermiformis is a vestigial, tubular organ that arises from the inferior pole of the cecum, 2–2.5 cm inferior to the ileocecal junction. Normal appendix varies in length from 5 to 35 cm (average 8 cm) in adult. It is seen as a blind-ended tubular structure whose orifice is usually constant in position; however, its body and tip are located in various sites including retrocecal (behind the cecum), pelvic, subcecal (below the cecum), and pre- or postileal (anterior or posterior to the terminal ileum) [1, 2].

Appendicitis, an inflammation of the appendix, is the most common abdominal surgical emergency. It is believed to occur as a result of appendiceal luminal obstruction which is most commonly caused by a fecalith. Luminal bacteria multiply and invade the appendiceal wall as venous engorgement and subsequent arterial compromise result from the high intraluminal pressures. Finally, gangrene and perforation

occur. If the process evolves slowly, adjacent organs such as the terminal ileum, cecum, and omentum may wall off the appendiceal area so that a localized abscess will develop, whereas rapid progression of vascular impairment may cause perforation with free access to the peritoneal cavity.

Typically, the illness begins with vague midabdominal discomfort followed by nausea, anorexia, and indigestion and within several hours the pain migrates to the right lower quadrant. Examination at this point shows localized tenderness to one-finger palpation and perhaps slight muscular guarding. Rebound or percussion tenderness (the latter provides the same information more humanely) may be elicited in the same area [3].

However, there are various kinds of difficulties in the diagnosis of acute appendicitis. The classic sequence of symptoms occurs in only 66% of patients. And the diagnosis of acute appendicitis is particularly difficult in the very young and in the elderly. These are the groups in which diagnosis

is most often delayed and perforation most common due to the lack of classic symptoms. The highest incidence of false-positive diagnosis (20%) is in women between ages 20 and 40 and is attributable to pelvic inflammatory disease and other gynecologic conditions [4]. Diagnosing appendicitis in pregnancy also can be difficult because the nausea, vomiting, and abdominal pain of appendicitis can also be features of pregnancy and physical examination may not be reliable in them [5].

As a result, imaging diagnosis such as ultrasonography (US), computed tomography (CT), and magnetic resonance imaging (MRI) is essential in confirming or excluding the diagnosis of acute appendicitis in clinically suspected acute appendicitis. Among these modalities, US examination does a key role in the management of the clinically suspected appendicitis. US examination should be the first imaging test performed, particularly among the pediatric and young adult populations, who represent the main targets for appendicitis, as well as in pregnant patients. A positive US examination for appendicitis or an alternative diagnosis of possible gastrointestinal or urological origin, or a negative US, either showing a normal appendix or presenting low clinical suspicion of appendicitis, should lead to a final diagnosis. A negative or indeterminate examination with a strong clinical suspicion of appendicitis should be followed by a CT scan or alternatively, a MRI scan in a pregnant patient. A second US examination in a patient with persistent symptoms, especially if the first one was performed by a less-experienced imaging professional, is a valid alternative to a CT [6, 7].

Sonographic findings of acute appendicitis include outer appendiceal diameter enlarged to 6 mm or greater under compression, intraluminal fluid, lack of compressibility, visualization of appendicolith, increased color signals along its wall, cecal wall thickening, periileal lymph nodes, and peritoneal fluid. Among these findings, a threshold 6 mm diameter of the appendix under compression is the most accurate US finding for appendicitis [8]. Thus, the critical point, 6 mm of the diameter of the appendix is a crucial factor in decision making for appendectomy. As a result, the measurement error of 1 mm near the critical point may lead doctors to a serious misdiagnosis.

By the way, current naked-eye examination of the US images has some limitations in accurate measurement in cases of unclear delineation of the appendix with thick abdomen and in cases showing ill-defined borders of the appendix by surrounding tissues. Thus, the need for more clear delineation and accurate extraction of appendix from surrounding tissues is always present to radiologists in the field of abdominal imaging.

Thus, there are growing needs for an intelligent decision tool for more accurate diagnosis by artificial intelligence technology. Unfortunately, there are few tools for the practitioners to use with credibility up to date. A preliminary study applies several histogram thresholding methods in detecting appendix [9] but that method is weak when the brightness contrast is not very high and will have potential information loss in edge linking procedure. Our previous study [10] used fuzzy logic in binarization procedure to enhance the brightness contrast and other studies used

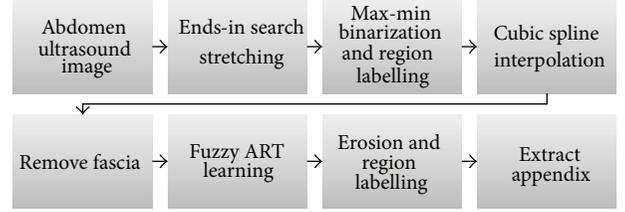


FIGURE 1: Process of Appendix extraction.

$K$ -means clustering [11] method to extract the target appendix as accurate as possible. With such a pixel clustering method, the appendix area has more contrast in extraction; however, the correct extraction rate is below satisfaction because patient's ascites of a significant size may mislead the system to extract it as a false positive appendix or when the shape of appendix is extraordinary, the system could not catch the correct fascia line that is the main predictor of appendix location of our logic.

Thus, in this paper, we propose a more efficient method to extract appendix area correctly by using fuzzy ART algorithm in the critical phase instead of  $K$ -means. Knowing that the appendix is located at the lower organ area below the bottom fascia line, we conduct a series of image processing techniques to find the fascia line correctly. Figure 1 demonstrates the overall process of our method.

The first step is to enhance the brightness contrast by Ends-in Search Stretching [12] and remove noises by Max-Min binarization and region labeling method. Then the fascia area is extracted with cubic spline interpolation [13]. The appendix area is then extracted from that image by applying fuzzy ART algorithm as explained in Section 3.

## 2. Removing Fascia Area

First, we apply End-in Search Stretching [12] to enhance the brightness contrast with the formula (1). Since the abdomen image is usually dark, it may not be sufficient to discriminate fascia, muscles, and other areas as it is given. Consider the following:

$$S(x, y) = \begin{cases} 0 & P(x, y) \leq \text{Min} \\ 255 \times \frac{P(x, y) - \text{Min}}{\text{Max} - \text{Min}} & \text{Min} < P(x, y) < \text{Max} \\ 255 & P(x, y) \geq \text{Max}, \end{cases} \quad (1)$$

where Min and Max are thresholds,  $P(x, y)$  denotes the brightness value in the original image, and  $S(x, y)$  denotes the result. Figure 2 shows the effect of Ends-in Search Stretching. Two different input images shown in Figure 2(a) become more brightness contrast as shown in Figure 2(b). Throughout the paper, we will show two example images at the same time in this section so that the effect of our image processing subsystems can be identified more clearly.

From Figure 2(b), we apply Max-Min binarization and repetitive region labeling method [14] to connect related

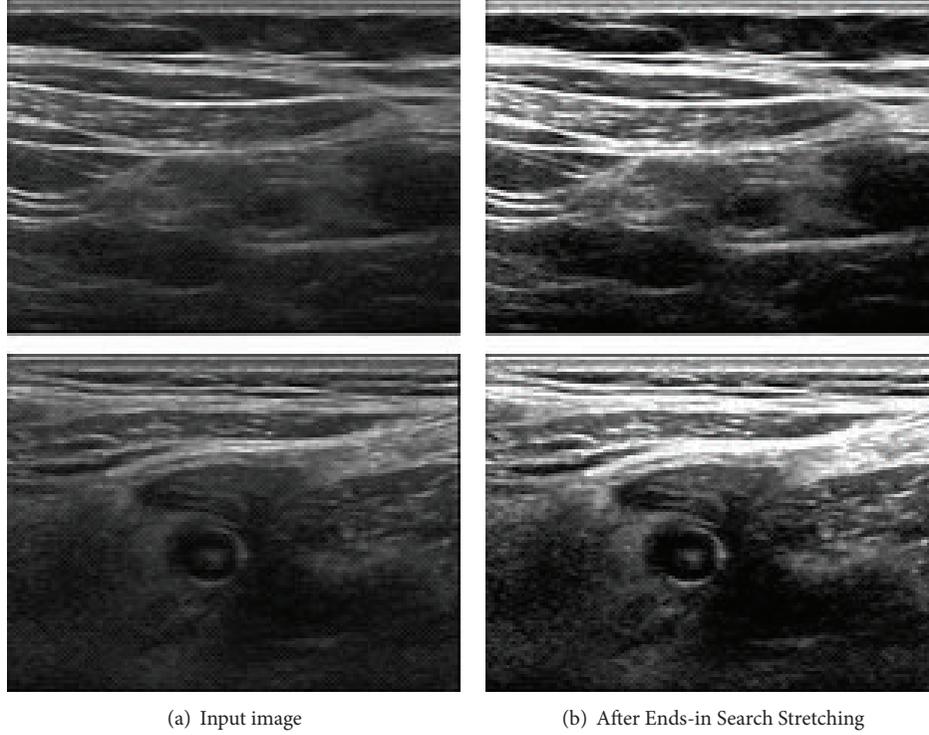


FIGURE 2: Effect of Ends-in Search Stretching.

pixels. If the connected object is too short, we remove them as noise. The result is as shown in Figure 3(b) and the experimental threshold in this paper is 1500.

Unfortunately, the binarized noise-removed image may have disconnected fascia area at the bottom due to the brightness difference of that area. In order to reconnect them, we apply cubic spline interpolation [13].

Cubic spline interpolation connects two points on the boundary when formula (2) is satisfied:

$$\begin{aligned}
 S''_{n-1}(x_n) &= S''_0(x_0) = 0 \\
 S''_{i-1}(x_i) &= S''_i(x_i) \quad i = 1, 2, \dots, n-1 \\
 S'_{i-1}(x_i) &= S'_i(x_i) \quad i = 1, 2, \dots, n-1 \\
 S_{i-1}(x_i) &= S_i(x_i) = y_i \quad i = 1, 2, \dots, n-1.
 \end{aligned} \tag{2}$$

First two equations are hard constraints but the next two are soft constraints for smoothing. Since we know that the target appendix is located below the bottom of fascia, cubic spline interpolation is applied to the bottom fascia line and the effect is shown in Figure 4.

### 3. Extract Appendix Object from Image

In order to extract the candidate area of appendix, we apply an unsupervised neural network learning algorithm called fuzzy ART [15] to the image obtained from Section 2.

The general characteristics of fuzzy ART can be summarized as follows:

- (1) an unsupervised real time learning algorithm that does not have the target value,
- (2) it creates a new cluster or merges existing clusters according to the similarity between input pattern and current set of clusters.

The main reason we adopt an unsupervised learning algorithm for our system is that the supervised learning suffers from frequent relearning of the learned patterns. Fuzzy ART algorithm is relatively immune to that problem since it has incremental learning capability and is also proven to be stable in learning [16].

The process of fuzzy ART learning in this problem can be summarized as in Figure 5.

The similarity of the input patterns is computed as formula (3) where  $\wedge$  denotes fuzzy logic Min operator:

$$\frac{\|w_{j^*i}(n) \wedge x_i\|}{\|\beta_{\max}\|} > \rho, \tag{3}$$

where  $x_i$  denotes the normalized brightness value of current pixel within  $[0, 1]$  and  $\beta_{\max}$  denotes the normalized maximum brightness value. And  $w_{j^*i}(n)$  denotes the weight of connection between all  $i$  nodes in the input layer and the node  $j$  chosen as a winner node in the cluster layer and  $\rho$  denotes the vigilance parameter.

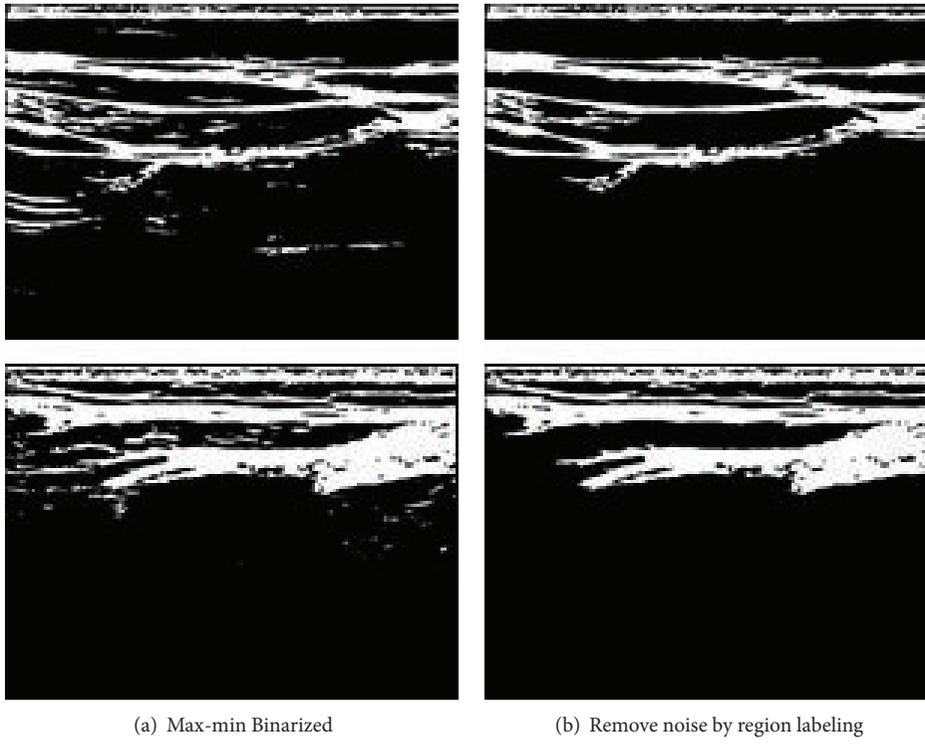


FIGURE 3: Effect of Region Labeling.

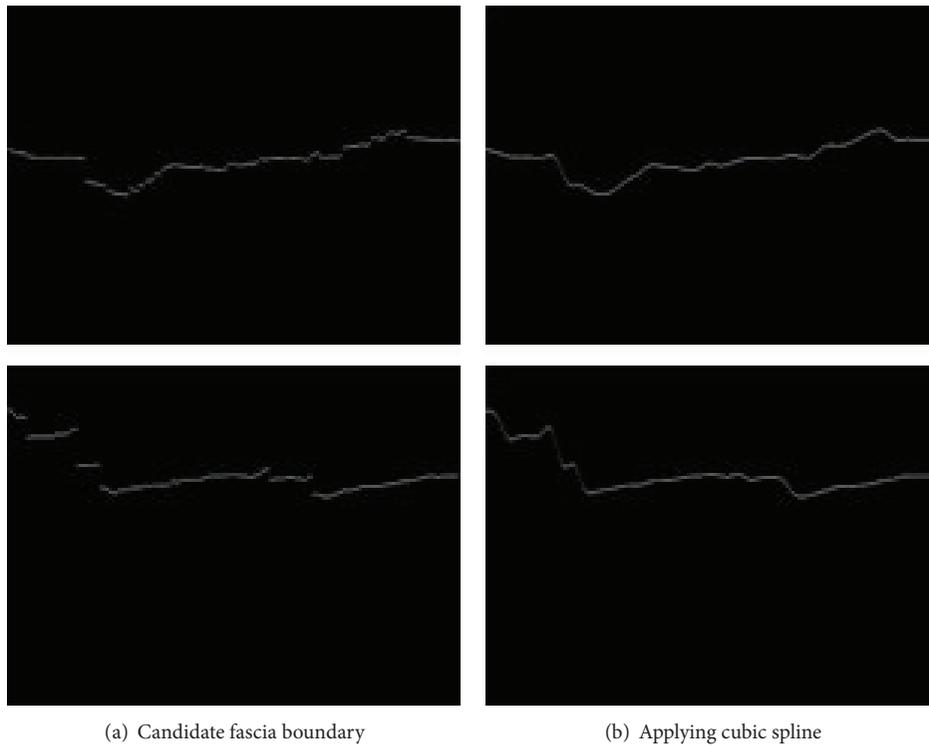


FIGURE 4: The effect of cubic spline interpolation.

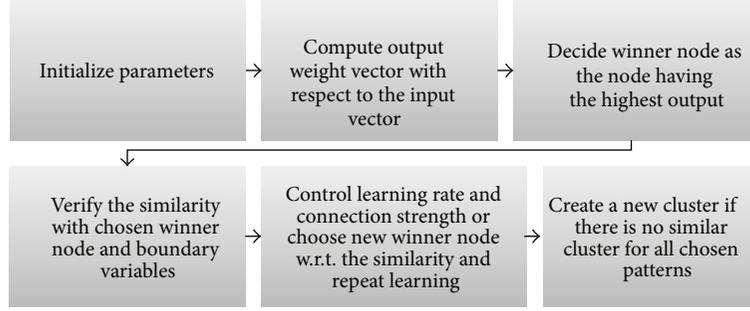


FIGURE 5: Process diagram for fuzzy ART.

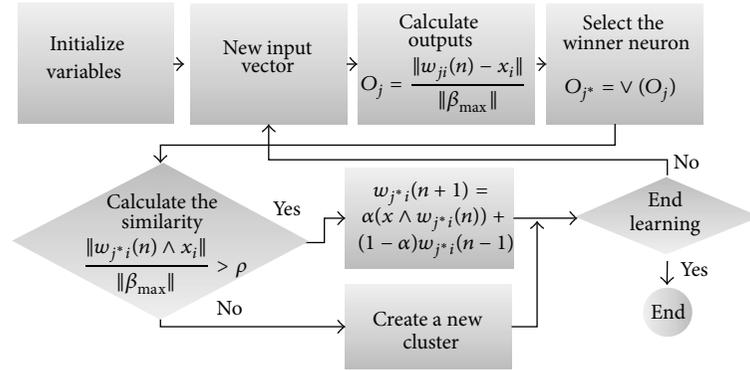


FIGURE 6: Distance based fuzzy ART process flow diagram.

The output ( $O_j$ ) is computed as formula (4) and the winner node is determined as having the maximum output as shown in formula (5):

$$O_j = \frac{\|w_{ji}(n) - x_i\|}{\|\beta_{\max}\|}, \quad (4)$$

$$O_{j^*} = v(O_j). \quad (5)$$

Then the weight  $w_{j^*i}(n+1)$  is controlled as formula (6):

$$w_{j^*i}(n+1) = \alpha(x \wedge w_{j^*i}(n)) + (1-\alpha)w_{j^*i}(n-1), \quad (6)$$

where the learning parameter  $\alpha$  is a real value between 0 and 1.

Then, the detailed algorithmic description of fuzzy ART process flow diagram is shown in Figure 6.

The main reason that we adopt fuzzy ART in extracting appendix area is to avoid cases when patient's ascites with a significant size are falsely classified as appendix. In such cases, the brightness difference itself is not sufficient to discriminate those two objects differently. That is one of the limitations of the previous research [11] using  $K$ -means pixel clustering. That problem is caused by the nature of  $K$ -means in that the clusters are based on random brightness values as centers.

The characteristics of fuzzy ART, meanwhile, is to determine single winner node with respect to the minimum error

between input and weighted pattern based on Euclidean distance metric. Thus, it is possible to discriminate ascites from appendix in brightness by checking boundary variables in fuzzy ART algorithm.

Knowing that the appendix has the shape of oval with low brightness, we apply erosion operation and region labeling method to extract the right area for the appendix with noise removal.

## 4. Experiment and Analysis

The system is implemented in Visual Studio 2010 C# with Intel(R) Core(TM) i7-2600 CPU @ 3.40 GHz and 4 GB RAM PC. Forty images containing appendicitis supplied by Busan Paik Hospital and Busan National University Medical Center are used in this experiment. The actual system gives some characteristic features of extracted appendix as shown in Figure 7. Four different example appendicitis cases extracted by our proposed method are demonstrated in Figure 8.

In Figure 9, we provide a visual comparison between the proposed method and the reimplemented version of previous  $K$ -means clustering based method [10]. By the nature of the algorithm,  $K$ -means pixel clustering is sensitive to the initial brightness values. From the same input image shown in Figure 9(a),  $K$ -means tends to cause clustering errors as shown in Figure 9(c) when there exists sufficiently large



FIGURE 7: Screenshot of proposed software.

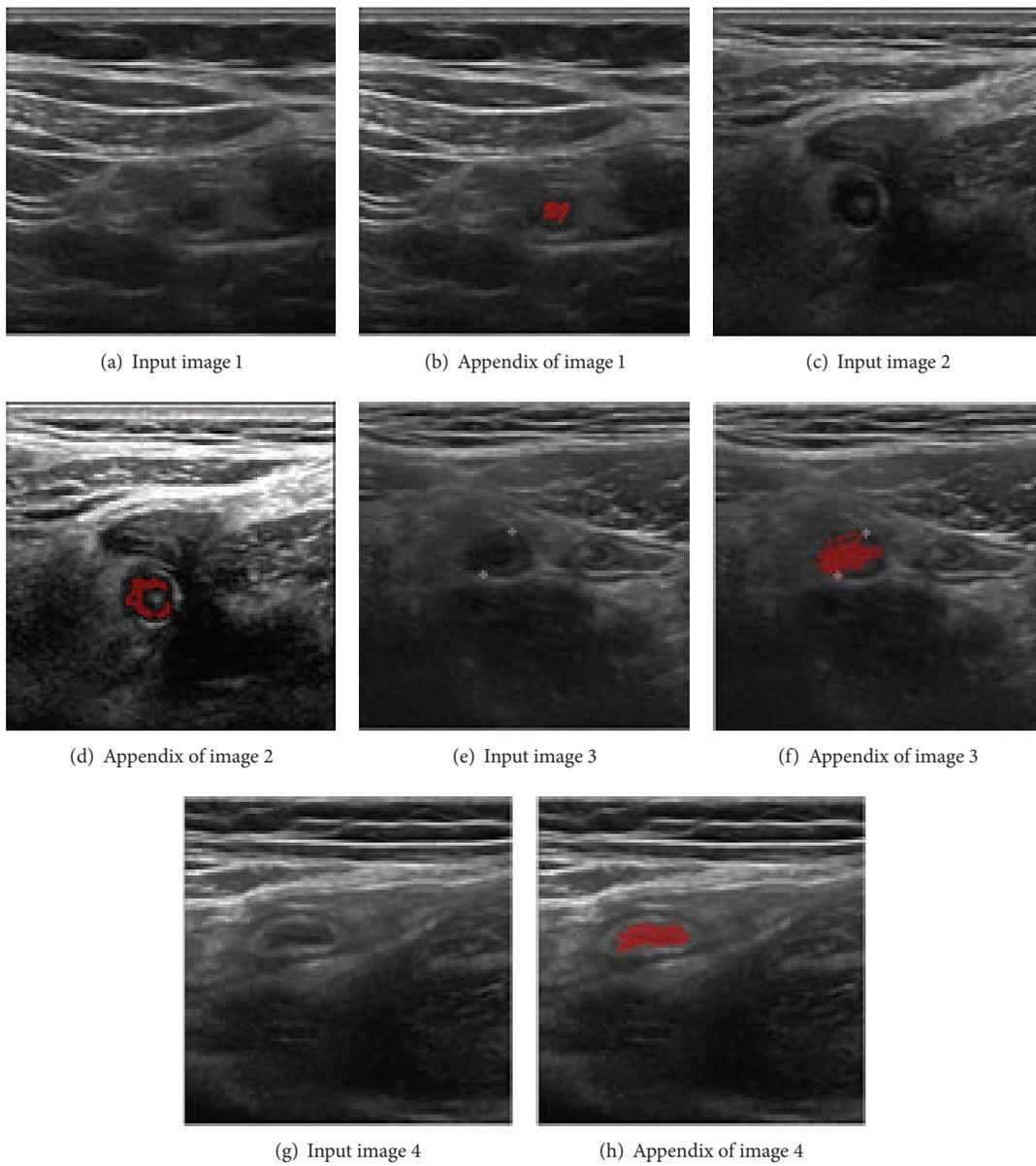


FIGURE 8: Examples of experiment result.

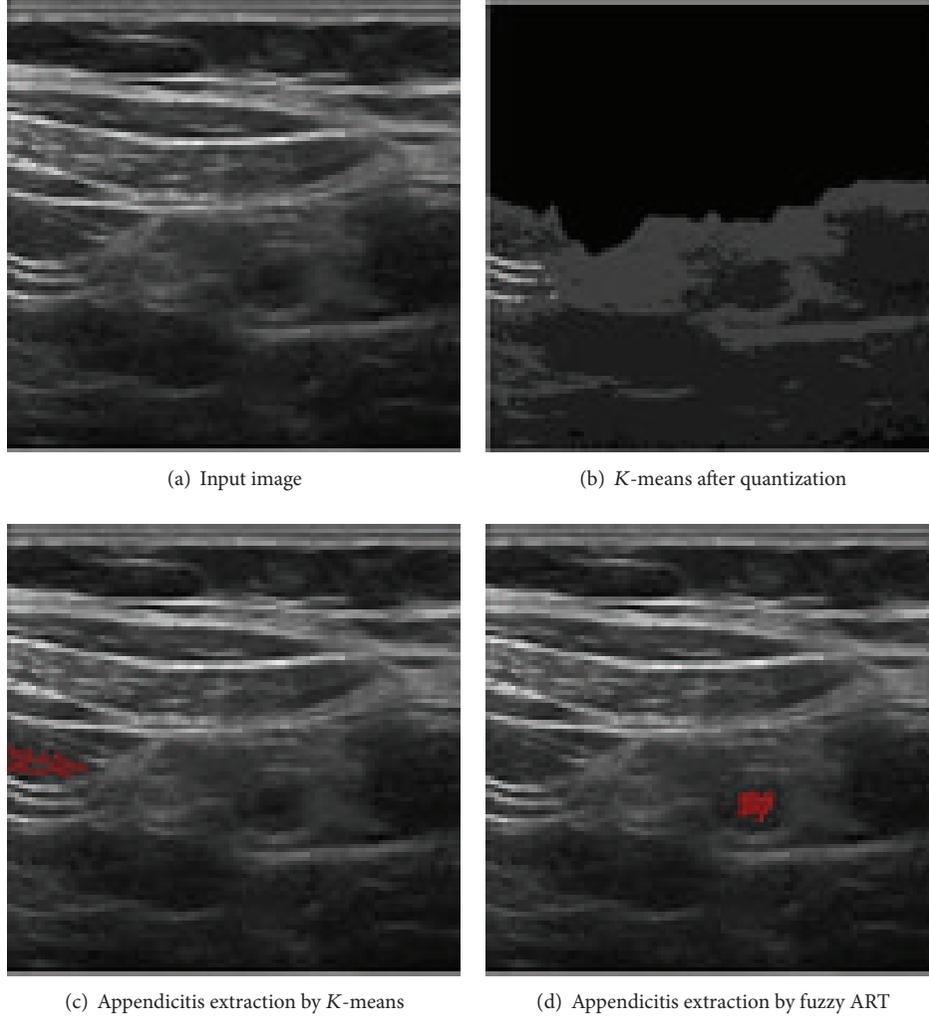


FIGURE 9: False positive case of *K*-means compared with proposed fuzzy ART in extraction.

noise not removed in the quantization process shown in Figure 9(b).

Our proposed method based on the fuzzy ART algorithm is, however, relatively immune to that brightness sensitivity and extract appendix effectively as shown in Figure 9(d) from the same input shown in Figure 9(a). In fuzzy ART process, even though the brightness contrast between the center and the observed pixel is small, the proposed system takes into account the current information with its previous state in weight control process so that clustering errors are minimized.

Also, the fuzziness of the proposed method is crucial in appendicitis extraction. Even within the same ART family, if we use ART2 algorithm [17] without fuzzy control, the performance is less than satisfactory. Figure 10 demonstrates the comparison of quantization and binarization by ART2 and proposed fuzzy ART.

Quantization result by ART2 shown in Figure 10(a) shows failed candidate area extraction of the target appendix. That is due to the sensitivity of vigilance parameter setting by the nature of ART2. In the clustering process, ART2 tends

TABLE 1: Extraction accuracy comparison (40 cases).

Method	True positive	True negative	False positive	False negative
<i>K</i> -means based	27	0	13	0
ART2 based	33	0	7	0
Proposed system	38	0	2	0

to have too many clusters than desired with respect to the setting of vigilance parameters; thus, similar characteristics of the appendix may belong to different clusters. Specifically, for the example shown in Figure 10, ART2 had 14 clusters whereas our proposed method had 8 clusters and resulted in much better result as shown in Figure 10(b). That is because fuzzy ART is relatively not sensitive to the setting of vigilance parameter; thus, our method shows more stable performance. It is worse when the binarization process is performed after such quantization. As one can see in Figure 10(c), the binarization process based on the incorrect ART2 clustering of Figure 10(a) results in the black area that the appendix

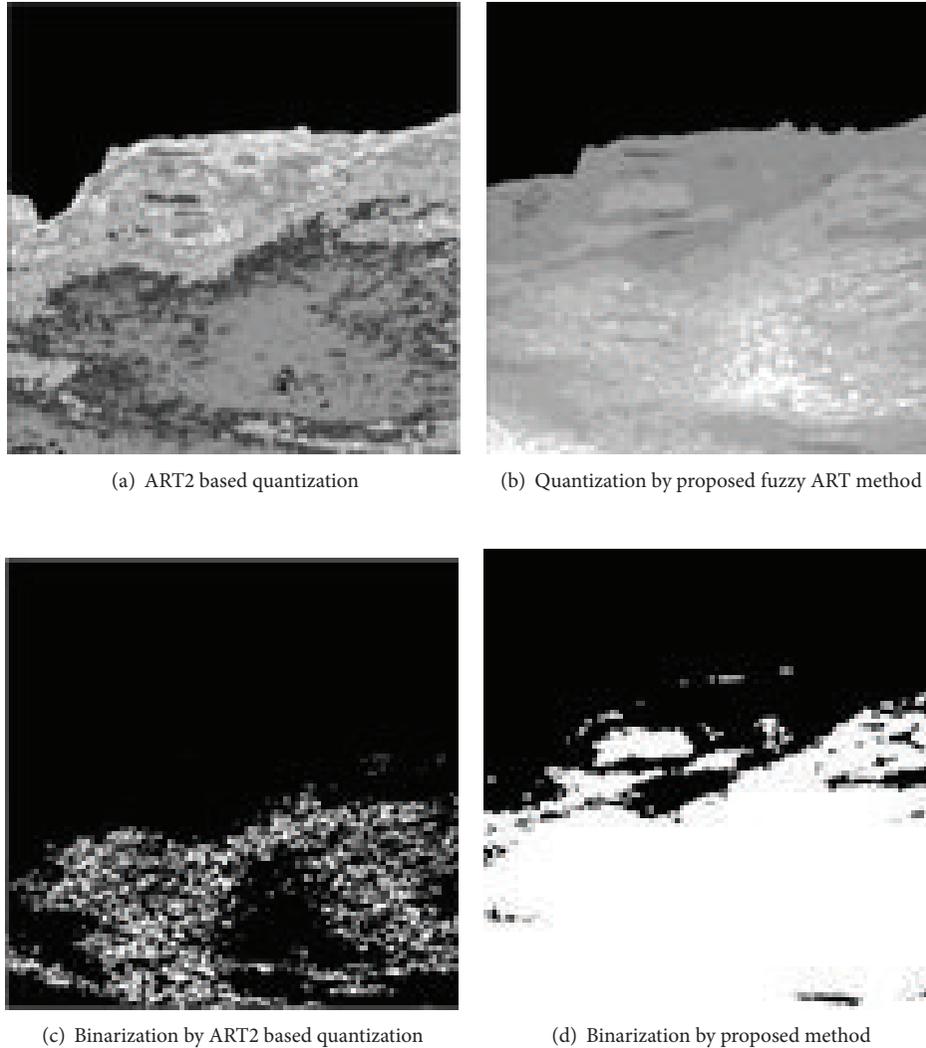


FIGURE 10: Quantization and binarization by ART2 and by proposed method.

and the background are inseparable whereas our proposed method has little difficulty to extract the appendix area correctly as shown in Figure 10(d). Thus we may conclude that the fuzzy control of clustering process is essential for the performance stability.

In conjunction with the false positive problem caused by ascites detected from previous study [10], the proposed system shows much better extraction rate as shown in Table 1. The success and failure decision of the automatic extraction by the system is determined by field expert.

Since all 40 images used in the experiment contain appendicitis, there is no true negative case and all three methods conclude an extraction of appendicitis; thus, there is no false negative either. The sensitivity of the appendicitis extraction is greatly improved from 67.5% (*K*-means based) and 82.5% (ART2 based) to 95% by the proposed method.

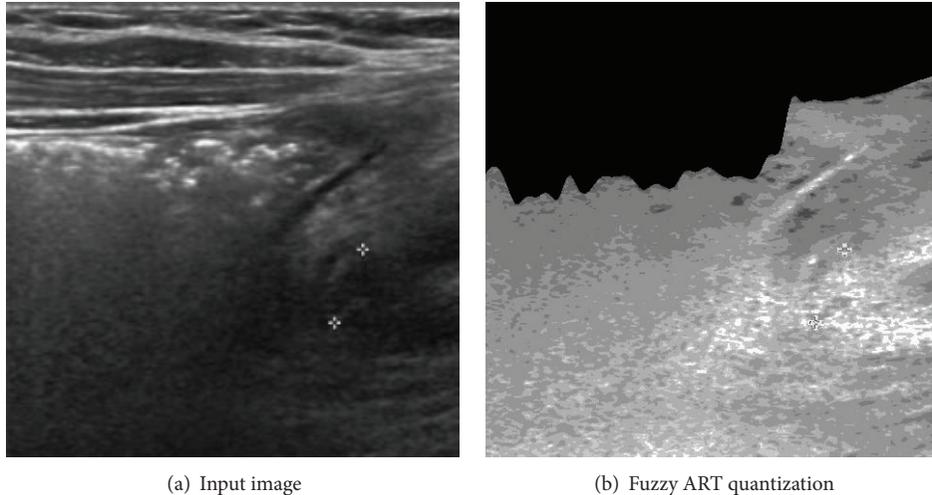
From the related literature, our result is much better than previous appendix segmentation from ultrasound image by histogram thresholding which was greatly sensitive to the

probing position [9] and the extraction accuracy is very close to that of CT result reported (33 out of 34 cases) [5].

Unfortunately, there are two failed extraction cases by our proposed method as well. As one can see from Figure 11(a), there is almost no distinction between the appendicitis area and its neighbor background. Although our proposed method applies careful brightness contrast enhancement procedures, the quantized result as shown in Figure 11(b) gives little clue in clustering and as a result, such clustering errors cause false extraction of appendicitis. In order to overcome this huddle, we need other morphological attributes in consideration such as the elliptical shape of the appendix of a certain size in the future research.

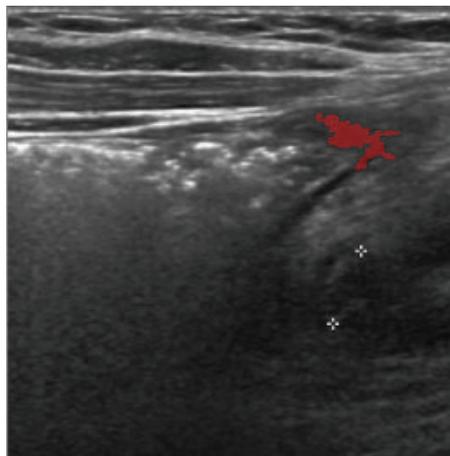
## 5. Conclusion

In this paper, we propose a method to extract appendicitis from ultrasound image automatically with various image processing techniques and fuzzy ART learning algorithm.



(a) Input image

(b) Fuzzy ART quantization



(c) False positive extraction by proposed method

FIGURE 11: Failed appendicitis extraction case by proposed method.

Knowing that the appendix is below fascia area; thus, we try to find fascia area first and remove it from our region of interest. In this part, the bottom fascia lines were carefully treated with cubic spline interpolation such that the lines were connected correctly. Many other image processing techniques such as Ends-in Search Stretching, Max-Min binarization, and region growing labeling are used to enhance the brightness contrast, remove noises, and connect pixels. Then the fuzzy ART learning is applied to classify pixels into the same objects of their labels such that the appendicitis could be extracted based on its morphological features.

Extracted appendix results were shown to medical experts and the proposed method shows a clear improvement from the previous study [10] in that our method is now successfully discriminate patient's ascites of a significant size case from appendix.

Developing automatic appendicitis diagnosis software to assist medical doctors is the final goal of our research and we believe that the proposed method in this paper can be an important building block of such effort.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] P. L. Allen, G. M. Baxter, and M. J. Weston, *Clinical Ultrasound*, Churchill Livingstone, 3rd edition, 2011.
- [2] S. Strandiring, N. R. Borley, P. Collins, A. R. Crossman, and M. A. Gatzoulis, *Gray's Anatomy: The Anatomical Basis of Clinical Practice*, Elsevier, 40th edition, 2008.
- [3] D. L. Longo, A. S. Fauci, D. L. Kasper, S. L. Hauser, J. Jameson, and J. Loscalzo, *Harrison's Principles of Internal Medicine*, McGraw-Hill, New York, NY, USA, 18th edition, 2012.
- [4] G. M. Doherty, *CURRENT Diagnosis & Treatment: Surgery*, McGraw-Hill, New York, NY, USA, 13th edition, 2010.
- [5] G. M. Israel, N. Malguria, S. McCarthy, J. Copel, and J. Weinreb, "MRI vs. ultrasound for suspected appendicitis during pregnancy," *Journal of Magnetic Resonance Imaging*, vol. 28, no. 2, pp. 428–433, 2008.

- [6] D. Gaitini, "Imaging acute appendicitis: state of the art," *Journal of Clinical Imaging Science*, vol. 1, no. 49, Article ID 85778, 2011.
- [7] K. M. Jang, K. Lee, M.-J. Kim et al., "What is the complementary role of ultrasound evaluation in the diagnosis of acute appendicitis after CT?" *European Journal of Radiology*, vol. 74, no. 1, pp. 71–76, 2010.
- [8] N. Kessler, C. Cyteval, B. Gailix et al., "Appendicitis: evaluation of sensitivity, specificity, and predictive values of US, doppler us, and laboratory findings," *Radiology*, vol. 230, no. 2, pp. 472–478, 2004.
- [9] M. Wider, Y. Myint, and E. Supriyanto, "Comparison of histogram thresholding methods for ultrasound appendix image extraction," *NAUN International Journal of Computers*, vol. 5, no. 11, pp. 542–549, 2011.
- [10] S. I. Park and K. B. Kim, "Extraction of appendix from ultrasonographic images with fuzzy binarization technique," *International Journal of Bio-Science and Bio-Technology*, vol. 5, no. 4, pp. 139–148, 2013.
- [11] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, no. 9, pp. 1277–1294, 1993.
- [12] K. B. Kim, D. H. Song, Y. H. Joo, H. J. Lee, and Y. W. Woo, "Extraction of muscle areas from ultrasonographic images using refined histogram stretching and fascia information," *Journal of Computational and Theoretical Nanoscience*, vol. 7, no. 5, pp. 921–926, 2010.
- [13] A. S. Ravi Kanth and Y. N. Reddy, "Cubic spline for a class of singular two-point boundary value problems," *Applied Mathematics and Computation*, vol. 170, no. 2, pp. 733–740, 2005.
- [14] K. B. Kim, H. J. Lee, D. H. Song, and Y. W. Woo, "Extracting fascia and analysis of muscles from ultrasound images with FCM-based quantization technology," *Neural Network World*, vol. 20, no. 3, pp. 405–416, 2010.
- [15] K. B. Kim and J. W. Kim, "Self health diagnosis system with Korean traditional medicine using fuzzy art and fuzzy inference rules," in *Intelligent Information and Database Systems*, vol. 7198 of *Lecture Notes in Computer Science*, pp. 326–335, Springer, 2012.
- [16] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, New York, NY, USA, 1995.
- [17] K. B. Kim, S. K. Je, and Y. J. Kim, "Enhanced RBF network by using ART2 algorithm and fuzzy control method," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 88, no. 6, pp. 1497–1501, 2005.

## Research Article

# Segmentation and Tracking of Lymphocytes Based on Modified Active Contour Models in Phase Contrast Microscopy Images

Yali Huang<sup>1,2,3</sup> and Zhiwen Liu<sup>1</sup>

<sup>1</sup>*Institute of Signal and Image Processing, School of Information and Electronics, Beijing Institute of Technology (BIT),  
5 South Zhongguancun Street, Haidian District, Beijing 100081, China*

<sup>2</sup>*College of Electronics and Information Engineering, Hebei University, Baoding, China*

<sup>3</sup>*Key Laboratory of Digital Medical Engineering of Hebei Province, Baoding, China*

Correspondence should be addressed to Zhiwen Liu; [zwliu@bit.edu.cn](mailto:zwliu@bit.edu.cn)

Received 8 August 2014; Accepted 3 January 2015

Academic Editor: Shahram Shirani

Copyright © 2015 Y. Huang and Z. Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The paper proposes an improved active contour model for segmenting and tracking accurate boundaries of the single lymphocyte in phase-contrast microscopic images. Active contour models have been widely used in object segmentation and tracking. However, current external-force-inspired methods are weak at handling low-contrast edges and suffer from initialization sensitivity. In order to segment low-contrast boundaries, we combine the region information of the object, extracted by morphology gray-scale reconstruction, and the edge information, extracted by the Laplacian of Gaussian filter, to obtain an improved feature map to compute the external force field for the evolution of active contours. To alleviate initial location sensitivity, we set the initial contour close to the real boundaries by performing morphological image processing. The proposed method was tested on live lymphocyte images acquired through the phase-contrast microscope from the blood samples of mice, and comparative experimental results showed the advantages of the proposed method in terms of the accuracy and the speed. Tracking experiments showed that the proposed method can accurately segment and track lymphocyte boundaries in microscopic images over time even in the presence of low-contrast edges, which will provide a good prerequisite for the quantitative analysis of lymphocyte morphology and motility.

## 1. Introduction

The study of cell morphology and motility in microscopic images is essential to understand and treat various biological processes [1, 2]. As is well known, lymphocytes are involved in immune response. Clinicians observe that lymphocytes are highly deformable objects in special conditions, especially the graft rejection occurring. Quantitative analysis of lymphocyte morphology and motility is very meaningful in immune response research. The segmentation and tracking of lymphocyte boundaries is one of the prerequisites for the quantitative analysis of cell morphology and motility [3, 4]. Manual segmentation is subjective, time-consuming, and prone to errors. Therefore, automatic segmentation and tracking of cells are desired, and many such methods have been proposed over the years.

Traditional methods for image segmentation, such as thresholding, region growing, and watershed, could generate

incorrect boundaries of objects since only local information is taken into consideration, while active contour models can segment, match, and track the object by exploiting (bottom-up) constraints derived from the image data together with (top-down) a priori knowledge about the location, the size, or the shape [5]. Hence they have been extensively studied and used in medical image segmentation ever since the introduction of active contour models in [6]. Zimmer et al. used parametric active contour models to segment and track migration cells in microscopic videos [7]. Acton's research group detected and tracked leukocytes based on the shape- and size-constrained active contour models [8, 9]. Meijering et al. applied the modified level set method to track cells with time-lapse fluorescence microscope [4]. Seroussi et al. used the directional gradient vector flow snakes to segment and track live cells in phase-contrast microscopic images [10].

