

Editorial

Computational Intelligence Techniques in Medicine

**Ezequiel López-Rubio,¹ David A. Elizondo,² Martin Grootveld,³
José M. Jerez,¹ and Rafael M. Luque-Baena⁴**

¹*Department of Computer Languages and Computer Science, University of Málaga, 29071 Málaga, Spain*

²*DIGITS, Faculty of Technology, De Montfort University, The Gateway, Leicester LE1 9BH, UK*

³*Leicester School of Pharmacy, Faculty of Health and Life Sciences, De Montfort University, The Gateway, Leicester LE1 9BH, UK*

⁴*Department of Computer and Network Systems Engineering, University of Extremadura, 06800 Mérida, Spain*

Correspondence should be addressed to Ezequiel López-Rubio; ezeqlr@lcc.uma.es

Received 9 December 2014; Accepted 9 December 2014

Copyright © 2015 Ezequiel López-Rubio 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.

1. Introduction

The advent of the information age, also commonly known as the digital age, has made a profound impact on health sciences. Vast amounts of datasets now flow through the different stages of healthcare organizations, and there is a major requirement to extract knowledge and employ it to improve these centres in all respects.

Intelligent computer systems provide support to health professionals involved both in the medical and managerial contexts. Amongst these systems, computational intelligence approaches have gained increasing popularity given their ability to cope with large amounts of clinical data and uncertain information.

The goal of this special issue is to offer a broad view of this exciting field, the ever-growing importance of which is driven by the increasing availability of data and computational power.

2. Computational Intelligence

Computational intelligence is based on biologically inspired computational algorithms. The key pillars that compose this field are neural networks, genetic algorithms, and fuzzy systems. Neural networks are algorithms that can be used for function approximation or classification problems [1, 2]. They include supervised, unsupervised, and reinforcement learning [3]. Genetic algorithms, however, [4] are search

algorithms inspired by biological genetics. They rely on two main operators, cross-over and mutation. Populations of individuals representing solutions to the problem are created over several generations. The algorithm uses a random-guided approach to optimize problems based on a fitness function. Fuzzy logic [5] is based on fuzzy set theory [6] in order to encompass reasoning that is fluid or approximate rather than fixed and exact. Fuzzy logic variables have “truth” values ranging in degree between 0 and 1 which can also handle “partial truth”.

Computational intelligence techniques have been successfully used in many “real-world” applications in a variety of engineering problems. They can also be used in the medical research domain.

3. Organisation and Plan of the Special Issue

This special issue presents a compilation of research papers describing novel recent strategies and developments on the use of computational intelligence techniques in medicine. It will be of interest to researchers interested in industry, academics, and also to postgraduate students interested in the latest advances and developments in the field of computational intelligence in medicine. Its contents are summarized next.

One of the best-known application fields of these techniques is computer-aided diagnosis (CAD). At this point, two types of methods must be distinguished: those which aim

at providing clinicians with enhanced diagnostic procedures that can lead to improved human decisions, and those that intend to provide a second opinion; that is, given a set of signs, they can indicate or confirm one or more possible disease processes.

Many methods belonging to the first type correspond to image processing algorithms. Indeed, the automated segmentation of magnetic resonance images (MRI) of the brain is employed to determine pathological regions and to plan image-guided surgery. On the other hand, high-resolution computed tomography has become a standard tool to diagnose diffuse lung diseases, and sparse representations can help to identify abnormal patterns. Segmentation is also useful at the microscopic scale, where intracellular calcium variation can be measured by a marked controlled watershed transform. Additionally, the emergence of a number of differing medical imaging techniques brings our attention to the problem of fusing these images. An approach based on the Nonsubsampled Contourlet Transform, which is a directional multiresolution image representation method, is also proposed.

Other diagnostic enhancement procedures are not related to image processing. In this issue, the most relevant criteria to diagnose Guillain-Barré syndrome are studied by the Partitions around Medoids clustering algorithm, which is able to manage both categorical and numerical data since it only requires the distance matrix amongst the samples. The reactivity of blood pressure (BP) to talking episodes is predicted by a hybrid system composed of several soft computing approaches, specifically neural networks, an adaptive neurofuzzy inference system, and support vector machines (SVMs). This increases the precision of BP measures for clinical and research purposes. Screening for prediabetes is carried out by means of a combination of neural networks and SVMs, which is aimed at early intervention to avoid the serious complications associated with the disease. Validation procedures for electroencephalographic (EEG) signal processing by principal component analysis (PCA) are also proposed, which is a valuable tool for many neurological studies. EEG signals are also analyzed by mixed-norm regularization for sensor selection in brain computer interfaces. Finally, the problem of privacy preservation in self-helped medical diagnosis is addressed, where secure two-party computation in wireless sensor networks is ensured in order to develop a system where the patient inserts a health card into an automated teller machine and obtains a diagnostic report.

The second class of methods includes classification systems which are most useful for the early diagnosis of diseases for which there is currently no definitive diagnostic test. Computer-aided screening for Alzheimer's disease based on neuropsychological rating scales is carried out by a hybrid approach which combines rough sets, genetic algorithms, and Bayesian networks, and an expert system for multicriteria decision support in psychotic disorders is proposed. Clinical decision support systems are not limited to diagnosis, as noted in the osteoporosis case, where the system can also recommend treatments and assess the fracture risk.

Management and planning represent the other sections of healthcare organizations that can also benefit from

the applications of computational intelligence. Particle swarm optimization can be employed to enhance blood assignment in blood banks, whilst social network simulation can help to develop effective preventative and interventional strategies for AIDS epidemics.

Last but not least, computational intelligence techniques are also useful tools for probing the improvement of drug manufacturing processes, as demonstrated in another proposal where neural networks are applied to drug tablet visual tracking for high speed mass production.

Ezequiel López-Rubio
David A. Elizondo
Martin Grootveld
José M. Jerez
Rafael M. Luque-Baena

References

- [1] K. Gurney, *An Introduction to Neural Networks*, Routledge, New York, NY, USA, 1997.
- [2] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, Upper Saddle River, NJ, USA, 1999.
- [3] T. Masters, *Practical Neural Network Recipes in C++*, Academic Press, New York, NY, USA, 1993.
- [4] D. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, New York, NY, USA, 1989.
- [5] L. A. Zadeh, "Fuzzy sets," *Information and Computation*, vol. 8, pp. 338–353, 1965.
- [6] L. A. Zadeh, "Fuzzy algorithms," *Information and Computation*, vol. 12, no. 2, pp. 94–102, 1968.

Research Article

A Sparse Representation Based Method to Classify Pulmonary Patterns of Diffuse Lung Diseases

Wei Zhao,¹ Rui Xu,² Yasushi Hirano,¹ Rie Tachibana,³ and Shoji Kido¹

¹Medical Engineering Science, Graduate School of Medicine, Yamaguchi University, Ube 755-8611, Japan

²Global Innovation Research Organization, Ritsumeikan University, Kusatsu 525-8577, Japan

³National Institute of Technology, Oshima College, Oshima 742-2193, Japan

Correspondence should be addressed to Rui Xu; xurui@fc.ritsumei.ac.jp

Received 26 June 2014; Revised 21 August 2014; Accepted 8 September 2014

Academic Editor: José M. Jerez

Copyright © 2015 Wei Zhao 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 applied and optimized the sparse representation (SR) approaches in the computer-aided diagnosis (CAD) to classify normal tissues and five kinds of diffuse lung disease (DLD) patterns: consolidation, ground-glass opacity, honeycombing, emphysema, and nodule. By using the K-SVD which is based on the singular value decomposition (SVD) and orthogonal matching pursuit (OMP), it can achieve a satisfied recognition rate, but too much time was spent in the experiment. To reduce the runtime of the method, the K-Means algorithm was substituted for the K-SVD, and the OMP was simplified by searching the desired atoms at one time (OMP₁). We proposed three SR based methods for evaluation: SR1 (K-SVD+OMP), SR2 (K-Means+OMP), and SR3 (K-Means+OMP₁). 1161 volumes of interest (VOIs) were used to optimize the parameters and train each method, and 1049 VOIs were adopted to evaluate the performances of the methods. The SR based methods were powerful to recognize the DLD patterns (SR1: 96.1%, SR2: 95.6%, SR3: 96.4%) and significantly better than the baseline methods. Furthermore, when the K-Means and OMP₁ were applied, the runtime of the SR based methods can be reduced by 98.2% and 55.2%, respectively. Therefore, we thought that the method using the K-Means and OMP₁ (SR3) was efficient for the CAD of the DLDs.

1. Introduction

Diffuse lung diseases (DLDs) refer to a series of abnormalities that spread out in large areas of the lungs [1]. With the development of the medical imaging technology, at present the high-resolution computed tomography (HRCT) is thought to be the best tool for the diagnosis of the DLDs, because the pulmonary patterns can be accurately analyzed on the HRCT images [2–4]. However, the interpretation of the DLD patterns mainly depends on the radiologists' individual experiences. It is reported that the agreements between the radiologists' first choices were only moderate [5]. So, the subjective differences between the radiologists may lead to the misdiagnosis. Furthermore, the HRCT produces large numbers of axial slices in each scan, which is a big burden for the radiologists. Considering the above reasons, a computer-aided diagnosis (CAD) method is required to provide the radiologists with a "second opinion" for the diagnosis of the DLDs [6–8].

In the past ten years, researchers have proposed several CAD systems to classify the DLD patterns, and most of the conventional methods aim to develop the discriminative features for the classification. For example, Park et al. adopted the statistical moments of the histograms and gray-level run-length matrices (GLRLM) to represent the textural information of the pulmonary patterns [9]. Wang et al. thought that the GLRLM could be partitioned into four areas with clear physical meaning, which can be used to measure the pulmonary textures [10]. Although the features based on the textural information have an excellent performance on the classification of the DLD patterns, these features are still difficult to recognize the pulmonary patterns with inhomogeneous textures. Therefore, researchers have combined the texture-based measures with the geometrical information to design the features with higher discriminative power. In the study [11], the pulmonary patterns were determined by six kinds of physical features, three based on the CT values (mean and standard deviation of CT values, air

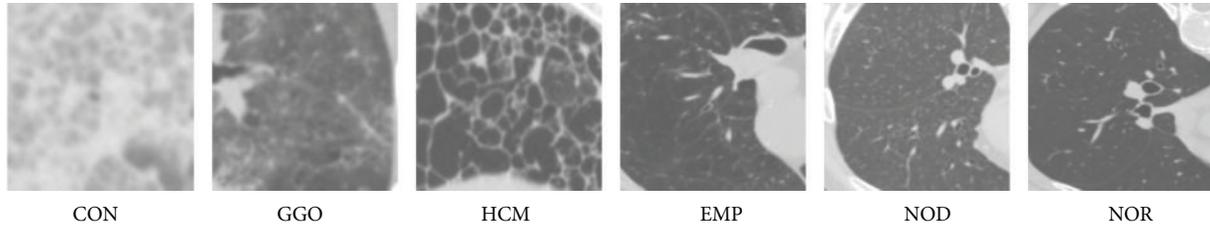


FIGURE 1: Images of six kinds of pulmonary patterns: consolidation (CON), ground-glass opacity (GGO), honeycombing (HCM), emphysema (EMP), nodule (NOD), and normal tissues (NOR).

density components) and three based on the geometrical information (nodular components, line components, and multilocular components). Uppaluri et al. adopted the texture features and geometric fractal dimension (GFD) to classify the pulmonary patterns, where the GFD was used to measure the roughness of textures [12]. In the work [13], the measures based on the histogram, gradient, gray-level cooccurrence matrix (GLCM), and GLRLM were used for texture analysis, and the measures based on the top-hat transformation and clusters of low attenuation areas were used to analyze the shape information. Besides, the local binary pattern (LBP) was employed to quantitatively measure the normal tissues and two subtypes of the emphysema [14].

In this paper, the sparse representation approaches were introduced to recognize the DLD patterns. The main idea of the sparse representation is to approximate the example by a weighted linear combination of a small number of key features (atoms), which are selected from an overcomplete dictionary. It is thought that the sparse representation can improve the performance of the image classification [15–17]. Firstly, the images could be treated as a distribution of a set of representative features, so the sparse representation can encode the semantic information of the images. Secondly, the number of atoms in the dictionary is greater than the dimensionality of the input examples, which means that the approximation of the example is not unique. So, it can find a relative better approximation among the various combinations of atoms. Thirdly, the sparse representation is shown to be robust in the presence of the noise. Due to these advantages, the sparse representation approaches have been applied in the CAD recently. For example, Liu et al. developed a sparse representation based method to detect the colon polyp and lung nodule [18]. Vo and Sowmya trained discriminative dictionaries to classify four kinds of the pulmonary patterns [19]. In the work [20], the dictionary of the texton was learned and used to recognize the normal tissues and three subtypes of the emphysema.

In this work, by adopting the two of the most popular algorithms, the singular value decomposition (SVD) based K-SVD algorithm [21] and orthogonal matching pursuit (OMP), we proposed a sparse representation based method to classify the normal tissues (NOR) and five kinds of the DLD patterns, including the consolidation (CON), ground-glass opacity (GGO), honeycombing (HCM), emphysema (EMP), and nodule (NOD). Figure 1 gives the images of the six kinds of the pulmonary patterns. According to our knowledge,

there is no work aimed at applying the sparse representation approaches to analyze these six kinds of the pulmonary patterns. The proposed method using the K-SVD and OMP achieved a high classification accuracy (greater than 95%) in the experiments, which was thought to be of great potential by the radiologists. However, the runtime of this method was relatively long. Therefore, we also tried to reduce the runtime of the sparse representation based method. Considering that the operation of the K-SVD and OMP spent the most time on the training and testing, respectively, we employed the K-Means to replace the K-SVD and used a simple version of the OMP which was named OMP_1 in the paper. Experimental results show that the replacement of the K-SVD and OMP by the K-Means and OMP_1 can reduce the runtime of the method while keeping the classification accuracy.

There are two major differences against a preliminary version of this work [23]. Firstly, we not only adopted the sparse representation approaches for the classification but also optimized the dictionary learning and sparse coding in this work. Secondly, we changed the experimental data to make the number of training and testing samples approximately equal. This paper is organized as follows. In Section 2, we describe the proposed methods. The experimental results are given and discussed in Section 3. Finally, we conclude the paper in Section 4.

2. Proposed Methods

In this research, we adopted and optimized the sparse representation approaches to classify the normal tissues and five kinds of the DLD patterns on HRCT images. Figure 2 gives the framework of our methods. In the training stage, firstly huge numbers of local features were extracted from the training volumes of interest (VOIs) and used to train an overcomplete dictionary. Secondly, the sparse representation of the local features was calculated according to the given dictionary, and the VOI-level descriptors of the training VOIs were generated by the procedure named spatial pooling. Finally, the descriptors were used to train a support vector machine (SVM) classifier. In the testing stage, after extracting the local features on the testing VOI, the learned dictionary was adopted to calculate the sparse representation of the local features, and then the VOI-level descriptor was generated. At last, the descriptor was fed into the trained classifier and the result was given. In order to easily understand the paper,

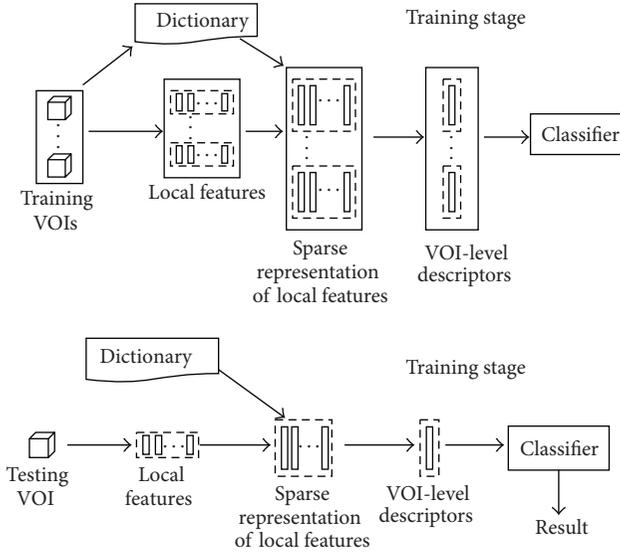


FIGURE 2: The framework of our methods.

we introduce the sparse representation and its optimization at first, and then we describe the other parts of our methods.

2.1. Sparse Representation. Let m examples and the normalized overcomplete dictionary be $\mathbf{y}_i \in \mathbb{R}^n$, $i = 1, 2, \dots, m$, and $\mathbf{D} \in \mathbb{R}^{n \times k}$, $n \ll k$, respectively, the sparse representation of m examples $\mathbf{a}_i \in \mathbb{R}^k$ can be formulated as

$$\min_{\mathbf{D}, \mathbf{a}} \sum_{i=1}^m \|\mathbf{y}_i - \mathbf{D}\mathbf{a}_i\|_2^2 \quad \text{subject to } \|\mathbf{a}_i\|_0 \leq T, \quad (1)$$

where the $\|\cdot\|_0$ means the l^0 -norm indicating the sparsity of the vector (number of nonzero entries in the vector), and T is the threshold of the sparsity. It could be found that the \mathbf{a}_i can be thought of as the coefficients of the atoms. There were two main components in the operation: (1) training an overcomplete dictionary \mathbf{D} (dictionary learning); (2) calculating the sparse representation of the input example \mathbf{a} according to a given dictionary (sparse coding).

By adopting the K-SVD and OMP for the dictionary learning and sparse coding, respectively, we proposed a method that was called SRI in the paper. The K-SVD trained the dictionary by alternatively updating the coefficients with the fixed dictionary (sparse coding stage) and then updating the dictionary with the fixed coefficients (dictionary updating stage) until the stop condition was met.

- (1) In the sparse coding stage, it was recommended to use the OMP, a greedy technique [21]. In the beginning, the solution support was empty and an initial residual vector was evaluated by the input example. At each iteration, the atom that had the largest inner product with the residual vector was added to the support. Then the sparse approximation of the example was calculated according to the support, and the residual was updated. These processes were repeated until

the number of atoms in the support was greater than the sparsity threshold.

- (2) In the dictionary updating stage, the columns of the \mathbf{D} (atoms of the dictionary) were updated sequentially. When the i th atom was being updated, the reconstruction matrix except the current atom was restricted by choosing the examples which were reconstructed by using the current atom and then decomposed by the SVD. The first left-singular vector was adopted to update the atom.

After training the dictionary by the K-SVD, the SRI also used the OMP to calculate the sparse representation of local features, the same as the sparse coding stage in the K-SVD. There were two parameters in the sparse representation, the number of atoms and the sparsity in the approximation. We adjusted the number of atoms from 500 to 3000 with an interval of 500 and the sparsity from 2 to 14 with an interval of 2 in the experiments. We present the way of optimizing the parameters in Section 3.2.

2.2. Optimization of Sparse Representation. The runtime of the CAD method is also an important criterion in the clinical practice. The SRI that uses the K-SVD and OMP (see Section 2.1) can achieve a high classification accuracy, but the runtime of the SRI was relatively long (see Section 3.3). Considering that the dictionary learning and sparse coding spent the most time on the training and testing, respectively, another aim of our research was to optimize these two steps.

Although it might be useless in the clinical workflow, we thought that the optimization of the dictionary learning can be convenient for the developers to update the existing methods. It is reported that the K-Means algorithm can achieve a competitive performance on the natural image classification with the K-SVD when the same number of atoms was used [24]. Moreover, the K-Means could be treated as a simple version of the K-SVD. In the dictionary updating stage, the average operation and SVD were adopted to update the atoms by the K-Means and K-SVD, respectively. And in the sparse coding stage, the K-Means set the coefficient of the closest atom to 1 (the values of other coefficients were 0), while the sophisticated OMP algorithm was adopted in the K-SVD. It can be deduced that the K-Means would need a shorter runtime than the K-SVD to train the dictionary. Furthermore, the K-Means can be efficiently implemented by using the k -dimensional tree (k -d tree) technique. Due to the above reasons, we tried to adopt the K-Means as a substitute of the K-SVD to train the dictionary.

In the SRI, the OMP was adopted as the solver of the sparse coding, which iteratively calculated the sparse approximation of examples, and only one atom was added to the support at each iteration. It can be deduced that the runtime of the method may be reduced by selecting enough atoms at one time. So, this approach was named OMP₁ in the paper. After arranging the inner products of the atoms and example in a descending order, the first sufficient numbers of atoms were treated as the solution support and then used to calculate the sparse representation of the example. Algorithm 1 gives the OMP₁. Although the residual error of

<p>Input: The example $\mathbf{y} \in \mathbb{R}^n$ and normalized dictionary $\mathbf{D} \in \mathbb{R}^{n \times k}$, $n \ll k$</p> <p>Output: The sparse representation of the example vector $\mathbf{a} \in \mathbb{R}^k$</p> <p>Initialization: Initial solution support $\Omega = \phi$, and sparsity constraint T</p> <p>Computation:</p> <p>(1) Choose T atoms that have the largest inner products with example $\mathbf{y} \cdot \mathbf{d}_1 \leq \mathbf{y} \cdot \mathbf{d}_2 \leq \dots \leq \mathbf{y} \cdot \mathbf{d}_T \leq \dots \leq \mathbf{y} \cdot \mathbf{d}_k$</p> <p>(2) Add the selected atoms to the solution support $\Omega = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_T\}$</p> <p>(3) Compute the sparse representation of the input example $\mathbf{a} = (\Omega^T \Omega)^{-1} \Omega \mathbf{y}$</p>
--

ALGORITHM 1: The OMP₁ algorithm.

TABLE 1: Three proposed methods for evaluation of sparse representation approaches.

(a) Proposed methods		
Method	Dictionary learning	Sparse coding
SR1	K -SVD	OMP
SR2	K -Means	OMP
SR3	K -Means	OMP ₁

(b) Experiments on the evaluation of sparse representation approaches	
Evaluation	Comparison of methods
K -SVD versus K -Means	SR1 and SR2
OMP versus OMP ₁	SR2 and SR3

the OMP₁ would be larger than the OMP, the performance of this approach can be ensured under a certain sparsity constraint [25].

In order to examine the performances of the sparse representation approaches, we constructed another two sparse representation based methods: SR2 (K -Means+OMP) and SR3 (K -Means+OMP₁) in this work. The replacement of the K -SVD by the K -Means was evaluated by comparing the SR1 and SR2 (both of the two methods adopted the OMP for the sparse coding), and the substitution of the OMP by the OMP₁ was evaluated by comparing the SR2 and SR3 (both of the two methods adopted the K -Means for the dictionary learning). Table 1 summarizes the three sparse representation based methods and the experiments on the evaluation of the sparse representation approaches.

2.3. Calculation of Local Features. It is thought that the DLD patterns can be featured by a combination of CT values and measures based on the geometrical information. In this work, we used the local features proposed in the work [26] which adopted the eigenvalues of the Hessian matrix to measure the geometrical information. The local features were calculated at each sampling point on the VOI as the following procedures. Firstly, a cubic-shape patch was constructed by sampling on the VOI whose center was located on the sampling point, and four kinds of the statistical moments were calculated on this patch: mean, standard deviation, skewness, and kurtosis.

Then the eigenvalues of the Hessian matrix were calculated for each voxel within the patch. Let the eigenvalues be λ_1, λ_2 , and λ_3 , $\lambda_1 \geq \lambda_2 \geq \lambda_3$. We arranged the eigenvalues in the order of the position. So, three new patches were constructed whose components were λ_1, λ_2 , and λ_3 respectively, and the same moments were calculated on these three eigenvalues based patches. Finally, the moments calculated on all four patches were concentrated into a 16-dimensional vector as the feature vector. In the experiments, the step of the sampling points was set to $4 \times 4 \times 4$. And the size of the patch was a parameter, which was adjusted from $2 \times 2 \times 2$ to $6 \times 6 \times 6$. The way of tuning the parameter is described in Section 3.2.

2.4. Spatial Pooling. The procedure of the spatial pooling was used to summarize the sparse representation of the local features over the regions into a VOI-level descriptor for each VOI. These descriptors were used as the input vectors of the classifier. We adopted one of the most popular choices, the average pooling in the work, which could be seen as an average operation of the vectors. Let $\mathbf{z} \in \mathbb{R}^k$ be the VOI-level descriptor, let $\mathbf{a} \in \mathbb{R}^k$ be the sparse representation vectors, and let $\{\cdot\}_t$ be the t th element of the vector. The average pooling of m vectors is given by

$$\mathbf{z}_t = \frac{1}{m} \sum_{i=1}^m \mathbf{a}_{it}, \quad t = 1, 2, \dots, k. \quad (2)$$

2.5. Classification. In the research, we adopted the support vector machine (SVM) as the classifier to recognize the descriptors generated in the spatial pooling. We used a version named LIBSVM [27]. It is reported that the sparse representation based classification with the linear kernel can achieve a competitive performance and smaller computational cost than the nonlinear kernels [16]. So, we employed the LIBSVM with a linear kernel. The kernel is given by

$$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j, \quad (3)$$

where \mathbf{x}_i and \mathbf{x}_j are both descriptors. Because the SVM was originally designed as the binary (two-class) classifier, the LIBSVM adopted the one-against-one technique to extend the binary SVM classifier for the multiclass tasks. There is one

parameter in the classifier: soft-margin penalty C . The way of adjusting the parameter is described in Section 3.2.

3. Experiments and Results

3.1. Data. We obtained 117 scans from 117 subjects from Tokushima University Hospital in Japan. All HRCT scans were acquired by Toshiba Aquilion 16-row multislice CT when edge-enhanced filtering was not applied. A tube voltage of 120 kVp and current of 250 mAs were used. The resolution of scans was 512×512 , and the in-plane resolution was about 0.6 mm. The slice thickness was 1 mm.

The VOIs were constructed according to the following procedures. (1) All scans were reviewed by a radiologist, and a maximum of three axial slices was selected from the top, middle, and bottom parts of the lungs, respectively, in each scan. Only one kind of the pulmonary pattern dominantly existed on each selected slice, and the radiologist should indicate what the dominant texture was and where it existed. (2) Another two radiologists reviewed the results of the first radiologist. Only the slices which were thought to be correct by both radiologists were selected. (3) The regions of the pulmonary patterns on the selected slices were marked by all three radiologists, respectively, and the common regions chosen by the radiologists were saved. (4) The grids with a size of 32×32 were overlaid on the slices, and the square-shaped patches were constructed where the regions marked by the radiologists should take more than 70% area of the patches. (6) The VOIs with a size of $32 \times 32 \times 32$ were constructed. The patches were treated as the central-axial slice of VOIs.

3.2. Experimental Setting. In the experiment, we separated the VOIs into two independent sets. One set (1161 VOIs) was adopted as the training set to optimize the parameters of the methods and then train the methods with the optimal parameters. The other set (1049 VOIs) was used as the testing set to evaluate the performances of the methods. There was no cross subject between the two sets. The number of VOIs of each type of patterns for the training and testing is summarized in Table 2. All methods were operated on the server with a 2.8 GHz Intel Core i7 CPU and 24 GB RAM.

There were four kinds of parameters in the proposed methods: the size of cube-shape patches, the number of atoms, the sparsity of the sparse representation, and the parameter related to the classifier. We tuned the values of the patch size from $2 \times 2 \times 2$ to $6 \times 6 \times 6$ with a step of $1 \times 1 \times 1$, the number of atoms from 500 to 3000 with an interval of 500, and the sparsity from 2 to 14 with an interval of 2. The parameter of the SVM classifier was set to $2^{-2}, 2^{-1}, \dots, 2^{11}, 2^{12}$. These parameters were simultaneously optimized by a 20-fold cross-validation test on the training set. The combination of the parameters which achieved the best overall accuracy in the cross-validation test was chosen as the optimal parameters. The results of the proposed methods in the cross-validation were given in Figure 3. Figure 3(a) shows that when the patch size was nearly to the step of sampling point ($4 \times 4 \times 4$), the overall accuracy was near its maximum. Figure 3(b) illustrates that the raising of

TABLE 2: Number of VOIs in the training and testing set.

	CON	GGO	HCM	EMP	NOD	NOR	Total
Training set	49	170	221	323	113	285	1161
Testing set	45	160	204	275	92	273	1049

the number of atoms can improve the performance of the methods. Figure 3(c) shows that, with the increasing of the sparsity, the overall accuracy of the SR1 and SR2 remained, but the SR3 was decreased.

3.3. Three Kinds of Baseline Methods. We compared the proposed methods with three kinds of state-of-the-art published techniques, which were called SDF [11], CSE [28], and BOW [29], respectively. The parameters of the baseline methods were optimized in the same way as the proposed methods.

- (1) In the work [11], the pulmonary patterns were determined by the six kinds of specially designed features. So, this method was called SDF in the paper. These six features were mean and standard deviation of CT values, air density components, nodular components, line components, and multilocular components. A three-layered artificial neural network (ANN) with back-propagation algorithm was adopted as the classifier. In the work [11], the number of hidden units in the ANN was empirically set to 10. We adjusted the number of hidden units from 5 to 30 with an interval of 5 in the experiments. Because 2D regions of interest (ROIs) were required by the SDF, we used the central slices in the axial direction of the VOIs as the ROIs in the experiments.
- (2) In the work [28], the signature of the VOI was used for the classification. The signature was defined as the centroids and the weights of the clusters (number of voxels in the clusters), and the K-means algorithm was used to calculate the centroids of the clusters. In order to reduce the computational cost, the canonical signatures for each class were generated by combining and reclustering the signatures of the training data. The earth mover’s distance (EMD) approach was adopted to measure the similarity between the two signatures, and the nearest neighbor (NN) was employed as the classifier. In the classification, the VOIs were recognized by comparing the signatures of the VOIs with the canonical signatures. Because the canonical signatures and earth mover’s distance (EMD) were used, this method was called CSE in the paper. The CSE had only one parameter: the number of clusters. Considering that the large value was suggested to be avoided, we adjusted the number of clusters from 5 to 60 with a step of 5.
- (3) The work [29] adopted a model named “bag-of-words” (also named bag-of-features) to generate the VOI-level descriptors, so this method was called BOW in the paper. The main idea of the bag-of-words was to train a code-book (dictionary) at first and then use the histograms of the words (atoms)

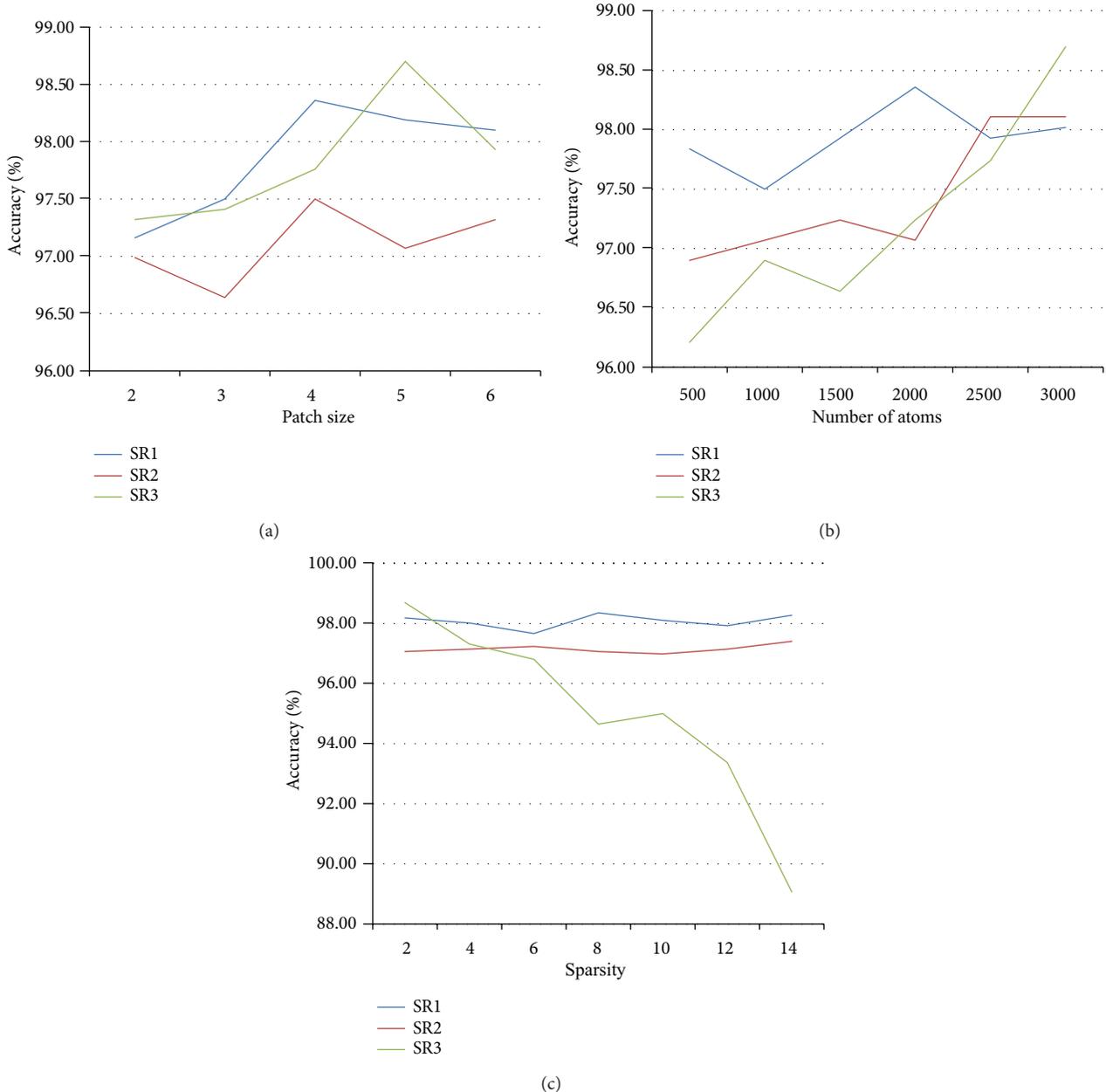


FIGURE 3: Overall accuracy of the proposed methods in the cross-validation; (a) the patch size was adjusted when the number of atoms and sparsity parameters were fixed; (b) the number of atoms was adjusted when the patch size and sparsity were fixed; (c) the sparsity was adjusted when the patch size and number of atoms were fixed.

in the code-book to represent the images. These histograms could be used as the input vectors of the classifier. In the experiments, the K-Means algorithm was adopted to cluster the local features, and the centers of the clusters were saved as the words of the code-book. The number of words was adjusted from 50 to 400 with an interval of 50. The local features adopted in the work [29] were the same as the proposed methods, so we adjusted the values of the patch size from $2 \times 2 \times 2$ to $6 \times 6 \times 6$ with a step of $1 \times 1 \times 1$, the same as proposed methods. The SVM was adopted as the classifier. Considering that the χ^2

kernel achieved the best result in the work [29], the LIBSVM was employed with the χ^2 kernel. Equation (4) gives the χ^2 kernel, where α is the parameter for the kernel and \mathbf{x}_i and \mathbf{x}_j are both histograms with k -bins:

$$H(\mathbf{x}_i, \mathbf{x}_j) = \exp \left[-\alpha \sum_{t=1}^k \frac{(\mathbf{x}_{it} - \mathbf{x}_{jt})^2}{\mathbf{x}_{it} + \mathbf{x}_{jt}} \right]. \quad (4)$$

The possible values of the soft-margin penalty and α were set to be $2^{-2}, 2^{-1}, \dots, 2^{11}, 2^{12}$ and $2^{-10}, 2^{-9}, \dots, 2^1$, respectively.

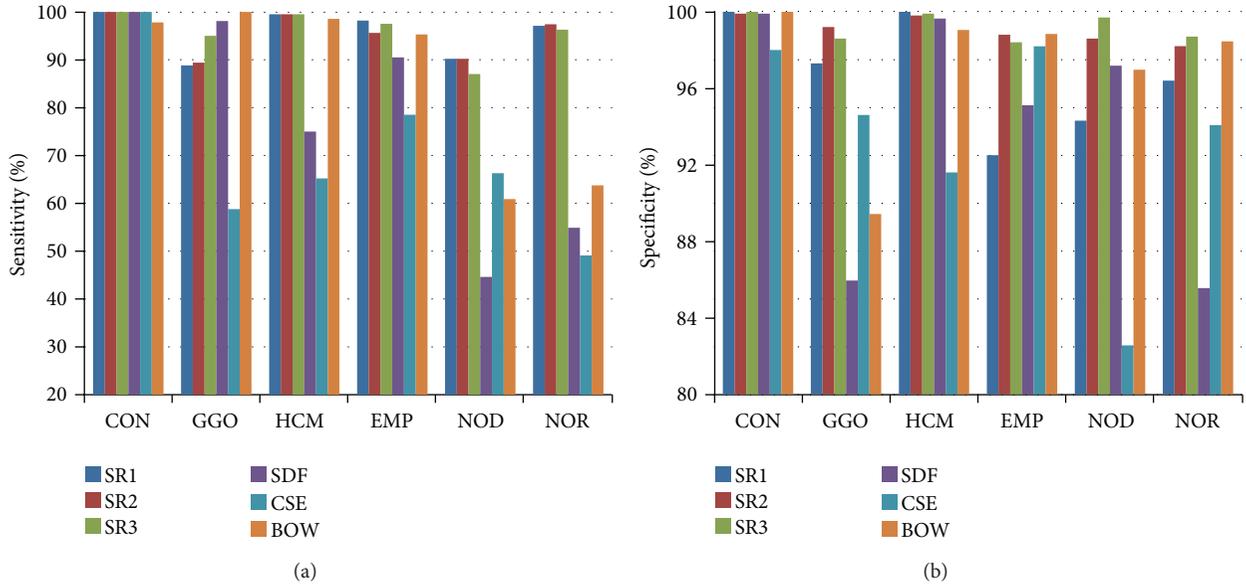


FIGURE 4: Sensitivity and specificity of each method for the pulmonary patterns with optimal parameters.

TABLE 3: Overall accuracy of each method with optimal parameters.

Method	Overall accuracy	Optimal parameter
SR1	96.1%	Patch size: $4 \times 4 \times 4$, number of atoms: 2000, sparsity: 8, C : 0.5
SR2	95.6%	Patch size: $5 \times 5 \times 5$, number of atoms: 2000, sparsity: 2, C : 0.25
SR3	96.4%	Patch size: $3 \times 3 \times 3$, number of atoms: 3000, sparsity: 2, C : 16
SDF	75.8%	Number of hidden units: 10
CSE	65.1%	Number of clusters: 25
BOW	85.5%	Patch size: $3 \times 3 \times 3$, number of atoms: 300, g : 1.0, C : 4.0

3.4. Experimental Results. Table 3 gives the overall accuracy of each method with the optimal parameters on the testing set. The sparse representation based methods achieved better results than the baseline methods (SR1: 96.1%, SR2: 95.6%, SR3: 96.4% versus SDF: 75.8%, CSE: 65.1%, BOW: 85.5%). Figure 4 shows that the sensitivity and specificity of the proposed methods for each pulmonary pattern were all beyond 90%, better than the baseline methods. Additionally, Table 4 shows that the P values of the statistical differences (calculated by the McNemar’s test) for the proposed methods against the baseline methods were all smaller than 0.0001, which means that there were significant differences between the methods.

On the other hand, Table 5 compares the runtime of the proposed methods with the optimal parameters. When the K-SVD was replaced by the K-Means, the runtime of the dictionary learning can be decreased by 98.2% (SR1: 13520 s versus SR2: 241 s). When the OMP_1 was substituted

for the OMP, the average runtime of recognizing one VOI can be decreased by 55.2% (SR2: 0.29 s versus SR3: 0.13 s).

3.5. Discussion. Experimental results show that the sparse representation based methods had a good performance on the classification of the six kinds of the pulmonary patterns, which were thought to be of great potential for the clinical application by the radiologists. Furthermore, the replacement of the K-SVD and OMP by the K-Means and OMP_1 can save the runtime of the method while keeping the classification accuracy. Therefore, we thought that the SR3 which adopted the K-Means and OMP_1 was efficient in the CAD of the DLDs.

It is thought that the images could be treated as a distribution of a set of representative features, so the sparse representation can extract the important information of examples while removing the irrelevant details, which is advantageous for the classification. Although the textures of the DLD patterns on the HRCT images are complex, the sparse representation approaches are able to produce the descriptors with enough discriminating power. So, the proposed methods achieved good results in the experiments.

However, the performance of the sparse representation based methods on the classification of the GGO and NOD on the HRCT image was relatively worse. The appearance of the GGO on the HRCT image is a hazy increased in the pulmonary attenuation (“whiter” than the normal pulmonary parenchyma, but “blackier” than the soft tissues such as vessels). So, the extent of the GGO would affect the recognition. Figures 5(a) and 5(b) give two examples of the GGO which were misclassified to be NOD and EMP, respectively. Compared with the surrounding normal tissues, the abnormal extent is relatively low in the VOIs. For the NOD, the recognition would be affected by the number of nodular opacities. Figure 5(c) shows an example of

TABLE 4: *P* value of statistical difference between pairs of methods.

	SR1	SR2	SR3	SDF	CSE	BOW
SR1	—	0.57	0.75	<0.0001	<0.0001	<0.0001
SR2	0.57	—	0.31	<0.0001	<0.0001	<0.0001
SR3	0.75	0.31	—	<0.0001	<0.0001	<0.0001
SDF	<0.0001	<0.0001	<0.0001	—	<0.0001	<0.0001
CSE	<0.0001	<0.0001	<0.0001	<0.0001	—	<0.0001
BOW	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	—

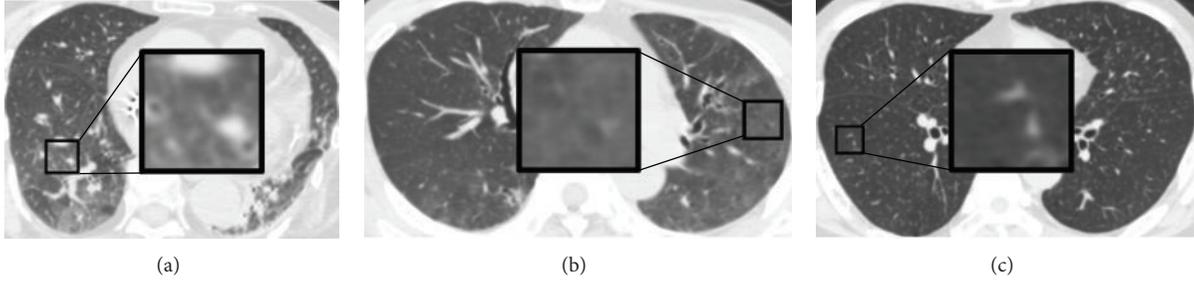


FIGURE 5: Misclassified examples by the proposed methods. (a) An example of the GGO which was misclassified to be the NOD. (b) An example of the GGO which was misclassified to be the EMP. (c) An example of the NOD which was misclassified to be the NOR.

TABLE 5: Runtime of proposed methods with the optimal parameters.

Methods	Time of dictionary learning	Time of recognizing one VOI
SR1	13520 s (<i>K</i> -SVD)	1.27 s (OMP)
SR2	241 s (<i>K</i> -Means)	0.29 s (OMP)
SR3	350 s (<i>K</i> -Means)	0.13 s (OMP ₁)

the NOD which was misclassified to be the NOR. The reason may be the few nodular opacities in the VOI.

For the CAD system, it is also important to reduce the runtime of the method while keeping the classification accuracy. Considering that the dictionary learning and sparse coding spent the most time on the training and testing, respectively, we tried to optimize these two stages. In order to reduce the runtime of the dictionary learning, we used the *K*-Means to train the dictionary, which could be seen as a simple version of the *K*-SVD. Figure 6 compares the SR1 (using the *K*-SVD) and SR2 (using the *K*-Means). It can be found that the two methods had similar classification accuracies. Furthermore, the runtime of the dictionary learning by the *K*-SVD was nearly 50 times as long as the *K*-Means when the same number of atoms was used. It is demonstrated that the replacement of the *K*-SVD by the *K*-Means can considerably decrease the runtime and not affect the classification accuracy.

For the optimization of the sparse coding, although the runtime of recognizing one VOI in our experiment seemed not very long, the CAD system will be used to analyze the whole lungs of patients in the clinical practice, which can be divided into tens of thousands of VOIs. So, a small reduction of the runtime in the experiment (classify

the individual VOIs) is meaningful which can lead to a remarkable decrease in the actual practice (recognize the whole lungs of patients). In order to reduce the runtime of the sparse coding, we applied a simple version of OMP, which selected the desired number of atoms at one time instead of the iterative calculation (OMP₁). Figure 7 compares the SR2 (using the OMP) and SR3 (using the OMP₁) when the same parameters were used. The recognition rates of the SR3 were similar to the SR2 when the sparsity was small (2 and 4). And the SR3 spent shorter runtime than the SR2. It is demonstrated that the application of the OMP₁ with a high sparsity can achieve a good result and reduce the runtime of the method.

We compared the proposed methods with the SDF due to the two reasons. The first one was that the SDF had been successfully applied to classify most kinds of pulmonary patterns, including normal tissues and six kinds of the DLD patterns. The second one was that the features extracted from the images were directly used as the input vectors of the classifier without a “sparse coding” step. Unfortunately, the performance of the SDF was not satisfied in the experiments. We thought that the classification may be affected by detecting the geometrical-based components (nodular, linear, and multilocular component), which is still a difficult problem in the CAD, especially for the images of the severe DLDs.

The CSE was slightly similar to the SR2 and SR3. Firstly, the *K*-Means algorithm was adopted in all three methods. Secondly, the signatures of the VOIs, which were used as the input vectors of the classifier, were generated according to the local features. It could seem as a “coding” step, but not the sparse coding. Therefore, the CSE was used to compare with the proposed methods. The CSE produced the worst result in the experiments. The reason for the bad performance may be that the NN classifier is naive comparing to the SVM.

TABLE 6: Comparison of SR3 and BOW.

	Overall accuracy	Time of dictionary learning	Time of recognizing one VOI
SR3	96.4%	350 s (K-Means)	0.13 s
BOW	85.5%	70 s (K-Means)	0.013 s

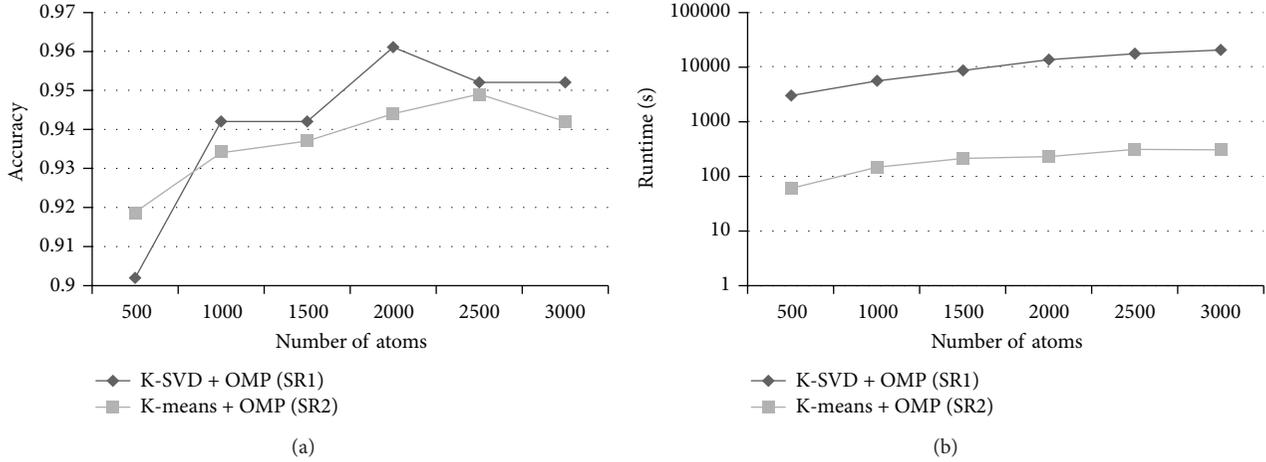


FIGURE 6: The classification accuracy and runtime of the dictionary learning of SR1 (K-SVD+OMP) and SR2 (K-Means+OMP). The patch size was $4 \times 4 \times 4$ and the number of nonzero entries was 8. Replacement of the K-SVD by the K-Means can reduce the runtime of the method while keeping the classification accuracy.

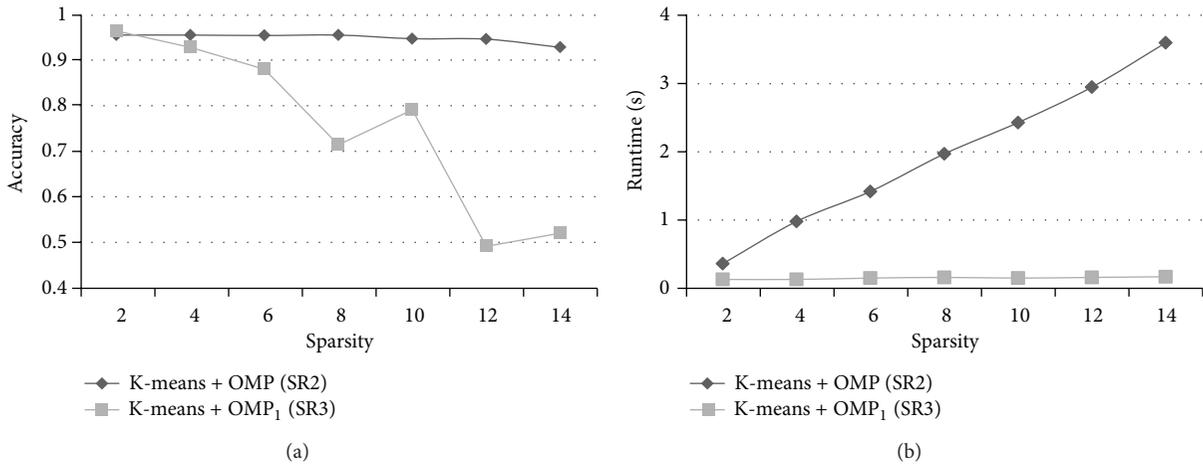


FIGURE 7: The classification accuracy and runtime of recognizing one VOI (except feature extraction) of SR2 (K-Means+OMP) and SR3 (K-Means+OMP₁). The patch size was $3 \times 3 \times 3$ and the number of nonzero atoms was 3000. Replacement of the OMP by the OMP₁ can reduce the runtime of the method and achieve good performance when the value of the sparsity is small.

The bag-of-words is a popular model for the image classification, and the bag-of-words based methods have achieved good results in the previous works. The bag-of-words model could be treated as a special version of the sparse representation, which was implemented with an extremely strict constraint on the sparsity. In the bag-of-words, only one atom was used to approximate the example, and the coefficient of the selected atom was fixed at 1. The work [16] thought that this constraint was too restrictive, so it would produce a large reconstruction error. For the sparse coding strategy, the sparsity constraint was relaxed by allowing a small number of atoms to describe the examples.

Although more time would cost, the sparse coding approach can achieve a fine reconstruction. Therefore, it can reserve more important information of the examples, which was advantageous for the classification. On the other hand, the experimental data adopted in our experiments was different from the previous work [29]. It also would affect the classification of the BOW. Table 6 compares the overall accuracy and runtime of the SR3 and BOW. The BOW spent little time on both the dictionary learning and recognizing. However, the SR3 achieved a significantly better overall accuracy. Figure 8 shows two example images of the NOR which were correctly classified by the SR3 but falsely recognized as the NOD by

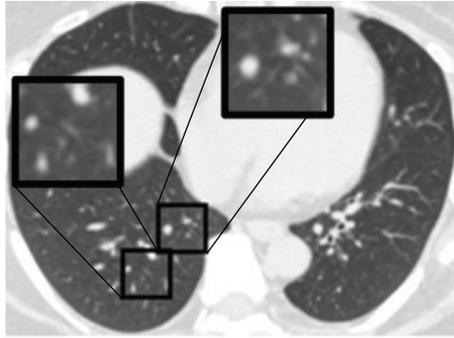


FIGURE 8: Example images of the NOR which were classified to the NOD by the BOW method.

the BOW. The reason of the misclassification may be that the appearance of these two VOIs was similar to the NOD. There were many structures with high CT values (“whiter” than the normal pulmonary parenchyma) such as vessels in the VOIs.