Generally speaking, there are two types of active contours categorized by the representation and implementation:

the parametric and the geometric active contours. The former type usually establishes energy function composed of internal and external energy terms. The latter is represented implicitly as the zero crossings of level set function, which can tackle topological changes elegantly at the cost of higher computational complexity [11]. Segmentation methods based on the parametric active contour models are used in our study since the observed lymphocytes are free of topological changes.

A number of methods for tracking cells have been developed over the past decades [4, 12]. In general, they can be divided into two categories according to the tracking strategies. The first type, based on the “first detect, then track” principle, initially detects the object in the first frame and then establishes the link between the detected objects from frame to frame based on certain criteria [13]. The second category of algorithms based on the integrated segmentation and tracking scheme is often referred to as active contour models. In active contour models, segmentation and tracking are performed simultaneously by fitting the model to the image data, and the result of the contour evolution in the previous frame is used as the initial contour of the subsequent frame. The advantage of these algorithms is that all available information from the previous frame can be directly incorporated into the segmentation of the subsequent frame. So tracking is realized by segmenting the object from frame to frame. In this paper, we adopt the second category of algorithms (active contour models) to segment and track cell boundaries.

Active contour models confront two challenges. First, inaccurate segmentations may occur when the edges are low-contrast or noise-contaminated. Second, active contour models are usually sensitive to the initial position. To address these difficulties, three types of methods have been proposed: edge-based models (where the energy optimization is driven by boundary information of the image), region-based models (where the energy optimization is driven by region information of the image), and hybrid models. Many external forces were proposed in different applications, such as balloons force [14], gradient vector flow (GVF) [15], virtual electric field [16], and external force using vector field convolution (VFC) [17]. The VFC snakes, within which an external force is computed by convolving the edge map with a user-defined vector field kernel, are more robust to noise and have less computational costs, compared with the classic GVF snakes. However, the initialization flexibility is still restricted. The initial contour, which was shown to be of importance [18, 19], evolves according to the external force field, and the external force field is computed via the edge map. The idea of the proposed external force field in this paper is inspired by [19], which introduced a modified feature map based on Harris detector for VFC snakes.

We note that the aforementioned methods of segmentation and tracking have focused on tracking the cell instead of the accurate extraction of cell boundaries. In order to segment and track the accurate boundaries of lymphocytes in the image sequence, we propose a novel feature map based on morphological gray-scale reconstruction and the LoG filter (MGRL) to compute external force field for the evolution of active contours. We make use of the region information of

the given image by applying the morphological gray-scale reconstruction [20] and make use of the edge information through the LoG filter. To alleviate initialization sensitivity and reduce the number of iterations, the active contour is initialized close to the phase halo by an initial segmentation.

The rest of the paper is organized as follows. Section 2 reviews several kinds of active contour models and their external force fields. Section 3 proposes the improved external force based on MGRL-feature map, and the numerical implementation is given. The initialization of the lymphocyte contour is also introduced in this section. In Section 4, we apply active contour models with the proposed external force field to segment and track lymphocyte boundaries and compare experimental results with that of other methods. The conclusions are given in Section 5.

## 2. Related Work

Active contours are curves defined in an image domain that can move under the influence of internal forces and external forces [6]. Mathematically, the active contour is defined by a parametric contour  $C(s) = [x(s), y(s)]$ ,  $s \in [0, 1]$  and evolves within the spatial domain of an image to minimize the following energy function:

$$E_{\text{snakes}} = \int_0^1 \left[ \frac{1}{2} \left( \alpha |C^{(1)}(s)|^2 + \beta |C^{(2)}(s)|^2 \right) + E_{\text{ext}}(C(s)) \right] ds, \quad (1)$$

where superscript “(p)” denotes the  $p$ th order derivative and  $\alpha$  and  $\beta$  are weighting parameters representing the degrees of elasticity and rigidity of the active contour, respectively. The former makes the contour behave like an elastic string, while the latter makes the contour behave like a rigid rod [6]. The external energy  $E_{\text{ext}}$ , which represents the image constraints, is defined to move the active contour toward an object boundary or other desired features. Using the calculus of variations [21], an active contour that minimizes (1) must satisfy the Euler equation

$$\alpha \cdot C^{(2)}(s) - \beta \cdot C^{(4)}(s) - \nabla E_{\text{ext}}(C(s)) = 0. \quad (2)$$

The solution of (2) is obtained by calculating the steady state solution of the following gradient flow:

$$C_t(s, t) = \alpha \cdot C^{(2)}(s, t) - \beta \cdot C^{(4)}(s, t) - \nabla E_{\text{ext}}(C(s, t)). \quad (3)$$

Xu and Prince defined the gradient vector flow as the external force for the evolution of active contour [15]. They proposed the method by replacing  $-\nabla E_{\text{ext}}(C(s))$  in (2) with the vector field  $\mathbf{V}(x, y) = [V_1(x, y), V_2(x, y)]$ , which is computed by minimizing the function

$$E_{\text{ext}} = \iint \mu \left( |\nabla V_1|^2 + |\nabla V_2|^2 \right) + |\nabla f|^2 |V - \nabla f|^2 dx dy, \quad (4)$$

where  $\mu$  is a regularization parameter. One of the generally used forms of edge map  $f$  is

$$f(x, y) = |\nabla G_\sigma(x, y) * I(x, y)|, \quad (5)$$

where  $G_\sigma$  is the Gaussian function with the standard deviation (STD) $\sigma$  and  $\sigma = 3$  in the study.

Li and Acton proposed another external force field for the evolution of active contours [17], in which the external force  $-\nabla E_{\text{ext}}(C(s))$  in (2) is replaced with the external force field  $\mathbf{V}_{\text{vfc}}(x, y)$ :

$$\mathbf{V}_{\text{vfc}}(x, y) = f(x, y) * \mathbf{k}(x, y), \quad (6)$$

where  $\mathbf{k}(x, y)$  is the vector field kernel:  $\mathbf{k}(x, y) = (\sqrt{x^2 + y^2} + \varepsilon)^{-r} \cdot \mathbf{n}(x, y)$ ;  $\mathbf{n}(x, y)$  is the unit vector pointing to the kernel origin,  $\mathbf{n}(x, y) = [-x/\sqrt{x^2 + y^2}, -y/\sqrt{x^2 + y^2}]$ ;  $r$  is a positive parameter to control the decrease,  $r = 2$  in the study; and  $\varepsilon$  is a small positive number.

In active contour models, the external force, which is computed from the edge map, determines the evolution of the active contour. A good edge map (feature map) should emphasize the normal and the low-contrast edges equally. In order to reduce the effects of nonuniform edge intensities and highlight the low-contrast edges in phase contrast microscopic images, we propose a novel feature map instead of the traditional edge map to compute the external force, which is given as follows.

### 3. The Proposed Method

*3.1. The Improved External Force Field Based on MGRL-Feature Map.* The proposed feature map based on MGR and the LoG filter (MGRL-feature map) is

$$f_{\text{MGRL}}(x, y) = F_l \left[ \nabla^2 \left( G_\sigma(x, y) * \rho_{I(x, y)}(J(x, y)) \right) \right], \quad (7)$$

where  $F_l(\cdot)$  is a low pass filter, which is realized by removing all connected region that have fewer than  $P$  (15~20) pixels from a binary image;  $\nabla^2(\cdot)$  is the Laplacian operator;  $G_\sigma(x, y) * \rho_{I(x, y)}(J(x, y))$  is the convolution of the Gaussian function  $G_\sigma(x, y)$  and  $\rho_{I(x, y)}(J(x, y))$ .  $\rho_{I(x, y)}(J(x, y))$  denotes the MGR of the  $I(x, y)$  from the  $J(x, y)$ .  $I(x, y)$  is the region of interest (ROI). The MGR process can fill the ‘‘holes’’ induced by the intensity inhomogeneity in the image, which is defined as follows [20]. Let  $J$  and  $I$  be two gray-scale images defined on the same discrete domain  $D$ , where  $J(p), I(p) \in \{0, 1, \dots, N-1\}$  and  $J(p) \leq I(p)$  for each pixel  $p \in D$ . The MGR  $\rho_I(J)$  of  $I$  from  $J$  is given by  $\rho_I(J)(p) = \max\{k \in [0, N-1] \mid p \in \rho_{T_k(I)}(T_k(J))\}, \forall p \in D$ , where  $T_k(\cdot)$  is the successive thresholds; for  $k = 0, 1, \dots, N-1$ ,  $T_k(I) = \{p \in D \mid I(p) \geq k\}$ , these sets satisfy the following inclusion relationship  $T_k(I) \subseteq T_{k-1}(I), \forall k \in [1, N-1]$ .

We emphasize edges to look for zero crossings by filtering the MGR image  $\rho_I(J)$  with the LoG filter. The STD of the LoG filter is 2, and the size is 13-by-13. Since the LoG filter detects many fragmentary edges simultaneously, we use a low pass filter  $F_l(\cdot)$  to omit the high frequency edges.

The external force field based on MGRL-feature map is

$$\mathbf{V}_{\text{vfc\_MGRL}}(x, y) = f_{\text{MGRL}}(x, y) * \mathbf{k}(x, y). \quad (8)$$

In our study, phase contrast imaging is used for cell images acquisition in order to observe single lymphocyte

morphology and motility over a long period of time. There is only one target-lymphocyte in the center of the view in each microscopic sequence, and the video is recorded by the help of the clinicians. An example frame is shown as Figure 1(a), in which the ROI containing the object-lymphocyte is marked as in the rectangular by the user. Two different external force fields of the marked rectangular are shown in Figures 1(b) and 1(c), respectively, which indicate that the external force field computed from the proposed feature map is sparser than that from the traditional edge map. Since the high frequency components of the feature map are removed by a low pass filter, the external force vectors only appear at the edges, which can accelerate the evolution of the active contour.

*3.2. Numerical Implementation.* The VFC snakes based on the improved external force minimize the following energy function:

$$E = \int_0^1 \left[ \frac{1}{2} \left( \alpha |C^{(1)}(s)|^2 + \beta |C^{(2)}(s)|^2 \right) + E_{\text{vfc\_MGRL}}(C(s)) \right] ds, \quad (9)$$

where  $E_{\text{vfc\_MGRL}}$  denotes the improved external energy. Using the calculus of the variations [21], the minimization of (9) must satisfy the Euler equation

$$\alpha \cdot C^{(2)}(x, y) - \beta \cdot C^{(4)}(x, y) + \mathbf{V}_{\text{vfc\_MGRL}}(x, y) = 0, \quad (10)$$

where  $\mathbf{V}_{\text{vfc\_MGRL}}(x, y) = [u(x, y), v(x, y)]$  denotes the improved external force field derived from  $E_{\text{vfc\_MGRL}}$ . Equation (10) is equivalent to the following expression:

$$\begin{aligned} \alpha \frac{d^2 x}{ds^2} - \beta \frac{d^4 x}{ds^4} + u(x, y) &= 0, \\ \alpha \frac{d^2 y}{ds^2} - \beta \frac{d^4 y}{ds^4} + v(x, y) &= 0. \end{aligned} \quad (11)$$

The solution of (11) is obtained by calculating the following gradient descent equation:

$$\begin{aligned} x_t &= \alpha x^{(2)} - \beta x^{(4)} + u(x, y), \\ y_t &= \alpha y^{(2)} - \beta y^{(4)} + v(x, y). \end{aligned} \quad (12)$$

Using a finite difference approach on a discrete grid, our iterative solution to (12) is as follows [6]:

$$\begin{aligned} x_i^{n+1} &= x_i^n + \gamma \left[ \alpha \cdot \frac{d^2(x_i^n)}{ds^2} - \beta \cdot \frac{d^4(x_i^n)}{ds^4} + u(x_i, y_i) \right], \\ y_i^{n+1} &= y_i^n + \gamma \left[ \alpha \cdot \frac{d^2(y_i^n)}{ds^2} - \beta \cdot \frac{d^4(y_i^n)}{ds^4} + v(x_i, y_i) \right], \end{aligned} \quad (13)$$

where  $n$  corresponds to discrete time and  $\gamma$  denotes the time step for each interaction.  $(x_i, y_i)$  ( $i = 1, 2, \dots, N$ ) is the discrete approximation of  $(x(s_i), y(s_i))$  by discretizing the interval  $[0, 1]$  into  $N-1$  equispaced subintervals of length  $h = 1/(N-1)$ . Replace the derivative by the difference; then we get the second-order derivative of  $x(s)$ ,  $d^2 x(s)/ds^2 \approx (x_{i+1} - 2x_i + x_{i-1})/h^2$ . The higher-order derivative can be obtained in this way.

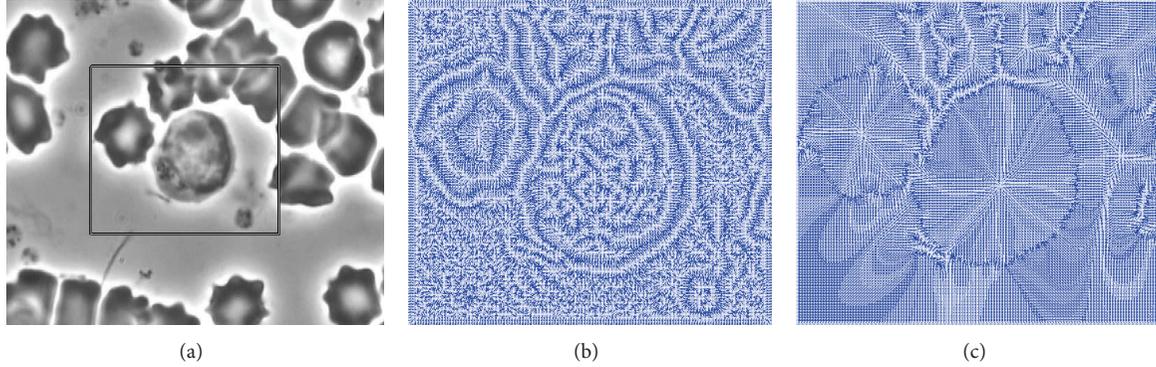


FIGURE 1: The original image and two different external force fields. (a) Cell image captured by the phase-contrast microscope, and the lymphocyte is in the center of the view. The ROI contains the target-lymphocyte, marked as in the rectangular. (b) The traditional external force field based on (5). (c) The proposed external force field based on (7).

**3.3. Initial Contour.** The active contour models may converge to an incorrect boundary if the initial contour is far from the real boundary. To alleviate the initialization sensitivity, we initialize the contour close to the real boundary by an initial segmentation. In the study, the initial contour of the object is determined by the following steps.

- (1) Select a rectangle ROI within the image by the user, as shown in Figure 2(a).
- (2) Apply MGR to the ROI, as shown in Figure 2(b). The result shows that the lymphocyte region is distinct from other red, white cells and the background.
- (3) Obtain the binary image from Figure 2(b) through thresholding techniques. The appropriate threshold is selected based on the intensity histogram distributions. The assumptions are as follows: there is only one target-lymphocyte in each image/frame, and there is no overlap occurrence between the target-lymphocyte and other cells. So the largest intensity distribution in the intensity histogram is selected as the optimal intensity threshold. Suppose the optimal threshold is  $\text{thr}$ ; the image binarization processing is defined as (14). The binarized result is shown in Figure 2(c):

$$B(x, y) = \begin{cases} 1, & I(x, y) = \text{thr} \\ 0, & \text{others.} \end{cases} \quad (14)$$

- (4) After the binarization, morphological operations, including open and close, are applied to obtain the initial segmentation. The flat disk-shaped structuring element is used, and the specified radius is 3. And then we obtain the coarse binary image of the object. The initial contour, which is close to the real edge, is extracted from the binary image, as shown in Figure 2(d).

## 4. Experiments and Results

In the study, the image sequence of live cells were obtained by an optical phase-contrast microscope at a magnification of 16,000 from blood samples which were collected from the tails of mice (6–8 weeks old, 20–22 grams heavy). Note that in the experiments, there is only one target-lymphocyte in each image, which is at the center of the view and separate from other red and white cells.

**4.1. Segmentation of Accurate Lymphocyte Boundaries.** We randomly selected 25 phase-contrast microscopic images. The result of manual segmentation is used as the ground truth in many studies although they have limitations. Therefore the average of the manual segmentation results by three experts is used as the ground truth in our study. To validate the performance the proposed method, we compared segmentation results with the other three types of widely used active contours: GVF snakes, edge-based geodesic active contour (GAC), region-based Chan, and Vese active contour model (CV). In the validation of GVF snakes, the active contour evolution equation is similar to (13), except that it uses gradient vector field as the external force. The evolution of GAC is  $\partial\phi/\partial t = g \cdot (k+c)|\nabla\phi| + \nabla g \cdot \nabla\phi$ , where  $c$  is the balloon force,  $k$  is the curvature, and  $g(\cdot)$  is the edge function. The evolution of CV is  $\partial\phi/\partial t = \delta(\phi)[\mu_1 \cdot k - \nu - \lambda_1(I - c_1)^2 + \lambda_2(I - c_2)^2]$ , where  $c_1$  and  $c_2$  are two constants which are the average intensities inside and outside the contour, respectively;  $\mu_1 \geq 0$ ,  $\nu \geq 0$ ,  $\lambda_1 > 0$ ,  $\lambda_2 > 0$  are fixed parameters. The detailed parameters settings of the four methods (GVF, GAC, CV, and the improved VFC) are listed in Table 1. Figure 3 shows the comparison among GAC, GVF, CV, the proposed method, and the manual segmentation results. The initial contours of the active contour models were obtained by the initial segmentation as mentioned above (see Section 3.3). The iteration numbers and the execution time are shown in Table 1. The experiments were conducted on a 2.93 GHz CPU, 4.00 G RAM computer. As we can see, the results of GVF and GAC suffered from the edge leaks outside of cells due to the

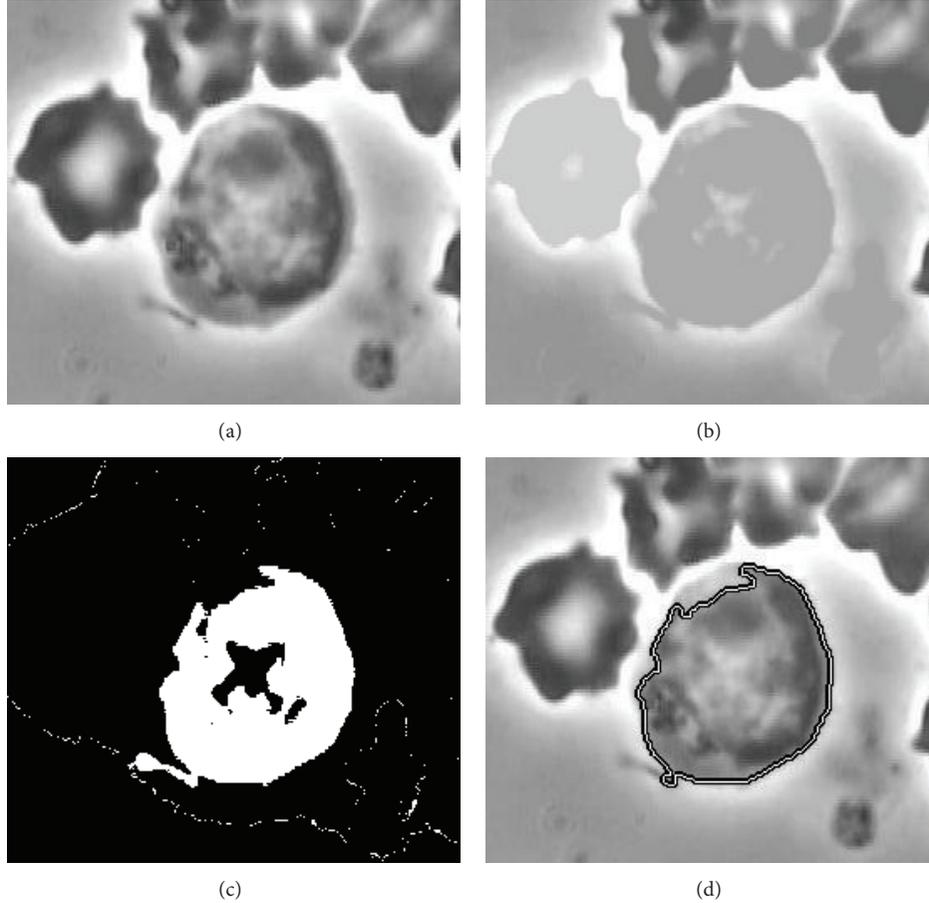


FIGURE 2: The procedure of obtaining the initial contour. (a) The ROI. (b) MGR of the ROI. (c) The coarse binary image of (b). (d) The initial contour extracted from the binary image.

low-contrast boundaries. The CV method failed to segment cell images with intensity variation inside the cell. The proposed method VFC\_MGRL can find the accurate boundaries at low-contrast edges due to the improved external force field. The segmentation result of the proposed method is close to the expert's manual segmentation result, even better in the details. For manual segmentation by the expert, it is hard to segment the cellular protrusion consistently, but the proposed method can converge to the details of the edges.

To quantitatively evaluate the segmentation, the Jaccard coefficient (JC) is employed to measure the similarity between the segmentation result and the ground truth, which is defined as the ratio between the size of the intersection of the sets and the size of their union; namely,  $JC = (A \cap B) / (A \cup B)$ , where  $A$  denotes the segmentation result and  $B$  is the ground truth. Figure 4 shows the detailed evaluation results for the 25 images separately. Parameters settings are shown in Table 1. It is important to note that the proposed VFC\_MGRL snakes method outperform their traditional counterparts in most cases.

To compare the performance of different segmentation methods, the statistical analysis of segmentation results by

different methods are shown in Table 2, in which STD means standard deviation.

*4.2. Tracking of Accurate Lymphocyte Boundaries.* Active contour models allow us to solve both segmentation and tracking problems simultaneously. The concept is that lymphocyte tracking is realized by lymphocyte segmentation frame by frame. The initial contour of the lymphocyte in the first frame is obtained by an initial segmentation, and for the second frame, the initial contour is obtained from the result of the contour evolution in the previous frame. That is to say, the final contour of the previous frame is regarded as the initial contour of the current frame during the evolution of active contours. The procedure of segmentation and tracking is described in Figure 5(a).

*Step 1.* Read the first frame of the video, and then choose the ROI containing the target-lymphocyte by the user, as shown in Figure 1(a).

*Step 2.* Extract the initial contour close to the ground truth. In the first frame of the video, this is realized by MGR and

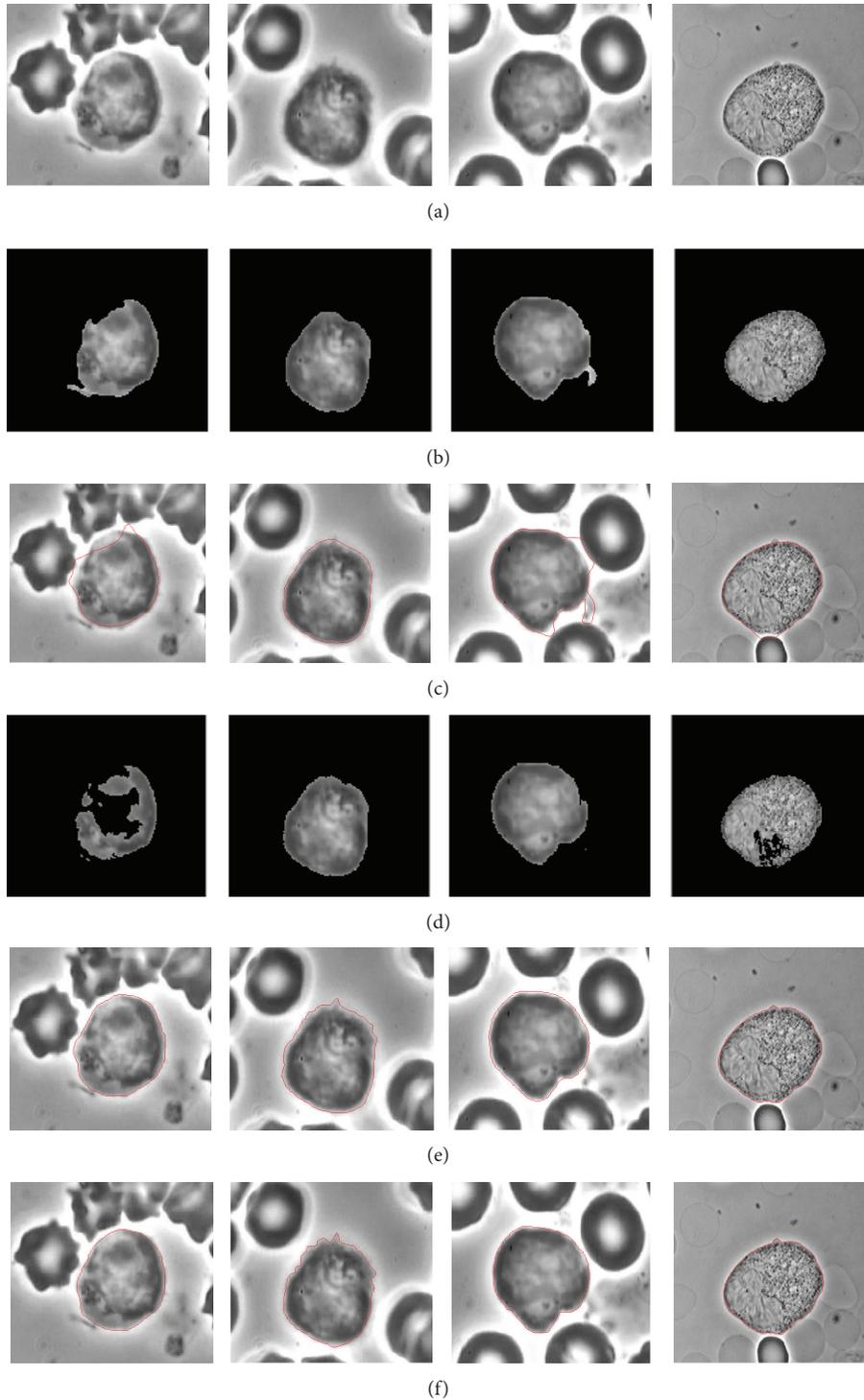


FIGURE 3: Segmentation of sample lymphocytes. The first line: the ROI (Img 1, Img 2, Img 3, and Img 4). The second, third, fourth, and fifth lines present the results for GAC, GVE, CV, and the proposed VFC\_MGRL method. The last line is the ground truth.

thresholding techniques (as introduced in Section 3.3); in the subsequent frames, the initial contour is obtained from the previous frame since the vibration of the lymphocyte position is not distinct between the successive frames.

*Step 3.* Compute the improved external force field of the ROI, and then implement the evolution of active contours according to the initial contour. The result of the evolution of the active contour is the final boundary of the lymphocyte.

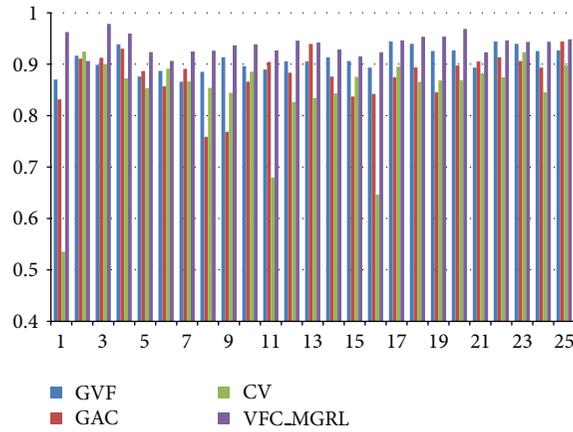


FIGURE 4: Detailed evaluation results. Horizontal axis shows the numbered images used for evaluation. Separate bars indicate the results of different methods: blue (the first bar) is GVF; red (the second bar) is GAC; green (the third bar) is CV; purple (the fourth bar) is the proposed VFC\_MGRL method.

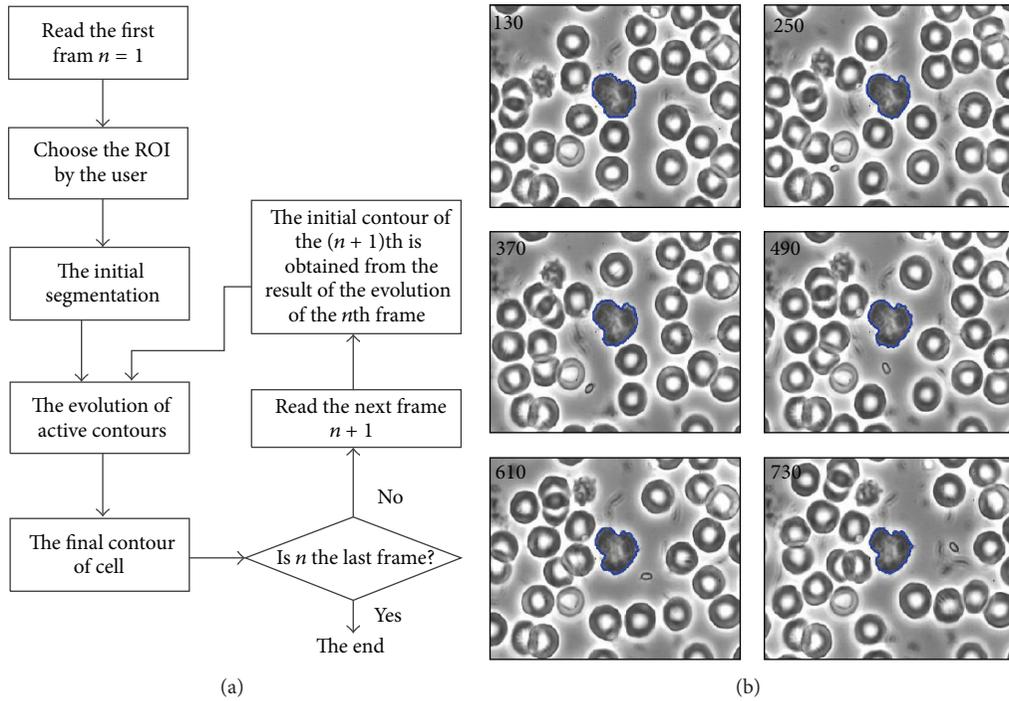


FIGURE 5: (a) The workflow of segmentation and tracking of the target-cell in an image sequence. (b) Segmentation and tracking of cell by using VFC\_MGRL snakes at frame 130, 250, 370, 490, 610, and 730, respectively. The frame interval is 0.04 seconds.

TABLE 1: Parameters settings and execution times for images 1–4 of the four kinds of active contour methods.

	Parameters settings	Number of iterations	Run time (seconds)			
			Image 1	Image 2	Image 3	Image 4
GVF	$r = 1, \alpha = 0.8, \beta = 0.2, \mu = 0.1$	100	2.26	2.25	2.39	2.17
GAC	$dt = 0.1, c = 1$	200	4.08	3.99	4.16	4.05
CV	$dt = 0.1, \mu_1 = 0.2, \nu = 0, \lambda_1 = \lambda_2 = 1$	200	2.92	2.63	2.63	2.72
VFC_MGRL	$r = 1, \alpha = 2, \beta = 0.2, \text{log filter : sigma} = 2, \text{size} = 13$	100	1.71	1.78	2.05	1.91

TABLE 2: The mean and standard deviation of JC by different segmentation methods.

	GVF	GAC	CV	VFC_MGRL
Mean	0.9091	0.8787	0.8422	0.9388
STD	0.0235	0.0460	0.0898	0.0185

*Step 4.* If the current frame is the last frame, the lymphocyte tracking procedure is over; if not, go to the next frame of the video repeating Step 3 till the last frame.

The accurate boundary of the lymphocyte is then obtained from frame to frame by VFC\_MGRL snakes.

One result of lymphocyte tracking was shown in Figure 5(b), which shows that the proposed algorithm can accurately track the target-lymphocyte boundaries and follow the dynamic change of lymphocyte shape over time in a semiautomated fashion. In the first frame, the ROI was chosen by the user; in the subsequent frames, the object boundaries were segmented and tracked automatically.

## 5. Conclusions

The paper proposes a VFC\_MGRL active contour model for segmenting and tracking of accurate boundaries of the single lymphocyte in phase contrast microscopic images. In a video, starting from the initial contour, the active contour converges to the accurate boundary according to the improved external force field. The MGRL-feature map can make full use of the given image by incorporating the advantages of MGR and the LoG filter. Therefore, it can be used for defining an efficient external force field when detecting low-contrast boundaries. To alleviate initialization sensitivity problem, the initial contour of the first frame is abstracted around the ground truth by an initial segmentation; the initial contour of the subsequent frame is obtained from the previous frame. The approach is tested on phase contrast microscopic images and performs better than other methods, which will provide a good prerequisite for the quantitative analysis of lymphocyte morphology and motility.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (61271112) and was also supported by the Midwest Universities Comprehensive Strength Promotion Project. The authors would like to thank the president of Beijing You'an Hospital, Ning Li, for providing the microscopic cell images.

## References

- [1] J. R. Swedlow, "Advanced hardware and software tools for fast multidimensional imaging of living cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 37, pp. 16005–16006, 2010.
- [2] S. Chen, M. Zhao, G. Wu, C. Yao, and J. Zhang, "Recent advances in morphological cell image analysis," *Computational and Mathematical Methods in Medicine*, vol. 2012, Article ID 101536, 10 pages, 2012.
- [3] X. An, Z. Liu, Y. Shi, N. Li, Y. Wang, and S. Joshi, "Modeling dynamic cellular morphology in images," in *Proceedings of the 15th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI '12)*, pp. 340–347, 2012.
- [4] E. Meijering, O. Dzyubachyk, I. Smal, and W. A. van Cappellen, "Tracking in cell and developmental biology," *Seminars in Cell and Developmental Biology*, vol. 20, no. 8, pp. 894–902, 2009.
- [5] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: a survey," *Medical Image Analysis*, vol. 1, no. 2, pp. 91–108, 1996.
- [6] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [7] C. Zimmer, E. Labruyère, V. Meas-Yedid, N. Guillén, and J.-C. Olivo-Marin, "Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: a tool for cell-based drug testing," *IEEE Transactions on Medical Imaging*, vol. 21, no. 10, pp. 1212–1221, 2002.
- [8] N. Ray, S. T. Acton, and K. Ley, "Tracking leukocytes in vivo with shape and size constrained active contours," *IEEE Transactions on Medical Imaging*, vol. 21, no. 10, pp. 1222–1235, 2002.
- [9] D. P. Mukherjee, N. Ray, and S. T. Acton, "Level set analysis for leukocyte detection and tracking," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 562–572, 2004.
- [10] I. Seroussi, D. Veikherman, N. Ofer, S. Yehudai-Resheff, and K. Keren, "Segmentation and tracking of live cells in phase-contrast images using directional gradient vector flow for snakes," *Journal of Microscopy*, vol. 247, no. 2, pp. 137–146, 2012.
- [11] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
- [12] O. Dzyubachyk, W. A. van Cappellen, J. Essers, W. J. Niessen, and E. Meijering, "Advanced level-set-based cell tracking in time-lapse fluorescence microscopy," *IEEE Transactions on Medical Imaging*, vol. 29, no. 3, pp. 852–867, 2010.
- [13] F. P. Ferrie, M. D. Levine, and S. W. Zucker, "Cell tracking: a modeling and minimization approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 4, no. 3, pp. 277–291, 1982.
- [14] L. D. Cohen, "On active contour models and balloons," *CVGIP: Image Understanding*, vol. 53, no. 2, pp. 211–218, 1991.
- [15] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359–369, 1998.
- [16] H. K. Park and M. J. Chung, "External force of snake: virtual electric field," *Electronics Letters*, vol. 38, no. 24, pp. 1500–1502, 2002.
- [17] B. Li and S. T. Acton, "Active contour external force using vector field convolution for image segmentation," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2096–2106, 2007.

- [18] C. Li, J. Liu, and M. D. Fox, "Segmentation of external force field for automatic initialization and splitting of snakes," *Pattern Recognition*, vol. 38, no. 11, pp. 1947–1960, 2005.
- [19] A. Kovacs and T. Sziranyi, "Harris function based active contour external force for image segmentation," *Pattern Recognition Letters*, vol. 33, no. 9, pp. 1180–1187, 2012.
- [20] L. Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *IEEE Transactions on Image Processing*, vol. 2, no. 2, pp. 176–201, 1993.
- [21] D. Lao, *Fundamentals of the Calculus of Variations*, National Defence Industry Press, Beijing, China, 2007.

## Research Article

# Fractional Diffusion Based Modelling and Prediction of Human Brain Response to External Stimuli

Hamidreza Namazi<sup>1</sup> and Vladimir V. Kulish<sup>2</sup>

<sup>1</sup>Department of Mechanical Engineering, Faculty of Engineering, Universiti Malaysia Sarawak, 94300 Kuching, Sarawak, Malaysia

<sup>2</sup>School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798

Correspondence should be addressed to Hamidreza Namazi; [nhamidreza@feng.unimas.my](mailto:nhamidreza@feng.unimas.my)

Received 2 November 2014; Accepted 17 January 2015

Academic Editor: Shahram Shirani

Copyright © 2015 H. Namazi and V. V. Kulish. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Human brain response is the result of the overall ability of the brain in analyzing different internal and external stimuli and thus making the proper decisions. During the last decades scientists have discovered more about this phenomenon and proposed some models based on computational, biological, or neuropsychological methods. Despite some advances in studies related to this area of the brain research, there were fewer efforts which have been done on the mathematical modeling of the human brain response to external stimuli. This research is devoted to the modeling and prediction of the human EEG signal, as an alert state of overall human brain activity monitoring, upon receiving external stimuli, based on fractional diffusion equations. The results of this modeling show very good agreement with the real human EEG signal and thus this model can be used for many types of applications such as prediction of seizure onset in patient with epilepsy.

## 1. Introduction

Brain as the most complex organ in the human body controls all bodies' actions/reactions by receiving different stimuli through the nervous system. Any stimulus stronger than the threshold stimulus is translated by the number of sensory neurons generating information about the stimulus and the frequency of the action potentials. After the action potential has been generated, it travels through the neural network to the brain. In various sections of the network and the brain, integration of the signals takes place. Different areas of the brain respond depending on the kind and location of stimuli. The brain sends out signals which generate the response mechanism.

During many years, numerous studies related to the brain response to external stimuli have been reported by scientists. Some researchers studied the brain response to different kinds of stimuli without proposing any model. In case of visual stimuli, we can mention the work done by Kaneoke et al. in analyzing the effect of the visual stimulus size on the human brain response using magnetoencephalography (MEG) [1]; see also [2, 3]. Other groups of researchers

investigated the effect of auditory stimuli on the brain response. For instance, Will and Berg studied and compared the brain responses to periodic stimulations, silence, and random noise using electroencephalography (EEG) [4]; see also [5, 6]. Olfactory stimuli also were the main focus of some researchers. Sutani et al. investigated the brain response to pleasant and unpleasant olfactory stimuli using MEG signals. They found out that the MEG signals have recorded from frontal/prefrontal cortical areas of the brain has some differences in case of pleasant versus unpleasant stimuli [7]; see also [8, 9]. Different works have been reported on the investigation of the brain response to other kinds of stimulus such as emotional stimuli [10, 11] and pain stimuli [12, 13].

On the other hand, some scientists proposed some models of the human brain activity. On the microscopic level, the work done by Freeman in the modeling of the EEG arising from the olfactory bulb of animals during the perception of odors is noteworthy. He developed a set of nonlinear equations for this response which generates EEG like pattern [14, 15]. In another work Seetharaman et al. proposed a mathematical model for generation and propagation of action potential in a node of Ranvier and they called it as

the phase-lagging model of single action potential [16]. When the microscopic models are extended to a macrolevel, then different methods are employed. Many of these models assume the cortical region to be a continuum. Liley et al. developed a set of nonlinear continuum field equations which described the macroscopic dynamics of neural activity in the cortical region [17]. These equations were used by Steyn-Ross et al. who introduced noise terms into them to give a set of stochastic partial differential equations (SPDEs). They also converted the equations governed by Liley et al. into linearized ODEs. This model could predict the substantial increase in low frequency power at the critical points of induction and emergence. They later used this model to study the electrical activity of an anaesthetized cortex [18–21]. Kramer et al. started with the equations given by Steyn-Ross and coworkers and neglected the spatial variation and the stochastic input. They believed that this gave rise to a set of ordinary differential equations (ODEs) for the modeling of the cortical activity. They showed that the results obtained from the SPDE model agree with clinical data in an approximate way [22], but they also stated that the spatial sampling of the cortex was poor because of inherent shortcomings in the equipment used. Kulish and Chan have suggested a novel method for the modeling of the brain response using fundamental laws of nature like energy conservation and the least action principle. The model equation obtained has been solved and the results show a good agreement with real EEGs [23].

Despite rapid advances in the studies related to the analysis of the human brain response, there has been less progress in the mathematical modeling of the human brain activity due to external stimuli. Yet, it seems that the contemporary level of developments in physics and mathematics makes establishing quantitative correlations between external physical stimuli and the brain responses to those stimuli.

This paper attempts to introduce a new mathematical model of the human brain response to external stimuli based on the fractional diffusion equation. At first we talk about the macroscopic level of brain organization and EEG signal as an alert state of human behavior monitoring. Then, by introducing fractional diffusion equations and considering the EEG signal as a fractal time series we model the EEG signal using mathematical equations. This model is then solved and after discussing different parameters in the model we provide some results and discuss about the model's solution in details. Some concluding remarks are provided at the end.

## 2. Macroscopic Level of Brain Organization and the EEG Signal

In order to study the human neural activity, one can consider different levels of brain organization at different scales in time and space from a single neuron (microscopic level) to the whole brain organization (macroscopic level). In fact, the whole brain activity cannot be observed by measuring the activity of just a single neuron as far it is informative as it contributes to study the entire population of which it is a member.

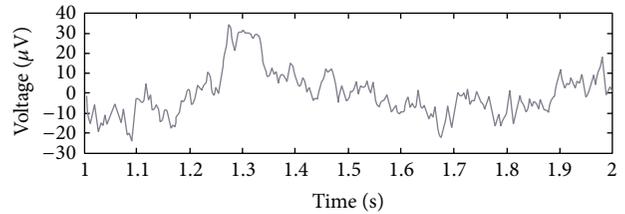


FIGURE 1: EEG signal of a subject in response to an external stimulus.

In this research, we focus on the macroscopic level of brain organization. Macroscopic level of brain organization refers to the level of neural assemblies' population in which each neural assembly interacts with other neural assemblies in close and distant cortical areas, exhibits spatial-temporal behaviour, and paints the human behaviour [24]. The functional behaviour of the brain is encoded in these spatial-temporal structures and can be extracted from the macroscopic quantities dynamics observed by EEG signal mostly [25].

During many years scientists have studied the human behavior by recording and analysis of the EEG signals from different areas of the brain. The EEG signal is the composition of different frequency bands (oscillatory activities-Alpha, Beta, ...) which are structured coordinately (spatially, temporally). In fact, this signal has different characteristics that can be used in order to study the human brain response to external stimuli in our research. For instance, Figure 1 shows the EEG signal of a subject who received an external stimulus at  $t = 1$  s. In this figure the brain response to external stimulus can be seen as a sudden upward deflection at about  $t = 1.25$  s after the application of stimulus at  $t = 1$  s.

## 3. Fractional Diffusion Equation

Here, in order to develop our model we start with a simple equation:

$$\frac{\partial F}{\partial t} = D \nabla^2 F, \quad (1)$$

where

$$\nabla^2 F = \frac{\partial^2 F}{\partial x^2}. \quad (2)$$

Equation (1) is the well-known diffusion equation where the coefficient  $D$  is the diffusivity of the medium for the property  $F$ . In fact, this equation arises in many descriptions of biological and physical phenomena, including Brownian motion [26], gradient driven chemical diffusion (with Fick's law), and heat transfer (heat diffusion with Fourier's Law). It is noteworthy that (1) is written in the Euclidean space.

On the other hand, the diffusion process can also be studied for fractals in fractal space. Fractals (such as random walk) may be defined as self-similar geometric objects whose scaling exponent (dimension) satisfies the Szpilrajn inequality:

$$\aleph \geq D_T, \quad (3)$$

where  $\aleph$  is the scaling exponent (dimension) of the object and  $D_T$  is its topological dimension, that is, Euclidean dimension of units from which the fractal object is built. In fact, Fractal and Euclidean geometries are conjugate approaches to the geometry of natural forms. Fractal geometry builds complex objects by applying simple processes to complex building blocks; Euclidean geometry uses simpler building blocks but frequently requires complex building processes.

Considering the diffusion process in the case of fractals, (1) is changed to (4) which is called the fractional diffusion equation [27]:

$$\frac{\partial^{2H} F}{\partial t^{2H}} = C^{2(2H-1)} D^{2(1-H)} \nabla^2 F, \quad (4)$$

where  $\nabla^2 F = \partial^2 F / \partial x^2$ .

In (4),  $C$  is the speed of propagation and  $H$  is the Hurst exponent with the value within the range  $0 \leq H \leq 1$  that brings predictability of signal into account. In fact, the Hurst exponent can be viewed as the probability of the diffusion process being persistent in a certain given direction. Note that the case  $H = 1/2$ , which corresponds to a nonfractal diffusion process, leads to the well-known classical equation. Observe also that if  $H = 0$ , (4) degenerates into the Poisson equation,  $\partial^2 F / \partial t^2 = C^{-2} D^2 \nabla^2 F$ ; that is, there is no preferred direction of random walk in this case, while the case  $H = 1$  leads to the wave equation,  $\partial^2 F / \partial t^2 = C^2 \nabla^2 F$ .

It is now necessary to make a very important remark. It is possible to consider a Brownian motion type process as a process which takes place in an Euclidean space (see the right hand side of (4)), considering the temporal dimension,  $t$ , of the process as fractal time which, for the same diffusion coefficient,  $D$ , either slows down (in case of  $H < 1/2$ ) or speeds up (in case of  $H > 1/2$ ) the process in question. This can be described with the generalized diffusion equation (4) in which the time coordinate appears as a fractal quantity. The generalized diffusion equation is a fractional PDE of order  $2H$  with respect to time.

Taking all these conjectures into account, (4) is a generalized form for describing the diffusion process when the time becomes fractal.

#### 4. Fractional Diffusion Model of EEG Signal

In this section, it is aimed to show that EEG signal as a fractal time series (but not stochastic), which represents a transient record of a random walk process, can be modeled by the solution of the fractional diffusion equation:

$$\frac{\partial^{2H} V}{\partial t^{2H}} = C^{2(2H-1)} D_{\text{eff}}^{2(1-H)} \frac{\partial^2 V}{\partial \eta^2}, \quad (5)$$

where the term  $V$ , which stands for the brain response, is the voltage fluctuations resulting from ionic current flows within the neurons of the brain. The impulse propagation speed within the neural network, which is a finite quantity, is represented by  $C$ . The term  $D_{\text{eff}}$ , the effective diffusion coefficient, as the property of neural tissue, is related to neuron's resistance to the electrical impulse as it travels over the nerve.



FIGURE 2: Upward and downward deflection in the signal from one point to the next point.

The term  $H$ , as a time variable parameter, corresponds to the Hurst exponent that brings the predictability of signal into account.

The direction of deflection at each moment in the signal can be studied by computing the Hurst exponent (Figure 2). The Hurst exponent is an indicator of the long term memory of the process generating the signal and thus it is the measure of the predictability of the signal.

As it was mentioned before the Hurst exponent can have any value between 0 and 1, where the value that it gains in each moment determines the behavior of the next deflection in the signal.

Firstly, if the Hurst exponent has a value between 0 and 0.5, it means that the process is antipersistent; that is, the trend of the value of the process at the next instant will be opposite to the trend in the previous instant. Secondly, a value of  $H$  between 0.5 and 1 means that the process is persistent; that is, the trend of the value of the process at the next instant will be the same as the trend in the previous instant. Finally, if  $H = 0.5$ , the process is considered to be truly random (e.g., Brownian motion). It means that there is absolutely no correlation between any values of the process.

One of the interesting points about the fractional diffusion model (5) is that it accounts for a finite time lag (reaction time) between any given disturbance (stimulus) and the brain response to it (human action/reaction) based on the assumption that no instantaneous propagation of information is possible within the brain. This effect is considered during the derivation of this model, but it is substituted by  $C$ ,  $D$  according to the equation:

$$\tau = \frac{D_{\text{eff}}}{C^2}. \quad (6)$$

Equation (6) is the formula for relaxation time in hyperbolic reaction diffusion equations [28].

In this research we consider the brain in an informational space where  $\eta$  is the spatial dimension of this space. In fact, each two neurons in informational space, which have informational interlink, make an informational channel between themselves and the length of this channel is called the informational distance  $\eta$  where

$$0 \leq \eta < \infty. \quad (7)$$

It is noteworthy that informational distance is not same as the concept of spatial distance. Two neurons, which are close together, may not exchange information which means  $\eta \sim \infty$ . On the other hand, two neurons which are far apart may be closely interlinked to each other and exchange a lot of information which makes  $\eta$  very small. Thus,  $\eta$  has value in the range  $0 \leq \eta \leq \infty$  [29].

The exchange of information between two neurons in the brain happens when there is a potential difference along the channel. In fact, the potential difference between two neurons causes an information flux and the exchange of information from the neuron with higher informational potential value to the one with the lower value continues till the gradient becomes zero.

The fractional diffusion equation is valid for the time scale:

$$t \geq 0. \quad (8)$$

Also, regarding (5),

$$V(\eta, 0) = V_0 \quad (9)$$

is the initial condition at  $t = 0$ . Also, the solution of the model is required to be bounded as  $\eta \rightarrow \infty$ . Otherwise, the conservation of energy principle would be violated. In other words,

$$\lim_{\eta \rightarrow +\infty} V(\eta, t) = \text{const} < \infty \quad (10)$$

is the boundary condition in the case of infinite domain.

In order to solve the above fractional diffusion model the method proposed by Oldham and Spanier is employed here [30].

Upon introducing the excess value  $\widehat{V}(\eta, t) = V(\eta, t) - V_0$ , so the initial condition imposed on  $\widehat{V}$  is  $\widehat{V}(\eta, 0) = 0$ .

In order to solve (5) we apply the Laplace transform with respect to time,  $t$ ; then we have

$$C^{2(2H-1)} D_{\text{eff}}^{2(1-H)} \frac{\partial^2 Y}{\partial \eta^2} - s^{2H} Y = 0, \quad (11)$$

where  $s$  is the Laplace transform variable and  $Y$  denotes the Laplace transform of the excess value  $\widehat{V}$ .

Equation (11) is a second-order ordinary differential equation (ODE), where the general solution is calculated as

$$Y(\eta; s) = A_1(s) e^{-\eta s^H / [C^{2(2H-1)} D_{\text{eff}}^{(1-H)}]} + A_2(s) e^{\eta s^H / [C^{2(2H-1)} D_{\text{eff}}^{(1-H)}]}, \quad (12)$$

where  $A_1(s)$  and  $A_2(s)$  are two arbitrary functions. However, since the solution is to be bounded for all  $\eta$ , the second arbitrary function,  $A_2(s)$ , must be identically zero, so the solution (12) is changed to

$$Y(\eta; s) = A_1(s) e^{-\eta s^H / [C^{2(2H-1)} D_{\text{eff}}^{(1-H)}]}. \quad (13)$$

Upon differentiating (13) with respect to  $\eta$ ,

$$\frac{dY(\eta; s)}{d\eta} = -\frac{A_1(s) s^H e^{-\eta s^H / [C^{2(2H-1)} D_{\text{eff}}^{(1-H)}]}}{[C^{2(2H-1)} D_{\text{eff}}^{(1-H)}]}. \quad (14)$$

After comparing (13) and (14),  $A_1(s)$  can be eliminated; then it can be written as

$$Y(\eta; s) = -s^{-H} C^{2(2H-1)} D_{\text{eff}}^{(1-H)} \frac{dY(\eta; s)}{d\eta}. \quad (15)$$

By taking the inverse Laplace transform of (15) and restoring the original variables, then we have

$$V(\eta, t) = V_0 - C^{2(2H-1)} D_{\text{eff}}^{(1-H)} \frac{\partial^{-H}}{\partial t^{-H}} \left( \frac{\partial V}{\partial \eta} \right) \quad (16)$$

which is written in terms of fractional derivative of order  $-H$  with respect to  $t$ .

Using the definition of fractional derivative [30], namely,

$$\frac{\partial^\alpha f}{\partial t^\alpha} = \frac{1}{\Gamma(-\alpha)} \int_0^t \frac{f(\xi) d\xi}{(t-\xi)^{\alpha+1}}, \quad \text{Re}(\alpha) < 0, \quad (17)$$

where  $\Gamma(\alpha)$  is the Gamma function, and noticing the Fick's law,

$$-\frac{\partial V}{\partial \eta} = \frac{\varphi}{D_{\text{eff}}}, \quad (18)$$

where  $\varphi$  represents the flux, and (16) can be written as

$$V(\eta, t) = V_0 + C^{2(2H-1)} D_{\text{eff}}^{(-H)} \frac{1}{\Gamma(H)} \int_0^t \frac{\varphi(\eta, \xi) d\xi}{(t-\xi)^{1-H}}. \quad (19)$$

Since the function  $\varphi(\eta, t)$  represents the flux it can be equated with the external influence acting on the system (external stimulus).

Equation (19) provides the relationship between the nonequilibrium value,  $V(\eta, t)$ , and the external influence acting on the system,  $\varphi(\eta, \xi)$ . This equation is valid for every location within the domain (including the boundary) at every moment.

Since  $H$  is a nonnegative parameter, it follows from (19) that the value of  $V$ , on the average, increases with the time according to the power law as  $t^H$ , provided of course that the fluctuations are small in comparison with the averaged influence. Note that, for  $H = 1/2$ ,  $\Gamma(1/2) = \pi^{1/2}$  and (19) yields the well-known diffusion (random walk) growth given by  $t^{1/2}$ .

An external influence can be modeled by a Gaussian pulse; that is,

$$\varphi(x, t) = \varphi_0(x, t) \exp\left[-\frac{(t-t^*)^2}{\sigma^2}\right], \quad (20)$$

where  $t^*$  denotes the moment, at which the Gaussian pulse reaches its maximal value  $\varphi_0(x, t)$ , whereas  $\sigma$  is the standard deviation of the Gaussian pulse.

In case of different types of stimuli, depending on the size and duration, the parameters in (20) will have different values. In the case of many concurrently external stimuli, a series of Gaussian pulses is considered.

It is noteworthy that (20) is not a compulsory form of external stimulus where other formulas might be used, provided that they meet the principle demands discussed in this research. This mathematical form is chosen because its capability was examined in the modeling of generation of action potential in a neuron [16].

In the next section we provide a formulation for computing the effective diffusion coefficient at each moment that can be used in (19) in order to compute the brain response to external stimuli.

## 5. A Phase-Lagging Diffusion Based Model of the Diffusion Coefficient

In order to analyze the behavior of fractal time series, the value and the direction of each fluctuation in the signal should be analyzed. The direction of each fluctuation can be found by computing the value of the Hurst exponent at the previous point of the signal.

In order to know about the value of the signal, we should know about different parameters which are appeared in the fractional diffusion model. By looking at (19) and replacing the value of  $\varphi = -D_{\text{eff}}(\partial V/\partial \eta)$ , we have

$$V(\eta, t) = V_0 - C^{(2H-1)} D_{\text{eff}}^{(1-H)} \frac{1}{\Gamma(H)} \int_0^t \frac{\partial V}{\partial \eta} \frac{d\xi}{(t-\xi)^{1-H}}. \quad (21)$$

The value of the Hurst exponent which is used in (21) can be computed using MATLAB based on Rescaled Range Analysis method [31, 32]. Then, in order to compute the value of the signal in (21) two parameters  $C$  and  $D$  should be known. In order to do this, first, a relationship between  $D$  and  $H$  can be made, and then by using the formula for relaxation time, (6),  $C$  can be replaced by  $D$  and  $\tau$ .

In order to make a relation between  $D_{\text{eff}}$  and  $H$ , the phase-lagging model of action potential can be used [16]:

$$V(x, t) = V_0 - \left(\frac{D}{\tau}\right)^{1/2} \int_0^t \frac{\partial V}{\partial x} \mathfrak{Z}_0\left(\frac{t-\xi}{2\tau}\right) \exp\left(-\frac{t-\xi}{2\tau}\right) d\xi. \quad (22)$$

In (22) the diffusivity term  $D$  is related to the resistance of the neuron to the electrical impulse.  $D$  is the property of the neural tissue and will dampen the impulse as it travels over the nerve. The term  $\tau$  is the reaction time and  $\mathfrak{Z}_0(z)$  is the zero-order modified Bessel function.

Considering the phase-lagging model of action potential in the whole brain scale then, this model and the fractional diffusion model explain the same phenomenon. Thus, the mathematical equations that belong to these two models can be equal.

By writing this equality,

$$\begin{aligned} V_0 - \sqrt{\frac{D}{\tau}} \int_0^t \frac{\partial V}{\partial \eta} \mathfrak{Z}_0\left(\frac{t-\xi}{2\tau}\right) \exp\left(-\frac{t-\xi}{2\tau}\right) d\xi \\ = V_0 - C^{(2H-1)} D_{\text{eff}}^{(1-H)} \frac{1}{\Gamma(H)} \int_0^t \frac{\partial V}{\partial \eta} \frac{d\xi}{(t-\xi)^{1-H}}, \end{aligned} \quad (23)$$

where  $D$  is the diffusion coefficient in the phase lagging model of action potential and  $D_{\text{eff}}$  is the diffusion coefficient in the fractional diffusion model.

Considering  $C = (D_{\text{eff}}/\tau)^{1/2}$  and removing  $V_0$  from both sides of the equation,

$$\begin{aligned} \sqrt{\frac{D}{\tau}} \int_0^t \frac{\partial V}{\partial \eta} \mathfrak{Z}_0\left(\frac{t-\xi}{2\tau}\right) \exp\left(-\frac{t-\xi}{2\tau}\right) d\xi \\ - \tau^{1/2-H} \sqrt{D_{\text{eff}}} \frac{1}{\Gamma(H)} \int_0^t \frac{\partial V}{\partial \eta} \frac{d\xi}{(t-\xi)^{1-H}} = 0. \end{aligned} \quad (24)$$

Then, we can write

$$\begin{aligned} \int_0^t \left\{ \sqrt{\frac{D}{\tau}} \frac{\partial V}{\partial \eta} \mathfrak{Z}_0\left(\frac{t-\xi}{2\tau}\right) \exp\left(-\frac{t-\xi}{2\tau}\right) \right. \\ \left. - \tau^{1/2-H} \sqrt{D_{\text{eff}}} \frac{1}{\Gamma(H)} \frac{\partial V}{\partial \eta} \frac{1}{(t-\xi)^{1-H}} \right\} d\xi = 0. \end{aligned} \quad (25)$$

Because (25) is valid for all values of  $t > 0$  then we can remove the integral:

$$\begin{aligned} \sqrt{\frac{D}{\tau}} \frac{\partial V}{\partial \eta} \mathfrak{Z}_0\left(\frac{t}{2\tau}\right) \exp\left(-\frac{t}{2\tau}\right) \\ - \tau^{1/2-H} \sqrt{D_{\text{eff}}} \frac{1}{\Gamma(H)} \frac{\partial V}{\partial \eta} \frac{1}{(t)^{1-H}} = 0. \end{aligned} \quad (26)$$

Thus,

$$\begin{aligned} \sqrt{\frac{D}{\tau}} \frac{\partial V}{\partial \eta} \mathfrak{Z}_0\left(\frac{t}{2\tau}\right) \exp\left(-\frac{t}{2\tau}\right) \\ = \tau^{1/2-H} \sqrt{D_{\text{eff}}} \frac{1}{\Gamma(H)} \frac{\partial V}{\partial \eta} \frac{1}{(t)^{1-H}}. \end{aligned} \quad (27)$$

Diving both sides of (27) by  $\partial V/\partial \eta \neq 0$  and introducing a new variable  $z = t/2\tau$  we have

$$\frac{D_{\text{eff}}}{D} = \left[ 2^{1-H} \Gamma(H) z^{1-H} \mathfrak{Z}_0(z) \exp(-z) \right]^2. \quad (28)$$

Considering a constant value for  $D$ , in order to study the variation of  $D_{\text{eff}}/D$  versus  $z$ , Figures 3 and 4 are provided. It is noteworthy in order to analyze the value of the signal that it is only needed to concentrate on the calculations of  $D_{\text{eff}}$  and, accordingly,  $V$  in  $0 \leq H \leq 0.5$  or  $0.5 \leq H \leq 1$  because as it is known, having the probability of  $H$  on one side yields the probability of  $H' = 1-H$  in the reverse side, but with the same value of fluctuation. In this research the span  $0.5 \leq H \leq 1$  is taken as the reference. So, for instance, in order to compute the value of the signal in a point with  $H = 0.2$  it is needed just to compute the value of the signal with  $H = 0.8$ .

As it can be seen in Figures 3 and 4, first the value of dimensionless diffusivity increases but after that as time increases ( $z \rightarrow \infty$ ), dimensionless diffusivity tends a constant value. In order to describe this behavior, (28) can be analyzed when  $z \rightarrow \infty$ :

$$\lim_{z \rightarrow \infty} \frac{D_{\text{eff}}}{D} = \lim_{z \rightarrow \infty} \left[ 2^{1-H} \Gamma(H) z^{1-H} \mathfrak{Z}_0(z) \exp(-z) \right]^2. \quad (29)$$

$\lim_{z \rightarrow \infty} \mathfrak{Z}_0(z)$  is computed as

$$\lim_{z \rightarrow \infty} \mathfrak{Z}_0(z) = \frac{\exp(z)}{\sqrt{2\pi z}} = \exp(z) 2^{-1/2} \pi^{-1/2} z^{-1/2}. \quad (30)$$

By substituting (30) into (29),

$$\lim_{z \rightarrow \infty} \frac{D_{\text{eff}}}{D} = \lim_{z \rightarrow \infty} \left[ \frac{2^{1-2H} \Gamma^2(H)}{\pi} z^{1-2H} \right]. \quad (31)$$

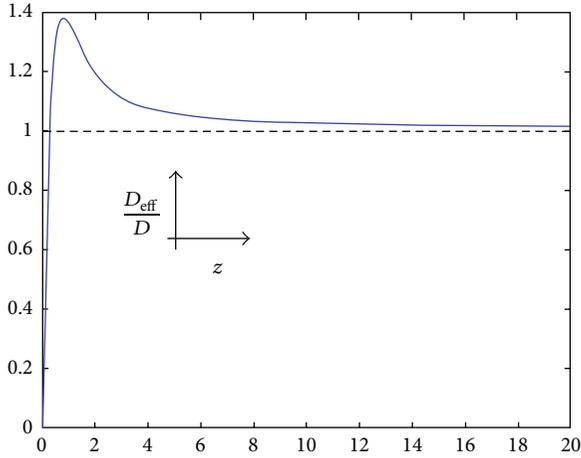


FIGURE 3: Dependence of the dimensionless diffusivity,  $D_{\text{eff}}/D$ , on the dimensionless temporal variable,  $z$ , for  $H = 0.5$ .

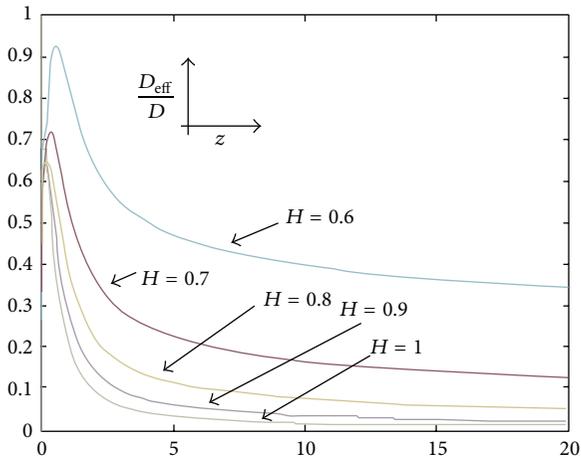


FIGURE 4: Dependence of the dimensionless diffusivity,  $D_{\text{eff}}/D$ , on the dimensionless temporal variable,  $z$ , in case of  $0.5 < H \leq 1$ .

Considering (31) in case of  $H = 0.5$

$$\lim_{z \rightarrow \infty} \frac{D_{\text{eff}}}{D} = \lim_{z \rightarrow \infty} [1] = 1 \quad (32)$$

which is the same as the trends observed in Figure 3.

By considering (31) in case of  $0.5 < H \leq 1$  we have

$$\lim_{z \rightarrow \infty} \left[ \frac{2^{1-2H} \Gamma^2(H)}{\pi} z^{1-2H} \right] = \lim_{z \rightarrow \infty} \left[ X \cdot \frac{1}{z^{2H-1}} \right]. \quad (33)$$

Thus,

$$\lim_{z \rightarrow \infty} \frac{D_{\text{eff}}}{D} = 0 \quad (34)$$

which is the same as the trend observed in Figure 4 for values of  $H$  where  $0.5 < H \leq 1$ . So, in both cases ( $H = 0.5$  and  $0.5 < H \leq 1$ ) the dimensionless diffusivity tends to a constant value as time increase.

So, in cases of Figures 3 and 4, before the maximum point of graph, as time goes on, the value of the dimensionless diffusivity increases, but after passing the maximum point, the dimensionless diffusivity shows the opposite behavior and, after some time, the dimensionless diffusivity tends to a constant value which is 0 and 1 in the cases of  $0.5 < H \leq 1$  and  $H = 0.5$ , respectively.

As  $D_{\text{eff}}$  has been considered as a time dependent parameter and  $D$  as a constant, thus from previous discussion it can be concluded that when human senses a stimulus, the diffusion of this stimulus to human brain increases but after some time, the diffusion of the stimulus decreases. In fact, these results validate and show the importance of (28) for computing the value of  $D_{\text{eff}}$ .

Thus, by computing the value of  $D_{\text{eff}}$  at each time moment and substituting its value in the fractional diffusion model the value of the signal at each time moment can be computed.

## 6. Result and Discussion

In this section using the fractional diffusion model the human brain response to a visual external stimulus is modeled and compared with the real EEG signal.

**6.1. Subjects.** In this research the experiments were carried out on 6 voluntary healthy students (21–24 years old, 3 males and 3 females). Prior to the experiment, each subject was examined and interviewed by a physician to ensure that no neurological deficit, pain condition, or medication affects the EEG. All subjects were right handed. Informed consent was obtained from each subject after the nature of the study was fully explained.

All procedures were approved by the Internal Review Board of the University and the approval for experimentation involving human subjects was issued. It is noteworthy that the identity of all subjects remains confidential.

**6.2. EEG Recording.** The EEG data used in this research were collected using Mindset 24 device, a 24-channel topographic neuromapping instrument, which can measure 24 channels of data with the sampling frequency of 256 Hz.

In an electrically shielded, acoustically isolated, and dimly illuminated room a visual stimulus applied on subjects. It should be mentioned that it is endeavoured to insulate the subjects from all other external stimuli. This ensures that the response measured in the EEG signals is primarily due to the stimulus applied. In the experiment subjects were watching a checkerboard pattern (see Figure 5) on the monitor of a computer from the distance of 130 cm.

The stimulus was the checker reversal. After one second the reversed pattern (see Figure 6) was displayed and then the original pattern was displayed again. Thus, the stimulus was applied at  $t = 1$  s. The interstimulus interval of 5 s was chosen in these experiments.

Mindmeld 24 software is used for the collection of data using Mindset 24 machine. The software gives data in the form of .bin files which can be processed to give text files (.txt) that are required for further processing.

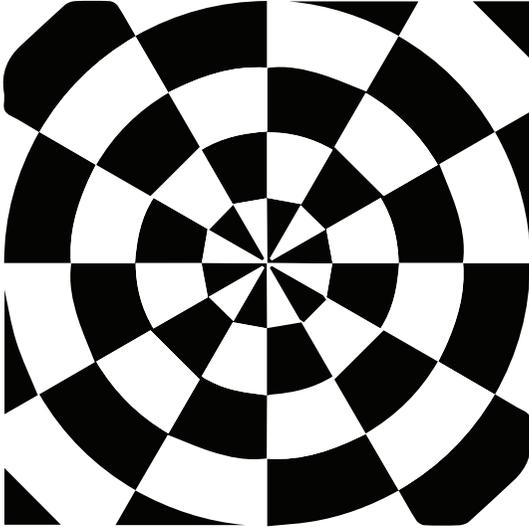


FIGURE 5: Checkerboard pattern.

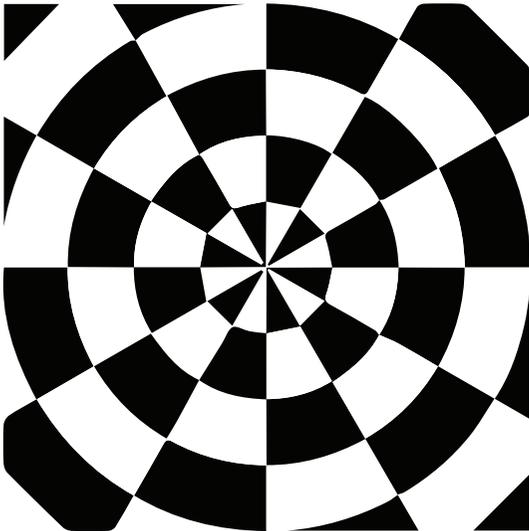


FIGURE 6: The reversed pattern as the visual stimulus.

Although EEG data are recorded from 24 electrodes, in this research the analysis is done on the data governed from the left occipital (O1) electrode (near to the location of the visual primary sensory area). This electrode was chosen based on the nearest place to the visual sensory area which shows the strongest response that can be seen in the signal recorded from this electrode compared to other electrodes. The electrode impedance was kept lower than 5 K $\Omega$ .

A bipolar electrooculogram (EOG, vertical and horizontal) was recorded for off-line artifact rejection. After bandpass filtering in the range of 0.1–70 Hz, 2 seconds of data (256 data before stimulation and 256 data after stimulation) was saved. It means that there are 256 values of voltage collected every second. Prestimulation is defined as the status before the application of the stimulus. On the other hand, poststimulation is defined as the status after the application

of the stimulus. As it was mentioned previously the stimulus was applied at  $t = 1$  s.

It is noteworthy that choosing higher sampling rate will result in more pre- and poststimulation data. Having more data will result in more precise computation of Hurst Exponent, diffusion coefficient, and accordingly the signal and its parameters such as response initiation time.

In the first week 40 trials were collected from each subject in one day. The data collections were repeated after a week for each subject in order to examine the reproducibility of the results from experiments. By repeating the experiments in the second week totally 80 trials were collected. After visual inspection of data collected from each subject and rejection of trials with artifacts, 40 trials free of artifacts were selected for future analysis. It is noteworthy mentioning that physician monitored the subjects during all experiments.

**6.3. Data Analysis.** A set of codes was written in MATLAB software in order to compute all required parameters which were discussed before.

As the recorded data were noisy, the EEG signals were filtered using the Wavelet toolbox in MATLAB and then were processed by the methodology discussed in this section.

The value of  $V_0$  can be read from the record of the signal at the moment the stimulus is applied to the subject,  $t = 1$  s. The initial value of  $H$  is computed for 1 second of the recorded data before the application of the stimulus to the subject. In order to compute the Hurst exponent, as it was mentioned previously the Rescaled Range Analysis method is employed, which is widely used by statisticians.

It is required to compute the Hurst exponent value in each moment in order to analyze the generated signal. At the first step, the program computes the Hurst exponent for the recorded EEG signal and the predicted signal and generates two time series in one figure.

In each moment, the program computes  $D_{\text{eff}}$  using (28). Also, as it was mentioned previously that, for all analysis performed here, a single stimulus is considered and thus a single Gaussian pulse is modeled using (20). By substituting the required parameters in (35), the program computes the value of the signal in each moment and then plots the modeled signal in a figure together with the recorded EEG signal (after stimulation):

$$V(\eta, t) = V_0 + \tau^{1/2-H} \frac{1}{\sqrt{D_{\text{eff}}}} \frac{1}{\Gamma(H)} \int_0^t \frac{\varphi(\eta, \xi) d\xi}{(t-\xi)^{1-H}}. \quad (35)$$

It is noteworthy that (35) is governed by substituting  $C = (D_{\text{eff}}/\tau)^{1/2}$  into (19).

The values of some required parameters are listed in Table 1.

As it was mentioned previously, one second of EEG data was recorded, and using these data as the reference, one second of EEG time series is predicted. After that, the modeled signal is analyzed in terms of the initiation time for the response, the response duration, and the peak to peak voltage. In this research in order to see the response fluctuations clearly, the real or the modelled signals after the poststimulation are averaged in each case. It means that for

TABLE 1: Values of required parameters.

Variable	Value	Units
$D$	$6.5 \times 10^{-4}$	$\text{m}^2/\text{s}$
$\varphi_0(x, t)$	1	$\text{V} \cdot \text{m}/\text{s}$
$t^*$	0.002	s
$\sigma$	0.001	s

each subject 40 selected trials are used as the input to the model and, accordingly, the grand average of all modeled signals (40 signals for each subject) is presented. Then, the generated plots for the real or the predicted signals are the result of averaging over all selected trials in each case.

Here it should be mentioned that in the analyses of the EEG signals plots the initiation time for the response and the duration are chosen based on the literature notes which consider the major positive or negative pole. For instance, the fluctuations which have voltage in the range of 5 to 10  $\mu\text{V}$  or  $-5$  to  $-10 \mu\text{V}$  are related with the response to the stimulus. Thus, for instance, in order to have a feature of the response duration, the time span between the first peak and the last peak which have the voltage values within one of these two ranges is considered.

The grand average of the recorded EEG signals and the grand average of the predicted signals using the fractional diffusion model over all selected trials for 1 second post-stimulation are shown in Figure 7.

As it can be seen for different subjects in Figure 7, the recorded signal (black solid line) and the predicted signal (red dashed line) show the similar behavior. In both cases the brain response to the stimulus starts with a positive peak ( $P$ ) after the application of the stimulus to the subject where its amplitude goes further than 5  $\mu\text{V}$ . This response causes the signal's voltage to fluctuate in a bigger span. Following the positive peak a negative rebound ( $N$ ) can be seen in the plot. In fact, the response to the stimulus terminates at this negative peak, after which the brain goes back to its normal status during rest, without any big deflection in the signal.

As an example, for subject 1 in cases of the real and predicted signals the response to the stimulus starts with a positive peak ( $P$ ) at about 118 ms and about 127 ms, respectively, after the application of the stimulus to the subject. This response causes the signal's voltage to fluctuate in a bigger span. Following the positive peak a negative rebound ( $N$ ) at about  $t = 1.170$  s and  $t = 1.174$  s can be seen in the plot in cases of the real and predicted signals, respectively. In fact, the response to the stimulus terminates at this point, after which the brain goes back to its normal status during rest, without any big deflection in the signal.

The values of peak to peak voltage, the initiation time for the response (the initial peak of the response), and the response duration (the time difference between the first and the last peaks within the response duration) for the recorded EEG signal and the predicted signal in case of different subjects are provided in Table 2.

As it can be seen in Table 2 for all subjects the predicted initiation time for the response, the response duration, and the peak to peak voltage have very close values with their

related values in the recorded EEG signals. Thus, it can be said that the predicted signal resembles the real EEG signal within the response duration in the cases of the initiation time for the response, the peak to peak voltage, and the response duration. Moreover, in order to study the uncertainty and predictability of the model's solution, the Hurst exponent variations for the recorded EEG signals and the predicted signals over all selected trials for 1 second after stimulation are shown in Figure 8.

The high correlation between the values of the real signals and also the predicted signals can be realized by looking at the values of the Hurst exponents. For instance, in the case of subject 1, the value of the Hurst exponent is distributed between 0.900 and 0.943 for the recorded EEG signal (black solid line) and between 0.900 and 0.952 for the predicted signals (red dashed line). Thus, the low uncertainty of the prediction can be confirmed, and it can be said the signal is predicted well, because the Hurst exponent values are not close to  $H = 0.5$ , which stand for a truly random process. This behavior can be seen in the Hurst exponent plots for all subjects.

Also, as it can be seen in the Hurst exponent plot, the value of the Hurst exponent in the case of the real EEG signal and the predicted signal experiences a sudden upward deflection. For instance, in the case of subject 1, the value of the Hurst exponent is decreasing in the time span of  $t = 1$  s to about  $t = 1.118$  s and  $t = 1$  s to about  $t = 1.127$  s, respectively, for real EEG signal and the predicted signal; after that, a sudden upward deflection can be seen, which stands for experiencing the visual stimulus, and again the trend shows the same behavior. The overall decreasing behavior stands for the phenomenon that when a longer time span is considered, the less the human brain "remembers" its initial state. The same behavior can be seen in other Hurst exponent plots.

By analyzing the behaviors which have been seen in Figures 7 and 8 it can be said that on one hand the uncertainty of the prediction was low and on the other hand, the accuracy of the prediction was very good as the predicted signal resembles the real EEG signal.

All the analyses which have been done in this research show that EEG signals can be modeled by the solution of fractional partial differential equations and, thus, the behavior of system modeled by means of such equations can, in principle, not only be predicted but also quantifies.

## 7. Conclusion

In this paper we introduced a new mathematical model which quantifies the human brain response to external stimuli. We developed this model by applying the fractional diffusion equation to human EEG signals. The model generates a multifractal time series which shows a quantitative concurrence with the real EEG signals. Using this model we successfully predicted the EEG signals of different subjects upon receiving a visual stimulus. This model shall be further applied in case of different external stimuli where the results can be verified against the real EEG signal which means the prediction of the human behavior by forecasting the EEG

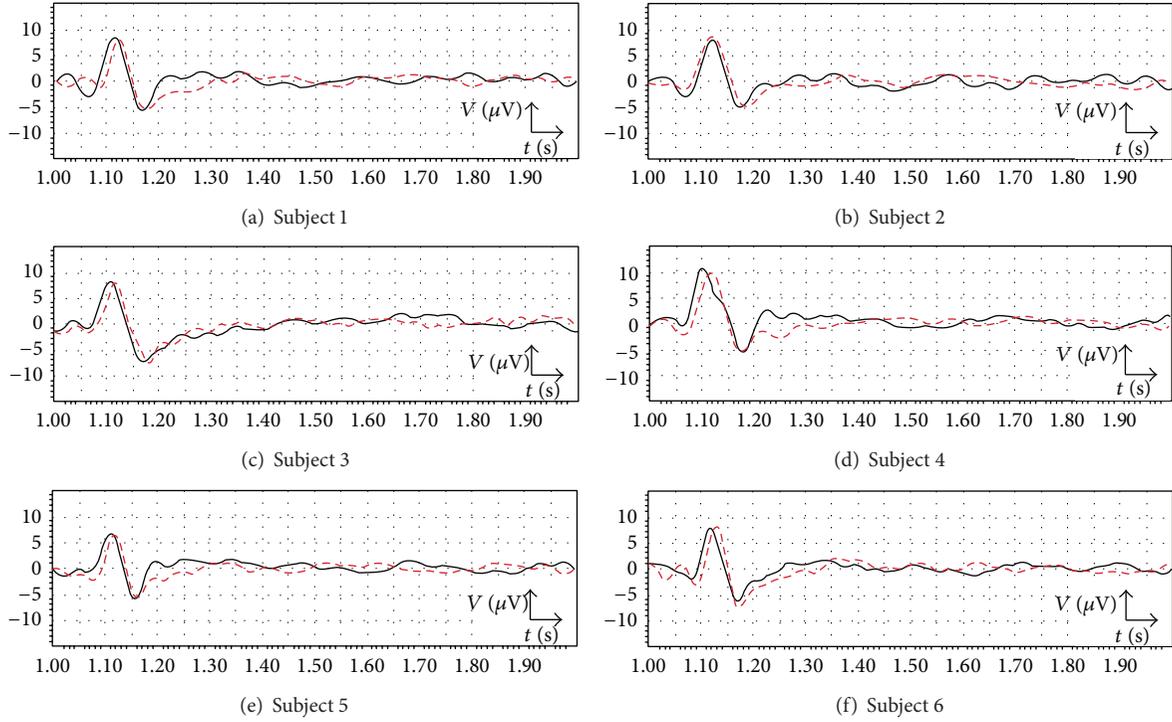


FIGURE 7: The grand average of the recorded EEG signals (black solid line) and the grand average of the predicted signals (red dashed line) for 1 second after stimulation in the case of the visual stimulus, subject 1 to subject 6.

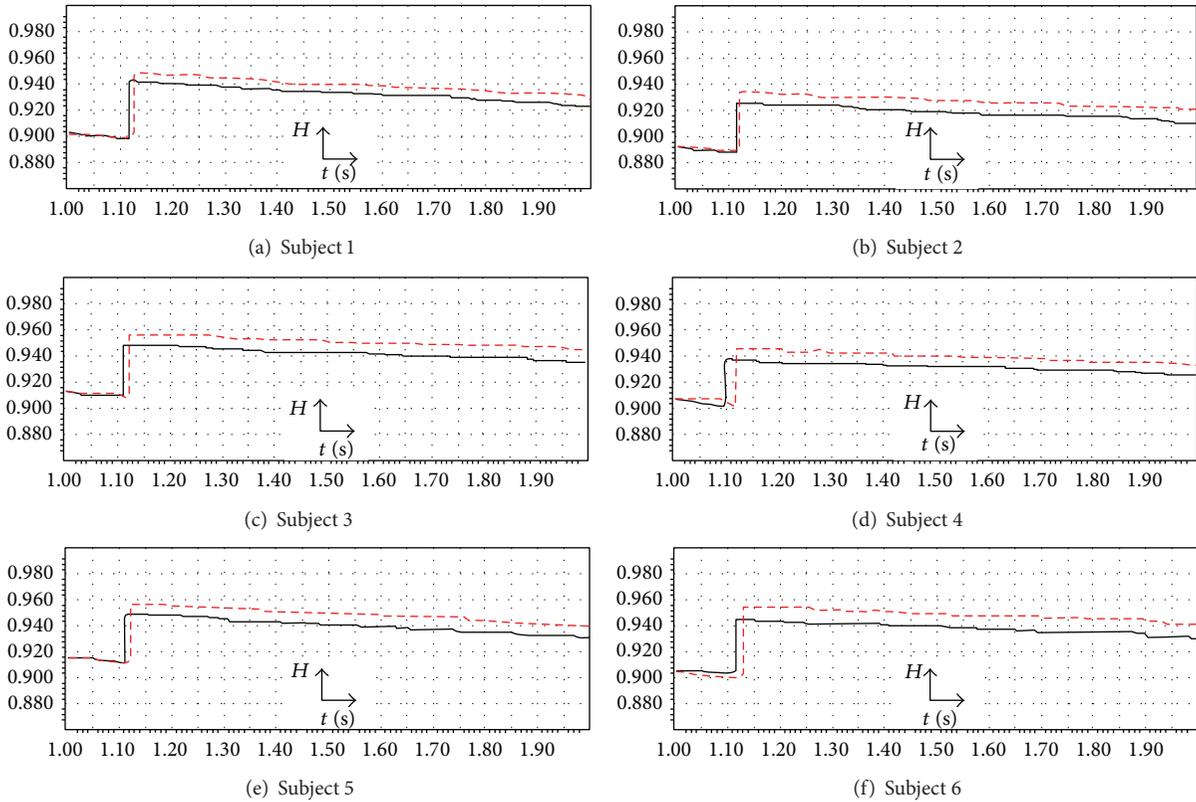


FIGURE 8: The grand average of the Hurst exponent variations for the recorded EEG signals (black solid line) and the grand average of the Hurst exponent variations for the predicted signals (red dashed line) for 1 second after stimulation in the case of the visual stimulus, subject 1 to subject 6.

TABLE 2: Comparison between the real and the predicted signals.

Subject	The initiation time for the response (ms)		Response duration (s)		Peak to peak voltage ( $\mu\text{V}$ )	
	Real	Predicted	Real	Predicted	Real	Predicted
1	P118	P127	0.052	0.047	13.97	13.07
2	P120	P120	0.054	0.067	12.86	13.75
3	P110	P120	0.058	0.060	15.24	15.24
4	P100	P120	0.078	0.057	16.28	15.50
5	P112	P121	0.043	0.041	12.50	11.81
6	P120	P130	0.050	0.044	13.85	15.50
Average	<b>113</b>	<b>123</b>	<b>0.055</b>	<b>0.052</b>	<b>14.11</b>	<b>14.14</b>

signal. On the other hand this model also can be employed in order to predict different abnormal brain activities, such as epileptic seizure, by at least some seconds before the time of occurrence. If so, a seizure warning and the expected time of this epilepsy occurrence can be generated, leading to the future monitoring of this disease.

### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

### References

- [1] Y. Kaneoke, T. Urakawa, and R. Kakigi, "Visual motion direction is represented in population-level neural response as measured by magnetoencephalography," *Neuroscience*, vol. 160, no. 3, pp. 676–687, 2009.
- [2] B. Lee, Y. Kaneoke, R. Kakigi, and Y. Sakai, "Human brain response to visual stimulus between lower/upper visual fields and cerebral hemispheres," *International Journal of Psychophysiology*, vol. 74, no. 2, pp. 81–87, 2009.
- [3] A. Sorrentino, L. Parkkonen, M. Piana et al., "Modulation of brain and behavioural responses to cognitive visual stimuli with varying signal-to-noise ratios," *Clinical Neurophysiology*, vol. 117, no. 5, pp. 1098–1105, 2006.
- [4] U. Will and E. Berg, "Brain wave synchronization and entrainment to periodic acoustic stimuli," *Neuroscience Letters*, vol. 424, no. 1, pp. 55–60, 2007.
- [5] G. Haiman, H. Pratt, and A. Miller, "Brain responses to verbal stimuli among multiple sclerosis patients with pseudobulbar affect," *Journal of the Neurological Sciences*, vol. 271, no. 1-2, pp. 137–147, 2008.
- [6] H. Eswaran, R. Draganova, and H. Preissl, "Auditory evoked responses: a tool to assess the fetal neurological activity," *Applied Acoustics*, vol. 68, no. 3, pp. 270–280, 2007.
- [7] K. Sutani, S. Iwaki, M. Yamaguchi, K. Uchida, and M. Tonoike, "Neuromagnetic brain responses evoked by pleasant and unpleasant olfactory stimuli," *International Congress Series*, vol. 1300, pp. 391–394, 2007.
- [8] W. G. Sannita, "Stimulus-specific oscillatory responses of the brain: a time/frequency related coding process," *Clinical Neurophysiology*, vol. 111, no. 4, pp. 565–583, 2000.
- [9] A. Miyanari, Y. Kaneoke, Y. Noguchi et al., "Human brain activation in response to olfactory stimulation by intravenous administration of odorants," *Neuroscience Letters*, vol. 423, no. 1, pp. 6–11, 2007.
- [10] K. Matz, M. Junghöfer, T. Elbert, K. Weber, C. Wienbruch, and B. Rockstroh, "Adverse experiences in childhood influence brain responses to emotional stimuli in adult psychiatric patients," *International Journal of Psychophysiology*, vol. 75, no. 3, pp. 277–286, 2010.
- [11] K. L. Phan, D. A. Fitzgerald, K. Gao, G. J. Moore, M. E. Tancer, and S. Posse, "Real-time fMRI of cortico-limbic brain activity during emotional processing," *NeuroReport*, vol. 15, no. 3, pp. 527–532, 2004.
- [12] P. G. Mullins, L. M. Rowland, R. E. Jung, and W. L. Sibbitt Jr., "A novel technique to study the brain's response to pain: proton magnetic resonance spectroscopy," *NeuroImage*, vol. 26, no. 2, pp. 642–646, 2005.
- [13] D. Simon, K. D. Craig, W. H. R. Miltner, and P. Rainville, "Brain responses to dynamic facial expressions of pain," *Pain*, vol. 126, no. 1-3, pp. 309–318, 2006.
- [14] W. Freeman, *Predictions on Neocortical Dynamics Derived from Studies in Paleocortex*, Birkhäuser Boston, Boston, Mass, USA, 1991.
- [15] W. J. Freeman, "Simulation of chaotic EEG patterns with a dynamic model of the olfactory system," *Biological Cybernetics*, vol. 56, no. 2-3, pp. 139–150, 1987.
- [16] K. Seetharaman, H. Namazi, and V. V. Kulsih, "Phase lagging model of brain response to external stimuli-modeling of single action potential," *Computers in Biology and Medicine*, vol. 42, no. 8, pp. 857–862, 2012.
- [17] D. T. J. Liley, P. J. Cadusch, and J. J. Wright, "A continuum theory of electro-cortical activity," *Neurocomputing*, vol. 26-27, pp. 795–800, 1999.
- [18] D. A. Steyn-Ross, M. L. Steyn-Ross, L. C. Wilcocks, and J. W. Sleigh, "Toward a theory of the general-anesthetic-induced phase transition of the cerebral cortex. II. Numerical simulations, spectral entropy, and correlation times," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 64, no. 1, p. 1918, 2001.
- [19] M. L. Steyn-Ross, D. A. Steyn-Ross, J. W. Sleigh, and L. C. Wilcocks, "Toward a theory of the general-anesthetic-induced phase transition of the cerebral cortex. I. A thermodynamics analogy," *Physical Review E*, vol. 64, no. 1, Article ID 011917, 2001.
- [20] M. L. Steyn-Ross, D. A. Steyn-Ross, J. W. Sleigh, and D. T. J. Liley, "Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: evidence for a general anesthetic-induced phase transition," *Physical Review E: Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 60, no. 6, pp. 7299–7311, 1999.
- [21] M. L. Steyn-Ross, D. A. Steyn-Ross, J. W. Sleigh, and D. R. Whiting, "Theoretical predictions for spatial covariance of the

- electroencephalographic signal during the anesthetic-induced phase transition: increased correlation length and emergence of spatial self-organization,” *Physical Review E—Statistical, Non-linear, and Soft Matter Physics*, vol. 68, no. 2, 18 pages, 2003.
- [22] M. A. Kramer, H. E. Kirsch, and A. J. Szeri, “Pathological pattern formation and cortical propagation of epileptic seizures,” *Journal of the Royal Society Interface*, vol. 2, no. 2, pp. 113–127, 2005.
- [23] V. Kulish and W. Chan, “Modeling of brain response to external stimuli,” in *Proceedings of the 7th NTU-SGH Symposium*, Singapore, 2005.
- [24] A. A. Fingelkurts, A. A. Fingelkurts, and C. F. H. Neves, “Natural world physical, brain operational, and mind phenomenal space-time,” *Physics of Life Reviews*, vol. 7, no. 2, pp. 195–249, 2010.
- [25] V. K. Jirsa and H. Haken, “Field theory of electromagnetic brain activity,” *Physical Review Letters*, vol. 77, no. 5, pp. 960–963, 1996.
- [26] A. Einstein, “Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen,” *Annalen der Physik*, vol. 322, no. 8, pp. 549–560, 1905, Reprinted in A. Einstein, “Investigations on the theory of the Brownian movement,” edited with notes by R. Fürth, translated by A. D. Cowper, London: Methuen, 1926.
- [27] V. V. Kulish, *Partial Differential Equations*, Pearson, Singapore, 2nd edition, 2010.
- [28] Y. Demirel, *Nonequilibrium Thermodynamics: Transport and Rate Processes in Physical, Chemical and Biological Systems*, Elsevier, Amsterdam, The Netherlands, 2007.
- [29] C. H. Bennett, P. Gács, M. Li, P. M. Vitányi, and W. H. Zurek, “Information distance,” *IEEE Transactions on Information Theory*, vol. 44, no. 4, pp. 1407–1423, 1998.
- [30] K. B. Oldham and J. Spanier, *The Fractional Calculus*, Academic Press, New York, NY, USA, 1974.
- [31] H. E. Hurst, “The long-term storage capacity of reservoirs,” *Transactions of the American Society of Civil Engineers*, vol. 116, pp. 770–799, 1951.
- [32] B. B. Mandelbrot and J. R. Wallis, “Noah, Joseph and operational hydrology,” *Water Resources Research*, vol. 4, no. 5, pp. 909–918, 1968.

## Research Article

# Color Enhancement in Endoscopic Images Using Adaptive Sigmoid Function and Space Variant Color Reproduction

**Mohammad S. Imtiaz and Khan A. Wahid**

*Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK, Canada S7N 5A9*

Correspondence should be addressed to Khan A. Wahid; [khan.wahid@usask.ca](mailto:khan.wahid@usask.ca)

Received 10 September 2014; Accepted 25 December 2014

Academic Editor: Kevin Ward

Copyright © 2015 M. S. Imtiaz and K. A. Wahid. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Modern endoscopes play an important role in diagnosing various gastrointestinal (GI) tract related diseases. The improved visual quality of endoscopic images can provide better diagnosis. This paper presents an efficient color image enhancement method for endoscopic images. It is achieved in two stages: image enhancement at gray level followed by space variant chrominance mapping color reproduction. Image enhancement is achieved by performing adaptive sigmoid function and uniform distribution of sigmoid pixels. Secondly, a space variant chrominance mapping color reproduction is used to generate new chrominance components. The proposed method is used on low contrast color white light images (WLI) to enhance and highlight the vascular and mucosa structures of the GI tract. The method is also used to colorize grayscale narrow band images (NBI) and video frames. The focus value and color enhancement factor show that the enhancement level in the processed image is greatly increased compared to the original endoscopic image. The overall contrast level of the processed image is higher than the original image. The color similarity test has proved that the proposed method does not add any additional color which is not present in the original image. The algorithm has low complexity with an execution speed faster than other related methods.

## 1. Introduction

Visual quality of color images plays an important role in medical image diagnosis. Wireless capsule endoscopy (WCE) is an established methodology that offers medical doctors the capability of examining the interior of the small intestine with a noninvasive procedure [1]. However, due to power and hardware limitations, the image quality in WCE is lower than high definition wired endoscopy [2]. Some GI tract related diseases, such as stomach and colon cancers and ulcerative colitis, are now of great threats to human's health [1]. Different such GI diseases can be prevented and cured by means of early detection. Despite several benefits of WCE, the images acquired by this technique are often not clear enough to see the mucosa structure, tissue and vascular characteristics of the digestive tract compared with traditional endoscope, which effects the detection accuracy and increase the miss rate during clinical diagnosis [1, 3–5]. This is why new

techniques are being constantly persuaded to enhance certain mucosal or vascular characteristics so that abnormal growths can be visualized better.

There are both in-chip and postprocessing systems that can enhance certain mucosal or vascular characteristics. Among the in-chip technologies, narrow band imaging (NBI) [6] and autofluorescence imaging (AFI) [7] are worth mentioning. There are two types of NBI systems: one is the RGB sequential illumination system, where narrow spectra of red, green, and blue lights centered on 415 nm, 445 nm, and 500 nm, respectively, are used for tissue illumination [8]. In another type of NBI system, a band-pass filter with bandwidths of 30 nm and central wavelengths of 415 nm (for blue) and 540 nm (for green) is used to generate NBI images [6]. On the other hand, in AFI system, a special rotating color filter wheel is used in front of the xenon light source to sequentially generate blue light (390–470 nm) and green light (540–560 nm) for tissue illumination [7]. All

of these techniques eventually increase the hardware complexity and power consumption of the endoscopic system. Virtual chromoendoscopy (CE) in contrast is a postprocessing system that decomposes images into various wavelengths and produces reconstructed image with enhanced mucosal surface [9]. Several researchers concluded that NBI appears to be a less time-consuming and equally effective alternative to CE for the detection of neoplasia, but with higher miss rate [3]. Additionally, neither NBI nor CE can improve the adenoma detection or reduce miss rates during screening colonoscopy. No difference has been observed in diagnostic efficacy between these two types of systems [4, 10].

There are some other global and adaptive techniques to enhance contrast and texture information of an image that is, adaptive histogram equalization (AHE) [11], contrast-limited adaptive histogram equalization (CLAHE) [12], high boost filtering (HBF) [13], brightness preserving dynamic fuzzy histogram equalization (BPDFHE) [14]. AHE applies locally varying grayscale transformation to each small blocks of the image, thus requiring the determination of the block size [15]. CLAHE operates on small regions in the image, often called tiles, instead of the entire image, based on user assigned parameters. Finally, the neighboring tiles are combined using bilinear interpolation to eliminate artificial included boundaries.

Two drawbacks of this technique are noise enhancement in smooth regions and image dependency of the contrast gain limit [15]. HBF emphasizes high frequency components without eliminating the low frequency. It may add distortions in the smoothing regions due to over filtering. BPDFHE is the modification of the brightness preserving dynamic histogram equalization (BPDHE) [16] that preserves the brightness and improves contrast enhancement abilities while reducing its computational complexity. However, it introduces additional artifacts depending on the variation of gray level distribution [17] which may lead to inaccurate diagnosis.

In this paper, a versatile endoscopic image enhancement and color reproduction method is proposed which can improve the detection rate of anomalies present in GI images. The image enhancement is achieved in two stages: image enhancement at gray level followed by space variant chrominance mapping color reproduction. Image enhancement is achieved in two steps using adaptive sigmoid function and uniform distribution of sigmoid pixels. This is somewhat similar to our previous work [18], where the enhancement is achieved by applying histogram equalization followed by adaptive sigmoid function; this can however enhance the desired mucosa and vascular features but cannot preserve the brightness of the image. As a result, in this work modified adaptive sigmoid function using precalculated gain and cutoff value is applied first to preserve the brightness of the gray image. The contrast level is enhanced in the next stage using histogram equalization.

Secondly, space variant color reproduction is achieved by generating a real color map by transferring and modifying old chrominance values either from theme image or input image. The proposed method can be useful in the following scenarios.

- (i) In white light imaging (WLI), white light is used for illuminating the GI tract and color images are generated by the endoscope. Using the proposed method, any low-contrast color WLI image can be enhanced at grayscale level and then be colorized with its original color, which can help the gastroenterologists to better inspect the vascular and mucosa structures.
- (ii) It can be used in colorizing a grayscale image using the tone of a different color theme image. This is useful when only grayscale image is available (the corresponding color image is either not available or distorted). Secondly, it is useful in saving power and bandwidth during transmission in wireless capsule endoscopy (WCE). Instead of transmitting all color images from the electronic capsule, it can only transmit one color image followed by 3 or 4 grayscale images. Using the proposed method, these grayscale images can be later colorized using the first color image as the theme image.
- (iii) In narrow band imaging (NBI), lights of 415 nm and 540 nm wavelengths are used to illuminate the mucosa surface; the reflected light from the mucosa is captured in a monochromic CCD image sensor [19]. The grayscale images from the CCD image sensor are then passed to an image processor where a pseudocolor is added to the images [20]. Using the proposed method, the grayscale NBI images can be further enhanced for better visibility of the mucosa structure; pseudocolors can then be added using the tone of any color theme image.

## 2. Proposed Method

The proposed method consists of two stages: image enhancement and space-variant chrominance mapping based color reproduction. The method is shown in Figure 1. The stages are briefly discussed below.

*2.1. Image Enhancement.* At first, the color endoscopic image is converted into  $YC_bC_r$  color space using (1). Here,  $Y$  is luminance or luma and  $C_b$  and  $C_r$  are chrominance components. The color space conversion allows us to process different luma pixels to enhance vascular features and chrominance pixels for color reproduction. Consider

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}. \quad (1)$$

Here,  $Y$  is considered grayscale image. After conversion, the proposed method normalizes grayscale image and each chrominance plane between 0 and 1 using (2). Consider

$$N_{\text{norm}}(x) = \frac{x - x_{\min}}{x_{\max}}. \quad (2)$$

Here,  $x_{\min}$  and  $x_{\max}$  are minimum and maximum pixel values. Later, the normalized grayscale image is enhanced using adaptive sigmoid function and uniform distribution.

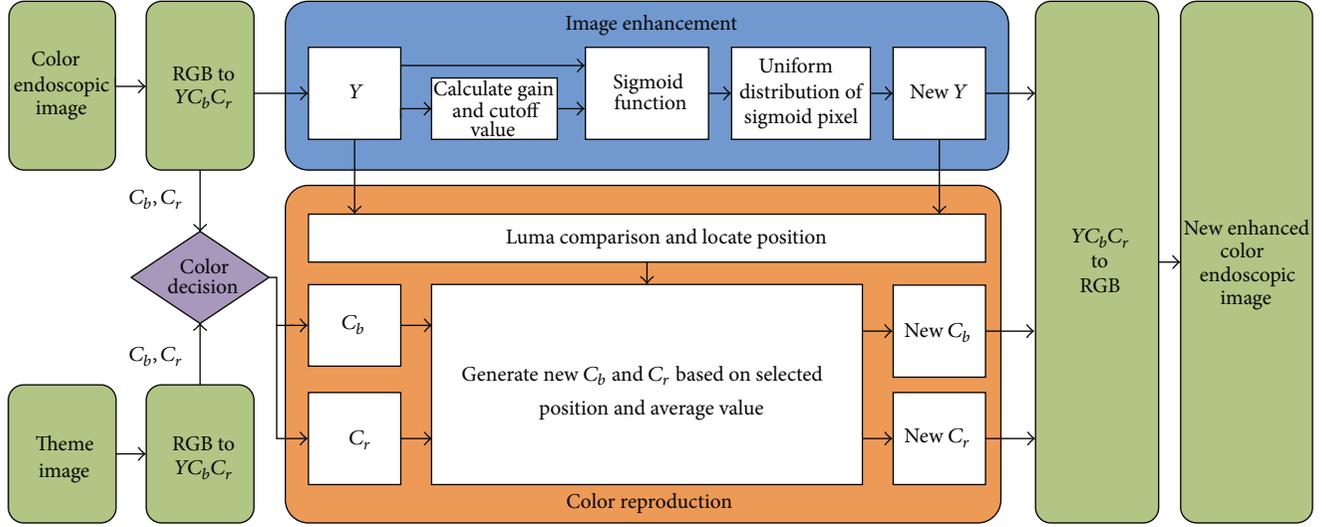


FIGURE 1: Proposed color image enhancement method.

**2.1.1. Adaptive Sigmoid Function.** The proposed method uses contrast manipulation techniques for image enhancement. Generally, contrast manipulation technique can be performed either globally or adaptively. Global techniques apply a transformation to all image pixels, while adaptive techniques use an input-output transformation that varies adaptively with local image characteristics. Our method transforms the pixel values adaptively using sigmoid function.

In general, a sigmoid function is real valued and differentiable, having either a nonnegative or nonpositive first derivative that is bell shaped. It has been used in several researches related to image processing [25–27]. Using  $x$  for the input, the sigmoid function is given below:

$$f(x) = \frac{1}{(1 + e^{-x})}. \quad (3)$$

In the training mode, we have observed that in a certain exponent the image highlights some vascular characteristics and mucosa structure, which are not clearly visible in the original image. To control the exponent, we have introduced two coefficients in the sigmoid function. Using  $x$  for the input,  $g$  for gain, and  $k$  for cutoff, the modified sigmoid function is expressed below:

$$f(x) = \frac{1}{(1 + e^{g(k-x)})}. \quad (4)$$

The cutoff value determines the midpoint of the input curve and the gain controls the amount of bending. These two parameters give us the control to train the proposed method to generate a certain exponent that highlights some vascular characteristics. Let,  $x = 0, 0.1, 0.2, \dots, 1$  normalized image pixel values where sigmoid function (4) is applied. Figure 2 presents the sigmoid curve of input pixel values based on different cutoff and gain.

These parameters (gain and cutoff) can control the overall brightness and contrast level of the image too. The cutoff

value controls the amount of brightness and the gain controls the consecutive difference between pixels. To maintain the exponent into desired level, we have proposed algorithms to generate cutoff and gain value. Based on the input pixel values, (5) generate specific cutoff and gain value. Later on, these values are used in (4) to generate the sigmoid image. Consider

$$k = \frac{\sum_{i=1}^n x_i}{n}, \quad (5)$$

$$g = A \times \log\left(\frac{S_m}{S_n}\right) \times \frac{\sum_{i=1}^n x_i}{n},$$

where  $A = 100$ ,  $S_m = 6$ ,  $S_n = 5$ ,  $x_i$  is the pixel values of  $i$ th position and  $n$  is the number of pixel. These values are heuristically collected from simulation. First of all, we processed endoscopic images in different combination of gain and cutoff values. The images are collected from Gastrolab [28] and Atlas [29] database and have comments from gastroenterologist; as a result, they can be sub-divided into different disease categories. Figure 3 shows some examples of the original and corresponding sigmoid images. The abnormalities in the images may be identified, but not the tissue and vascular characterization (as marked with an arrow in Figures 3(a) and 3(c)). It is noted that mucosa structure, tissue and vascular characteristics are important since by analyzing them the status of gastric glands and pits can be investigated [30–32].

During simulation, we observed that in certain cases, with gain in a range of 7.5–8.5 and cutoff in a range of 0.4–0.5, the tissue and vascular characterization are highly visible. To keep the gain and cutoff in that desired range, we propose (5). For better illustration, we have presented sigmoid images processed with different combination of gain and cutoff values in Table 1. Here, the effects on images for different combination of gain and cutoff values are observed. For example, Image #1 and #5 have low intensity; image #2

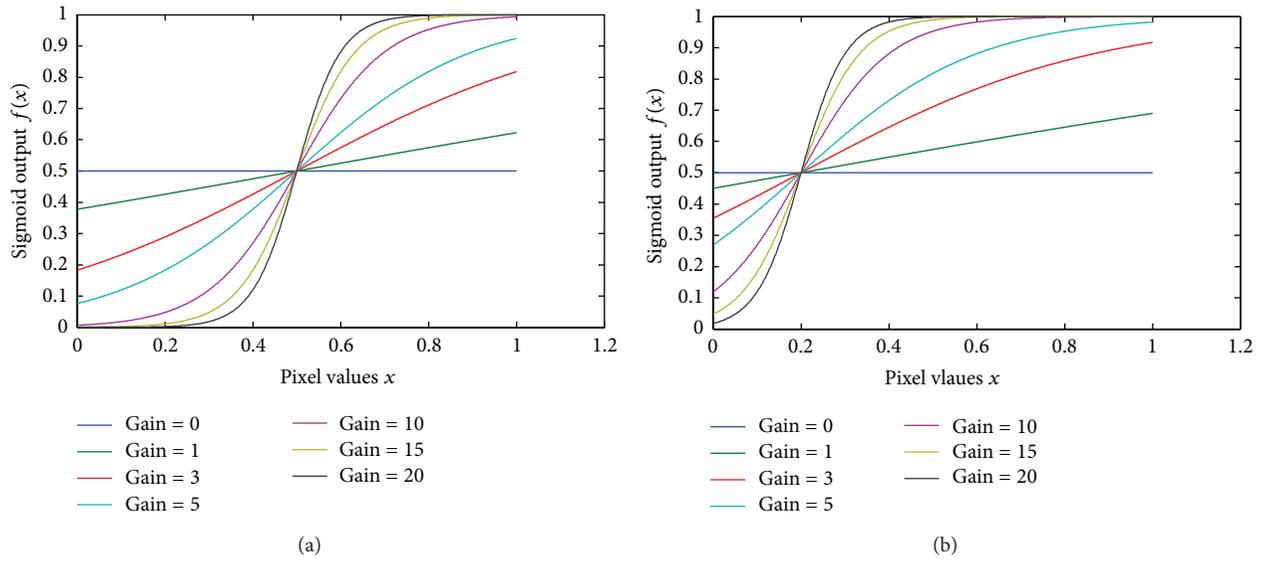


FIGURE 2: Sigmoid effect on pixel for different gain values (a) with 0.5 cutoff; (b) with 0.2 cutoff.

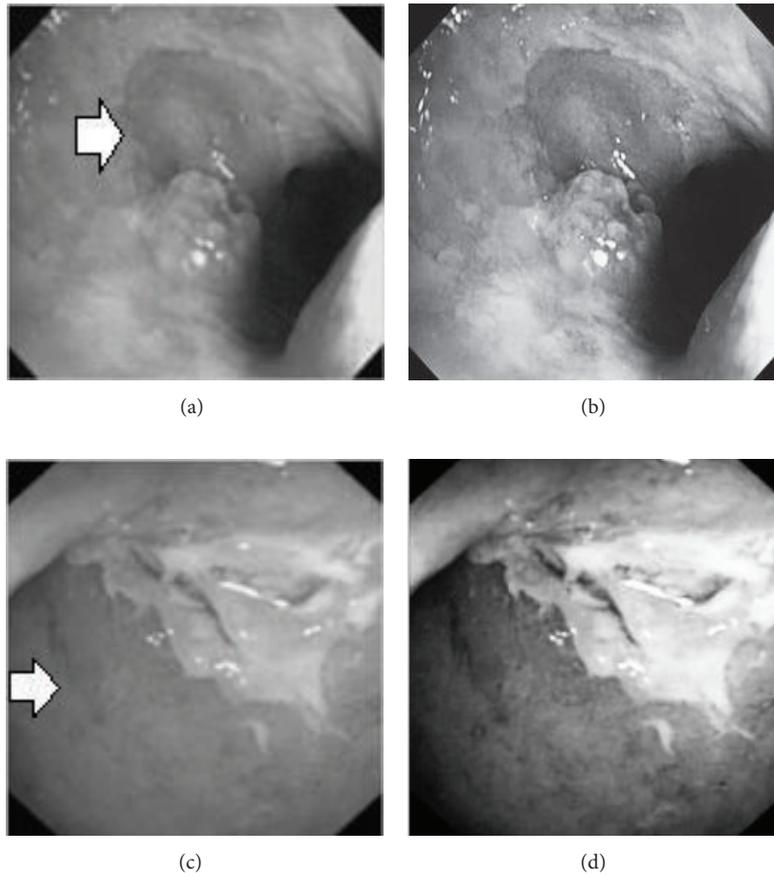
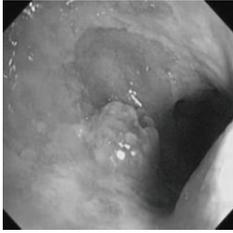
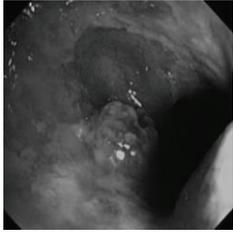
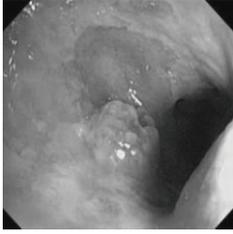
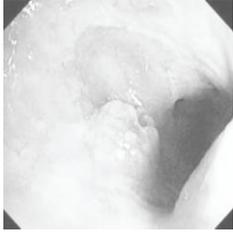
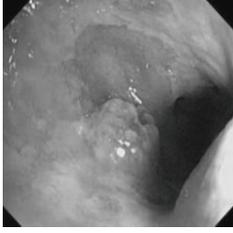
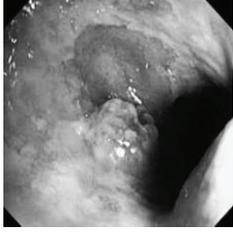
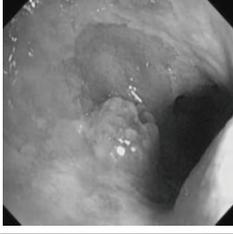
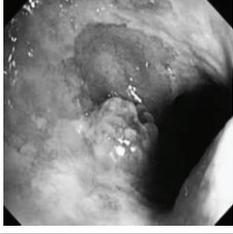
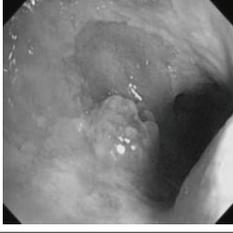
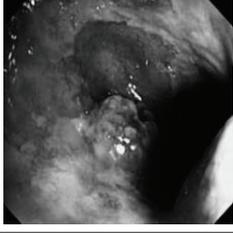


FIGURE 3: (a) Original image with defected polyp and (c) Crohn's disease; (b) and (d) adaptive sigmoid images of (a) and (c), respectively.

TABLE 1: Sigmoid image with different combination of gain ( $g$ ) and cutoff values ( $k$ ).

Number	Original gray scale image	Sigmoid image	Used gain and cutoff values
1			$g = 5$ $k = 0.7$
2			$g = 6$ $k = 0.1$
3			$g = 8$ $k = 0.5$
4			$g = 7.5$ $k = 0.4$
5			$g = 9$ $k = 0.6$

has high brightness; image #3 and #4 have highlighted tissue and vascular characterization.

**2.1.2. Uniform Distribution of Sigmoid Pixels.** In the next stage, the sigmoid pixels are uniformly distributed to increase the contrast level. It helps to visualize the vascular characteristic of darker part of an adaptive sigmoid image. It is employed by effectively spreading out the most frequent intensities.

Let,  $f$  be a given sigmoid image represented as  $i$  by  $j$  matrix of integer pixel intensities ranging from 0 to 255. Let,  $p$

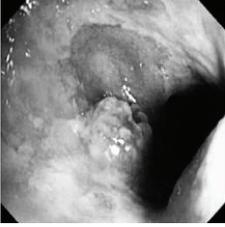
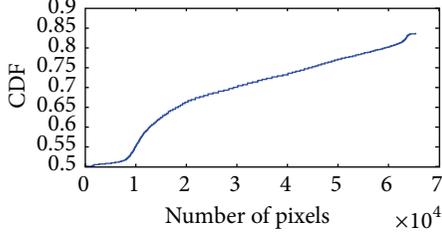
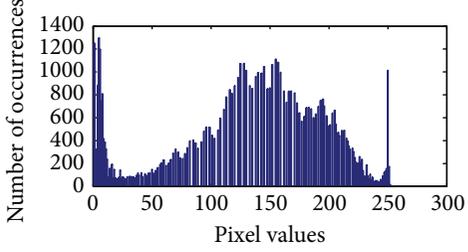
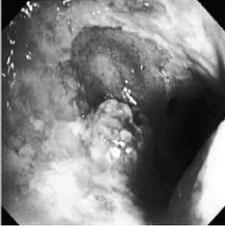
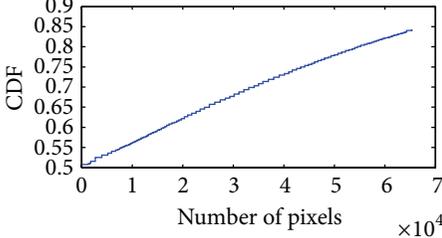
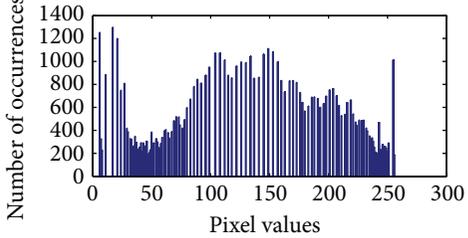
denotes the normalized histogram of  $f$  with bin for possible intensities. So,

$$p_m = \frac{\text{Number of pixels with intensity } m}{\text{total number of pixels}}, \quad (6)$$

where  $m = 0, 1, 2, \dots, 255$ . The uniformly distributed sigmoid image  $\Psi$  is defined as,

$$\Psi_{i,j} = \text{floor} \left( (L-1) \sum_{n=0}^{f_{i,j}} p_n \right), \quad (7)$$

TABLE 2: The Comparison between processed sigmoid image and uniformly distributed sigmoid image.

Image type	Cumulative distribution function (CDF)	Histogram of image
 Adaptive sigmoid image		
 Uniformly distributed sigmoid image		

where floor() maps to the largest integer but lesser than the number. Normally, the cumulative distribution function (CDF) of an image does not form a horizontal line, that means, the pixel values are not equally likely to occur. In the proposed method, a uniform distribution of sigmoid pixels is achieved by applying (6) and (7); this technique is similar to global histogram equalization. Table 2 shows the visual comparison of uniform distribution of sigmoid pixels. This uniformly distributed sigmoid image  $\Psi_{i,j}$  is later treated as new enhanced grayscale image ( $\Psi$ ).

**2.2. Color Reproduction.** In the second stage of the proposed method, we apply color reproduction. It is a computer-assisted process of adding color to a monochrome image [33, 34]. In the proposed method, it is possible to retrieve the original color with a better tone or add pseudocolor using a theme image. This choice is controlled by the user through the “color decision” module (see Figure 1) which selects the chrominance components.

*Case 1.* To retrieve original color, we first create new  $C_b$  and  $C_r$  planes by matching the original  $C_b$  and  $C_r$  values for corresponding  $Y$  pixels from the original grayscale image. First of all, the positions of all  $C_b$  and  $C_r$  values in the plane for a particular  $Y$  pixel are identified as expressed by (8):

$$[m, n] = \text{locate}(Y - Y_{i,j}). \quad (8)$$

Here,  $Y$  is normalized grayscale image,  $Y_{i,j}$  is a pixel of normalized grayscale image and  $[m, n]$  holds one or multiple positions. These positions will allow us to generate new chrominance planes. Two scenarios may occur: (a) if only one chrominance value is found, it places that value in the corresponding positions in the new  $C_b$  and  $C_r$  planes. (b) Otherwise, if multiple chrominance values are found, it

generates a new chrominance value using (9) and places it in the corresponding positions of the new  $C_b$  and  $C_r$  planes

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}. \quad (9)$$

These steps continue until all pixels of the grayscale image are scanned. The new  $C_b$  and  $C_r$  will have the same dimension of the original grayscale image. Later, the enhanced grayscale image ( $\Psi$ ) and the new  $C_b$  and  $C_r$  images are converted back to RGB image using (10)

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.164 & 0 & 1.596 \\ 1.164 & -0.392 & -0.813 \\ 1.164 & 2.017 & 0 \end{bmatrix} \begin{bmatrix} \Psi - 16 \\ C_b^{\text{new}} - 128 \\ C_r^{\text{new}} - 128 \end{bmatrix}. \quad (10)$$

*Case 2.* To add pseudocolor, a theme image is required. It is applicable when only grayscale image or no color information is available. As the color information in an endoscopic image dictates clinical decision, the selection of theme image is very important. The theme image must be selected from the nearby location or region of GI tract. After selecting a proper theme color image, it is converted into  $YC_bC_r$  space. Then, we create new  $C_b$  and  $C_r$  planes by matching the chrominance values of the theme image for the corresponding enhanced pixel ( $\Psi_{i,j}$ ). Now, similar procedure as given in (8) is followed to find the new  $C_b$  and  $C_r$  planes (given in (11))

$$[m, n] = \text{locate}(Y_t - \Psi_{i,j}). \quad (11)$$

Here,  $Y_t$  is normalized theme grayscale image,  $\Psi_{i,j}$  is a pixel of enhanced grayscale image and  $[m, n]$  holds one or multiple locations. These locations allow us to generate the new chrominance plane with respect to the enhanced and theme grayscale images. Here, the chrominance values are generated from the  $C_b$  and  $C_r$  planes of the theme image.

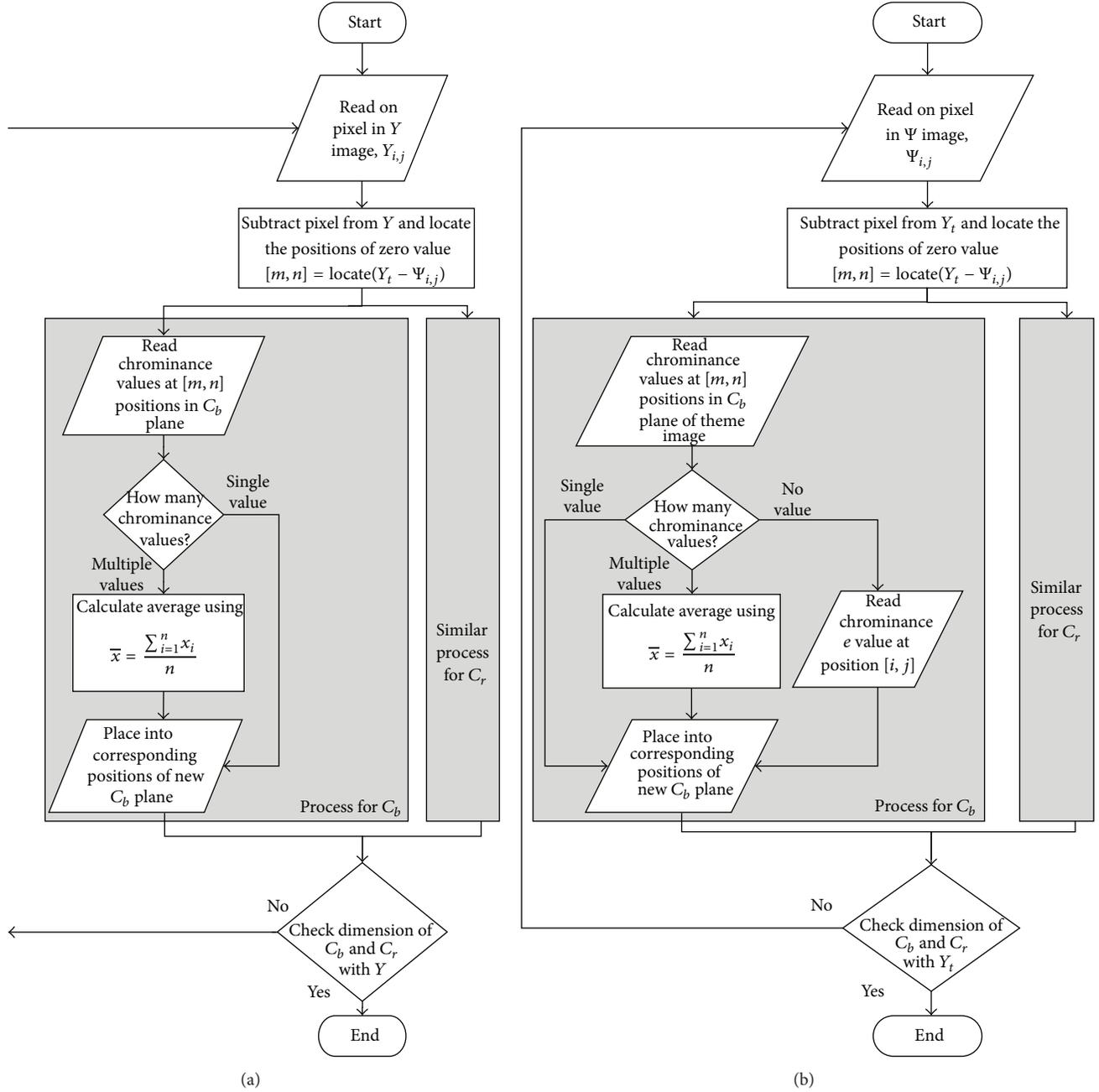


FIGURE 4: Flow chart: (a) Case 1: to retrieve original color; (b) Case 2: to add pseudocolor from theme image.

Three scenarios may occur: (a) if only one chrominance value is found, it places that value in the corresponding positions in the new  $C_b$  and  $C_r$  planes. (b) if multiple chrominance values are found, it generates a new chrominance value using (9) and places it in the corresponding positions of the new  $C_b$  and  $C_r$  planes (c) if no chrominance value is found, it reads the chrominance value respect to the positions of  $\Psi_{i,j}$  in theme  $C_b$  and  $C_r$  planes and places it in the corresponding position of the new  $C_b$  and  $C_r$  planes. These steps continue until all pixels of the enhanced grayscale image are scanned. The new  $C_b$  and  $C_r$  will have the same dimension of the original grayscale image. Later, the enhanced grayscale image

( $\Psi$ ) and the new  $C_b$  and  $C_r$  images are converted back to RGB image using (10).

In Figure 4, the flow chart of the color reproduction algorithm is presented. Some reconstructed images for the two cases are shown in Figures 5 and 6. It can be seen that the proposed method enhances color information in all reconstructed images.

### 3. Results and Discussion

In order to evaluate the performance of the proposed algorithm, we have applied it to several endoscopic images

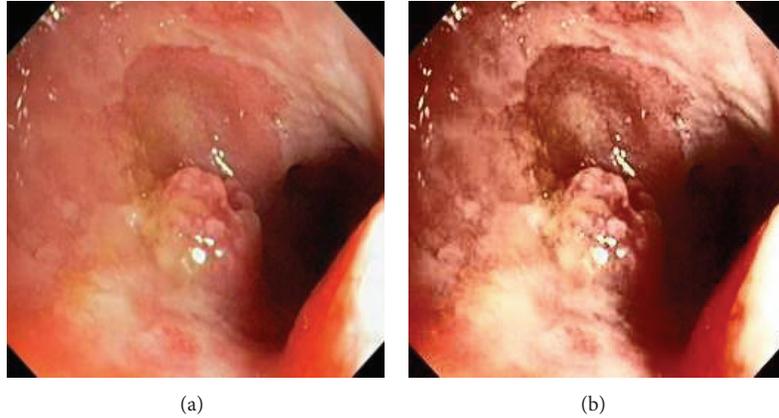


FIGURE 5: (a) Original image and (b) enhanced color image (color reproduced from original image).

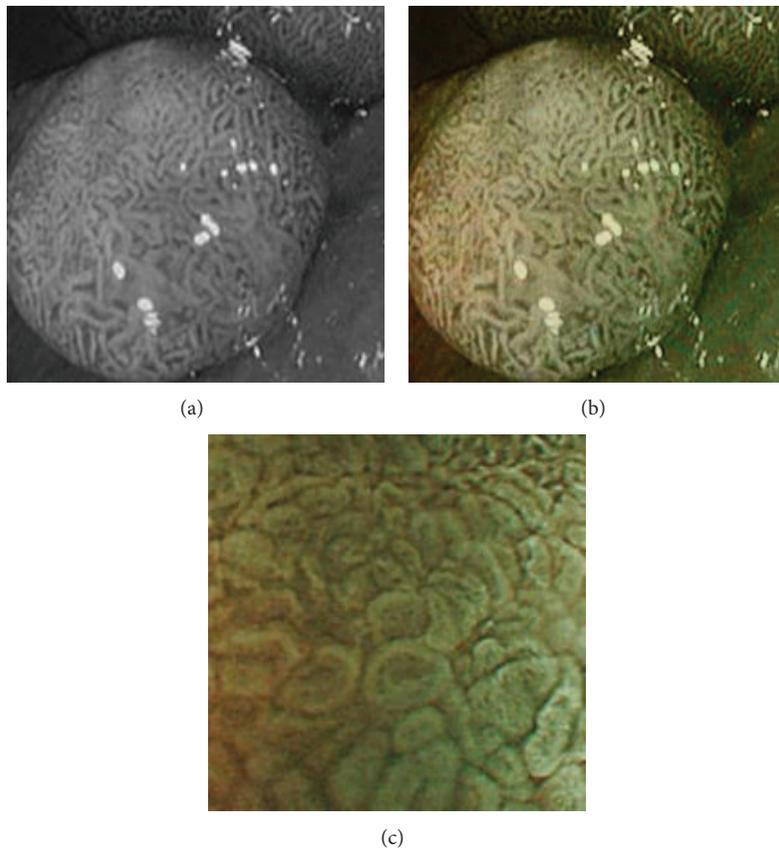


FIGURE 6: (a) Original grayscale image (no color information is available) (b) enhanced color image (color reproduced from theme image shown in (c)).

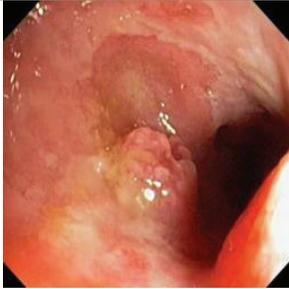
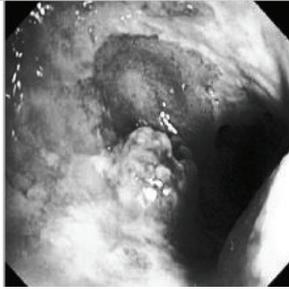
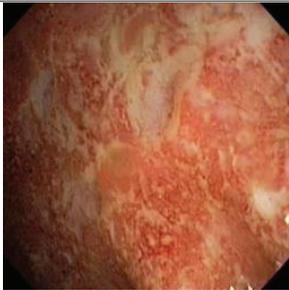
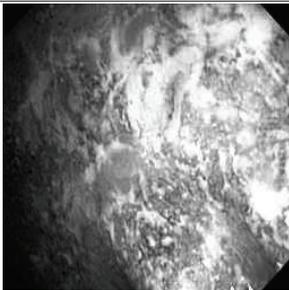
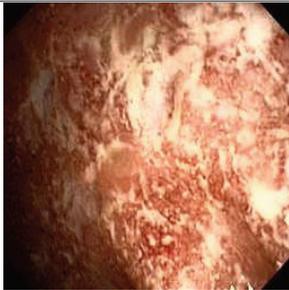
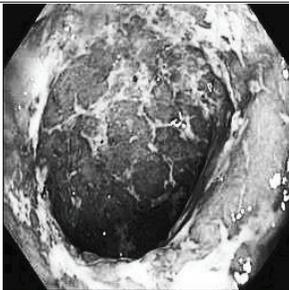
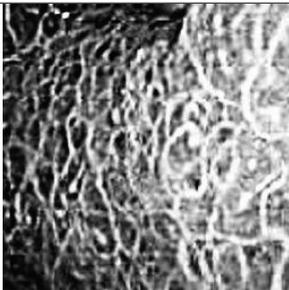
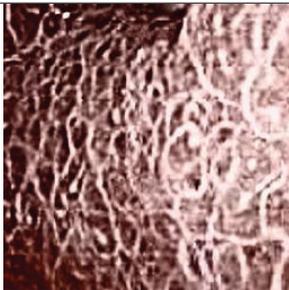
collected from Gastrolab [28] and Atlas [29]. The results are summarized below in four categories.