4. Conclusion

In this research, the sparse representation approaches were applied and optimized for the classification of the normal tissues and five kinds of the DLD patterns. By using the K-SVD and OMP, it achieved a satisfied recognition rate but spent too much time in the experiment. So, we tried to replace the K-SVD by the K-Means and substitute the OMP by a simple version of the OMP, which selected a sufficient number of atoms at one time (OMP₁). Experimental results showed that the performances of the sparse representation based methods were significantly better than the baseline methods (SR1: 96.1%, SR2: 95.6%, and SR3: 96.4% versus SDF: 75.8%, CSE: 65.1%, and BOW: 85.5%). Furthermore, when the K-SVD was replaced by the K-Means, the runtime of the dictionary learning was reduced by 98.2% (SR1: 13520 s versus SR2: 241 s). And when the OMP₁ was substituted for the OMP, the average runtime of recognizing one VOI was decreased by 55.2% (SR2: 0.29 s versus SR3: 0.13 s). Therefore, we concluded that the method using the K-Means and OMP₁ (SR3) was efficient for the CAD of the DLDs. We will apply the SR3 in the clinical practice in future research.

Conflict of Interests

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

Acknowledgment

This work was supported by MEXT KAKENHI Grant nos. 21103008 and 26108009.

References

- [1] W. Thurlbeck, M. Roberta, R. N. Muller, and E. Rosenow, *Diffuse Diseases of the Lung: A Team Approach*, Mosby, 1991.
- [2] W. Webb, N. Lüller, and D. P. Naidich, *High Resolution CT of the Lung*, Lippincott Williams & Wilkins, Philadelphia, Pa, USA, 4th edition, 2008.
- [3] P. Grenier, D. Valeyre, P. Cluzel, M. W. Brauner, S. Lenoir, and C. Chastang, “Chronic diffuse interstitial lung disease: diagnostic value of chest radiography and high-resolution CT,” *Radiology*, vol. 179, no. 1, pp. 123–132, 1991.
- [4] M. Zompatori, N. Sverzellati, V. Poletti et al., “High-resolution CT in diagnosis of diffuse infiltrative lung disease,” *Seminars in Ultrasound, CT and MRI*, vol. 26, no. 5, pp. 332–347, 2005.
- [5] Z. A. Aziz, A. U. Wells, D. M. Hansell et al., “HRCT diagnosis of diffuse parenchymal lung disease: inter-observer variation,” *Thorax*, vol. 59, no. 6, pp. 506–511, 2004.
- [6] U. Bağcı, M. Bray, J. Caban, J. Yao, and D. J. Mollura, “Computer-assisted detection of infectious lung diseases: a review,” *Computerized Medical Imaging and Graphics*, vol. 36, no. 1, pp. 72–87, 2012.
- [7] K. Doi, “Computer-aided diagnosis in medical imaging: historical review, current status and future potential,” *Computerized Medical Imaging and Graphics*, vol. 31, no. 4-5, pp. 198–211, 2007.
- [8] M. L. Giger, H.-P. Chan, and J. Boone, “Anniversary paper: History and status of CAD and quantitative image analysis: the role of Medical Physics and AAPM,” *Medical Physics*, vol. 35, no. 12, pp. 5799–5820, 2008.
- [9] S. C. Park, J. Tan, X. Wang et al., “Computer-aided detection of early interstitial lung diseases using low-dose CT images,” *Physics in Medicine and Biology*, vol. 56, no. 4, pp. 1139–1153, 2011.
- [10] J. Wang, F. Li, K. Doi, and Q. Li, “A novel scheme for detection of diffuse lung disease in mdct by use of statistical texture features,” in *Medical Imaging 2009: Computer-Aided Diagnosis*, vol. 7260 of *Proceedings of SPIE*, 726039, Lake Buena Vista, Fla, USA, February 2009.
- [11] Y. Uchiyama, S. Katsuragawa, H. Abe et al., “Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography,” *Medical Physics*, vol. 30, no. 9, pp. 2440–2454, 2003.
- [12] R. Uppaluri, E. A. Hoffman, M. Sonka, P. G. Hartley, G. W. Hunninghake, and G. McLennan, “Computer recognition of regional lung disease patterns,” *American Journal of Respiratory and Critical Care Medicine*, vol. 160, no. 2, pp. 648–654, 1999.
- [13] O. P. Sang, B. S. Joon, N. Kim et al., “Feasibility of automated quantification of regional disease patterns depicted on high-resolution computed tomography in patients with various diffuse lung diseases,” *Korean Journal of Radiology*, vol. 10, no. 5, pp. 455–463, 2009.
- [14] L. Sørensen, S. B. Shaker, and M. De Bruijne, “Quantitative analysis of pulmonary emphysema using local binary patterns,” *IEEE Transactions on Medical Imaging*, vol. 29, no. 2, pp. 559–569, 2010.
- [15] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang, and S. Yan, “Sparse representation for computer vision and pattern recognition,” *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1031–1044, 2010.
- [16] J. Yang, K. Yu, Y. Gong, and T. Huang, “Linear spatial pyramid matching using sparse coding for image classification,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR ’09)*, pp. 1794–1801, June 2009.
- [17] B. Wohlberg, “Noise sensitivity of sparse signal representations: reconstruction error bounds for the inverse problem,” *IEEE*

- Transactions on Signal Processing*, vol. 51, no. 12, pp. 3053–3060, 2003.
- [18] M. Liu, L. Lu, X. Ye, S. Yu, and M. Salganicoff, “Sparse classification for computer aided diagnosis using learned dictionaries,” in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2011*, G. Fichtinger, A. Martel, and T. Peters, Eds., vol. 6893 of *Lecture Notes in Computer Science*, pp. 41–48, Springer, Berlin, Germany, 2011.
- [19] K. T. Vo and A. Sowmya, “Multiscale sparse representation of high-resolution computed tomography (HRCT) lung images for diffuse lung disease classification,” in *Proceedings of the 18th IEEE International Conference on Image Processing (ICIP '11)*, pp. 441–444, Brussels, Belgium, September 2011.
- [20] M. Zhang, X. Zhou, S. Goshima et al., “An application to pulmonary emphysema classification based on model of texton learning by sparse representation,” in *Medical Imaging: Computer-Aided Diagnosis*, vol. 831534 of *Proceedings of SPIE*, 2012.
- [21] M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation,” *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [22] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, “Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition,” in *Proceedings of the 27th Asilomar Conference on Signals, Systems & Computers*, pp. 40–44, November 1993.
- [23] W. Zhao, R. Xu, Y. Hirano, R. Tachibana, and S. Kido, “Classification of diffuse lung diseases patterns by a sparse representation based method on HRCT images,” in *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '13)*, pp. 5457–5460, July 2013.
- [24] J. Zepeda, E. Kijak, and C. Guillemot, “SIFT-based local image description using sparse representations,” in *Proceedings of the IEEE International Workshop on Multimedia Signal Processing (MMSp '09)*, pp. 1–6, October 2009.
- [25] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer, New York, NY, USA, 1st edition, 2010.
- [26] R. Xu, Y. Hirano, R. Tachibana, and S. Kido, “Classification of diffuse lung disease patterns on high-resolution computed tomography by a bag of words approach,” in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2011*, vol. 6893 of *Lecture Notes in Computer Science*, pp. 193–190, Springer, Berlin, Germany, 2011.
- [27] C. C. Chang and C. J. Lin, “Libsvm: a library for support vector machines,” *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, article 27, 2011.
- [28] V. A. Zavaletta, B. J. Bartholmai, and R. A. Robb, “High resolution multi-detector ct -aided tissue analysis and quantification of lung fibrosis,” *Academic Radiology*, vol. 14, no. 7, pp. 772–787, 2007.
- [29] R. Xu, Y. Hirano, R. Tachibana, and S. Kido, “A bag-of-features approach to classify six types of pulmonary textures on high-resolution computed tomography,” *IEICE Transactions on Information and Systems*, vol. E96-D, no. 4, pp. 845–855, 2013.

Research Article

Handling Diagnosis of Schizophrenia by a Hybrid Method

**Luciano Comin Nunes, Plácido Rogério Pinheiro,
Tarcísio Pequeno Cavalcante, and Mirian Caliope Dantas Pinheiro**

*Graduate Program in Applied Informatics, University of Fortaleza (UNIFOR), Avenida Washington Soares, 1321,
Bl J Sl 30, 60811-905 Fortaleza, Brazil*

Correspondence should be addressed to Plácido Rogério Pinheiro; placidrp@uol.com.br

Received 6 June 2014; Revised 26 August 2014; Accepted 8 September 2014

Academic Editor: José M. Jerez

Copyright © 2015 Luciano Comin Nunes 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.

Psychotics disorders, most commonly known as schizophrenia, have incapacitated professionals in different sectors of activities. Those disorders have caused damage in a microlevel to the individual and his/her family and in a macrolevel to the economic and production system of the country. The lack of early and sometimes very late diagnosis has provided reactive measures, when the professional is already showing psychological signs of incapacity to work. This study aims to help the early diagnosis of psychotics' disorders with a hybrid proposal of an expert system that is integrated to structured methodologies in decision support (multicriteria decision analysis: MCDA) and knowledge structured representations into production rules and probabilities (artificial intelligence: AI).

1. Introduction

Psychological disorders have been identified among the main causes of absence from work environments, in view of the influence that those disorders may have on the immune system of people, generating diseases to them. However, some factors hinder a proper diagnosis; for example, a person suffering from psychological disorder offer resistance to accept such a situation, while a healthy person can simulate some symptoms, in an attempt to circumvent a situation to their advantage [1].

In the set of psychological disorders, the psychotic disorders are more incapacitating, characterized by behaviors as those at schizophrenia. A terrible consequence for an individual who is suffering from psychotic disorders is the image of invalidity that the individual presents for the society in which he lives. By more than the, individual strives to erase the negative image formed in respect of his person, the prejudices and the damage are difficult to reverse.

It is usually observed in the workplace as people react with a teammate who suffers from a psychotic disorder. Even if renowned experts are treating him or he/she is

ingesting drugs of last generation, many of his work team observes him/her with suspicion. A mutilation of a limb, a serious infection controlled by medication, and even a cancer eradicated can provide chances for reinstatement of an individual to social life and to the work. However, the stigma of being a carrier of a psychotic disorder usually invalidates the professional who suffers from it. Aiming at the propose of a system for early diagnosis of psychotic disorders, a hybrid model is presented by combining multicriteria methodology for decision support and expert system. Importantly, hybrid models have been used to support decision-making and finding diagnostics for diseases [2–5].

Considering the greater organicity proportioned as well as being a version of the experience of mental health professionals, it was decided to use the model of the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders, DSM-IV*, recommended by the American Psychiatric Association. Also, in the study relations of forces between symptoms and causes of psychotics disorders were established, in order to form a classification, aiming at formatting rules to be used in knowledge base of expert system to assist in establishing the desired diagnostic process.

2. A Methodology for Classification of Diseases

2.1. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. As previously mentioned, this study was based on the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*; due to this manual, much is already known in the area of mental health, in several parts of the world, and thus, it itself constitutes a reference book for many. The *DSM-IV* describes typical patterns of behavior, thought, and emotion, allowing categorizing the diagnosis. The categories are prototypes. A patient who is suffering events related to the prototype is said to be having the disorder relating to that prototype. Qualifiers are sometimes used, to moderate or severe form of the disorder. Each category of disorders has a numeric code that is used by providers of health services for administrative purposes. The structure of the *DSM-IV* is axial; that is, it shows the diagnostic types of mental disorders arranged in five axes, as listed below:

- (i) Axis I: clinical disorders, mental disorders, developmental disorders, and learning disorders are included here;
- (ii) Axis II: personality disorders and mental retardation;
- (iii) Axis III: medical conditions or acute physical disorders;
- (iv) Axis IV: psychosocial or environmental factors that are correlated with the disturbances;
- (v) Axis V: global assessment of function (global assessment of functioning) or global assessment scale for children (children's global assessment scale) for youth under the age of 18 (on a scale of 0 to 100).

2.2. A Classification of Diseases Applied to Psychological Contextualization. Based on the premise that it is possible to identify causes and symptoms to psychotic disorders, we seek early diagnosis of these disorders. To this end, in this study related "control events" that correlate with symptoms and causes that show mentioned psychotic disorders were presented [5]. For each of these "control events" that appear in psychotic disorders, a confidence factor that indicates the degree of influence on the outcome of the indicative diagnosis of these disorders is assigned. The tree shown in Figure 1, formatted using the HiView software, summarizes the major psychological disorders that affect people. Specified disorders represent some classes presented in the *DSM-IV*, among which the category of psychotic disorders stands out in Figure 2, subject of this paper.

2.3. Psychotic Disorders. More commonly known as schizophrenia, term of Greek origin, that means "split mind," for the individual who suffers this disease there is dissociation between his/her thinking and the reality in which he/she lives. Thus, the individual with this psychological disorder acts as if he lived in a world that exists only in his mind, disconnected from reality. Psychotic disorders are considered the most severe and disabling form of mental disorders. These disorders usually emerge in adolescence or early

adulthood and rarely after 45 years of age [1]. For this type of psychological disorder, in this paper itself, *DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition)* is used, which contemplates two groups: positive symptoms and negative symptoms. Many experts and scholars have proposed this system of categorization into two groups of schizophrenia on the subject, among which worth mentioning [7–11]. The *DSM-V* version, released by the American Psychiatric Association (APA) in May 2013, which presents a different categorization for this disease, will be used in future works.

2.3.1. Positive Symptoms Psychotic Disorders. This disorder type's main characteristics are as follows [1]:

- (i) Frequent manifestations of delusions or bizarre thoughts, that is, outside the sociocultural context of the individual, and as much as it itself argues against using including corroborative veracity of the delirium's elements, the individual who suffers from this disorder does not accept the explanations;
- (ii) hallucinations or sensory perceptions without the presence of a being or object relative to the manifested sense, such as voices and sounds directed only to the individual who is suffering from the disorder;
- (iii) laughing and crying easily, without explanation;
- (iv) manias to be suffering persecution;
- (v) disorganization of thought, barring participation in dialogues or expressions of related ideas.

In Figure 3, the main control events, which correlate with psychotic disorders positive symptoms, are provided. The numerical values in Figure 3 represent the importance of each event in the occurrence of this disorder; that is, the higher the value is, the more influence this event has in establishing the diagnosis for this disorder. These values were assigned based on a scale of importance of each event to the occurrence of each disorder. That scale, detailed in Table 1, was created from reports made by experts in the psychiatry and psychology areas, about the degree of importance of the events of control in psychological disorders, object of the present study.

2.3.2. Psychotic Disorders Negative Symptoms. People suffering from this disorder exhibit behavior with the following characteristics [1]:

- (i) the search for the life in closure, total isolation, and avoiding social interaction;
- (ii) affective flattening, that is, gradual reduction of affective relationships;
- (iii) abulia, or loss of willingness to engage in any activities; reaching the extreme situation of staying motionless, the individual eyes are fixed at anything;
- (iv) poor thoughts of ideas manifested by monosyllabic vocabulary.

In Figure 4, it is possible to observe control events associated with psychotic disorders negative symptoms. In the scale,

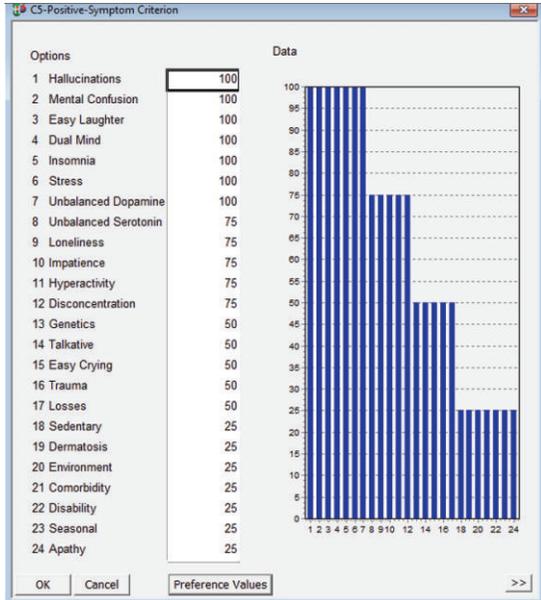


FIGURE 3: Events correlated psychotic disorders positive symptoms. Source: formatted by the author from HiView software.

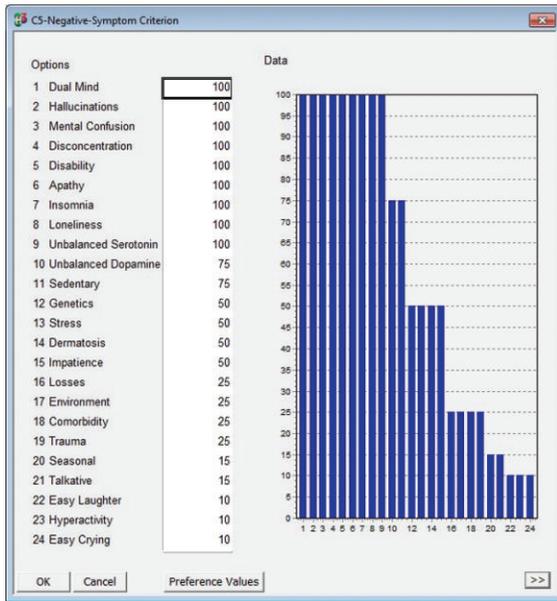


FIGURE 4: Events correlated psychotic disorders negative symptoms. Source: formatted by the author from HiView software.



FIGURE 5: Simplified view of the system architecture, based on a hybrid model of information technologies. Source: formatted by the author.

(iii) Recommendation: in this phase sensitivity analyses and robustness are made to verify those changes, in the parameters of the evaluation model, affecting the final result. It is a key phase that helps to generate knowledge about the problem and thus increases confidence in the results obtained from the decision-maker.

3.2. *Decision Support in the Area of Mental Health.* Initially, it would be worth noting that, searching the literature of decision support in health, some studies reported hybrid models used to aid decision-making and finding diagnostics for diseases; among these studies highlight the following: [3, 5, 13, 14].

According to [14], decision analysis is intended to enhance the quality of decisions and communication among physicians, patients, and other healthcare professionals. Unquestionably, many decisions in healthcare are complex, with multiple factors affecting the decision of a particular outcome. It has been heard or has had the experience of some of the deficiencies caused by a decision taken improperly. People of all ages and all states of health have been affected by this type of practice of establishing a diagnosis [15].

It is clear therefore that the decision-making in seeking diagnosis for a health disorder should not be treated informally, practice somewhat common these days, most likely due to the short time spent in search processes diagnosis, especially in the public health service. This line may lead to risks of compromising the quality of diagnosis as well as the health of sick individuals and, consequently, the health system as a whole, generating waste resources and increasing spending. In this trail of problems, it can also raise a culture of self-diagnosis, where lay people self-medicate, because the patients observe that health professionals do not give proper attention to their problems and act by trial error/hit. Therefore, individuals are induced to mimic the health area professionals, through the use of the practice of the error/hit, complicating the search by the diagnostic process.

This work aims, among other things, to show the positive impact of the use of information technology in support-establishing clinical diagnoses of psychotic disorders. Szolovits [16] describes various approaches that try to present a great way of using information technology in healthcare, particularly through systems supporting diagnosis of mental disorders. The motivations for the development of such studies, as also described by Szolovits [17], are readily visible because excellence in health care is of paramount importance to society.

3.3. *Problems of Decision Support and Methodologies Multicriteria.* To establish a diagnosis of psychology and psychiatry, studies are needed, experiments and observations which are aware of the symptoms that lead to the cause of the disorder evidenced. A major difficulty in reaching a correct diagnosis is the complexity of factors evidenced, for, besides the large amount of information, the expert needs to take into account cultural, biological, psychosocial issues, quality of information, signs, and symptoms common to many

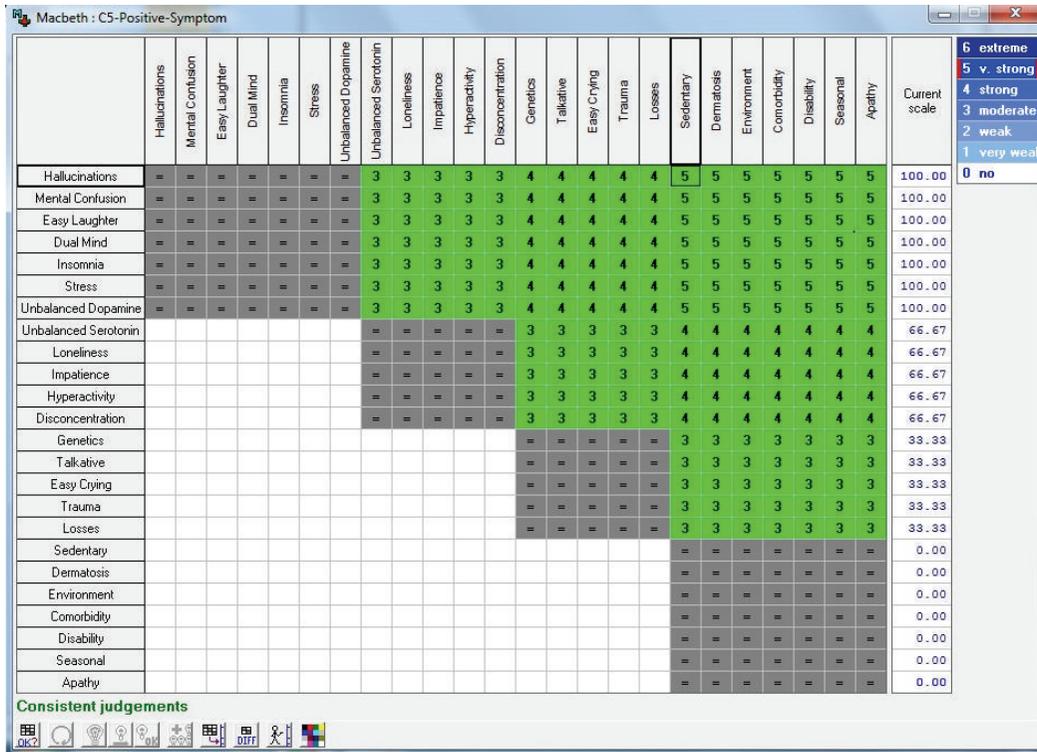


FIGURE 6: Matrix of judgment and values difference in attractiveness between the “control events” of the psychotic disorders positive symptoms. Source: formatted by the author from HiView software.

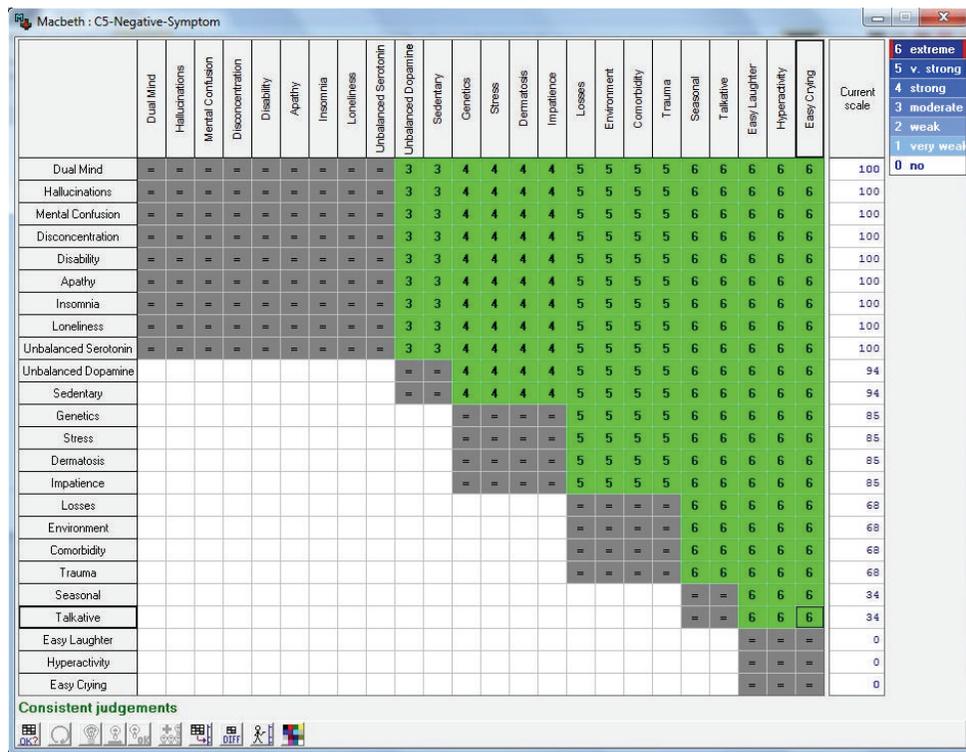


FIGURE 7: Matrix of judgment and values difference in attractiveness between the “control events” of the psychotic disorders negative symptoms. Source: formatted by the author from HiView software.

diseases. This complicates the decision process, turning it into something complex, and it also takes into account unstructured information. The methodology to aid multicriteria decision-making has much to add to the diagnostic process in psychology and psychiatry. In order to put at the disposal of the decision-maker, methods and tools enable them to design the control events in question as well as prioritizing these events for proper classification.

3.4. Multicriteria Methodology

3.4.1. Basics Multicriteria Decision Aiding. Problems involving multiple criteria have several agents and concepts, reasoning its definition which is merely didactic [18]. Some of them are discriminated below. Decision-maker agent has the power and responsibility to ratify the decision, assuming the consequences of this act, whether positive or negative. The decision-maker can be an individual or a group of people establishing the boundaries of the problem, specifying the objectives to be achieved and issues opinions. Not all decision-makers have the power of choice. So it is important to distinguish the degree of influence on the decision-makers' process [6, 19].

Analyst is an agent who interprets and quantifies the opinions of decision-makers, conducts the structuring of the problem, formulates a mathematical model, and presents the results to the decision. The analyst has a duty to act in continuous dialogue and interact with decision-makers, in a constant learning process.

Model is a simplified representation or interpretation of reality through rules and mathematical operations that allow transforming preferences and opinions of decision-makers in quantitative result.

Alternative is known as potential action. The subject of the decision or that is directed to support the decision [20]. It is identified early in the decision process or during this and could become a solution to the problem under study [6].

Criterion represents the preferences of the decision-maker under certain viewpoint of him/her. It is presented as a function g , defined on a set A , which assigns value's sort of set A [21].

According to [21], when a problem has multiple criteria, they are defined as $g_1, g_2, \dots, g_j, \dots, g_n$. The evaluation of an action "a" according to the criterion "j" is represented by $g_j(a)$. The representation of different viewpoints (aspects, factors, and characteristics), with the help of a family $F = \{g_1, \dots, g_j, \dots, g_n\}$ of criteria, constitutes one of the most delicate parts in formulation of decision problems.

Relation of dominance is a relationship that occurs when two elements "a" and "b" belong to the set A ; "a" dominates "b" (aDb) if and only if $j = 1, 2, \dots, n$ wherein at least one of the inequalities is preferably narrow. It may be noted that the dominance relation of "a" and "b" is characterized by being a strict partial order, being an asymmetric relation and transitive. If "a" dominates "b," "a" is greater than "b" in all criteria of the problem [21].

Efficient Action—The action (or alternative) "a" is considered efficient if and only if no other action set A , dominates. The efficient set of actions A , where A maybe the dominance relation is empty and is generally regarded as a set that contains the interesting actions, to be analyzed in greater depth, even though they lack good reasons to disregard the inefficient [21].

Decision matrix termed as the evaluation matrix, where each row explicitly measures evaluations of alternative i with respect to n criteria considered. Each column in turn expresses evaluations of the measurements of the m alternatives with respect to criterion j . Assuming that a_{ij} represents the evaluation of alternative (or action) A_i , then belonging to the set of potential actions A , $[a_{ij}]$, according to the criterion g_j , one could construct a similar matrix to that shown in Table 2 [6].

3.4.2. Multicriteria Decision Aiding. The method MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique) is an approach to multicriteria decision support. The research that initiated the method was performed by Antonio Carlos Bana e Costa and JC Vansnick in the early 90s [12]. This methodology has emerged in response to the question of how to build an interval scale of preferences from a set of options without forcing decision-makers to produce their preferences numerically directly. This approach allows you to assign scores to each alternative through a paired comparison. Given two alternatives, the decision-maker must express the most attractive in the case have greater confidence, and what degree this attractiveness on a scale that has semantic correspondence with an ordinal scale. The program itself makes the analysis of cardinal consistency (transitivity) and semantics (relations between differences), suggesting, in the event of inconsistency, how to resolve it. It is still available to the decision-maker to graphically adjust the value of the scores assigned, within permitted ranges. According to [12], the construction of cardinal value scale is only after this adjustment, with the introduction of specialist expertise, which is characterized. The difference of attractiveness is very important in this methodology.

According to [22], the MACBETH method, when the decision-maker is asked to value judgments about potential actions (alternatives) in a given situation are realized, it will do in terms of attraction he feels for this alternative. This task is defined on the construction of the role, criterion v_j , such that

- (i) for $a, b \in A$, $v(a) > v(b)$, if and only if, for the evaluator, a is more attractive (locally) than b (aPb);
- (ii) any positive difference, $v(a) > v(b)$, represents the numerical difference in value between "a" and "b," with "aPb" being always in terms of a fundamental point of view j (FPV_j) or criterion "j";
- (iii) Then, for each "a," "b," "c," "d" $\in A$, being judged a more attractive than "b," and "c" more attractive than "d", it is clear that $v(a) - v(b) > v(c) - v(d)$, if and only if the difference of attractiveness between a and b is

TABLE 2: Decision matrix.

Criteria →	g_1	g_2	...	g_j	...	g_n
Limit →	q_1, P_1	q_2, P_2	...	q_j, P_j	...	q_n, P_n
Alternatives ↓						
A_1	a_{11}	a_{12}	...	a_{1j}	...	a_{1n}
A_2	a_{21}	a_{22}	...	a_{2j}	...	a_{2n}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
A_i	a_{i1}	a_{i2}	...	a_{ij}	...	a_{in}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
A_m	a_{m1}	a_{m2}	...	a_{mj}	...	a_{mn}

Source: [6].

greater than the difference in attractiveness between “c” and “d”.

For the MACBETH method is important, the following is reasoning. Given the impacts ij (a) and ij (b) of two potential actions “a” and “b,” according to a fundamental point of view FPV_j , being judged a more attractive than “b,” the difference in attractiveness between “a” and “b” is judged as “null,” “very weak,” “weak,” “moderate,” “strong,” “very strong,” or “extreme.” It introduced a scale formed by different semantic categories in attractiveness; the size is not necessarily equal, to facilitate the interaction between the decision-maker and analyst. The semantic categories, C_k , $k = 1, \dots, 6$, are represented as follows [12]:

- (i) C_1 is very weak difference of attractiveness (or between zero and weak) → $C_1 = [s_1, s_2]$ and $s_1 = 0$;
- (ii) C_2 is weak difference of attractiveness → $C_2 =]s_2, s_3]$;
- (iii) C_3 is moderate difference of attractiveness (or between weak and strong) → $C_3 =]s_3, s_4]$;
- (iv) C_4 is strong difference of attractiveness → $C_4 =]s_4, s_5]$;
- (v) C_5 is strong difference of attractiveness (or between strong and extreme) → $C_5 =]s_5, s_6]$;
- (vi) C_6 is extreme difference of attractiveness → $C_6 =]s_6, +[$.

To facilitate the expression of absolute judgments of difference in attractiveness between the pairs of alternatives, it is useful to construct the arrays of value judgments [23].

Considering one software to support multicriteria decision, the M-MACBETH allows the structuring values of trees, construction of the rating criteria, the development of functions values, the weight of criteria, and extensive sensitivity analyses and robustness on value intrinsic and relative options.

3.4.3. Application of MACBETH Method and Software HiView. In this paper the following concepts, techniques, methods, and tools have been applied. Method MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique) was applied in supporting decision-making. It has

been chosen because it is a method that has been specializing in the use of hierarchy of importance of components of multicriteria events, printing speed in decision-making, especially in the area of human health. The MACBETH method has been used for defining the existing attractiveness between events of control psychological disorders, to simplify the professional judgment of the decision-maker since the entire set of alternatives does not need to be evaluated simultaneously. It is notable, however, that the difficulty for a decision-maker to remain consistent is, mainly, when the number of alternatives and criteria increases. To circumvent this problem, the method makes analysis cardinal consistency and semantics and also suggests, when necessary, the contour shape.

To facilitate the handling of the concepts of the MACBETH method, there are two software programs to implement and run the MACBETH method: the HiView software and the M-MACBETH software. The HiView software was chosen for the generation of arrays of judgment of the problem, to conduct several sensitivity analyses and robustness of the results of the model application, offering numerous graphical representations that facilitate the preparation of a report justifying the recommendations developed. Built from the MACBETH approach, the HiView tool is crucial for evaluation of models based on this method. Thus, the HiView software executes the functionalities of the M-MACBETH software, with the objective of verifying the consistency of the information and the potential of the methodology employed. Analyze the trend of local and global results of actions when it does vary replacement rates [24, 25]. The HiView provides decision-makers with professional confirmation of their judgments or even allows some values that are not in line with their expectations, validating the data model and consolidating the credibility of this being changed.

3.4.4. Application of the Control Events into the Methodology to Support Decision-Making. After undergoing methodology to support decision-making (MCDA—multicriteria decision aid), implemented in this work through HiView software, which also runs the MACBETH method, the “control events” retro mentioned, were adjusted so as to obtain the matrix of constant value judgment of Figures 6 and 7, which allows the visualization of degrees of attractiveness between events as well as the “current scale” of “confidence factors,” besides indicating if the results are coherent states “consistent judgments.” As for “confidence factors,” Figure 8 gives an idea of how the decision-maker can adjust these degrees of attractiveness between “control events,” pointing a ruler at the boundaries between the various degrees of attractiveness of these events. The new levels of attractiveness indicated on the ruler are reflected in the current range of the matrix as well as adjustments in the confidence of control events factors, as can be seen in Figure 8.

In the present work, the export of the control events and degrees of confidence will be made indirectly and through the collection of answers to questions correlated to each control event. That is, the control events analyzed using the MACBETH methodology, implemented in the HiView

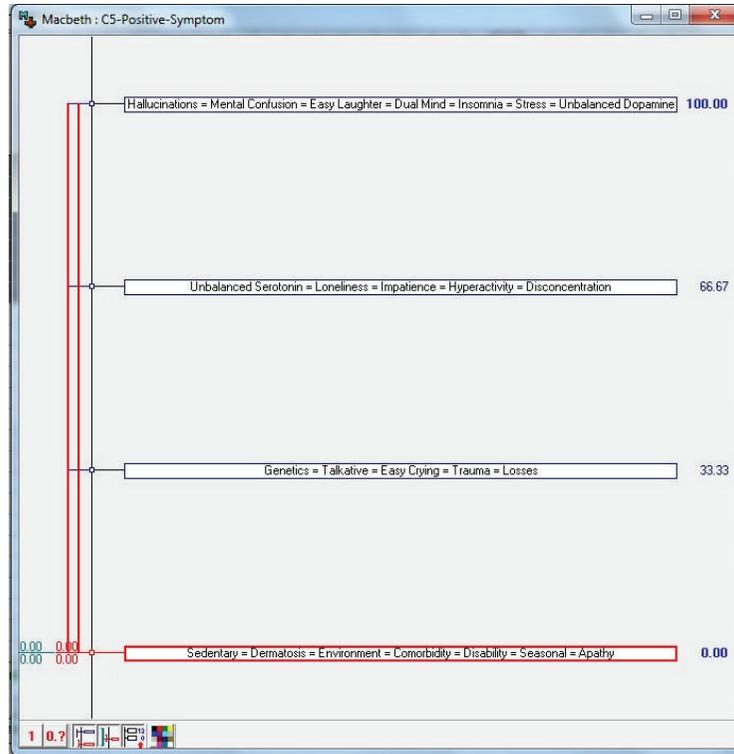


FIGURE 8: Control events with a new degree of confidence within allowed ranges. Source: formatted by the author from HiView software.

software, are associated with the variables treated by expert system, which can be of two quantities: common variables and objective variables. The common variables correspond to the symptoms and causes of the disorders studied. Already, the objective variables correspond to the final diagnosis to be found for the disorder.

Furthermore, as stated in the analysis of disorders seen in this study, each control event has a percentage of influence in relation to the disorder and in relation to the other control events. This percentage, derived from the scale discriminated in Table 1, will compose the degree of confidence in expert system, which will use the control event as an important part to the diagnosis. Thus, for each set of answers to the questions that appear in the user interface, the expert system links this set of responses to the degree of confidence of the control events in order to point the diagnosis. After constructing the array of value judgment, it is possible to do in the HiView software a sensitivity analysis as presented in Figure 8. This analysis allows changing the values of the degrees of confidence of the events of controls.

4. Specialist System Applied to the Diagnosis of Psychological Disorders

4.1. Specialist System: Contextualization. Expert systems are associated with the term artificial intelligence (AI). This expression was first used at a summer conference at Dartmouth College, USA, when researchers John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon

met in order to conduct a study on this subject, which had been baptized with the expression that generated warm and controversial debates during the aforementioned conference and beyond it. Mentioned expression stayed, however, without a formal definition because, first of all, it was imperative to define, too, and formally, what intelligence would be. There are two main strands of research for the intelligent building systems: connectionist and symbolic. The first aims at the modeling of human intelligence by simulating the components of the brain, that is, their neurons and their interconnections. This proposal was first formalized in 1943, when the neuropsychologist McCulloch and, the logical, Pitts proposed a first logical mathematical model for a neuron. In turn, the second part followed the logic tradition and took McCarthy Newell and his principal investigators. Thus, knowledge-based systems, or expert systems, are built mainly with rules that reproduce the knowledge of experts in the fields of human knowledge and are used to solve particular problems in specific domains. The area of human health has been one of the most allotted portions by expert systems, due to being possessed by classic problems that need systems with these potentialities. It is important to note that there are characteristics that indicate whether a given problem should be solved by this information technology or not. To facilitate the process of analyzing a given problem, certain conditions must be met so that they can add value in identifying and choosing which expert systems technology as follows:

- (i) Expert systems that require professional experts in knowledge related to the problem, considering that

this knowledge will form the knowledge base that will be used in solving problems;

- (ii) activities to be performed that require the participation of groups of experts who, when alone, do not have sufficient knowledge to perform them;
- (iii) activities that require knowledge of the details that, if overlooked, cause degradation of performance;
- (iv) activities that show large differences in performance between experts from the group study the problem;
- (v) number of experts available to solve the problem which is insufficient.

The criteria referenced in the previous paragraph provide important respects, for example, a significant increase in the productivity of a business decision maker, in performing specialized tasks when assisted by an intelligent system. It is noteworthy that the portability of these expert systems, being capable of developing and being used in microcomputers, is currently deciding factor. This characteristic causes these systems to become accessible and affordable. In general, systems with automated reasoning can be used banks incorporating existing data in the organization, or being incorporated into the set of tools available in the databases. Many work areas can be helped by structured expert systems solutions as these are efficient applications for information management. Providing tools to support decision-making in this case goes further than providing graphs and tables to the user, provide them a north in the identifying of their needs, simulating scenarios and allowing greater accuracy and reliability in the solutions of their problems. Expert systems are, therefore, computer programs that provide solutions to certain problems, in the same way that human experts offer under the same conditions. The most common architecture of expert systems involves production rules structuring simply a set of conditions in the IF... THEN... style, with the possibility of inclusion of logical connectives relating the attributes within the scope of knowledge and the use of probabilities.

4.2. The Expert SINTA. The software Expert SINTA was created by a group of scholars at the Federal University of Ceará (UFC), and the State University of Ceará (UECE) called Group SINTA (Sistemas INTeligentes Aplicados or Applied Intelligent Systems) [26, 27]. It was developed using Borland Delphi technology. That is a computational tool that uses artificial intelligence techniques for automatic generation of expert systems. It uses a knowledge representation model based on production rules and probabilities, with the main objective of simplifying the construction work of expert systems through the use of a shared inference machine. The automatic construction of screens and menus, treatment of probabilistic rules production, and use of sensitive explanations to the context of the knowledge base modeled. An expert system based on this type of model is very useful in classification problems. The user responds to a sequence of menus, and the system provides answers that fit the framework identified by the user. Some of the main features of the Expert SINTA are use of backward chaining; use of confidence factors; tools for debugging; opportunity to include

online help for each knowledge base. The expert systems are generated in the Expert SINTA using the architecture described in Figure 9.

The Expert SINTA aims to simplify the steps of creating a full expert system. For this, a basic inference engine is already offered, based on backward chaining. The backward chaining excels in problems where there are a large number of conclusions that can be achieved, but the number of ways in which they can be reached is not great, in problems where it cannot meet an acceptable number of facts before starting to search for answers. The backward chaining is also more intuitive for the developer, because it is based on recursion, an elegant and efficient way of programming, where the logic programming itself is directed. At no time, however, it ceases to recognize that the forward chaining has advantages in certain occasions. Moreover, the calculations of values of differences in degree of attractiveness and judgment needed to build the array of values generated by the MACBETH method, as shown in Figures 6 and 7. In Figures 10 to 19 the steps to building an expert system, with aim at assisting in the final diagnosis are presented. In Figure 10 there is an illustration of how the user in the Expert SINTA should proceed to set the precedence of the logical operators that will be used by the inference machine, that is, if $(A \text{ and } B)$ or C/A and $(B \text{ or } C)$.

The choice of the minimum value for the confidence factor is shown in Figure 11. Besides this, there is the possibility of placing a password for the database that will be created. This option is most commonly used as the basis of expert knowledge is confidential.

Figure 12 presents information about the knowledge base to be built. This window will appear after the user starts using the expert system.

On the other hand, Figure 13 shows how to define the variables in the Expert SINTA. These variables will feed the expert system knowledge base and, in the specific case, for each control event detailed variable will be created.

It is possible to see in Figure 14 the definition of objective variables, that is, those which point the final diagnosis in the expert system.

Moreover, Figure 15 shows how the interface is defined in the expert system for collecting events with respective degrees of confidence. For each variable, an issue that aims to collect symptoms of the disorder as well as a degree of confidence about this symptom felt by the patient is created.

At the time of creation of each rule, the Expert SINTA allows this rule to be inserted for desired preference. Moreover, Figure 16 shows the list of logical rules constructed from the rule editor of the Expert SINTA.

The editor of rules, which constitute the logical reasoning of the expert system, can be seen in Figure 17. Mentioned rules point the final diagnosis to the user of system. Note that the rules follow the following structure: IF... THEN... (SE... ENTÃO...).

During the execution of the expert system developed with Expert SINTA, the user can interact through graphical interfaces as shown in Figure 18. The interaction through

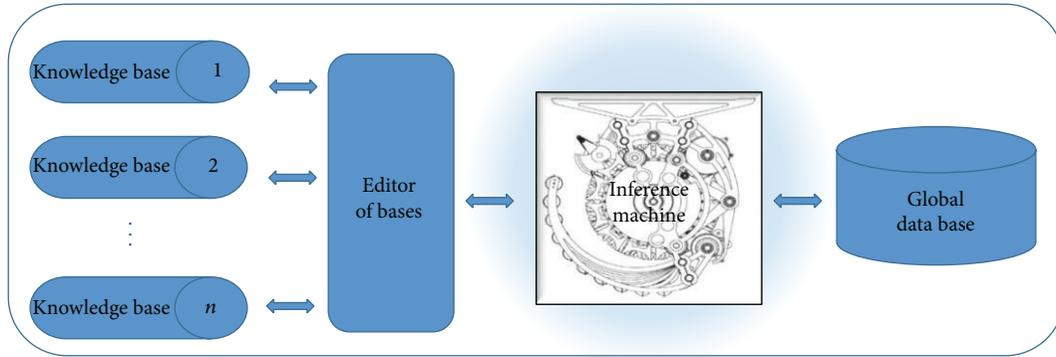


FIGURE 9: Simplified architecture of Expert SINTA. Source: adapted by the author from information of the LIA-UFC. Knowledge base comprises representative information of facts and rules that the expert uses. Editor of bases is the means by which the shell allows the implementation of desired knowledge bases. Inference machine is the part of the expert system responsible for logical deductions about the knowledge base. Global database consists of the evidence presented by the expert system user during a consultation.

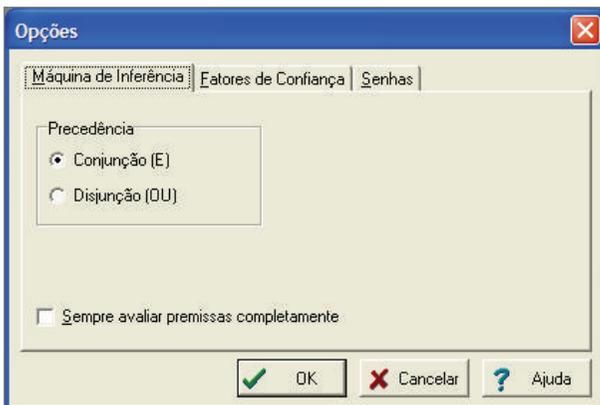


FIGURE 10: Precedence of logical operators. Source: formatted by the author from software Expert SINTA.

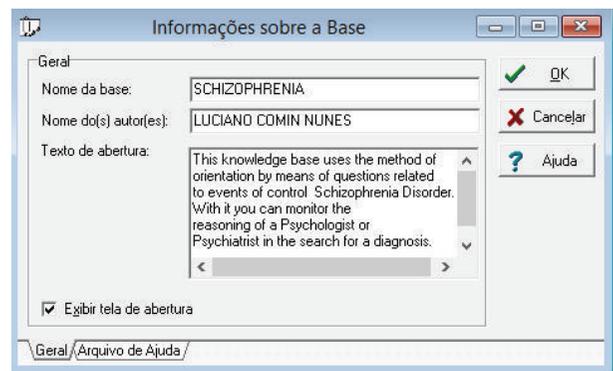


FIGURE 12: General information about the knowledge base of the expert system. Source: formatted by the author from software Expert SINTA.

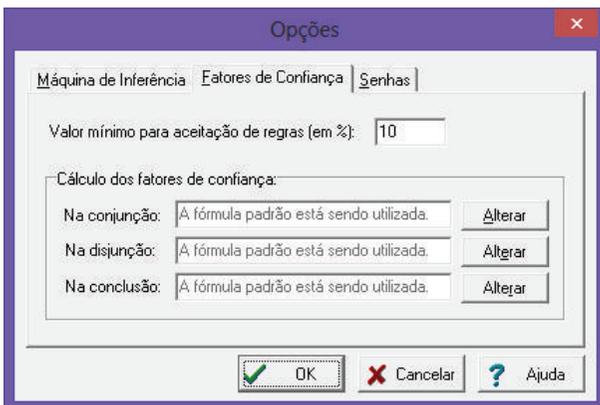


FIGURE 11: Setting minimum and indicative password. Source: formatted by the author from software Expert SINTA.

these interfaces enables the collection of values, which will feed the variables and factors of trust used in the system specialist.

Figure 19 shows the diagnosis reached by the expert system after inferring the information supplied to him by the fed interfaces as well as preexisting knowledge bases in their information.

The Expert SINTA provides for each expert system the path of logical reasoning trod of the execution of the respective expert system. Figure 20 provides an overview of this track, which helps in analyzing the results after obtaining a diagnosis.

5. Conclusion and Future Works

Many smart techniques were incorporated into the decision-making process. These methods are based on technologies that use concepts of artificial intelligence (AI), such as systems expert's genetic algorithms [26, 27], neural networks [28], intelligent agents, reasoning based on case studies [2, 29], and fuzzy logic [30].

Despite these technological advances, much more needs to be done to automate decision-making, especially when it involved a multicriteria analysis.

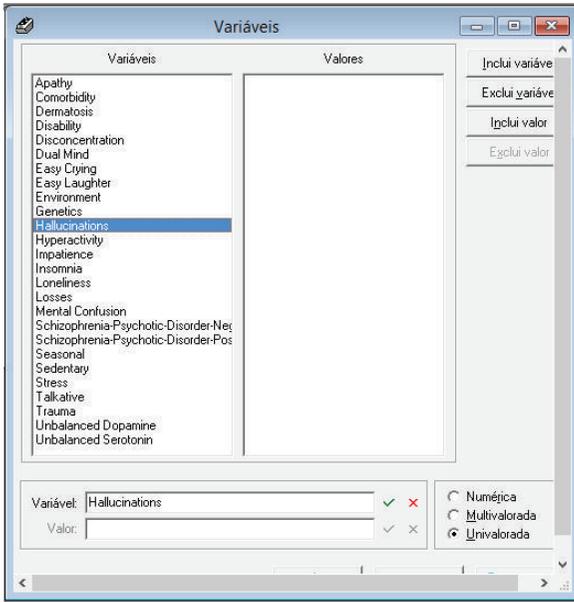


FIGURE 13: Definition of variables in Expert SINTA. Source: formatted by the author from software Expert SINTA.

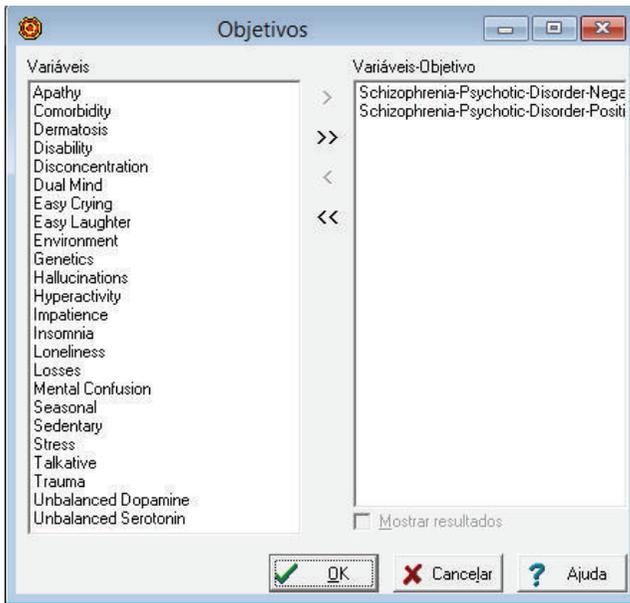


FIGURE 14: Definition of variables goal. Source: formatted by the author from software Expert SINTA.

In the present study, we observed that the information generated in the multicriteria analysis methodology was used in the algorithm in order to transform them into variables with respective degrees of confidence, which were processed by the expert system inference engine through use of IF... THEN... rules, with the aim of using this information, pointing your diagnosis, incorporating it to their knowledge base. The automation of this process of transition between multicriteria methodology and expert system is a challenge

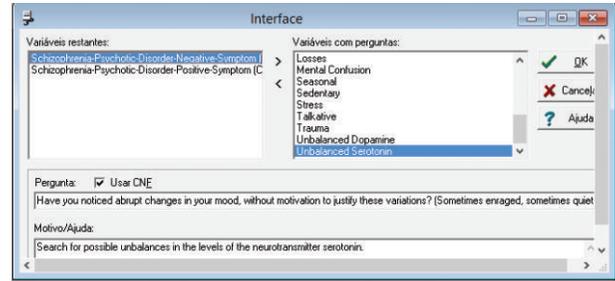


FIGURE 15: Creating user interfaces. Source: formatted by the author from software Expert SINTA.

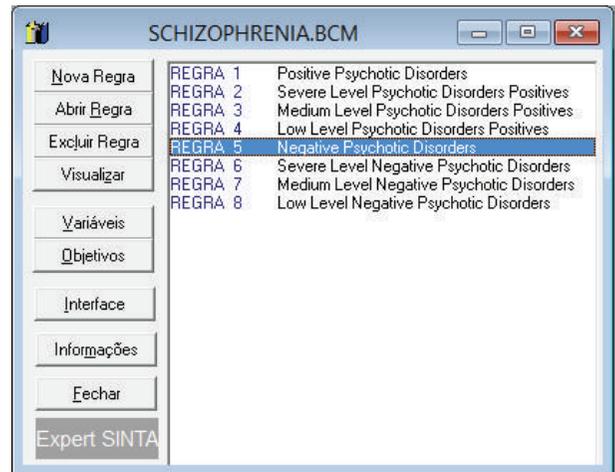


FIGURE 16: Creating logical rules. Source: formatted by the author from software Expert SINTA.