**3.1. Category 1: Low-Contrast Color Images.** In this case, the input image is first enhanced on gray level and then color added. The chrominance values of the original input image are used for color reproduction. As a result, the output image has similar color tone with enhanced features as shown in

Table 3. It can be seen from the table that the vascular and other mucosa structures are better visible and highlighted in the output images, which can help the gastroenterologists in better diagnosis.

**3.2. Category 2: Low-Contrast Grayscale Images.** In this case, we show examples where low-contrast grayscale images are used (i.e., color information is not available for these images).

TABLE 3: Category 1: Enhancement of colored WLI images (where input image is used as theme image).

Number	Input original image (color)	Enhanced grayscale image	Output enhanced color image
1			
	(a)	(b)	(c)
2			
	(d)	(e)	(f)
3			
	(g)	(h)	(i)
4			
	(j)	(k)	(l)

The grayscale images are first enhanced and then colorized using a theme image. The choice of the theme image is important as it may add color distortion if not properly chosen. As a result, we choose a theme image from the same or similar physical location of the GI tract. The results of the enhanced color images are shown in Table 4 along with the corresponding theme images.

3.3. *Category 3: Raw NBI Images.* In the next experiment, we applied our algorithm on several NBI images (grayscale

in nature) as shown in Table 5. The raw NBI images are enhanced first and then a color theme image is used to generate pseudocolor. The theme images are chosen the same way as described before. We can see from the table that the output images have much better visibility of the mucosa structure compared to the grayscale images.

3.4. *Category 4: Image Transmission in WCE.* The proposed color generation method is very useful in saving power consumption during transmission in wireless capsule endoscopy

TABLE 4: Category 2: Enhancement of grayscale WLI images (here, no original color image is available, so color image from similar location of the GI tract is used as theme image).

Number	Input original image (grayscale)	Enhanced grayscale image	Theme image	Output enhanced color image
1				
2				
3				

TABLE 4: Continued.

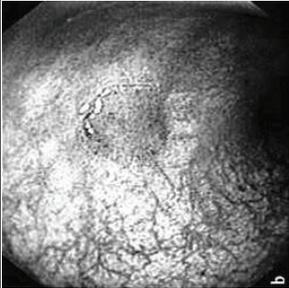
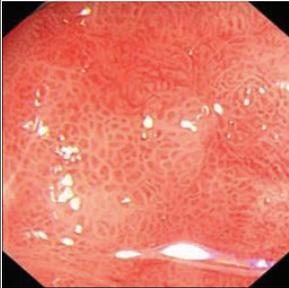
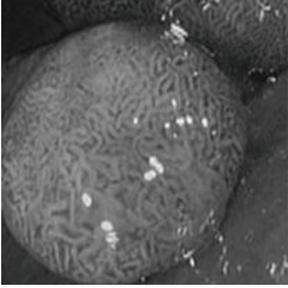
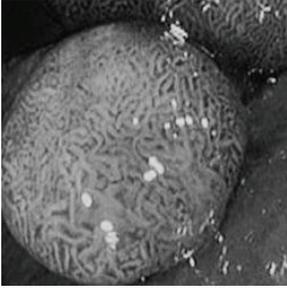
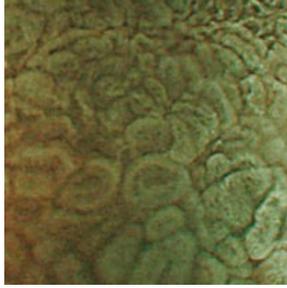
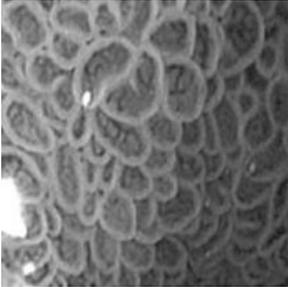
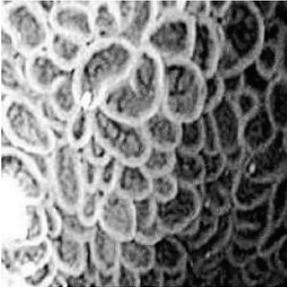
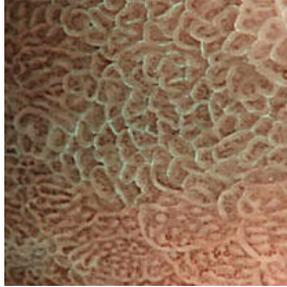
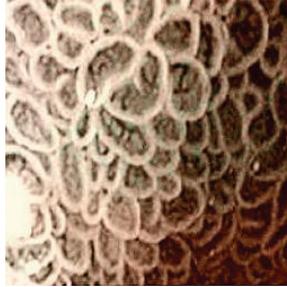
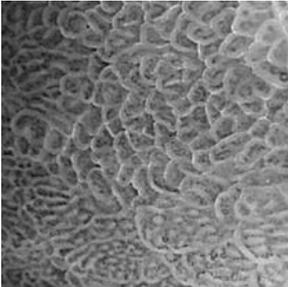
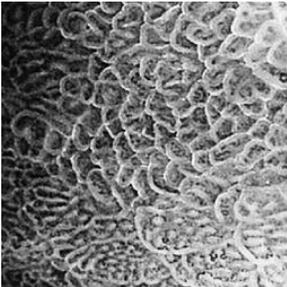
Number	Input original image (grayscale)	Enhanced grayscale image	Theme image	Output enhanced color image
4	 (m)	 (n)	 (o)	 (p)

TABLE 5: Category 3: Enhancement and color reproduction of grayscale NBI images.

Number	Input RAW NBI image	Enhanced grayscale image	Theme image	Output enhanced color image
1				
2				
3				

(WCE). Instead of transmitting all color images from the electronic capsule (which takes 24 bits per pixel per image), it can only transmit one color image at the beginning followed by a defined number of grayscale images (8 bits per pixel per image). Using the proposed method, these grayscale images will be later colorized using the first color image at the receiver. In Table 5, we show the results of such case where the R, G and B components of frame 1 are transmitted first. Then only the luminance (Y) components of frame 2, 3, 4, and 5 are transmitted. At the receiver, these frames 2–5 are reconstructed using the proposed color reproduction method taking frame 1 as the theme image. Later on, the color reconstructed images are compared with the original color video sequences. In conventional case, the R, G, and B components of all frames are transmitted. For the given case, for five frames, it will require a total of 120 bits per pixel (i.e.,  $24 \times 5$ ). On the other hand, using the proposed method, it will require only 56 bits per pixel (i.e.,  $24 + 8 + 8 + 8 + 8$ ) which results in a saving of 53% during the transmission. More saving will be achieved using the number of grayscale frames

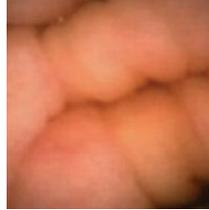
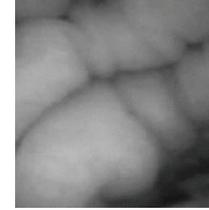
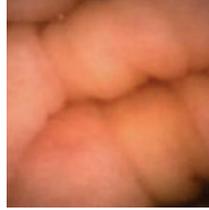
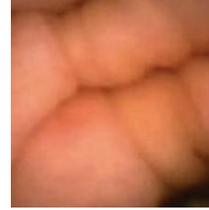
is increased. The original color video frames are also shown in Table 6 for comparison. Here we see that the reconstructed output images have the same color as compared with the original color video frames with a power saving of 53%.

It should be noted here that, the previous work [18] was only applied to color images whereas the proposed method can be applied to both color and grayscale images. As a result, low-contrast gray (category 2) and NBI raw (category 3) images can be colorized using the method using a theme image. This feature also makes the algorithm helpful in saving power during WCE image transmission (category 4).

#### 4. Performance Analysis

In the following section, the performance of the proposed scheme is evaluated using focus value, statistic of visual representation, measurement of uniform distribution, color similarity test, color enhancement factor (CEF) and time complexity. The results are discussed below.

TABLE 6: Category 4: Reproduction of color frames from grayscale frame in WCE video; no enhancement was applied.

Frame number	1	2	3	4	5
Original video frames					
Input images					
Proposed color images (Frame number 1 used as theme image)					

**4.1. Focus Value.** In our method, image enhancement is achieved by adaptive sigmoid function and uniform distribution of sigmoid pixels. As a result, the overall information of sharp counters and contrast is increased. These changes of an image are evaluated using focus value [35]. Focus value is a mathematical representation of the ratio of AC and DC energy values of a Discrete Cosine Transform (DCT) of an image [36]. Let  $E_{AC}$  be the AC values and  $E_{DC}$  the DC value of a DCT image.  $E_{AC}$  values carry the information related to high frequency component (i.e., changes of contrast level, sharp counters and crisp edges) of an image. On the other hand,  $E_{DC}$  value carries only the information related to low frequency components (i.e., luminance or brightness). The expressions are given below:

$$E_{DC} = (F_{DC}(u, v))^2, \quad (12)$$

$$E_{AC} = \sum_{u=1}^n \sum_{v=1}^m (F_{AC}(u, v) - \overline{F_{AC}(u, v)})^2.$$

Here,  $u$  and  $v$  represent the row and column of the DCT image,  $F_{DC}$  is the DC part and  $F_{AC}$  is the AC part of DCT image. The resultant of the ratio of  $E_{AC}$  and  $E_{DC}$  is the focus value  $F_S$  as given by

$$F_S = \frac{E_{AC}}{E_{DC}}. \quad (13)$$

If the overall information of sharp counters, crisp edges and contrast of enhanced image is higher than the original image, then  $F_S$  of the enhanced image will be higher than

that of the original image and vice versa. We have compared our method in terms of focus value using 60 sample images with other methods like AHE [11], CLAHE [12], HBF [13] and BPDFHE [14]. The results are presented in Table 7. Here, we see that the focus values of the proposed method are relatively higher compared to the other methods.

**4.2. Statistic of Visual Representation.** Next, we used statistic of visual representation [37] to measure the contrast and intensity distortion between two images. Equations (14) represent statistic visual representation. Consider

$$C = \frac{\sigma_{out} - \sigma_{in}}{\sigma_{in}}, \quad (14)$$

$$L = \frac{L_{out} - L_{in}}{L_{in}},$$

where  $\sigma_{out}$  and  $L_{out}$  are the variance and mean of enhanced image;  $\sigma_{in}$  and  $L_{in}$  are the variance and mean of original image, respectively. Here,  $C$  defines the percentage of increment or decrement of contrast level and  $L$  defines the percentage of increment or decrement in intensity level. In our experiment, we used 60 grayscale images. The results are presented in Table 8. We can see that the  $C$  and  $L$  of the first image using proposed method are 1.0636 and 0.0716, which means that the contrast and intensity level of proposed image are 103.6 and 7.16 times higher than the original image, respectively. Here, the negative sign denotes the decrement. It is noticeable that the proposed method's contrast level and intensity level are higher compare to the other method.

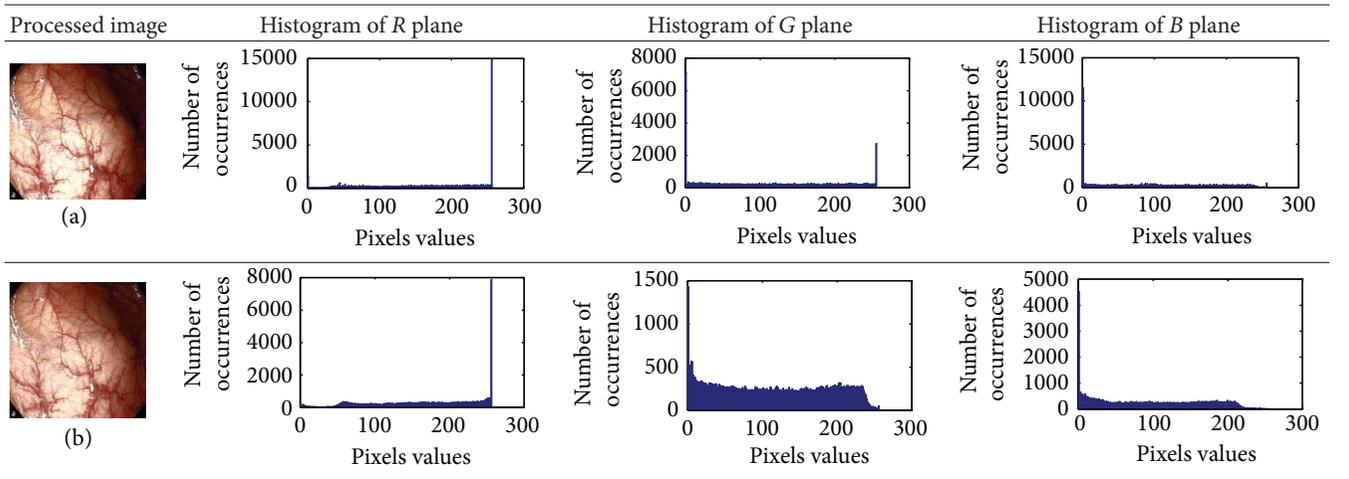
TABLE 7: Comparisons of focus value with other related works.

Image number	Focus value					
	Original	AHE [11]	CLAHE [12]	HBF [13]	BPDFHE [14]	Proposed
1	15.64	17.50	22.96	22.86	22.04	<b>42.10</b>
2	12.52	16.75	20.14	19.33	17.11	<b>42.52</b>
3	18.01	19.20	23.18	22.40	22.00	<b>41.57</b>
4	13.66	18.85	22.01	20.97	20.51	<b>42.00</b>
Average of 60 endoscopic images	13.77	19.08	21.11	20.49	19.49	<b>41.17</b>

TABLE 8: Comparisons of statistic of visual representation with other related works.

Image number	Contrast measurement					Intensity measurement				
	AHE [11]	CLAHE [12]	HBF [13]	BPDFHE [14]	Proposed	AHE [11]	CLAHE [12]	HBF [13]	BPDFHE [14]	Proposed
1	0.3373	0.4735	0.4285	0.2025	<b>1.0636</b>	-0.1781	0.0598	0.0254	0.0065	<b>0.0716</b>
2	0.2767	0.6352	0.2395	0.3608	<b>2.4684</b>	-0.0312	0.1207	0.0095	0.0092	<b>0.2526</b>
3	0.1205	0.2935	0.2395	0.1192	<b>1.12294</b>	-0.1122	0.0410	0.0095	-0.0012	<b>0.1322</b>
4	0.3401	1.1598	0.8235	0.7206	<b>2.0349</b>	-0.1186	-0.0291	0.0134	-0.0100	<b>0.0218</b>
Average of 60 endoscopic images	0.2554	0.6411	0.4149	0.3371	<b>1.6147</b>	-0.1203	0.0381	0.0197	0.0174	<b>0.1748</b>

TABLE 9: Histogram of R, G, and B planes in terms of uniform distribution (a) without and (b) with color reproduction.



4.3. *Measurement of Uniform Distribution.* Here, we calculate the uniform distribution of R, G, and B planes by calculating entropy [38, 39]. The more the uniform distribution of color planes, the better the color enhancement. The entropy of  $n$  distributed signals is defined by

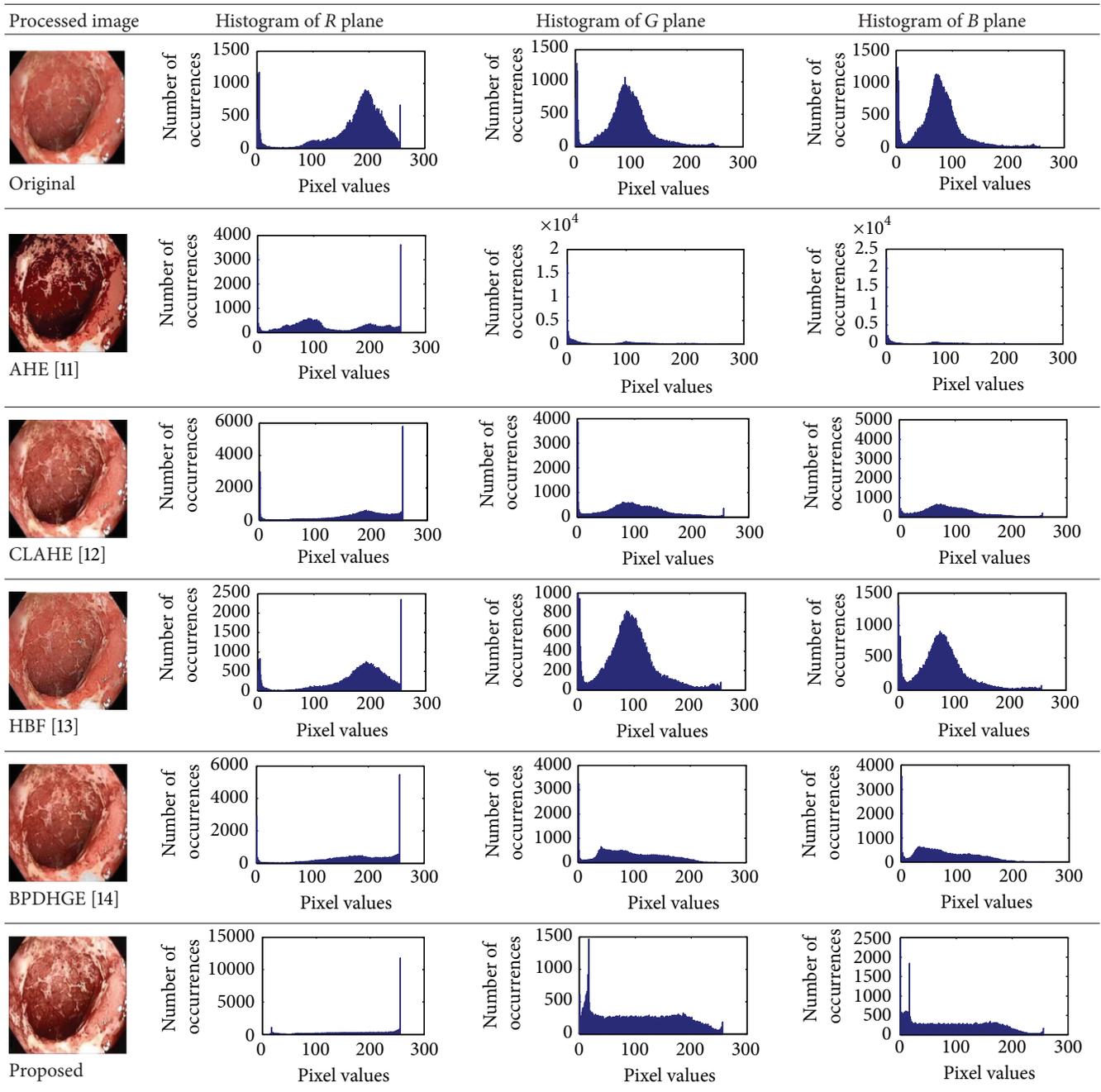
$$\begin{aligned}
 H(x_1, x_2, \dots, x_n) &= -\sum_{x_1} \sum_{x_2} \dots \sum_{x_n} P(x_1, x_2, \dots, x_n) \times \log_2 P(x_1, x_2, \dots, x_n). \\
 & \quad (15)
 \end{aligned}$$

First, we have showed the advantage of using proposed color reproduction in Table 9. Here, in image (a), we used the proposed image enhancement algorithm on the luminance plane and left the chrominance planes unchanged. In image (b), we applied proposed image enhancement on luminance

and color reproduction on the chrominance planes. From both images, it is noticeable that the image in (a) without color reproduction does not preserve brightness and shows imbalance saturation level. On contrary, the image in (b) with color reproduction has much balanced saturation and it preserves the overall brightness. It happens because  $YCbCr$  is a nonuniform and nonorthogonal color space. That is why we need to manipulate both luminance and chrominance in such a way that the correlation does not break and preserve the brightness along with the color saturation level. Additionally, our method achieves a higher entropy value, which means that it produces a more uniform histogram. The entropy value of image (b) is 7.6237 which is higher than that of image (a) that is 7.4961.

Table 10 shows the performance comparison with other related methods. In shows that the proposed method

TABLE 10: Comparison of the histogram of  $R$ ,  $G$ , and  $B$  planes in terms of uniform distribution.



produces images with enhanced and highlighted mucosa structures. The results are also summarized in Table 11.

**4.4. Color Similarity Test.** To validate the results statistically, the color similarity between the original and color reproduced images is evaluated using several performance metrics such as, CIE94 delta- $E$  color difference [40], mean structure similarity index (MSSIM) [41] and structure and hue similarity (SHSIM) [42]. The purpose is to show that our color reproduction method does not add any additional color. CIE94 is used to measure the color differences between

processed and original image in LAB color space. In CIE94,  $\Delta E_{94}^* \approx 2.3$  indicates that the color difference between two images is the lowest. MSSIM are used to measure color similarity in the chrominance planes in  $Y C_b C_r$  color space. SHSIM is used to measure the hue and structure similarity between processed and original image in HSV color space. Here, we have used 60 trial images to evaluate the color similarity index. The results are compared with other color reproduction methods and presented in Table 12. It can be seen that the average MSSIM and SHSIM indices are higher than others in our scheme with a color difference  $\Delta E_{94}^*$  close

TABLE 11: Comparisons of the measurement of uniform distribution based on entropy with other related works.

Image number	Measurement of entropy					Proposed
	Original	AHE [11]	CLAHE [12]	HBF [13]	BPDFHE [14]	
1	7.25	6.68	7.38	7.33	7.41	<b>7.46</b>
2	7.21	6.47	7.48	7.33	7.44	<b>7.57</b>
3	7.11	6.38	7.34	7.31	7.38	<b>7.51</b>
4	6.52	4.87	7.11	6.77	7.51	<b>7.70</b>
Average of 60 endoscopic images	7.41	7.01	7.68	7.31	7.75	<b>7.91</b>

TABLE 12: Color similarity assessment.

Similarity between two color images		$\Delta E_{94}^*$	MSSIM	SHSIM
Image 1	Image 2			
	Proposed	<b>2.97</b>	<b>0.9851</b>	<b>0.9992</b>
Original image	[21]	4.1	0.8605	0.8578
	[22]	0.6	0.8714	0.8001
	[23, 24]	3.01	0.9567	0.9942

TABLE 13: Comparisons of CEF indices with other enhancement works.

Image number	AHE [11]	CLAHE [12]	HBF [13]	BPDFHE [14]	Proposed
1	0.9620	1.1197	1.0355	1.0149	<b>1.7329</b>
2	0.9623	1.0989	1.0077	1.0037	<b>1.8897</b>
3	0.9446	1.1716	1.0751	1.0547	<b>1.8812</b>
4	0.9946	1.2009	1.1913	1.1003	<b>1.7443</b>
Average of 60 endoscopic images	0.9661	1.1574	1.0614	1.0411	<b>1.7477</b>

TABLE 14: Comparisons of CEF indices with other color reproduction works.

Image number	[23, 24]	[21]	[22]	Proposed
1	1.1059	0.5987	0.5118	<b>1.7784</b>
2	1.0991	0.7481	0.5997	<b>1.6599</b>
3	1.1007	0.6187	0.6001	<b>1.8413</b>
4	1.1972	0.5249	0.4991	<b>1.7749</b>
Average of 60 endoscopic images	1.1391	0.5149	0.4977	<b>1.5621</b>

to 2.3. All these values indicate that the colorized images are very close to the original images.

**4.5. Color Enhancement Factor (CEF).** We have also evaluated our scheme in terms of color enhancement. Here, we have used a no-reference performance metric called colorfulness matrix (CM) [43]. The CM measurement is based on the mean and standard deviations of two axes opponent color representation with,  $\alpha = R - G$  and  $\beta = (1/2)(R + G) - B$ . The metric is defined as

$$CM = \sqrt{\sigma_{\alpha}^2 + \sigma_{\beta}^2} + 0.3\sqrt{\mu_{\alpha}^2 + \mu_{\beta}^2}, \quad (16)$$

where  $\sigma_{\alpha}$  and  $\sigma_{\beta}$  are standard deviations of  $\alpha$  and  $\beta$ , respectively. Similarly,  $\mu_{\alpha}$  and  $\mu_{\beta}$  are their means. However, in our comparison, we have used the ratio of CMs between the enhanced and original image for observing the color enhancement factor (CEF). If  $CEF < 1$ , than the original image is better compared to the enhanced image in terms of color image enhancement. CEF with value 1 indicates that there is no difference between the enhanced and original image in terms of color enhancement. The results have been presented in Tables 13 and 14. Here we can see that CEF values of the proposed method are highest compared to other enhancement methods which indicates that our scheme performs better in terms of color enhancement. Figure 7 shows some reconstructed images.

TABLE 15: Comparison of simulation speed between proposed method and other related works.

Methodology	Image size	Step 1	Step 2	Total time (sec.)
		Enhancement time (sec.)	Color reproduction time (sec.)	
Proposed	$256 \times 256$	<b>1.909</b>	<b>20.614</b>	<b>22.523</b>
	$512 \times 256$	<b>2.597</b>	<b>40.119</b>	<b>42.716</b>
	$512 \times 512$	<b>3.182</b>	<b>81.779</b>	<b>84.961</b>
[23, 24]	$256 \times 256$	0.39	28.255	28.645
	$512 \times 256$	1.519	72.956	74.475
	$512 \times 512$	1.525	113.31	114.835
[21]	$256 \times 256$	—	64.461	64.461
	$512 \times 256$	—	117.921	117.921
	$512 \times 512$	—	232.146	235.149
[22]	$256 \times 256$	—	0.19	0.19
	$512 \times 256$	—	0.81	0.81
	$512 \times 512$	—	0.47	0.47

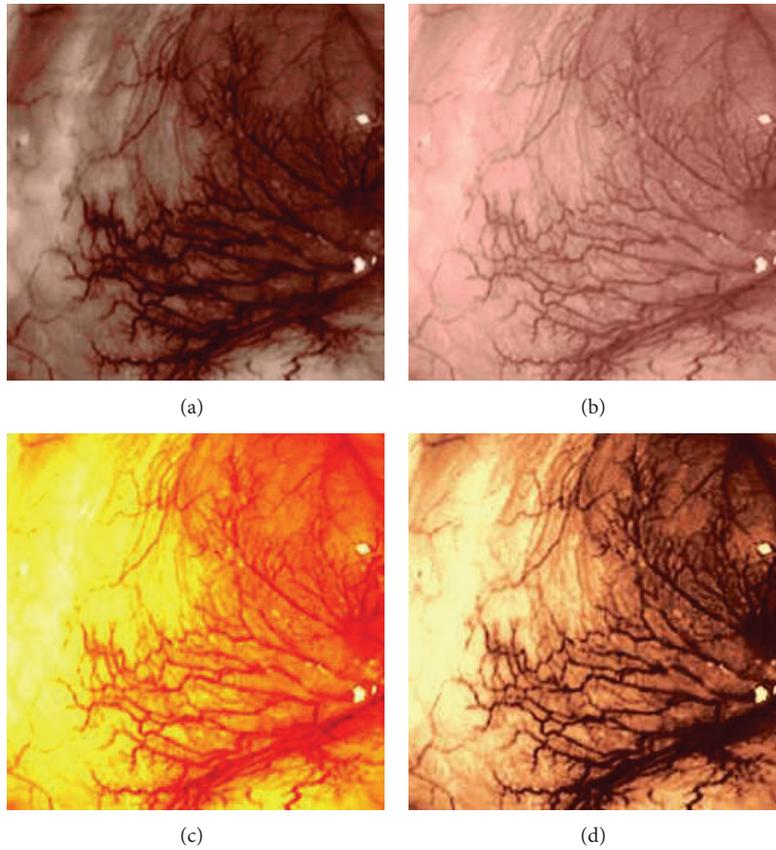


FIGURE 7: Enhanced color images using different color reproduction algorithms. (a) References [23, 24]. (b) Reference [21]. (c) Reference [22]. (d) Proposed.

**4.6. Algorithm Complexity.** The time required to generate an enhanced color image for different image sizes using the proposed method and other related works [21–24] are shown in Table 15. The experiment was conducted on a PC having Intel (R) Pentium(R) dual CPU @ 2.00 GHz and 6 GB of RAM. Here, it is noticeable that the proposed method is the fastest method including both image enhancement and

color reproduction. For an image of  $n$  pixels, the proposed algorithm has linear computational time complexity,  $O(n)$ . The average simulation time of proposed method for  $256 \times 256$  images is approximately 22 seconds and for  $512 \times 512$  images is approximately 85 seconds. The work in [21, 23, 24] have significantly higher execution time when compared with the proposed method. Although the execution time of [22] is

lower than ours, the quality of the color reproduction is much worse as shown in Figure 7.

## 5. Conclusion

In this paper, we have presented an image enhancement and color reproduction method for endoscopic images. The work focuses on enhancing the mucosa structures present in endoscopic image. The proposed color image enhancement is achieved in two stages: image enhancement at gray level followed by space variant chrominance mapping color reproduction. Image enhancement is achieved in two steps: adaptive sigmoid function and uniform distribution of sigmoid pixels. Secondly, space variant color reproduction is performed by generating a real color map by transferring and modifying old chrominance values either from theme image or input image. The quality of the generated enhanced colored images is evaluated using several standard performance metrics, which show that the features are highlighted on the new processed images.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

The authors would like to acknowledge Grand Challenges Canada Star in Global Health, Natural Science and Engineering Research Council of Canada (NSERC), Canada Foundation for Innovation (CFI), and Western Economic Diversification (WED) Canada for their support to this research work.

## References

- [1] A. Brownsey and J. Michalek, *Wireless Capsule Endoscopy*, The American Society for Gastrointestinal Endoscopy, 2010.
- [2] A. Brownsey and J. Michalek, *High Definition Scopes, Narrow Band Imaging, Chromoendoscopy*, American Society for Gastrointestinal Endoscopy, 2010.
- [3] H.-M. Chiu, C.-Y. Chang, C.-C. Chen et al., "A prospective comparative study of narrow-band imaging, chromoendoscopy, and conventional colonoscopy in the diagnosis of colorectal neoplasia," *Gut*, vol. 56, no. 3, pp. 373–379, 2007.
- [4] W. Curvers, L. Baak, R. Kiesslich et al., "Chromoendoscopy and narrow-band imaging compared with high-resolution magnification endoscopy in Barrett's Esophagus," *Gastroenterology*, vol. 134, no. 3, pp. 670–679, 2008.
- [5] S. Liangpunsakul, V. Chadalawada, D. K. Rex, D. Maglinte, and J. Lappas, "Wireless capsule endoscopy detects small bowel ulcers in patients with normal results from state of the art enteroclysis," *The American Journal of Gastroenterology*, vol. 98, no. 6, pp. 1295–1298, 2003.
- [6] K. Kuznetsov, R. Lambert, and J.-F. Rey, "Narrow-band imaging: potential and limitations," *Endoscopy*, vol. 38, no. 1, pp. 76–81, 2006.
- [7] S. Schmitz-Valckenberg, F. G. Holz, A. C. Bird, and R. F. Spaide, "Fundus autofluorescence imaging: review and perspectives," *Retina*, vol. 28, no. 3, pp. 385–409, 2008.
- [8] T. Kaltenbach, Y. Sano, S. Friedland, and R. Soetikno, "American Gastroenterological Association (AGA) Institute technology assessment on image-enhanced endoscopy," *Gastroenterology*, vol. 134, no. 1, pp. 327–340, 2008.
- [9] J. Pohl, A. May, T. Rabenstein, O. Pech, and C. Ell, "Computed virtual chromoendoscopy: a new tool for enhancing tissue surface structures," *Endoscopy*, vol. 39, no. 1, pp. 80–83, 2007.
- [10] S. J. Chung, D. Kim, J. H. Song et al., "Comparison of detection and miss rates of narrow band imaging, flexible spectral imaging chromoendoscopy and white light at screening colonoscopy: a randomised controlled back-to-back study," *Gut*, vol. 63, no. 5, pp. 785–791, 2014.
- [11] S. M. Pizer, E. P. Amburn, J. D. Austin et al., "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355–368, 1987.
- [12] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*, pp. 474–485, Academic Press Professional, 1994.
- [13] R. Srivastava, J. R. P. Gupta, H. Parthasarthy, and S. Srivastava, "PDE based unsharp masking, crispering and high boost filtering of digital images," *Communications in Computer and Information Science*, vol. 40, pp. 8–13, 2009.
- [14] D. Sheet, H. Garud, A. Suveer, M. Mahadevappa, and J. Chatterjee, "Brightness preserving dynamic fuzzy histogram equalization," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 4, pp. 2475–2480, 2010.
- [15] A. Carbonaro and P. Zingaretti, "A comprehensive approach to image-contrast enhancement," in *Proceedings of the 10th International Conference on Image Analysis and Processing (ICIAP '99)*, pp. 241–246, Venice, Italy, September 1999.
- [16] H. Ibrahim and N. S. P. Kong, "Brightness preserving dynamic histogram equalization for image contrast enhancement," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 4, pp. 1752–1758, 2007.
- [17] V. Magudeeswaran and C. G. Ravichandran, "Fuzzy logic-based histogram equalization for image contrast enhancement," *Mathematical Problems in Engineering*, vol. 2013, Article ID 891864, 10 pages, 2013.
- [18] M. S. Imtiaz and K. A. Wahid, "Image enhancement and space-variant color reproduction method for endoscopic images using adaptive sigmoid function," in *Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '14)*, pp. 3905–3908, August 2014.
- [19] L.-R. Dung and Y.-Y. Wu, "A wireless narrowband imaging chip for capsule endoscope," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 4, no. 6, pp. 462–468, 2010.
- [20] Olympus, "Evis Lucera Spectrum family brochure," 2014, [http://www.olympus.co.uk/medical/en/medical.systems/mediacentre/media\\_detail.7450.jsp](http://www.olympus.co.uk/medical/en/medical.systems/mediacentre/media_detail.7450.jsp).
- [21] T. Welsh, M. Ashikhmin, and K. Mueller, "Transferring color to greyscale images," *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 277–280, 2002.
- [22] V. Korostyshevskiy, *Grayscale to RGB Converter*, MATLAB Central File Exchange, 2006, <http://www.mathworks.com/matlabcentral/fileexchange/13312-grayscale-to-rgb-converter>.
- [23] M. S. Imtiaz, T. H. Khan, and K. A. Wahid, "New color image enhancement method for endoscopic images," in *Proceedings of the 2nd International Conference on Advances in Electrical Engineering (ICAEE '13)*, pp. 263–266, December 2013.

- [24] M. S. Imtiaz and K. Wahid, "A color reproduction method with image enhancement for endoscopic images," in *Proceedings of the 2nd Middle East Conference on Biomedical Engineering (MECBME '14)*, pp. 135–138, Doha, Qatar, February 2014.
- [25] Saruchi, "Adaptive sigmoid function to enhance low contrast images," *International Journal of Computer Applications*, vol. 55, no. 4, pp. 45–49, 2012.
- [26] A. Plaza, J. A. Benediktsson, J. W. Boardman et al., "Recent advances in techniques for hyperspectral image processing," *Remote Sensing of Environment*, vol. 113, no. 1, pp. S110–S122, 2009.
- [27] N. Hassan and N. Akamatsu, "A new approach for contrast enhancement using sigmoid function," *The International Arab Journal of Information Technology*, vol. 1, no. 2, pp. 221–225, 2004.
- [28] *Gastrolab—The Gastrointestinal Site*, 1996, <http://www.gastro-lab.net/index.htm>.
- [29] Atlanta South Gastroenterology, *Atlas of Gastrointestinal Endoscopy*, Atlanta South Gastroenterology, 1996, <http://www.endoatlas.com/index.html>.
- [30] A. Allen and D. Snary, "The structure and function of gastric mucus," *Gut*, vol. 13, no. 8, pp. 666–672, 1972.
- [31] R. Ishihara, T. Inoue, N. Hanaoka et al., "Autofluorescence imaging endoscopy for screening of esophageal squamous mucosal high-grade neoplasia: a phase II study," *Journal of Gastroenterology and Hepatology*, vol. 27, no. 1, pp. 86–90, 2012.
- [32] B. Lovisa, *Endoscopic fluorescence imaging: spectral optimization and in vivo characterization of positive sites by magnifying vascular imaging [Ph.D. thesis]*, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, 2010.
- [33] B. A. Wandell and L. D. Silverstein, "Digital color reproduction," in *The Science of Color*, S. K. Shevell, Ed., Optical Society of America, Washington, DC, USA, 2nd edition, 2003.
- [34] J. A. Yule, G. G. Field, and J. A. C. Yule, "Color reproduction," in *Color: An Introduction to Practice and Principles*, pp. 167–186, John Wiley & Sons, New York, NY, USA, 2012.
- [35] X. Xu, Y. Wang, J. Tang, X. Zhang, and X. Liu, "Robust automatic focus algorithm for low contrast images using a new contrast measure," *Sensors*, vol. 11, no. 9, pp. 8281–8294, 2011.
- [36] C.-H. Shen and H. H. Chen, "Robust focus measure for low-contrast images," in *Proceedings of the International Conference on Consumer Electronics (ICCE '06)*, pp. 69–70, January 2006.
- [37] B. Balas, L. Nakano, and R. Rosenholtz, "A summary-statistic representation in peripheral vision explains visual crowding," *Journal of Vision*, vol. 9, no. 12, pp. 1–18, 2009.
- [38] C. D. Manning and H. Schütze, *Foundations of Statistical Natural Language Processing*, MIT Press, Cambridge, Mass, USA, 1st edition, 1999.
- [39] H. Ney, S. Martin, and F. Wessel, "Statistical language modeling using leaving-one-out," in *Corpus-Based Methods in Language and Speech Processing*, vol. 2 of *Text, Speech and Language Technology*, pp. 174–207, Springer, Berlin, Germany, 1997.
- [40] A. R. Robertson, "Historical development of CIE recommended color difference equation," *Color Research & Application*, vol. 15, no. 3, pp. 167–170, 1990.
- [41] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [42] Y. Shi, Y. Ding, R. Zhang, and J. Li, "Structure and hue similarity for color image quality assessment," in *Proceedings of the International Conference on Electronic Computer Technology (ICECT '09)*, pp. 329–333, February 2009.
- [43] S. E. Süssstrunk and S. Winkler, "Color image quality on the Internet," in *5th IS&T/SPIE Electronic Imaging 2004: Internet Imaging*, pp. 118–131, International Society of Optics and Photonics, San Jose, Calif, USA, January 2004.

## Research Article

# A Unified Framework for Brain Segmentation in MR Images

**S. Yazdani,<sup>1</sup> R. Yusof,<sup>1</sup> A. Karimian,<sup>2</sup> A. H. Riazi,<sup>3</sup> and M. Bennamoun<sup>4</sup>**

<sup>1</sup>Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia, 54100 Jalan Semarak, Kuala Lumpur, Malaysia

<sup>2</sup>Department of Biomedical Engineering, Faculty of Engineering, University of Isfahan, Isfahan 81745, Iran

<sup>3</sup>Control and Intelligent Processing Center of Excellence, School of Electrical and Computer Engineering, University College of Engineering, University of Tehran, Tehran 14174, Iran

<sup>4</sup>School of Computer Science and Software Engineering, The University of Western Australia, Perth, WA 6907, Australia

Correspondence should be addressed to R. Yusof; rubiyah.kl@utm.my

Received 4 August 2014; Revised 7 November 2014; Accepted 18 November 2014

Academic Editor: Kayvan Najarian

Copyright © 2015 S. Yazdani et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Brain MRI segmentation is an important issue for discovering the brain structure and diagnosis of subtle anatomical changes in different brain diseases. However, due to several artifacts brain tissue segmentation remains a challenging task. The aim of this paper is to improve the automatic segmentation of brain into gray matter, white matter, and cerebrospinal fluid in magnetic resonance images (MRI). We proposed an automatic hybrid image segmentation method that integrates the modified statistical expectation-maximization (EM) method and the spatial information combined with support vector machine (SVM). The combined method has more accurate results than what can be achieved with its individual techniques that is demonstrated through experiments on both real data and simulated images. Experiments are carried out on both synthetic and real MRI. The results of proposed technique are evaluated against manual segmentation results and other methods based on real T1-weighted scans from Internet Brain Segmentation Repository (IBSR) and simulated images from BrainWeb. The Kappa index is calculated to assess the performance of the proposed framework relative to the ground truth and expert segmentations. The results demonstrate that the proposed combined method has satisfactory results on both simulated MRI and real brain datasets.

## 1. Introduction

Visualization and three-dimensional (3D) processing of medical images are rapidly growing fields of study. In particular accurate and robust technique for image segmentation is a research topic which has been one of the core problems in medical image analysis for years. In particular, the segmentation of brain MR images aiming to assign each voxel to a specific tissue class has received considerable attention.

Multimodality imaging techniques are valuable to medical and clinical studies, as well as other fields. Positron emission tomography [1], computed tomography (CT), magnetic resonance imaging (MRI), digital mammography, ultrasound or single photon emission computed tomography (SPECT), and X-ray provide effective ways for the representation of the subject's anatomy [2, 3]. High spatial resolution and good soft-tissue contrast in MR brain images and also recent

progress in MRI systems make them suitable for the realization of this goal [4, 5]. In medical and clinical research on brain structures, the description of tissue size is an important aspect of therapy that should be performed accurately.

Most procedures rely on a slice-by-slice interactive input of human knowledge that is very labor intensive and time consuming. These methods suffer from inter- and intraobserver variability [4]. Intraobserver variability occurs when the same users make various choices on different occasions, producing different results each time [6, 7]. Interobserver variation occurs when different users make different selections, which affect the segmentation results [8]. This generally leads to the need for reliable and accurate segmentation of MRI brain images and also to define tumors or lesions if present [9, 10].

In addition, a robust segmentation of lesions is a very important stage for diagnosing disease [11, 12], monitoring

treatment, investigating disease progress, and computer-integrated surgery. It is also of noticeable interest to study regional volumes of white matter (WM) and gray matter [13] across several developmental stages of the brain [5, 14, 15]. In the context of neuroimaging, automatic three-dimensional (3D) segmentation of brain MR images into WM, GM, and cerebrospinal fluid (CSF) has received an enormous amount of attention, as it is extremely important for quantitative analysis of MR images. In this paper, we used some techniques for brain segmentation into WM, GM, and CSF.

There has been a wide range of automatic segmentation techniques proposed in the literature. The main problems found in the automatic segmentation of MR images derive from the fact that the intensities of images are not necessarily constant for each tissue class [16].

Among fuzzy clustering algorithms, Fuzzy C-Mean is a powerful technique that has been extensively used in MR image segmentation [17] in which voxels are partially classified into various classes using different memberships for each class [18, 19]. Fuzzy C-Mean was first conceived by Dunn [20] and generalizes  $K$ -means algorithm to allow soft segmentation [21]. Pham extended the standard FCM technique to deal with brain MR images corrupted by bias field. The greatest drawback of FCM is its sensitivity to noise. MR images always include a considerable amount of noise, leading to further degradation with segmentation. Many extensions of FCM have been reported in the literature to overcome its drawbacks, but most of them still have some problems [22].