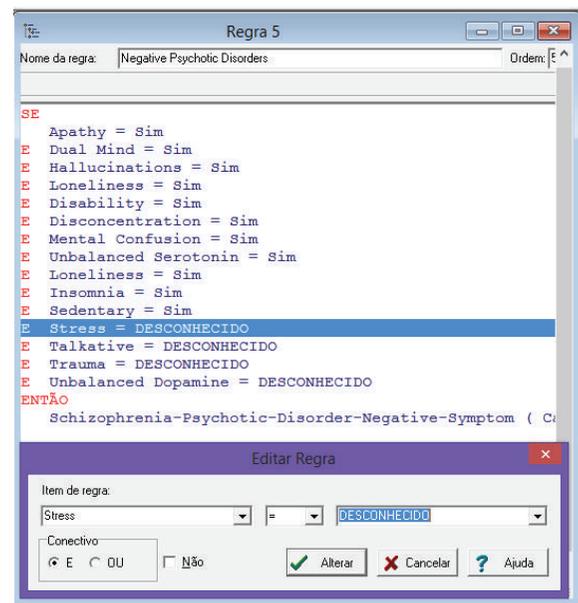


FIGURE 17: Definition of logical rules. Source: formatted by the author from software Expert SINTA.

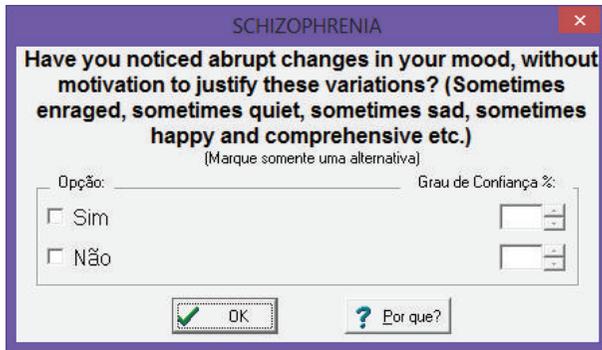


FIGURE 18: Input data through the user interface. Source: formatted by the author from software Expert SINTA.

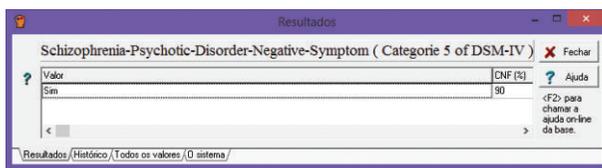


FIGURE 19: Possibility of diagnosis. Source: formatted by the author from software Expert SINTA.

that remains to be solved by the construction of a bridge that enables this integration. Unfortunately, such a task was not possible within this study, given the limited resources of time and availability of technical staff. Thus, the test performed in this study tried to show a hybrid model, using a connection manually. Therefore, we tried to present the feasibility of integration between the mentioned technologies: multicriteria methodology to support decision-making and expert system generated by Expert SINTA. With this proposal it is expected to have contributed to the quality of automated diagnosis.

Finally, it was realized that this field of work is still largely unexplored, fitting entities related to research, encouraging greatly having more concrete results, emphasizing compensatory be returns on that investment, because organizations have required based expert systems solutions.

It is suggested, therefore, in order to improve the model proposed in this paper, to take the following measures:

- (i) Use of some other multicriteria methodologies, such as influence diagram and Bayesian networks;
- (ii) Improved interface with the user, the Expert SINTA including among its features the export and import of data files containing control events and degrees of confidence;
- (iii) Formatting and automation of generic questionnaires covering control events and degrees of confidence.

On the other hand, given that the psychiatry proves to be one of the specialties of medicine that formalized consciously and comprehensively the diagnostic process for the disorders of his study and domain and by virtue, the interest of the current *DSM-V* formulators, in categories and dimensional classification of mental disorders be compatible with neuroscientific



FIGURE 20: Trail to assist in analysis of outcomes after diagnosis. Source: formatted by the author from software Expert SINTA.

and genetic foundations that itself want to consolidate, the authors plan in future work, include in the proposal hybrid information technology, beyond recommendations contained in the *DSM-V*, quantitative and dimensional model with image processing and scales of values, as a way to enhance the diagnostic process currently based on categorization of types of schizophrenia with events of control based in symptoms, causes and historical of this disease.

Conflict of Interests

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

Acknowledgment

The second author is thankful to National Council of Technological and Scientific Development (CNPq) via Grant no. 475239/2012-1.

References

- [1] K. Huffman, M. Vernoy, and J. Vernoy, *Psychology in Action*, chapter 14, John Wiley & Sons, New York, NY, USA, 2000.
- [2] A. K. A. de Castro, P. R. Pinheiro, M. C. D. Pinheiro, and I. Tamanini, "Towards the applied hybrid model in decision making: a neuropsychological diagnosis of Alzheimer's disease study case," *International Journal of Computational Intelligence Systems*, vol. 4, no. 1, pp. 89–99, 2011.
- [3] A. C. Menezes, P. R. Pinheiro, M. C. D. Pinheiro, and T. P. Cavalcante, "Towards the applied hybrid model in decision making: support the early diagnosis of type 2 diabetes," *Lecture Notes in Computer Science*, vol. 7473, pp. 648–655, 2012.
- [4] L. C. Nunes, P. R. Pinheiro, T. C. Pequeno, and M. C. D. Pinheiro, "Support tool in the diagnosis of major depressive disorder," *Communications in Computer and Information Science*, vol. 112, no. 2, pp. 136–145, 2010.
- [5] L. C. Nunes, P. R. Pinheiro, T. C. Pequeno, and M. C. D. Pinheiro, "Toward an application to psychological disorders diagnosis," *Advances in Experimental Medicine and Biology*, vol. 696, pp. 573–580, 2011.

- [6] L. F. A. M. Gomes, C. F. S. Gomes, and A. T. Almeida, *Tomada de Decisão Gerencial: Enfoque Multicritério*, Atlas, São Paulo, Brazil, 2nd edition, 2006.
- [7] R. Toomey, S. V. Faraone, J. C. Simpson, and M. T. Tsuang, "Negative, positive, and disorganized symptom dimensions in schizophrenia, major depression, and bipolar disorder," *Journal of Nervous and Mental Disease*, vol. 186, no. 8, pp. 470–476, 1998.
- [8] I. I. Gottesman, *Schizophrenia Genesis: The Origins of Madness*, Freeman, New York, NY, USA, 1991.
- [9] B. Kirkpatrick, R. Ram, X. F. Amador et al., "Summer birth and the deficit syndrome of schizophrenia," *The American Journal of Psychiatry*, vol. 155, no. 9, pp. 1221–1226, 1998.
- [10] N. C. Andreasen, M. Flaum, V. W. Swayze II, G. Tyrrell, and S. Arndt, "Positive and negative symptoms in schizophrenia. A critical reappraisal," *Archives of General Psychiatry*, vol. 47, no. 7, pp. 615–621, 1990.
- [11] N. C. Andreasen, S. Arndt, V. Swayze II et al., "Thalamic abnormalities in schizophrenia visualized through magnetic resonance image averaging," *Science*, vol. 266, no. 5183, pp. 294–298, 1994.
- [12] C. A. Bana E Costa, J.-M. De Corte, and J.-C. Vansnick, "MACBETH," *International Journal of Information Technology and Decision Making*, vol. 11, no. 2, pp. 359–387, 2012.
- [13] J. M. J. Aragonés, J. I. P. Sánchez, J. M. Doña, and E. Alba, "A neuro-fuzzy decision model for prognosis of breast cancer relapse," in *Current Topics in Artificial Intelligence*, vol. 3040 of *Lecture Notes in Computer Science*, pp. 638–645, Springer, 2004.
- [14] M. G. M. Hunink, P. P. Glasziou, J. E. Siegel et al., *Decision Making in Health and Medicine: Integrating Evidence and Values*, Cambridge University Press, 2001.
- [15] O. I. Reznik, *The Secrets of Medical Decision Making: How to Avoid Becoming a Victim of the Health Care Machine*, Loving Healing Press, 2006.
- [16] P. Szolovits, "How can computer programs reason?" in *CARDIOSTIM '88*, Mônaco, France, 1988.
- [17] P. Szolovits, "Uncertainty and decisions in medical informatics," *Methods of Information in Medicine*, vol. 34, no. 1-2, pp. 111–121, 1995.
- [18] J. C. C. B. S. D. Mello, E. G. Gomes, F. R. Leta, and R. B. V. Pessolani, "Conceitos Básicos do Apoio Multicritério à Decisão e sua Aplicação no Projeto Aerodesign," *Engevista*, vol. 5, no. 8, pp. 22–35, 2003.
- [19] D. C. Morais, *Modelagem Multicritério em Grupo para Planejamento Estratégico do Controle de Perdas no Abastecimento de Água [Tese Doutorado]*, Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Pernambuco, Recife, Brazil, 2006.
- [20] B. Roy, "Paradigms and challenges," in *Multiple Criteria Decision Analysis: State of the Art Surveys Series: International Series in Operations Research & Management Science*, J. Figueira, S. Greco, and M. Ehrgott, Eds., vol. 78, pp. 3–24, 2005.
- [21] P. Vincke, *Multicriteria Decision-Aid*, John Wiley & Sons, Chichester, UK, 1992.
- [22] R. Plácido, G. C. Pinheiro Gilberto, and A. K. A. de Castro, "A Estruturação do Modelo Multicritério para Produção de Jornal," *Pesquisa Operacional*, vol. 28, pp. 203–216, 2008.
- [23] C. A. Bana e Costa and M. P. Chagas, "A career choice problem: an example of how to use MACBETH to build a quantitative value model based on qualitative value judgments," *European Journal of Operational Research*, vol. 153, no. 2, pp. 323–331, 2004.
- [24] S. Barclay, *HIVIEW Software Package*, London School of Business, London, UK, 1984.
- [25] A. K. A. de Castro, P. R. Pinheiro, and M. C. D. Pinheiro, "A hybrid model for aiding in decision making for the neuropsychological diagnosis of Alzheimer's disease," in *Rough Sets and Current Trends in Computing*, vol. 5306 of *Lecture Notes in Computer Science*, pp. 495–504, Springer, Berlin, Germany, 2008.
- [26] A. T. Brasil, P. R. Pinheiro, and A. L. V. Coelho, "Towards the early diagnosis of Alzheimer's disease via a multicriteria classification model," in *Proceedings of the 5th Evolutionary Multi-Criterion Optimization (EMO '09)*, vol. 5467 of *Lecture Notes in Computer Science*, pp. 393–406, 2009.
- [27] J. H. M. Nogueira, J. F. L. Alcântara, R. C. de Andrade, R. B. Silva, and R. S. Silvestre, "Grupo SINTA—Sistemas INTeligentes Aplicados," Universidade Federal do Ceará (UFC) e Universidade Estadual do Ceará (UECE), 1995.
- [28] S. Chattopadhyay, D. K. Pratihar, and S. C. De Sarkar, "Fuzzy-logic-based screening and prediction of adult psychoses: a novel approach," *IEEE Transactions on Systems, Man, and Cybernetics A: Systems and Humans*, vol. 39, no. 2, pp. 381–387, 2009.
- [29] I. Tamanini and P. R. Pinheiro, "Reducing incomparability in multicriteria decision analysis: an extension of the ZAPROS method," *Pesquisa Operacional*, vol. 31, no. 2, pp. 251–270, 2011.
- [30] J. V. Rodado, M. J. Rodado, J. M. Chamizo et al., "Study of psychotic disorders with artificial neural networks," *European Psychiatry*, vol. 15, 2000.

Research Article

A Hybrid Intelligent Diagnosis Approach for Quick Screening of Alzheimer's Disease Based on Multiple Neuropsychological Rating Scales

Ziming Yin,¹ Yinhong Zhao,² Xudong Lu,¹ and Huilong Duan¹

¹College of Biomedical Engineering & Instrument Science, Zhejiang University, Hangzhou, Zhejiang 31002, China

²China National Center for Biotechnology Development, Building D, No. 16, Xisihuanzhonglu, Haidian District, Beijing 100036, China

Correspondence should be addressed to Xudong Lu; lxldxd66@gmail.com

Received 6 June 2014; Revised 20 November 2014; Accepted 20 November 2014

Academic Editor: José M. Jerez

Copyright © 2015 Ziming Yin 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.

Neuropsychological testing is an effective means for the screening of Alzheimer's disease. Multiple neuropsychological rating scales should be used together to get subjects' comprehensive cognitive state due to the limitation of a single scale, but it is difficult to operate in primary clinical settings because of the inadequacy of time and qualified clinicians. Aiming at identifying AD's stages more accurately and conveniently in screening, we proposed a computer-aided diagnosis approach based on critical items extracted from multiple neuropsychological scales. The proposed hybrid intelligent approach combines the strengths of rough sets, genetic algorithm, and Bayesian network. There are two stages: one is attributes reduction technique based on rough sets and genetic algorithm, which can find out the most discriminative items for AD diagnosis in scales; the other is uncertain reasoning technique based on Bayesian network, which can forecast the probability of suffering from AD. The experimental data set consists of 500 cases collected by a top hospital in China and each case is determined by the expert panel. The results showed that the proposed approach could not only reduce items drastically with the same classification precision, but also perform better on identifying different stages of AD comparing with other existing scales.

1. Introduction

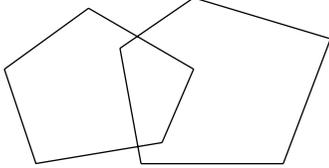
Alzheimer's disease (AD) is a degenerative senile dementia characterized by memory loss and cognitive functions disorders, and it is also one of the main types of senile dementia [1]. As AD has a slow onset and no highly specific diagnostic indicators at the early stage of the disease, it is particularly challenging for primary clinicians to identify transition points (from the asymptomatic phase to the symptomatic predementia phase to dementia onset) for individual patients [2, 3]. It is, nevertheless, important to identify these transition points between different stages, because studies [4] have proved that targeted therapies may help slow down the progress of the disease and improve quality of life for patients and their families.

Due to the lack of advanced medical facilities (advanced imaging and cerebrospinal fluid measures), the screening

of AD usually depends on the use of neuropsychological rating scales in primary clinics. Various neuropsychological rating scales, which are considered as a reliable and valid standardized testing tool, have been designed for cognitive abilities screening, and many of them have yielded good results as decision-making tools, such as minimal state examination (MMSE) [5], clinical dementia rating (CDR) [6], Montreal Cognitive Assessment (MoCA) [7], Geriatric Depression Scale (GDS) [8], and Activity of Daily Living Scale (ADL) [9]. Table 1 and Figure 1 are two most commonly used rating scales (the MMSE and the MoCA) in clinical practice.

However, each neuropsychological rating scale has its emphasis and limitation. A previous study has shown some scales do not perform well in one or more cognitive domains [10]. Multiple neuropsychological rating scales can cover more comprehensive cognitive domains. Therefore, multiple scales should be used together in order to get patients'

TABLE 1: Minimal state examination (MMSE).

Orientation	Year Month Day Date Time	___/5
	Country Town District Hospital Ward	___/5
Registration	Examiner names 3 objects (e.g. apple, table, and penny). Patient asked to repeat (1 point for each correct answer). THEN patient to learn the 3 names repeating until correct.	___/3
Attention and calculation	Subtract 7 from 100 and then repeat from result. Continue 5 times: 100 93 86 79 65 Alternative: spell "WORLD" backwards-"DLROW"	___/5
Recall	Ask for names of 3 objects learned earlier.	___/3
Language	Name a pencil and watch	___/2
	Repeat "No fits, ands, or buts"	___/1
	Give a 3-stage command. Score 1 for each stage. E.g., "Place index finger of right hand on your nose and then on your left ear"	___/3
	Ask patient to read and obey a written command on a piece of paper stating "Close your eyes"	___/1
	Ask patient to write a sentence. Score if it is sensible and has a subject and a verb	___/1
Copying	Ask the patient to copy a pair of intersecting pentagons:	
		___/1
Total		___/30

comprehensive cognitive status, which can help doctors to make correct diagnosis. However, this will bring two challenges: (1) neuropsychological testing requires highly trained assessors [11], while most primary clinicians are not qualified to conduct a full mental status examination or interpret a battery of scales' score; it is difficult for them to offer exact judgments about the examinee's cognitive state [12]. (2) Neuropsychological testing is quiet time consuming; the elders cooperate well only for short periods with the limitation of vitality and cognition [13], so long-time testing will bring negative impact on the quality of neuropsychological testing. Thus, we can conclude that the screening of AD in primary clinics should be based on the criteria that can get maximum accuracy in a convenient way within limited time.

To solve the above-mentioned challenges, identifying the items with the best ability to distinguishing AD (called critical items for short) from a battery of commonly used rating scales may help improve the efficiency of cognitive abilities screening. Then, a well-performance decision-making model, while the previously selected items can be taken as its input, may help primary clinicians improve diagnostic accuracy in routine clinical practice. So in this paper, we suggest dealing with the screening of AD by means of a two-stage hybrid intelligent approach based on multineuropsychological rating scales analysis: in Stage 1, use a genetic algorithm-rough sets (GA-RS) model to identify critical items, and in Stage 2, use a Bayesian network to develop a diagnosis assisting model of AD based on the selected items. This hybrid intelligent technique takes the advantage of attributes reduction of rough set theory requiring no prior knowledge and the uncertain reasoning ability of Bayesian network to build a relatively convenient and accurate decision-making model for primary clinicians.

The rest of this paper is organized as follows: Section 2 introduces the related work; Section 3 introduces basic concepts behind rough set theory, genetic algorithm, and Bayesian network; Section 4 presents the proposed approach including the proposed GA-RS attributes reduction algorithm applied in AD and the Bayesian network model constructed for AD diagnosis; Section 5 describes the evaluation results of the proposed model; then in Section 6, the discussions on some benefits and limitations of the proposed approach in the clinical environment are made; Section 7 draws conclusions and future work.

2. Related Work

Since the application of multiple neuropsychological rating scales is very time consuming and challenging for primary care physicians, many researchers have been trying to find out the most effective screening method for clinical application. For example, some scholars have tried to simplify the MMSE, which is the most widely used scale for screening dementia. Lou et al. [14] reported a 16-item simplified version of the original MMSE with high sensitivity and specificity. Callahan et al. [15] designed a six-item screener, which derived from the MMSE, the Blessed Dementia Rating Scale (BDRS), and the Word List Recall. And its sensitivity and specificity for a diagnosis of dementia reached 88.7% and 88.0%. In addition to the study of the MMSE, some researchers also have studied the combination of multiple neuropsychological rating scales. For example, Chen et al. [16] proposed an eight-item test, subtracted from the MMSE, the Clock Drawing Test (CDT), and the Instrumental Activities of Daily Living Scale (IADL). The evaluation result revealed that it was a sufficient and simple tool for the screening of early dementia in primary

Montreal cognitive assessment (MOCA)		Name:		Date of birth:			
Version 7.1 Original version		Education:		Date:			
Sex:							
Visuospatial/executive		Copy cube	Draw CLOCK (ten past eleven) 3 points			Points	
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	___/5	
			Contour Numbers Hands				
Naming						___/3	
Memory	Read list of words, subject must repeat them. Do 2 trials, even if 1st trail is successful. Do a recall after 5 minutes	Face	Velvet	Church	Daisy	Red	No points
	1st trail						
	2nd trail						
Attention	Read list of digits (1 digit/s)	Subject has to repeat them in the forward order <input type="checkbox"/> 2 1 8 5 4				___/2	
		Subject has to repeat them in the backward order <input type="checkbox"/> 7 4 2					
	Read list of letters. The subject must tap with his hand at each letter A. No points if >2 errors	<input type="checkbox"/> F B C M N A A J K L B A F A K D E A A A J A M O F A A B				___/1	
	Serial 7 subtraction starting at 100	<input type="checkbox"/> 93	<input type="checkbox"/> 86	<input type="checkbox"/> 79	<input type="checkbox"/> 72	<input type="checkbox"/> 65	___/3
	4 or 5 correct subtractions: 3 pts, 2 or 3 correct: 2 pts, 1 correct: 1 pt, 0 correct: 0 pt						
Language	Repeat: lonely know that John is the one to help today. <input type="checkbox"/>					___/2	
	The cat always hid under the couch when dogs were in the room. <input type="checkbox"/>						
	Fluency/name maximum number of words in one minute that begin with the letter F. <input type="checkbox"/> ____ (N ≥ 11 words)					___/1	
Abstraction	Similarity between e.g. banana-orange = fruit <input type="checkbox"/> train-bicycle <input type="checkbox"/> watch-ruler <input type="checkbox"/>					___/2	
Delayed recall	Has to recall words WITH NO CUE	Face	Velvet	Church	Daisy	Red	Points for UNCUED recall only
		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Optional	Category cue						
	Multiple choice cue						
Orientation	<input type="checkbox"/> Date	<input type="checkbox"/> Month	<input type="checkbox"/> Year	<input type="checkbox"/> Day	<input type="checkbox"/> Place	<input type="checkbox"/> City	___/6
Normal ≥ 26/30 Total						___/30	
Add 1 point if ≤12 year edu							

FIGURE I: Montreal Cognitive Assessment (MOCA).

care clinics. Besides, other studies [17–20] have reported various screening methods such as instruments designed for detection of memory, attention-executive function, visuospatial ability, and interview with reliable informants. After a general survey of the above studies, we found that all

these researchers selected items according to their own clinical intuition regarding domains of impairment commonly encountered in AD rather than objective data analysis based on past cases. This selection was influenced by subjective factor easily. After selecting discriminative items, a large

number of experiments need to be done for confirming weight coefficient of each item. For instance, the MoCA team spent over 5 years modifying the MoCA for clinical use [7]. So a method based on data mining may bring a new thought to searching the most discriminative items from multiple neuropsychological rating scales. It is based on large-scale objective clinical evidences, and, moreover, it can not only improve efficiency of identifying critical items but also adjust weight coefficient of items automatically.

The computer-aided diagnosis of AD is always the hot topic of research in the last few years. DeFigueiredo et al. [21] presented an algorithm based on the analysis of computer tomography image (CT) data from brain. This algorithm used an optimal interpolative neural network to classify individuals into four different groups (i.e., clinically diagnosed groups of elderly normal, demented, AD, and vascular dementia subjects). Ramírez et al. [22] showed a computer-aided diagnosis (CAD) system for the diagnosis of AD. His method is based on partial least squares regression model and a random forest predictor. The analyzed data is from single photon emission computed tomography (SPECT). Further works on SPECT data have reported high AD classification accuracy [23–25]. Duchesne et al. [26] presented their work on automated computer classification in Alzheimer’s dementia using the context of cross-sectional analysis of magnetic resonance images (MRI). Daliri [27] also presented an automated method for diagnosing AD from brain MR images. In addition to volumetric MRI, diffusion tensor imaging (DTI) has increasingly been used to detect microstructural brain differences in AD. Graña et al. [28] obtained discriminant features from two scalar measures of DTI data and used support vector machine (SVM) as a classifier. From the view point of methodology, most of the studies focused on finding the microstructural differences using image analysis based on supervised machine learning algorithms. Although the analysis based on imaging is becoming an important research trend of AD studies, the technique is difficult to be applied in primary clinical settings because access to imaging equipment may be limited there.

Based on above-mentioned analysis, we proposed a computer-aided quick screening method for AD based on multiple neuropsychological rating scales. We especially used data mining technique to reduce the items of scales and lower the barriers to applying this method in primary clinical settings.

3. Preliminary

This section addresses some basic concepts needed for the remainder of the discussion. We introduce rough set theory first and then discuss genetic algorithm and Bayesian network, so as to set up a necessary context for describing our approach.

3.1. Rough Set Theory. Rough set theory [29–31] is a mathematical approach to deal with imprecision, vagueness, and uncertainty. The main difference between rough set theory and other mathematical tools for dealing with uncertain

problems is that rough set theory does not need any prior information beyond the problem itself. Rough set theory is an important method for attribute reduction, where attributes that do not contribute to the classification of the given training data can be identified and removed.

Rough set theory is based on the establishment of equivalence classes within the given training data. All the data tuples forming an equivalence class are indiscernible; that is, the samples are identical with respect to the attributes describing the data. Formally, the indiscernibility relation is defined as follows:

$$\text{IND}(B) = \{(x, y) \in O^2 \mid \forall a \in B, f(x, a) = f(y, a)\}, \quad (1)$$

where A is a nonempty finite set of attributes and O is a nonempty finite set of objects. Given any $B \subseteq A$, relation $\text{IND}(B)$ induces a partition of O , which is denoted by $O/\text{IND}(B)$, where an element from $O/\text{IND}(B)$ is called an equivalence class.

It is common that some classes cannot be distinguished by given real-world data in terms of the available attributes. Rough sets can be used to approximately or “roughly” define such classes. A rough set definition for a given class X , $X \subseteq O$, is approximately by two sets—a lower approximation of X and upper approximation of X . The lower approximation of X consists of all the data tuples that, based on the knowledge of the attributes, are certain to belong to X without ambiguity. The upper approximation of X consists of all the tuples that, based on the knowledge of the attributes, cannot be described as not belonging to X . So the lower and upper approximation of X , is defined, respectively, as

$$\underline{X}(B) = \{x \in U \mid [x]_B \subseteq X\}, \quad (2)$$

$$\overline{X}(B) = \{x \in U \mid [x]_B \cap X \neq \Phi\},$$

where $x \in O$ and the pair $(\underline{X}(B), \overline{X}(B))$ is called the rough set with respect to X . The set $BN_B(x) = \overline{X}(B) - \underline{X}(B)$ is called the boundary region of X . Let $[x]_B$ denote the equivalence class of relation $\text{IND}(B)$ that contains element x .

An important concept in attributes reduction is dependency of attributes, which can be defined in the following way.

For A and B , B depends on A in degree γ_B ,

$$\gamma_B = \frac{|\text{POS}_A(B)|}{|O|}, \quad (3)$$

$$\text{POS}_A(B) = \bigcup_{X \in O/\text{IND}(B)} \underline{X}_A, \quad (4)$$

where $\gamma_B = 1$ means B depends totally on A . $0 < \gamma_B < 1$ indicates B depends partially on A , and if $\gamma_B = 0$, B is totally independent on A .

Formally, a reduct is a subset of attributes which can fully characterize the knowledge in the database. Let $\text{RED}(B)$ be the minimal subset of A and $\text{CORE}(B)$ the set of attributes which cannot be eliminated and the intersection of all reducts. If the attribute in $\text{CORE}(B)$ is removed, the ability to classify objects into the elementary classes of B will decrease:

$$\text{CORE}(B) = \bigcap_{R_i \in \text{RED}(B)} R_i, \quad i = 1, 2, 3, \dots \quad (5)$$

3.2. Genetic Algorithm. The genetic algorithm (GA) [32–34] is an optimized algorithm based on the Darwinian principle of natural selection. It can be used with other data mining techniques for optimization and performance amelioration. The genetic algorithm process starts with the randomly generated and encoded initial population, which includes several hundreds or thousands of potential solutions to the problem. Each encoded individual in the population is called chromosome and each bit in the chromosome is called gene and has a value. The next step is called genetic operators. The most widespread genetic operators include selection, crossover, and mutation. Each chromosome in the population is evaluated by user-defined fitness function. The higher a chromosome's fitness value is, the more likely it is to produce offspring. In this way the overall fitness of the population is guaranteed to increase and those with weak fitness will be eliminated gradually. Crossover forms new chromosomes for the population by exchanging a fixed part between two chromosomes. The chromosomes most often used for crossover are those destined to be eliminated from the population. Mutation can be applied by randomly flipping bits (or attribute values) within a single chromosome to avoid the local optima. New offspring is reevaluated by fitness function to search the solution. The whole process is repeated until reaching the prespecified number of generations or the desired level of fitness.

3.3. Bayesian Network. Bayesian network is an acyclic directed graph for representing probabilistic relationships among a set of random variables [35]. Trained Bayesian networks can be used for classification. There are two key elements of a Bayesian network: (1) a directed acyclic graph (DAG) encoding the dependence relationships among a set of variables and (2) a probability table associating each node to its immediate parent nodes. Each node in the directed acyclic graph represents actual attributes given in the data. Each arc represents a probabilistic dependence. If an arc is drawn from a node X to a node Y , then X is a parent of Y , and Y is a descendant of X [36]. Bayesian network has one conditional probability table (CPT) for each node. The CPT for a node Y specifies the conditional distribution $P(Y \mid \text{Parents}(Y))$, where $\text{Parents}(Y)$ are the parents of Y . A node within the network can be selected as an “output” node, representing a class label attribute. There may be more than one output node. Given a set of variables, the network can be used to compute the probabilities of the presence of various classes, rather than return a single class label. There have been some works with applications using Bayesian network in diagnosis of AD [37, 38].

4. Methods

4.1. Overview. In this section, we present the formation process of AD diagnosis assisting model with the proposed hybrid intelligent method, which consists of two steps: in Step 1, use a genetic algorithm-rough sets (GA-RS) model to identify critical items, and in Step 2, use a Bayesian network to build a diagnosis assisting model of AD based on selected items.

In the first step, finding critical items from multiple neuropsychological scales is the problem of attributes reduction, which is also a classical problem in machine learning. Rough set theory is a useful attributes reduction method in machine learning. It can find the shortest or minimal reducts while keeping high-quality classification performance [39]. However, current rough set approaches to attributes reduction are inadequate to find optimal reductions as no perfect heuristic can guarantee optimality. Optimal attribute reduction has been proved to be a NP-hard problem [40]. So, stochastic approaches provide a promising attributes reduction mechanism, like genetic algorithm (GA). GA is an effective and robust method for solving both constrained and unconstrained multiparameter optimization problems that is based on natural selection. Many literatures have combined rough set theory and genetic algorithm for solving machine learning problems in a variety of domains [41–44]. We present a genetic algorithm for reduct set computation which is very fast and gives a good approximation in the AD field.

As exact causes and the mechanism of AD remain uncertain, there is currently no method to ensure the AD presence except for an autopsy. Clinicians can only make the diagnosis called “probable or possible AD” in clinical environment, especially in the early stage of this disease. The proposed approach applies Bayesian network that has strong reasoning ability in solving uncertain problems to build the decision-making model and predict the AD probability rather than offer a definitive diagnosis. So it is more conducive to the practical application of the proposed approach.

4.2. Attributes Reduction Based on Genetic Algorithm and Rough Set Theory. As mentioned above, GA-RS is used to identify critical items from a battery of rating scales. Each step of the algorithm is described as follows.

4.2.1. Chromosome Representation. Because genetic algorithm cannot deal with data in solution space directly, we must represent them as binary strings of length M which is the number of the condition attributes by encoding. Binary encoding is simple and easy to operate. Each binary string is called a chromosome, in which “1” means that the corresponding attribute is selected and “0” means not. Attributes in *Core* should take “1”, and remain the same in the whole process of evolution, since genetic search starts from the *Core*.

4.2.2. Fitness Function. Fitness function is a user-defined function which is used to measure each chromosome's optimization calculation in the groups. The fitness value of each chromosome represents suitability for the environment. In this paper, we expect the “best” chromosome could have the minimal length and the strongest classification performance as the algorithm proceeds. So the fitness function is defined as follows:

$$\begin{aligned} F(x) &= \beta \cdot f(x) + p(x) \\ &= \beta \cdot \left(1 - \frac{\text{card}(x)}{\text{card}(C)} \right) + \frac{\text{card}(\text{POS}_x(D))}{\text{card}(\text{POS}_C(D))}, \end{aligned} \quad (6)$$

where $\text{card}(x)$ is the number of “1” in chromosome, which means the number of condition attributes contained by chromosome; $\text{card}(C)$ is the length of chromosome, which is the total number of condition attributes; $f(x) = 1 - \text{card}(x)/\text{card}(C)$ indicates the chromosome x that is not included in the proportion of condition attributes. $p(x)$ indicates the distinction ability of attribute x .

4.2.3. Selection Method. Select chromosomes based on their fitness values from the current population to produce offspring for the new population. Tournament selection is used, which means the higher the fitness value is, the higher probability of that chromosome is selected for reproduction. This step is repeated until the number of chromosomes selected is equal to the number of the population.

4.2.4. Crossover and Mutation. One-point crossover method is used to reproduce with a probability of P_c . In mutation process, we first select a chromosome to be mutated with probability P_m and then replace a single gene of the chromosome from “1” to “0” or from “0” to “1” randomly.

4.2.5. Elitist Strategy. We take the elite strategy [45] to preserve the best individual of the fitness function value. Copy the individual of highest fitness value in the current generation to the next generation, unaltered.

The detail of the whole algorithm is as follows.

Input. Decision table $IS = \langle O, A, V, f \rangle$; O is a nonempty finite set of objects. A is a nonempty finite set of attributes: $A = C \cup D$, C is the set of condition attributes, and D is the set of decision attributes. $V = \bigcup_{a \in A} V_a$; V_a is the set of values of attribute $a \in A$. $f : O \times A \rightarrow V$ is an information function so that, for any $a \in A$ and $x \in O$, $f(x, a) \in V_a$.

Output. There is an attributes reduction R of decision table.

Steps

Step 1. Calculate the dependency $\gamma_c(D)$ between decision attributes set D and condition attributes set C by formula (3).

Step 2. Let $\text{Core}(C) = \varphi$, to get rid of each attribute $c \in C$ one by one, if $\gamma_{C-c} \neq \gamma_C$, $\text{Core}(C) = \text{Core}(C) \cup \{c\}$ which means the core is $\text{Core}(C)$; if $\gamma_{\text{Core}(C)}(D) = \gamma_C(D)$, then the core is minimal attributes reduction and if not, go to Step 3.

Step 3. Generate m binary strings with length n randomly, which can be seen as the initial population. n is the number of the condition attributes. “1” means that the corresponding attribute is present, and “0” indicates not. For attributes in core, corresponding position is “1” and for others, corresponding position is “1” or “0” randomly.

Step 4. Calculate the fitness value for each individual by formula (6) and select individuals by tournament selection.

Step 5. Perform crossover operation according to the crossover probability P_c , using single-point crossover mode.

TABLE 2: Parameter settings of GA.

Population size	1000
Number of generations	500
Initialization method	Binary method
Percentage of elite	0.2
Selection method	Tournament selection
Crossover method	Uniform crossover
Crossover ratio	0.5
Mutation method	Single-point mutation
Mutation ratio	0.03

Step 6. Perform mutation operation according to the mutation probability P_m . We basically bit mutation strategy while the corresponding bit of attributes in the *Core* does not change.

Step 7. Select the individuals with the best fitness values to be offspring of the current generation. This strategy is to guarantee the best chromosome could carry over to the next generation.

Step 8. Repeat the genetic operation until either one of the following conditions is satisfied: (1) the maximum number of generations is achieved or (2) the fitness value of the best individual for the present generation no longer changes during several successive generations.

Step 9. Convert the best individual to condition attribute and get the final result.

The whole computation steps are shown in Figure 2.

Parametric settings of genetic algorithm are as follows: population scale $N = 1000$, crossover ratio $P_c = 0.5$, mutation ratio $P_m = 0.03$, and the largest number of iterations is 500, just as demonstrated in Table 2.

The fitness function employed in this paper controls the chromosomes that evolve in the direction of the minimum reduction while keeping the classification performance: the higher the $\text{card}(x)$ is, the smaller the $f(x)$ is; the larger $p(x)$, the more dependence between the condition attribute C and decision attribute D . This algorithm ensures the two requirements, so the result is the optimal solution of the problem.

In our approach, attributes reduction mentioned above is not the final goal but an intermediate process and core technology of AD diagnosis assisting for clinician in primary clinic. An uncertainty inference model for AD should be built after attributes reduction, which will be discussed in next section.

4.3. Bayesian Network Model for AD Diagnosis. Based on the above step, we attempt to construct the structural model for AD diagnosis. These selected items can be represented as input variables of the model. Since there is strong diagnostic uncertainty earlier in the disease process, an uncertainty inference model must be built. A popular modeling tool for complex uncertain domains is a Bayesian network.

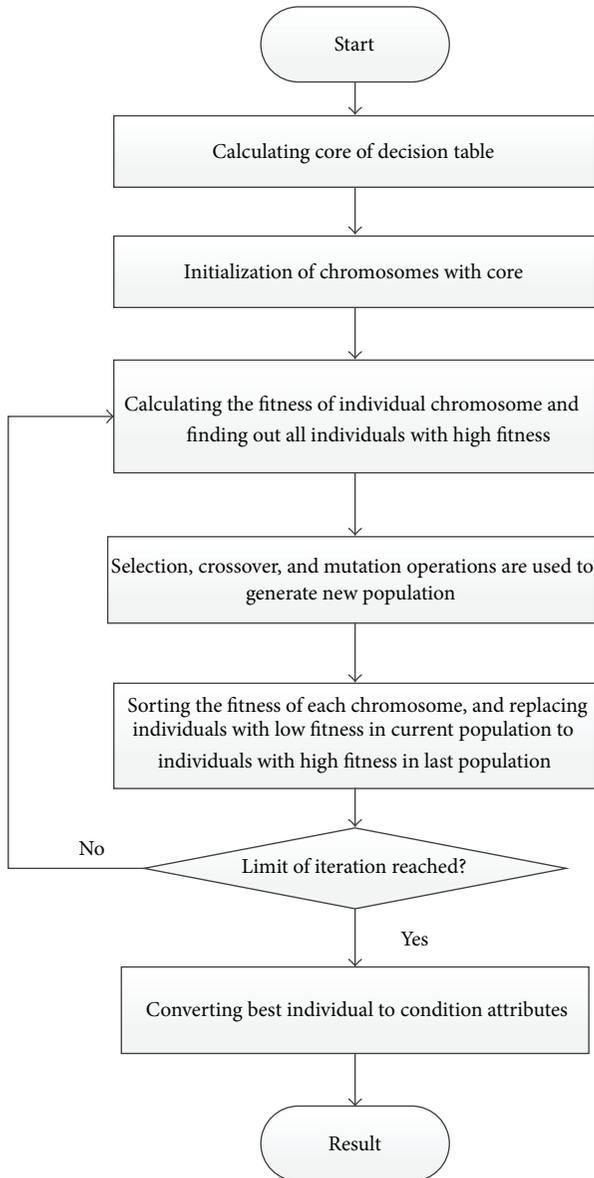


FIGURE 2: Detailed algorithm flow of GA-RS.

5. Experiment and Results

5.1. Data Collection. The experimental data set is composed of 500 consecutive historical cases collected by the neurology department of a certain top hospital in China from 2009 to 2014. Each case is a series of scale scores belonging to one subject, and each subject has only one case. All neuropsychological tests were conducted by trained neuropsychologists and administered on the same day. The mean age of subjects is 74.4 (range, 51–92); 59.5 of the subjects' percent were female. These 500 historical cases have the following characteristics.

- (1) All these 11 neuropsychological rating scales are selected from a large number of scales by leading experts in neurology, including the MMSE, the

MoCA, the CDR, the GDS, the ADL, the Word-List Learning, the figure copying, the new word discriminating, the trail making test, the similarity, and the perception. All these neuropsychological rating scales are commonly used instruments for screening cognitive or noncognitive impairment in the clinical diagnosis.

- (2) Each scale consists of a series of items. In total, there are 101 testing items in these scales. Some items are straightforward Q & A pattern, for instance, "What is the date?" Some others need the subject to do some actions, "Please read this and do what it says. (Show subjects the following words on the stimulus form: Close your eyes.)" Each of the tests scores points if it is answered correctly.

To ensure the correctness of diagnosis of each case, an expert panel group composed of three neuropsychologists was set up, and the diagnosis of each case was determined by the panel. The diagnosis of experts not only depended on an objective neuropsychological testing, but also on the history-taking from the patient and a knowledgeable informant. Their diagnosis was regarded as the gold standard. In the current study, the diagnosis of cases could be divided into three types: patients with AD, patients meeting criteria for mild cognitive impairment [46] (called MCI for short, which is regarded as the predementia stage of AD), and the elderly subjects with normal cognition, in which, the number of each type is 33.5%, 37.7%, and 28.8%, respectively.

Parts of cases are given in Table 3. In the table, each column is one testing item of scales, for instance, Time Orientation, Place Orientation and Repetition belong to the MMSE while Visuospatial Skills belongs to the MoCA. They are regarded as the condition attributes. The last column, Result, is the decision attribute (the diagnosis of each patient).

5.2. Experimental Design and Results. To verify the feasibility and validity of the proposed approach, the performance of proposed approach can be measured by the following evaluations: (1) reduction ratio on testing duration and reduction ratio on quantity of items; (2) comparison with multiple classifiers; (3) comparison of classification accuracy before and after reduction; (4) the performance of classification compared with two existing cognitive screening scales.

We applied two evaluation methods to prove the reliability of our experimental results, one was 10-fold cross-validation, and the other was 0.632 bootstrap. Obtain computing results by averaging after executing 10 times. Recall rate, precision rate, and accuracy were selected as the performance evaluation metrics. The obtained results of the experiment are to be presented and discussed in the next section.

5.3. Reduction Result. After attributes reduction, 10 items were selected finally, which are listed in Table 4.

Some items represent one test, such as Figure_Copy, Figure_Short_Memory, Figure_Delay_Memory, Word_Delay_Recall, and Reading_Comprehension, while others are a set of several tests, which cannot be separated (because these

TABLE 3: Part of AD dataset.

Fact	Time orientation	Place orientation	Repetition	Visuospatial skills	...	Result
1	5	5	1	5	...	Normal
2	4	5	1	5	...	Normal
3	4	5	1	3	...	MCI
4	1	4	0	3	...	AD
⋮	⋮	⋮	⋮	⋮	⋮	⋮
500	0	0	1	1	...	AD

TABLE 4: Reduction results by GA-RS.

Result	Source
Reading_Comprehension	MMSE
Visuospatial_Execution	MoCA
Naming	MoCA
Attention	MoCA
Figure_Copy	Figure copying test
Figure_Short_Memory	Figure copying test
Figure_Delay_Memory	Figure copying test
IADL	ADL
Word_Delay_Recall	Word-List Learning test
Word_AVG	Word-List Learning test

tests significantly correlate with one another and must be performed together), such as Visuospatial_Execution, IADL, Naming, Attention, and Word_AVG.

5.3.1. Reduction Ratio. We used reduction ratio including the reduction ratio of testing duration and the reduction ratio of quantity of items as measurable metrics. Assume that the number of condition attributes before and after reduction is m and n , respectively. The reduction ratio r is defined:

$$r = \frac{(m - n)}{m} \times 100\%. \quad (7)$$

Before reduction, the number of items is 34, while only 10 items left after reduction using the proposed method, so we can conclude that the reduction ratio is 70.59%. The experimental results indicate that using GA-RS to select subset can reduce items dramatically.

Similarly, the reduction ratio of testing duration can also be calculated using formula (7). In clinical practice, the duration of finishing these scales varies a lot, which depends on the subject's state of cognitive impairment. According to [47, 48], the performance time for the MMSE and the MoCA is 13.4 minutes and 14.8 minutes on average, respectively. Based on the past experience, a skilled clinician administers the scale for more than one hour to complete all the 11 scales mentioned above. By using the proposed model, clinicians do not have to finish all the scales but only need to complete the selected testing items. Hence the test duration is reduced greatly and ranges from 12 to 15 minutes with a mean time of 13.5 minutes and a standard deviation of 2.3 minutes.

5.3.2. Comparison with Multiple Classifiers. The constructed Bayesian network structure is presented in Figure 3.

We compared some common used classifiers with Bayesian network in order to select the well-performed classifier. All these classifiers had the same input items (the items selected by the above step). The result of comparison is as shown in Table 5.

From Table 5, we could see that the Bayesian network performed best in the four classifiers. Then, in order to further prove the effectiveness of results, we compared the four groups using Friedman test to see if a significant difference emerged (Table 6). Table 6 shows the mean rank for each classifier and Table 7 shows the result of the Friedman test.

In Table 7, $P = 0.001 < 0.05$, there was a significant difference between these four classifiers. However, the cross-validation estimate of prediction error may lead to a high variability in results. In order to validate the result further, 0.632 bootstrap was used to evaluate the performance, as shown in Table 8.

Similar result came from Table 8 when compared with Table 5. It suggested that Bayesian network performed better than other three classifiers. And Friedman test was also performed, as shown in Tables 9 and 10. We found that there existed a significant difference between the four groups as well. Because of low variance with only moderate bias, the result got by 0.632 bootstrap was selected as the final result.

5.3.3. Classification Performance before and after Reduction.

In order to evaluate the validity of attributes reduction, we used Bayesian network algorithm to compute the classification performance before and after attributes reduction, respectively, and to check whether or not the classification performance had changed.

Each subject had been given the probability of each classification. The highest probability was regarded as the diagnosis of the model. Table 11 presents a summary of the classification results before and after reduction.

From the variance of recall rate and precision rate after attributes reduction as shown in Table 11, we found that the recall rate and precision rate of each group decreased a little, but less than 3.05%. We analyzed the result data using Wilcoxon Signed-Rank Test. The calculated P value was 0.853 and larger than 0.05, so the null hypothesis was true, which means that there was no significant statistical difference between these two methods. In conclusion, the comparative experimental results indicated that the proposed

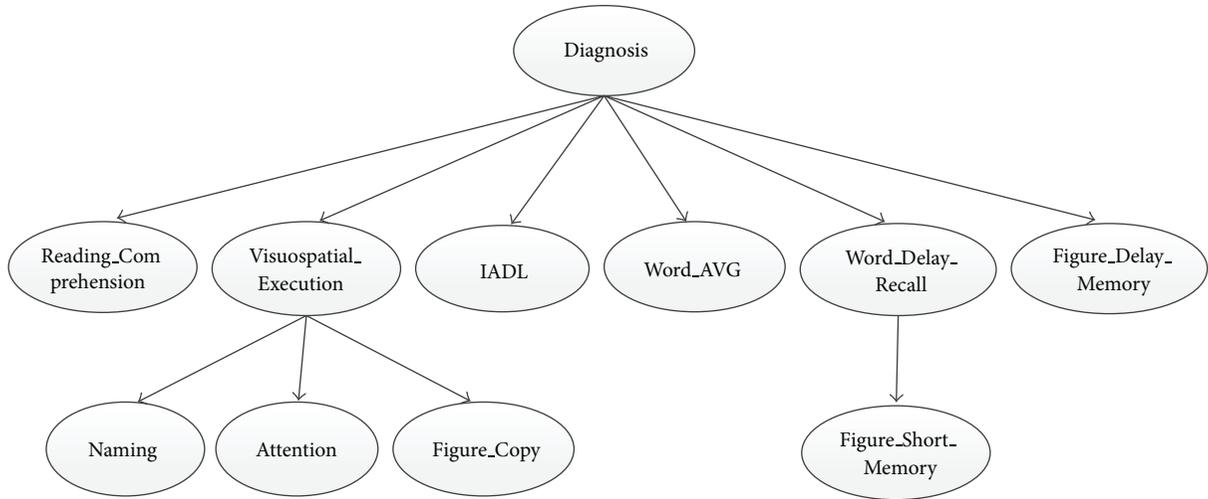


FIGURE 3: AD-related diagnosis assisting model.

TABLE 5: The comparison of classifiers by 10-fold validation.

	Normal		MCI		AD		ACC%
	R (%)	P (%)	R (%)	P (%)	R (%)	P (%)	
BN	85.41	87.23	78.95	73.17	90.7	95.12	85.27
NN	77.27	79.07	69.05	64.44	86.05	90.24	77.52
C4.5	81.82	63.16	42.86	48.65	72.09	88.57	65.89
SMO	86.36	74.51	64.29	77.14	93.02	93.02	81.35

R: recall rate; P: precision rate; ACC: accuracy; BN: Bayesian network; NN: neural network.

TABLE 6: Ranks in Friedman tests.

	Mean rank
Bayesian network	2.58
Decision tree (C4.5)	2.27
Neural network	2.50
SMO	2.66

TABLE 7: Test statistics in Friedman tests.

Chi-square	15.789
df	3
Asymp.Sig.	0.001

method could find the shortest or minimal reducts while keeping high-quality classification performance.

5.3.4. *Comparison with Comprehensive Cognitive Screening Scales.* Comprehensive cognitive screening scales measure all important aspects of cognitive function, such as memory, language, visuospatial skills, attention, and executive function. The most commonly used comprehensive cognitive screening scales include the MMSE and the MoCA. Our computer-aided model also covers multiple aspects of cognitive screening, so the comparison of recall rate and precision rate between our model and these two scales is needed based on the same dataset in order to evaluate the validity of our model.

TABLE 8: The comparison of classifier by 0.632 bootstrap.

	Normal		MCI		AD		ACC%
	R (%)	P (%)	R (%)	P (%)	R (%)	P (%)	
BN	84.55	84.91	75	76.39	91.94	89.81	83.73
NN	78.99	71.76	60.63	67.68	91.55	91.98	76.94
C4.5	71.90	66.67	46.45	50.26	79.15	80.29	66.24
SMO	84.71	79.46	63.03	69.63	89.57	87.91	79.61

R: recall rate; P: precision rate; ACC: accuracy; BN: Bayesian network; NN: neural network.

TABLE 9: Ranks in Friedman tests.

	Mean rank
Bayesian network	2.59
C4.5	2.30
Neural network	2.54
SMO	2.57

TABLE 10: Test statistics in Friedman tests.

Chi-square	48.694
df	3
Asymp.Sig.	0.000

TABLE 11: Comparison of the classification performance before and after attributes reduction by Bayesian network.

	Normal		MCI		AD		ACC%
	R (%)	P (%)	R (%)	P (%)	R (%)	P (%)	
Before	86.36	84.44	76.19	76.19	90.7	92.86	84.50
After	84.55	84.91	75	76.39	91.94	89.81	83.73

R: recall rate; P: precision rate; ACC: accuracy.

The MMSE is a questionnaire test that is used to screen cognitive impairment. The total score is 30. If the score is greater than or equal to 27 points, it means the subject has

TABLE 12: Comparison with the MMSE and the MoCA.

	Normal		MCI		AD		ACC%
	R (%)	P (%)	R (%)	P (%)	R (%)	P (%)	
MMSE	93.88	67.65	45.83	61.11	84.29	93.65	74.29
MoCA	68.89	86.11	81.82	60.00	80.39	93.18	77.14
Model	84.55	84.91	75.00	76.39	91.94	89.81	83.57

R: recall rate; P: precision rate; ACC: accuracy.

a normal cognition. Below 27 points, scores can be divided into several stages: severe cognitive impairment (≤ 9 points), moderate cognitive impairment (10–18 points), and mild (19–24 points) cognitive impairment [49]. The MoCA is also a one-page 30-point test developed as a brief cognitive screening tool to detect mild-moderate cognitive impairment. The suggested cut-off score on MoCA is 26, which yielded the best balance between sensitivity and specificity for the MCI and AD groups. We applied these criteria to our dataset, and the result is showed in Table 12.

From Table 12, we found that the MMSE did not perform well as a screening instrument for MCI due to the lack of sensitivity to MCI [50]. Some researchers believe the low sensitivity of the MMSE comes from the emphasis placed on language items and a paucity of visuospatial items [51]. However, visuospatial skills and executive function had been retained through our attributes reduction algorithm. So our model performed much better than the MMSE on detecting MCI, which is a very significant stage of AD. In general, our model is a more effective tool for identifying different stages of AD than the MMSE.

The MoCA is designed for the detection of MCI; that is to say, it is developed to screen patients who has cognitive impairment complaints but still performed in the normal range on the MMSE. The MoCA is more sensitive on detecting MCI than the MMSE, because the MoCA focuses more on tasks of frontal executive functioning and attention. Our model retained these key parts of the MoCA, so its performance on detecting MCI was close to that of the MoCA. Above all, our model was advantageous when identifying multiple transition points between different stages of AD, and it was not only designed for screening MCI. Compared with the MoCA, our model had distinct advance on differentiating normal and AD while almost keeping the sensitivity of detecting MCI. This was more helpful for primary clinicians to take target care and therapies.

We also performed Friedman test on these three groups and the actual result of the Friedman test is shown in Tables 13 and 14. From the result, we can see that there is an overall statistically significant difference between the mean ranks of these three groups.

6. Discussion

This study proposed a computer-aided diagnosis model of AD applied in primary clinics. In order to solve especially the problem that cognitive screening based on multiple neuropsychological scales was time consuming, GA-RS algorithm was used to identify the most related items from numerous rating scales while ensuring a satisfactory accuracy

TABLE 13: Ranks in Friedman tests.

	Mean rank
MMSE	1.94
MoCA	1.98
Model	2.08

TABLE 14: Test statistics in Friedman tests.

Chi-square	3.756
df	2
Asymp.Sig.	0.043

of classification. The experimental results validated the effectiveness of the proposed approach.

The benefits of this study are listed as follows.

- (i) The proposed approach is suitable to be applied in the primary clinics, because the clinician's day is labor-intensive, and the mean time of clinical interviewing is limited for each patient. Clinicians are eager to have a tool for the screening of AD without spending too much time. The proposed approach reduces the testing time for each patient while keeping classification accuracy. Hence, such computer-aided approach is applicable in clinical practice.
- (ii) As an important merit of the proposed approach, it is a computer-aided diagnostic tool based on multiple neuropsychological rating scales, rather than neuroimaging, or biomarker. To the best of our knowledge, there exist few reports providing a computer-aided diagnosis method that is completely based on multiple neuropsychological rating scales. Thus our approach is suitable to be popularized in the primary clinics which have no advanced imaging and biological molecular equipment.
- (iii) It is estimated by the specialists that there are 68 relevant AD scales in the world [52], most of which have established the normative data and interpretation of scores in different countries. However, there have been seldom studies on how to employ so many rating scales to give a comprehensive diagnosis. For instance, if a patient "MMSE = 27", "MoCA = 23", and "ADL = 26," then what is the comprehensive status of the patient? It is relatively difficult for young general practitioners to make diagnosis in primary clinics. Our approach provides a new thought to solve this problem in hopes of supplementing the research in this field.
- (iv) The proposed approach is based on neuropsychological rating scales. Any disease that has no specific golden criteria and needs a long test by rating scales can try this method.

It should also be mentioned that there remain some limits to the approach proposed in this paper.

First of all, the chosen data might be bias, as all the cases of the study were collected only in one hospital. Secondly, there are some other types of dementia, such as Lewy body

dementia and vascular dementia, which are difficult to be distinguished from AD for young practitioners. The cases of these diseases are not included in our dataset, so clinicians must differentiate diagnostic methods for these diseases when using our model. Thirdly, the classification performance of Bayesian Network does not perform as well as expected, more machine learning algorithm can be tried to improve the classification performance. Finally, the number of subjects is still limited and more subjects are necessary in order to generalize the results to a larger population.

7. Conclusion and Future Work

The increasing aging population has led to a high increase in the prevalence of AD. Due to the fact that targeted care and therapies may slow down the progression of disease, the identification of different stages of AD is very important. In this paper, we proposed a computer-aided diagnosis method for AD based on analyzing the practical scores of rating scales. We especially identified the most discriminative items based on rough set theory and genetic algorithm. The selected items cover multiple cognitive domains and can be administered generally within 15 minutes. So it is user-friendly and is quickly administered, it may be appropriate use in primary clinics where assessment time is often limited. By comparing the classification performance, the result showed that the approach can effectively reduce the representation space of the attributes whilst hardly decreasing classification precision. The data also indicated that it has satisfactory reliability for both MCI and AD comparing with other existed cognitive screening scales.

Without doubt, opportunities for future research are abundant. First, we plan to further evaluate the built model with a perspective study in a real clinical setting. Second, more rating scales for specific dementias are going to be involved in the training set data and more comprehensive model for senile dementia will be built in the future work. Based on above work, a “three-level medical service network” for AD is going to be built in the near future and different computer-aided diagnosis tools for each level hospital will be developed; for example, the simple cognitive screening tool helps clinicians in primary clinics to judge whether patients suffer from cognitive impairment; the advanced cognitive assessment tool helps clinicians in second class hospitals to estimate the severity of cognitive impairment; the comprehensive assisted diagnosis tool is designed for clinicians in top hospitals to differentiate the types of dementia. The setup of such network will improve diagnosis accuracy of AD greatly and reduce the burden on public health care resource.

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 High Technology Research and Development Program of China (863 Program)

under Grant no. 2012AA02A601 and the National Science and Technology Major Project of China under Grant no. 2013ZX03005012.

References

- [1] C. S. Porto, H. C. Fichman, P. Caramelli, V. S. Bahia, and R. Nitrini, “Brazilian version of the Mattis dementia rating scale: diagnosis of mild dementia in Alzheimer’s disease,” *Arquivos de Neuro-Psiquiatria*, vol. 61, no. 2B, pp. 339–345, 2003.
- [2] R. A. Sperling, P. S. Aisen, L. A. Beckett et al., “Toward defining the preclinical stages of Alzheimer’s disease: recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease,” *Alzheimer’s & Dementia*, vol. 7, no. 3, pp. 280–292, 2011.
- [3] G. M. McKhann, D. S. Knopman, H. Chertkow et al., “The diagnosis of dementia due to Alzheimer’s disease: recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease,” *Alzheimer’s and Dementia*, vol. 7, no. 3, pp. 263–269, 2011.
- [4] J. Rhodes-Kropf, “Palliative care for patients with Alzheimer’s dementia: advance care planning across transition points,” in *Choices in Palliative Care*, pp. 144–156, Springer, 2007.
- [5] M. F. Folstein, S. E. Folstein, and P. R. McHugh, ““Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician,” *Journal of Psychiatric Research*, vol. 12, no. 3, pp. 189–198, 1975.
- [6] C. P. Hughes, L. Berg, and W. L. Danziger, “A new clinical scale for the staging of dementia,” *The British Journal of Psychiatry*, vol. 140, no. 6, pp. 566–572, 1982.
- [7] Z. S. Nasreddine, N. A. Phillips, V. Bédirian et al., “The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment,” *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695–699, 2005.
- [8] J. A. Yesavage, T. L. Brink, T. L. Rose et al., “Development and validation of a geriatric depression screening scale: a preliminary report,” *Journal of Psychiatric Research*, vol. 17, no. 1, pp. 37–49, 1982.
- [9] P. Self-Maintenance, *Assessment of Older People: Self-Maintaining and Instrumental Activities of Daily Living*, 1969.
- [10] T. N. Tombaugh and N. J. McIntyre, “The mini-mental state examination: a comprehensive review,” *Journal of the American Geriatrics Society*, vol. 40, no. 9, pp. 922–935, 1992.
- [11] C. Zadikoff, S. H. Fox, D. F. Tang-Wai et al., “A comparison of the mini mental state exam to the Montreal cognitive assessment in identifying cognitive deficits in Parkinson’s disease,” *Movement Disorders*, vol. 23, no. 2, pp. 297–299, 2008.
- [12] S. Borson, J. Scanlan, J. Hummel, K. Gibbs, M. Lessig, and E. Zuhr, “Implementing routine cognitive screening of older adults in primary care: process and impact on physician behavior,” *Journal of General Internal Medicine*, vol. 22, no. 6, pp. 811–817, 2007.
- [13] H. Halstead, “A psychometric study of senility,” *The British Journal of Psychiatry*, vol. 89, no. 376–377, pp. 363–373, 1943.
- [14] M.-F. Lou, Y.-T. Dai, G.-S. Huang, and P.-J. Yu, “Identifying the most efficient items from the Mini-Mental State Examination for cognitive function assessment in older Taiwanese patients,” *Journal of Clinical Nursing*, vol. 16, no. 3, pp. 502–508, 2007.
- [15] C. M. Callahan, F. W. Unverzagt, S. L. Hui, A. J. Perkins, and H. C. Hendrie, “Six-item screener to identify cognitive impairment

- among potential subjects for clinical research,” *Medical Care*, vol. 40, no. 9, pp. 771–781, 2002.
- [16] C.-Y. Chen, K.-K. Leung, and C.-Y. Chen, “A quick dementia screening tool for primary care physicians,” *Archives of Gerontology and Geriatrics*, vol. 53, no. 1, pp. 100–103, 2011.
- [17] R. B. Lipton, M. J. Katz, G. Kuslansky et al., “Screening for dementia by telephone using the memory impairment screen,” *Journal of the American Geriatrics Society*, vol. 51, no. 10, pp. 1382–1390, 2003.
- [18] A. F. Jorm, “The Informant Questionnaire on cognitive decline in the elderly (IQCODE): a review,” *International Psychogeriatrics*, vol. 16, no. 3, pp. 275–293, 2004.
- [19] S. Borson, J. M. Scanlan, J. Watanabe, S.-P. Tu, and M. Lessig, “Simplifying detection of cognitive impairment: comparison of the Mini-Cog and Mini-Mental State examination in a multiethnic sample,” *Journal of the American Geriatrics Society*, vol. 53, no. 5, pp. 871–874, 2005.
- [20] S. Kilada, A. Gamaldo, E. A. Grant, A. Moghekar, J. C. Morris, and R. J. O’Brien, “Brief screening tests for the diagnosis of dementia: comparison with the Mini-Mental State Exam,” *Alzheimer Disease and Associated Disorders*, vol. 19, no. 1, pp. 8–16, 2005.
- [21] R. J. P. DeFigueiredo, W. R. Shankle, A. Maccato et al., “Neural-network-based classification of cognitively normal, demented, Alzheimer disease and vascular dementia from single photon emission with computed tomography image data from brain,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 12, pp. 5530–5534, 1995.
- [22] J. Ramírez, J. M. Górriz, F. Segovia et al., “Computer aided diagnosis system for the Alzheimer’s disease based on partial least squares and random forest SPECT image classification,” *Neuroscience Letters*, vol. 472, no. 2, pp. 99–103, 2010.
- [23] R. Chaves, J. Ramírez, J. M. Górriz et al., “SVM-based computer-aided diagnosis of the Alzheimer’s disease using t-test NMSE feature selection with feature correlation weighting,” *Neuroscience Letters*, vol. 461, no. 3, pp. 293–297, 2009.
- [24] P. Padilla, J. M. Górriz, J. Ramírez et al., “Analysis of SPECT brain images for the diagnosis of Alzheimer’s disease based on NMF for feature extraction,” *Neuroscience Letters*, vol. 479, no. 3, pp. 192–196, 2010.
- [25] F. Segovia, J. M. Górriz, J. Ramírez et al., “Classification of functional brain images using a GMM-based multi-variate approach,” *Neuroscience Letters*, vol. 474, no. 1, pp. 58–62, 2010.
- [26] S. Duchesne, A. Caroli, C. Geroldi, C. Barillot, G. B. Frisoni, and D. L. Collins, “MRI-based automated computer classification of probable AD versus normal controls,” *IEEE Transactions on Medical Imaging*, vol. 27, no. 4, pp. 509–520, 2008.
- [27] M. R. Daliri, “Automated diagnosis of alzheimer disease using the scale-invariant feature transforms in magnetic resonance images,” *Journal of Medical Systems*, vol. 36, no. 2, pp. 995–1000, 2012.
- [28] M. Graña, M. Termenon, A. Savio et al., “Computer Aided Diagnosis system for Alzheimer Disease using brain Diffusion Tensor Imaging features selected by Pearson’s correlation,” *Neuroscience Letters*, vol. 502, no. 3, pp. 225–229, 2011.
- [29] Z. Pawlak, “Rough sets,” *International Journal of Computer & Information Sciences*, vol. 11, no. 5, pp. 341–356, 1982.
- [30] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data*, vol. 9, Springer, 1991.
- [31] Z. Pawlak, “Rough set approach to knowledge-based decision support,” *European Journal of Operational Research*, vol. 99, no. 1, pp. 48–57, 1997.
- [32] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Menlo Park, Calif, USA, 1989.
- [33] F. Ahmad, N. A. Mat Isa, Z. Hussain, and M. K. Osman, “Intelligent medical disease diagnosis using improved hybrid genetic algorithm—multilayer perceptron network,” *Journal of Medical Systems*, vol. 37, no. 2, article 9934, 2013.
- [34] M. R. Daliri, “A hybrid automatic system for the diagnosis of lung cancer based on genetic algorithm and fuzzy extreme learning machines,” *Journal of Medical Systems*, vol. 36, no. 2, pp. 1001–1005, 2012.
- [35] J. Peral, *Probabilistic Reasoning in Intelligent Systems*, vol. 12, Morgan Kaufmann, San Mateo, Calif, USA, 1988.
- [36] C. B. Hall, J. Ying, L. Kuo, and R. B. Lipton, “Bayesian and profile likelihood change point methods for modeling cognitive function over time,” *Computational Statistics & Data Analysis*, vol. 42, no. 1, pp. 91–109, 2003.
- [37] P. Bélisle, L. Joseph, D. B. Wolfson, and X. Zhou, “Bayesian estimation of cognitive decline in patients with Alzheimer’s disease,” *The Canadian Journal of Statistics*, vol. 30, no. 1, pp. 37–54, 2002.
- [38] M. J. Prince, “Predicting the onset of Alzheimer’s disease using Bayes’ theorem,” *American Journal of Epidemiology*, vol. 143, no. 3, pp. 301–308, 1996.
- [39] R. W. Swiniarski and A. Skowron, “Rough set methods in feature selection and recognition,” *Pattern Recognition Letters*, vol. 24, no. 6, pp. 833–849, 2003.
- [40] A. Skowron and C. Rauszer, “The discernibility matrices and functions in information systems,” in *Intelligent Decision Support*, pp. 331–362, Springer, 1992.
- [41] C.-S. Ong, J.-J. Huang, and G.-H. Tzeng, “Building credit scoring models using genetic programming,” *Expert Systems with Applications*, vol. 29, no. 1, pp. 41–47, 2005.
- [42] W.-Y. Liang and C.-C. Huang, “The generic genetic algorithm incorporates with rough set theory—an application of the web services composition,” *Expert Systems with Applications*, vol. 36, no. 3, pp. 5549–5556, 2009.
- [43] T. E. McKee and T. Lensberg, “Genetic programming and rough sets: a hybrid approach to bankruptcy classification,” *European Journal of Operational Research*, vol. 138, no. 2, pp. 436–451, 2002.
- [44] H. I. Elshazly, N. I. Ghali, A. M. E. Korany, and A. E. Hassanien, “Rough sets and genetic algorithms: a hybrid approach to breast cancer classification,” in *Proceedings of the World Congress on Information and Communication Technologies (WICT’12)*, pp. 260–265, IEEE, Trivandrum, India, November 2012.
- [45] H. Shimodaira, “A new genetic algorithm using large mutation rates and population-elitist selection (GALME),” in *Proceedings Eighth IEEE International Conference on Tools with Artificial Intelligence (ICTAI ’96)*, 1996.
- [46] S. Gauthier, B. Reisberg, M. Zaudig et al., “Mild cognitive impairment,” *The Lancet*, vol. 367, no. 9518, pp. 1262–1270, 2006.
- [47] D. W. Molloy and T. I. M. Standish, “A guide to the standardized Mini-Mental State Examination,” *International Psychogeriatrics*, vol. 9, no. 1, pp. 87–94, 1997.
- [48] A. Aggarwal and E. Kean, “Comparison of the folstein mini mental state examination (MMSE) to the montreal cognitive assessment (MoCA) as a cognitive screening tool in an inpatient rehabilitation setting gradient plasticity,” *Neuroscience & Medicine*, vol. 1, no. 2, 2010.

- [49] D. Mungas, "In-office mental status testing: a practical guide," *Geriatrics*, vol. 46, no. 7, pp. 54–66, 1991.
- [50] T. Smith, N. Gildeh, and C. Holmes, "The Montreal cognitive assessment: validity and utility in a memory clinic setting," *Canadian Journal of Psychiatry*, vol. 52, no. 5, pp. 329–332, 2007.
- [51] I. de Koning, F. van Kooten, and P. J. Koudstaal, "Value of screening instruments in the diagnosis of post-stroke dementia," *Pathophysiology of Haemostasis and Thrombosis*, vol. 28, no. 3-4, pp. 158–166, 1998.
- [52] P. Robert, S. Ferris, S. Gauthier, R. Ihl, B. Winblad, and F. Tennigkeit, "Review of Alzheimer's disease scales: is there a need for a new multi-domain scale for therapy evaluation in medical practice?" *Alzheimer's Research and Therapy*, vol. 2, no. 4, article 24, 2010.

Review Article

MRI Segmentation of the Human Brain: Challenges, Methods, and Applications

Ivana Despotović, Bart Goossens, and Wilfried Philips

Department of Telecommunications and Information Processing TELIN-IPI-iMinds, Ghent University, St-Pietersnieuwstraat 41, 9000 Ghent, Belgium

Correspondence should be addressed to Ivana Despotović; ivek@telin.ugent.be

Received 27 June 2014; Revised 11 September 2014; Accepted 1 October 2014

Academic Editor: Rafael M. Luque-Baena

Copyright © 2015 Ivana Despotović 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.

Image segmentation is one of the most important tasks in medical image analysis and is often the first and the most critical step in many clinical applications. In brain MRI analysis, image segmentation is commonly used for measuring and visualizing the brain's anatomical structures, for analyzing brain changes, for delineating pathological regions, and for surgical planning and image-guided interventions. In the last few decades, various segmentation techniques of different accuracy and degree of complexity have been developed and reported in the literature. In this paper we review the most popular methods commonly used for brain MRI segmentation. We highlight differences between them and discuss their capabilities, advantages, and limitations. To address the complexity and challenges of the brain MRI segmentation problem, we first introduce the basic concepts of image segmentation. Then, we explain different MRI preprocessing steps including image registration, bias field correction, and removal of nonbrain tissue. Finally, after reviewing different brain MRI segmentation methods, we discuss the validation problem in brain MRI segmentation.

1. Introduction

Over the last few decades, the rapid development of noninvasive brain imaging technologies has opened new horizons in analysing and studying the brain anatomy and function. Enormous progress in accessing brain injury and exploring brain anatomy has been made using magnetic resonance imaging (MRI). The advances in brain MR imaging have also provided large amount of data with an increasingly high level of quality. The analysis of these large and complex MRI datasets has become a tedious and complex task for clinicians, who have to manually extract important information. This manual analysis is often time-consuming and prone to errors due to various inter- or intraoperator variability studies. These difficulties in brain MRI data analysis required inventions in computerized methods to improve disease diagnosis and testing. Nowadays, computerized methods for MR image segmentation, registration, and visualization have been extensively used to assist doctors in qualitative diagnosis.

Brain MRI segmentation is an essential task in many clinical applications because it influences the outcome of the entire analysis. This is because different processing steps rely on accurate segmentation of anatomical regions. For example, MRI segmentation is commonly used for measuring and visualizing different brain structures, for delineating lesions, for analysing brain development, and for image-guided interventions and surgical planning. This diversity of image processing applications has led to development of various segmentation techniques of different accuracy and degree of complexity.

In this paper we review the most popular methods commonly used for brain MRI segmentation. We highlight differences between them and discuss their capabilities, advantages, and limitations. To introduce the reader to the complexity of the brain MRI segmentation problem and address its challenges, we first introduce the basic concepts of image segmentation. This includes defining 2D and 3D images, describing an image segmentation problem and image features, and introducing MRI intensity

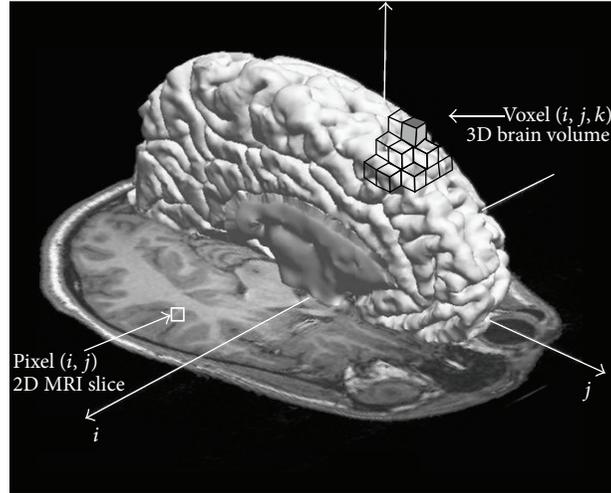


FIGURE 1: Illustration of image elements in the MRI of the brain. An image pixel (i, j) is represented with the square in the 2D MRI slice and an image voxel (x, y, z) is represented as the cube in 3D space.

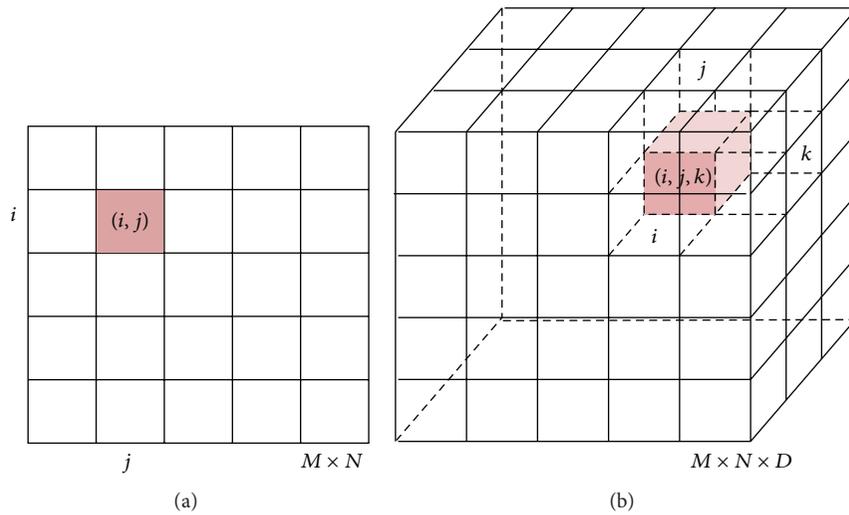


FIGURE 2: Illustration of image elements in 2D and 3D space. (a) In 2D space image elements (pixels) are represented with lattice nodes depicted as a square. (b) In 3D space image elements (voxels) are represented with lattice nodes depicted as a cube.

distributions of the brain tissue. Then, we explain different MRI preprocessing steps including image registration, bias field correction, and removal of nonbrain tissue. Finally, after reviewing different brain MRI segmentation methods, we discuss the validation problem in brain MRI segmentation.

2. Basic Concepts

2.1. 2D and 3D Images. An image can be defined as a function $I(i, j)$ in 2D space or $I(i, j, k)$ in 3D space, where $i = 0, \dots, M - 1$, $j = 1, \dots, N - 1$, and $k = 0, \dots, D - 1$ denote spatial coordinates. The values (or amplitudes) of the functions $I(i, j)$ and $I(i, j, k)$ are intensity values and are typically represented by a gray value $\{0, \dots, 255\}$ in MRI of the

brain; see Figure 1. Every image consists of a finite set of image elements called pixels in 2D space or voxels in 3D space. Each image element is uniquely specified by its intensity value and its coordinates (i, j) for pixels and (i, j, k) for voxels, where i is the image row number, j is the image column number, and k is the slice number in a volumetric stack; see Figure 2.

To each image element is assigned a single value based on the average magnetic resonance characteristics present in the tissue corresponding to that element. The size of the element determines the spatial resolution, or the fineness of detail that can be distinguished in an image. Voxel/pixel sizes vary depending on imaging parameters, magnet strength, the time allowed for acquisition, and other factors, but often in standard MRI studies voxel sizes are on the order of 1-2 mm. Greater spatial resolution can be obtained with a longer

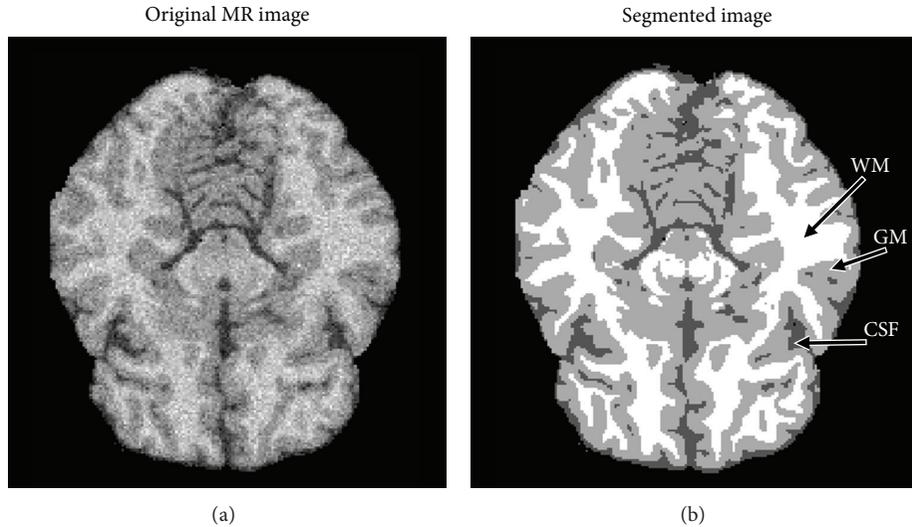


FIGURE 3: An example of the brain MRI segmentation with an original MR image (a) and segmented image with three labels: WM, GM, and CSF (b).

scanning time, but this must be weighed against patient discomfort. In adult brain MRI studies image acquisition time is around 20 min, while in pediatric MRI studies image acquisition time is limited to between 5 and 15 min.

2.2. Image Segmentation. The goal of image segmentation is to divide an image into a set of semantically meaningful, homogeneous, and nonoverlapping regions of similar attributes such as intensity, depth, color, or texture. The segmentation result is either an image of labels identifying each homogeneous region or a set of contours which describe the region boundaries.

Fundamental components of structural brain MRI analysis include the classification of MRI data into specific tissue types and the identification and description of specific anatomical structures. Classification means to assign to each element in the image a tissue class, where the classes are defined in advance. The problems of segmentation and classification are interlinked because segmentation implies a classification, while a classifier implicitly segments an image. In the case of brain MRI, image elements are typically classified into three main tissue types: white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF); see Figure 3. The segmentation results are further used in different applications such as for analyzing anatomical structures, for studying pathological regions, for surgical planning, and for visualization.

Image segmentation can be performed on 2D images, sequences of 2D images, or 3D volumetric imagery. Most of the image segmentation research has focused on 2D images. If the data is defined in 3D space (e.g., obtained from a series of MRI images), then typically each image “slice” is segmented individually in a “slice-by-slice” manner. This type of segmenting 3D image volumes often requires a postprocessing step to connect segmented 2D slices into a 3D volume or a continuous surface. Furthermore, the resulting segmentation can contain inconsistencies and nonsmooth

surface due to omitting important anatomical information in 3D space. Therefore, the development of 3D segmentation algorithms is desired for more accurate segmentation of volumetric imagery. The main difference between 2D and 3D image segmentation is in the processing elements, pixels/voxels, respectively, and their 2D or 3D neighborhoods (see Section 2.3) over which image features are calculated (see Section 2.4). In practice, 2D image segmentation methods can be extended to 3D space, but often with the cost of an increased complexity of the method and slower computational time.

2.3. Modeling the Spatial Context. The use of spatial context or neighborhood information is of great importance in brain MRI segmentation. Unless the image is simply random noise, the intensity of an image pixel/voxel is highly statistically dependent on the gray intensities of its neighbors (surrounding pixels/voxels). Markov random field (MRF) theory provides a basis for modeling local properties of an image, where the global image properties follow the local interactions. MRF models have been successfully integrated in various brain MRI segmentation methods to decrease misclassification errors due to image noise [1–3].

First, let us introduce some notations. As has been described in Section 2.1, every pixel (or voxel) in an image can be represented with one node in the lattice \mathcal{P} . Let x_i represent an intensity value of a single pixel (or voxel) with a position i in an image $\vec{x} = (x_1, \dots, x_m)$ defined over a finite lattice \mathcal{P} , where m is the total number of image elements ($m = MN$ for a 2D image and $m = MND$ for a 3D image). Let $\mathcal{N} = \{\mathcal{N}_i \mid \forall i \in \mathcal{P}\}$ denote a neighborhood system for a lattice \mathcal{P} , where \mathcal{N}_i represent a small neighborhood around i , not including x_i .

The nodes (pixels/voxels) in a lattice \mathcal{P} are related to one another via neighborhood system \mathcal{N} that can be defined as

$$\mathcal{N} = \{\mathcal{N}_i \mid \forall i \in \mathcal{P}\}. \quad (1)$$

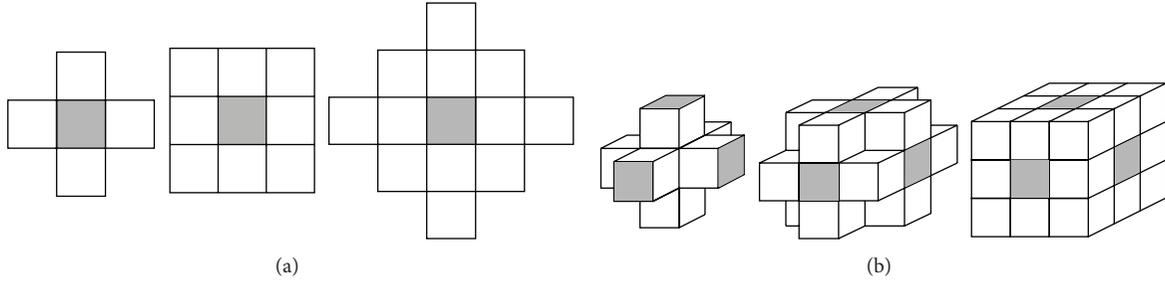


FIGURE 4: (a) 2D and (b) 3D neighborhood configuration for the first, second, and third order, respectively.

The neighboring relationship has the following properties:

- (i) a node i does not belong to its own neighborhood: $i \notin \mathcal{N}_i$;
- (ii) the neighboring relationship is mutual:

$$i \in \mathcal{N}_{i'} \iff i' \in \mathcal{N}_i. \quad (2)$$

The set of neighbors of i can be defined as the set of surrounding nodes within a radius of \sqrt{r} from the center i :

$$\mathcal{N}_i = \{i' \in \mathcal{P} \mid [\text{dist}(\text{pixel}_i, \text{pixel}_{i'})]^2 \leq r, i' \neq i\}, \quad (3)$$

where $\text{dist}(a, b)$ is the Euclidean distance between neighboring pixels a and b and $r \in \mathbb{Z} : r \geq 0$ is an integer number.

The first and the second order neighborhoods are the most commonly used neighborhoods in image segmentation. The first order neighborhood consists of 4 nearest nodes in a 2D image and 6 nearest nodes in a 3D image, while the second order neighborhood consists of 8 nearest nodes in a 2D image and 18 nearest nodes in a 3D image; see Figure 4.

Markov random field model can be represented with a graph $\mathcal{G} \triangleq (\mathcal{P}, \mathcal{N})$, where \mathcal{P} represents the nodes and \mathcal{N} determines the links (also called edges) that connect the nodes according to the neighborhood relationship. Such graph structure corresponds to an image, where nodes correspond to pixels (or voxels) and the links connecting the nodes represent the contextual dependency between pixels (or voxels). More reading about MRF can be found in [4].

2.4. Image Features. Image features represent distinctive characteristics of an object or an image structure to be segmented. Features rely on numerical measurements, including quantitative visual appearance and shape descriptors, that can help to discriminate between the structures of interest and their background. The outcome of image segmentation highly depends on appropriate feature selection (choosing the most relevant features) and accurate feature extraction.

Typically, statistical approach is used for feature extraction and classification in MRI, where pattern/texture is defined by a set of statistically extracted features represented as a vector in multidimensional feature space. The statistical features are based on first and second order statistics of gray level intensities in an image. First order features are derived from the image grey value histogram and include the intensity, mean, median, and standard deviation of the pixel values.

Since these features do not incorporate any information on the spatial distribution of the pixel values, they are often used in combination with second order features. Second order descriptors are used to describe image texture and are typically computed using gray level cooccurrence matrix [5]. First and second order features are often called appearance features in the literature. This is because the visual appearance of an object of interest is typically associated with its pixel or voxel intensities (gray values in brain MRI) and spatial interaction between intensities (intensity cooccurrence) in an image.

Image segmentation based on individual pixel/voxel intensities (first order features) is feasible only when intensities of an object of interest and its background differ to a large extent. Then, the complete object or the majority of its pixels/voxels can be separated from the background by simply comparing the intensity values to the threshold (the intensity value that clearly separates the object from the background). The threshold is derived from the overall intensity distribution of the image. In the presence of image noise and other imaging artifacts, first order features are not sufficient for accurate brain MRI segmentation. In this case more powerful second order discriminative features have to be used that include spatial interaction between intensities. For instance, the appearance of tumour lesion in brain MRI can be associated with spatial patterns of local pixel/voxel intensity variations or empirical probability distributions of intensity cooccurrences. In the spatial interaction models each intensity depends on a subset of the neighboring intensities; see Figures 5 and 4. The most popular models that can capture local spatial interactions between pixels/voxels intensities are MRF models [4].

Additionally, image segmentation performance can be also improved by incorporating probabilistic prior shape models, which have been extensively used in medical image segmentation [6–10]. The probabilistic prior shape models specify an average shape and variation of an object of interest and are typically estimated from a population of coaligned images of the object (training data sets) [11].

One of the most popular features for image segmentation is edges. Edges refer to boundaries of an object surface where the intensities change sharply [12]. Such changes are typically detected by thresholding the first and second order spatial derivatives of the intensities (the intensity gradient and Laplacian). However, edges detected in this way are sensitive to image noise [13] and often require image smoothing as a preprocessing step [14, 15].

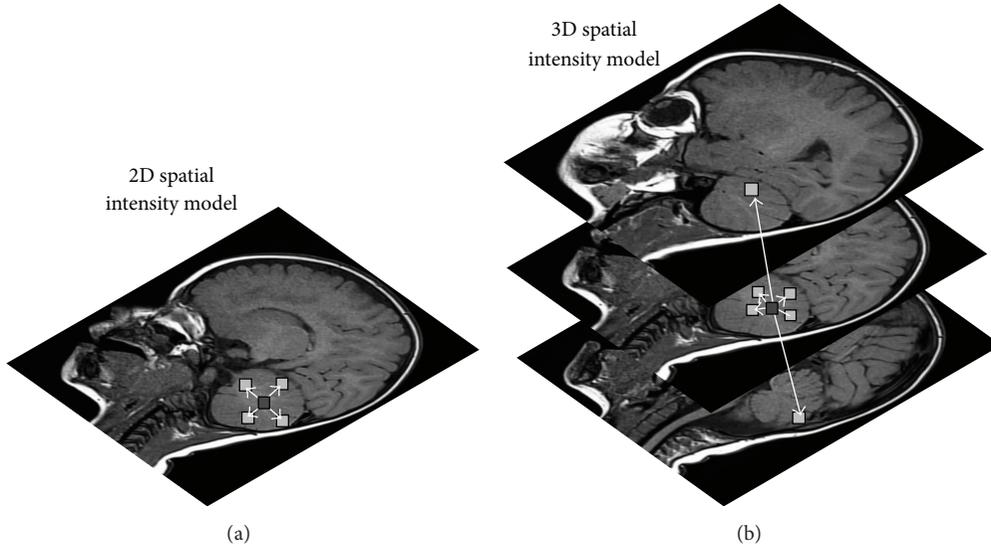


FIGURE 5: Illustration of 2D (a) and 3D (b) spatial interactions between neighboring pixel/voxel intensities.

Another more robust method for edge detected is the phase congruency method [16, 17], which is a frequency-based method for feature detection. This feature detection method, using local phase and energy, is based on a plausible model of how mammals detect edges suggested by Morrone and Owens [18] and successfully explains the psychophysical effect of human feature perception. Instead of searching the pixels/voxels in the image with sharp intensity changes, features such as step edges, lines, and corners are detected at points where the Fourier components of the image are maximally in phase.

2.5. Intensity Distribution in Brain MRI. The intensity of brain tissue is one of the most important features for brain MRI segmentation. However, when intensity values are corrupted by MRI artifacts such as image noise, partial volume effect (PVE), and bias field effect, intensity-based segmentation algorithms will lead to wrong results. Thus, to obtain relevant and accurate segmentation results, very often several preprocessing steps are necessary to prepare MRI data. For instance, it is necessary to remove background voxels, extract brain tissue, perform image registration for multimodal segmentation, and remove the bias field effect; see Figure 6.

In the case when the bias field, nonbrain structures (e.g., the skull and the scalp) and background voxels are removed, the histogram of the adult brain MRI has three main peaks corresponding to the three main tissue classes; see Figure 8(a). In the healthy adult brain, the intensity variation within tissue is small and the intensities inside the brain can be considered to be a piecewise constant intensity function, corrupted by noise and PVE. The PVE describes the loss of small tissue regions due to the limited resolution of the MRI scanner. It means that one pixel/voxel lies in the interface between two (or more) classes and is a mix of different tissues. This problem is even more critical in imaging of the small

neonatal brain. The correction of PVE will be addressed in Section 4.6.

It has been shown that the noise in the magnitude images is governed by a Rician distribution, based on the assumption that the noise on the real and imaginary channels is Gaussian [19]. The probability density function for a Rician distribution is defined as

$$f_{\text{Rice}}(x) = \frac{x}{\sigma^2} \exp\left(-\frac{(x^2 + \nu^2)}{2\sigma^2}\right) I_0\left(\frac{x\nu}{\sigma^2}\right), \quad (4)$$

where x is the measured pixel/voxel intensity, ν is the image pixel/voxel intensity in the absence of noise, σ is the standard deviation of the Gaussian noise in the real and the imaginary images, and I_0 is the zero-order modified Bessel function of the first kind. The Rician probability density function (PDF) is plotted in Figure 7(a) for several values of the signal-to-noise ratio (SNR), where the SNR is defined as ν/σ (the power ratio between the signal and the background noise).

A special case of the Rician distribution is in image regions where only noise is present and $\text{SNR} = \nu/\sigma = 0$ (e.g., in the dark background areas of an MRI where no NMR signal is present). This special case of the Rician distribution where $\nu = 0$ and $I_0 = 1$ is also known as the Rayleigh distribution:

$$f_{\text{Rayleigh}}(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (5)$$

In the image regions where the NMR signal is present and $\text{SNR} \geq 3$, the noise distribution approximates a Gaussian distribution; see Figure 7. Thus, the problem of Rician noise in the brain MRI is often simplified in practice by assuming the Gaussian distribution for the noise:

$$f_{\text{Gauss}}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad (6)$$

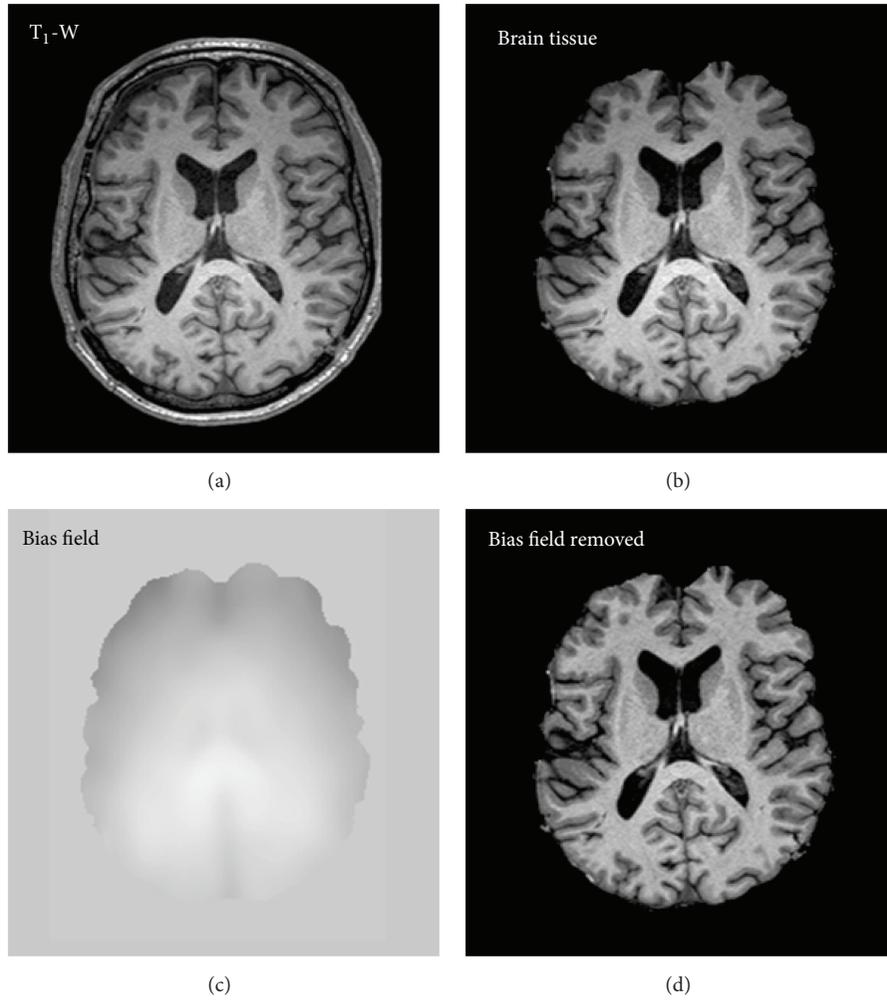


FIGURE 6: Preprocessing steps: (a) the original T₁-W MR image of the adult brain; (b) the brain tissue image after removing non-brain structures; (c) the bias field; (d) the brain tissue image after bias field correction.

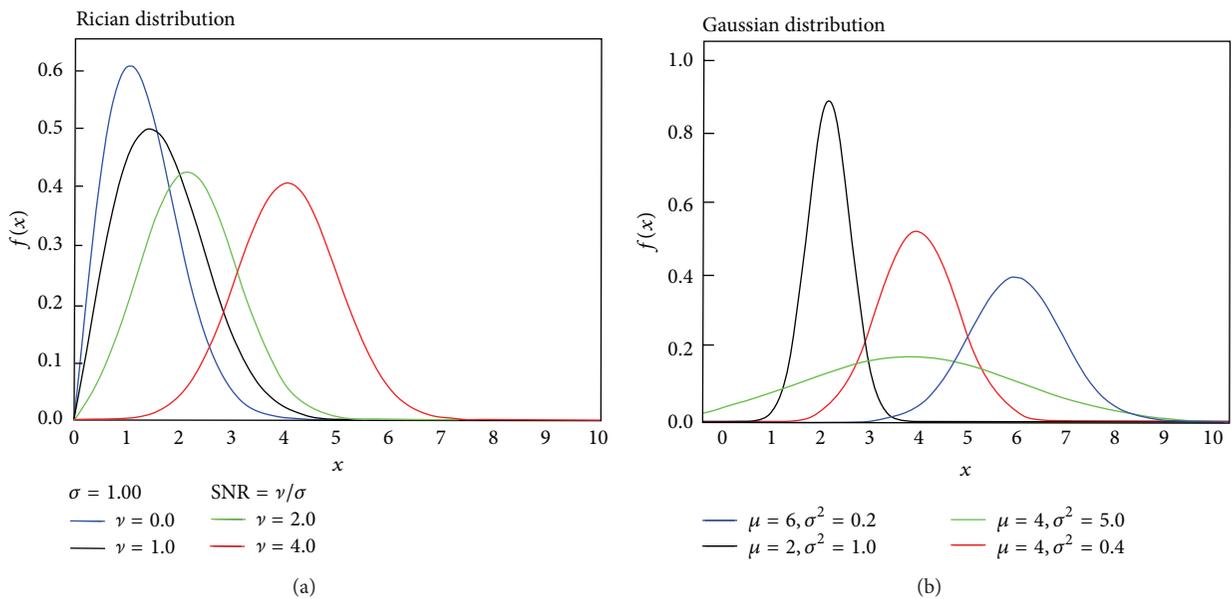


FIGURE 7: (a) The PDF for the Rician distribution. (b) The PDF for the Gaussian distribution.

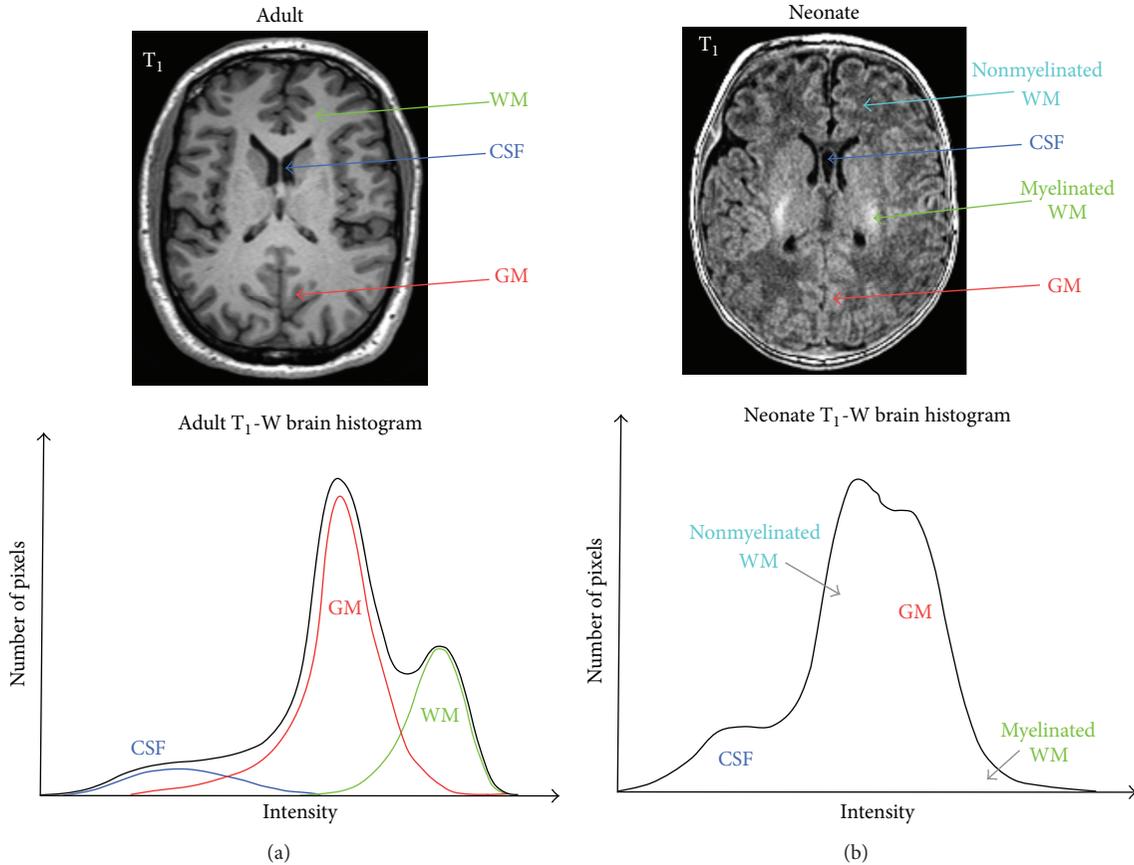


FIGURE 8: (a) Histogram of a bias-corrected T_1 -W MRI of an adult brain. Histograms of the tissue classes are based on manual segmentation and distributions slightly differ from the Gaussian distribution due to partial volume effect. (b) Histogram of a 1.5 T T_1 -W MRI of a neonatal brain. The difference between the neonatal and the adult brain histogram is the existence of the myelinated and nonmyelinated WM in neonates, which are separated with GM intensities. Since nonmyelinated WM is more dominant than myelinated WM, T_1 -W MRI shows inverted WM/GM intensities in neonates in comparison to adults.

where x , σ , and μ are the intensity, the standard deviation, and the mean value, respectively. Due to this approximation, the histogram of a bias-corrected brain MRI in the presence of noise can be described with a Gaussian mixture model (GMM), where each tissue class (WM, GM, and CSF) is modeled by a Gaussian distribution. However, in the presence of partial volume effects the tissue intensity distributions slightly diverge from a Gaussian distribution, as can be seen from the histogram in Figure 8(a) where histograms of the tissue classes are based on manual segmentation. Recently, the α -stable distribution mixture model is also suggested as an alternative to the Gaussian mixture model to model the histogram of MRI data for more complex MRI segmentation [20]. Note that the α -stable distribution is a generalization of the Gaussian distribution.

The MRI intensity distribution of the neonatal brain is more complex because the intensity variability within tissue cannot be neglected due to the process of myelination. The histogram of 1.5 T T_1 -W MRI of the neonatal brain is shown in Figure 8(b). The difference between the neonatal and the adult brain histogram is the existence of the myelinated and nonmyelinated WM in neonates, which are separated with

GM intensities. Since nonmyelinated WM is more dominant than myelinated WM, T_1 -W MRI shows inverted WM/GM intensities in neonates in comparison to adults.

2.5.1. T_1 -W and T_2 -W Intensity Distribution. It can be noted from the 1D histogram of the bias-corrected T_1 -W MRI of an adult brain in Figure 8(a) that there is an overlap between different tissue classes. Also, it can be seen that an overlap between WM and GM tissue is higher than between GM and CSF. This overlap between the class distributions can cause ambiguities in the decision boundaries when intensity-based segmentation methods are used [21]. However, many researchers showed that adding additional MRI sequences with different contrast properties (e.g., T_2 -W MRI, Proton Density MRI) can improve intensity-based segmentation and help separate the class distributions [22–24]; see Figure 13.

3. MRI Preprocessing

After MRI acquisition several preprocessing steps are necessary to prepare MR images for segmentation; see Figure 6. The most important steps include MRI bias field correction,

image registration (in the case of multimodal image analysis), and removal of nonbrain tissue (also called a brain extraction).

3.1. Bias Field Correction. The bias field, also called the intensity inhomogeneity, is a low-frequency spatially varying MRI artifact causing a smooth signal intensity variation within tissue of the same physical properties; see Figure 6. The bias field arises from spatial inhomogeneity of the magnetic field, variations in the sensitivity of the reception coil, and the interaction between the magnetic field and the human body [25, 26]. The bias field is dependent on the strength of the magnetic field. When MR images are scanned at 0.5 T, the bias field is almost invisible and can be neglected. However, when MR images are acquired with modern high-field MR scanners with a magnetic field strength of 1.5 T, 3 T, or higher, the bias field is strong enough to cause problems and considerably affect MRI analysis. In practice, trained medical experts can make visual MRI analysis to certain levels of intensity inhomogeneity (10%–30%) [26]. In contrast, the performance of automatic MRI analysis and intensity-based segmentation methods decreases greatly in the presence of the bias field; see Figure 9. This is because most of the segmentation algorithms assume intensity homogeneity within each class. Therefore, the correction of the bias field is an important step for the efficient segmentation and registration of brain MRI.

The bias field is typically modeled as low-frequency multiplicative field [26, 27]. Suppose that we place all image elements $I(i, j, k)$, $i = 0, \dots, M - 1$, $j = 0, \dots, N - 1$, and $k = 0, \dots, D - 1$, into an $m \times 1$ column vector $\vec{x} = (x_1, \dots, x_m)$, where x_i , $i = 1, \dots, m$, represents the observed intensity of the i th voxel and $m = MND$ is the total number of image elements. The degradation effect of each image voxel x_i can be expressed as

$$x_i = x'_i b_i, \quad i = 1, \dots, MND, \quad (7)$$

where x'_i is an ideal intensity of the i th voxel and b_i is an unknown smoothly varying bias field. The problem of eliminating the bias field is the task of estimating b_i .

If the intensities of MRI are logarithmically transformed, the multiplicative bias field becomes an additive bias field as follows:

$$\log(x_i) = \log(x'_i) + \log(b_i). \quad (8)$$

This simplified multiplicative model is used in most state-of-the-art bias correction methods to represent the bias field [26–28]. However, in reality there are certain limitations to the correctness of this model. Even though the model is consistent with the variations arising from the sensitivity of the receiver coil, the relationship between the measured and true intensities in MRI is more complicated. This is due to nonuniformity of the induced currents and spatial inhomogeneity of the excitation field, which depends on the geometry and electromagnetic properties of the subject as well as the coil polarization and pulse sequence [26]. In spite of these difficulties, the multiplicative low-frequency model is successfully used in practice to model the intensity inhomogeneity in brain MRI.

In the literature, various methods have been proposed to correct the bias field in MRI. One of the earliest methods proposed to correct the bias field is based on the manual labeling of the brain tissue voxels, which are then used to reconstruct the bias field in form of a parametric surface. The main disadvantage of this surface fitting method is the need for manual interaction. The bias field can be also estimated and corrected by using low-pass filtering [29], but this approach can introduce additional artifacts in the image because it also removes the low-frequency component of the true image data. Both the surface fitting method and the low-pass method can be improved and made fully automatic if they are coupled with automatic segmentation of the brain [30, 31]. Other approaches for the bias field correction include minimizing the image entropy [32], fitting the histogram of the local neighbourhood to global histogram of the image [28], maximizing the high-frequency content of the image [26], and using a registered template image [27].

The template is an image/volume which encodes the average probability of finding different kinds of tissues at each spatial location. The anatomical template is obtained by normalizing, aligning, and averaging of anatomical images from several different subjects. All the images are normalized in a standard stereotaxic space such as the Montreal Neurological Institute (MNI space) [33]. MNI is widely used to provide a common reference for the 3D localization of functional activation foci and anatomical structures, enabling the comparison of results obtained across different studies. The standard probabilistic atlas of the human brain consists of a template and three tissue probability maps for WM, GM, and CSF [33]. The tissue probability maps are obtained by normalizing and averaging a number of segmented subjects. The probabilistic atlas then describes the anatomical variability of the brain.

Image registration is a necessary step for the inclusion of probabilistic atlases as a prior knowledge of the brain anatomy into the segmentation method. A probabilistic atlas is often used to initialize and constrain the segmentation process. The prior knowledge of the brain anatomical structures can increase the robustness and accuracy of a segmentation method; see Section 4.3.