Many researchers segmented brain MRI by applying an artificial neural network (ANN). In comparison with FCM, the FCM algorithm was shown to be worse for abnormal brain with tumor, edema, and so forth and superior on normal brain.

Lemieux et al. have segmented brain MR images into WM, GM, and CSF using Gaussian mixtures and morphological operations [5, 23]. Homomorphic filtering techniques for eliminating the effect of the bias field have been commonly used because of their easy implementation. The problem is that this method is effective only on low contrast images and some researchers reported undesirable artifacts with this method.

One of the main drawbacks of classifiers and clusters is that they do not contain contextual information. The classification of a voxel is quite independent of all the other voxels. One solution for dealing with this problem is applying Markov random fields (MRFs), which is a statistical model in the group of random field methods. In the literature, 3D MRF has been used for tissue classification, which assumes a Gibbs prior to the Gaussian mixtures. It is equivalent to a Gibbs joint probability distribution, which is defined by an energy function [5, 24].

Generally classification techniques dealing with MR images can be divided into two categories: parametric and nonparametric methods. The parametric approaches usually make the assumption that the tissues of brain follow a Gaussian distribution. The statistical model parameters usually are estimated applying a maximum a posteriori (MAP), maximum likelihood method and the expectation-maximization (EM) algorithm that is used for the optimization process [25, 26].

An iterative algorithm based on the EM method algorithm was proposed by Wells III et al. [27]. The algorithm is also designed for eliminating anatomical features of the image, along with intensity nonuniformity field estimations. The advantages of the EM algorithm are its ease of implementation, conceptual simplicity, and also the fact that each of the iterations improves the results. A main problem of EM method is that it is based on a Gaussian distribution model for the intensity distribution of images, which is not true, especially for noisy MRI. In this paper, we proposed a modified EM method as an initial segmentation stage. The proposed EM algorithm overcomes the shortcoming of the standard EM technique using asymmetric Gaussian.

In addition machine learning algorithms have proven to yield desirable results in many cases. The SVM method is considered as a good candidate because of its high generalization performance without the need of prior knowledge, even when the dimension of the input space is very high [28]. The SVM was first proposed by Vapnik and has since attracted a high degree of interest within the research community in the category of machine learning. Some papers have reported that the SVM generally is more able to deliver higher performance in terms of classification accuracy than the other classification techniques as SVMs do not suffer the limitations of limited samples and data dimensionality [29]. In our study, support vectors, which are critical for classification, are created by learning from the training samples, which are extracted from the previous stages.

The key aspect of the proposed automatic framework is that we divided the segmentation task into three stages, each of which extracts a different set of constraints of the problem, and also the output of each stage simplifies the one which follows it. The first step of our method after preprocessing steps falls into the category of statistical segmentation techniques and provides an intensity-based classification. In this stage we modified EM algorithm for modeling the underlying distributions of WM, GM, and CSF, which is an initial classification step. Then we extracted textural features from the target areas that include both nonoverlapped and overlapped voxels. By applying feature extraction method based on cooccurrence features prior to SVM algorithm, a reliable class labeling for the image will be generated, thereby facilitating the SVM step. Subsequently, we used SVM for margin classification and segmentation enhancement. The goal of using SVM is to assign a label to each overlapped voxel of the brain borders and to enhance the segmentation result.

It is demonstrated that a robust and accurate segmentation approach can be achieved to find optimal segmentations. That is demonstrated through experiments on both real data and simulated images. The rest of this paper is organized as follows.

In Section 2 we present the new automatic method for segmentation of brain tissues that combines three techniques with some new ideas and that is more robust than its individual components. We give a step-by-step explanation for estimating model parameters. In Section 3, we present experimental results of the proposed technique. In this section we discuss issues regarding verification of medical image segmentation and also present a comparison of our results on

simulated and real database. The segmentation performance is evaluated for the proposed method. Section 4 contains discussion and concluding remarks.

The proposed combination method is an accurate and fast way to find optimal segmentations, given the intensity models which incorporate the spatial coherence assumptions.

## 2. Materials and Methods

Our classification method involves three steps: modified EM based segmentation method, feature extraction, and nonlinear classification, which are detailed next. In the first step before brain classification we attempted to extract the brain from MR images. To compensate for the inhomogeneity and partial volume effects the preprocessing steps are applied prior to actual segmentation, which are explained as follows.

*2.1. Skull Stripping.* The first task for MRI analysis is to define brain and nonbrain voxels. This work is concerned with the predominant brain tissues: WM, GM, and CSF. The measured signal intensities of these predominant tissues may overlap other tissues, such as bone, fat, skin, dura, and muscle. This problem complicates reliable brain segmentation. Brain surface skull stripping is one of the important preprocessing steps for MRI segmentation. In this paper, we removed the skull, scalp, and other extraneous tissues of brain images by using the Brain Surface Extractor (BSE).

*2.2. Image Nonuniformity Compensation.* The magnetic susceptibility variations in the MR images cause the intensity nonuniformities (bias field) that prevent description of voxel tissue content based exclusively on image intensity [30]. Consequently segmentation and quantitative analyses of MR images require bias field correction. We applied the Bias Field Corrector (BFC) software to each of the images after skull stripping with BSE. The BFC is utilized to compensate for the intensity nonuniformity [26]. Both BSE and BFC are implemented in BrainSuite package (<http://brainsuite.usc.edu/>).

*2.3. Partial Volume Estimation.* Partial volume estimation (PVE) is caused by the finite spatial resolution of imaging devices, due to the complexity of human brain anatomy. This phenomenon is created in MR images when more than one tissue type occurs in a voxel, and a voxel is a combination of different tissues, such as WM and GM [27, 28]. The partial volume effect blurs the intensity distinction between tissue classes at the edges of the two tissues. The estimation of the amount of each tissue type within each voxel has received considerable interest in recent years. The PVE is an important stage when a robust and accurate segmentation is needed.

We used the trimmed minimum covariance determinant (TMCD) technique for the estimation of the parameters of the PV model in this paper [31]. This technique is based on trimmed minimum covariance determinant parameter estimation and MRF based tissue classification [29]. We computed the fractional tissue values for each image using the PVC software.

The results are three images of the three primary tissue types of CSF, GM, and WM, respectively, whose elements

reflect the proportion of the corresponding tissue type in each voxel. Also the algorithm classifies the voxels into the three primary tissue types and their partial volume mixtures (CSF/background, CSF/GM, and GM/WM). According to the partial volume classification, the voxels belonging to the pure CSF or CSF/background are removed from the skull stripped volume and the remaining volume will be processed for decomposition.

*2.4. Registration and Atlas Alignment.* In the preprocessing step we performed a spatial registration (alignment) of the input images. Image registration is the operation of aligning images to relate corresponding features. For most kinds of image processing on two or more images, it is required that the images are aligned, so that one voxel position represents the same anatomical position in all images. We performed affine registration with 12 degrees of freedom [30–32].

*2.5. Brain Segmentation.* In this paper, a novel algorithmic framework is proposed, in which we integrated different types of information, MR intensity, textural features, voxel location, and relationship with neighboring voxels, to improve the overall segmentation performance.

The first step is extended EM algorithm to initially segment the brain into three tissues. The results of modified EM for initial segmentation are superior to standard EM. The reasons that we combine modified EM to the next stages are as follows.

- (i) Since the voxels of brain regions, especially at the edges and borders, are not defined by unique intensities in MR images due to the presence of artifacts and overlapped voxels, further processing is also needed to ensure robust segmentation. Due to mentioned problems, some of the voxels that have been segmented in the first stage have two tissue types such as GM, CSF or GM, WM.
- (ii) Since the intensity information which is used in the first step is not sufficient to have powerful segmentation, textural features and spatial relationships of voxels are investigated in the next steps.

In the second step we extracted some textural features of non-overlapped regions and also to improve the SVM training process, some features of overlapped voxels were extracted randomly. Finally SVM algorithm is applied to identify overlapped voxels using extracted features. We also used SVM classifier for the brain margin classification and segmentation enhancement. In other words in the SVM stage we applied nonoverlapped and overlapped voxels for training stage and overlapped voxels for testing stage. Using overlap voxels in addition to nonoverlapped regions for SVM training step leads to more accurate segmentation. Therefore the SVM stage improves the segmentation results. It represents the results of each method which are then refined with the next method. In the next section three steps are described. A general overview of our method is shown in Figure 1.

*2.6. EM-Based Algorithm.* The expectation-maximization algorithm (EM) is an algorithm to find missing data based on

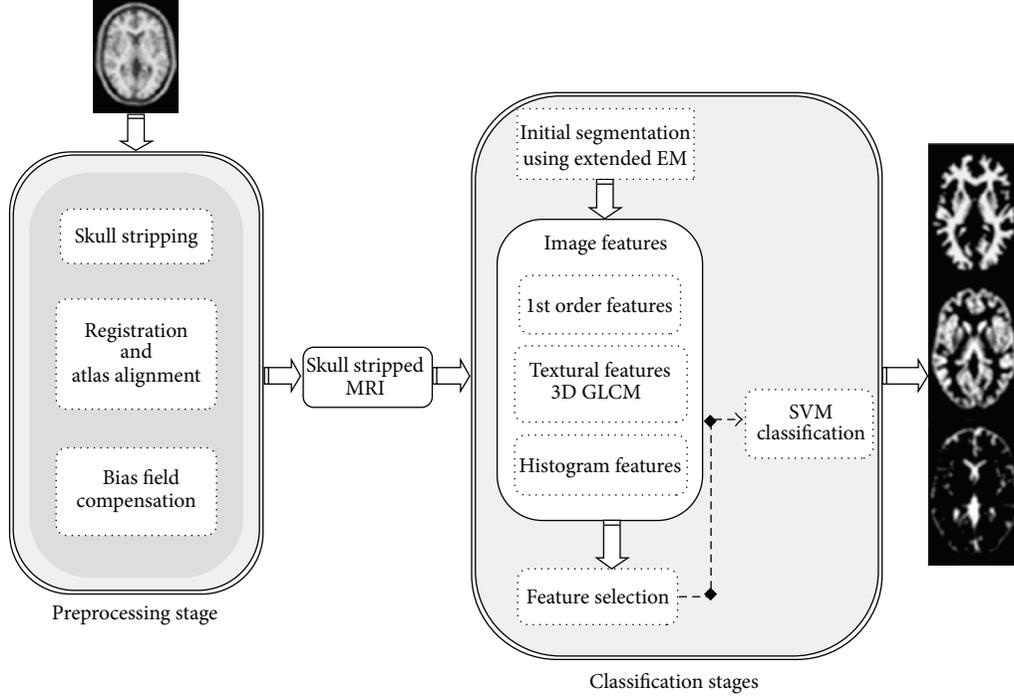


FIGURE 1: General overview of the proposed technique.

observed data and maximum likelihood parameter estimates. In automated model-based bias field correction of MR images of the brain the observed information is the intensities of the image, the missing data are the labels, and the parameters are the standard deviations and means of the Gaussian distribution, which is assumed for the intensity distribution of each tissue class. The EM algorithm is an iterative method, which interleaves two steps: the expectation step (E-step), which is the computation of posterior probabilities of each voxel belonging to each class (WM, GM, and CSF), and the maximization step with maximum likelihood estimation of the Gaussian distribution parameters. The maximum values are then taken as the new parameters [32].

The random observations are the intensity of the  $n$ th voxel in a brain region of interest [33]. Let  $\theta_i = \{\theta_1, \dots, \theta_k\}$  and  $\theta_i = (\mu_i, \sigma_i^2)$  demonstrate the parameters of Gaussian distributions. The Gaussian mixtures can be expressed to denote the parameters of  $K$  Gaussian distributions. The Gaussian mixtures of  $K$  tissue classes can be explained as

$$p(y_n | \omega, \theta) = \sum_{i=1}^k \omega_i \cdot p(y_n | \theta_i), \quad (1)$$

where  $\omega = \{\omega_i | 0 \leq \omega_i, \sum_i \omega_i = 1, i = 1, \dots, K\}$  are the weights of tissue classes, and

$$p(y_n | \theta_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left\{ -\frac{(y_n - \mu_i)^2}{2\sigma_i^2} \right\}, \quad (2)$$

$$\omega = \{\omega_1, \dots, \omega_k\}, \quad \theta = \{\theta_1, \dots, \theta_k\},$$

$$y = \{y_{nm} = 1, \dots, N\}.$$

The unknown parameters  $(\omega, \theta)$  can be estimated applying the maximum likelihood estimation (MLE) technique.  $N$  is the number of voxels, and  $y$  is the set of voxel intensities. The likelihood of the voxel intensity data with unknown parameters is shown as follows:

$$L(y | \omega, \theta) \equiv \prod_{n=1}^N p(y_n | \omega, \theta) = \prod_{n=1}^N \left\{ \sum_{i=1}^K \omega_i p(y_n | \theta_i) \right\}. \quad (3)$$

The ML estimates are created by solving for the parameters in the normal equations, which are derived from first partial derivatives of (3) that are equated to zero with respect to the unknown parameters  $(\omega, \theta)$ . The same notations are applied for the conditional probabilities and their estimates [4, 5].

Let  $p(i | y_n)$  be the posterior probability and the random observation belongs to the  $i$ th category. Consider

$$p(i | y_n, \omega^{(t)}, \theta^{(t)}) = \frac{\omega_i^{(t)} p(y_n | \theta_i^{(t)})}{\sum_j \omega_j^{(t)} p(y_n | \theta_j^{(t)})}. \quad (4)$$

The expectation step updates the posterior probability given the latest estimates of unknown parameters  $(\omega, \theta)$  which is in the  $t$ th iteration. The likelihood equations admit the posterior probability  $p(i | y_n, \omega^{(t)}, \theta^{(t)})$ . The maximization

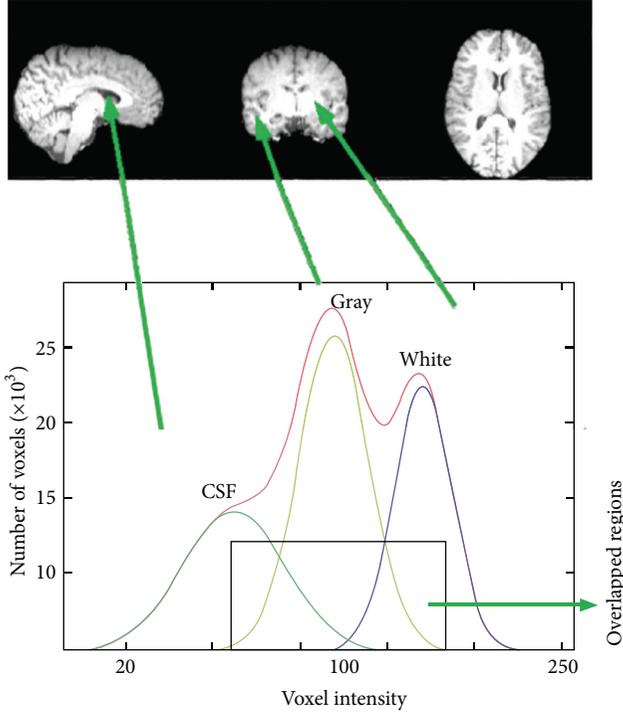


FIGURE 2: Illustration of symmetric Gaussian distribution and overlapped regions in the histogram.

step estimates  $\omega$  and  $\theta$  by inserting (4) into these equations [7]:

$$\begin{aligned} \omega_i^{(t+1)} &= \frac{\sum_n p(i | y_n, \omega^{(t)}, \theta^{(t)})}{\sum_j \sum_n p(j | y_n, \omega^{(t)}, \theta^{(t)})}, \\ \mu_i^{(t+1)} &= \frac{\sum_n p(i | y_n, \omega^{(t)}, \theta^{(t)}) \cdot y_n}{\sum_n p(i | y_n, \omega^{(t)}, \theta^{(t)})}, \\ \sigma_i^{2(t+1)} &= \frac{\sum_n p(i | y_n, \omega^{(t)}, \theta^{(t)}) \cdot (y_n - \mu_i^{(t+1)})^2}{\sum_n p(i | y_n, \omega^{(t)}, \theta^{(t)})}, \end{aligned} \quad (5)$$

where  $\mu_i$  and  $\sigma_i$  are the mean intensity and the standard deviation of the  $i$ th class, respectively. The algorithm simply continues iterating between E-step and M-step until  $\theta^{(t)}$  convergence to local maxima.

The segmentation result is then obtained by allocating the  $n$ th voxel to its tissue class with the maximum posterior probability in (4) [4, 5].

**2.7. Modified EM.** Image classification based on the EM method essentially models the intensities of an image as a finite mixture of  $K$  tissue classes. The classification based on standard EM may not recognize individual tissue types accurately. In other words the main shortcoming of the EM based techniques is that they are based on symmetric Gaussian distribution model for the intensity distribution of brain images (see Figure 2). That is not true in the real MRI, especially for noisy images. In real images the estimated

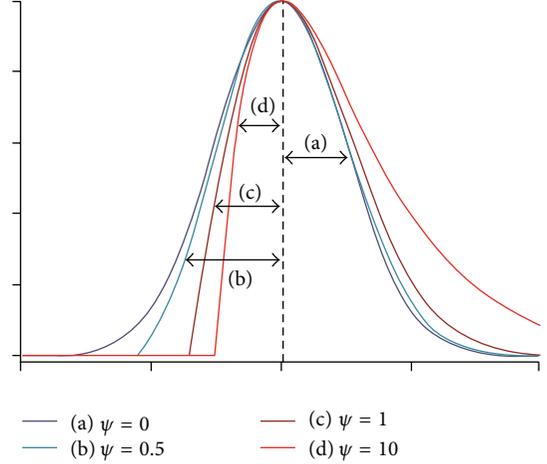


FIGURE 3: Skewed shapes of asymmetric distributions.

Gaussian distribution is not well fitted by the original histogram [5] because they do not have three normal Gaussians due to the existence of noise, artifacts, and overlapped Gaussians in the histogram (Figure 2).

Usually in real MR images the standard deviations of two sides of estimated Gaussian are different from each other. Thus the intensity distributions of brain tissues can vary asymmetrically in these images. Consequently the intensity of individual tissues may display skewed or spread shapes between brain images that may not be well fitted by a Gaussian distribution (Figure 3). For example, the CSF intensity on T1 brain images usually spreads across a wide range at the lower end of the histogram and displays an overlap with the GM tissue [5]. In this kind of asymmetric Gaussian distribution, if we use standard EM, the estimated standard deviation of one side is true and the other side is not accurate. Therefore the estimation of parameters is not completely acceptable.

Extended EM using asymmetric distribution is explained in detail as follows.

As mentioned before, EM algorithm is used to estimate mean and standard deviation of Gaussian distribution in two different steps (E-step and M-step) to reach the optimum Gaussian model. In the extended model, the asymmetric Gaussian has observation variable  $x$ , latent variable  $z$  ( $z \in R^d$ ), and the orthonormal matrix  $\phi \in R^{d \times d}$ :  $x = \phi z$  [34]. The latent variable is a different point between asymmetric Gaussian and standard EM method. The following distribution defines probability density functions for each  $z$  element. Consider

$$\mathcal{A}(z_i; \mu_i^z, \sigma_i^2, r_i) \equiv \frac{2}{\sqrt{2\pi}} \frac{1}{\sqrt{\sigma_i^2} (r_i + 1)} \begin{cases} \exp\left(-\frac{(z_i - \mu_i^z)^2}{2\sigma_i^2}\right) & \text{if } z_i > \mu_i^z, \\ \exp\left(-\frac{(z_i - \mu_i^z)^2}{2r_i^2 \sigma_i^2}\right) & \text{otherwise,} \end{cases} \quad (6)$$

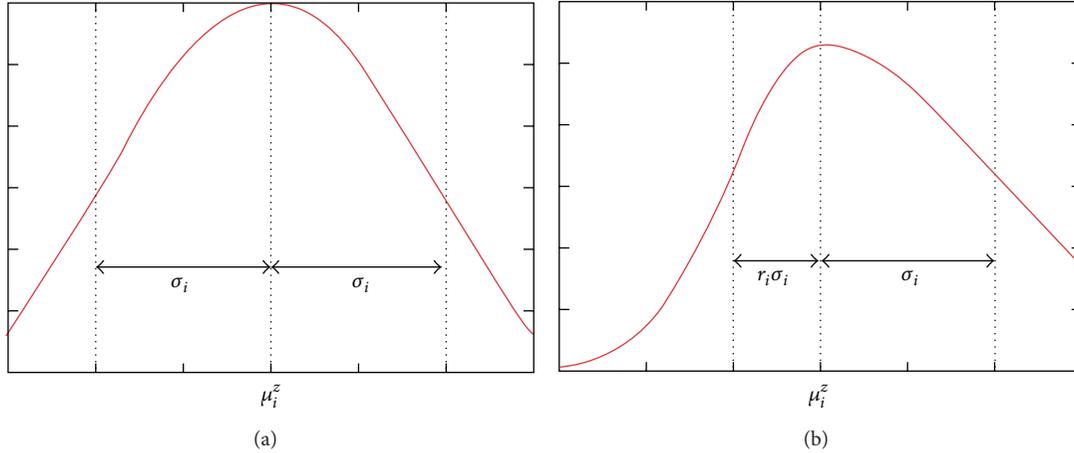


FIGURE 4: From (a) to (b) symmetric estimated Gaussian and asymmetric Gaussian of real image.

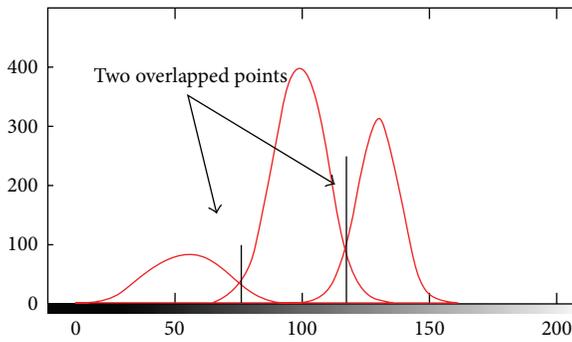


FIGURE 5: Illustration of two overlapped point.

where  $\mu_i$  and  $\sigma_i$  are estimated using standard EM and  $r_i$  is the linear coefficient between the standard deviation of two sides. Equation (6) demonstrates that the mentioned density model has an asymmetric distribution (see Figure 4). In this study, we suppose that the  $\mu_i$  and  $\sigma_i$  are the mean and standard deviation of one side of Gaussian, respectively, and the relationship between two standard deviations is as follows:

$$\sigma_{\text{second-side}} = r \cdot \sigma_{\text{first-side}}. \quad (7)$$

As mentioned before brain MR images have a mixture of three Gaussian distributions (7), each of which is estimated separately in this study. To compute the error, we separate each distribution in the histogram. Therefore, before error estimation we simply split the whole histogram into three separate Gaussians. These three Gaussians have overlapped at two points, in which sign of the gradient is changed and these two points are also local minimum (see Figure 5). It means that concavity at these two points is positive. Thus the overlapped points are the points in which the sign of first derivative is changed and sign of the second derivative is positive. This is an easy gradient-based method to separate three Gaussian distributions. In this step, we have two overlapped points in the histogram that give three distributions or three classes, which are WM, GM, and CSF.

In the next step, for each Gaussian, we fixed the estimated  $\mu_i$  and  $\sigma_{\text{first-side}}$  as the final value and then we determined the estimated standard deviation of the other side. To define the final value of  $\sigma_{\text{second-side}}$ , we applied an iterative error-based technique between the estimated Gaussian and real image histogram. Since we assumed that the distribution is (6), the standard deviation of one side of each Gaussian is proportional to the standard deviation of the other side (see (7)). In the next step, we supposed that the initial value of  $r$  is 1 ( $r = 1$ ); then  $r$  value is increased progressively. “ $r$ ” value should be increased step by step. The increasing trend of “ $r$ ” will be continued till the error reaches the threshold value. Experimentally this step is fixed to 10%. Smaller steps increase the processing time and accuracy but the percentage of increased accuracy is not significant. We defined the threshold value using RMS threshold method [35].

The error is calculated for each  $r$  that is a natural or real number. If error is increased dramatically by increasing  $r$ , the amount of  $r$  should be decreased from the initial value. If increasing  $r$  reduces the amount of error this process will be continued till the error has reached the threshold value (minimum error). The minimum calculated error demonstrates the best value of  $r$  and consequently the final value of standard deviation.

However the method requires a large number of iterations to reach the defined threshold or convergence. In other words the algorithm starts with some initial value of the parameters, one cycle between the E- and M-steps until  $\theta^{(t)}$  converges to a local maxima. In theory, the EM based methods are guaranteed to converge and they perform a ML estimation of the model parameters at a fast convergence rate. To accelerate the convergence rate of our algorithm, we provide a stopping criterion using the RMS error [35]. By applying consecutive RMS errors we are able to find when to progress or stop the procedure [35].

The minimum error is calculated for three tissue classes. In other words, based on the estimated error, we changed the standard deviation to obtain the best value. This is an iterative technique to reach the optimum parameters.

To compute the error, we considered all intensities of brain MR images. In other words, the estimated distribution and real image histogram should be compared at each intensity. The amount of error is the average difference between estimated distribution and original distribution at related intensity:

$$\text{Error} = \frac{1}{N} \sum_{i=1}^N (f_i - g_i), \quad (8)$$

where  $N$  is the number of intensities that consists of Gaussian distribution,  $f_i$  is the number of voxels in related intensity of real histogram, and  $g_i$  is the number of voxels in related intensity of estimated histogram. Consequently, in this section minimum error and optimum parameters are estimated. Finally, it is repeated two more times to provide the three distributions that best fit the histogram.

After the modified EM converges, the parameters that maximize the likelihood function are then applied to segment brain image into three tissue classes. In this step we compute the membership probability with the estimated parameters. The initial brain volume is updated by adding the voxels, which are labeled as brain regions (WM, GM, and CSF) based on the value of membership probability. The process continues by checking every neighboring voxel of an already labeled one, until the left and right brain volumes remain unaltered. The outermost voxels of corresponding unaltered brain volumes provide the final left and right brain borders. The coordinates of each labeled voxels are stored to avoid double-checking of neighboring voxels during the brain volume updating.

The modified algorithm is summarized as follows.

- (1) Choose the number of Gaussian distributions and separate them based on the gradient-based method.
- (2) Select one Gaussian distribution in the split histogram.
- (3) Parameter initialization is as follows.

E-step: estimate distribution over labels given a certain fixed model.

M-step: choose new parameters for model to maximize expected log-likelihood of observed data and hidden variables.

Outputs: these are  $\mu_i$  and  $\sigma_{\text{first-side}}$ .

- (4) Assume the following.
  - (a) The estimated mean is correct and the  $\sigma_1$  is the standard deviation of one side.
  - (b) Standard deviation of the other side is  $\sigma_2 = r \cdot \sigma_1$ .

$r$  changes 10 percent at each calculation.

The defined threshold determines when calculation should be stopped.

- (5) Calculate error between the estimated distribution and real image histogram and obtain the optimal value of  $r$ .

If the error has reached threshold (an acceptable error that is defined in the first stage) then estimation is complete.

If error is more than threshold, estimation will be continued.

- (6) Continue steps (4) and (5) until the convergence of the sequence of parameters is reached.
- (7) Keep the estimated distribution and estimate the next separated distribution.
- (8) Compute the membership probability with the estimated parameters.
- (9) Assign each voxel to the  $K$ th class.

However, because of the intensity similarity between GM and CSF or CSF and WM and overlapping problem in the brain MR images the next improvement stage is required to have powerful segmentation. To have a robust and accurate segmentation in the next step we will extract some textural features from the image and finally use SVM to improve the classification process.

**2.8. Feature Extraction.** The goal of feature extraction is to reduce the original dataset by extracting the most important features. Choosing the optimal features has a strong effect on classification results. Image intensities are the most prominent features for image segmentation. Using intensity information as the only features in MR image is not sufficient due to several reasons.

- (i) In some scans, the nonbrain voxels have a similar intensity to GM, WM, and CSF.
- (ii) The intensity of constructing brain tissues varies among different slices.
- (iii) In some slices, the intensity of different tissues is similar.

Therefore, we carried out texture analysis for describing texture of the images to have adequate features for accurate segmentation. We also extracted useful features such as first- and second-order texture information in this study to have an appropriate segmentation for all cases.

In the previous section voxel labeling is initially applied on each voxel of the brain using intensity information with some new ideas. Since overlapped voxels (voxels of brain borders) have two labels (GM, CSF or GM, WM) instead of one, to compensate for this problem, the overlapped voxels should be classified to identify which classes they exactly belong to. Thus, in this section to have robust and accurate brain segmentation, each overlapped voxel and its 18-connected neighbors are used as input for 3D statistical features extraction technique, which is an improvement stage. In other words the input of 3D GLCM is the target area in the rectangular region of interest [33] that is demonstrated in Figure 6.

One of the important issues in the field of image analysis is the question of how to determine the texture differences of complex images. These differences are often due to

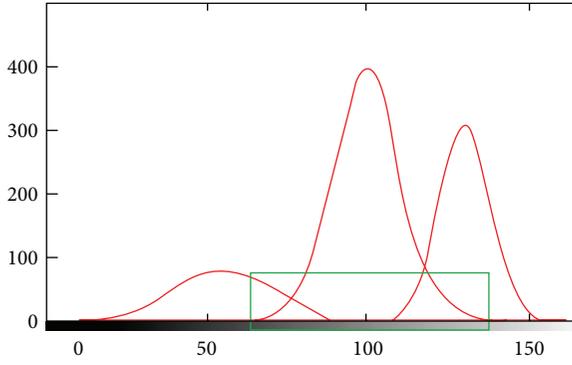
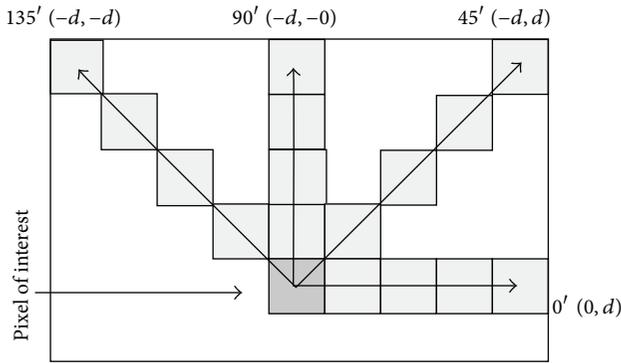


FIGURE 6: Target area or input data for 3D GLCM.

FIGURE 7: The spatial relationships of pixels, which are defined by the array of offsets, and  $d$  represents the distance from the pixel of interest.

the relative emplacement of pixels of various intensities. One way to describe these differences in the spatial relationships of voxels is using a GLCM.

The objective of this work is to generalize the concept of cooccurrence matrices to  $n$ -dimensional Euclidean spaces and to extract more features from the matrix. The GLCM matrix, defined as  $G_d^\phi(i, j)$ , is a square matrix (size  $N$ ), where  $N$  is the total number of voxels in the window and  $(i, j)$  entry represents the number of cooccurrences of gray levels  $i$  and  $j$  for voxels separated at a distance  $d$  in direction  $\Phi$ . In other words, the GLCM provides information on how often a gray level occurs at different directions. Usually, four directions are considered in the 2D case (Figure 7):  $\Phi = 0^\circ$ ,  $\Phi = 45^\circ$ ,  $\Phi = 90^\circ$ , and  $\Phi = 135^\circ$ , but in 3D images 13 directions are considered.

**2.9. 3D GLCM.** In this paper, we proposed 3D GLCM for feature extraction. Therefore, instead of square window ( $W \times W$ ), we considered cubes of size  $W \times W \times W$ . Selecting the window size is one of the main issues with this step, as it can define the discrimination capabilities of the extracted features. The choice of the window size plays an important role in the segmentation process. A small window decreases the computational burden and also enables resolution to capture the texture. Furthermore, large windows capture textural characteristics, but they increase the processing requirement

and memory. Moreover, the smaller windows reduce the processing time and make the results more accurate and vice versa for bigger windows. This way, we chose  $21 \times 21 \times 21$  windows as a trade-off between resolution and performance.

GLCM computation can be generalized as

$$G_d^\phi(i, j) = \sum_{z=1}^{V_z-d_z} \sum_{y=1}^{V_y-d_y} \sum_{x=1}^{V_x-d_x} \begin{cases} 1, & \text{if } (Q(x, y, z) = i) \\ & \wedge (Q(x + d_x, y + d_y, z + d_z) = j), \\ 0, & \text{otherwise,} \end{cases} \quad i, j = 1, \dots, N, \quad (9)$$

where  $v = (v_x, v_y, v_z)$  is the position of the voxel,  $N$  is the number of gray levels present in the images or subimages considered for GLCM calculation, and  $d = (d_x, d_y, d_z)$  is the distance in each direction.

The GLCM is a well-established tool for characterizing the spatial distribution, which includes second-order statistics of gray levels in an image. Second-order statistics are the texture of the image as they take into account the relationship among voxels in a window. An element at location  $(i, j)$  of the cooccurrence matrix signifies the joint probability density of the occurrence of gray levels in a specified direction  $\Phi$  and specified distance  $d$  from each other. The 3D cooccurrence matrix stores the number of cooccurrences of pairs of gray levels  $i$  and  $j$ , which are separated by a distance  $d$  (in this study,  $d = 1, 2, \dots, 5$  voxels) in 13 directions of the voxel of interest (VOI). In this paper, for each distance ( $d$ ) thirteen 3D cooccurrence matrix features were calculated from a sliding window ( $21 \times 21 \times 21$ ) within the brain volume, such as; angular second moment, contrast, correlation, variance, inverse different moment, and so forth.

In addition two first-order texture features (mean and standard deviation of each feature) over the thirteen cooccurrence matrices (corresponding to 13 directions) are calculated, comprising a total of 26 GLCM-based features for each distance  $d$ . In total, 130 features were calculated per VOI.

**2.10. Feature Selection.** As presented in previous sections second-order (textural) and first-order and histogram-based features are extracted from the image. Since using all the features does not provide the best results the next priority is to choose the subset of features most likely to recognize one tissue class from another. The challenge is that even a modest GLCM method with 3D and 4 $\theta$  values can create many more textural features than are suitable for the number of cases that will be subjected to classification. There are a number of techniques available for dimensionality reduction of features.

In this paper Stepwise Discriminant Analysis (SDA) that is a statistical approach is used to reduce the dimensions of the feature [33]. Discriminant Function Analysis undertakes the same task as multiple linear regressions by predicting the outcome. Multiple linear regression is limited to cases where

the dependent variable on the  $y$ -axis is an interval variable. Thus, the combination of predictors will create the estimated mean numerical  $Y$  values for given values of weighted combinations of  $X$  values. Discriminant Analysis (DA) is an earlier alternative to logistic regression that is recently mostly used.

After feature selection step the subset of features has been used to analyze the images on real and simulated database and to provide a powerful segmentation using the last step (SVM). We used SVM classifier especially for brain margin classification and segmentation enhancement.

*2.11. Segmentation Enhancement Using SVM.* In this paper brain pattern identification and characterization is achieved by initially classifying the brain volume into three classes based on the extended EM method. Due to the existence of artifacts and overlapped regions in the histogram of brain images the extended EM method is not enough for accurate segmentation. Therefore a feature extraction and then SVM classification are performed to obtain satisfactory results. For images without artifact the SVM step does not change the segmentation results. The problem is that all MRI images have artifacts to some degree and due to existence of these artifacts an improvement stage is necessary. Moreover as mentioned before, in brain images especially in borders there are overlapped voxels in the histogram. The existence of these overlapped voxels is inevitable in MRI and they make the segmentation inaccurate.

In our case, most of the false positive and false negative of WM are clearly located in the brain borders. In particular false positive and false negative of WM in brain borders may have a large influence on the relatively small total volume of WM. This problem is also similar for gray matter and CSF. In this study a postprocessing step is applied to compensate this problem. In other words, in the target area of Figure 6, when two Gaussians join each other, false positive and negative reach the maximum value and in the tails of Gaussians the false positive and false negative are decreased. In these areas the problem is that each overlapped voxel has two labels instead of one label and this problem degrades the accuracy of algorithm. In this paper to solve this problem we used SVM classifier as a postprocessing step, which is also a well known method to border identification.

SVM is currently considered a state-of-the-art method to solve binary classification problems. Because of generalization ability, SVMs have experienced great success in different applications [3, 36]. Since the SVMs attempt to maximize the separation margin, the generalization performance does not drop considerably even when the training data is scarce. SVMs work well for classification of the objects, which are not linearly separable. These objects are mapped into a high-dimensional feature space through kernel transformation. We can also replace or combine SVM with other classifiers to have better segmentation results [37].

Specifically in this study a support vector machine classifier is employed as an enhancement stage for segmentation to assign a label to each overlapped VOI sample of each class. This method is also used to rank computed features from the extracted features.

SVM is a linear discriminate classifier, which was developed in statistical machine learning theory by Vapnik as a linear binary classifier. In this stage SVM classifiers are trained for each brain tissue based on the set of extracted features from the target area. Most features are extracted from nonoverlapped regions. In addition to improving the SVM training process, some features of overlapped voxels are also extracted. In this section support vectors are briefly described.

The SVM classifiers require a training step to define a separating hyperplane for the data in the feature space. These hyperplanes separate various tissue classes so that the margin between the classes is the maximum margin.

The appealing characteristic of SVM algorithms is that they offer the possibility to apply a kernel function for transforming the data into a higher-dimensional feature space ( $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$ ). The kernel makes the data linearly separated with a maximum margin. For soft margin classification we used slack variables  $\xi_i$ . In this paper, to enable nonlinear decision functions, we used radial basis function (RBF) kernel for parameter selection of SVM classifier. Since the SVM algorithm is designed for two-class classification, to enable multiclass classification the classification is extended by one-against-the-others strategy. The SVM classifier assigns a label of brain tissue using extracted features. The features are extracted from a  $(21 \times 21 \times 21)$  VOI that is centered at the voxels being labeled. As mentioned in Section 2.8, the choice of the window size plays an important role in the classification process, as it may determine the discrimination capabilities of the extracted features.

In the SVM step, the sliding window moves along target area and labels the overlapped voxels in this region. SVM input involves overlapped and some nonoverlapped voxels in target area.

We performed the training process in two steps. In the first step, each subject is trained individually. To improve the training process we used overlapped and some nonoverlapped voxels as training data and overlapped voxels as test data. In the second step, we used all subjects to have an accurate and robust classifier. Because the problem is in the region of overlapped voxels and after the first step of segmentation, most of them have two labels, and we used these voxels for testing step. Finally in testing step voxels labeling is performed completely and each voxel belongs to one class.

In this section 12 subject of T1-weighted images of IBSR and 12 samples of BrainWeb datasets were applied to the training process. We used 8 subjects as training data and 4 remaining subjects to test the performance of the training process in each dataset. SVM training requires fixing the penalty term for misclassifications ( $C$ ). In this series of experiments, the  $C$  is set to 100. With an appropriate selection of metric within the RBF kernel, the “leap” in implementation did not occur, as normally expected by using RBF kernels.  $\chi^2$  or Laplacian RBF kernels decrease the Gaussian RBF error rate from around 30% down to less than 10%. This improvement is not only due to the selection of the proper metric, but also due to the suitable generalization of SVMs. The SVM classifier was trained for a total of 10 000 samples per training brain image that were randomly selected from the provided brain mask.

### 3. Experimental Results and Discussion

In this study to evaluate the performance of proposed technique we accomplished two sets of experiments, one on simulated images and another on real data. Since in these cases the ground truth (anatomical model or expert segmentation) is available, it is feasible to have a quantitative evaluation of the performance of method under different conditions and compare the results with the other state-of-the-art methods. To evaluate the performance of proposed algorithm on real images we compared the result with expert segmented images from IBSR dataset and finally compared the result ( $K$  indexes) with the other state-of-the-art methods. The evaluation result is presented in the next section.

Reliability of data collection is a component of overall confidence in an algorithm accuracy. The importance of rater reliability lies in the fact that it demonstrates the extent to which the data collected in the research study are true demonstration of the variables measured. Measurement of the extent to which data raters assign the same value to the same variable is called interrater reliability [38]. Although there have been different techniques to measure interrater reliability, it is measured as percent agreement, computed as the number of agreement scores divided by the total number of scores [38].

For this reason, in some studies in the literature the standard Jaccard similarity index of images is calculated. This metric measures the similarities between the two sets such as  $S1$  and  $S2$  as the ratio of the amount of their intersection divided by the amount of their union using (10) [5]. Two sets of  $S1$  and  $S2$ , indicating the created and gold standard segmentations, respectively. Consider

$$JS(S1, S2) = \frac{|S1 \cap S2|}{|S1 \cup S2|} * 100\%. \quad (10)$$

The other metric usually applied to compare the set similarity is Cohen's Kappa statistic or simply Kappa coefficient [39], which is defined as

$$\begin{aligned} k(S1, S2) &= \frac{|S1 \cap S2|}{1/2 (|S1| + |S2|)} \\ &= \frac{|S1 \cap S2|}{|S1 \cup S2| - 1/2 (|S1 \setminus S2| + |S2 \setminus S1|)} * 100\%. \end{aligned} \quad (11)$$

The Kappa is one of the most frequently used statistics in the literature to test interrater reliability [38]. This metric demonstrates that this is a special case of the  $k$  index, suitable for evaluation of image segmentation algorithm [39].  $K$  metric is larger than the Jaccard metric, except at 1 and 0. These metrics are related to each other by the function [40]

$$k = \frac{2J}{J+1}. \quad (12)$$

Both metrics agree that 1 means the two sets are similar and that 0 means the two sets are dissimilar. For the purpose of comparison, the mentioned metrics are consistent. For example, an increase in the  $k$  index means an increase in the Jaccard metric [40]. In this paper, the Kappa coefficient is defined for both phantom and real datasets and the results are presented in the next sections.

**3.1. Simulated Brain MRI.** As the ground truth is an image that is not known for the real data, the proposed algorithm has been first evaluated on simulated images. Knowing the anatomical model (ground truth) we can have a quantitative evaluation of the performance of the different methods and also compare them.

BrainWeb is a dataset providing simulated brain MR images for different acquisition parameters and acquisition modalities like T1 and T2 [40]. We applied the simulated MR scans of the head that are generated using the BrianWeb simulator (available from <http://www.bic.mni.mcgill.ca/brainweb/>) produced by the McConnell Brain Imaging Center at the Montreal Neurological Institute (MNI) [41]. Each MR image is provided with a ground truth that provides main tissue class labels for each voxel. For the technique, the considered BrainWeb images have been chosen with classical acquisition parameters (with respect to a standard brain MRI acquisition), namely, by considering T1-weighted images, with 1 mm resolution. The repetition time is equal to 18 ms and the echo time has been set to 10 ms. We used 18 synthetic volumes of  $181 * 217 * 181$ . BrainWeb dataset has six degrees of noise contamination (i.e., pn0, pn1, pn3, pn5, pn7, and pn9) and the bias field (intensity nonuniformity) can be specified at three different levels (i.e., rf0, rf20, and rf40). For both our labeled results and the ground truth labeling, we generated three-class labeling (see Figure 8). The Kappa coefficient is computed for WM and GM tissues for each volume compared to ground truth [42].