3.2. Image Registration. Image registration is the process of overlaying (spatially aligning) two or more images of the same content taken at different times, from different viewpoints, and/or by different sensors. Registration is required in medical image analysis for obtaining more complete information about the patient's health when using multimodal images (e.g., MRI, CT, PET, and SPECT) and for treatment verification by comparison of pre- and postintervention images. In medical image registration the term coregistration is used for intrasubject registration (the alignment multimodal images of the same subject), realignment is used for motion correction within the same subject, and normalization is used for intersubject registration when several population groups are studied.

Image registration involves finding the transformation between images so that corresponding image features are

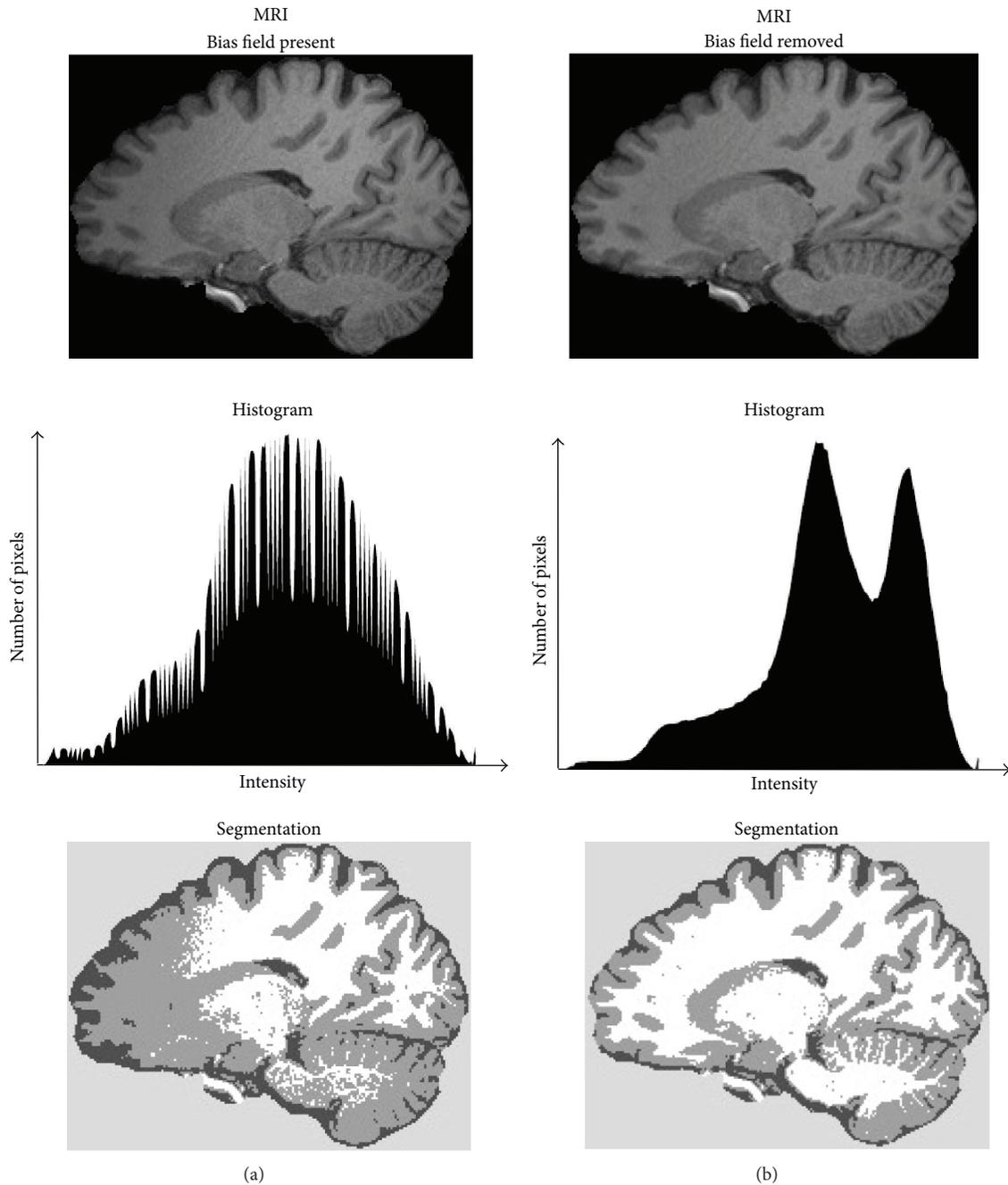


FIGURE 9: Influence of the bias field on brain MRI segmentation. (a) An example of the sagittal brain MRI slice with bias field is shown in the top of the figure. The image histogram is shown in the middle and the three-label segmentation in the bottom. (b) The bias-corrected MRI slice is shown in the top, the corresponding histogram in the middle, and three-label segmentation in the bottom.

spatially aligned. The spatial alignment is typically initialized using rigid or affine transformation [34]. A rigid transformation is a 6-parameter transformation composed of translation and rotation. If scaling and skewing are allowed, we obtain a 12-parameter affine transformation. A rigid registration is sufficient for intrasubject registration if the object of interest does not deform. This is a reasonable assumption for images of the brain if these are acquired at the same stage of brain development. However, if the task is to match

images belonging to either different subjects (intersubject registration) or the same subject at different stages of brain development (e.g., growth in children, changes related to ageing, or atrophy due to disease), a nonrigid registration of the images is required to obtain satisfactory results. The nonrigid registration algorithms are typically based on either physical models for transformation such as elastic [35] or fluid deformation models [36] or a linear combination of smooth basis functions [37] or free-form deformations [38].

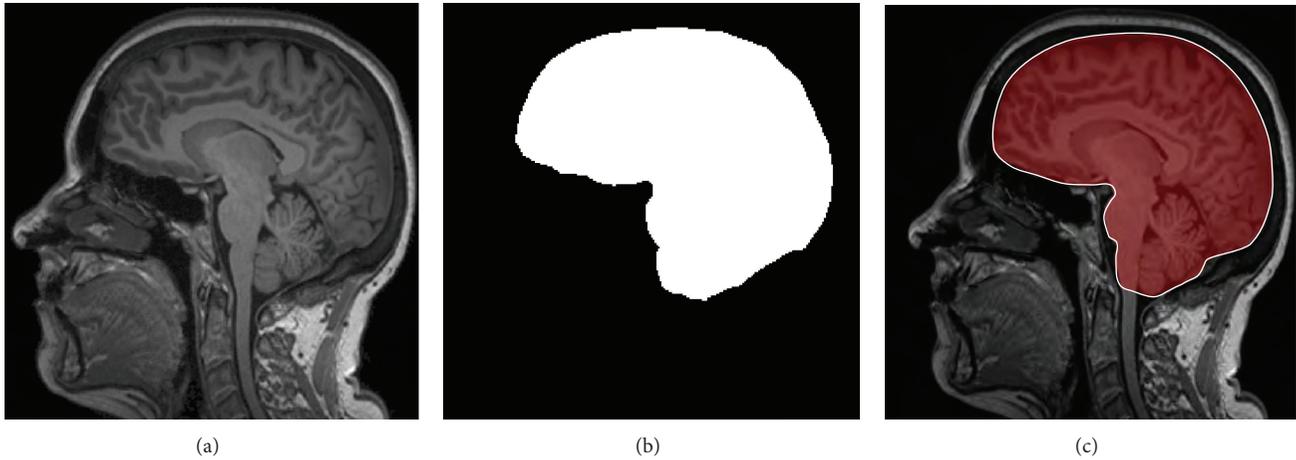


FIGURE 10: Result of brain extraction on a T_1 MR image in an axial plane. (a) shows the original T_1 -W MRI. (b) depicts the estimated brain mask. (c) presents an overlap of the brain mask and original MR image.

However, the problems in intersubject brain MRI registration will arise when brains include lesions or diseases, because it is not possible to match the same structures between healthy and diseased brains. A general review of registration techniques can be found in [39–41].

3.3. Removal of Nonbrain Tissue. Nonbrain tissues such as fat, skull, or neck have intensities overlapping with intensities of brain tissues. Therefore, the brain has to be extracted before brain segmentation methods can be used. This step classifies voxels as brain or nonbrain. The result can be either a new image with just brain voxels or a binary mask, which has a value of 1 for brain voxels and 0 for the rest of tissues. In general, the brain voxels comprise GM, WM, and CSF of the cerebral cortex and subcortical structures, including the brain stem and cerebellum. The scalp, dura matter, fat, skin, muscles, eyes, and bones are always classified as nonbrain voxels.

The common method for brain extraction is to use prior information of the brain anatomy. A deformable template can be registered with an image and nonbrain tissue is then removed by transferring the brain mask from the template [42]. However, brain extraction using a probabilistic atlas is usually not very accurate and can cause misclassification around the brain boundary. An alternative method for extracting the brain is the brain extraction tool (BET) [43, 44], which is part of the publicly available software package FSL. This method finds the center of gravity of the brain and then inflates a sphere until the brain boundary is found. It has been proven to work in practice on good quality T_1 -W and T_2 -W images of the adult brain. An example of the brain extraction is shown in Figure 10.

4. MRI Segmentation Methods

In general, MRI segmentation is not a trivial task, because acquired MR images are imperfect and are often corrupted by noise and other image artifacts. The diversity of image

processing applications has led to development of various techniques for image segmentation [45–54]. This is because there is no single method that can be suitable for all images, nor are all methods equally good for a particular type of image. For example, some of the methods use only the gray level histogram, while some integrate spatial image information to be robust for noisy environments. Some methods use probabilistic or fuzzy set theoretic approaches, while some additionally integrate prior knowledge (specific image formation model, e.g., MRI brain atlas) to further improve segmentation performance.

However, most of the segmentation methods developed for one class of images can be easily applied/extended to another class of images. For example, the theory of graph cuts, although firstly developed for binary images [55], can be modified and used for MRI segmentation of the brain tissue. Also, unsupervised fuzzy clustering [45, 56, 57] has been successfully applied in different areas such as remote sensing, geology, and medical, biological, and molecular imaging.

The segmentation methods, with application to brain MRI, may be grouped as follows:

- (i) manual segmentation;
- (ii) intensity-based methods (including thresholding, region growing, classification, and clustering);
- (iii) atlas-based methods;
- (iv) surface-based methods (including active contours and surfaces, and multiphase active contours);
- (v) hybrid segmentation methods.

4.1. Manual Segmentation. Manual segmentation refers to the process where a human operator (e.g., expert physician) segments and labels an image by hand. This segmentation is typically done in a “slice-by-slice” manner for 3D volumetric imagery. The manual method is believed to be the most accurate because of the difficulty to accurately and reliably delineate structures in medical images. The segmentation difficulties are related to image quality and artifacts.

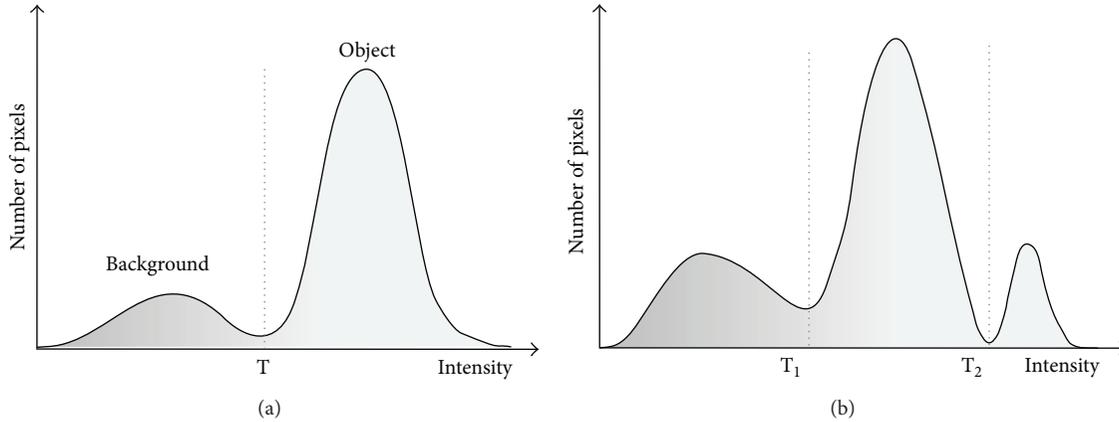


FIGURE 11: (a) Gray level histogram that can be partitioned by a single threshold. (b) Gray level histogram that can be partitioned by multiple thresholds.

Given the improvements achieved over the past years by imaging tools (e.g., MR scanners resolve images at millimetric resolution), the manual segmentation has become an intensive and time-consuming task. A trained operator typically has to go through around eighty 512×512 images, slice by slice, to extract the contours of the target structures. This manual segmentation is not only tedious but also particularly prone to errors, as assessed by various intra- or interoperator variability studies [58, 59]. Also, manual segmentation results are often difficult and even impossible to reproduce, because even experienced operators show significant variability with respect to their own previous delineation.

However, manual segmentation is still intensively used for defining a surrogate for true delineation (called “ground truth”) and quantitative evaluation of automated segmentation methods. Also, manual segmentation of different brain structures is a fundamental step in brain atlas formation and is used in atlas-based segmentation approaches [50, 51, 60].

For manual delineation, editing tools such as ITK-SNAP [61, 62] usually display 3D data in the form of 3 synchronized 2D orthogonal views (sagittal, coronal, and axial) onto which the operator draws the contour of the target structure. The output data therefore consists of a series of 2D contours from which a continuous 3D surface has to be extracted. This is a nontrivial postprocessing task and is prone to errors. For instance, due to interslice inconsistencies in segmentation, bumps in the reconstructed 3D surface are inevitable. More robust segmentation methods can usually be derived from true 3D structure models in that they can ensure globally smoother and more coherent surfaces across slices.

4.2. Intensity-Based Methods. Intensity-based segmentation methods classify individual pixels/voxels based on their intensity. In the case of the brain MRI, three main tissue classes, WM, GM, and CSF, can be distinguished based on intensity; see Figure 3. A more detailed classification is not possible because the intensity profiles of more detailed brain structures overlap. Even separation of the three main tissue classes based on intensity itself requires incorporating tools for dealing with artifacts in MRI, such as intensity

inhomogeneity, noise, and partial volume, as well as overlap in intensities of brain and nonbrain tissue (e.g., the scalp has the same intensities as brain tissues).

Several intensity-based techniques are available for tissue classification. The most common method is the use of intensity histogram of all of the voxels and fitting Gaussian functions to the distribution. The probability of a given intensity corresponding to a given type of tissue can thus be inferred and voxels are assigned to tissue types accordingly. Additionally incorporating neighbourhood information helps to give preference to spatially homogeneous regions in the resulting segmentation. This can significantly decrease misclassification due to random noise in the image [1]. Additionally, probabilistic atlases can be included in the classification to inform whether a given location in the brain is likely to contain WM, GM, or CSF voxels [50].

4.2.1. Thresholding. Thresholding is the simplest image segmentation method. A thresholding procedure uses the intensity histogram and attempts to determine intensity values, called thresholds τ , which separates the desired classes. The segmentation is then achieved by grouping all pixels between thresholds into one class; see Figure 11. The thresholding methods have many variations: global (single threshold) or local threshold (depending on the position in the image), multithresholding, adaptive thresholding, and so forth. In the case of a single global threshold, segmentation of an image $I(i, j)$ is defined as

$$I'(i, j) = \begin{cases} 1, & \text{if } I(i, j) > \tau, \\ 0, & \text{if } I(i, j) \leq \tau, \end{cases} \quad (9)$$

where $I'(i, j)$ is a segmented (thresholded) image, where pixels labeled with 1 correspond to object and pixels labeled with 0 correspond to background; see Figure 11(a).

Thresholding is fast and computationally efficient method but does not take into account the spatial characteristics of an image (neighborhood information). Thus thresholding is sensitive to noise and intensity inhomogeneities. In low-contrast images it tends to produce scattered groups of pixels

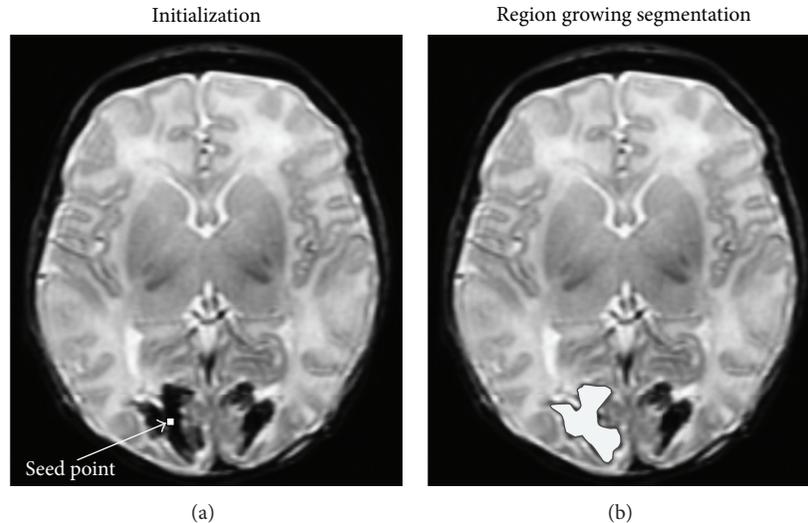


FIGURE 12: An example of region growing segmentation of a brain lesion. (a) In the initialization step, a seed point is manually selected in the lesion area. (b) The final segmentation result is a connected region and represents the lesion.

rather than connected regions and requires connectivity algorithms as a postprocessing step.

In general, threshold-based segmentation methods are not suitable for textured images. This is because the perceptual qualities of textured images are based on higher order interactions between image elements or objects in the scene. However, in brain MRI segmentation, thresholding can be used to separate background voxels from the brain tissue or to initialize the tissue classes in iterative segmentation methods such as fuzzy C -means clustering. A survey on thresholding techniques is provided in [63].

4.2.2. Region Growing. Region growing (also called region merging) is a technique for extracting a connected region of the image which consists of groups of pixels/voxels with similar intensities [64]. In its simplest form, region growing starts with a seed point (pixel/voxel) that belongs to the object of interest. The seed point can be manually selected by an operator or automatically initialised with a seed finding algorithm. Then, region growing examines all neighboring pixels/voxels and if their intensities are similar enough (satisfying a predefined uniformity or homogeneity criterion), they are added to the growing region. This procedure is repeated until no more pixels/voxels can be added to the region.

Region growing is suitable for segmentation of volumetric images which are composed of large connected homogeneous regions. Thus, it is successfully used in medical image analysis to segment different tissues, organs, or lesions from MR images. For example, it is used in brain MRI analysis for segmentation of brain vessels [65], brain tumour segmentation [66], or extraction of brain surface [67]. See an example of region growing segmentation in Figure 12.

The main disadvantage of the region growing method is its sensitivity to the initialization of seed point. By selecting a different seed point, the segmentation result can be completely different. If seed point and homogeneity criterion are not properly defined, the growing region can leak out

and merge with the regions that do not belong to the object of interest. Also, region growing is sensitive to noise and segmented regions in the presence of noise can become disconnected or have holes. On the other hand, separate regions can become connected in the presence of partial volume effects.

4.2.3. Classification Methods. Classification methods use data with known labels to partition image feature space. Image features are typically intensity values but can be also related to texture or other image properties. Classification methods can be both supervised and unsupervised. Supervised classification requires training images, which are manually segmented and then used as references for automatic segmentation of new images. Next to the manual interaction that is laborious and time-consuming, another disadvantage of supervised classification methods is that they generally do not take into account the neighborhood information and thus they are sensitive to noise. Also, the use of the same training set for a large number of images can lead to biased results, which do not take into account anatomical and physiological variability between different subjects.

One of the simplest classifiers is the nearest-neighbor classifier [68], where each pixel/voxel is classified in the same class as the training datum with the closest intensity. A generalization of this approach is the k -nearest-neighbor (k NN) classifier, where the pixel/voxel is classified according to the majority vote of the closest training data. The k NN classifier is considered a nonparametric classifier because it makes no underlying assumption about the statistical structure of the data. It is especially suitable if a large number of training data are available.

The k NN classification method was applied to brain MRI segmentation by Warfield et al. [69]. In addition to image intensities, Warfield used spatial localization of brain structures (classes) in form of a nonrigidly registered template as an additional feature to enhance the classification process.

The segmentation is then calculated in an iterative process by interleaving the segmentation refinement with updating the nonrigid alignment to the template. This procedure requires manual selection of a large number of training samples for each tissue class to train the k NN classifier. Due to the manual interaction in the training phase, the method is not fully automatic and the results depend on particular choice of the training set. Cocosco et al. [70] developed a method for the robust selection of training samples to make the k NN classification process fully automatic. This method is reported to deal well with anatomies which differ from the probabilistic atlas. However, it does not deal with the problem of natural intensity variation within each tissue class. Both methods require correction of the bias field as a preprocessing step.

One of the most commonly used parametric classifiers is the Bayesian classifier [30]. The Bayesian classifier models the probabilistic relationships between the attribute set and the class variables, which are then used for estimating the class probability of the unknown variable. This model involves Bayesian inference such as maximum a posteriori (MAP) estimation, where the goal is to estimate the label output image $\hat{\mathbf{x}}$ given the observed image \mathbf{y} by minimizing the posterior distribution $P(\mathbf{x} | \mathbf{y})$ of the possible labels \mathbf{x} :

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} P(\mathbf{x} | \mathbf{y}). \quad (10)$$

The Bayesian framework consists of three probability distributions: the *prior* distribution $P(\hat{\mathbf{x}})$, the *posterior* distribution $P(\mathbf{x} | \mathbf{y})$, and the *conditional* distribution $P(\mathbf{y} | \mathbf{x})$ (also called the *likelihood*). The prior distribution embodies the knowledge of likely configurations before an actual image is observed. The posterior distribution is derived after an observation has been made and the likelihood is defined as the probability of obtaining a particular observation given a set of model parameters.

The Bayes rule describes the relation between the posterior probability $P(\mathbf{x} | \mathbf{y})$, prior probability $P(\mathbf{x})$, and likelihood $P(\mathbf{y} | \mathbf{x})$ as follows:

$$P(\mathbf{x} | \mathbf{y}) = \frac{P(\mathbf{y} | \mathbf{x}) P(\mathbf{x})}{P(\mathbf{y})}. \quad (11)$$

Using definition (11), the MAP estimate can be written as

$$\begin{aligned} \hat{\mathbf{x}} &= \arg \max_{\mathbf{x}} \left(\frac{P(\mathbf{y} | \mathbf{x}) P(\mathbf{x})}{P(\mathbf{y})} \right) \\ &= \arg \max_{\mathbf{x}} (P(\mathbf{y} | \mathbf{x}) P(\mathbf{x})), \end{aligned} \quad (12)$$

where $P(\mathbf{y})$ can be omitted because it is a constant in the case when \mathbf{y} is known. Since in many cases the probability distributions have exponential functions, this computation can be simplified by using a logarithmic transform:

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} (\log P(\mathbf{y} | \mathbf{x}) + \log P(\mathbf{x})). \quad (13)$$

In the case of the brain MRI segmentation, often it is assumed that the pixel intensities are independent samples from a mixture of Gaussian probability distributions. Training data is collected by obtaining representative samples

from each component of the Gaussian mixture accordingly. Classification of new data is obtained by assigning each pixel to the class with the highest posterior probability.

Bayesian classifiers are used in the expectation-maximization (EM) segmentation methods which have been successfully implemented in several software packages used in the medical imaging community: SPM [2], FAST [3], FreeSurfer [21], and 3DSlicer [71]. All these methods implement segmentation and bias correction in the EM framework. They also include various additional improvements, such as nonrigid alignment of atlas [2], including neighbourhood information in the form of Markov random fields [3, 4] or using the α -stable distribution mixture model as a generalization of the GMM [20]. More details about the Bayes' theory can be found in [72].

4.2.4. Clustering Methods. Clustering methods are unsupervised segmentation methods that partition an image into clusters of pixels/voxels with similar intensities without using training images. In fact, clustering methods use the available image data to train themselves. The segmentation and training are done in parallel by iterating between two steps: data clustering and estimating the properties of each tissue class. The most commonly used clustering methods are the k -means clustering [73], the fuzzy C -means clustering [74, 75], and the expectation-maximisation (EM) method [1].

The k -means clustering method partitions the input data into k classes by iteratively computing a mean intensity for each class (also called centroid) and segmenting the image by classifying each pixel/voxel in the class with the closest centroid. The k -means clustering is also known as a hard classification method because it forces each pixel/voxel to belong exclusively to one class in each iteration. The fuzzy C -means clustering is soft classification method based on fuzzy set theory [76]. It is a generalization of the k -means clustering because it allows each pixel/voxel to belong to multiple classes according to a certain membership value.

The FCM clustering algorithm is based on minimizing the following objective function:

$$J_m = \sum_{i=1}^C \sum_{j=1}^N u_{ij}^m D_{ij}, \quad (14)$$

where N is the number of image elements that need to be partitioned into C clusters, u_{ij} is the membership function of the element \mathbf{x}_j (a feature vector at position j) belonging to the i th cluster, m is the weighting exponent that controls the fuzziness of the resulting partition (most often is set to $m = 2$, if $m = 1$ we have the k -means clustering), and D_{ij} is the similarity measure between \mathbf{x}_j and the i th cluster center \mathbf{v}_i . The most commonly used similarity measure is the squared Euclidean distance $D_{ij} = \|\mathbf{x}_j - \mathbf{v}_i\|^2$.

The objective function J_m (see (14)) is minimized under the following constraints: $u_{ij} \in [0, 1]$, $\sum_{i=1}^C u_{ij} = 1 \forall j$, and $0 < \sum_{j=1}^N u_{ij} < N \forall i$. Considering these constraints and calculating the first derivatives of J_m with respect to u_{ij} and \mathbf{v}_i

and setting them to zero result in the following two conditions for minimizing J_m :

$$u_{ij} = \left[\sum_{k=1}^C \left(\frac{D_{ij}}{D_{kj}} \right)^{1/(m-1)} \right]^{-1}, \quad (15)$$

$$\mathbf{v}_i = \frac{\sum_{j=1}^N u_{ij}^m \mathbf{x}_j}{\sum_{j=1}^N u_{ij}^m}.$$

The FCM algorithm iteratively optimizes J_m , by evaluating (15), until the following stop criterion is satisfied: $\max_{i \in [1, C]} \|\mathbf{v}_i^{(l)} - \mathbf{v}_i^{(l+1)}\|_{\infty} < \epsilon$, where l is the iteration index and $\|\cdot\|_{\infty}$ is the L_{∞} norm. Once a membership value u_{ij} for each class i is assigned to each pixel j , defuzzification of the fuzzy clusters $\{F_k\}_{k=1}^C$ into its crisp version $\{H_k\}_{k=1}^C$ is done by assigning the pixel to the class with the highest membership value as follows:

$$\max_{i \in [1, C]} (u_{ij}) = u_{kj} \implies \mathbf{x}_j \in H_k. \quad (16)$$

The EM method is an iterative method for finding maximum likelihood or MAP estimates of a statistical model. It has the same soft classification principle as FCM method but typically assumes that MRI intensities of different brain tissues can be represented with a Gaussian mixture model. Even though clustering methods do not require training images, they do require some initial parameters and the EM method has shown the highest sensitivity to initialization in comparison to fuzzy C -means and k -means methods [1].

In general, the EM segmentation framework can be described as follows.

EM Approach for Brain MRI Segmentation. Firstly, initialize the EM algorithm. In the case of brain MRI segmentation, the GMM is used to initially estimate model parameters. Then, iterate between expectation step (E-step) and maximization step (M-step) until convergence.

E-Step. Estimate the brain tissue segmentation given the current estimate of model parameters. This step can include the use of neighbourhood information (e.g., in the form of MRF modeling).

M-Step. Estimate the model parameters. This step can consist of a combination of the following steps.

- (1) Estimate the intensity distribution parameters for each tissue class.
- (2) Estimate the bias correction parameters.
- (3) Estimate the registration parameters for alignment of probabilistic atlas with the image.

As it is the case with classification methods, clustering methods initially do not incorporate spatial neighborhood information and thus they are sensitive to noise and intensity inhomogeneities. To improve the clustering performance for images corrupted by noise, many extensions of the clustering algorithms have been proposed [49, 57, 77–84]. The most

common approach is to include feature information (e.g., intensity values) of the neighboring pixels into the modified FCM objective function [77, 79] or into a similarity measure between cluster centers and image elements [80]. Ahmed et al. [77] modified the objective function of the standard FCM algorithm to allow the immediate neighbours of the pixel to influence its labeling. Chen and Zhang [79] proposed two improvements of the Ahmed et al. algorithm to reduce the computational time. On the other hand, to keep the continuity from the FCM algorithm, Shen et al. [80] introduced a new similarity measure that depends on spatial neighbourhood information, where the degree of the neighbourhood attraction is optimized by a neural network. The clustering performance can also be enhanced by combining pixel-wise fuzzy classification with preprocessing (noise cleaning in the original image) [49, 78] and postprocessing (noise cleaning on the classified data) [78].

An example of the multimodal T_1 -W and T_2 -W MRI clustering of the adult brain is shown in Figure 13(a). In general, the shape of joints T_1 -W and T_2 -W MRI intensity distributions of different tissue classes depends on the image quality (the presence of noise, PVE, etc.). The shape of the classified data depends on the applied segmentation method. In the example in Figure 13(a), there is a small overlap among classes due to the good quality MRI. Thus, the standard k -means clustering method is used to segment the brain tissue probability maps (see Section 4.2.4) and the final clusters are indicated with different colors in the scatter plot of T_1 -W and T_2 -W MRI in Figure 13(b). In general, when MRI artifacts are present and there is a significant overlap among tissue classes, the spatial information of the brain tissue is required to disambiguate the classification problem.

4.3. Atlas-Based Methods. If an atlas or template of the human brain for a specific population of interest is available, then atlas-based methods can be a powerful tool for brain MRI segmentation. The atlas contains information about the brain anatomy (e.g., it contains the information about the location of different brain structures) and it is used as a reference (a prior knowledge) for segmenting new images. The main advantage of these methods is the possibility to segment any brain structure available in the atlas without any additional cost. Conceptually, atlas-based approaches are similar to classifier methods, except that they are implemented in the spatial domain rather than in the feature space.

Before a probabilistic atlas can be used as a prior knowledge, it has to be aligned with the image to be segmented. Since the segmentation labels and the “ground truth” are known for the atlas, all atlas information is transferred to the target image after registration. Therefore, the performance of atlas-based methods is directly dependent on quality of the registration method used.

The traditional way of aligning the probabilistic atlas with the image is to use affine registration. Unfortunately, an affine alignment may not be sufficient if the brain anatomy of interest differs significantly from the average atlas anatomy. Pohl et al. therefore suggest aligning the atlas using nonrigid registration [85]. However, in their later work Pohl reports

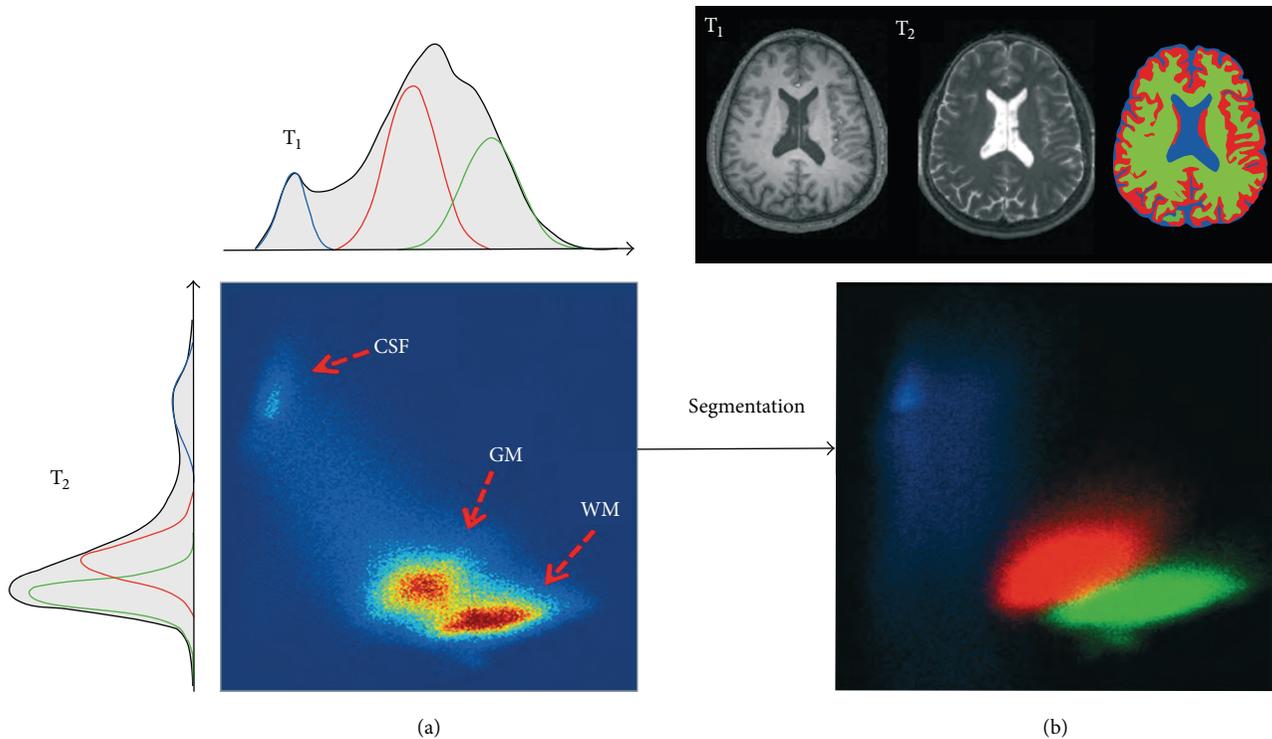


FIGURE 13: (a) Joint 2D intensity histogram of T_1 -W and T_2 -W MRI of the adult brain. The associated 1D histograms of each MRI modality are plotted on the left and top. Both individual histograms consist of three overlapped Gaussian distributions that approximate the expected tissue distribution of GM, WM, and CSF. (b) The scatter plot of the tissue intensities after applying tissue segmentation. The horizontal axis represents T_1 -W intensities and the vertical axis represents T_2 -W intensities. The red cloud corresponds to GM, the green to WM, and the blue to CSF.

difficulties in registering anatomical template with the image to be segmented using standard registration methods [86]. D'Agostino et al. developed a special similarity measure for registering probabilistic maps directly to the new image [87]. Recently, several methods have been developed which aim to overcome this problem by iteratively refining the segmentation and nonrigid registration of the probabilistic atlas at the same time. Ashburner and Friston developed a method for simultaneous segmentation, bias correction, and nonrigid registration of a probabilistic atlas [2].

However, even with nonrigid registration methods, accurate segmentation of complex structures is difficult due to anatomical variability. Also, atlas-guided segmentation in patients with brain deformations can be difficult and prone to errors, because the probabilistic atlas is based on a population of healthy subjects. For instance, in patients with brain lesions or a brain anatomy that significantly differs from the atlas template, the atlas alignment and the corresponding segmentation of the brain will fail or give inaccurate results. In these cases an atlas-based approach is not a suitable method for image segmentation.

An aligned probabilistic atlas can be also used as a good initial estimate of the segmentation, which is especially important for EM-based methods, as EM algorithm is guaranteed to converge to local, not global, maxima. In addition, most EM-based methods [2, 71] use the probabilistic atlas to constrain the segmentation process where again the correct

alignment of the probabilistic atlas is crucial for successful and accurate segmentation.

It is important to note that atlas-based MRI segmentation of the neonatal brain has become a research focus in recent years [42, 51, 53, 88]. MRI segmentation of the neonatal brain tissue is more complex than in adults due to fast growth process, complex anatomy of the developing brain, and often poor MRI quality. Therefore a probabilistic atlas of the newborn brain that contains the spatial variability of the tissue structure is used to segment different brain tissues such as brain cortex, myelinated and nonmyelinated white matter. However, a good atlas of the newborn brain is even more difficult to obtain than in adults, mainly due to the greater anatomical variations between subjects. Therefore, it was necessary to develop a dynamic, probabilistic atlas for any chosen stage of neonatal brain development (for ages of 29 to 44 weeks) [89].

4.4. Surface-Based Methods. In addition to intensity-based and atlas-based methods, there are a number of alternative brain MRI segmentation approaches. These approaches include surface-based methods, such as deformable models including active contours and surfaces [46, 52, 54, 61, 90–93].

4.4.1. Active Contours and Surfaces. Deformable models are also called active contours or snakes in 2D and active surfaces

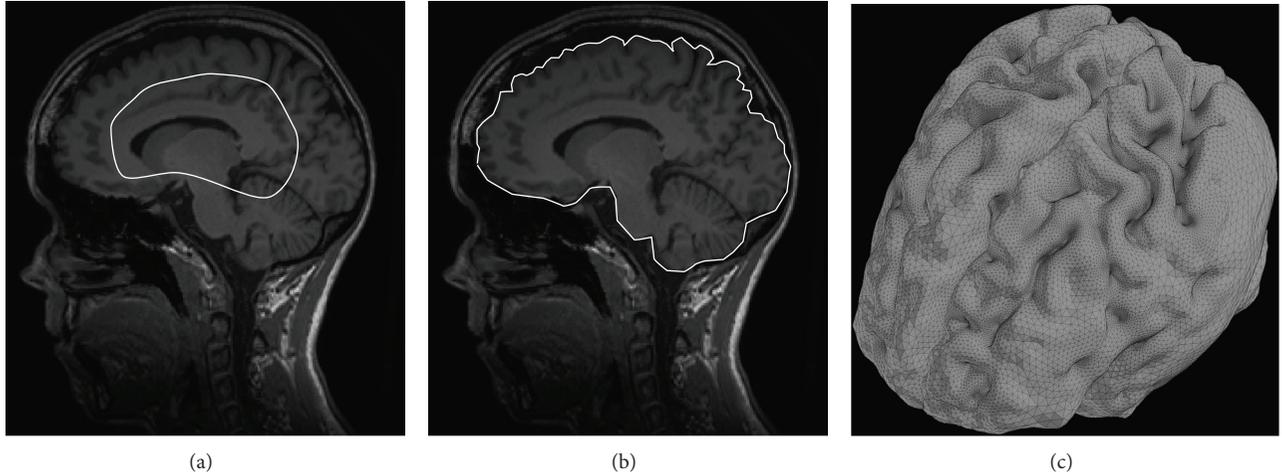


FIGURE 14: Segmentation of the brain surface using deformable models. (a) A closed curve is initialised inside the brain. (b) The segmentation result of the brain surface in 2D. (c) 3D surface of the brain.

or active balloons in 3D. Deformable models were introduced by Kass et al. [92] in 2D space and were further developed and generalized in 3D space by Terzopoulos et al. [93, 94]. Deformable models use closed parametric curves or surfaces for delineating region boundaries. The parametric curves and surfaces deform under the influence of external (or image) forces (controlled by the image attributes) and the internal forces, which control the surface regularity. In general, deformable models represent the fusion of geometry, physics, and approximation theory. Geometry is used to represent the shape of the object, physics defines constraints on how the shape may vary over time and space, and approximation theory provides mechanisms for fitting the models to measured data. A visual example of segmentation using deformable models is given in Figure 14.

To delineate a boundary of an object, first a closed curve or surface \mathcal{S} is placed near the desired boundary in an image. Then, internal and external forces are deforming the curve or surface in an iterative relaxation process where the energy functional is defined as

$$F(\mathcal{S}) = F_{\text{int}} + F_{\text{ext}}. \quad (17)$$

The internal forces F_{int} are computed from within the curve or surface to keep it smooth throughout the deformation. The external forces F_{ext} are usually derived from the image to deform the curve or surface towards the desired feature of interest.

In traditional deformable models, image forces come primarily from the local edge-based information (e.g., based on the gradients of sharp image intensities) [95–97]. However, such reliance on edge information makes deformable models sensitive to noise (e.g., deformable model can leak through noisy edges) and highly dependent on the initial estimate. There have been significant efforts to integrate more global region information into deformable models. The Mumford-Shah model [98] was one of the first region-based methods where the image is approximated using a smooth function inside the regions and not only at their

boundaries. Many variants of this model have been proposed later [46, 90]. For instance, Chan and Vese [46] presented a level set method which approximates an image with a constant function inside the regions. Li et al. proposed a region-based level set method for segmentation of MRI in the presence of intensity inhomogeneity. Furthermore, several hybrid deformable models had been later proposed to make use of both local (edge-based) and global (region-based) information [99–102]. Sometimes the image data are not sufficient to delineate the region of interest and thus prior knowledge has to be introduced [103].

4.4.2. Multiphase Active Contours. The most popular Chan-Vese level set method [46] has been successfully used in segmenting images with two distinctive regions (images with binary segmentation energies). In [104], Vese and Chan extended their binary segmentation energies to a multiphase level set formulation. In this way, multiple nonoverlapping regions with spatial consistency and varying characteristics (such as the mean intensities of regions) could be represented with multiple level set functions. This multiphase level set approach was attractive for segmentation of brain MR images which typically have multiple regions of interest with different characteristics. Starting from the Vese and Chan method [104], different extension to multiphase active contours had been developed [54, 105–108]. The advantage of multiphase active contours to other approaches is their robustness to image variations, adaptive energy functionals, topological flexibility, and accurate boundaries.

Traditionally, active contours methods have nonconvex energy minimization due to a gradient descent formulation. In this way, energy minimization converges to undesirable local minima and results in erroneous segmentations. Also, the traditional level set implementation has slower convergence due to discretization errors and the well-known reinitialization requirement. One of the first convex approaches to the two-phase active contours segmentation was proposed by Chan et al. [109]. Later on, several extensions to

the more challenging multiphase problem have been proposed [110, 111]. A lot of recent work has been dedicated to developing new multiphase active contours not only with a convex formulation, but also with a reduced computational complexity [54, 105, 107]. Note that globally convex methods are initialization independent.

4.5. Hybrid Segmentation Methods. New application-specific brain MRI segmentation problems are emerging and new methods are continuously explored and introduced. Since a selection of the most appropriate technique for a given application is often a difficult task, a combination of several techniques may be necessary to obtain the segmentation goal. Therefore, hybrid or combined segmentation methods have been used extensively in different brain MRI segmentation applications [78, 84, 112–120]. The main idea is to combine different complementary segmentation methods into a hybrid approach to avoid many of the disadvantages of each method alone and improve segmentation accuracy.

Here are some examples of the hybrid brain MRI segmentation methods. Kapur et al. [112] segmented different brain tissues in adults using 2D MRI by combining expectation-maximization segmentation, binary mathematical morphology, and active contours models. Masutani et al. [113] combined model-based region growing with morphological information of local shape to segment cerebral blood vessels. Warfield et al. [120] developed a combined 3D brain MRI segmentation algorithm which iterates between a classification step to identify tissues and an elastic matching step to align a template of normal brain anatomy with the classified tissues. Elastic matching step can generate image segmentation by registering an anatomical atlas to a patient scan.

Furthermore, an unsupervised global-to-local brain MRI segmentation is developed by Xue et al. [78]. They combined minimum error global thresholding and a spatial-feature-based FCM clustering to segment 3D MRI in a “slice-by-slice” manner. In the work of Vijayakumar and Gharpure [117], a hybrid MRI segmentation method, based on artificial neural networks (ANN), is proposed for segmenting tumor lesions, edema, cysts, necrosis, and normal tissue in T2 and FLAIR MRI. More recently, Ortiz et al. [119] suggested an improved brain MRI segmentation method using self-organizing maps (a particular case of ANN) and entropy-gradient clustering.

Hybrid segmentation methods are also used for the neonatal brain segmentation [115, 118]. For example, Despotovic et al. [115] proposed a hybrid strategy to segment the brain volume in neonates using T_1 -W and T_2 -W MRI by combining thresholding, active contours, FCM clustering, and morphological operations. Later on, Gui et al. [118] proposed a morphology-driven automatic segmentation method to segment different anatomical regions of the neonatal brain.

The main drawback of hybrid (combined) segmentation methods is often the increased complexity in comparison with each single method integrated into a hybrid one. This includes a lower computational time and a higher number of different parameters that needs to be tuned for a specific application. Therefore, a hybrid segmentation method should be carefully and wisely designed to give efficient and good quality segmentation.

4.6. Partial Volume Effect Correction. As mentioned in Section 2.5, the PVE problem is one of the most common problems in brain MRI segmentation. The PVE describes the loss of small tissue regions because of the limited resolution of the MRI scanner and it is seen on MRI scans as a mix of different tissues in a single pixel/voxel. This effect can cause the misclassification of pixel/voxel that lies in the transition between two (or more) tissues classes. Several methods have been proposed to address the problem of PVE in MRI segmentation for both adult and neonatal brains [3, 42, 103, 121, 122].

One of the first approaches for the partial volume correction is the method by Santago and Gage [123]. They assumed a uniform prior probability for mixed (nonpure) tissues and calculated the intensity distribution of partial volume by minimizing the distance between a model and an image histogram. Another approach is proposed by Nocera and Gee [124], where they used MRF to obtain spatial smooth variations of tissue mixing proportions and a MAP estimation is then computed for partial volume segmentation. Later on, Zhang et al. [3] employ a hidden Markov random field with a finite mixture model to overcome possible PVE and bias field distortions. In the work by Van Leemput et al. [122], a statistical uniform framework for partial volume segmentation is presented without using a heuristic assumption for the prior distribution of mix proportions. They used a parametric statistical image model where each voxel belongs to a single tissue type and introduced an additional downsampling step to cause partial volumes along the borders between tissues. Then, they estimated the tissue mixing proportions by using the expectation-maximization approach.

For all above-mentioned partial volume correction methods, very promising results have been reported for adult T_1 -W MRI. However, these methods rely on the fact that the intensity levels of partial volumes in adult MRI images do not predominantly overlap with the characteristic intensity of any pure tissue class. This assumption is not possible for neonatal MRI due to the inverted graywhite matter contrast and great tissue overlap due to the existence of myelinated WM and nonmyelinated WM; see Figure 8(b). Therefore, several approaches have been proposed to address the PVE in neonatal MRI [42, 103]. Xue et al. [42] proposed method based on expectation-maximization, Markov random field, and atlas information to remove mislabeled voxels and correct errors caused by PVE. They implemented a label propagation strategy to mask off deep GM and myelinated WM, which enabled the segmentation of cortical GM and nonmyelinated WM. In the paper of Wang et al. [103], a rather simple and effective scheme is used to deal with the PVE problem and is based on the anatomical observation that the misclassified WM voxels are surrounded by the CSF and GM and that mislabeled CSF voxels are unconnected from the true WM volume.

5. Validation of Brain MRI Segmentation

Validation and quantitative comparison of different segmentation methods are a general problem in medical image analysis. It requires a “ground truth” or gold standard to

which the outcome of the segmentation method can be compared. Unfortunately, the “ground truth” does not exist for the analysis of in vivo acquired data in humans. Thus, the “ground truth” of the real patients is typically generated after image acquisition.

In brain MRI analysis, the “ground truth” for the real patient data is usually made by one or more expert physicians who need to manually analyze and segment anatomical structures of interest; see Section 4.1. Although this is the only way to validate the real patient MRI data, this validation must be critically considered because the manual segmentation is prone to errors, highly subjective, and difficult to reproduce (even by the same expert) [59]. Also, this type of validation is not always available because it is time-consuming and depends on the human operator. Therefore, few alternative validation methods evolved in the praxes to validate the accuracy of the segmentation algorithms. The most popular validation methods include the use of software simulations and phantoms.

In software simulations, the artificial MR images are generated with computer programs that simulate the real acquisition process. In this way the “ground truth” is known and the influence of different acquisition parameters and imaging artifacts can be controlled and examined independently. This type of validation is very flexible and easily accessible by different researchers and can be performed with little effort. However, a drawback of this validation is that software simulators cannot take into account all factors that might influence the real image acquisition and the simulated images are only an approximation of the real images.

Since software simulations have certain limitations, validation of new segmentation methods can be done using human-like phantoms, whose physical properties (e.g., geometry of the tissue structures and material properties) are known and are similar to the in vivo properties. The phantom images are generated using the MRI scanner and are more realistic than images generated with software simulations. On the other hand, the phantom images do not offer the flexibility of the software simulations and imaging is more expensive and labour intensive.

The most popular simulated images used for validation of brain MRI segmentation methods are designed by Collins et al. [125] and are also known as a realistic digital brain phantom or simply BrainWeb. Images are freely available online and easily accessible for all researchers to test the performance of the new segmentation methods. The BrainWeb data consists of $181 \times 217 \times 181$ voxel matrix with a resolution of $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$ and is available for different additive noise levels. The noise level (expressed in percentages) is relative to the average real and imaginary values of the overall brightness of the tissue class. The noise is generated using a pseudorandom Gaussian noise, which is added to both real and imaginary components before the final magnitude value of the simulated MR image is computed.

Beside the phantom BrainWeb data, the most popular repository with real MRI data used for validation of brain MRI segmentation methods is the Internet Brain Segmentation Repository (IBSR) [126]. The IBSR repository is also freely available online. It consists of 20 real T_1 -W MRI

brain data sets and manually guided expert segmentation results, which are used as a “ground truth” segmentation. Each MRI volume consists of about 60 coronal T_1 -W slices with the interslice resolution of 3.1 mm (thickness between consecutive slices).

To quantify the overlap between the MRI segmentation and the given “ground truth,” several similarity measures are used in the literature. One of the most popular measures used often with BrainWeb data is the Dice coefficient ρ_i [127]:

$$\rho_i = \frac{2|A_i \cap B_i|}{|A_i| + |B_i|}, \quad (18)$$

where i stands for a tissue type, A_i and B_i denote the set of pixels labeled into i by the “ground truth” and MRI segmentation, respectively, and $|A_i|$ denotes the number of elements in A_i . The Dice coefficient is in the range $0 \leq \rho_i \leq 1$ and has value 0 if there is no overlap between the two segmentations and 1 if both segmentations are identical.

The Tanimoto coefficient (also known as the Jaccard index) is often used as a similarity measure with real IBSR data. The Tanimoto coefficient $\mathcal{T}(i)$ for each tissue type i is defined as follows:

$$\mathcal{T}_i = \frac{|A_i \cap B_i|}{|A_i| + |B_i| - |A_i \cap B_i|}, \quad (19)$$

where A_i and B_i denote the set of pixels labeled into i by the “ground truth” and the segmentation method, respectively, and $|A_i|$ denotes the number of elements in A_i . Note that $\mathcal{T}(i) \leq \rho(i)$ and $0 \leq \mathcal{T}(i) \leq 1$.

More readings about similarity measures for evaluation and validation in medical image analysis can be found in [128].

6. Discussion and Conclusions

Image segmentation is an important step in many medical applications involving 3D visualization, computer-aided diagnosis, measurements, and registration. This paper has provided a brief introduction to the fundamental concepts of MRI segmentation of the human brain and methods that are commonly used.

In Section 2, we have defined the basic concepts necessary for understanding MRI segmentation methods, such as 2D and 3D image definition, image features, and brain MRI intensity distributions. Following this, preprocessing steps necessary to prepare images for MRI segmentation have been described in Section 3. The most important steps include bias field correction, image registration, and removal of nonbrain tissues or brain extraction. The correction of intensity inhomogeneity is an important step for the efficient segmentation and registration of brain MRI. Image registration is required in brain MRI segmentation for the alignment of multimodal images of the same subject or several population groups taken at different times and from different viewpoints.

Due to the rapid development of medical image modalities, new application-specific segmentation problems are emerging and new methods are continuously explored and

introduced. Selection of the most appropriate technique for a given application is a difficult task. In many cases, a combination of several techniques may be necessary to obtain the segmentation goal. Very often integration of multimodal information (acquired from different modalities or over time) can help to segment structures that otherwise could not be detected on single images.

The most popular image segmentation methods that are used for brain MRI segmentation have been reviewed and discussed in Section 4. Newer methods are usually designed to bring more accurate results by incorporating 3D neighborhood information and prior information from atlases. As a consequence, the segmentation process often becomes more complex and time-consuming. The likely future research will still focus not only on developing more accurate and noise-robust methods, but also on improving the computational speed of segmentation methods. Computational efficiency will be particularly important in real-time processing applications such as computer guided surgery.

Probably one of the most important questions concerning medical image segmentation is its use in real clinical settings. It is undeniable that computerized segmentation methods have shown their potentials and applicability in computer-aided diagnosis and therapy planning. It is expected that in the near future they will also become essential tools in real clinical settings, particularly in qualitative diagnosis and where 3D reconstruction and visualization of the anatomical structures are important.

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 FWO-Vlaanderen under Grant G.0341.07 “Data fusion of multimodal information using advanced signal processing, segmentation and registration techniques” and by IWT project NeoGuard.

References

- [1] D. L. Pham, C. Xu, and J. L. Prince, “Current methods in medical image segmentation,” *Annual Review of Biomedical Engineering*, vol. 2, no. 2000, pp. 315–337, 2000.
- [2] J. Ashburner and K. J. Friston, “Unified segmentation,” *NeuroImage*, vol. 26, no. 3, pp. 839–851, 2005.
- [3] 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.
- [4] S. Z. Li, *Markov Random Field Modeling in Computer Vision*, Springer, New York, NY, USA, 1995.
- [5] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.
- [6] P. Yan, S. Xu, B. Turkbey, and J. Kruecker, “Discrete deformable model guided by partial active shape model for trus image segmentation,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 5, pp. 1158–1166, 2010.
- [7] O. Ecabert, J. Peters, H. Schramm et al., “Automatic model-based segmentation of the heart in CT images,” *IEEE Transactions on Medical Imaging*, vol. 27, no. 9, pp. 1189–1201, 2008.
- [8] J. Yang and J. S. Duncan, “3D image segmentation of deformable objects with joint shape-intensity prior models using level sets,” *Medical Image Analysis*, vol. 8, no. 3, pp. 285–294, 2004.
- [9] X. Tao, J. L. Prince, and C. Davatzikos, “Using a statistical shape model to extract sulcal curves on the outer cortex of the human brain,” *IEEE Transactions on Medical Imaging*, vol. 21, no. 5, pp. 513–524, 2002.
- [10] N. Duta and M. Sonka, “Segmentation and interpretation of MR brain images: an improved active shape model,” *IEEE Transactions on Medical Imaging*, vol. 17, no. 6, pp. 1049–1062, 1998.
- [11] R. Davies, C. Twining, and C. Taylor, *Statistical Models of Shape Optimisation and Evaluation*, Springer, London, UK, 2008.
- [12] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Pearson Education, 2008.
- [13] J. Rogowska, “Overview and fundamentals of medical image segmentation,” in *Handbook of Medical Image Processing and Analysis*, I. Bankman, Ed., Elsevier, Amsterdam, The Netherlands, 2000.
- [14] J. Canny, “A computational approach to edge detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
- [15] D. Marr and E. Hildreth, “Theory of edge detection,” *Proceedings of the Royal Society of London—Biological Sciences*, vol. 207, no. 1167, pp. 187–217, 1980.
- [16] P. Kovese, “Image feature from phase congruency,” *Journal of Computer Vision Research*, vol. 1, no. 3, pp. 1–26, 1999.
- [17] P. Kovese, “Edges are not just steps,” in *Proceedings of the 5th Asian Conference on Computer Vision (ACCV '02)*, pp. 822–827, Melbourne, Australia, 2002.
- [18] M. C. Morrone and R. A. Owens, “Feature detection from local energy,” *Pattern Recognition Letters*, vol. 6, no. 5, pp. 303–313, 1987.
- [19] H. Gudbjartsson and S. Patz, “The Rician distribution of noisy MRI data,” *Magnetic Resonance in Medicine*, vol. 34, no. 6, pp. 910–914, 1995.
- [20] D. Salas-González, J. M. Górriz, J. Ramírez, M. Schloegl, E. W. Lang, and A. Ortiz, “Parameterization of the distribution of white and grey matter in MRI using-stable distribution,” *Computers in Biology and Medicine*, vol. 43, pp. 559–567, 2013.
- [21] B. Fischl, D. H. Salat, E. Busa et al., “Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain,” *Neuron*, vol. 33, no. 3, pp. 341–355, 2002.
- [22] P. Maillard, N. Delcroix, F. Crivello et al., “An automated procedure for the assessment of white matter hyperintensities by multispectral (T1, T2, PD) MRI and an evaluation of its between-centre reproducibility based on two large community databases,” *Neuroradiology*, vol. 50, no. 1, pp. 31–42, 2008.
- [23] A. Mayer and H. Greenspan, “An adaptive mean-shift framework for MRI brain segmentation,” *IEEE Transactions on Medical Imaging*, vol. 28, no. 8, pp. 1238–1250, 2009.
- [24] C. R. Traynora, G. J. Barkerb, W. R. Crumb, S. C. R. Williamsb, and M. P. Richardsona, “Segmentation of the thalamus in MRI based on T1 and T2,” *NeuroImage*, vol. 56, no. 1, pp. 939–950, 2011.

- [25] C. M. Collins, W. Liu, W. Schreiber, Q. X. Yang, and M. B. Smith, "Central brightening due to constructive interference with, without, and despite dielectric resonance," *Journal of Magnetic Resonance Imaging*, vol. 21, no. 2, pp. 192–196, 2005.
- [26] J. G. Sied, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Transactions on Medical Imaging*, vol. 17, no. 1, pp. 87–97, 1998.
- [27] E. B. Lewis and N. C. Fox, "Correction of differential intensity inhomogeneity in longitudinal MR images," *NeuroImage*, vol. 23, no. 1, pp. 75–83, 2004.
- [28] 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.
- [29] M. S. Cohen, R. M. DuBois, and M. M. Zeneih, "Rapid and effective correction of RF inhomogeneity for high field magnetic resonance imaging," *Human Brain Mapping*, vol. 10, no. 4, pp. 204–211, 2000.
- [30] 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.
- [31] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based bias field correction of MR images of the brain," *IEEE Transactions on Medical Imaging*, vol. 18, no. 10, pp. 885–896, 1999.
- [32] J.-F. Mangin, "Entropy minimization for automatic correction of intensity nonuniformity," in *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA '00)*, pp. 162–169, Hilton Head Island, SC, USA, June 2000.
- [33] A. C. Evans, D. L. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, and T. M. Peters, "3D statistical neuroanatomical models from 305 MRI volumes," in *Proceedings of the IEEE Nuclear Science Symposium & Medical Imaging Conference*, pp. 1813–1817, November 1993.
- [34] J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, *Medical Image Registration*, CRC Press, New York, NY, USA, 2001.
- [35] D. Shen and C. Davatzikos, "HAMMER: hierarchical attribute matching mechanism for elastic registration," *IEEE Transactions on Medical Imaging*, vol. 21, no. 11, pp. 1421–1439, 2002.
- [36] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "A viscous fluid model for multimodal non-rigid image registration using mutual information," in *Proceedings of the Medical Image Computing and Computer Assisted Intervention (MICCAI '02)*, pp. 541–548, 2002.
- [37] J. Ashburner and K. J. Friston, "Nonlinear spatial normalization using basis functions," *Human Brain Mapping*, vol. 7, no. 4, pp. 254–266, 1999.
- [38] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast mr images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.
- [39] J. M. Fitzpatrick, D. L. G. Hill, and C. R. Maurer, "Image registration," in *Handbook of Medical Imaging: Medical Image Processing and Analysis*, M. Sonka and J. M. Fitzpatrick, Eds., vol. 2, SPIE Publications, 2004.
- [40] W. R. Crum, T. Hartkens, and D. L. G. Hill, "Non-rigid image registration: theory and practice," *British Journal of Radiology*, vol. 77, no. 2, pp. S140–S153, 2004.
- [41] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.
- [42] H. Xue, L. Srinivasan, S. Jiang et al., "Automatic segmentation and reconstruction of the cortex from neonatal MRI," *NeuroImage*, vol. 38, no. 3, pp. 461–477, 2007.
- [43] S. M. Smith, "Fast robust automated brain extraction," *Human Brain Mapping*, vol. 17, no. 3, pp. 143–155, 2002.
- [44] M. Battaglini, S. M. Smith, S. Brogi, and N. de Stefano, "Enhanced brain extraction improves the accuracy of brain atrophy estimation," *NeuroImage*, vol. 40, no. 2, pp. 583–589, 2008.
- [45] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, no. 9, pp. 1277–1294, 1993.
- [46] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
- [47] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient N-D image segmentation," *International Journal of Computer Vision*, vol. 70, no. 2, pp. 109–131, 2006.
- [48] R. Unnikrishnan, C. Pantofaru, and M. Hebert, "Toward objective evaluation of image segmentation algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 929–944, 2007.
- [49] W. Cai, S. Chen, and D. Zhang, "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation," *Pattern Recognition*, vol. 40, no. 3, pp. 825–838, 2007.
- [50] M. Prastawa, *An MRI segmentation framework for brains with anatomical deviations [Ph.D. thesis]*, University of North Carolina at Chapel Hill, 2007.
- [51] F. Shi, D. Shen, P.-T. Yap et al., "CENTS: cortical enhanced neonatal tissue segmentation," *Human Brain Mapping*, vol. 32, no. 3, pp. 382–396, 2011.
- [52] L. Wang, F. Shi, G. Li et al., "Segmentation of neonatal brain MR images using patch-driven level sets," *NeuroImage*, vol. 84, pp. 141–158, 2014.
- [53] N. I. Weisenfeld and S. K. Warfield, "Automatic segmentation of newborn brain MRI," *NeuroImage*, vol. 47, no. 2, pp. 564–572, 2009.
- [54] J. C. Moreno, V. B. Surya Prasath, H. Proença, and K. Palaniappan, "Fast and globally convex multiphase active contours for brain MRI segmentation," *Computer Vision and Image Understanding*, vol. 125, pp. 237–250, 2014.
- [55] D. M. Greig, B. T. Porteous, and A. H. Seheult, "Exact maximum a posteriori estimation for binary images," *Journal of the Royal Statistical Society B*, vol. 51, no. 2, pp. 271–279, 1989.
- [56] J. de Oliveira and W. Pedrycz, *Advances in Fuzzy Clustering and Its Applications*, John Wiley & Sons, New York, NY, USA, 2007.
- [57] I. Despotović, E. Vansteenkiste, and W. Philips, "Spatially coherent fuzzy clustering for accurate and noise-robust image segmentation," *IEEE Signal Processing Letters*, vol. 20, no. 4, pp. 295–298, 2013.
- [58] D. C. Collier, S. S. C. Burnett, M. Amin et al., "Assessment of consistency in contouring of normal-tissue anatomic structures," *Journal of Applied Clinical Medical Physics*, vol. 4, no. 1, pp. 17–24, 2003.
- [59] E. Vansteenkiste, *Quantitative analysis of ultrasound images of the preterm brain [Ph.D. dissertation]*, Ghent University, 2007.
- [60] M. Murgasova, *Segmentation of brain MRI during early childhood [Ph.D. thesis]*, Imperial College London, 2008.

- [61] P. A. Yushkevich, J. Piven, H. C. Hazlett et al., "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability," *NeuroImage*, vol. 31, no. 3, pp. 1116–1128, 2006.
- [62] ITK-SNAP, 2009, <http://www.itksnap.org/>.
- [63] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 146–168, 2004.
- [64] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Computer Vision, Graphics, and Image Processing*, vol. 29, no. 1, pp. 100–132, 1985.
- [65] N. Passat, C. Ronse, J. Baruthio, J.-P. Armspach, C. Maillot, and C. Jahn, "Region-growing segmentation of brain vessels: an atlas-based automatic approach," *Journal of Magnetic Resonance Imaging*, vol. 21, no. 6, pp. 715–725, 2005.
- [66] T. Weglinski and A. Fabijanska, "Brain tumor segmentation from MRI data sets using region growing approach," in *Proceedings of the 7th International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH '11)*, pp. 185–188, May 2011.
- [67] M. del Fresno, M. Vénere, and A. Clause, "A combined region growing and deformable model method for extraction of closed surfaces in 3D CT and MRI scans," *Computerized Medical Imaging and Graphics*, vol. 33, no. 5, pp. 369–376, 2009.
- [68] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, John Wiley & Sons, New York, NY, USA, 2nd edition, 2001.
- [69] S. K. Warfield, F. A. Kaus, M. Jolesz, and R. Kikinis, "Adaptive template moderated spatially varying statistical classification," in *Proceedings of the Medical Image Computing and Computer Assisted Intervention (MICCAI '98)*, pp. 431–438, 1998.
- [70] C. A. Cocosco, A. P. Zijdenbos, and A. C. Evans, "A fully automatic and robust brain MRI tissue classification method," *Medical Image Analysis*, vol. 7, no. 4, pp. 513–527, 2003.
- [71] K. M. Pohl, J. Fisher, W. E. L. Grimson, R. Kikinis, and W. M. Wells, "A Bayesian model for joint segmentation and registration," *NeuroImage*, vol. 31, no. 1, pp. 228–239, 2006.
- [72] C. W. Therrien, *Decision, Estimation, and Classification: An Introduction to Pattern Recognition and Related Topics*, John Wiley & Sons, New York, NY, USA, 1989.
- [73] G. B. Coleman and H. C. Andrews, "Image segmentation by clustering," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 773–785, 1979.
- [74] 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.
- [75] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, NY, USA, 1981.
- [76] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [77] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty, "A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data," *IEEE Transactions on Medical Imaging*, vol. 21, no. 3, pp. 193–199, 2002.
- [78] J.-H. Xue, A. Pizurica, W. Philips, E. Kerre, R. van de Walle, and I. Lemahieu, "An integrated method of adaptive enhancement for unsupervised segmentation of MRI brain images," *Pattern Recognition Letters*, vol. 24, no. 15, pp. 2549–2560, 2003.
- [79] S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. 34, no. 4, pp. 1907–1916, 2004.
- [80] 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.
- [81] K.-S. Chuang, H.-L. Tzeng, S. Chen, J. Wu, and T.-J. Chen, "Fuzzy c-means clustering with spatial information for image segmentation," *Computerized Medical Imaging and Graphics*, vol. 30, no. 1, pp. 9–15, 2006.
- [82] Z. M. Wang, Y. C. Soh, Q. Song, and K. Sim, "Adaptive spatial information-theoretic clustering for image segmentation," *Pattern Recognition*, vol. 42, no. 9, pp. 2029–2044, 2009.
- [83] I. Despotović, B. Goossens, E. Vansteenkiste, and W. Philips, "T1- and T2-weighted spatially constrained fuzzy C-means clustering for brain MRI segmentation," in *Medical Image Processing, Proceedings of SPIE*, p. 9, San Diego, Calif, USA, February 2010.
- [84] B. N. Li, C. K. Chui, S. Chang, and S. H. Ong, "Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation," *Computers in Biology and Medicine*, vol. 41, no. 1, pp. 1–10, 2011.
- [85] K. M. Pohl, W. M. Wells, A. Guimond et al., "Incorporating non-rigid registration into expectation maximization algorithm to segment MR images," in *Proceedings of the Medical Image Computing and Computer Assisted Intervention (MICCAI '02)*, pp. 564–572, 2002.
- [86] K. M. Pohl, *Prior information for brain parcellation [Ph.D. thesis]*, Massachusetts Institute of Technology, 2005.
- [87] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "Non-rigid atlas-to-image registration by minimization of class-conditional image entropy," in *Proceedings of the 7th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '04)*, pp. 745–753, September 2004.
- [88] M. Prastawa, J. H. Gilmore, W. Lin, and G. Gerig, "Automatic segmentation of MR images of the developing newborn brain," *Medical Image Analysis*, vol. 9, no. 5, pp. 457–466, 2005.
- [89] M. Kuklisova-Murgasova, P. Aljabar, L. Srinivasan et al., "A dynamic 4D probabilistic atlas of the developing brain," *NeuroImage*, vol. 54, no. 4, pp. 2750–2763, 2011.
- [90] C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas, and J. C. Gore, "A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 2007–2016, 2011.
- [91] D. Rivest-Hénault and M. Cheriet, "Unsupervised MRI segmentation of brain tissues using a local linear model and level set," *Magnetic Resonance Imaging*, vol. 29, no. 2, pp. 243–259, 2011.
- [92] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [93] D. Terzopoulos, A. Witkin, and M. Kass, "Constraints on deformable models: recovering 3D shape and nonrigid motion," *Artificial Intelligence Journal*, vol. 36, no. 1, pp. 91–123, 1988.
- [94] L. D. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2-D and 3-D images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1131–1147, 1993.
- [95] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzy, "Gradient flows and geometric active contour models,"

- in *Proceedings of the 5th International Conference on Computer Vision (ICCV '95)*, pp. 810–815, 1995.
- [96] R. Malladi, J. A. Sethian, and B. C. Vemuri, “Shape modeling with front propagation: a level set approach,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 158–175, 1995.
- [97] V. Caselles, R. Kimmel, and G. Sapiro, “Geodesic active contours,” *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.
- [98] D. Mumford and J. Shah, “Optimal approximations by piecewise smooth functions and associated variational problems,” *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, 1989.
- [99] A. Huang, R. Abugharbieh, and R. Tam, “A hybrid geometric-statistical deformable model for automated 3-D segmentation in brain MRI,” *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 7, pp. 1838–1848, 2009.
- [100] R. Kimmel, “Fast edge integration,” in *Geometric Level Set Methods in Imaging, Vision, and Graphics*, S. Osher and N. Paragios, Eds., Springer, New York, NY, USA, 2003.
- [101] C. Sagiv, N. A. Sochen, and Y. Y. Zeevi, “Integrated active contours for texture segmentation,” *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1633–1646, 2006.
- [102] P. Mesejo, A. Valsecchi, L. Marrakchi-Kacem, S. Cagnoni, and S. Damas, “Biomedical image segmentation using geometric deformable models and metaheuristics,” *Computerized Medical Imaging and Graphics*, 2014.
- [103] L. Wang, F. Shi, W. Lin, J. H. Gilmore, and D. Shen, “Automatic segmentation of neonatal images using convex optimization and coupled level sets,” *NeuroImage*, vol. 58, no. 3, pp. 805–817, 2011.
- [104] L. A. Vese and T. F. Chan, “A multiphase level set framework for image segmentation using the Mumford and Shah model,” *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.
- [105] E. S. Brown, T. F. Chan, and X. Bresson, “Completely convex formulation of the Chan-Vese image segmentation model,” *International Journal of Computer Vision*, vol. 98, no. 1, pp. 103–121, 2012.
- [106] M. S. Keegan, B. Sandberg, and T. F. Chan, “A multiphase logic framework for multichannel image segmentation,” *Inverse Problems and Imaging*, vol. 6, no. 1, pp. 95–110, 2012.
- [107] S. H. Kang and R. March, “Existence and regularity of minimizers of a functional for unsupervised multiphase segmentation,” *Nonlinear Analysis: Theory, Methods & Applications*, vol. 76, pp. 181–201, 2013.
- [108] 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.
- [109] 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.
- [110] A. Chambolle, D. Cremers, and T. Pock, “A convex approach to minimal partitions,” *SIAM Journal on Imaging Sciences*, vol. 5, no. 4, pp. 1113–1158, 2012.
- [111] E. Bae, J. Yuan, and X.-C. Tai, “Global minimization for continuous multiphase partitioning problems using a dual approach,” *International Journal of Computer Vision*, vol. 92, no. 1, pp. 112–129, 2011.
- [112] 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.
- [113] Y. Masutani, T. Schiemann, and K. H. Hohne, “Vascular shape segmentation and structure extraction using a shape-based region-growing model,” in *Medical Image Computing and Computer-Assisted Intervention—MICCAI'98: Proceedings of the 1st International Conference Cambridge, MA, USA, October 11–13, 1998*, vol. 1496 of *Lecture Notes in Computer Science*, pp. 1242–1249, Springer, Berlin, Germany, 1998.
- [114] Z. Tu, K. L. Narr, P. Dollar, I. Dinov, P. M. Thompson, and A. W. Toga, “Brain anatomical structure segmentation by hybrid discriminative/generative models,” *IEEE Transactions on Medical Imaging*, vol. 27, no. 4, pp. 495–508, 2008.
- [115] I. Despotovic, E. Vansteenkiste, and W. Philips, “Brain volume segmentation in newborn infants using multi-modal MRI with a low inter-slice resolution,” in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10)*, pp. 5038–5041, Buenos Aires, Argentina, September 2010.
- [116] J. M. Löjtönen, R. Wolz, J. R. Koikkalainen et al., “Fast and robust multi-atlas segmentation of brain magnetic resonance images,” *NeuroImage*, vol. 49, no. 3, pp. 2352–2365, 2010.
- [117] C. Vijayakumar and D. Gharpure, “Development of image-processing software for automatic segmentation of brain tumors in MR images,” *Journal of Medical Physics*, vol. 36, no. 3, pp. 147–158, 2011.
- [118] L. Gui, R. Lisowski, T. Faundez, P. S. Hüppi, F. Lazeyras, and M. Kocher, “Morphology-driven automatic segmentation of MR images of the neonatal brain,” *Medical Image Analysis*, vol. 16, no. 8, pp. 1565–1579, 2012.
- [119] A. Ortiz, J. M. Gorriz, J. Ramirez, and D. Salas-Gonzalez, “Improving MR brain image segmentation using self-organising maps and entropy-gradient clustering,” *Information Sciences*, vol. 262, pp. 117–136, 2014.
- [120] S. K. Warfield, M. Kaus, F. A. Jolesz, and R. Kikinis, “Adaptive, template moderated, spatially varying statistical classification,” *Medical Image Analysis*, vol. 4, no. 1, pp. 43–55, 2000.
- [121] 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.
- [122] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, “A unifying framework for partial volume segmentation of brain MR images,” *IEEE Transactions on Medical Imaging*, vol. 22, no. 1, pp. 105–119, 2003.
- [123] P. Santago and D. H. Gage, “Quantification of MR brain images by mixture density and partial volume modeling,” *IEEE Transactions on Medical Imaging*, vol. 12, no. 3, pp. 566–574, 1993.
- [124] L. Nocera and J. C. Gee, “Robust partial volume tissue classification of cerebral MRI scans,” in *Medical Imaging: Image Processing*, vol. 3034 of *Proceedings of SPIE*, pp. 312–322, February 1997.
- [125] 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.
- [126] IBSR, “The Internet Brain Segmentation Repository,” 2013, <http://www.nitrc.org/projects/ibsr>.

- [127] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [128] W. R. Crum, O. Camara, and D. L. G. Hill, "Generalized overlap measures for evaluation and validation in medical image analysis," *IEEE Transactions on Medical Imaging*, vol. 25, no. 11, pp. 1451–1461, 2006.

Research Article

Particle Swarm Optimization Algorithm for Optimizing Assignment of Blood in Blood Banking System

Micheal O. Olusanya, Martins A. Arasomwan, and Aderemi O. Adewumi

School of Mathematics, Statistics, and Computer Science, University of Kwazulu-Natal, Private Bag Box X54001, Durban 4000, South Africa

Correspondence should be addressed to Aderemi O. Adewumi; laremtj@gmail.com

Received 9 June 2014; Accepted 20 October 2014

Academic Editor: David A. Elizondo

Copyright © 2015 Micheal O. Olusanya 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 reports the performance of particle swarm optimization (PSO) for the assignment of blood to meet patients' blood transfusion requests for blood transfusion. While the drive for blood donation lingers, there is need for effective and efficient management of available blood in blood banking systems. Moreover, inherent danger of transfusing wrong blood types to patients, unnecessary importation of blood units from external sources, and wastage of blood products due to nonusage necessitate the development of mathematical models and techniques for effective handling of blood distribution among available blood types in order to minimize wastages and importation from external sources. This gives rise to the blood assignment problem (BAP) introduced recently in literature. We propose a queue and multiple knapsack models with PSO-based solution to address this challenge. Simulation is based on sets of randomly generated data that mimic real-world population distribution of blood types. Results obtained show the efficiency of the proposed algorithm for BAP with no blood units wasted and very low importation, where necessary, from outside the blood bank. The result therefore can serve as a benchmark and basis for decision support tools for real-life deployment.

1. Introduction

Blood is a living tissue of unique medical value to the human body [1]. It is responsible for carrying the substances needed for healthy living to various parts of the body and carries away those not needed as wastes. A normal blood is made up of four different components, namely: red cells, white cells, platelets, and plasma [1], which serve separate functions in human organism and different use in the medical treatment of patients in the hospitals or health clinics. The need for blood every day in the hospitals for various reasons and uses in the treatment of patients has made it become of high demand. As a result, there exists collection sites where blood is collected from donors, processed, and shipped to a Hospital Blood Bank to be stored and be available to meet the demands for transfusions to patients. Handling blood transfusion in hospitals involves some complexities due to blood compatibility issues and stochastic nature of the daily demands for blood. Therefore, assignment of blood can be

seen as NP-Hard problem. As a limited resource with limited shelf life, there is need for its efficient management during the process of assigning it to patients from hospitals blood banks. The component of blood considered in this paper is the red cells and it is made up of four main types, namely: A, B, AB, and O. For convenience, blood type AB will be denoted by C in this work. With the Rhesus factor [2], the number of blood types is doubled, resulting in A^+ , A^- , B^+ , B^- , C^+ , C^- , O^+ , and O^- . Optimizing the assignment of these components is the focus of this paper.

The assignment problem is one of the basic combinatorial optimization problems in the fields of discrete optimization. Many optimization techniques have been used to solve various assignment problems including vehicle assignment problem, transportation problem, task allocation problem (TAP), and blood assignment problem (BAP) [3–5]. For instance, Jean et al. [3] used PSO to solve the problem of allocating a set of cabs to some customers with the goal of minimizing the distance traveled by the fleet with result that

showed that PSO is capable of achieving optimal results. Similarly, a discrete PSO (DPSO) and genetic algorithm (GA) were compared for the problems of finding optimal solution to the allocation of the expected number of people in flooded areas to various types of available vehicle in an evacuation process [6] with DPSO having better performance than GA. DPSO has also shown great performance with different degrees of difficult knapsack problems [7, 8]. In addition, experimental results in [9] showed that DPSO algorithm is highly efficient for solving the multiple knapsack problems (MKP) even with large problem instances. This motivates the investigation of DPSO in solving the BAP defined in this paper.

BAP recently stated attracting interests among meta-heuristics and optimization researchers [2, 10–14] as the sourcing for blood for transfusion purpose is currently posing global challenge. An early work reported in [10] modelled the problem as a multiobjective linear programming model to help determine the best assignment of blood from donors to requests. The model was based on variables that represent the units of blood coming from both within the blood bank and outside the system and assigned available blood on daily basis depending on demand. A case study of Italian Red Cross System in Rome was considered. The model is however incapable of handling global request of different blood types at the same time. In [2] a dynamical system model was developed to handle the management of blood in a blood bank but without considering Rhesus factors. The model in [2] was improved upon in [11] where the authors modelled BAP as a MKP and compared the performances of five different (meta)heuristics in solving the problem. The goal was to minimize the amount of blood unit imported from outside the blood banking system. To further improve on the work in [11], a new cross and mutation techniques were introduced into the algorithm in [12] with the result that GA was reported as more efficient than other compared techniques. In [13], two local search optimization techniques, namely, dynamic programming and greedy randomized adaptive search procedure (GRASP) were tested for the BAP with GRASP performing better than dynamic programming. Tabu search and simulated annealing with their hybrid were also tested on BAP [14] where the hybrid showed better results than the individual techniques.

The problem of assigning blood in blood bank system principally involves attending to requests for blood and accepting donated blood from donors with the lifespan of each blood type of critical considerations. These activities are on daily basis and stochastic in nature. In situations where there is not enough volume of compatible blood type(s) to meet demands, such blood types are imported from other sources outside the blood bank system and this could be very expensive. All these put together show that the objective of assignment of blood in blood banks is to meet the daily demand for blood and to minimize the volume as well as the number of blood types imported from external source(s). Also, the volume of expired blood should be minimized. There is no report of the application of any swarm intelligence algorithms to this new domain of BAP. This paper is therefore set to investigate the application of a well-known swarm intelligence algorithm, particle swarm optimization (PSO), to the BAP.

TABLE 1: Blood type compatibility with Rhesus factor.

Receiver	Donor							
	A ⁺	A ⁻	B ⁺	B ⁻	C ⁺	C ⁻	O ⁺	O ⁻
A ⁺	YES	NO	NO	NO	NO	NO	YES	YES
A ⁻	NO	YES	NO	NO	NO	NO	NO	YES
B ⁺	NO	NO	YES	YES	NO	NO	YES	YES
B ⁻	NO	NO	NO	YES	NO	NO	NO	YES
C ⁺	YES							
C ⁻	NO	YES	NO	YES	NO	YES	NO	YES
O ⁺	NO	NO	NO	NO	NO	NO	YES	YES
O ⁻	NO	YES						

PSO was introduced in [15, 16]. It is a fast and efficient technique used in solving complex and simple optimization problems. This technique is nature-inspired and draws inspiration from social behavior of animals, like birds and fishes. Thus, it belongs to the swarm intelligence group of evolutionary computation techniques. From the period of its inception, it has been used by researchers to solve a wide range of simple and complex optimization problems [17–24].

In the sections that follow, Section 2 gives the description of BAP considered in this paper. Section 3 describes the method that was applied in solving the problem. Section 4 focuses on the numerical simulations conducted which comprises the analysis and discussion of results. Finally, Section 5 concludes the paper and identifies some possible areas of future research.

2. Problem Description for Blood Assignment

The blood bank stores and issues the appropriate blood units to satisfy transfusion requests. On each day the bank receives a random number of transfusion requests for each blood type and each request for a random number of units. Once a request is received, the appropriate number of units of that type is removed from the bank upon successful cross-matching. We will define demand to be the number of units requested and usage to be the number of units transfused. Any units which are not used within their shelf life are considered expired and are discarded from the bank. Normally, a patient should be transfused his or her own blood type when there is a need for it. However, there could be situations when the blood type in the bank may not be enough to meet the blood units of the type requested. In order to address such challenges compatible blood types are supplied instead, where possible. The compatibility between blood types is presented in Table 1 (in this paper, blood types “C⁺” and “C⁻” refer to AB⁺ and AB⁻ throughout and are thus chosen for typo simplicity). In the table, “YES” means that there is compatibility between the donor and receiver, while “NO” means there is no compatibility. Also, in the table, it is evident that O⁻ is the universal donor, while C⁺ is the universal receiver. In Republic of South Africa, 86% of the population has positive Rhesus factor, while 14% have negative Rhesus factor leading to the repartition presented

TABLE 2: Proportion of blood types in South Africa².

Blood type	A ⁺	A ⁻	B ⁺	B ⁻	C ⁺	C ⁻	O ⁺	O ⁻
Proportion (%)	32	5	12	2	3	1	39	7

²Source: South African National Blood Service—<http://www.sanbs.org.za/index.php/donors/what-s-your-type>.

in Table 2 [2, 11–14]. Further description of the BAP can be found in [2, 10, 12].

From the preceding description, it is evident that the assignment of blood in blood banks is a complex problem. Furthermore, it becomes more complex because blood products have limited shelf life. As a result, there is therefore the need for algorithms that could efficiently assign available blood resources in the blood bank to receivers, in order to minimize the blood units imported from external source(s) into the blood bank and at the same time minimize wastage of blood unit because of their limited shelf life.

2.1. Objective Function. The objective function of the BAP is given in (4). The aim is to minimize the volume of red blood cells imported into the bank over a period of n days:

$$\min \sum_{t=1}^n I_{\text{Total}}(t), \quad (1)$$

where $I_{\text{Total}}(t) = I_{O^+}(t) + I_{O^-}(t) + I_{A^+}(t) + I_{A^-}(t) + I_{B^+}(t) + I_{B^-}(t) + I_{C^+}(t) + I_{C^-}(t)$ and I_{O^+} , I_{O^-} , I_{A^+} , I_{A^-} , I_{B^+} , I_{B^-} , I_{C^+} , and I_{C^-} are volumes of the respective blood types imported in day t .

2.2. Constraints. Description of different constraints that are to be satisfied in implementing the proposed algorithm to solve the BAP is as follows.

- (i) The total volume of blood requested must be satisfied for each day, either by using the ones in the bank or by importing blood from external sources.
- (ii) The volume of blood assigned each day must not exceed the total volume available in the bank.
- (iii) The minimum volume of each blood type must not be less than zero.
- (iv) Each unit of blood types must not exceed 30 days of shelf life; otherwise, it should be incinerated.

2.3. Variables. Explained below are the descriptions of possible variables that can be used in the algorithm and could be represented using any meaningful variable names:

- (i) volume of blood types in storage at day t ;
- (ii) proportion of blood types in the population of the region considered;
- (iii) donations of blood into the bank at day t ;
- (iv) requests for unit blood at time t ;
- (v) volume of incinerated expired blood at day t ;
- (vi) volume of blood supplied at day t ;
- (vii) volume of blood imported from external source(s) into the bank at day t .

2.4. Updating Blood Volume. On daily basis, the volumes of the blood types are updated relative to the volumes supplied, donated, and expired as stated in (15).

2.5. Assumptions. The model was defined with the following assumptions.

- (i) There is sufficient supply from the external source(s) of blood, when units are required to be imported.
- (ii) The desired level of the volumes of each blood type is relative to the proportional of blood distribution in the region under consideration.
- (iii) If emergency arises, no optimization may be performed, but patients straightaway receive O⁻ blood.
- (iv) The validity date of the blood products is 30 days.

3. Methodology

In the daily management of the blood bank, the demand and supply of blood units vary; thus, there is a need for an effective method to determine the assignment of blood units for the purpose of good management. The assignment of blood units would be considered optimal if there is no importation of units from external sources; however, this may not be a possibility. Therefore, an algorithm that could be able to assign blood units relative to the available volume of blood units and volume requested with no or minimum units imported from outside the blood bank will suffice.

Three techniques are combined to provide solution to the BAP considered and thus to determine (near) optimal assignment of blood units relative to demands. They are particle swarm optimization which is the optimization technique used, multiple knapsack assignment technique which is used to handle the cross-matching of blood types in order to satisfy requests and stabilize the proportions of blood types stored in the bank, and queuing technique which is used to monitor the expiration date of the units of each blood type. Also described in this section is a bottom-up technique that was used to assign the blood units before importing any blood types from external source(s).

3.1. Particle Swarm Optimization Technique. The PSO as a population-based and stochastic technique needs a swarm of particles to carry out its optimization process [15, 16]. Using the search range defined for a problem, values are randomly generated for the decision variables. These values are then used to randomly distribute the particles in the solution search space before the technique begins its iterative process. During the optimization process, each particle communicates its new discoveries to others and this in turn determines subsequent moves of the particles in the search space. In each attempt of iteration, each particle makes use of two major pieces of information: its personal experience and the experiences of reachable neighbours to guide its search. Furthermore, the objective function of the problem being optimized is used to evaluate the quality of the discovery of each particle. Given an n -dimensional space, each particle is characterized by the position vector $X_i = (x_{i1}, \dots, x_{in})$ and

Begin PSO Algorithm**Input:** f : the function to optimize s : the swarm size d : the problem dimension X_{\min}, X_{\max} : decision variable search range V_{\min}, V_{\max} : particle velocity limits**Output:** x^* : the best particle position found (global best) f^* : the best fitness value found**Initialize:** position $x_i = (x_{i1}, \dots, x_{id})$ and velocity $v_i = (v_{i1}, \dots, v_{id})$, for all particles in problem space evaluate $f(x_i)$ in d variables and get $pbest_i, (i = 1, \dots, s)$ $gbest \leftarrow$ best of $pbest_i$ **While** stopping criteria is false **do** Compute inertia weight (ω) if it is not a constant Repeat for s times Repeat for d times update v_i for particle using (2)

validate for velocity boundaries using Algorithm 3

 update x_i for particle using (3)

validate for position boundaries using Algorithm 2

 compute $f(x_i)$ End Repeat for d compute $f(x_i)$ obtain new $pbest_i$ If $f(x_i) < f(pbest_i)$ then $pbest_i \leftarrow x_i$ If $f(x_i) < f(gbest)$ then $gbest \leftarrow x_i$ $f(gbest) \leftarrow f(x_i)$

end if

 End Repeat for s **End while** $x^* \leftarrow gbest$ $f^* \leftarrow f(gbest)$ Return x^* and f^* **End PSO Algorithm**

ALGORITHM 1

the velocity vector $V_i = (v_{i1}, \dots, v_{im})$. When the particles are searching for optimum solution in the search space, their velocities and positions are updated using (2) and (3), respectively:

$$V_i(t+1) = \omega V_i(t) + c_1 r_1 (P_i - X_i) + c_2 r_2 (P_g - X_i), \quad (2)$$

$$X_i(t+1) = X(t) + V_i(t+1). \quad (3)$$

In (2), P_i and P_g are vectors representing the i th particle personal best position and swarm global best position, respectively; r_1 and r_2 are random numbers in the interval $[0, 1]$, while c_1 and c_2 are acceleration coefficients called cognitive and social scaling parameters which determine the extent of the random forces in the direction of P_i and P_g . The parameter t represents iteration index, while ω is the inertia weight which was introduced in [25]. It regulates the particle's velocity and helps to balance the global and local search abilities of PSO. In the original PSO algorithm, though not explicitly added to (2), the parameter ω implicitly had a value of 1. A general framework of PSO algorithm for continuous problems is presented in Algorithm 1.

Although the PSO was initially developed for continuous optimization problems, it has in several occasions been customized for discrete problems [19, 20]. Customizing PSO for discrete problems was initially proposed by Kennedy and Eberhart in [26] where they defined trajectories and velocities of particles in terms of changes of probabilities, where the particles move in a state space restricted to 0 and 1 on each dimension, with a certain probability computed using the sigmoid limiting transformation as shown in

$$x_{ij} = \begin{cases} 1, & \text{if } \mu \leq S(v_{ij}) \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

where x_{ij} is the value in the j th dimension of the position of particle i , v_{ij} is the value in the j th dimension of the velocity of particle i , μ is a value selected from a uniform distribution in $[0, 1]$, and $S(v_{ij})$ is as given in

$$S(v_{ij}) = \frac{1}{1 + e^{-v_{ij}}}. \quad (5)$$

In this paper, some customization was made to the PSO algorithm before it was applied to the BAP. The different parts that were customized are as follows.

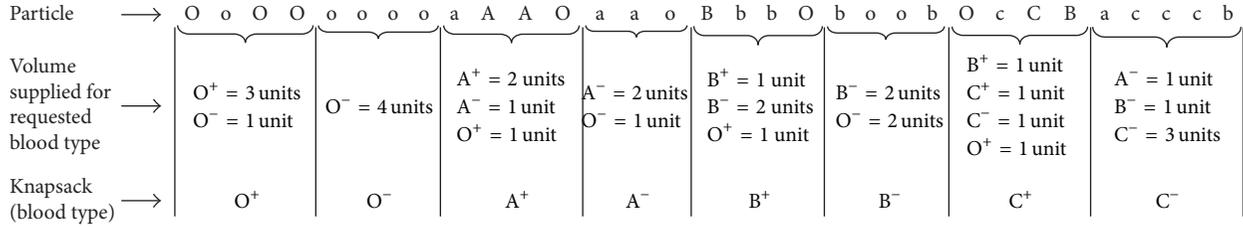


FIGURE 1: Exemplified particle representation in the proposed discrete PSO algorithm.

TABLE 3: Transformed particle used for computation in (2).

0.07	0.04	0.07	0.07	0.04	0.04	0.04	0.04	0.07	0.15	0.15	0.07	0.07	0.07	0.04	0.15	0.07	0.07	0.07	0.07	0.04	0.04	0.07	0.07	0.15	0.3	0.15	0.07	0.15	0.15	0.15	0.07
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	------	------	------	------	------	------

TABLE 4: Values and proportions of the different blood types.

Blood type	A ⁺	A ⁻	B ⁺	B ⁻	C ⁺	C ⁻	O ⁺	O ⁻
Value	4/27 = 0.15	2/27 = 0.07	4/27 = 0.15	2/27 = 0.07	8/27 = 0.3	4/27 = 0.15	2/27 = 0.07	1/27 = 0.04
Proportion (%)	32	5	12	2	3	1	39	7

3.1.1. *Particle Representation.* Each particle is encoded by a long string of characters representing each of the blood types. The string is divided into groups, with each group representing the content of each of the 8 knapsacks for each blood type. The characters in each group show which units of blood have been placed into each knapsack. Furthermore, each character represents 1 unit of positive blood type if it is uppercase (e.g., “A” for “A⁺”), but represents 1 unit of negative blood type if it is lowercase (e.g., “a” for “A⁻”). An example is shown in Figure 1.

The example in Figure 1 shows an interpretation of the representation of a typical particle in the swarm. The particle is of size (dimension) 32 and it represents the combination of all the units of requested blood types. The first 4 dimensions represent the number of units (i.e., 4) of O⁺ that was requested and the supply is made up of 3 units of O⁺ and 1 unit of O⁻. The last 5 dimensions also represent the number of units (i.e., 5) of C⁻ that was requested and the supply is made up of 1 unit of A⁻, 1 unit of B⁻, and 3 units of C⁻. The same idea of interpretation applies to the remaining dimensions of the particle. In the course of implementation of the technique, the dimension of the particles and the size of the various knapsacks vary each day relative to the changing requests received by the blood bank.

3.1.2. *Transformation of Particles.* From Figure 1, it is clear that the values of the dimensions of the particle are alphabets. Therefore, X_i , P_i , and P_g will contain alphabets too and these cannot be directly used to compute $V_i(t + 1)$ in (2). For the velocity of each particle to be computed, these alphabets should be converted to numbers. There may be different methods to do this; however, the method used was to replace the alphabet (blood type) in each dimension of the particle with its value (see Table 4). In other words, the particle represented in Figure 1 becomes as represented in Table 3.

P_i and P_g are also transformed likewise, depending on the contents of their respective dimensions; hence, (2) can be computed to obtain a value for $V_i(t + 1)$.

3.1.3. *Obtaining the Positions of Particles.* The value obtained for $V_i(t + 1)$ is then transformed using (5) to get a value between 0 and 1. This value is then used to determine the blood type to be imported into the blood bank when there is a shortage of blood to meet requests. In other words, the different compatible blood types (including the blood type requested) that could be used to meet the request are given equal chances to be imported from external source(s) into the blood bank. This means that if there are two compatible blood types, each is given 50% chance to be imported, if there are four compatible blood types, each is given 25% chance of being imported, and so forth. Before the beginning of next iteration, each particle is reverted back to a string of alphabets of blood types representing the new potential solution.

3.2. *Multiple Knapsack Model.* The multiple knapsack problems involve selecting any of n items and packing them into m knapsacks of different capacity c in order to obtain the largest sum of profit [11]. The standard model for multiple knapsack problems is stated in (6)–(9). Consider

$$\text{maximize } \sum_{i=1}^m \sum_{j=1}^n p_j x_{ij} \tag{6}$$

$$\text{Subject to } \sum_{j=1}^n w_j x_{ij} \leq c_i \quad i = 1, \dots, m, \tag{7}$$

$$\sum_{i=1}^m x_{ij} \leq 1 \quad j = 1, \dots, n, \tag{8}$$

$$x_{ij} \in \{0, 1\} \quad i = 1, \dots, m; \quad j = 1, \dots, n. \tag{9}$$

Equation (6) is the objective function, which is to maximize the profit of the number of items placed in all the knapsacks; (7) is the constraint that ensures that the capacity of each knapsack is not exceeded; (8) is the constraint that ensures that each item can only be placed in any of the knapsacks once

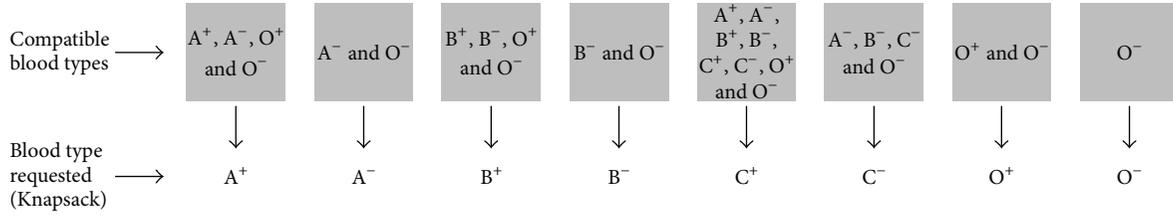


FIGURE 2: Representing blood type compatibilities as knapsack problem.

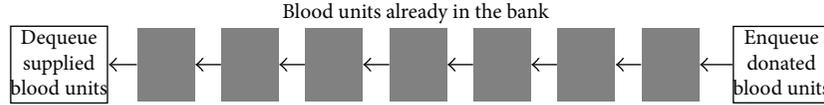


FIGURE 3: Representation of how blood units are stored in the blood bank.

and (9) indicates whether an item j is placed into knapsack i or not.

Application of the Technique from Multiple Knapsack Problems (MKP). Inspired by the various successes in the application of MKP to different problems, the technique was applied to the BAP considered in this paper with some modifications. Equations (10)–(13) represent the way MKP was applied. In the case of the assignment of blood, units of blood types stored in the bank are the items, while different requests for the blood types are the knapsacks and the number of blood types compatible with each blood type is the capacity of each knapsack. Consider

$$\text{minimize } \sum_{i=1}^m \sum_{j=0}^n p_j x_{ij} \quad (10)$$

$$\text{Subject to } T = S + \sum_{j=0}^n p_j \quad (11)$$

$$x_{ij} \in \{0, 1\} \quad i = 1, \dots, m; \quad j = 1, \dots, n \quad (12)$$

$$S, x_{ij} \geq 0, \quad i = 1, \dots, m; \quad j = 1, \dots, n, \quad (13)$$

where T is the total units of blood supplied from the blood bank and/or imported from external source(s), S is the total unit of blood supplied from the bank in each day, p_j is units (or volume) of blood type j coming from external sources, m is the total number of requests (number of knapsacks) to be met, and n is the total number of blood types available in the bank for supply (meet requests).

Equation (11) certifies that the total blood units supplied are given by the amount of blood units from the bank and external sources. Equation (12) shows whether there is a supply of blood “1” or not “0.” Equation (13) is a nonnegativity constraint; that is, all the values S and x_{ij} are positive integers.

In applying this technique to the BAP, some additional constraints are to be satisfied. These constraints are stated as follows.

- (i) In each day, there are 8 sets of requests for the various blood types (A⁺, A⁻, B⁺, B⁻, C⁺, C⁻, O⁺, and O⁻) which correspond to 8 knapsacks.
- (ii) To satisfy requests, units of blood within the bank or imported from external source(s) must be placed in each knapsack.
- (iii) Each knapsack can only be satisfied with a compatible unit of blood. This is illustrated in Figure 2.

It should be noted that the weight w_j of every unit of blood of any of the blood type is equal to 1. As a result, selecting the most profitable units that would satisfy each request in order to maximize profit is of uttermost interest.

3.3. Queuing Technique. In the daily management of the blood bank, it is ensured that the shelf life of each unit of blood is not exceeded in order to minimize wastage. An efficient way which can be monitored is to use the queue technique. Queue is a list-like structure that provides restricted access to its elements. Elements may only be inserted at the back (enqueue) and removed from the front (dequeue). It implements the First-In, First-Out (FIFO) operations; thus, queues release their elements in order of arrival. This means that when units of blood are supplied from the bank, they are dequeued but enqueued when units of blood are donated into the bank. This is illustrated in Figure 3, using a single-linked list. It should be noted that each of the blood type is represented separately.

Represented in (14)–(16) is the way the stack in Figure 3 is operated in the algorithm:

$$S_j = B_{\text{tot}} - \bigcup_{i=0}^r R_{ij} \quad j = 1, \dots, n \quad (14)$$

$$B_{\text{tot}}(t+1) = B_{\text{tot}}(t) \cup \bigcup_{x=0}^d D_{xy} \quad y = 1, \dots, m \quad (15)$$

$$R_{ij}, D_{xy} \geq 0. \quad (16)$$

In (14) and (15), S_j is the total blood units of type j supplied by the bank, r is the total number of requests of all the blood

TABLE 5: Various values used for defining the different datasets used in the simulation.

Dataset	Initial total volume of blood in bank	Running time (days)	Upper and lower bounds for requests (%)	Upper and lower bounds for donations (%)	Remark
1	500	90	[25, 75]	[25, 75]	These datasets were used to test the algorithm with different initial volume
2	1000	90	[25, 75]	[25, 75]	
3	2000	90	[25, 75]	[25, 75]	
4	1000	90	[25, 75]	[30, 75]	These datasets were used to test the algorithm when the ratios of requests to donations are unequal
5	1000	90	[30, 75]	[25, 75]	
6	500	365	[25, 75]	[25, 75]	These datasets were used to test the volume of blood units that will expire when the bank is run for a year
7	1000	365	[25, 75]	[25, 75]	

types, R_{ij} is the total blood units of type j requested, B_{tot} is the total units of all the blood types in the bank, t represents the current day of blood bank operation, d is the total units of all the blood types donated, and D_{xy} is the total blood units of type y donated.

Expiration of Blood Units. As mentioned earlier, the queue is used to monitor the shelf life of each unit of the different blood type. Any of the blood units of a particular blood type that exceeds its shelf life are removed from the queue. This is represented in (16). Consider

$$B_{tp}(t+1) = B_{tp}(t) - \bigcup_{i=1}^z (E_i, f), \quad \forall f > \text{life}, \quad (17)$$

where $tp \in \{A^+, A^-, B^+, B^-, C^+, C^-, O^+, O^-\}$ is the blood type, B_{tp} is the total blood units of type tp on queue, t is the current day of operation, and E_i is the expired blood unit i if its life, f , on the queue is greater than its shelf life, life .

3.4. Bottom-Up Technique. Some blood types are compatible while others are not, as shown in Figure 1. Thus, when there is a request for some units of a particular blood type, the request could be met using the same blood type if there are enough of its units in the bank. If there are not enough units (shortage) of the type requested, available units of compatible types are used to meet up the remaining units. In this case, a bottom-up approach was used in the course of assigning blood to meet the request. This means that, whenever blood is to be assigned, the requested blood type is considered first; when there is a shortage the next available compatible blood type is assigned, and so forth. Blood type O^- is the last option, being the universal donor; if it is not available, then blood will be imported into the bank from external source(s) relative to the request. This approach is exemplified as follows.

Given that $typREQ$ is the number of units requested of blood type typ and $typVOL$ is the number of units of the same blood type that is available in the bank to meet the request, if $typREQ > typVOL$, then there is a shortage of the blood type to meet the request.

4. Numerical Simulations

The proposed PSO algorithm proposed for solving the considered problem was investigated, while implementing

the modified multiple knapsack problem, with a number of simulations with different datasets. As a result of the nonavailability of real-life data, randomly generated data were used. The stochastic nature of blood donation and requests in real-life scenario informed the usage of the fictitious randomly generated data. Whenever real-life data becomes available, they can be substituted for the fictitious data. The application software was developed in Microsoft Visual C# programming language.

4.1. Data Generation. All the data used were randomly generated based on the uniform random number generator. Given a specified (by the user) upper and lower bounds, random volume of requests and donation for each blood type is computed in each day. This provided the opportunity of testing the blood assignment system developed under various scenarios of varied volumes of requested and donated-blood types. Furthermore, an initial volume of all the blood types in the bank was also specified and used to calculate respective volume for each blood type using some blood type proportion. In this case, the blood type proportion in South Africa was used (see Table 1). All the parameters and their respective values as used in the algorithm to generate the testing data are presented in Table 5.

4.2. Parameter Setting. Presented in this section are the definitions of all the parameters used in the course of implementing the proposed model and algorithm. In Table 4, the different blood types and their respective value and proportion are presented. In Table 4, "Value" was computed based on the information in Figure 1. The numerator is the number of compatible blood types that can be received by the respective "Blood type," while the denominator is the sum of all the compatible blood types.

For the PSO technique, the values for c_1 and c_2 were, respectively, set to 1.7 and parameter ω was set to 0.715, as recommended in [27]. The parameters r_1 and r_2 were randomly generated using the uniform random number generator. A swarm size of 50 particles was used and the maximum number of iteration was set to 1000.

4.3. Results and Discussion. Presented in Tables 6, 7, 8, 9, 10, 11, and 12 are the average blood units that were available in, requested for, supplied by, and imported into the blood

TABLE 6: Average volume of blood units during a runtime of 90 days with 500 blood units as initial volume (using dataset 1).

Average volume	O ⁺	O ⁻	A ⁺	A ⁻	B ⁺	B ⁻	C ⁺	C ⁻
Available in bank	114.06	17.36	109.57	19.47	41.63	8.31	21.02	4.96
Requested	66.79	11.87	54.43	8.41	19.87	3.31	5.08	1.77
Supplied	65.33	11.97	54.32	8.49	19.87	3.42	5.08	1.57
Imported into bank	0.42	1.03	0.00	0.01	0.00	0.01	0.00	0.00

TABLE 7: Average volume of blood units during a runtime of 90 days with 1000 blood units as initial volume (using dataset 2).