To point out the contribution of the proposed method, we compared the proposed method with fuzzy and nonfuzzy methods with different Rician noise and 20% inhomogeneity as shown in Table 1. The results of each technique are averaged over the 18 volumes. The fuzzy methods are FCM [40] and NL-FCM [43] and nonfuzzy methods are EM, SPM 5 [4, 25, 44], HMC [45], and Fast [45]. SPM5, FCM, EM, and Fast are free available reference software for brain MRI segmentation. We carried out experiments to define the robustness to noise for the proposed technique with BrainWeb T1-weighted images.

In this study, as indicated in the first row of table, the  $k$  index of each method is defined by percent value. The percent value is computed by multiplying the numeric value of the  $k$  indexes by 100.

The average Kappa indexes over all 18 volumes in different methods for WM and GM are shown in Figure 9. As it can be observed from Table 1 and the graphs in Figure 9 the proposed method has notable superiority over the other methods specifically in image volumes that are seriously contaminated by random noise (i.e., pn0–pn9). For example, the presented Markov random field and Fast methods in the table are superior to our method in low-level noise, but in high-level noise the proposed method is superior. In addition Table 1 shows that the FCM based algorithms are not a reliable method in noisy image applications. The proposed method also presents satisfactory results in comparison with standard EM due to accurate demonstration of intensity distribution and using other features for segmentation. The average Kappa indexes in different levels of noise for WM are EM = 88.36, SPM 5 = 91.07, FCM = 91.4, HMC = 94.6, NL-FCM = 90.86,

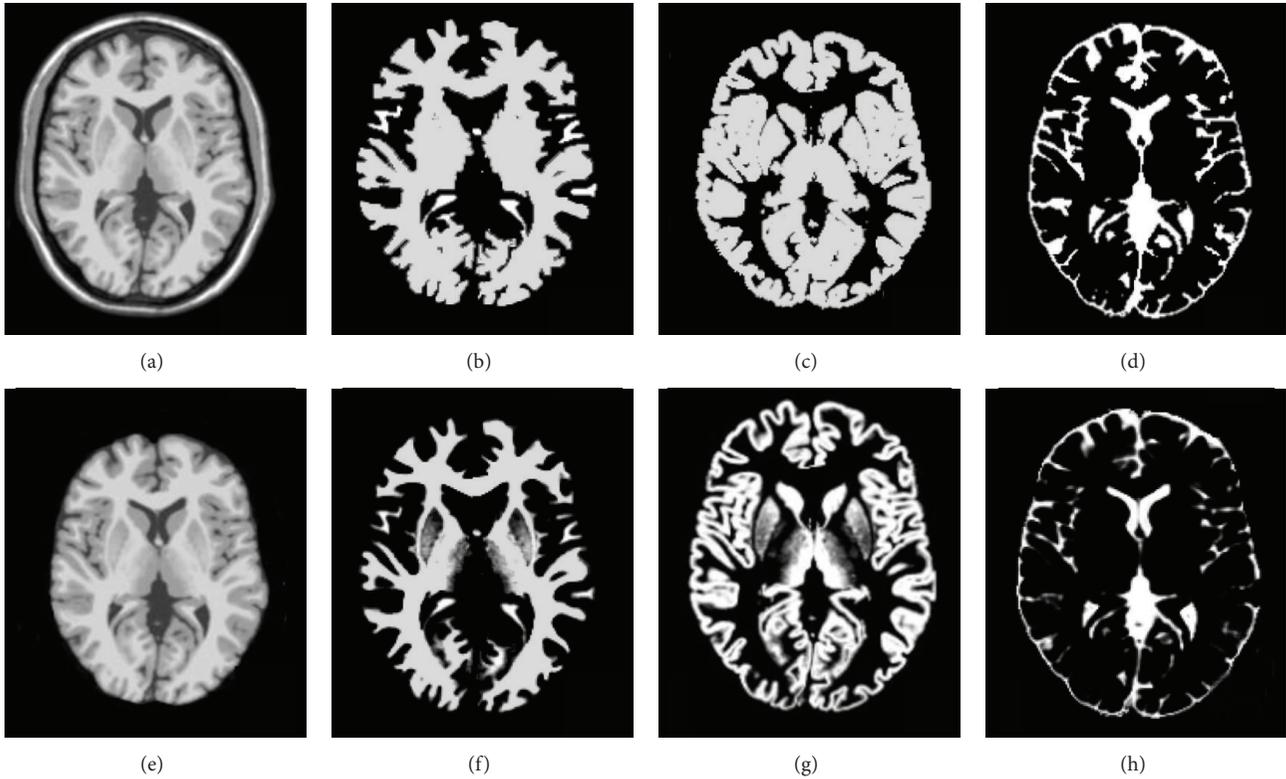


FIGURE 8: Results of segmentations on the BrainWeb images (a, e), synthetic image, and the extracted brain image, respectively (b, c, and d). Estimated WM, GM, and CSF, respectively (f, g, and h). The ground truth images of WM, GM, and CSF, respectively.

TABLE 1: The Kappa index for the 18 volumes on the BrainWeb database with different Rician noise levels and 20% inhomogeneity.

Methods	White matter (%)					Grey matter (%)						
	Noise level	0	1	5	7	9	0	1	3	5	7	9
EM		86.1	91.5	92.2	90.1	86.4	83.1	90.8	92.5	92	89.1	84.2
SPM 5		91.05	94.2	93.6	90.2	86.3	91.2	93.4	93.3	92.1	90	86.6
HMC		97.8	97.7	93.9	92.3	91.7	97	96.5	95.1	93.7	91.6	90.3
Fast		97.4	96.8	94.8	94.3	91.9	96	95.8	95.3	93.8	91.5	91.1
FCM		97	96	92	88	84	97	96	94	91	87	83
NL-FCM		95.6	94.2	91.5	89.8	83.2	95.4	94.1	93.8	92.9	89.9	79.3
Proposed method		97	95	94.9	94.4	92.2	95.9	95.7	95.3	93.8	92.1	91.2

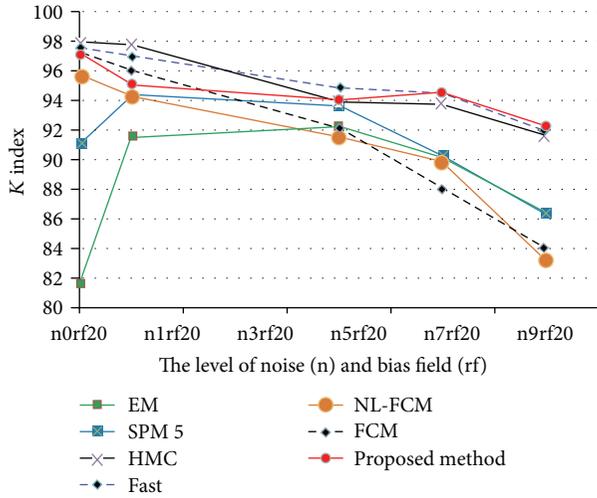
Fast = 94.72, and the proposed method = 94.46. The average Kappa indexes of GM segmentation are EM = 88.61, SPM 5 = 91.1, FCM = 91.3, HMC = 94.03, NL-FCM = 90.9, NL-Fast = 93.9, and the proposed method = 94.

**3.2. Real Brain MRI.** The proposed algorithm is also applied to real MR images, which are obtained from the Internet Brain Segmentation Repository (IBSR), which are available at <http://www.nitrc.org/projects/ibsr>. These brain image datasets and their manual segmentations were provided by the Center for Morphometric Analysis at Massachusetts General Hospital. This dataset of images is a set of 18 3D brain images with expert segmented volumes. The resolution of these images varies from  $0.8 \times 0.8 \times 1.5$  to  $1.0 \times 1.0 \times 1.5$  mm and they have a size of  $256 \times 256 \times 128$  voxels. The computation time of IBSR dataset is then globally similar to that one of a BrainWeb image.

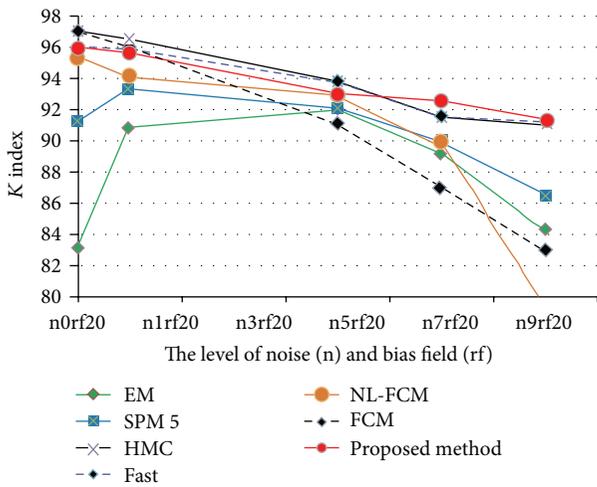
Dealing with real data, we are faced with problems using Brain Surface Extractor (BSE) in separating the brain from nonbrain tissues. Indeed, some nonbrain regions still appear in the images, degrading the segmentation results. In order to handle this issue we used the atlas to separate nonbrain tissues.

For visual evaluation, two slices are selected from IBSR dataset and our method was applied on them. Figures 10(a) and 10(b) display two original selected slices. The manual segmented results are shown in Figures 10(c) and 10(d), respectively. The corresponding segmentation results achieved by our method are shown in Figures 10(e) and 10(f), respectively. Figure 11 demonstrates the other slice of the original IBSR volume after brain segmentation.

Comparison with manual segmented images demonstrates that the proposed method outputs provide satisfactory results because the similarity between manual segmented and



(a)



(b)

FIGURE 9: Average Kappa indexes of the simulated images. ((a) to (b)) The average Kappa indexes for WM segmentation. The average Kappa indexes for GM segmentation.

automatic segmented images could be observed (see Figures 10 and 11). For more analysis, our method is applied on all 18 real images of IBSR and to measure the similarities between the two sets we calculated  $K$  indexes of images and compared them with the other state-of-the-art methods. The results of our comparison are satisfactory based on the  $K$  index values (see Tables 1 and 2).

Since IBSR is generally used brain MR images for the validation of tissue segmentation, the results of the mentioned algorithm can be compared to those obtained by the other state-of-the-art techniques, particularly the following ones: hidden Markov chains (HMC) [43, 46], expectation-maximization (EM), statistical parametric mapping (SPM 5) [4, 45], Fuzzy  $C$ -Means (FCM), Nonlocal Fuzzy  $C$ -Means (NL-FCM) [45].

Based on these considerations, the overlap measures are computed for WM and GM and the average results obtained

TABLE 2: Mean and standard deviation of the Kappa index for different segmentation methods.

Methods	White matter (%)		Grey matter (%)	
	Mean	Standard deviation	Mean	Standard deviation
HMC	86.53	1.73	79.94	5.57
EM	85.87	2.27	78.94	5.68
SPM 5	85.27	5.52	78.7	13.98
FCM	85.6	3.81	83.21	4.03
NL-FCM	84.68	3.38	78.84	4.07
Proposed method	85.90	2.89	82.21	3.95

in the 18 cases are compared to the ones of these other techniques.

Since the brain images in the IBSR dataset are segmented only into pure tissue classes, our segmentation results are converted into three classes (WM, GM, and CSF). As the IBSR ground truth includes only internal CSF while our technique also defines sulcal CSF, we do not report results for CSF. The quantitative mean results are also presented in Table 2 while the results for each method are depicted.

From the measures of Table 2, it appears that all the segmentation algorithms considered in these experiments provide approximately similar results for the WM. When considering the results, the use of the proposed method globally leads to better results than the other state-of-the-art techniques (in terms of both the mean value and standard deviation).

Figure 12 presents the Kappa index for the 18 volumes from the IBSR database. The values of overlap measures in Figure 12 are based on published results and free available reference software for brain MRI segmentation.

The average Kappa index of WM segmentations is slightly improved in IBSR dataset. On the other hand the average Kappa indexes of GM segmentations suggest that the improvement is substantial. This variation may be attributed to the various spatial complexities of WM and GM. However, for both WM and GM, voxels of the same tissue class are connected to one another; GM tissue is inherently much more tortuous than WM tissue. Together with the fact that the standard Gaussian model does not precisely demonstrate the intensity distributions of real images, the proposed technique tends to improve the segmentation. In other words, the better performance of the proposed method over standard EM can be attributed to its accurate demonstration of intensity distribution. On average, however, the proposed technique still outperforms other competing techniques in classifying GM voxels. The average Kappa indexes of IBSR images for WM are EM = 86, SPM 5 = 85.3, FCM = 85.4, HMC = 86.91, NL-FCM = 84.83, and the proposed method = 87.20. The average Kappa indexes of GM segmentation are EM = 78.72, SPM 5 = 78.6, FCM = 83.16, HMC = 80.36, NL-FCM = 78.9, and the proposed method = 84.40. Furthermore, brain extraction step can cause differences in the final results of segmentation in terms of the Kappa index, as the number of voxels in the segmentation references can differ depending on the brain extraction algorithm.

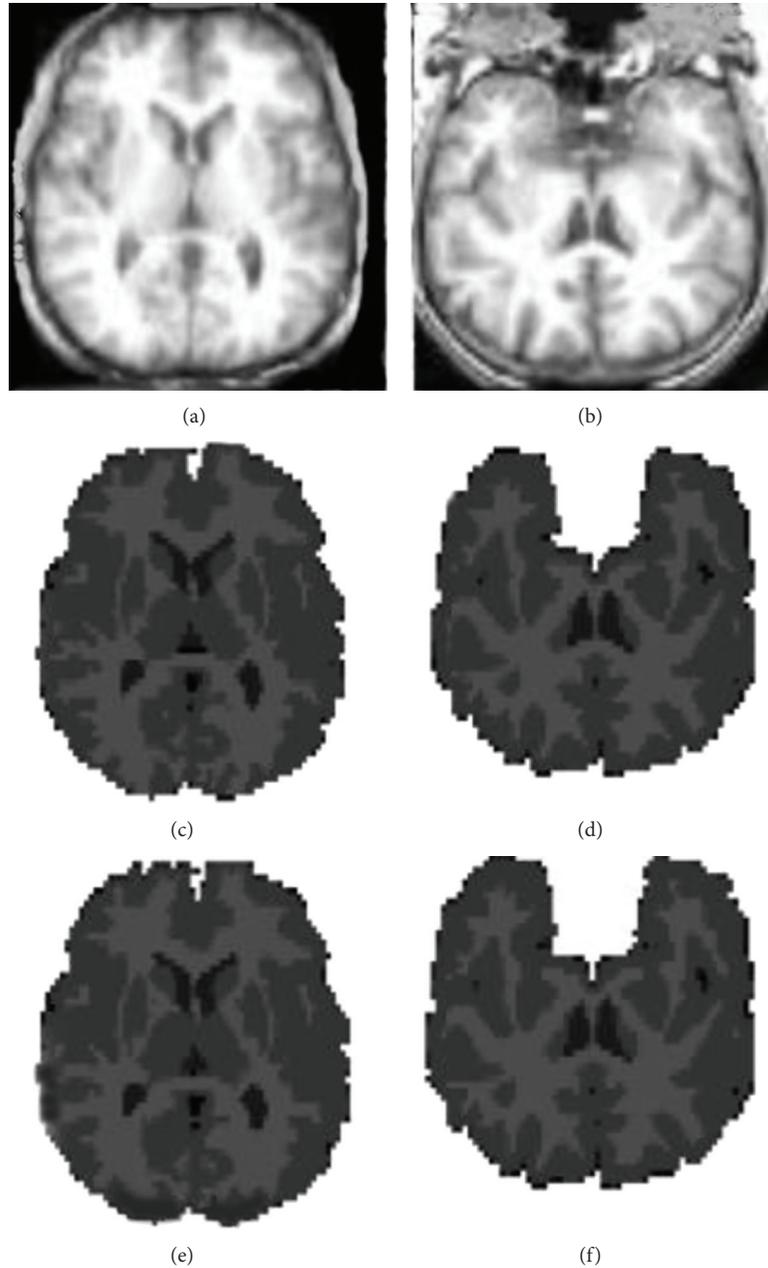


FIGURE 10: Proposed algorithm applied to IBSR database. Brain MRI slices of IBSR database. Expert segmentation. (e, f) Results of the segmentation of proposed method.

As statistical analysis we considered standard deviation and mean of  $k$  index for 18 real images and different methods. The robustness and accuracy of techniques could be evaluated by the amount of mean and standard deviation. In other words the larger mean leads to more accurate result and the smaller standard deviation leads to robustness (see Figure 13). In Figure 13 the vertical lines demonstrate the standard deviation and the blue graph indicates the mean overlap rate of different techniques for 18 real MRI. Figure 13 and Table 2 display that the proposed algorithm outperforms other competing methods.

In WM segmentation the mean overlap measure of our method is 88.35, which is 2% to 5% higher than other methods. In addition the standard deviation of  $k$  index of proposed method is 1.70, which is 1% to 4% less than other techniques. In addition, in terms of GM classification, the results of our algorithm are significantly better than WM segmentation.

Moreover, brain extraction step may cause differences in the results of brain classification in terms of the  $K$  index, as the number of voxels in the segmentation references may vary depending on the brain extraction method.

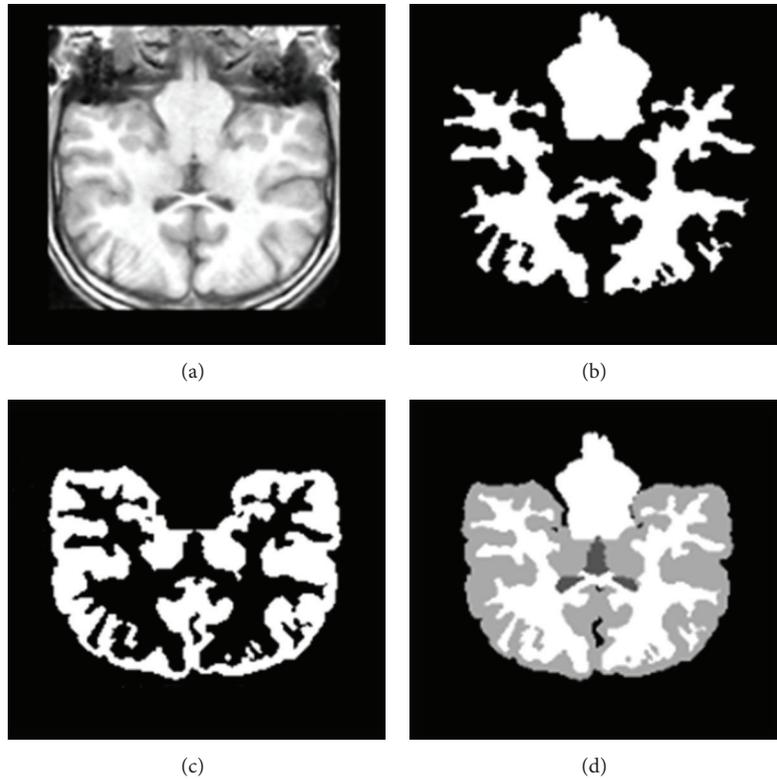


FIGURE 11: Results of segmentations on the T1-weighted IBSR image; (a) original image; (b) estimated WM image; (c) estimated GM image; (d) estimated CSF image.

#### 4. Conclusion

In this paper, we proposed a new automatic algorithmic framework for brain tissue segmentation using a novel combination of modified EM, 3D GLCM, and SVM.

Since using intensity information as the only feature in MR image is not sufficient to have a robust segmentation, in the proposed technique the spatial information and intensity information are used in different stages. The intensity information is used for the initial segmentation by EM based method. After feature extraction we used the target area from the first segmentation step to train SVM. Consequently to decrease the training and testing time of SVM and to have an accurate segmentation, we used location information as well as MRI intensity information as input features. The support vector machine classifier is employed to assign a label to each overlapped VOI sample of each class.

On the other hand, most statistical segmentation techniques in the literature have assumed that the intensity distribution of each tissue type is Gaussian distribution. However, the manual segmentation results provided by the IBSR dataset suggested that intensity distributions of brain tissues could vary asymmetrically. Thus an initial segmentation of the brain image into primitive regions is set by applying a modified EM method. In this step we assumed that the real brain MR images have asymmetric Gaussian distribution. The proposed EM algorithm overcomes the shortcoming of the standard EM technique using asymmetric Gaussians.

This step demonstrates a new method to overcome the problems with estimating the symmetric standard deviation of each Gaussian in the histogram. The experimental results indicate that the combination of the statistical and the machine learning based segmentation methods can enhance the overall segmentation performance, compared with each component individually. This is because the proposed method takes advantages of the classification ability of machine learning method in addition to the MR intensity and location information, which are consequential information to classify the brain in a 3D MRI into the multiple classes. In this paper to improve the overall segmentation performance different types of features are integrated, which are textural features, MR intensity, relationship with neighboring voxels, and voxel location. Robustness to noise and simplicity are two advantages of proposed framework. The results are independent of registration step and it makes our algorithm faster than other registration-based methods. In addition because our method is designed to run in Matlab, it is not platform-dependent and it can be run in both Linux and Windows operating systems.

In order to assess the proposed approach, it has been applied to brain tissue MR segmentation using real and simulated data, producing satisfactory results with respect to segmentation performance. The experimental results demonstrate that the integration of machine learning and statistical based segmentation techniques can improve the overall segmentation performance, in comparison with its individual

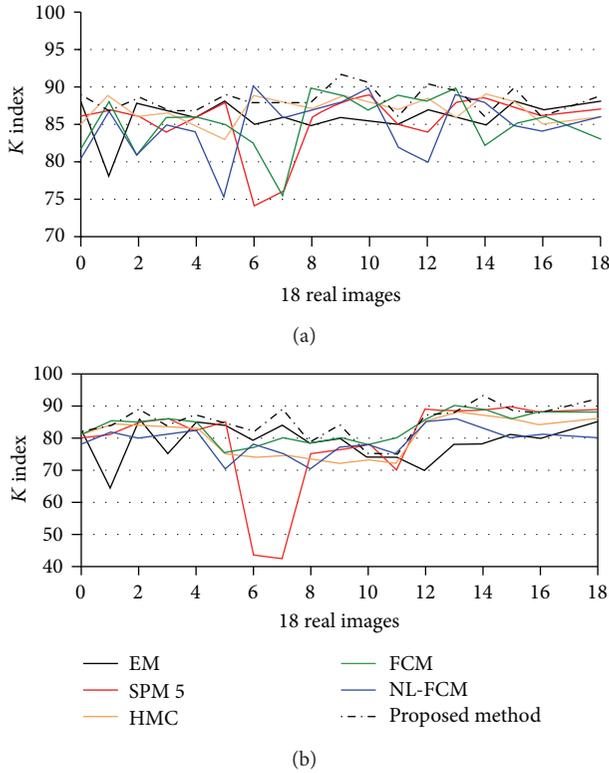


FIGURE 12: Kappa index calculated throughout the images in the IBSR database. ((a) to (b)) The Kappa indexes for WM segmentation. The Kappa indexes for GM segmentation.

components. This improvement is because the proposed technique takes advantage of the classification ability of machine learning algorithm in addition to the location and voxel intensity information, which are consequential information for brain MRI segmentation into the different classes.

Experiments on real data from the IBSR and synthetic images from BrainWeb have indicated that the proposed method achieves higher Kappa indexes compared with other methods currently in use. Incorporating spatial techniques such as 3D GLCM into the proposed approach could lead to interesting alternatives.

The proposed method not only preserves simplicity, but also has the potential to be generalized to multivariate versions adapted for segmentation applying multimodality images (e.g., T1, T2, and PD images). Experiments were first performed on different noise (up to 9% Rician noise) and 20% inhomogeneity BrainWeb MR images. These experiments show the precision and robustness of our method in the presence of different levels of noise and bias field. Additional experiments run on real MR images from the IBSR database have demonstrated that this method reliably extracts brain tissues with accuracy comparable to state-of-the-art techniques.

In terms of application, the proposed technique can be useful in the case of low contrast images (challenged by inherently low contrast tissue boundaries), for example, in the study of the developing human fetus. Extension of

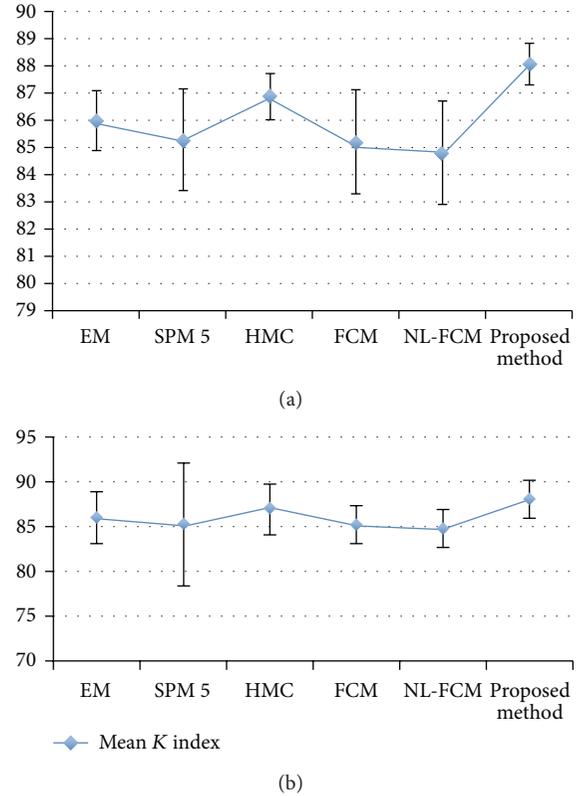


FIGURE 13: Mean and standard deviation of K index for segmentation methods in Figure 12. ((a) to (b)) WM graph, GM graph.

the proposed method for tumor and disease detection is the next challenging task for the future.

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

**Acknowledgments**

The authors would like to acknowledge the financial support from Research University grant of the Ministry of Higher Education of Malaysia (MOHE) under Project Grant GUP-04H40. Also, thanks are due to the Research Management Center (RMC) of Universiti Teknologi Malaysia (UTM) for research environment to complete this work.

**References**

- [1] K. Shmueli, J. A. de Zwart, P. van Gelderen, T.-Q. Li, S. J. Dodd, and J. H. Duyn, "Magnetic susceptibility mapping of brain tissue in vivo using MRI phase data," *Magnetic Resonance in Medicine*, vol. 62, no. 6, pp. 1510–1522, 2009.
- [2] A. Karimian, S. Yazdani, and M. A. Askari, "Reducing the absorbed dose in analogue radiography of infant chest images by improving the image quality, using image processing techniques," *Radiation Protection Dosimetry*, vol. 147, no. 1-2, pp. 176–179, 2011.

- [3] M. Mansourvar, M. A. Ismail, T. Herawan, R. G. Raj, S. Abdul Kareem, and F. H. Nasaruddin, "Automated bone age assessment: motivation, taxonomies, and challenges," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 391626, 11 pages, 2013.
- [4] L. P. Clarke, R. P. Velthuizen, M. A. Camacho et al., "MRI segmentation: methods and applications," *Magnetic Resonance Imaging*, vol. 13, no. 3, pp. 343–368, 1995.
- [5] J.-D. Lee, H.-R. Su, P. E. Cheng et al., "MR image segmentation using a power transformation approach," *IEEE Transactions on Medical Imaging*, vol. 28, no. 6, pp. 894–905, 2009.
- [6] A. M. MacEira, J. Cosín-Sales, M. Roughton, S. K. Prasad, and D. J. Pennell, "Reference right atrial dimensions and volume estimation by steady state free precession cardiovascular magnetic resonance," *Journal of Cardiovascular Magnetic Resonance*, vol. 15, article 29, 2013.
- [7] T. P. Hellebust, K. Tanderup, C. Lervåg et al., "Dosimetric impact of interobserver variability in MRI-based delineation for cervical cancer brachytherapy," *Radiotherapy and Oncology*, vol. 107, no. 1, pp. 13–19, 2013.
- [8] R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, "Automatic anatomical brain MRI segmentation combining label propagation and decision fusion," *NeuroImage*, vol. 33, no. 1, pp. 115–126, 2006.
- [9] Z.-X. Ji, Q.-S. Sun, and D.-S. Xia, "A framework with modified fast FCM for brain MR images segmentation," *Pattern Recognition*, vol. 44, no. 5, pp. 999–1013, 2011.
- [10] C. W. Kanaly, D. Ding, A. I. Mehta et al., "A novel method for volumetric MRI response assessment of enhancing brain tumors," *PLoS ONE*, vol. 6, no. 1, Article ID e16031, 2011.
- [11] M. Weygandt, K. Hackmack, C. Pfüller et al., "MRI pattern recognition in multiple sclerosis normal-appearing brain areas," *PLoS ONE*, vol. 6, no. 6, Article ID e21138, 2011.
- [12] T. Jubault, S. M. Brambati, C. Degroot et al., "Regional brain stem atrophy in idiopathic Parkinson's disease detected by anatomical MRI," *PLoS ONE*, vol. 4, no. 12, Article ID e8247, 2009.
- [13] N. Zhang, S. Ruan, S. Lebonvallet, Q. Liao, and Y. Zhu, "Kernel feature selection to fuse multi-spectral MRI images for brain tumor segmentation," *Computer Vision and Image Understanding*, vol. 115, no. 2, pp. 256–269, 2011.
- [14] A. Vita, L. de Peri, G. Deste, and E. Sacchetti, "Progressive loss of cortical gray matter in schizophrenia: a meta-analysis and meta-regression of longitudinal MRI studies," *Translational Psychiatry*, vol. 2, article e190, 2012.
- [15] J. Tang, Y. Liao, B. Zhou et al., "Decrease in temporal gyrus gray matter volume in first-episode, early onset schizophrenia: an MRI study," *PLoS ONE*, vol. 7, no. 7, Article ID e40247, 2012.
- [16] J. L. Marroquin, B. C. Vemuri, S. Botello, F. Calderon, and A. Fernandez-Bouzas, "An accurate and efficient Bayesian method for automatic segmentation of brain MRI," *IEEE Transactions on Medical Imaging*, vol. 21, no. 8, pp. 934–945, 2002.
- [17] M. S. M. Rahim, A. Norouzi, I. B. M. Amin, and A. E. Rad, "Current methods in medical image segmentation and its application on knee bone," *Archives Des Sciences Journal*, vol. 65, no. 9, 2012.
- [18] X. Yang and B. Fei, "A multiscale and multiblock fuzzy C-means classification method for brain MR images," *Medical Physics*, vol. 38, no. 6, pp. 2879–2891, 2011.
- [19] M. Bhattacharya, A. Das, and M. Chandana, "GA-based multiresolution fusion of segmented brain images using PD-, T1- and T2-weighted MR modalities," *Neural Computing and Applications*, vol. 21, no. 6, pp. 1433–1447, 2012.
- [20] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *Journal of Cybernetics. Transactions of the American Society for Cybernetics*, vol. 3, no. 3, pp. 32–57, 1973.
- [21] H. M. Mofteh, A. T. Azar, E. T. Al-Shammari, N. I. Ghali, A. E. Hassanien, and M. Shoman, "Adaptive k-means clustering algorithm for MR breast image segmentation," *Neural Computing and Applications*, vol. 24, no. 7-8, pp. 1917–1928, 2014.
- [22] S. Shen, W. Sandham, M. Granat, and A. Sterr, "MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization," *IEEE Transactions on Information Technology in Biomedicine*, vol. 9, no. 3, pp. 459–467, 2005.
- [23] L. Lemieux, A. Hammers, T. Mackinnon, and R. S. N. Liu, "Automatic segmentation of the brain and intracranial cerebrospinal fluid in T1-weighted volume MRI scans of the head, and its application to serial cerebral and intracranial volumetry," *Magnetic Resonance in Medicine*, vol. 49, no. 5, pp. 872–884, 2003.
- [24] K. Held, E. R. Kops, B. J. Krause, W. M. Wells III, R. Kikinis, and H.-W. Müller-Gärtner, "Markov random field segmentation of brain MR images," *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, pp. 878–886, 1997.
- [25] B. Caldairou, N. Passat, P. A. Habas, C. Studholme, and F. Rousseau, "A non-local fuzzy segmentation method: application to brain MRI," *Pattern Recognition*, vol. 44, no. 9, pp. 1916–1927, 2011.
- [26] A. Abdullah, A. Hirayama, S. Yatsushiro, M. Matsumae, and K. Kuroda, "Cerebrospinal fluid image segmentation using spatial fuzzy clustering method with improved evolutionary expectation maximization," in *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '13)*, pp. 3359–3362, IEEE, Osaka, Japan, July 2013.
- [27] W. M. Wells III, W. E. L. Crimson, R. Kikinis, and F. A. Jolesz, "Adaptive segmentation of MRI data," *IEEE Transactions on Medical Imaging*, vol. 15, no. 4, pp. 429–442, 1996.
- [28] O. Chapelle, P. Haffner, and V. N. Vapnik, "Support vector machines for histogram-based image classification," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 1055–1064, 1999.
- [29] D. K. Srivastava and L. Bhambhu, "Data classification using support vector machine," *Journal of Theoretical and Applied Information Technology*, vol. 12, no. 1, pp. 1–7, 2010.
- [30] F. Akram, J. H. Kim, H. U. Lim, and K. N. Choi, "Segmentation of intensity inhomogeneous brain MR images using active contours," *Computational and Mathematical Methods in Medicine*, vol. 2014, Article ID 194614, 14 pages, 2014.
- [31] J. Tohka, A. Zijdenbos, and A. Evans, "Fast and robust parameter estimation for statistical partial volume models in brain MRI," *NeuroImage*, vol. 23, no. 1, pp. 84–97, 2004.
- [32] T. Kapur, W. Eric, L. Grimson, W. M. Wells III, and R. Kikinis, "Segmentation of brain tissue from magnetic resonance images," *Medical Image Analysis*, vol. 1, no. 2, pp. 109–127, 1996.
- [33] A. A. Willette, V. D. Calhoun, J. M. Egan, and D. Kapogiannis, "Prognostic classification of mild cognitive impairment and Alzheimer's disease: MRI independent component analysis," *Psychiatry Research: Neuroimaging*, vol. 224, no. 2, pp. 81–88, 2014.

- [34] T. Kato, S. Omachi, and H. Aso, "Asymmetric gaussian and its application to pattern recognition," in *Structural, Syntactic, and Statistical Pattern Recognition*, pp. 405–413, Springer, Berlin, Germany, 2002.
- [35] A. Ryan, B. Mora, and M. Chen, "On the implementation and analysis of expectation maximization algorithms with stopping criterion," in *Proceedings of the 19th IEEE International Conference on Image Processing (ICIP '12)*, pp. 2393–2396, IEEE, Orlando, Fla, USA, October 2012.
- [36] S. B. Choi, W. J. Kim, T. K. Yoo et al., "Screening for prediabetes using machine learning models," *Computational and Mathematical Methods in Medicine*, vol. 2014, Article ID 618976, 8 pages, 2014.
- [37] A. Van Opbroek, F. Van der Lijn, and M. De Bruijne, "Automated brain-tissue segmentation by multi-feature SVM classification," *The MIDAS Journal*. In press.
- [38] M. L. McHugh, "Interrater reliability: the kappa statistic," *Biochemia Medica*, vol. 22, no. 3, pp. 276–282, 2012.
- [39] M. S. Atkins and B. T. Mackiewicz, "Fully automatic segmentation of the brain in MRI," *IEEE Transactions on Medical Imaging*, vol. 17, no. 1, pp. 98–107, 1998.
- [40] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy, "Magnetic resonance image tissue classification using a partial volume model," *NeuroImage*, vol. 13, no. 5, pp. 856–876, 2001.
- [41] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, and A. C. Evans, "Brainweb: online interface to a 3D MRI simulated brain database," *NeuroImage*, vol. 5, no. 4, p. S425, 1997.
- [42] D. L. Collins, A. P. Zijdenbos, V. Kollokian et al., "Design and construction of a realistic digital brain phantom," *IEEE Transactions on Medical Imaging*, vol. 17, no. 3, pp. 463–468, 1998.
- [43] J. Ghasemi, R. Ghaderi, M. R. Karami Mollaei, and S. A. Hojatoleslami, "A novel fuzzy Dempster-Shafer inference system for brain MRI segmentation," *Information Sciences*, vol. 223, pp. 205–220, 2013.
- [44] A. Ortiz, A. A. Palacio, J. M. Górriz, J. Ramírez, and D. Salas-González, "Segmentation of brain MRI using SOM-FCM-based method and 3D statistical descriptors," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 638563, 12 pages, 2013.
- [45] S. Bricq, C. Collet, and J. P. Armspach, "Unifying framework for multimodal brain MRI segmentation based on Hidden Markov Chains," *Medical Image Analysis*, vol. 12, no. 6, pp. 639–652, 2008.
- [46] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 45–57, 2001.

## Research Article

# Developing a New Computer-Aided Clinical Decision Support System for Prediction of Successful Postcardioversion Patients with Persistent Atrial Fibrillation

Mark Sterling,<sup>1</sup> David T. Huang,<sup>2</sup> and Behnaz Ghoraani<sup>1</sup>

<sup>1</sup>Biomedical Engineering Department, Rochester Institute of Technology, Rochester, NY 14623, USA

<sup>2</sup>University of Rochester Medical Center, School of Medicine and Dentistry, Rochester, NY 14642, USA

Correspondence should be addressed to Behnaz Ghoraani; [bxgbme@rit.edu](mailto:bxgbme@rit.edu)

Received 26 September 2014; Revised 12 February 2015; Accepted 17 February 2015

Academic Editor: Shahram Shirani

Copyright © 2015 Mark Sterling et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a new algorithm to predict the outcome of direct-current electric (DCE) cardioversion for atrial fibrillation (AF) patients. AF is the most common cardiac arrhythmia and DCE cardioversion is a noninvasive treatment to end AF and return the patient to sinus rhythm (SR). Unfortunately, there is a high risk of AF recurrence in persistent AF patients; hence clinically it is important to predict the DCE outcome in order to avoid the procedure's side effects. This study develops a feature extraction and classification framework to predict AF recurrence patients from the underlying structure of atrial activity (AA). A multiresolution signal decomposition technique, based on matching pursuit (MP), was used to project the AA over a dictionary of wavelets. Seven novel features were derived from the decompositions and were employed in a quadratic discrimination analysis classification to predict the success of post-DCE cardioversion in 40 patients with persistent AF. The proposed algorithm achieved 100% sensitivity and 95% specificity, indicating that the proposed computational approach captures detailed structural information about the underlying AA and could provide reliable information for effective management of AF.

## 1. Introduction

Atrial fibrillation (AF), the most common abnormal rhythm of the heart, is associated with significant morbidity and mortality and increases the risk of heart failure and stroke [1]. AF is the disorganized propagation of electrical activity in the atrium that prevents organized contractions. As a result, the atrial depolarization wavefront, the P-wave, measured during sinus rhythm (SR) devolves into a series of fibrillatory waves in the surface electrocardiogram (ECG). AF is known to be progressive in nature [2, 3]. The disease tends to worsen over time and the resistance to therapy increases. Paroxysmal AF is defined by self-terminating AF episodes that last no longer than seven days. Persistent AF is defined by AF episodes which lasts longer than seven days and typically requires medical intervention to be terminated. Lastly, if AF is sustained for over a year and all attempts to eliminate AF fail, the AF is defined as Permanent AF. Given the progressive nature of AF and potential risks of different AF therapies, it

is critical to identify if a given therapy is effective. This could provide invaluable information for effective management of AF.

There are a variety of treatment options for AF, including both pharmacological and electrical cardioversion and also surgical methods. Direct-current electric (DCE) cardioversion is one noninvasive treatment for AF that applies controlled transthoracic electrical shocks synchronized to the R-wave of the patient [1] in order to end AF and return the patient to SR. The DCE cardioversion treatment may be either immediately unsuccessful or there may be a recurrence of AF in the following months, which means that AF cannot be terminated using the DCE cardioversion therapy. It was reported that this procedure is successful in around 80–100% of the patients; however, only 20–40% maintain SR within one year after the therapy [4]. Hence, a reliable test that could accurately predict the likelihood of SR maintenance after DCE cardioversion is important in order to weigh the benefits versus potential risks such as postshock

TABLE 1: Clinical characteristics of the study population with respect to rhythm at follow-up.

Variable	AF-Free	AF-Relapse	P value
Age	68 ± 7	69 ± 7	N.S.
Male	16	17	N.S.
AF duration	128 ± 94	210 ± 94	N.S.
Other heart diseases			
Hypertension	6	7	N.S.
Ischemic heart disease	5	2	N.S.
Congestive heart failure	1	6	0.002
Valvular disease	1	3	0.04
Left atrial diameter	49 ± 6	51 ± 6	N.S.
	Cardioactive drugs		
β-blocker	12	11	N.S.
Sotalol	3	3	N.S.
Class III antiarrhythmic agent	1	1	N.S.
Digitalis	2	7	0.05
Calcium channel blocker	3	5	N.S.

bradycardia, malignant ventricular arrhythmias, and atrial thromboembolism [5]. Therefore, the objective of the present work is to develop a novel computational approach to analyze the electrocardiogram of AF patients before application of DCE cardioversion and predict the success of the therapy. Such a predictor could provide an important computer-aided clinical decision support system for therapy management of AF patients.

Over the past decade, several studies have attempted clinical and electrophysiological parameters to predict SR maintenance after DCE cardioversion of AF [6–11]. A central notion in AF therapy management is that irregularity of fibrillatory wave signals reflects the severity of the disease in an individual. Thus, several studies measured organization of atrial activity (AA) from the surface ECG as a measure of SR maintenance. Some of these algorithms include fibrillatory rate [6, 12, 13], harmonic decay [6], and entropy [14, 15]. However, none of the existing methods has been used in the routine clinical AF therapy management [16]. In the present study, we investigate AF organization beyond what has been performed in literature so far. Our method studies both the morphology and frequency of the fibrillatory waves during AF in an attempt to provide a strong and yet meaningful predictor for sinus rhythm maintenance after electric cardioversion. We apply a signal decomposition technique to examine the structure of AA at different decomposition levels for the purpose of prediction of the outcome of the DCE cardioversion in persistent AF.

A preprocessing technique is applied to extract the AA from the ECG. The matching pursuit (MP) technique [17] is used to decompose the AA signal into multiresolution time-frequency (TF) decompositions. The MP decomposition consists of a combination of wavelet atoms with two wavelet types (i.e., Coiflet1 and Symlet2) and 6 scales ( $S_0$  to  $S_5$ ). We investigate the type and scale of the wavelet types and scales that most accurately capture the structural changes relevant to SR maintenance and propose seven new MP features. Using a quadratic discriminant analysis (QDA) classification

technique and leave-one-out cross validation, we evaluate the developed MP features on a database containing ECG from persistent patients who underwent DCE cardioversion. Details of the algorithm are outlined in Section 2; a validation of the features against clinical data and discussion is provided in Section 3. The paper is concluded in Section 4.