Average volume	O ⁺	O ⁻	A ⁺	A ⁻	B ⁺	B ⁻	C ⁺	C ⁻
Available in bank	231.34	34.59	216.72	40.36	80.79	12.56	45.67	15.77
Requested	134.59	23.87	109.73	16.88	40.10	6.68	10.23	3.31
Supplied	131.66	24.19	109.47	17.09	39.96	6.81	10.23	2.97
Imported into bank	0.90	2.08	0.00	0.02	0.00	0.00	0.00	0.02

TABLE 8: Average volume of blood units during a runtime of 90 days with 2000 blood units as initial volume (using dataset 3).

Average volume	O ⁺	O ⁻	A ⁺	A ⁻	B ⁺	B ⁻	C ⁺	C ⁻
Available in bank	460.20	69.17	430.09	80.11	159.83	26.31	90.77	30.54
Requested	267.53	47.46	218.10	33.68	79.71	13.24	20.42	6.54
Supplied	261.86	48.09	217.62	33.91	79.47	13.67	20.42	5.90
Imported into bank	1.69	4.01	0.00	0.06	0.00	0.00	0.00	0.00

bank over the period of 90 and 365 days, using all the datasets. Figures 4–7 show the various curves associated with the management of the blood bank as observed during the simulations.

From the results, it is evident that managing the blood bank with each of the datasets over the number of specified periods (90 and 365 days) led to average importation of very few blood units from external sources. The importation became necessary when there were shortages of compatible blood types to meet requests. In Table 10, A⁺ blood type was imported into the blood bank when dataset 5 was used (with unequal ratio of requests to donations of blood units). For A⁻ blood type, apart from Table 9 (using dataset 4), some of its units were imported into the bank when the other datasets were used and the average number of units imported was in increasing order with small variations. Using any of the datasets, no units of B⁺ were imported. Some units of B⁻ were imported in Tables 6, 10, 11, and 12 (using datasets 1, 5, 6, and 7). No C⁺ blood type was imported into the bank using any of the datasets. Some units of C⁻ blood type were imported in Tables 7, 11, and 12 (using datasets 2, 6, and 7).

Much importation was observed with O⁺ and O⁻ blood types, but with O⁻, being a universal donor, the importation was more compared to others. Also, when there is a shortage of O⁻ type, no other blood types could be used as alternative(s) for it and thus it must be imported into the bank. The very low level of average importation of blood units from external source(s) is an indication that the bank was efficiently managed. This efficient management was a result of combining the efforts of the multiple knapsack problem technique used to implement the cross-matching between blood types and PSO technique used to spread importation of blood types (where necessary) in assigning blood types to meet requests. Using all the tested datasets, no blood units exceeded their shelf life; thus, no wastage of blood products was experienced. The reason for this is that the cross-matching method helps to quickly use the older blood units which have been on queue to meet requests while the new ones donated into the blood bank are enqueued.

Presented in Figures 4 and 5 are the curves showing the various volume levels of the different blood types in the blood bank when the bank was operated for 90 and 365 days, respectively. The curves show the corresponding total units of blood available, requested, supplied, and imported in each day of operating the bank. In figures, it is evident that minimal units of blood were imported from external source(s) into the bank. In the first few days no units were imported, but between 75th and 85th days a higher number of blood units were imported because the number of units requested was higher than the available blood units relative to the blood types requested. From the graphs, it appears that at the point of importation the volume of available blood in the bank was high and there could be the question, why should there be importation? The reason for importation is that the entire volume of blood in the bank comprised all the units of all blood types and the other units of blood available at the point of importation were not compatible with the blood types requested; therefore, they could not be supplied or used to meet the requests.

Also, in Figure 5, the curves show the corresponding total units of blood available, requested, supplied, and imported in each day of operating the bank for a period of 365 days. In this case, the total units of blood imported (though little) were higher than when the bank was operated for 90 days. In the graphs, it is observed that the highest units of blood were imported between the 340th and 350th days. At this period, some of the available blood types were already exhausted and there were needs for importation of blood units from external source(s) to meet the requests.

Presented in Figure 6 are the curves showing the number of blood units imported when the algorithm was tested with all the datasets. The highest importation of blood units occurred when dataset 4 was used (Figure 6(a)) and when dataset 7 was used (Figure 6(b)); these occurred around the 80th and 350th days, respectively. In Figure 7, the ratios of total units of blood imported to total units of blood requested for both 90 and 365 days of operating the blood bank are presented. For the graphs, the algorithm appears not to have performed very well with datasets 5 and 7.

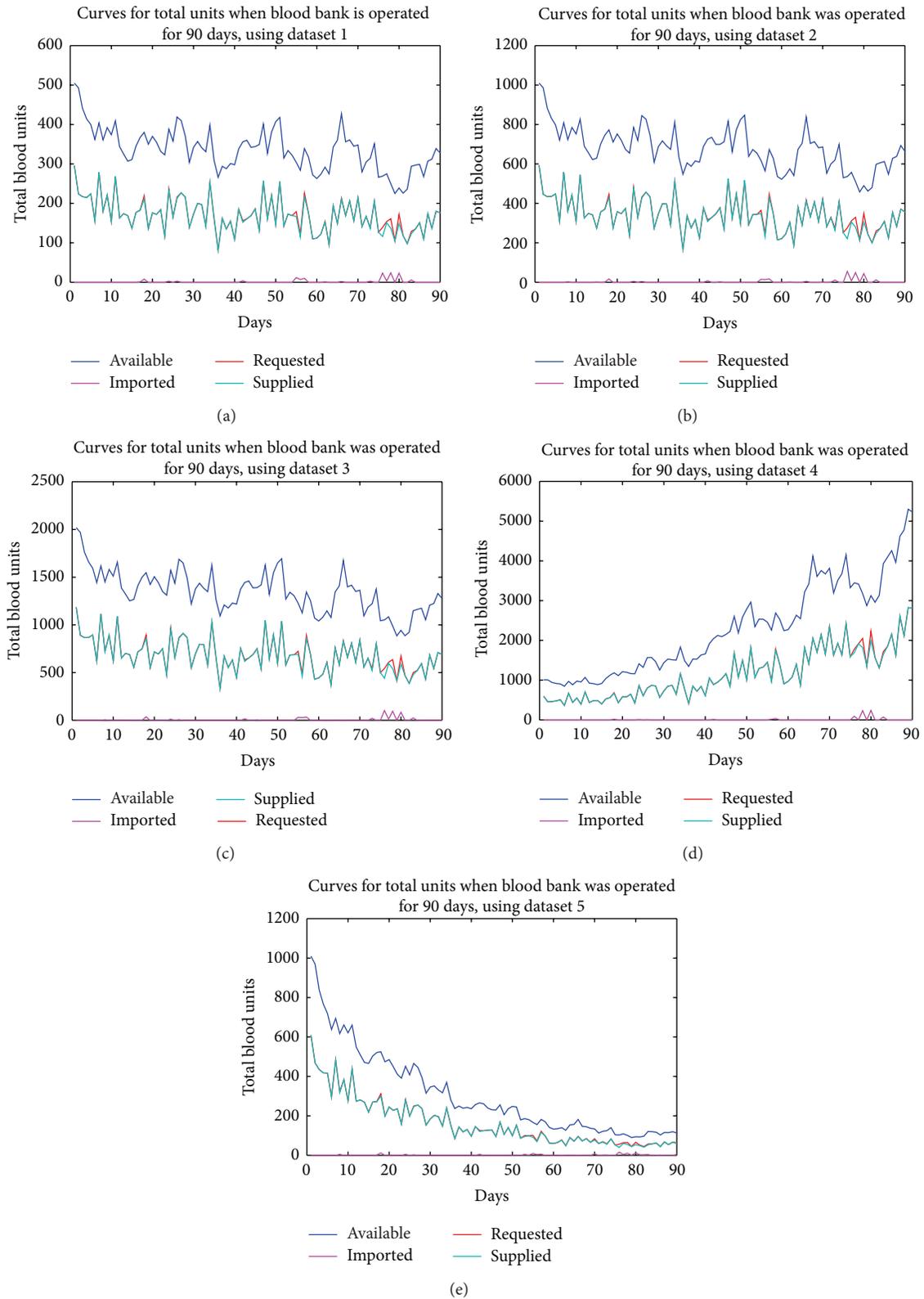


FIGURE 4: Various curves showing the different blood levels in the blood bank over 90 days, using datasets from 1 to 5.

TABLE 9: Average volume of blood units during a runtime of 90 days with 1000 blood units as initial volume (using dataset 4).

Average volume	O ⁺	O ⁻	A ⁺	A ⁻	B ⁺	B ⁻	C ⁺	C ⁻
Available in bank	759.21	126.86	773.10	146.51	260.02	46.24	112.67	62.60
Requested	449.09	80.66	372.94	56.67	135.39	22.91	34.90	10.08
Supplied	439.63	82.37	372.94	56.63	135.39	23.17	34.90	9.59
Imported into bank	2.29	5.72	0.00	0.00	0.00	0.00	0.00	0.00

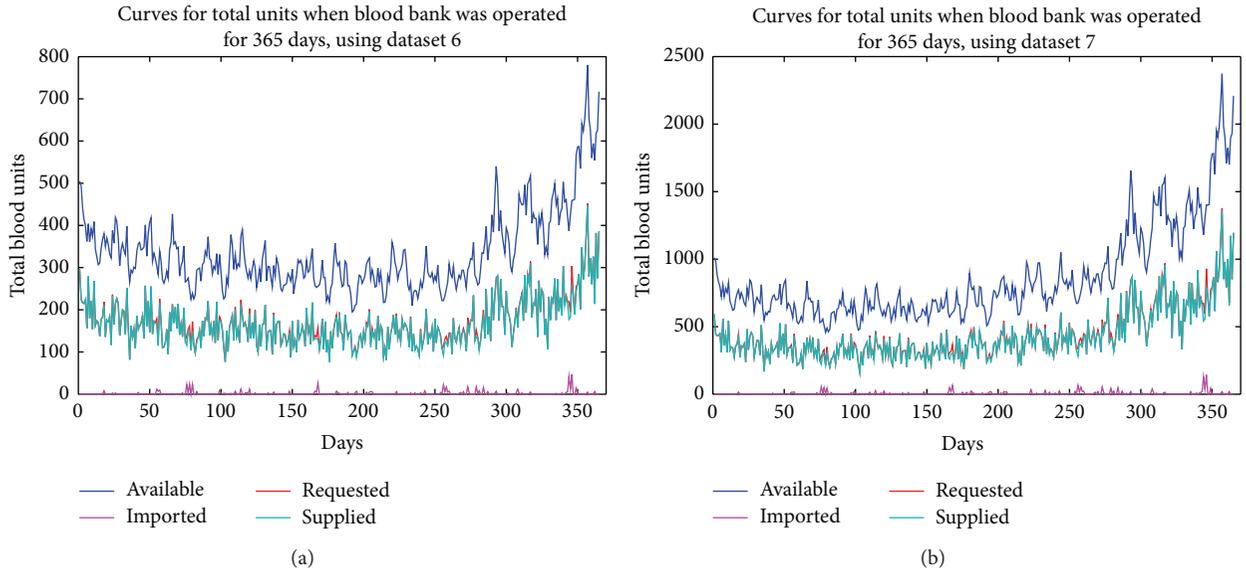


FIGURE 5: Importation rate of blood units using datasets 6 and 7.

TABLE 10: Average volume of blood units during a runtime of 90 days with 1000 blood units as initial volume (using dataset 5).

Average volume	O ⁺	O ⁻	A ⁺	A ⁻	B ⁺	B ⁻	C ⁺	C ⁻
Available in bank	108.80	15.39	82.01	14.78	41.89	6.63	29.31	6.51
Requested	64.07	11.33	51.89	8.11	18.93	3.12	4.79	1.69
Supplied	63.24	11.24	51.52	8.07	18.84	3.28	4.79	1.50
Imported into bank	0.33	0.91	0.02	0.17	0.00	0.01	0.00	0.00

TABLE 11: Average volume of blood units during a runtime of 365 days with 500 blood units as initial volume (using dataset 6).

Average volume	O ⁺	O ⁻	A ⁺	A ⁻	B ⁺	B ⁻	C ⁺	C ⁻
Available in bank	116.71	16.04	88.33	14.69	54.68	6.43	28.35	10.55
Requested	66.05	11.79	53.84	8.37	20.01	3.36	5.07	1.72
Supplied	65.27	11.77	53.04	8.38	19.89	3.37	5.07	1.61
Imported into bank	0.41	1.03	0.00	0.30	0.00	0.05	0.00	0.01

5. Conclusion

The problem of assigning of blood units by blood banks to meet requests for blood transfusion in hospitals has been considered in this paper and an efficient method of handling this problem was proposed. The method that was proposed combined the efforts of the PSO technique and multiple knapsack problem. The multiple knapsack problem was modified to reflect some additional constraints that needed to be satisfied for the blood bank to be efficiently managed. Using the queue, multiple knapsack problem, and optimization (PSO) techniques, the total units of blood types imported into the bank were greatly minimized and no wastage was experienced.

TABLE 12: Average volume of blood units during a runtime of 365 days with 1000 blood units as initial volume (using dataset 7).

Average volume	O ⁺	O ⁻	A ⁺	A ⁻	B ⁺	B ⁻	C ⁺	C ⁻
Available in bank	286.50	39.22	220.99	35.68	140.24	14.68	80.41	43.34
Requested	169.22	30.23	137.99	21.49	51.36	8.65	12.93	4.32
Supplied	166.98	30.05	135.69	21.36	51.05	8.58	12.93	4.09
Imported into bank	1.22	3.08	0.00	0.98	0.00	0.17	0.00	0.01

Only the red blood cells were considered in the paper and no real-life data were used but randomly generated data.

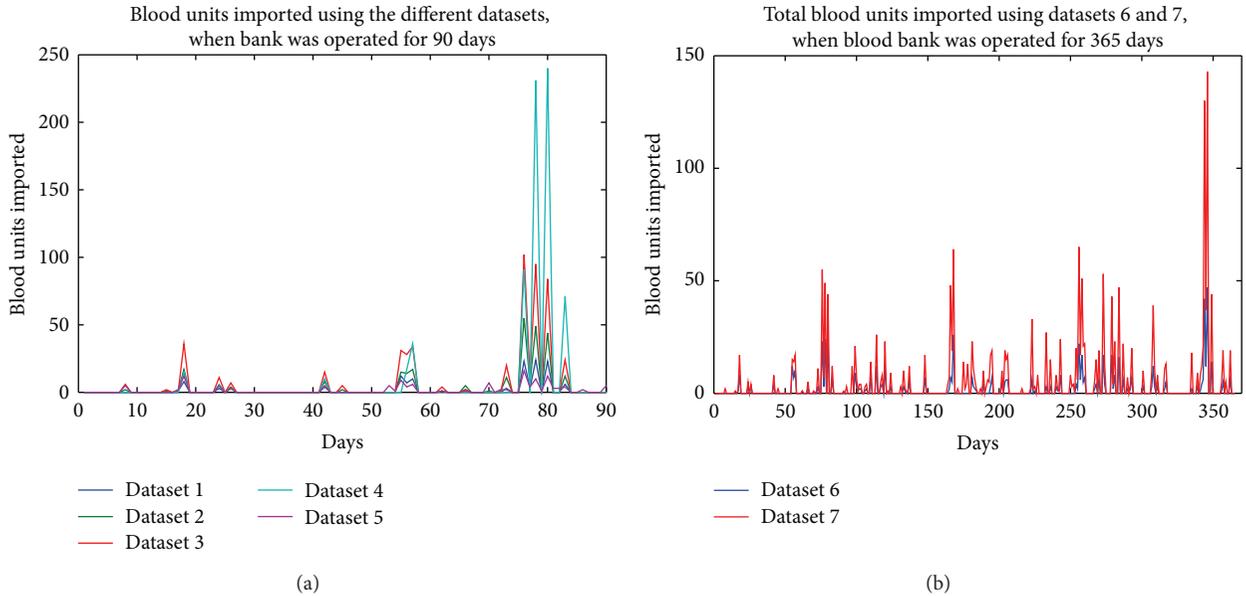


FIGURE 6: Importation rate of blood units using the different datasets.

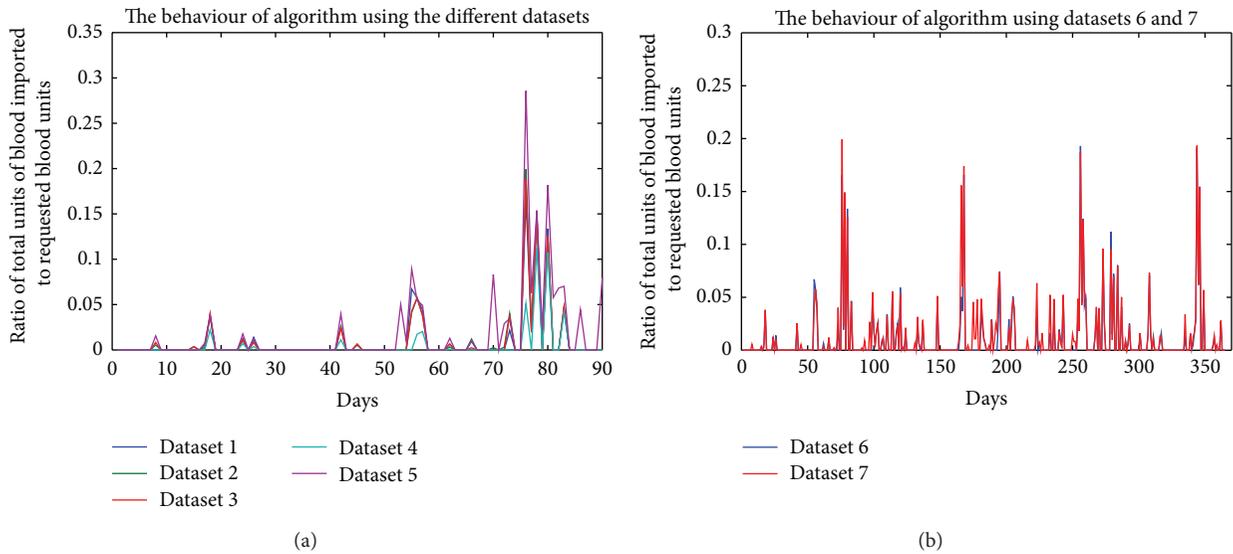


FIGURE 7: Performance measurement of the algorithm using the different datasets.

Therefore, with respect to future work, other components of the blood could be considered and the proposed method could be tested with real-life data.

Conflict of Interests

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

References

[1] G. P. Prastacos, *Blood-Management System: An Overview of Theory and Practice*, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1980.

[2] J. P. F. Charpin and A. O. Adewumi, "Optimal assignment of blood in a blood banking System," Tech. Rep., Mathematics in Industry Study Group (MISG), 2011.

[3] L. Jean, R. P. Myriam, and A. A. Celso, "Particle swarm optimization applied to task assignment problem," in *Proceedings of the 10th Brazilian Congress on Computational Intelligence (CBIC '11)*, Fortaleza, Brazil, November 2011.

[4] A. O. Adewumi, J. O. A. Ayeni, and E. P. Fasina, "Some comments on task allocation problem in distributed computing systems," *Journal of Computer Science and Its Application, Nigeria*, vol. 9, no. 1, pp. 136–145, 2003.

[5] A. O. Adewumi, *Some improved genetic-algorithms based heuristics for global optimization with innovative applications [Ph.D. thesis]*, School of Computational and Applied Mathematics,

- University of Witwatersrand, Johannesburg, South Africa, 2010.
- [6] M. Yusuff, J. Ariffin, and A. Mohamed, "Solving vehicle assignment problem using evolutionary computation," in *Advances in Swarm Intelligence*, vol. 6145 of *Lecture Notes in Computer Science*, pp. 523–532, 2010.
 - [7] J. C. Bansal and K. Deep, "A modified binary particle swarm optimization for knapsack problems," *Applied Mathematics and Computation*, vol. 218, no. 22, pp. 11042–11061, 2012.
 - [8] B. Ye, J. Sun, and W.-B. Xu, *Solving the Hard Knapsack Problems with a Binary Particle Swarm Approach*, vol. 4115 of *Lecture Notes in Computer Science*, 2006.
 - [9] Z. Ren and J. Wang, "A discrete particle swarm optimization for solving multiple knapsack problems," in *Proceedings of the 5th International Conference on Natural Computation (ICNC '09)*, vol. 3, pp. 166–170, Tianjin, China, August 2009.
 - [10] V. Angelis, N. Ricciardi, and G. Storchi, *Optimizing Blood Assignment in a Donation-Transfusion System*, Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università A degli Studi di Roma 'La Sapienza', Rome, Italy, 2001.
 - [11] A. Adewumi, N. Budlender, and M. Olusanya, "Optimizing the assignment of blood in a blood banking system: some initial results," in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC '12)*, pp. 751–756, Brisbane, Australia, June 2012.
 - [12] E. Dufourq, M. O. Olusanya, and A. O. Adewumi, "Studies in metaheuristics for the blood assignment problem," in *Proceedings of the Operation Research Society of South Africa (ORSSA '12)*, September 2012.
 - [13] K. Igwe, M. Olusanya, and A. Adewumi, "On the performance of GRASP and dynamic programming for the blood assignment problem," in *Proceedings of the IEEE Global Humanitarian Technology Conference (GHTC '13)*, pp. 221–225, San Jose, Calif, USA, October 2013.
 - [14] M. O. Olusanya and A. O. Adewumi, "Using Metaheuristic techniques to optimize the Blood Assignment Problem," in *Proceedings of the 4th IEEE International Advance Computing Conference (IACC '14)*, pp. 1331–1336, February 2014.
 - [15] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, pp. 1942–1948, Perth, Australia, December 1995.
 - [16] R. Eberhart and J. Kennedy, "New optimizer using particle swarm theory," in *Proceedings of the 6th International Symposium on Micro Machine and Human Science (MHS '95)*, pp. 39–43, Nagoya, Japan, October 1995.
 - [17] B. Khokhar, K. P. S. Parmar, and S. Dahiya, "An efficient particle swarm optimization with time varying acceleration coefficients to solve economic dispatch problem with valve point loading," *Energy and Power*, vol. 2, no. 4, pp. 74–80, 2012.
 - [18] J. D. L. Silva, *Metaheuristic and multiobjective approaches for space allocation [Ph.D. thesis]*, University of Nottingham, Nottingham, UK, 2003.
 - [19] M. M. Mansour, S. F. Mekhamer, and N. E.-S. El-Kharbawe, "A modified particle swarm optimizer for the coordination of directional overcurrent relays," *IEEE Transactions on Power Delivery*, vol. 22, no. 3, pp. 1400–1410, 2007.
 - [20] V. N. Dieu, P. Schegner, and W. Ongsakul, "A newly improved particle swarm optimization for economic dispatch with valve point loading effects," in *Proceedings of the IEEE Power and Energy Society General Meeting*, pp. 1–8, San Diego, Calif, USA, July 2011.
 - [21] Y. Ma, C. Jiang, Z. Hou, and C. Wang, "The formulation of the optimal strategies for the electricity producers based on the particle swarm optimization algorithm," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1663–1671, 2006.
 - [22] A. A. Martins and A. A. Oluyinka, "An Adaptive Velocity Particle Swarm Optimization for high-dimensional function optimization," in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC '13)*, pp. 2352–2359, Cancún, Mexico, June 2013.
 - [23] M. A. Arasomwan and A. O. Adewumi, "Improved particle swarm optimization with a collective local Unimodal search for continuous optimization problems," *The Scientific World Journal*, vol. 2014, Article ID 798129, 23 pages, 2014.
 - [24] M. A. Arasomwan and A. O. Adewumi, "On the performance of linear decreasing inertia weight particle swarm optimization for global optimization," *The Scientific World Journal*, vol. 2013, Article ID 860289, 12 pages, 2013.
 - [25] Y. H. Shi and R. C. Eberhart, "A modified particle swarm optimizer," in *Proceedings of the IEEE International Conference on Evolutionary Computation (ICEC '98)*, pp. 69–73, Anchorage, Alaska, USA, May 1998.
 - [26] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation*, vol. 5, pp. 4104–4108, Orlando, Fla, USA, October 1997.
 - [27] M. Jiang, Y. P. Luo, and S. Y. Yang, "Particle swarm optimization-stochastic trajectory analysis and parameter selection," in *Swarm Intelligence, Focus on Ant and Particle Swarm Optimization*, F. T. S. Chan and M. K. Tiwari, Eds., pp. 179–198, InTech, 2007.

Research Article

A Clinical Decision Support System for the Diagnosis, Fracture Risks and Treatment of Osteoporosis

Bjarni V. Halldorsson,¹ Aron Hjalti Bjornsson,^{2,3} Haukur Tyr Gudmundsson,^{2,4} Elvar Orn Birgisson,⁵ Bjorn Runar Ludviksson,^{4,6} and Bjorn Gudbjornsson^{2,4}

¹*Institute of Biomedical and Neural Engineering, School of Science and Engineering, Reykjavik University, 101 Reykjavik, Iceland*

²*Centre for Rheumatology Research, University Hospital, 101 Reykjavik, Iceland*

³*Faculty of Medicine, Debrecen University, Debrecen 4032, Hungary*

⁴*Faculty of Medicine, University of Iceland, 101 Reykjavik, Iceland*

⁵*Osteoporosis Clinic, Akureyri Hospital, 600 Akureyri, Iceland*

⁶*Department of Immunology, University Hospital, 101 Reykjavik, Iceland*

Correspondence should be addressed to Bjarni V. Halldorsson; bjarnivh@ru.is

Received 18 May 2014; Revised 31 August 2014; Accepted 9 October 2014

Academic Editor: Martin Grootveld

Copyright © 2015 Bjarni V. Halldorsson 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.

Expanding medical knowledge increases the potential risk of medical errors in clinical practice. We present, OPAD, a clinical decision support system in the field of the medical care of osteoporosis. We utilize clinical information from international guidelines and experts in the field of osteoporosis. Physicians are provided with user interface to insert standard patient data, from which OPAD provides instant diagnostic comments, 10-year risk of fragility fracture, treatment options for the given case, and when to offer a follow-up DXA-evaluation. Thus, the medical decision making is standardized according to the best expert knowledge at any given time. OPAD was evaluated in a set of 308 randomly selected individuals. OPAD's ten-year fracture risk computation is nearly identical to FRAX ($r = 0.988$). In 58% of cases OPAD recommended DXA evaluation at the present time. Following a DXA measurement in all individuals, 71% of those that were recommended to have DXA at the present time received recommendation for further investigation or specific treatment by the OPAD. In only 5.9% of individuals in which DXA was not recommended, the result of the BMD measurement changed the recommendations given by OPAD.

1. Introduction

According to the International Osteoporosis Foundation (IOF), one in three women and one in five men will experience an osteoporotic fracture later in their life; thus, globally an osteoporotic fracture is estimated to occur every third second [1–3]. In that context, nine million North Americans have osteoporosis and it is estimated that 43 million have low bone mass measured by DXA (<http://nof.org/news/1648>), a precursor of osteoporosis, known as osteopenia. The National Osteoporosis Foundation (NOF) in the USA has estimated that, in USA alone, osteoporosis related fractures number two million every year resulting in an annual cost of \$19 billion [4]. Due to the increasing numbers of elderly people,

the number of fractures will increase to three million in 2025, resulting in a \$25.3 billion annual cost in the USA [5]. Correctly prescribed bone protective treatment may reduce the fracture risk by 30–50% in only three years [6]. It is therefore critical to identify individuals at risk of a fragility fracture and offer them the treatment protocols to decrease the number of fractures in coming years.

The diagnosis of osteoporosis has traditionally relied on bone mineral density (BMD) measurement [7–9]. However, a number of other factors affect a physicians' decision to treat osteoporosis, including family history, lifestyle, and various medical factors [10]. The incorporation of all of these factors into a treatment recommendation and a decision to perform a BMD scan can be a formidable task that requires considerable

expert knowledge and experience. Here we describe the development of a clinical decision support system that automates the dissemination of this knowledge.

Currently available risk calculator systems for osteoporosis, for example, FRAX [11], Garvan [12], and Qfracture [13], provide the user with ten-year probability of fragility fracture related to osteoporosis. While this information may be useful for a number of users, fracture risk is known to be difficult for the layperson to interpret [14] and even for nonexpert health professionals. Furthermore, it is not immediately evident how this information is to be interpreted, for example, for treatment decision in daily clinical practice. Clinical decision support systems (CDSS) have been suggested as a means of disseminating knowledge of best practice to physicians [15], electronic medical reminders have been shown to improve osteoporosis management after fractures [16], and Kastner and Straus [17] have concluded that “multi-component tools that are targeted to physicians and patients may be effective for supporting clinical decision making in osteoporosis disease management.”

We have developed a clinical decision support system that gives the 10-year fracture probability due to osteoporosis for the individual patient, lifestyle recommendations, and recommendations as to whether and when a BMD scan is recommended. Furthermore, the system will identify patients at risk of fractures and would benefit from specific preventive medical treatment. Our osteoporosis adviser (OPAD) is designed for those with medical knowledge such as general physicians and clinical nurse specialists.

In this paper, we evaluate the reliability of the OPAD system by comparing its 10-year fracture probability with the probability given by FRAX and by assessing the quality of its BMD scanning recommendations, that is, whether those that were recommended to have a BMD measurement benefited from the measurement.

2. Materials and Methods

2.1. The Osteoporosis Advisor (OPAD). We have designed an expert system to assist in the diagnosis and treatment of osteoporosis. The software takes as input a set of clinically relevant parameters from which the 10-year fracture risk is computed based on published country specific data [8, 11, 18]. The output of the program is as follows: the 10-year fracture risk of an individual, lifestyle and treatment recommendations, and lastly a suggestion for the time when a follow-up BMD scan should take place. The clinically relevant parameters and the computed risk for fracture are used as input into an expert system which gives specific recommendations for each case with respect to lifestyle changes and treatment options.

The software further outputs immediately relevant information to the user: a risk group for the individual (low, medium, or high risk for fracture compared to age-matched controls [18, 19]) and a diagnosis of osteopenia, osteoporosis, or manifested osteoporosis (osteoporosis with fracture), as well as for glucocorticosteroid induced osteoporosis. Diagnosis of osteoporosis is made according to the WHO definitions [20],

TABLE 1: Patient attributes used by the osteoporosis advisor.

Age
Bone mineral density (T value)
Ethnicity
Gender
Previous osteoporotic related fracture
Parent hip fracture
Current smoking
Current use of glucocorticosteroids for more than three months
Rheumatoid arthritis
Secondary osteoporosis
Alcohol: 3 or more units per day
Hormone replacement therapy
Regular exercise
Sufficient calcium intake
Sufficient vitamin D intake

where osteoporosis is diagnosed when BMD results in a T value of -2.5 or lower; that is, the BMD is ≥ 2.5 SD below the mean of young individuals of the same sex and reach, and osteopenia is diagnosed when the T value is between -1 and -2.5 .

2.2. Design of Expert System. The design of the system uses a knowledge mapping approach. Expert physicians were queried to determine the clinically relevant parameters for the recommendation of osteoporosis treatment and recommendations for BMD measurements. A group of different specialists who were all interested in osteoporosis (rheumatologist, endocrinologist, general practitioner, and geriatrician) participated in the process. The Intellix Advisor [21, 22] was used for knowledge capture in the model.

The Intellix Advisor allows for active acquisition of knowledge. A set of instances is input into the software and from these input examples the software can be told either to construct a neural network based model based on the examples or to ask for more examples that consist of patients not considered by the model. In our approach patients were added to the model until a diagnosis or recommendations could be made for every possible tested patient. The model is implemented as a lookup table and for every new patient diagnosed a patient with the same characteristics is found in our database.

The diagnosis for a patient includes four distinct pieces of information: 10-year fracture risk, lifestyle recommendations, treatment recommendations, and recommendation of the time for the next BMD measurement or follow-up evaluation of the individual patient. Initially a set of clinically relevant parameters was determined (see Table 1).

2.3. Knowledge Capture of Treatment Recommendations. The OPAD system follows the frame of international guidelines, for example, the Scottish Intercollegiate Guidelines Network (SIGN) on osteoporosis (<http://sign.ac.uk/pdf/sign71.pdf>),

TABLE 2: Patient attributes used for the recommendation of osteoporosis treatment.

Attribute	Type
Gender	Male/female
GIOP	Yes/no
Fragility fracture	Yes/no
Fracture risk	High/medium/low
Treatment	Yes/no
Secondary osteoporosis	Yes/no
T value	Numerical
Age	Numerical
Menopause status	Before/<3 years/>3 years
Diagnosis	None/osteopenia/osteoporosis/manifest osteoporosis/GIOP

and the system also takes into account regional differences, that is, local or national guidelines. However, the system predominantly relies on knowledge capture process of the expert panels, as guidelines never cover all cases. In the end a total of fifteen different treatment recommendations were initially identified as possible recommended treatment options for osteoporosis. The recommended treatments ranged from no treatment to specific recommendations for which drug class was the most appropriate, either as a preventive measure or as a treatment for manifest osteoporosis.

Table 2 lists the clinically relevant patient attributes determined by our expert physicians. Different medical specialists with expertise in osteoporosis reviewed a list of cases that had recently visited the osteoporosis clinic at the University Hospital in Reykjavik (LSH), Iceland. For each patient the physicians reviewed their clinical decision and were then asked to determine which of the clinically relevant parameters influenced his decision. The clinically relevant information and the diagnoses were entered into the Intellix Advisor, leaving those clinical relevant parameters that were not relevant for the decision as “do not care.” After considering the set of real patient cases and their detailed experts’ reviews, the Intellix Advisor software constructed a set of virtual patients having clinical characteristics not observed among the list of cases already considered. This process was continued until the space of all possible patients was covered; a decision could be reached for all possible patients that could enter the clinic, independent of clinical characteristics. As a result a total of 80 relevant rules were constructed in the final outcome of the system.

2.4. Time until Next DXA Measurement. The recommended time until the next DXA measurement was also determined by using a knowledge mapping process. A list of clinically relevant pieces of information was determined which can be seen in Table 3.

Seven different recommendations were made for the time for the next BMD measurement, listed in Table 4. The construction of the clinical decision model then followed

TABLE 3: Attributes used to determine time until next DXA measurement.

Attribute	Type
Gender	Male/female
Menopause	Before, <3 years, >3 years
Risk group	High/medium/low
Treatment	Yes/no
Changes in BMD measured by DXA	No DXA/improving/unknown or losing/fast losing/neutral
Glucocorticosteroids	Yes/no

TABLE 4: The possible recommendations for the next time for a BMD scan.

At menopause
At the age of 65
Now
In 1-2 years
In 3 years
In 5 years
DXA not recommended

the same protocol as described above for the treatment recommendations. In the end a total of 87 rules were determined to be clinically relevant.

2.5. Capture of Disease Risk Models. The computed 10-year risk of fracture was based on the World Health Organization (WHO) fracture risk assessment recommendations [19, 20]. Thus, the fracture risk predications delivered by our system are comparable to the results given by FRAX [18]. The output of the FRAX in collaboration with NOGG recommendation guidelines also includes the classification of individuals into high, medium, and low risk individuals [23]. The FRAX recommendation guidelines are given as a set of text tables, with the risk of fracture and confidence interval for the risk of fracture given as a function of age and the number of risk components an individual has. As the tables only give fracture risk probabilities for selected age groups, an interpolation is done to compute the fracture risk probabilities for other age groups.

2.6. Model Testing. In order to validate that the treatment recommendations presented to the end user agreed with the treatment recommendations originally determined for each patient a quality control module was developed. Thus, built on top of our clinical database a set of 300 virtual quality control patients was created which were used to automatically verify the correctness of the system recommendations for real life clinical information given for each case. Experiments verified that the results of these patients agreed in both the model created and the interface to the end user.

2.7. Test in Real Life. As the WHO recommendation guidelines are not given with a closed form formula we compared the fracture risk computation given by our OPAD system with the fracture risk computation given by FRAX. We selected consecutive 308 individuals from the out-patient osteoporosis clinic at LSH, who visited the clinic from the 1st of January 2012 onwards. We compared the ten-year fracture risk computed using the OPAD and the ten-year fracture risk computed using recommendations given by our model with those given by FRAX. Linear regression was run to compare the results between the two systems.

We also reevaluated the same group of 308 individuals with our OPAD system with respect to the need of DXA at the present time or later; that is, the risk evaluation was done without the DXA results (T value). We then analyzed whether the DXA result, that is, in those cases where OPAD did not recommend DXA at the given time, influenced the automatic treatment and the follow-up recommendations given by the OPAD compared to the recommendations by the experts, who had access to all the clinical data.

2.8. Ethics Statement. The Data Protection Authority (S5680) and the National Bioethics Committee of Iceland approved the study protocol (VSNb2010050008). The original patients' data were hosted by the University Hospital, Reykjavik, Iceland. The data were anonymized before being provided to the researchers.

3. Results

3.1. Correlation between Risk Evaluation by OPAD and FRAX. Of the 308 cases, 39 were males and 269 were females, with a mean age of 61 years (15–89). Figure 1 shows the 10-year fracture risk when DXA is included presented by the two different programs, that is, OPAD and FRAX. We obtained a correlation of $r = 0.988$ with a mean paired difference of 0.7678% (SD = 1.9946%) when comparing individual patient prior to DXA evaluation and $r = 0.977$, mean difference 1.8285% (SD = 2.8%) when DXA result were included in the risk evaluation or a near perfect correlation between these two risk calculators.

3.2. OPAD and Next DXA. These 308 patients were reevaluated by the OPAD with respect to the need for a DXA at the present time or later, that is, before they underwent their DXA evaluation.

In 178 cases (58%), out of these 308 cases, the OPAD system recommended DXA evaluation at the present time. Following DXA measurement in these 178 cases, 91 (51%) of those received OPAD recommendation on specific treatment options, where 5 patients (3%) were recommended to continue with their treatment. Additional 31 patients (17%) received recommendation on consulting specialist in osteoporosis. Meanwhile, only 51 of these 178 cases (29%) received general prevention measurement recommendations. Thus, the DXA investigation performed according to the recommendation of the OPAD system seems to influence the clinical decision-making process.

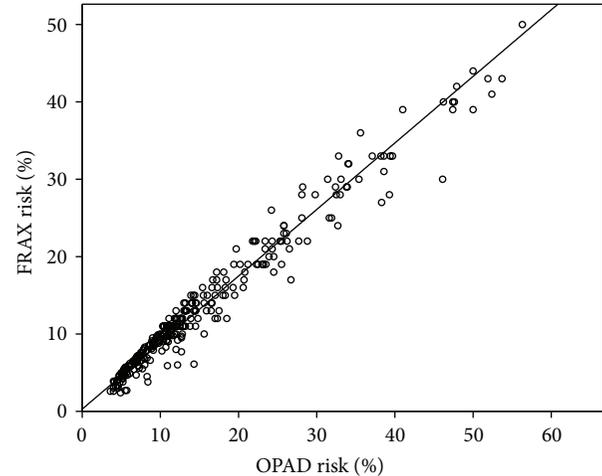


FIGURE 1: Ten-year risk for major osteoporotic fracture computed using the osteoporosis advisor (OPAD; x -axis), compared to the fracture risk computed using FRAX (y -axis).

Out of the 308 original cases, 102 cases (33%) came for their DXA even though the OPAD system would have recommended that they should have their DXA at the age of 65; that is, every third patient who came in for a measurement did not need the DXA evaluation at the present time according to the OPAD system.

3.3. Influence of DXA on OPAD Recommendations. In only six of these 102 cases (5.9%), did the OPAD system change its recommendation following the BMD measurement? In four cases the OPAD recommended specific bone protective treatment, due to the fact that these four individuals were diagnosed with osteoporosis with a significantly increased risk of fragility fracture, that is, with a 10-year fracture risk in the range from 9.1% to 14.3%. In two additional cases the OPAD changed its recommendation to continuation of already taken measurements.

In a further 22 cases of these 308 cases (7.1%) the OPAD system recommended DXA within 1–3 years depending on various clinical circumstances. Independent of these recommendations all patients received a DXA measurement, like other patients in this study and only one of these individuals received a different recommendation following the DXA evaluation.

4. Discussion

Osteoporosis is a disorder affecting the density and infrastructure of the bone mass [20]. In its worst outcome it can lead to the so-called fragility fractures [24], most frequently seen in the vertebral spine, wrist, and the hip. Due to their critical location, such fractures most often result in a debilitating outcome for those affected individuals leading to a significantly negative impact on not only their quality of life but also individual life expectancy [25]. Despite an effective drug treatment being readily available for affected patients, it has been recognized that the majority of individuals at risk

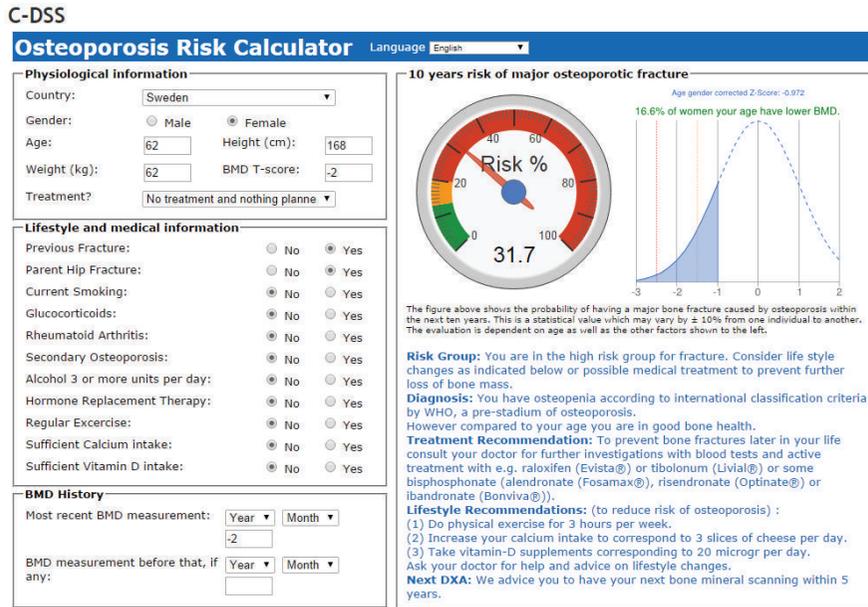


FIGURE 2: A 62-year-old Swedish female patient with history of fracture whose mother also had a history of hip fracture. She did not have other medical risk factors and she had osteopenia according to a recent DXA evaluation with a T -score of -2.0 .

and also those already affected with osteoporosis and increased risk of fragility fracture are not correctly diagnosed or go unnoticed and therefore miss potentially lifesaving therapeutic measures [6]. Thus, even patients who have suffered fragility fractures and have been exposed to the healthcare system are not identified and treated with the proper preventive regimen they deserve. In this context, if correctly implemented clinical decision support systems such as the one presented here (OPAD) should have the potential to improve both public and healthcare workers awareness of osteoporosis [15]. Such measures not only would improve and streamline the diagnosis process but would also become a valid diagnosis aid to secure the optimal treatment and outcome of our osteoporotic patients.

Bone mineral density (BMD), that is, bone mass, may be measured with several methods, for example, quantitative ultrasound (QUS), peripheral quantitative computer tomography (pQCT), and dual energy X-ray absorptiometry (DXA) which is the golden standard of bone mass measurement. Most professional interest groups, for example, NOF and IOF, have published recommendations on when to use DXA for BMD measurement [26], while others research groups supported by the National Institutes of Health have made an effort to analyze the need of rescreening in postmenopausal women [27]. As osteoporosis is a silent disease until the fracture occurs, BMD measurement is the only method to diagnose osteoporosis prior to the fracture. However, the access to DXA-machines is limited in most countries, and it is therefore important to find those at risk of osteoporosis for DXA-evaluation and those individuals who receive the greatest benefits from the investigation. Thus, improving the cost efficiency by correctly identifying those at risk and those who should be referred for BMD should lead to shorter

waiting lists for specialist referrals and improved diagnostic accuracy. Our finding that more than one third of BMD tested patients could be identified a priori to be at no risk of osteoporosis related fractures, thus, did not benefit from the DXA evaluation.

Several risk calculator tools for bone fractures have been presented, but only FRAX has been recommended and supported by the World Health Organization (WHO). FRAX was developed for the calculation of the 10-year fracture risk for hip fracture or a major osteoporotic fracture (clinical spine, forearm, hip, or shoulder fracture) based on certain risk factors, with or without results of DXA measurement of the hip. FRAX offers country specific values for several countries in Europe, North and Latin America, the Middle East, Africa, and Oceania, or a total of 52 country specific datasets [18]. However, FRAX does not give any specific diagnosis or treatment options, and neither have they reported their algorithms used to derive the 10-year fracture risk. Further analysis of our 10-year fracture risk in context to the given recommendation by our OPAD is in progress, before we can make our methodology public in detail.

The busy clinician may have difficulties in interpreting the risk value figure for each patient in hectic daily clinical practice. With this in mind we have extended the information provided by OPAD, by giving a specific diagnosis, that is, osteoporosis or osteopenia, and specific recommendations on prevention, time of next DXA, and treatment options according to international guidelines and experts knowledge [10]. In addition, our results are presented on an interactive riskometer, which gives a comparison to the background population of the same sex and age, both graphically and by calculation of the Z value for the individual patient (please see Figure 2).

Although physicians subscribe to several medical journals, presenting thousands of articles yearly, they have difficulty in keeping up to date in all areas in their daily practice. The more complicated medical world amplifies the risk of diagnosis errors. OPAD allows “best practice” in osteoporosis risk evaluation of fragility fractures and treatment to be captured, distributed, and automated in a simple bedside manner for the busy practicing physician and other health care providers, including nurses working in fracture liaison services, as now highly recommended by IOF [28]. Although the OPAD system presented in this present study reflects Swedish data, it runs in ten different national specific datasets. Improvements in treatment alternatives or changes in clinical guidelines can easily be incorporated into the OPAD system; even country specific guidelines can be internalized. Further studies are needed to analyze whether a clinical approach with the help of digital CDSS tools, such as OPAD, may not only improve individual care, but also become highly cost effective. Such studies need to involve primary care, fracture clinics, and in-hospital fracture liaison services. Furthermore, survey among primary care doctors regarding their use of and evaluation of the OPAD system, including their assessment of how to implement clinical decision system, such as OPAD, was carried on fracture risk management and treatment in daily clinical primary care praxis.

We conclude that OPAD is accurate in respect to fracture risk probability evaluation and may presumptively be cost effective in fracture liaison services. However, cost-benefit studies are needed in the field of osteoporosis preventive care and CDSS.

Conflict of Interests

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

Authors' Contribution

Aron Hjalti Bjornsson, Bjorn Gudbjornsson, Bjorn Runar Ludviksson, Elvar Orn Birgisson, and Haukur Tyr Gudmundsson collected the patient data used in the paper. Aron Hjalti Bjornsson, Bjarni V. Halldorsson, and Haukur Tyr Gudmundsson analyzed the data collected. Bjarni V. Halldorsson, Bjorn Runar Ludviksson, and Bjorn Gudbjornsson wrote the paper. All authors read and approved the final version of the paper.

Acknowledgments

The authors would like to thank Klaus Bjorn Jensen, Thorsteinn Geirsson, Dagrun Jonasdottir, Elias R. Ragnarsson, and Orn Bardur Jonsson for their contribution to this work.

References

- [1] A. Randell, P. N. Sambrook, T. V. Nguyen et al., “Direct clinical and welfare costs of osteoporotic fractures in elderly men and women,” *Osteoporosis International*, vol. 5, no. 6, pp. 427–432, 1995.
- [2] J. Y. Reginster and N. Burlet, “Osteoporosis: a still increasing prevalence,” *Bone*, vol. 38, pp. S4–S9, 2006.
- [3] L. Melton, E. Chrischilles, C. Cooper, A. W. Lane, and B. L. Riggs, “Perspective: how many women have osteoporosis?” *Journal of Bone and Mineral Research*, vol. 7, no. 9, pp. 1005–1010, 1992.
- [4] J. Waalen, “Current and emerging therapies for the treatment of osteoporosis,” *Journal of Experimental Pharmacology*, vol. 2, pp. 121–134, 2010.
- [5] E. Hernlund, A. Svedbom, M. Ivergård et al., “Osteoporosis in the European Union: medical management, epidemiology and economic burden,” *Archives of Osteoporosis*, vol. 8, article 136, 2013.
- [6] S. Das and J. Crockett, “Osteoporosis—a current view of pharmacological prevention and treatment,” *Journal of Drug Design, Development and Therapy*, vol. 7, pp. 435–448, 2013.
- [7] J. A. Kanis, “Diagnosis of osteoporosis and assessment of fracture risk,” *The Lancet*, vol. 359, no. 9321, pp. 1929–1936, 2002.
- [8] J. C. Martin and D. M. Reid, “Appendicular measurements in screening women for low axial bone mineral density,” *The British Journal of Radiology*, vol. 69, no. 819, pp. 234–240, 1996.
- [9] “Osteoporosis in postmenopausal women: diagnosis and monitoring,” <http://www.ncbi.nlm.nih.gov/books/NBK11863/>.
- [10] “SIGN Guidelines: Management of Osteoporosis,” <http://www.sign.ac.uk/guidelines/fulltext/71/>.
- [11] J. A. Kanis, O. Johnell, A. Oden, B. Jonsson, A. Dawson, and W. Dere, “Risk of hip fracture derived from relative risks: an analysis applied to the population of Sweden,” *Osteoporosis International*, vol. 11, no. 2, pp. 120–127, 2000.
- [12] S. A. Frost, N. D. Nguyen, J. R. Center, J. A. Eisman, and T. V. Nguyen, “Timing of repeat BMD measurements: development of an absolute risk-based prognostic model,” *Journal of Bone and Mineral Research*, vol. 24, no. 11, pp. 1800–1807, 2009.
- [13] J. Hippisley-Cox and C. Coupland, “Derivation and validation of updated QFracture algorithm to predict risk of osteoporotic fracture in primary care in the United Kingdom: prospective open cohort study,” *The British Medical Journal*, vol. 344, Article ID e3427, 2012.
- [14] L. Sorensen, D. Gyrd-Hansen, I. S. Kristiansen, J. Nexøe, and J. B. Nielsen, “Laypersons’ understanding of relative risk reductions: randomised cross-sectional study,” *BMC Medical Informatics and Decision Making*, vol. 8, article 31, 2008.
- [15] J. B. Jones, W. F. Stewart, J. D. Darer, and D. F. Sittig, “Beyond the threshold: real-time use of evidence in practice,” *BMC Medical Informatics and Decision Making*, vol. 13, no. 1, article 47, 2013.
- [16] A. Feldstein, P. J. Elmer, D. H. Smith et al., “Electronic medical record reminder improves osteoporosis management after a fracture: a randomized, controlled trial,” *Journal of the American Geriatrics Society*, vol. 54, no. 3, pp. 450–457, 2006.
- [17] M. Kastner and S. E. Straus, “Clinical decision support tools for osteoporosis disease management: a systematic review of randomized controlled trials,” *Journal of General Internal Medicine*, vol. 23, no. 12, pp. 2095–2105, 2008.
- [18] J. A. Kanis, E. V. McCloskey, H. Johansson, A. Oden, O. Ström, and F. Borgström, “Development and use of FRAX in osteoporosis,” *Osteoporosis International*, vol. 21, no. 2, pp. 407–413, 2010.

- [19] J. A. Kanis, O. Johnell, A. Oden, H. Johansson, and E. McCloskey, "FRAX and the assessment of fracture probability in men and women from the UK," *Osteoporosis International*, vol. 19, no. 4, pp. 385–397, 2008.
- [20] WHO Study Group, "Assessment of osteoporosis at the primary health care level," *World Health Organization Technical Report Series*, vol. 843, pp. 1–129, 1994.
- [21] *Intellix A: Getting Started with Intellix Designer*, 2013.
- [22] C. Wang, J. T. Luxhøj, and J. Johansen, "Applying a knowledge management modeling tool for manufacturing vision (MV) development," *Industrial Management & Data Systems*, vol. 104, no. 9, pp. 735–743, 2004.
- [23] J. A. Kanis, E. V. McCloskey, H. Johansson et al., "Case finding for the management of osteoporosis with FRAX—assessment and intervention thresholds for the UK," *Osteoporosis International*, vol. 19, no. 10, pp. 1395–1408, 2008.
- [24] A. Warriner and K. Saag, "Osteoporosis diagnosis and medical treatment," *Orthopedic Clinics of North America*, vol. 44, no. 2, pp. 125–135, 2013.
- [25] G. S. Tajeu, E. Delzell, W. Smith et al., "Death, debility, and destitution following hip fracture," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 69, no. 3, pp. 346–353, 2014.
- [26] "Having a Bone Density Test," <http://www.nof.org/articles/743>.
- [27] M. L. Gourlay, J. P. Fine, J. S. Preisser et al., "Bone-density testing interval and transition to osteoporosis in older women," *The New England Journal of Medicine*, vol. 366, no. 3, pp. 225–233, 2012.
- [28] K. Akesson, D. Marsh, P. J. Mitchell et al., "Capture the fracture: a best practice framework and global campaign to break the fragility fracture cycle," *Osteoporosis International*, vol. 24, no. 8, pp. 2135–2152, 2013.