## 2. Methods

The proposed method consists of three stages: (i) preprocessing of the ECG signal, (ii) feature extraction, and (iii) classification and validation of the extracted features against clinical outcome. Figure 1 depicts the overall outline of the proposed method.

*2.1. ECG Database.* The ECG data [18] was obtained from 40 persistent AF patients who had a successful external DCE cardioversion therapy. The study was approved by the local ethics committee of the enrolling organization and complied with the Declaration of Helsinki. Prior to cardioversion, a 10-minute 12-lead ECG ( $f_s = 1$  kHz) was recorded for each patient. Twenty patients had maintained SR (AF-Free) after 2-week follow-up and 20 had a relapse of AF (AF-Relapse). The clinical characteristics of the patients, including medication and AF history, are given in Table 1. The proposed analysis was based on a single lead. Hence, we selected Lead VI, which has shown to provide the best atrial signal [19]. The proposed algorithm was applied on a 10-second long ECG signal, which was selected from the 10-minute ECG as explained in Section 2.2.

*2.2. Preprocessing.* The preprocessing stage is performed in three steps as follows.

*Noise and Baseline Wander Removal.* A bandpass filter with cutoff frequencies of 0.01 Hz and 50 Hz was used to remove the noise and baseline fluctuations in the ECG [20].

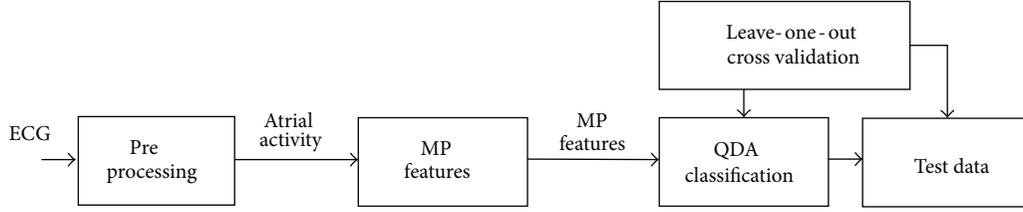


FIGURE 1: Overall outline of the study. Standard supervised learning approach is applied consisting of a feature extraction step followed by a classification step. Leave-one-out cross validation is used to evaluate the predictive power of our technique.

*Segment Selection.* Following the baseline removal step, the segments with a consistent QRST morphology are identified. In case of AF, it is common that the ECG contains more than one QRST morphology, which can increase the QRST residuals in the atrial activity extraction step and dilute the quality of AA for the further analysis. Hence in this step we identify the ECG segments that include steady QRST complexes. First, R-wave fiducial markers are placed at points of maximum absolute derivative on the QRST complexes. We construct a QRST template by averaging all of the QRST complexes in the ECG. We then compute the correlation between the QRST template and each beat and identify the segments with more than 90% correlation coefficient.

*Atrial Activity Extraction.* Several techniques have been used to cancel the QRST complexes and obtain the AA from the ECG [21, 22]. In this study, we employ the average beat subtraction method [23] which has been widely used in the literature. Using the QRST template that was computed in the previous step, at each fiducial marker, we fit the QRST template to the ECG and obtain the estimated QRST template from the ECG. Then, we subtract the estimated QRST template from the ECG to obtain the estimated AA signal. We evaluate the QRST removal by computing kurtosis as a measure of the AA estimation quality [24]. Finally, for each record we select a 10-second long excerpt with the lowest kurtosis. Figure 2 illustrates the selected segment for one of the records.

**2.3. Feature Extraction.** MP decomposition is applied to the extracted AA signal and the MP features are extracted from the MP expansion coefficients.

*Matching Pursuit Decomposition.* MP is an iterative signal decomposition technique that expresses a signal  $x(t)$  as a linear combination of functions selected from an overcomplete dictionary of TF basis functions [17]. The algorithm has been successful in creating high-resolution TF representations of biomedical signals [25–27]. In this study, we apply the MP algorithm to the AA signal obtained from the preprocessing step. Consider

$$x(t) = \sum_{m=1}^M b_m A_{(W_m, S_m, T_m)}(t) + R_x^M. \quad (1)$$

In (1),  $x(t)$  represents AA signal and  $A_{(W_m, S_m, T_m)}(t)$  is a wavelet with type, scale, and temporal location defined by

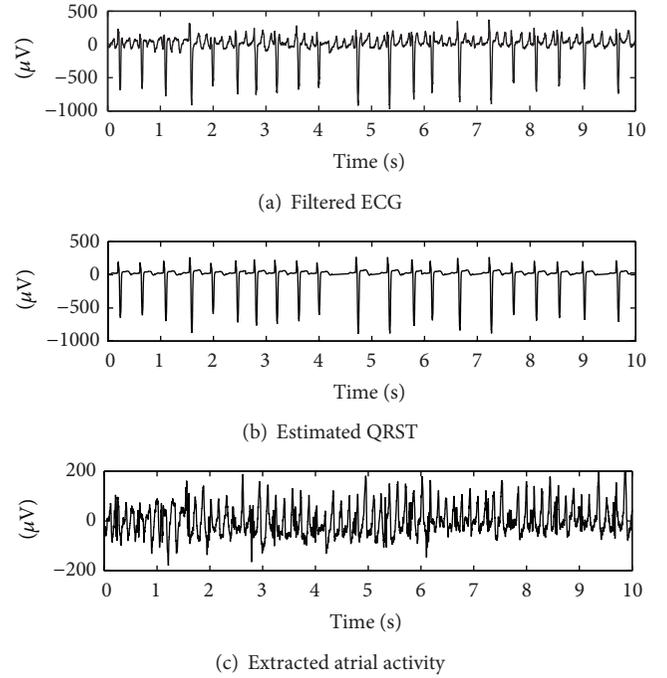


FIGURE 2: Illustrative example of an ECG waveform, the estimated QRST complex, and the extracted AA signal with kurtosis values of 10, 13, and  $-0.45$ , respectively.

$W_m$ ,  $S_m$ , and  $T_m$ , respectively.  $b_m$  is the expansion coefficient for  $A_{(W_m, S_m, T_m)}(t)$ ,  $M$  is the number of iterations that are performed, and  $R_x^M$  is the residue of  $x(t)$  after  $M$  iterations. In (1), the AA signal  $x(t)$  is projected onto an overcomplete dictionary of TF functions with a combination of different wavelet types and scales. At each iteration, the best correlated TF function is selected from the overcomplete dictionary by finding the maximum inner product of the current residue with each of the atoms in the dictionary ( $|\langle R_x^M, A_{(W_m, S_m, T_m)} \rangle|$ ). In the next iteration the residue is decomposed according to the same rules. After  $M$  iterations, the AA signal  $x(t)$  is expressed in the form of (1) where the first term on the right-hand side represents the decomposition of the original signal by the selected TF functions, and the second term is the residue at iteration  $M$ . For  $M$  large enough, it can be observed that the residue in (1) becomes negligibly small.

There are three ways of stopping the iterative process of MP. The iterations may proceed until the energy of

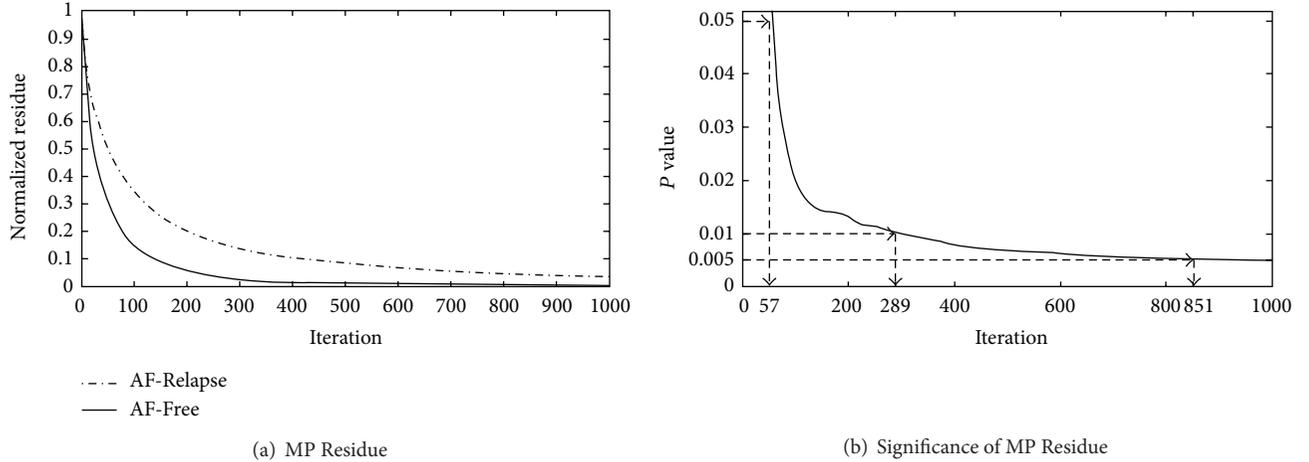


FIGURE 3: (a) Plots showing the behavior of the MP coefficients as a function of iteration number. This representation was used to find an appropriate number of iterations in our MP analysis.  $M$  was set equal to the average number of iterations required for the coefficients to reach less than 5% of their initial value. (b) The  $P$  values represent the statistical significance of the normalized expansion coefficient at each iteration. The vertical dashed lines mark the significance thresholds of 0.05, 0.01, and 0.005.

the residue is less than a threshold, the value of the most recent expansion coefficient is less than a threshold, or the number of iterations reaches a preassigned maximum. In this study, we used a combination of the last two stopping methods and determined a fixed iteration number based on the average number of iterations required for the expansion coefficients to reach less than 5% of their initial value. Based on this analysis, we found that after  $M = 1,000$  iterations, there is a negligible change in the expansion coefficients. Hence, we used  $M = 1,000$  as the fixed stopping criterion. A plot of the expansion coefficients for an AF-Free and AF-Relapse example is shown in Figure 3(a).

**MP Dictionary.** Two different wavelet types at six different scales ( $S_0$  to  $S_5$ ) are used in this study: Coiflet1 (Coif1) and Symlet2 (Sym2). We build a MP dictionary by pairing the two types of wavelets (i.e.,  $W_1$  and  $W_2$ ). Then the MP decomposition projects each AA signal over the combined MP dictionary. We depict an example of AF-Free and AF-Relapse signal along with the signal decompositions in Figures 4(a) and 4(b), respectively. The plots on the left-hand side show the reconstructed signals by combining the components corresponding to Coif1  $S_0$  and the right-hand side plots show the sum of the reconstructed signals related to  $\text{MPF}_{\text{Coif1},S_3}$ ,  $\text{MPF}_{\text{Coif1},S_4}$ ,  $\text{MPF}_{\text{Sym2},S_3}$ , and  $\text{MPF}_{\text{Sym2},S_4}$ .

**MP Features.** We performed the MP on each AA signal and obtained the decomposed wavelets and scales given by  $A_{W_m, S_m, T_m}$ ,  $m = 1, \dots, M$ . Thirteen MP features are extracted for each patient as explained in this section. However, only seven of these features contained a significant differentiation between the AF-Relapse and AF-Free data and were used in the final decision making algorithm.

The first MP feature was based on the expansion coefficient at  $M = 1,000$  iterations. As can be seen in Figure 3(a), we realized that AF-Free cases present a faster decay rate compared to AF-Relapse cases. Such a behavior was expected

as it can be hypothesized that the AF-Free cases present a more organized AA and are decomposed faster by the MP wavelets [28]. The AA signals from the AF-Relapse data contain more disorganized and incoherent structure and have a slower decay rate. Thus, we use the normalized expansion coefficient (i.e.,  $\text{MPF}_{\text{Residue}} = b_M/b_1$ ) at  $M = 1,000$  iteration as the MP feature representing the decomposition decay rate. The logarithm of the normalized coefficient expansion is taken to further spread out the data points. We performed an exploratory statistical test to investigate if the expansion coefficient at a smaller number of iterations (i.e.,  $b_m/b_1$ , where  $m < M$ ) was a more appropriate choice for the quantification of AA organization. Using the Mann-Whitney  $U$  test, we calculated the  $P$  value of the normalized coefficient expansion for  $m = 1$  to  $m = 1,000$ . As can be seen in Figure 3(b), the  $P$  value decreases as the iteration number increases. Any normalized expansion coefficient ( $b_m/b_1$ ) with  $m > 851$  can achieve a significant  $P$  value of  $< 0.005$ .

The other twelve MP features are extracted based on the decomposition results as follows. We build two matrices for each wavelet type ( $W_1$  and  $W_2$ ) in a given dictionary:  $O_{W_1}$  and  $O_{W_2}$ . These matrices which are called the *occupancy* matrices are constructed as follows:

$$O_{W_1}(i, j) = \begin{cases} 1 & \text{if } W_m = W_1 \\ 0 & \text{o.w.,} \end{cases} \quad (2)$$

$$O_{W_2}(i, j) = \begin{cases} 1 & \text{if } W_m = W_2 \\ 0 & \text{o.w.,} \end{cases}$$

for  $m = 1, \dots, M$ ,

where  $W_m$  represents the wavelet type with scale and temporal location of  $S_m$  and  $T_m$ , respectively,  $i = \{0, \dots, 5\}$  is the scale value of  $S_m$ , and  $j$  corresponds to the temporal location  $T_m$ . A graphical representation of this process is shown in

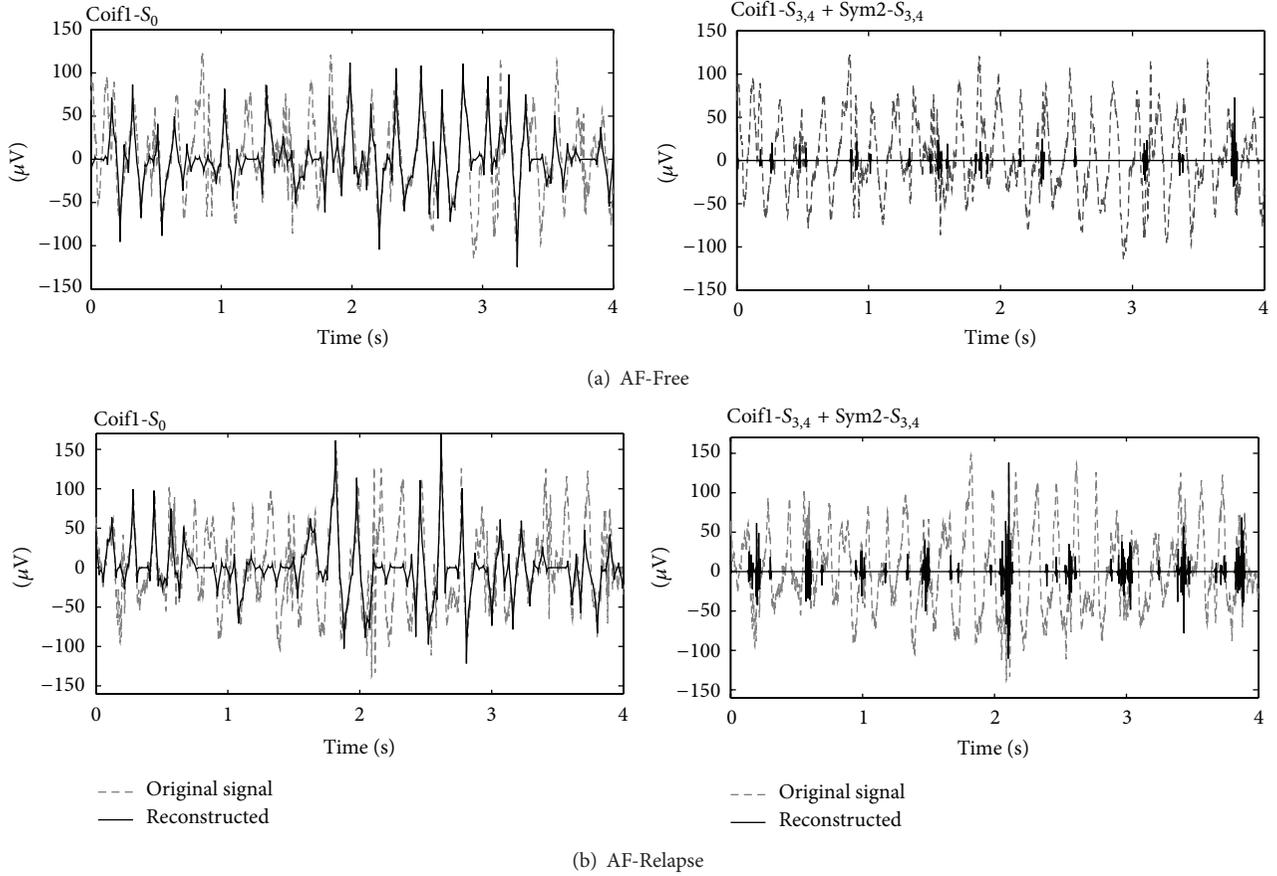


FIGURE 4: The reconstructed signals by  $\text{MPF}_{\text{Coif1},S_0}$  and the combined reconstructed signals by  $\text{MPF}_{\text{Coif1},S_3}$ ,  $\text{MPF}_{\text{Coif1},S_4}$ ,  $\text{MPF}_{\text{Sym2},S_3}$ , and  $\text{MPF}_{\text{Sym2},S_4}$  are displayed for an AF-Free (a) and AF-Relapse (b) case.

Figures 5 and 6 where two *occupancy* matrices of  $O_{W_1}$  and  $O_{W_2}$  are plotted for an example of AF-Free and an example of AF-Relapse, respectively. The plots display the analysis results of only 0.5 seconds of the AA data for visualization purposes. The first six rows show the probability of occupancy for Coif1 wavelets for scales  $S_0$  to  $S_5$ , and the next six rows show this information for the Sym2 wavelets. In this plot, each black circle implies the presence of a decomposition at the given time and scale. Twelve features are extracted from each dictionary by summing over time as follows:

$$\begin{aligned} \text{MPF}_{W_1,S_i} &= \sum_j O_{W_1}(i, j), \\ \text{MPF}_{W_2,S_i} &= \sum_j O_{W_2}(i, j), \end{aligned} \quad (3)$$

for  $i = 0, \dots, 5$ .

In (3), we obtain the features as the total presence of a given wavelet type and scale in an AA signal.

**MP Feature Selection.** The MP features proposed in this study were evaluated using an exploratory statistical analysis. The purpose was to ensure that any subsequent learning technique we applied to the data would not be burdened by

many irrelevant degrees of freedom. Thirteen MP features are extracted for each AA signal. We select the MP features that show a statistically significant correlation with the success of electric cardioversion. The statistical significance is determined for each MP feature using the Mann-Whitney  $U$  test, which is a nonparametric method for cases where the probability distribution of the data is not normal. This test is used in this study, because the MP features do not exhibit a Gaussian probability distribution. The Mann-Whitney  $U$  test results showed a statistical significance for only seven MP features. The values of these significant features are presented in Figure 7. The logarithm of the decay MP feature was taken to further spread out the data points. Because this feature had a different range than the other six features, it was shown in a separate plot. A total of seven MP features that are selected here are used in the classification stage:  $\{\text{MPF}_{\text{Coif1},S_0}, \text{MPF}_{\text{Coif1},S_3}, \text{MPF}_{\text{Coif1},S_4}, \text{MPF}_{\text{Sym2},S_2}, \text{MPF}_{\text{Sym2},S_3}, \text{MPF}_{\text{Sym2},S_4}, \text{MPF}_{\text{Residue}}\}$ .

**2.4. Classification.** A label of “0” or “1” corresponding to the AF-Free and AF-Relapse cases, respectively, was attached to each of the feature vectors derived from all of the AA signals. The learning algorithm chosen for this study uses the quadratic discriminant analysis (QDA) which separates

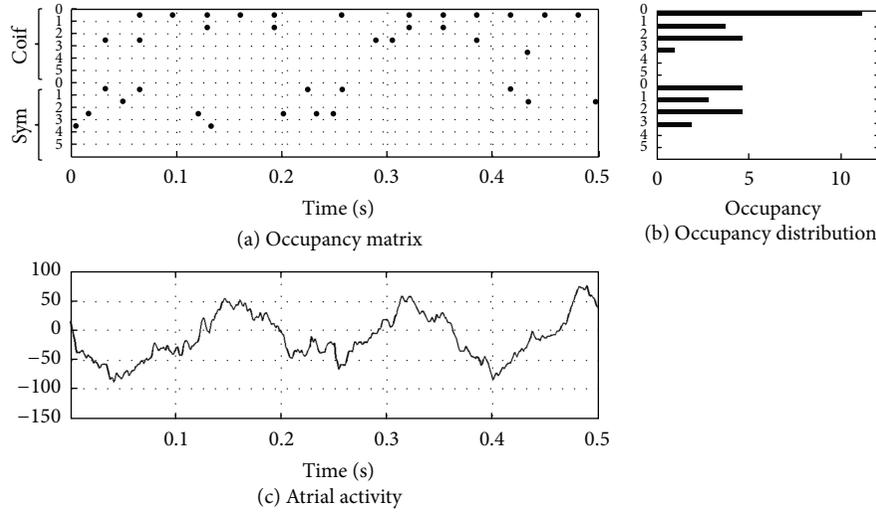


FIGURE 5: (a) The occupancy matrix is shown for an AF-Free case. Each black circle shows wherever there is a MP decomposition at a given time. The  $y$ -axis indicates the corresponding wavelet type and scale of each MP decomposition. (b) The occupancy distribution is shown for each wavelet type and scale. (c) The corresponding AA segments for the occupancy matrix. Only 0.5-second duration of the data is shown here for visualization purposes.

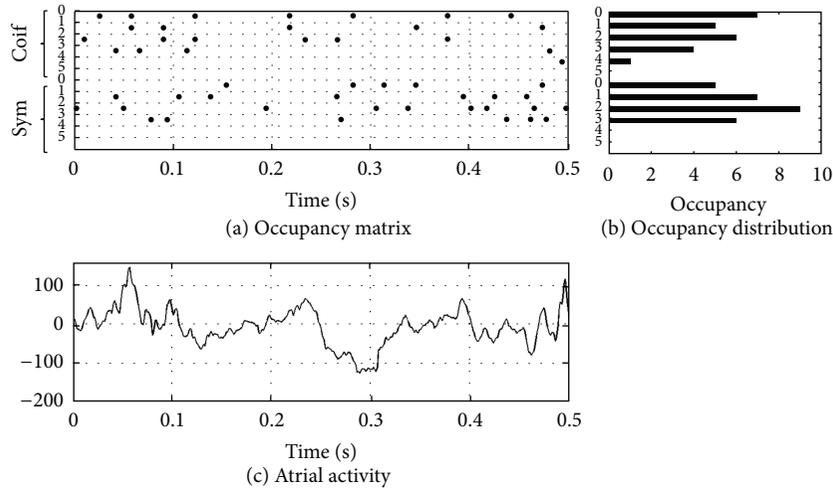


FIGURE 6: (a) The occupancy matrix is shown for a case of AF-Relapse. Each black circle shows wherever there is a MP decomposition at a given time. The  $y$ -axis indicates the corresponding wavelet type and scale of each MP decomposition. (b) The occupancy distribution is shown for each wavelet type and scale. (c) The corresponding AA segments for the occupancy matrix. Only 0.5-second duration of the data is shown here for visualization purposes.

the AF-Relapse and AF-Free feature vectors by a quadratic surface. To evaluate the classification performance, we used a leave-one-out cross validation procedure where the data of one patient was withheld in each trial. Hence, the classification procedure is repeated in 40 trials corresponding to each of the patients and the training sets consist of the feature vectors from the entire database with the exception of the single patient withheld. The feature vectors from the patient under study are the test data. At every trial, the posterior probabilities of the left-out data corresponding to the AF-Free and AF-Relapse classes are recorded. A final receiver operating curve (ROC) is obtained using

the collected posterior probabilities after all the 40 trials are completed.

### 3. Results

The proposed feature extraction and classification algorithm was applied to the dataset described in Section 2.1. The ROC of the QDA and leave-one-out cross validation is shown in Figure 8. According to this analysis, area under the curve (AUC) is 0.97, which is slightly higher than AUC of the linear discriminant analysis (LDA) method, which is 0.94. The best sensitivity and specificity values are 100% (20 out

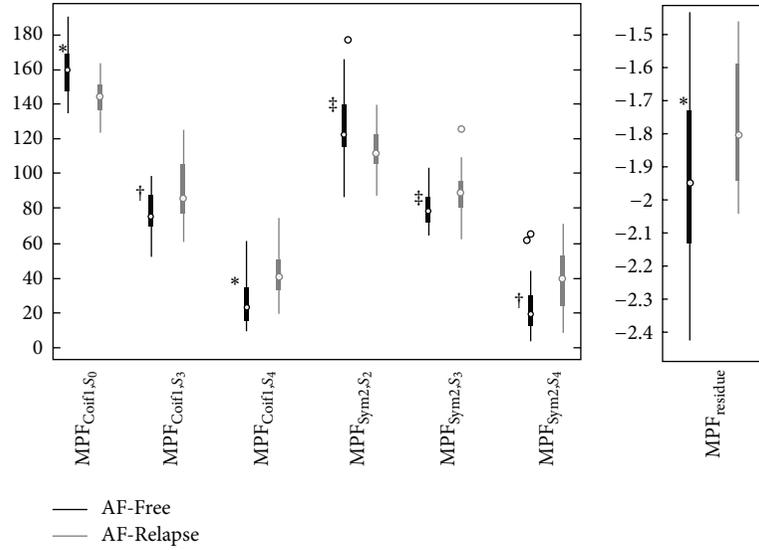


FIGURE 7: The seven statistically significant features are shown in this plot. \*  $P$  value  $< 0.005$ , †  $P$  value  $< 0.01$ , and ‡  $P$  value  $< 0.05$ .

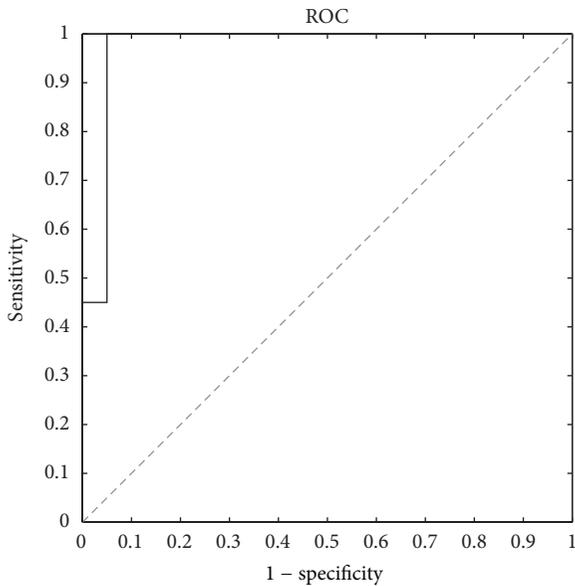


FIGURE 8: Receiver operating characteristic analysis of the QDA classification scores using leave-one-out cross validation. The AUC is 0.97 and the best sensitivity and specificity are 100% and 95%, respectively.

of 20) and 95% (19 out of 20), respectively. Except one case in the AF-Free class, all the data are perfectly classified using the novel features that are proposed in this paper.

**3.1. Noise and QRST Residual.** The preprocessing for the extraction of the AA signal is based on the average beat subtraction where the segments with a single morphology and the least QRST residuals are selected for the further analysis. It is common among AF patients that the ECG signal contains more than a single QRST morphology, which can

result in extensive amount of QRST residue in the estimated AA signal. Another common problem is the changes in the QRST due to respiration and mismatches in the alignment of the QRST template with each QRST complex. In order to avoid computational artifact in the AA signal, we employ a correlation-based analysis to only consider the data with insignificant variability in the QRST complex. In our dataset, only 4 records consisted of a single morphology (i.e., no beat with the correlation of less than 90% with the QRST template). The remaining records had  $32 \pm 30$  (8.5%  $\pm$  7.6%) beats that showed a correlation coefficient of less than 90% with the QRST template. In total,  $4.6 \pm 2.0$  ECG segments ( $66 \pm 42$  seconds long) were selected for the AA extraction. Kurtosis was computed for each segment and the 10-second long ECG that showed the least kurtosis value was selected from each record. The average kurtosis value for the selected segments was  $-0.80 \pm 1.01$ . We repeated the feature extraction and classification algorithm on randomly selected segments from each record (i.e., there was no constraint on the QRST correlation and kurtosis). The performance of the algorithm significantly dropped, which implies that selecting the noise and artifact-free segments is essential for the successful analysis of AA signals.

**3.2. Relationship of Wavelet Type and Cardioversion Outcomes.** Several observations can be made from the statistical analysis performed to select the significant MPF features. Scale 0 ( $S_0$ ) of the Coiflet1 wavelet and Scale 2 ( $S_2$ ) of the Symlet2 wavelet model (i.e., decompose) the activations in the AF-Free signals. This behavior can be seen in Figure 7 which shows the elevated activities of  $MPF_{Coif1,S_0}$  and  $MPF_{Sym2,S_2}$  for the AF-Free data compared to the AF-Relapse data. It can also be observed from the right-hand side plots in Figure 4 that the AF-Relapse signal contains more elevated activations in the higher scales (i.e.,  $S_3$  and  $S_4$ ) compared to the AF-Free signal. This can also be seen in Figure 7 where AF-Relapse

TABLE 2: Comparison of signal processing methods.

Method	Study size	Significance	Sensitivity	Specificity
P-wave duration 1997 [10]	35	0.001	73%	71%
Heart rate variability 2001 [29]	93	—	76%	90%
Fibrillatory rate 2003 [30]	44	0.021	—	—
Clustering of RR intervals 2004 [31]	66	0.034	—	—
P-wave duration 2005 [32]	118	0.0001	72%	77%
P-wave duration 2006 [8]	122493	0.02	90%	21%
Fibrillatory rate 2006 [13]	175	0.0001	79%	80%
Fibrillatory rate 2006 [6]	54	0.002	—	—
Harmonic decay 2006 [6]	54	0.0004	92%	47%
Sample entropy 2011 [14]	66	0.02	—	—
Wavelet transform 2007 [33]	30	—	100%	89%
P-wave dispersion 2011 [9]	26	0.001	86%	95%
Wavelet sample entropy 2008 [34]	40	—	95%	93%
Proposed MP-based method 2014	40	0.005	100%	95%

data presents a higher activity at  $MPF_{Coif1,S_3}$ ,  $MPF_{Coif1,S_4}$ ,  $MPF_{Sym2,S_3}$ , and  $MPF_{Sym2,S_4}$ . In addition, comparing Figures 5 and 6 one can see that the occupancy matrix and distribution of the AF-Free signal are more concentrated at the lower scales while the AF-Relapse signal demonstrates a wider distribution. Hence, our observation is aligned with the literature [2, 35, 36] supporting that worsening AF is associated with a more disorganized atrial signal in the surface ECG. Our analysis suggests that the elevated activation of the higher scale wavelets in the AF-Relapse cases may be used as a predictor of disorganization and disturbances in AA signals.

### 3.3. Comparison with Other Related Studies for AF Progression.

Table 2 lists results from a variety of previous studies proposed for the successful prediction of DCE cardioversion. It is worth mentioning that the results in this table were obtained using different datasets and the patient population used is important in explaining the differences among different studies. However, one may conclude that the proposed method provides a comparable if not better predictive capability compared to the other algorithms.

We assessed the performance of the atrial frequency rate (AFR) on our dataset as an important surface ECG statistic obtained during fibrillation. Previous studies, such as [13], show that AFR is correlated with endocardial measurements of cycle length. Moreover, as a measure of AA organization, AFR is significantly associated with risk of recurrence after therapy. Elevated AFR is generally understood to indicate a worsening of AF, perhaps associated with the progression of the disease via electrical remodeling [37]. We applied the AFR method to our dataset and found that the AFR was elevated in patients who had a recurrence of AF at follow-up ( $P = 0.012$ ) and resulted in 58% and 75% sensitivity and specificity. Therefore, our analysis supported the significance of AFR in predicting the successful DCE cardioversion; however, as it is evident from these results we need to combine the AFR with some additional features in order to provide a successful aggregate score. For example, in Figure 4 the AFR

did not find a significant difference between the AF-Free and AF-Relapse signals with the calculated AFR of 6.5 Hz and 6.4 Hz, respectively; however, the proposed multiresolution-based technique successfully differentiated the two cases. This result may be explained by the difference between the structures of the two methods. The proposed technique provides a distribution of the organization at different time and frequency scales while AFR is bounded by the time and frequency resolution of the Fourier Transform and can only provide the information about the overall frequency content over a given length of the signal. The performance of the proposed algorithm indicates that both the wavelet type and scale are important in predicting the successful postcardioversion patients. However, a larger population is required to further assess the success of the proposed MP-based analysis in a future study.

## 4. Conclusion

In this study, we proposed a novel analysis for the structure of the atrial activity to predict the success of DCE cardioversion AF therapy after 1 month following the therapy. We developed novel features from MP decomposition, performed a statistical evaluation, and selected 7 significant MP features. The extracted MP features were used in a quadratic discriminant analysis-based classification to predict the outcome of DCE cardioversion in our database. A leave-one-out evaluation demonstrated that our proposed algorithm provides a promising noninvasive indicator of the outcome with 100% and 95% sensitivity and specificity, respectively. Given the significant outcome, it may be concluded that our multiresolution-based signal decomposition technique yields novel insights into organization of the atrial activations that could improve the prediction of the successful postcardioversion patients. Further studies on wider databases could determine the reliability of the proposed computational approach as a new computer-aided clinical decision support system that could successfully predict the outcome of

DCE cardioversion and may potentially guide the care of AF patients.

## Conflict of Interests

The authors have no competing interests to declare.

## References

- [1] V. Fuster, L. E. Rydn, D. S. Cannom et al., "2011 ACCF/AHA/HRS focused updates incorporated into the ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation," *Journal of the American College of Cardiology*, vol. 57, no. 11, pp. e101–e198, 2011.
- [2] M. C. E. F. Wijffels, C. J. H. J. Kirchhof, R. Dorland, and M. A. Allesie, "Atrial fibrillation begets atrial fibrillation: a study in awake chronically instrumented goats," *Circulation*, vol. 92, no. 7, pp. 1954–1968, 1995.
- [3] T. Rostock, D. Steven, B. Lutomsky et al., "Atrial fibrillation begets atrial fibrillation in the pulmonary veins: on the impact of atrial fibrillation on the electrophysiological properties of the pulmonary veins in humans," *Journal of the American College of Cardiology*, vol. 51, no. 22, pp. 2153–2160, 2008.
- [4] T. Lundstrom and L. Ryden, "Chronic atrial fibrillation. Long-term results of direct current conversion," *Acta Medica Scandinavica*, vol. 223, no. 1, pp. 53–59, 1988.
- [5] P. Žohar, M. Kovačič, M. Brezocnik, and M. Podbregar, "Prediction of maintenance of sinus rhythm after electrical cardioversion of atrial fibrillation by non-deterministic modelling," *Europace*, vol. 7, no. 5, pp. 500–507, 2005.
- [6] F. Holmqvist, M. Stridh, J. E. P. Waktare et al., "Atrial fibrillation signal organization predicts sinus rhythm maintenance in patients undergoing cardioversion of atrial fibrillation," *Europace*, vol. 8, no. 8, pp. 559–565, 2006.
- [7] P. R. B. Barbosa, A. D. S. Bomfim, E. C. Barbosa et al., "Spectral turbulence analysis of the signal-averaged electrocardiogram of the atrial activation as predictor of recurrence of idiopathic and persistent atrial fibrillation," *International Journal of Cardiology*, vol. 107, no. 3, pp. 307–316, 2006.
- [8] M. H. Raitt, A. S. Volgman, R. G. Zoble et al., "Prediction of the recurrence of atrial fibrillation after cardioversion in the Atrial Fibrillation Follow-up Investigation of Rhythm Management (AFFIRM) study," *American Heart Journal*, vol. 151, no. 2, pp. 390–396, 2006.
- [9] N. Başar, M. Gürel, F. Ozcan et al., "Diagnostic accuracy of P-wave dispersion in prediction of maintenance of sinus rhythm after external cardioversion of atrial fibrillation," *Anadolu Kardiyoloji Dergisi*, vol. 11, no. 1, pp. 34–38, 2011.
- [10] G. Opolski, P. Ścisło, J. Stanisławska, A. Górecki, R. Steckiewicz, and A. Torbicki, "Detection of patients at risk for recurrence of atrial fibrillation after successful electrical cardioversion by signal-averaged p-wave ECG," *International Journal of Cardiology*, vol. 60, no. 2, pp. 181–185, 1997.
- [11] R. Alcaraz, J. J. Rieta, and F. Hornero, "Non-invasive atrial fibrillation organization follow-up under successive attempts of electrical cardioversion," *Medical & Biological Engineering & Computing*, vol. 47, no. 12, pp. 1247–1255, 2009.
- [12] M. Holm, S. Pehrson, M. Ingemansson et al., "Non-invasive assessment of the atrial cycle length during atrial fibrillation in man: introducing, validating and illustrating a new ECG method," *Cardiovascular Research*, vol. 38, no. 1, pp. 69–81, 1998.
- [13] F. Holmqvist, M. Stridh, J. E. P. Waktare, L. Sörnmo, S. B. Olsson, and C. J. Meurling, "Atrial fibrillatory rate and sinus rhythm maintenance in patients undergoing cardioversion of persistent atrial fibrillation," *European Heart Journal*, vol. 27, no. 18, pp. 2201–2207, 2006.
- [14] R. Petersson, F. Sandberg, P. G. Platonov, and F. Holmqvist, "Noninvasive estimation of organization in atrial fibrillation as a predictor of sinus rhythm maintenance," *Journal of Electrocardiology*, vol. 44, no. 2, pp. 171–175, 2011.
- [15] C. Vayá and J. J. Rieta, "Time and frequency series combination for non-invasive regularity analysis of atrial fibrillation," *Medical & Biological Engineering & Computing*, vol. 47, no. 7, pp. 687–696, 2009.
- [16] C. T. January, L. S. Wann, J. S. Alpert et al., "2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American college of cardiology/American heart association task force on practice guidelines and the heart rhythm society," *Journal of the American College of Cardiology*, vol. 64, no. 21, pp. 2246–2280, 2014.
- [17] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [18] J.-P. Couderc, "The Telemetric and Holter ECG Warehouse initiative (THEW): a data repository for the design, implementation and validation of ECG-related technologies," in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10)*, pp. 6252–6255, September 2010.
- [19] S. Matsuo, N. Lellouche, M. Wright et al., "Clinical predictors of termination and clinical outcome of catheter ablation for persistent atrial fibrillation," *Journal of the American College of Cardiology*, vol. 54, no. 9, pp. 788–795, 2009.
- [20] A. Bollmann, D. Husser, M. Stride et al., "Frequency measures obtained from the surface electrocardiogram in atrial fibrillation research and clinical decision-making," *Journal of Cardiovascular Electrophysiology*, vol. 14, no. 10, pp. S154–S161, 2003.
- [21] J. J. Rieta, F. Castells, C. Sanchez, V. Zarzoso, and J. Millet, "Atrial activity extraction for atrial fibrillation analysis using blind source separation," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, pp. 1176–1186, 2004.
- [22] M. Stridh and L. Sörnmo, "Spatiotemporal QRST cancellation techniques for analysis of atrial fibrillation," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 1, pp. 105–111, 2001.
- [23] J. Slocum, E. Byrom, L. McCarthy, A. Sahakian, and S. Swiryn, "Computer detection of atrioventricular dissociation from surface electrocardiograms during wide QRS complex tachycardias," *Circulation*, vol. 72, no. 5, pp. 1028–1036, 1985.
- [24] J. J. Rieta, F. Castells, C. Sánchez, V. Zarzoso, and J. Millet, "Atrial activity extraction for atrial fibrillation analysis using blind source separation," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, pp. 1176–1186, 2004.
- [25] K. Balasundaram, S. Masse, K. Nair, and K. Umapathy, "A classification scheme for ventricular arrhythmias using wavelets analysis," *Medical and Biological Engineering and Computing*, vol. 51, no. 1–2, pp. 153–164, 2013.
- [26] M. Kaleem, B. Ghoraani, A. Guergachi, and S. Krishnan, "Pathological speech signal analysis and classification using empirical mode decomposition," *Medical and Biological Engineering & Computing*, vol. 51, no. 7, pp. 811–821, 2013.
- [27] K. J. Blinowska, W. W. Jędrzejczak, and K. Kwaskiewicz, "Matching pursuit with asymmetric functions for signal decomposition

- and parameterization,” in *Proceedings of the 13th Mediterranean Conference on Medical Biological Engineering Computing*, vol. 41, pp. 791–794, 2014.
- [28] H. J. Sih, D. P. Zipes, E. J. Berbari, and J. E. Olgin, “A high-temporal resolution algorithm for quantifying organization during atrial fibrillation,” *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 4, pp. 440–450, 1999.
- [29] F. Lombardi, A. Colombo, B. Basilio et al., “Heart rate variability and early recurrence of atrial fibrillation after electrical cardioversion,” *Journal of the American College of Cardiology*, vol. 37, no. 1, pp. 157–162, 2001.
- [30] A. Bollmann, D. Husser, R. Steinert et al., “Echocardiographic and electrocardiographic predictors for atrial fibrillation recurrence following cardioversion,” *Journal of Cardiovascular Electrophysiology*, vol. 14, supplement 10, pp. S162–S165, 2003.
- [31] M. P. van den Berg, T. van Noord, J. Brouwer et al., “Clustering of RR intervals predicts effective electrical cardioversion for atrial fibrillation,” *Journal of Cardiovascular Electrophysiology*, vol. 15, no. 9, pp. 1027–1033, 2004.
- [32] M. Budeus, M. Hennersdorf, C. Perings, H. Wieneke, R. Erbel, and S. Sack, “Prediction of the recurrence of atrial fibrillation after successful cardioversion with p wave signal-averaged ECG,” *Annals of Noninvasive Electrocardiology*, vol. 10, no. 4, pp. 414–419, 2005.
- [33] J. N. Watson, P. S. Addison, N. Uchaipichat, A. S. Shah, and N. R. Grubb, “Wavelet transform analysis predicts outcome of DC cardioversion for atrial fibrillation patients,” *Computers in Biology and Medicine*, vol. 37, no. 4, pp. 517–523, 2007.
- [34] R. Alcaraz and J. J. Rieta, “A non-invasive method to predict electrical cardioversion outcome of persistent atrial fibrillation,” *Medical & Biological Engineering & Computing*, vol. 46, no. 7, pp. 625–635, 2008.
- [35] L. Faes, G. Nollo, R. Antolini, F. Gaita, and F. Ravelli, “A method for quantifying atrial fibrillation organization based on wave-morphology similarity,” *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 12, pp. 1504–1513, 2002.
- [36] T. H. Everett, L.-C. Kok, R. H. Vaughn, J. R. Moorman, and D. E. Haines, “Frequency domain algorithm for quantifying atrial fibrillation organization to increase defibrillation efficacy,” *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 9, pp. 969–978, 2001.
- [37] A. Bollmann, D. Husser, L. Mainardi et al., “Analysis of surface electrocardiograms in atrial fibrillation: techniques, research, and clinical applications,” *Europace*, vol. 8, no. 11, pp. 911–926, 2006.