Research Article

An Agent-Based Epidemic Simulation of Social Behaviors Affecting HIV Transmission among Taiwanese Homosexuals

Chung-Yuan Huang

Department of Computer Science and Information Engineering, School of Electrical and Computer Engineering, College of Engineering, Chang Gung University, 259 Wen Hwa 1st Road, Taoyuan 333, Taiwan

Correspondence should be addressed to Chung-Yuan Huang; gscott@mail.cgu.edu.tw

Received 28 May 2014; Accepted 30 October 2014

Academic Editor: Ezequiel López-Rubio

Copyright © 2015 Chung-Yuan Huang. 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.

Computational simulations are currently used to identify epidemic dynamics, to test potential prevention and intervention strategies, and to study the effects of social behaviors on HIV transmission. The author describes an agent-based epidemic simulation model of a network of individuals who participate in high-risk sexual practices, using number of partners, condom usage, and relationship length to distinguish between high- and low-risk populations. Two new concepts—free links and fixed links—are used to indicate tendencies among individuals who either have large numbers of short-term partners or stay in long-term monogamous relationships. An attempt was made to reproduce epidemic curves of reported HIV cases among male homosexuals in Taiwan prior to using the agent-based model to determine the effects of various policies on epidemic dynamics. Results suggest that when suitable adjustments are made based on available social survey statistics, the model accurately simulates real-world behaviors on a large scale.

1. Introduction

According to the World Health organization (WHO) [1], more than 4.7 million people in Asian countries were HIV-positive at the end of 2008, with 350,000 newly infected individuals and 330,000 AIDS-related deaths in that year alone. While the 2008 AIDS-related mortality rate in South-East and South Asia was approximately 12% lower than the 2004 peak, the number of deaths in 2008 was still more than three times higher than that recorded in 2000. Between 2000 and 2009, the annual number of newly identified HIV-positive Taiwanese people increased from 536 to 1,694 (Figure 1). Since the peak number in 2005, the annual incidence rate between 2006 and 2010 declined by between 5% and 34% (average 14.9%). Homosexual activity (“men who have sex with men” or MSM) is believed to be the primary mode of HIV transmission in Taiwan today—57% of all cases recorded in 2008 alone [2]. Developing an HIV simulation model to understand HIV epidemic dynamics within male homosexual communities in Taiwan is therefore considered important

for assessing the efficacies of HIV prevention efforts and intervention strategies associated with MSM activities [3, 4].

Some researchers have noted that the topological (connectivity) features of sexual networks exert considerable influences on HIV epidemic dynamics [5–9]. These features support analyses of the subtle details of HIV epidemic dynamics that population-based simulation approaches do not. Brailsford et al. [10], Paltiel et al. [11], and Peterson et al. [12] have applied specific complex network models to investigate HIV epidemic dynamics. Other researchers have addressed topological features and statistical distributions of sexual networks without focusing on individual social behaviors that affect HIV epidemic dynamics (see, e.g., [13–19]). They also show a tendency to overlook the flexibility offered by agent-based social simulation approaches.

An HIV epidemic simulation with the characteristics of power-law degree distribution and small-world phenomena for reproducing sexual networks was developed for this project to assess the efficacies of specific HIV prevention and intervention strategies in Taiwan associated with social

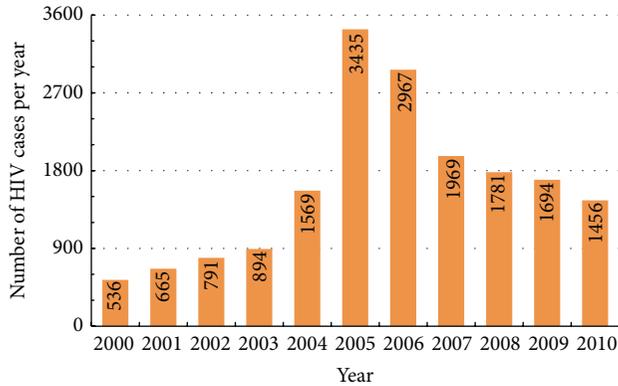


FIGURE 1: Taiwan HIV epidemic curves from 2000 to 2010.

behaviors. A specific emphasis in this paper will be MSM activity in saunas and bars in northern Taiwan, with agents in the model capable of modifying their behaviors to the degree that they affect all homosexuals residing in that part of Taiwan. Following the lead of previous sexual network studies [20–24], the proposed model focuses on scale-free properties depicting small-world phenomena. We found that, during one three-month period, the numbers of sex partners of highly active homosexual males in the cities of Taipei and Taichung followed a scale-free distribution on a log-log-axis. One conclusion of the present study is that by addressing specific individual behaviors and implementing rules derived from social survey statistics, global observations can be accurately identified by simulations without introducing unreasonably large numbers of societal rules, thus reflecting small-world properties that are inherent to scale-free networks. We also found that the clustering results can be used to support analyses of factors such as drug and condom usage.

2. Previous Sexual Network Studies

In terms of approach, epidemic simulations can be classified as top-down or bottom-up [25]. Top-down simulations start by specifying group or subpopulation characteristics and behaviors and modeling their relationships with other groups [26]. One problem is that mathematical equations in top-down epidemic simulations can quickly become so complex that analytical representations of multivariate and complex stochastic processes become exceptionally difficult to manage in terms of computation and stability. In contrast, bottom-up epidemic simulations are based on a combination of an agent-based model and an underlying social network, with nodes representing agents and edges representing contacts between them. Bottom-up epidemic simulations have been widely used to explore social macrolevel phenomena by specifying the microlevel characteristics and behaviors of individuals and their contacts in well-defined social networks over specific time periods [27–30]. The HIV epidemic simulation described in this paper is in the second category.

The history of social network studies by sociologists has produced numerous insights regarding specific social

contacts [31]. Typical top-down epidemic simulations such as the compartmental models proposed by Kermack and McKendrick [32] assume random contacts among individuals, especially in scenarios involving the airborne transmission of infections. In numerical simulations of sexually transmitted infections (STIs), randomness is deemphasized when addressing sexual contact networks and mixing patterns within populations, both considered important concepts in modeling sexual contact within well-defined subgroups [33, 34]. Assortative sexual mixing implies concordance between sex partners in terms of factors such as age, location of residence, ethnicity, socioeconomic status, and sex partner acquisition rate [35]. In contrast, disassortative mixing, which implies discordance in sex partner characteristics, allows the spread of STIs from groups with high STI prevalence (e.g., commercial sex workers) to members of other groups [36].

Wylie and Jolly’s [37] use of voluntarily given contact information and tracing via clinical contacts in Manitoba, Canada, is a rare example of a reconstructed real-world sexual contact network. A total of 4,544 STI-positive individuals were asked to identify their sexual partners—information that in most cases is very difficult to obtain. According to their data, sexual contact networks consist of numerous small clusters. An important task for any researcher is to determine cluster overlaps and to analyze how they support the spread of HIV. Note that since informants such as those in the Manitoba study are already infected, the data they offer are often viewed as suspect, influenced by subjective biases, and lacking in insightful detail [38].

Sexual networks are dynamic. Since connections evolve at different rates over time, concurrent relationships within networks must be considered [17]. Researchers have made many attempts to develop models that reflect various levels of fidelity among partners, as well as transmission routes among individuals in nonconcurrent relationships within specific time frames. More active individuals exert greater influence in terms of the spread of STIs within their networks. This stresses the importance of model precision in representing more sexually active subpopulations.

Some contemporary approaches to social network research are based on topological properties found in real-world human societies, including high degrees of local clustering and small average distances between nodes. In these models, STIs are passed between linked nodes that are few in number but high in variance. This architecture is closely associated with small-world networks [14, 15, 20–23], whose nodes have strong connections with their closest neighbors (thereby influencing local networks) and very few random links connecting distant locations. An interesting property of these networks is the dramatic decrease in average distance between random nodes due to the small number of long-distance links.

Complex networks with small-world characteristics can be classified according to the node degree statistical distribution $P(k)$ [39]. Three network classes that have been identified so far are single-scale, broadscale, and scale-free; sexual networks belong to the scale-free category [14, 15, 20–23]. Most scale-free network nodes have few connections with their adjacent nodes, and only small numbers of

nodes have large numbers of connections. A sexual network exhibiting power-law degree distribution properties serves as an example of how epidemics can spread quickly, with scope and speed determined by link distribution rules. A large number of nodes with multiple links significantly increase the speed of a spreading epidemic on a large scale, thus explaining why degree distribution is central to sexual network model construction. Accordingly, a major challenge in HIV simulation design and analysis is obtaining reliable data.

3. HIV in Taiwanese Homosexuals

According to statistics from the Taiwan Center of Disease Control (TCDC) [2] of the 6,850 known HIV-1-positive individuals living in Taiwan 2004, 488 were foreigners, 1,874 had developed AIDS, and the male : female ratio was just below 14 : 1. This represents a 15% annual increase between 2004 and 2008—well above the United Nations criteria for a “serious increase.” According to risk factor analysis results, 35.6% of the increase could be attributed to heterosexual sex, 45% to MSM activity involving either homosexuals or bisexuals, 6.2% to injecting drug use behaviors (primarily syringe-sharing), and 12% to “other.” In other words, sexual behaviors accounted for the large majority of HIV transmission routes in Taiwan during this time period.

Gay Taiwanese men frequently find partners in saunas, bars, and secluded parks. To investigate MSM-related HIV-1 infections, in 2003 and 2004, the AIDS Prevention and Research Center of National Yang-Ming University used anonymous questionnaires to collect data on sexual contacts among gay sauna customers [40]. As part of this project, the center distributed information on sexually transmitted diseases (STDs) to sauna businesses. Concerned about the fact that an insufficient number of patrons would feel motivated to join the study, the researchers also used a mix of purposive and snowball sampling to increase the number of participants. The primary goals of the two-year study were to offer anonymous HIV-1 antibody and syphilis tests and counseling and to investigate risk factors for specific STDs among gay sauna customers. The researchers also used epidemiological methods to determine HIV-1 subtypes and to analyze correlations between subtypes and risk factors. Data were also collected on demographics, sexual behavior self-cognition, and knowledge of risk factors associated with HIV-1 and syphilis among men taking part in MSM behaviors in saunas. Of the 1,101 men who participated, 1,000 turned in usable questionnaires. For 40% of the participants it was their first time to be tested for HIV-1. Chen et al.’s main conclusions were the following. (a) MSM activity is a high-risk category for HIV-1 and syphilis infections in Taiwan, and ongoing research is required to monitor seroprevalence rates and to investigate associated risk factors; (b) the rate of condom usage within Taiwan’s male homosexual community is unacceptably low, thus calling for intensive education efforts; and (c) a positive correlation exists between nonmedical drug use and HIV-1 infections, especially when drugs are used prior to visits to gay saunas (see also [2, 41, 42]).

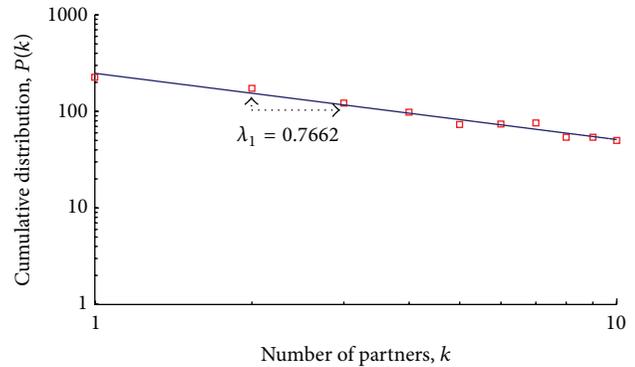


FIGURE 2: Scale-free distribution of sex partners among highly active homosexual males in Taipei/Taichung over three months (log-log axis). The low scaling exponent value ($\lambda = 0.7662$) indicates that individuals in this population tend to have many sexual encounters with different partners.

4. Simulation Model

In addition to saunas and bars, Chen et al.’s [40] questionnaire addressed structural and behavioral aspects of the male homosexual populations in Taipei and Taichung, Taiwan’s two largest cities. Topics included frequencies of visiting high-risk places, frequencies of change in sexual partners, behaviors leading to new connections, relationship durations, condom usage, and attitudes regarding HIV testing. Demographic data were deemphasized—no effort was made to look for associations between HIV infection and age, marital status, education level, or social behavior, since the study goal was to determine risk for entire networks to reflect the HIV epidemic among male homosexuals in Taiwan.

During model construction, high- and low-risk subpopulations were distinguished according to number of partners, condom usage, and a “faithfulness factor” indicating long-term monogamy. The term free link was used to describe situations in which a pair of agents has multiple partners and fixed links to describe two agents in a long term, multiyear fixed relationship. Topological features were incorporated into the network to take advantage of small-world, scale-free, and other complex network model properties. This required distributing agent links in a manner that ensures power-law degree distribution properties over a period of many years [14, 15, 21–23].

Sample data on cumulative numbers of sex partners during the preceding three months produced curves representative of a power-law pattern—specifically, with a low scaling exponent of 0.7662 for the number-of-partners distribution $k \geq 1$ (Figure 2). Scale-free networks show a cumulative power-law distribution decay in the form $P(k) \approx k^{-\lambda_1}$, where k denotes the distribution of number of partners and λ_1 the scaling exponent.

As stated above, the survey data only cover members of the high-risk subpopulation; the majority of male homosexuals in the two target cities might express a similar pattern, but with different scaling exponents and curve shapes. Since the goal was to achieve a power-law degree distribution among

TABLE 1: Simulation parameters for our proposed agent-based HIV epidemic model.

Attribute	Data type	Description
Population size	Integer	Total number of agents.
Alpha pop. size	Integer	
Beta pop. size	Integer	
Gamma pop. size	Integer	
Max contacts	Integer	Maximum number of sexual contacts for one agent in one month.
p	Real	
q	Real	
NC	Integer	Number of sexual contacts for one agent in one month.
Delta	Integer	Number of sexual contacts with fixed partners.
Epsilon	Integer	Number of sexual contacts with nonfixed partners.

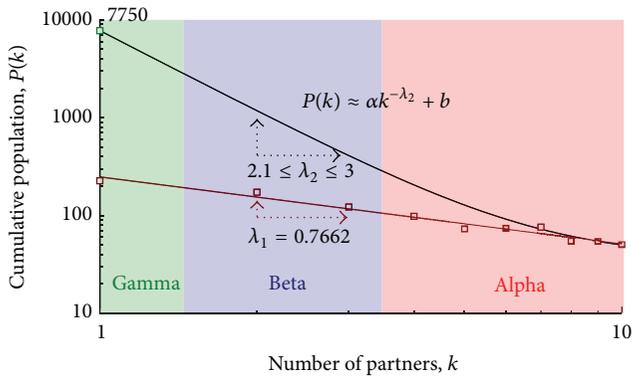


FIGURE 3: General distribution curves for the alpha, beta, and gamma subpopulations.

all simulation agents over a long time period, a Pf function or class of functions representing an accurate distribution of agent links was assumed. A plot of this function (generally expressed as $P(k) \approx \alpha k^{-\lambda_2} + b$) is shown in Figure 3.

Since individuals who have multiple sex partners are more likely to express different behaviors than those with only one partner, the curves can be divided into different sections representing subpopulations with unique profiles. The population in Figure 3 is divided into three clusters: alpha, in which a large number of sex partners are correlated with high-risk behaviors among a small number of individuals; beta, larger than alpha and characterized by a sexually active population expressing less extreme behaviors than alpha agents; and gamma, the largest population segment, representing the rest of the population. As suggested by the power-law pattern, each gamma individual has only one sex partner. The beta class was established because the gap between alpha and gamma patterns was considered too large to reflect real-world situations.

5. Design and Implementation of HIV Simulation Model

The proposed model consists of multiple layers, with each playing a role in describing and representing a specific

function or phenomenon. Layer definition determines the roles of other layers in the simulation. Layers communicate with each other to exchange information regarding their computation results. Five layers were established for this study: agent, agent behavior, links, contact frequency, and epidemic. Layers can be categorized as belonging to an agent-based model consisting of agents and agent behavior layers, a social network model consisting of links and contact frequency layers, or an epidemic model. As explained above, a social network model has spatial, temporal, and intrinsic properties that evolve over time. A mix of global and individual behaviors was established to simulate agent influences on whole populations, but in a manner that reflects intrinsic societal changes via behavioral changes among a large number of individual agents. Societal global structure is defined in the agent layer, spatial concepts are reflected in the links layer, temporal concepts appear in the contact frequency layer, and intrinsic changes are modeled in the agent behavior layer. Suites of simple epidemiological progress statuses were added to the epidemic layer.

5.1. Agent Layer. The agent layer can be variously defined in terms of population structure, agent pool (Table 1), or individual agent characteristics (Table 2). In this study it represents three sexual activity patterns. The minority of extremely active agents (who play an important simulation role due to their high-risk behaviors) are found in the layer core. The second pattern consists of a number of sexually active agents that is far from a majority but exceeds the first pattern—that is, higher-than-average sexual activity, but lower risk behaviors in terms of condom usage and relationship duration. The third pattern, consisting of the large majority of relatively inactive individuals in a population, represents a pool with infection potential, but at a much lower level of risk.

5.2. Agent Behavior Layer. This layer is defined as initial agent behavior and environmental adaptation (Figure 4). The first of two major behavioral changes that can be tested in an HIV epidemic simulation is an increase in condom usage—a parameter that serves as a moderating factor. However, according to the data used in this study, there is no indication of an overall increase in condom usage in Taiwan’s gay

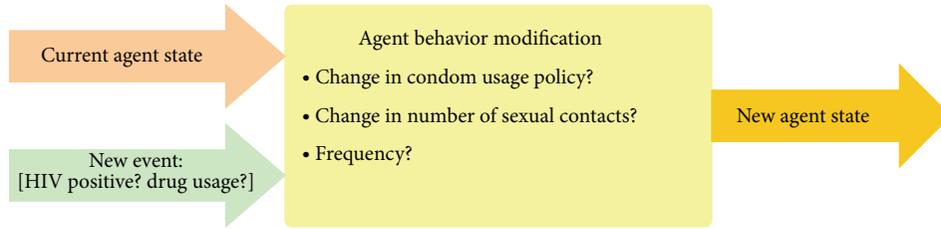


FIGURE 4: Layered agent behavior implementation.

TABLE 2: Agent attributes.

Attribute	Data type	Description
ID	Integer	Serial number identifying agent in model.
State	Symbol	Epidemiological progress status.
Timer	Integer	Number of months in each epidemiological progress state.
Type	Symbol	Subpopulation type.
Fixed partners	Integer	Number of an agent's fixed partners.
Free partners	Integer	Number of an agent's nonfixed partners.
Fixed links	Integer array	IDs of fixed partners.
Free links	Integer array	IDs of nonfixed partners.
Condom usage rate	Real	1 = always use condom and 0 = never use condom.
Drug usage rate	Real	Rate of drug use.

community, despite research showing that condom usage generally increases when agents are made aware of their HIV-positive statuses [40–42]. We assumed that (a) fixed link couples in long-term monogamous relationships were much less likely to use condoms than free link agents and (b) free link agents would be more careful during sexual encounters and more likely to insist that their partners use condoms.

The second behavioral change incorporated into simulations was willingness to be tested for HIV. It is difficult to track changes in testing behaviors due to the lack of statistical data over the past 20 years. Still, this factor is important in terms of behavioral decisions made by HIV-positive agents who are aware of their status, including condom usage and encouraging their peers and partners to be tested (Figure 4).

5.3. Links Layer. This layer reflects the spatial aspect of agent connections and determines the importance of epidemic size. Since new links support new contacts, they are considered a key HIV propagation vector. As mentioned in an earlier section, the majority of individuals in the model are involved in fixed, long-term relationships that can be measured in terms of years. Distinctions between free and fixed link relationships are important because agents generally make behavioral decisions based on context—frequency of visiting high risk locations, frequency of sex partners, relationship duration, condom usage, attitudes toward HIV testing, and so on. These distinctions also reflect disease propagation outside of high-risk subpopulations. Even though low-risk populations mostly consist of long-term monogamous relationships, a small number of free links can act as significant vectors for propagating the HIV virus. Once

the virus enters a gamma subpopulation, the spread of the disease may be slow, but the overall number of infected individuals can be large due to the population size. Precise knowledge of the gamma population size is unnecessary; it is assumed to be much larger than the alpha and beta populations. However, screening alpha and beta populations is very important because they represent the core of sexual activity.

Fixed links are distributed among agents in the same subpopulation, with most having at least one partner during each simulation. Free links are distributed within and across various subpopulations, with the number of potential links for any agent dependent upon the population in question: alpha agents have more free links than beta agents, and beta agents have more free links than gamma agents. According to Liljeros et al. [21–23], the number of sexual contacts for any individual follows a power-law curve. We therefore assumed that any free link distribution of contacts over the lifetime of a single agent in any population also reflects a power-law pattern.

Assortative and disassortative population patterns reflect mixes of free and fixed link agents. Statistically speaking, even if an alpha population is small, an alpha agent is more likely to have free links with other alpha agents, resulting in easily identifiable assortative patterns. The same is true within beta subpopulations and to a lesser degree between alpha and beta subpopulations. Gamma populations are most likely to reflect assortative patterns, although some disassortative patterns are bound to appear due to size, despite the low probability of any individual gamma agent having free links. Free links

in any gamma subpopulation are statistically disassortative, meaning that gamma agents with free links are very likely to connect with alpha or beta agents with numerous free links.

5.4. Contact Frequency Layer. The contact frequency layer reflects the temporal aspect of links between agents. Sexual networks are considered dynamic in two ways. The first is the number of new contacts established by an agent, with existing links lost when new links are created. It is important to understand the underlying mechanism, since creating a link with a new agent can expose a large portion of the network to HIV infection. The second aspect concerns existing link activity. Since some links are more active than others, the number of contacts with certain agents will be much larger than with other agents. For example, a fixed link between two gamma agents will show more activity than a free link in the same population. Even in an alpha population, a free link may only be activated once or twice. A trade-off exists between the number of links an agent can maintain and link duration: the more links an agent has, the weaker they are and the shorter their lengths are. In contrast, maintaining long-term fixed links may block the establishment of new links.

5.5. Epidemic Layer. The study goal was to model growth in the number of HIV cases based on available data for a sexual contact network consisting of homosexual males. An epidemic model consisting of S (susceptible) and I (infected) individuals was used to simulate epidemiological progress resulting in status change, creating two infected subcategories in the process: AIDS and HIV status (Figure 5). AIDS individuals have the virus but do not yet have the antibodies required for detection. Lack of certainty about HIV status can influence behavioral decisions regarding sexual contacts and condom usage. Infected agents are automatically added to the HIV status subcategory after 6–12 weeks. Although HIV can be detected as early as two weeks following infection, we could not assume that all populations have access to state-of-the-art testing facilities.

All uncontaminated agents have an S status prior to virus exposure. Sexual contact between two S agents, with or without condoms, is insufficient for changing the status of either one; the same is true for sexual contact between two infected or HIV-positive agents, but not for contact between susceptible and infected or HIV-positive agents. If a susceptible agent uses a condom, it will reduce the probability of infection, but not by 100%. Furthermore, unprotected sex between susceptible and infected agents does not automatically result in new infections. As noted in Table 1, p denotes the probability of status change following sexual contact between susceptible and infected agents who use condoms and q denotes the probability of a susceptible agent being contaminated by an infected agent during unprotected anal intercourse.

5.6. System Implementation. Implementation of the proposed five-layer simulation system entails communication between components to execute layer functions. A component diagram of the simulation framework is presented in Figure 6,

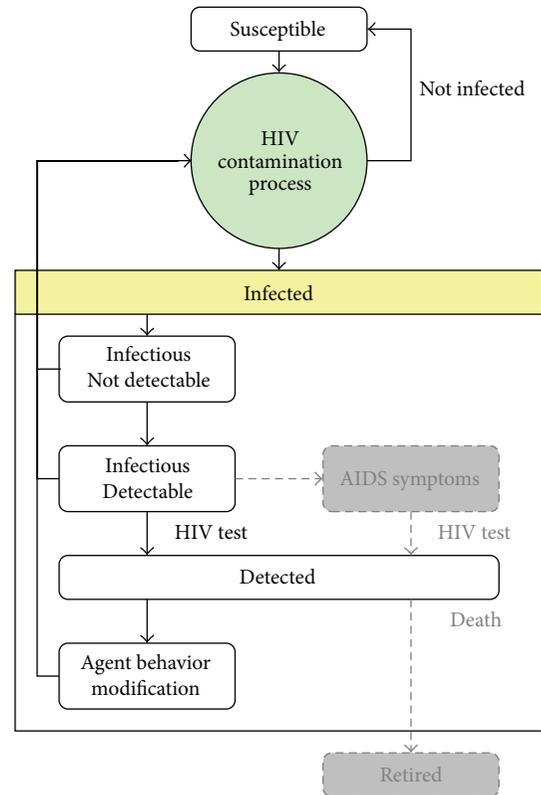


FIGURE 5: HIV and AIDS status development diagram. Our simulation is limited to HIV.

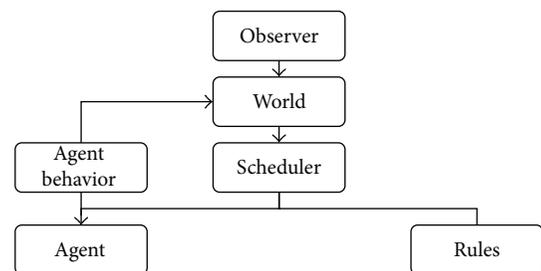


FIGURE 6: HIV epidemic simulation component diagram.

and descriptions of the various components are given in Table 3. Agent layer and behavior are treated as two separate components, with links, contact frequencies, and epidemic layers managed according to the rules component, which includes a power-law shaping rule. The total time complexity of an experiment is $O(\text{Population Size} \times \text{Max Contacts})$.

We used the Borland C++ Builder's visual component library and event-driven programming model to develop the user interface and input/output functions of the simulation system. In addition to providing statistical reports and charts based on HIV epidemic data, the system lets users observe agent infection spreading scenarios in real time. Following compilation and conversion into an executable application, this system can be run on Windows using Dynamic Linked Library files.

TABLE 3: Properties and methods for each simulation component.

Component	Properties	Methods
Observer	Higher in class hierarchy.	Uses information from classes lower in the hierarchy to create screen display.
World	Two-dimensional lattice containing agents.	Defines social network structure. Can be influenced by agent behavior.
Scheduler	Defines time in one-month units.	Defines events that happen to each agent during each time unit and ensures that rules are followed.
Rules	Rules applied to simulation pertaining to contamination, HIV evolution, epidemiological progress status, and so forth.	Communicates rules to schedulers and agents.
Agent	One agent represents one human individual.	Changes status according to epidemic model parameters, capable of interacting with other agents.
Agent behavior	Behavioral rules based on statistical data from surveys.	Controls agent behaviors (e.g., number of contacts, condom usage).

6. Simulation Results

Due to screening limitations, the number of reported cases should not be viewed as an accurate representation of the actual number of cases, meaning that reported curves may not accurately reflect the evolution of the HIV epidemic over a period of years. However, we did assume that the spike in the number of reported cases in Taiwan during the time period covered by this project reflects an increased awareness and willingness to be tested and that a significant number of individuals who have unknowingly carried the HIV virus for a period of time are included in recent statistics. The primary hypothesis is that testing a larger percentage of the population will result in an increase in the number of reported cases. In terms of epidemic simulation, agents who test positive for HIV are more likely to be cautious about their sexual behavior, eventually slowing growth in the number of simulated cases and reducing the number of reported cases. This scenario must be considered when comparing reported and simulated case numbers.

Data from Chen et al. [40], Lai [41], and Lai et al. [42] were used to set the number of links, contact frequency, and drug and condom usage parameters for the alpha and beta populations. The parameters discussed in this section reflect important changes in social behaviors over the past 20 years. However, due to insufficient data on HIV infections among Taiwanese homosexuals for the same period, we had to use rough estimates for the evolution of changes in those behaviors. We did attempt to determine the plausibility of change in agent behavior based on feedback from an earlier generation of homosexuals. For example, we heard one comment that a growing number of young gay men are openly declaring their homosexuality, which may influence changes in alpha and beta populations. Another important change is the sharp increase in drug usage among young adults between the ages of 20 and 30, meaning that researchers must address the question of how drug usage affects safe sex practices and other behaviors. We will review models for estimating the effects of these new factors in the following sections.

6.1. Dynamic Change in the Population. In the proposed model, the overall population was divided into alpha, beta, and gamma subpopulations. The alpha and beta clusters increased in size based on the growing number of Taiwanese men openly declaring their homosexuality during the past 20 years; however, the lack of hard data on this trend makes it impossible to establish an accurate figure for simulation purposes. Since the proposed model is based on recently gathered information, such data are required in order to achieve a close fit between a simulation and the actual number of reported HIV cases.

In Figure 7, the curve labeled “No Increase” represents the number of simulated HIV cases without considering alpha and beta subpopulation increases. The curve clearly shows fewer infections than the actual situation—the slope flattens out over the long term, which is one result of saturation in the number of infected agents in the most sexually active subpopulations. One way to increase the number of infected cases would be to increase the number of exchanges between members of the alpha and beta subpopulations with members of the gamma subpopulation, but doing so would contradict the assumption of low levels of sexual activity for gamma individuals with anyone but their fixed partners. An alternative approach is to increase the number of alpha and beta agents.

Model 1 in Figure 7 reflects the potential for significant change in subpopulation size. While such changes are irregular, they can exert strong influences on individual behaviors—note the number of male homosexuals who “came out” in Western countries in the 1960s and 1970s. In model 1 we assumed three spikes in Taiwan during the past 20 years that resulted in significant increases in alpha and beta subpopulations and respective increases in the numbers of simulated HIV cases [43].

Model 2 reflects the assumption that growth in all three subpopulations is the result of linear increases over many years. While it is unlikely that this trend occurred in homosexual communities, the assumption does underscore the point that such growth generally emerges from gradual

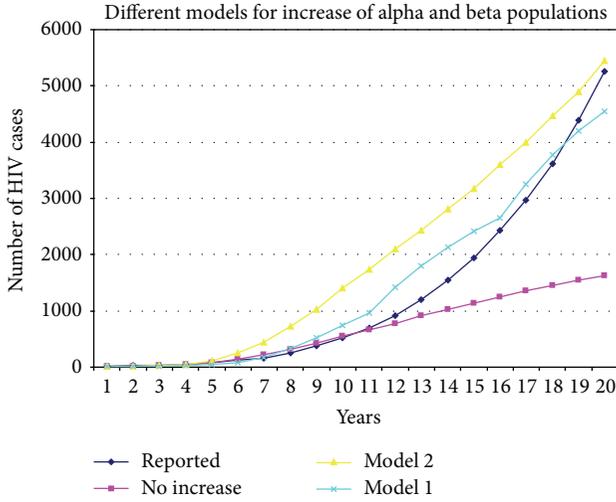


FIGURE 7: Results for two models of alpha and beta population increases.

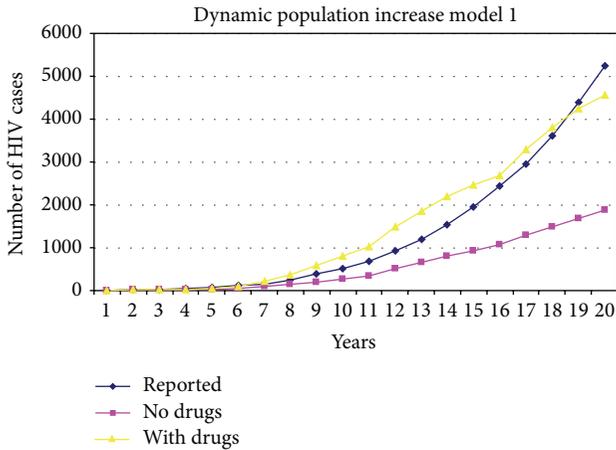


FIGURE 8: Comparison of model 1 based dynamic population increases with and without drug usage.

processes. The data indicate that simulations of dynamic increases in alpha and beta subpopulations produced more accurate results than simulations of static subpopulations. It is very likely that modeling this dynamic is dependent on cultural and spatial factors, thus making the model Taiwan-specific.

6.2. Impact of Drugs. TCDC-sponsored research conducted by Lai [41] and Lai et al. [42] confirms that nonmedicinal drug usage exerts a major impact on the spread of HIV in Taiwan. As stated earlier, we did not address hypodermic needle sharing behavior (which occurs among both heterosexuals and homosexuals) in the present study. A secondary impact can be traced to carelessness in practicing safe sex while being under the influence of drugs [40]. Figure 8 presents the results of two simulations, one considering drug usage and one not. Both simulations assume the same sharp increases in alpha and beta populations as stated in Figure 7, model 1. The results

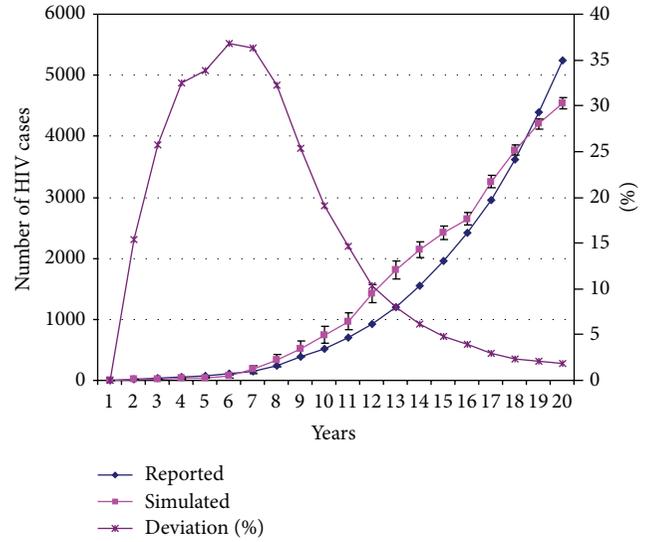


FIGURE 9: HIV infection simulation data (100 runs), including standard deviation at each point (left axis, “simulated”) and relative to the value of each point (right axis, “deviation” percentage).

indicate significant effects from drug usage, especially in the alpha and beta subpopulations.

6.3. Simulation Results Discussion. Certain simulation parameters (e.g., link distribution) were implemented as random variables. The basic time unit was one month, with results reported for each year. Results from a simulation with dynamic increases in alpha and beta subpopulations, drug usage, and HIV testing are presented in Figure 9. We executed 100 runs of each simulation in an effort to reduce statistical bias. Each point along simulated curves in the figure represents the standard deviation over 100 runs; error bars represent upper and lower point values. Each point on the deviation curve represents the standard deviation over 100 runs for each point pegged to 100% of the current point value, meaning that the curve represents the relative error percentage at each point in relationship to the graph’s right axis. Relative deviation from upper and lower values obtained over 100 runs was higher when the number of infected individuals was smaller, though it never exceeded 37% of the current point value. In terms of number of cases, the deviation between upper and lower values did not exceed 147 cases over an average of 1,421 for the 12th year. This is approximately 10% of the overall deviation, indicating a higher value of 1,495 infected cases and a lower value of 1,347. We also observed a connection between higher numbers of infected cases and lower dispersion levels. Combined, these results suggest that the simulations attained acceptable levels of validity in terms of reproducibility (since dispersion was limited) and credibility (since deviations across multiple runs did not affect global trends based on different policies).

Figure 10 presents simulation and deviation results for different policies with dynamic increases in alpha and beta populations using model 1. The patterns of the deviation curve shapes and values are similar to those for actual cases.

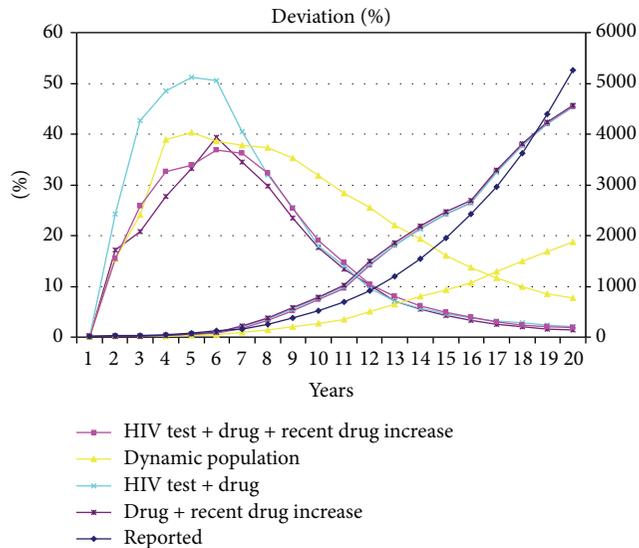


FIGURE 10: Simulation results (right axis) and percentage of relative standard deviation (left axis) for different policies. Dynamic population increases are based on model 1.

Accordingly, it is possible to provide good estimates from the simulation results across a range of accuracy values.

7. Conclusion

The original study goal was to establish a foundation for constructing a model capable of predicting the spread of HIV among homosexual males living in two cities in northern Taiwan. However, the project evolved into a more limited effort to create a reliable model for an agent-based simulation of a sexual activity network. It is assumed that an approach that focuses on individual agent behavior offers a new path for STI simulations. Furthermore, the results indicate that implementing rules derived from social survey statistics can, with some adjustment, accurately reflect global population behaviors that are observable in the real world. In other words, simulations in which specific behaviors are addressed on an individual level can reflect global observations without introducing unreasonably large numbers of societal rules. Such results would reflect the small-world properties inherent to scale-free networks. In the proposed model, the only rules applied at a societal level were increases in alpha and beta populations and increased drug usage. These rules were minimized to reflect global trends associated with cultural and individual tendencies. It is impossible to implement such behaviors on an agent level when a study population is as specific and limited as male homosexuals in two Taiwanese cities.

The model can be modified and improved using new data from field studies. It would be especially interesting to keep track of changes in Taiwan over time, since acknowledgment and limited acceptance of homosexuality represent a significant shift in societal attitudes. Clustering results by generation may provide insights into new behavioral trends as they emerge and improve our understanding of the effects

of such factors as drug and condom usage. The biggest challenge may be organizing and making a commitment to a research project that will require several decades to complete.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The work was supported in part by a Grant from the Republic of China National Science Council (MOST-103-2221-E-182-052). The work was supported in part by the High Speed Intelligent Communication (HSIC) Research Center, Chang Gung University, Taiwan.

References

- [1] World Health Organization, *AIDS Epidemic Update*, WHO, 2009.
- [2] Taiwan Centers for Disease Control (TCDC), *Statistics of Communicable Diseases and Surveillance Report, Republic of China, 2008*, TCDC, Taipei City, Taiwan, 2009, <http://www.cdc.gov.tw/uploads/files/201210/8bb92a35-39e7-4bbd-954a-6fbf303edb98.pdf>.
- [3] Y. Jia, M. H. Aliyu, and Z. J. Huang, "Dynamics of the HIV epidemic in MSM," vol. 2014, Article ID 497543, 3 pages, 2014.
- [4] X. Li, H. Lu, C. Cox et al., "Changing the landscape of the HIV epidemic among MSM in China: results from three consecutive respondent-driven sampling surveys from 2009 to 2011," *BioMed Research International*, vol. 2014, Article ID 563517, 10 pages, 2014.
- [5] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [6] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, no. 1, pp. 47–97, 2002.
- [7] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [8] M. E. J. Newman, "Models of the small world: a review," *Journal of Statistical Physics*, vol. 101, no. 3-4, pp. 819–841, 2000.
- [9] M. E. J. Newman and D. J. Watts, "Scaling and percolation in the small-world network model," *Physical Review E*, vol. 60, no. 6, pp. 7332–7342, 1999.
- [10] S. C. Brailsford, A. K. Shahani, R. B. Roy, and S. Sivapalan, "Simulation modelling for HIV infection and AIDS," *International Journal of Bio-Medical Computing*, vol. 31, no. 2, pp. 73–88, 1992.
- [11] A. D. Paltiel, J. A. Scharfstein, G. R. Seage III et al., "A Monte Carlo simulation of advanced HIV disease: application to prevention of CMV infection," *Medical Decision Making*, vol. 18, no. 2, pp. S93–S105, 1998.
- [12] D. Peterson, K. Willard, M. Altmann, L. Gatewood, and G. Davidson, "Monte Carlo simulation of HIV infection in an intravenous drug user community," *Journal of Acquired Immune Deficiency Syndromes*, vol. 3, no. 11, pp. 1086–1095, 1990.
- [13] P. S. Bearman, J. Moody, and K. Stovel, "Chains of affection: the structure of adolescent romantic and sexual networks," *American Journal of Sociology*, vol. 110, no. 1, pp. 44–91, 2004.

- [14] J. H. Jones and M. S. Handcock, "An assessment of preferential attachment as a mechanism for human sexual network formation," *Proceedings of the Royal Society B: Biological Sciences*, vol. 270, no. 1520, pp. 1123–1128, 2003.
- [15] J. H. Jones and M. S. Handcock, "Sexual contacts and epidemic thresholds," *Nature*, vol. 423, no. 6940, pp. 605–606, 2003.
- [16] A. S. Klondahl, J. J. Potterat, D. E. Woodhouse, J. B. Muth, S. Q. Muth, and W. W. Darrow, "Social networks and infectious disease: the Colorado Springs Study," *Social Science & Medicine*, vol. 38, no. 1, pp. 79–88, 1994.
- [17] M. Kretzschmar, "Sexual network structure and sexually transmitted disease prevention: a modeling perspective," *Sexually Transmitted Diseases*, vol. 27, no. 10, pp. 627–635, 2000.
- [18] M. Morris, "Sexual networks and HIV," *AIDS*, vol. 11, supplement A, pp. S209–S216, 1997.
- [19] J. J. Potterat, L. Phillips-Plummer, S. Q. Muth et al., "Risk network structure in the early epidemic phase of HIV transmission in Colorado Springs," *Sexually Transmitted Infections*, vol. 78, no. 1, pp. i159–i163, 2002.
- [20] P. Benkimoun, "Quarante millions de personnes vivent avec le virus du sida," *Le Monde*, p. A1, 2004.
- [21] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Åberg, "Social networks: the web of human sexual contacts," *Nature*, vol. 411, no. 6840, pp. 907–908, 2001.
- [22] F. Liljeros, C. R. Edling, and L. A. N. Amaral, "Sexual networks: implications for the transmission of sexually transmitted infections," *Microbes and Infection*, vol. 5, no. 2, pp. 189–196, 2003.
- [23] F. Liljeros, C. R. Edling, H. E. Stanley, Y. Åberg, and L. A. Amaral, "Distributions of number of sexual partnerships have power-law decaying tails and finite variance," Working Paper, 2003.
- [24] A. Schneeberger, C. H. Mercer, S. A. J. Gregson et al., "Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe," *Sexually Transmitted Diseases*, vol. 31, no. 6, pp. 380–387, 2004.
- [25] E. Teweldemedhin, T. Marwala, and C. Mueller, "Agent-based modelling: a case study in HIV epidemic," in *Proceedings of the 4th International Conference on Hybrid Intelligent Systems (HIS '04)*, pp. 154–159, Kitakyushu, Japan, December 2004.
- [26] P. Zvirinsky and J. Laansky, "A bottom-up approach to immune system modelling," in *Proceedings of Intelligent and Adaptive Systems in Medicine*, pp. 163–169, Prague, Czech Republic, 2003.
- [27] C. Y. Huang, C. T. Sun, J. L. Hsieh, and H. Lin, "Simulating SARS: small-world epidemiological modeling and public health policy assessments," *Journal of Artificial Societies and Social Simulation*, vol. 7, no. 4, 2004, <http://jasss.soc.surrey.ac.uk/7/4/2.html>.
- [28] C.-Y. Huang, C.-T. Sun, and H.-C. Lin, "Influence of local information on social simulations in small-world network models," *Journal of Artificial Societies and Social Simulation*, vol. 8, no. 4, 2005.
- [29] G. Xiaoguang and L. Zhicheng, "Study on multi-agent simulation of new product market diffusion," *System Engineering Theory and Practice*, vol. 12, pp. 60–63, 2003.
- [30] H. Zhou, "Research of the small-world character during Rumor's propagation," *Journal of Wuhan University of Science and Engineering*, vol. 1, pp. 113–116, 2005.
- [31] D. J. Watts, *Six Degrees: The Science of a Connected Age*, W. W. Norton and Company, New York, NY, USA, 2003.
- [32] W. O. Kermack and A. G. McKendrick, "Contributions to the mathematical theory of epidemics," *Proceedings of the Royal Society London*, vol. 115, no. 772, pp. 700–721, 1927.
- [33] J. A. Catania, J. T. Coates, S. Kegels, and T. M. Fullilove, "The population-based AMEN (AIDS in multi-ethnic neighborhoods) study," *The American Journal of Public Health*, vol. 82, no. 2, pp. 284–287, 1992.
- [34] M. Morris, "Data driven network models for the spread of infectious disease," in *Epidemic Models: Their Structure and Relation to Data*, D. Mollison, Ed., pp. 302–322, Cambridge University Press, Cambridge, UK, 1995.
- [35] S. Shiboski and N. S. Padian, "Population- and individual-based approaches to the design and analysis of epidemiologic studies of sexually transmitted disease transmission," *The Journal of Infectious Diseases*, vol. 174, supplement 2, pp. S188–S200, 1996.
- [36] S. Haraldsdottir, S. Gupta, and R. M. Anderson, "Preliminary studies of sexual networks in a male homosexual community in Iceland," *Journal of Acquired Immune Deficiency Syndromes*, vol. 5, no. 4, pp. 374–381, 1992.
- [37] J. L. Wylie and A. Jolly, "Patterns of chlamydia and gonorrhoea infection in sexual networks in Manitoba, Canada," *Sexually Transmitted Diseases*, vol. 28, no. 1, pp. 14–24, 2001.
- [38] P. V. Marsden, "Network data and measurement," *Annual Review of Sociology*, vol. 16, pp. 435–463, 1990.
- [39] L. A. N. Amaral, A. Scala, M. Barthélemy, and H. E. Stanley, "Classes of small-world networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 21, pp. 11149–11152, 2000.
- [40] M. A. Chen, S. F. Lai, Y. C. Lan, K. H. Chen, and Y. J. Chen, "Molecular epidemiology of HIV-1 infection among injecting drug users in Taiwan-report of an emergent situation," in *Proceedings of the 7th International Congress on AIDS in Asia and the Pacific*, Kobe, Japan, 2005.
- [41] S. F. Lai, *Molecular epidemiology of HIV-1 infection in men who have sex with men from gay Saunas in Taiwan [Ph.D. Dissertation]*, Institute of Public Health, School of Medicine, National Yang-Ming University, Taipei, Taiwan, 2003.
- [42] S. F. Lai, C. P. Hong, Y. C. Lan et al., "Molecular epidemiology of HIV-1 in men who have sex with men from gay Saunas in Taiwan from 2000 to 2003," in *Proceedings of the 15th International AIDS Conference*, Bangkok, Thailand, 2004.
- [43] T. Yamamoto and I. Satoko, "Fighting a rising tide: the response to AIDS in East Asia," Japan Center for International Exchange, 2006.

Research Article

Prediction of BP Reactivity to Talking Using Hybrid Soft Computing Approaches

Gurmanik Kaur, Ajat Shatru Arora, and Vijender Kumar Jain

Electrical and Instrumentation Engineering Department, Sant Longowal Institute of Engineering & Technology, Deemed University (Established by Government of India), Longowal, Sangrur District, Punjab 148106, India

Correspondence should be addressed to Gurmanik Kaur; mannsliet@gmail.com

Received 3 June 2014; Revised 25 August 2014; Accepted 29 August 2014; Published 21 September 2014

Academic Editor: Ezequiel López-Rubio

Copyright © 2014 Gurmanik Kaur 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.

High blood pressure (BP) is associated with an increased risk of cardiovascular diseases. Therefore, optimal precision in measurement of BP is appropriate in clinical and research studies. In this work, anthropometric characteristics including age, height, weight, body mass index (BMI), and arm circumference (AC) were used as independent predictor variables for the prediction of BP reactivity to talking. Principal component analysis (PCA) was fused with artificial neural network (ANN), adaptive neurofuzzy inference system (ANFIS), and least square-support vector machine (LS-SVM) model to remove the multicollinearity effect among anthropometric predictor variables. The statistical tests in terms of coefficient of determination (R^2), root mean square error (RMSE), and mean absolute percentage error (MAPE) revealed that PCA based LS-SVM (PCA-LS-SVM) model produced a more efficient prediction of BP reactivity as compared to other models. This assessment presents the importance and advantages posed by PCA fused prediction models for prediction of biological variables.

1. Introduction

Accurate measurement of BP is essential in epidemiological studies, in screening programmes, in research studies, and in clinical practice to classify individuals, to ascertain hypertension related risks (coronary heart disease, stroke, and kidney failure), and to guide management. Recommendations of several international organisations including the American Heart Association (AHA) [1], British Hypertension Society (BHS) [2], and European Society of Hypertension (ESH) [3] revealed that accuracy of BP measurements is highly associated with the conditions in which the measurements are taken. The observer should be aware of the considerable variability that may occur in BP due to various factors. However, it is not always feasible to control all the factors, but we can minimize their effect by taking them into account in reaching a decision [3].

In clinical practice, talking is one of the most common measurement disturbances influencing BP measurement accuracy [4]. It can contribute to elevated BP reading, termed BP reactivity to talking, that may result in the misdiagnosis of hypertension or in overestimation of the

severity of hypertension and may lead to overly aggressive therapy. Antihypertensive treatment may be unnecessary in the absence of concurrent cardiovascular risk factors [5].

In the past few years, several studies have quantified the effect of talking on BP. Zheng et al. [6] measured BP in healthy subjects under five different conditions including resting, deeper breathing, talking, and head and arm movement and proved that SBP and DBP changed significantly in comparison to the resting condition. Le Pailleur et al. [7] explored a sharp and significant increase in SBP and DBP of hypertensive subjects while talking. Le Pailleur et al. [8] showed an instantaneous rise in SBP and DBP of treated and untreated hypertensive patients under a period of stress talking and a period of counting aloud (active periods).

Zheng et al. [9] demonstrated significantly higher manual and automated MAPs with talking in healthy subjects. Lynch et al. [4] reported that verbal activity is consistently associated with marked elevations in both normotensive and hypertensive subjects. Tardy et al. [10] demonstrated that talking state increased the BP as compared to resting state of the subjects. Lynch et al. [11] described sudden extreme drop in blood pressure in both experimental and clinical situations when

a person is talking about or describing situations of hopelessness and helplessness. Long et al. [12] showed statistically significant increase in BP when speaking compared to when quiet. Hellmann and Grimm [13] investigated the effect of talking on subjects with one previous diastolic blood pressure reading of 90 mm Hg or more and not taking antihypertensive medicines. Blood pressure increased significantly under both talking conditions (reading neutral material for part of the procedure and reading neutral material continuously).

Epidemiological studies from different populations have explored a significant correlation between BP and anthropometric characteristics [14–16]. Therefore, anthropometric variables should be considered to attain an accurate measurement of BP. However, multicollinearity between anthropometric predictor variables has also been reported, which may result in “overfitting” of the prediction model [17–19]. One approach to dealing with multicollinearity is to use PCA, a statistical approach. By using PCA the original data set can be transformed into principal components (PCs) that are orthogonal and are able to explain the maximal variance of the data without losing any information [20, 21].

Soft computing covers computational techniques that offer somewhat “inexact” solutions of very complex problems through modeling and analysis with a tolerance of imprecision, uncertainty, partial truth, and approximation. The successful applications of soft computing approaches in biomedical studies suggest that the impact of soft computing will be felt increasingly in the coming years.

The fusion of a statistical and soft computing approach usually improves the training speed, enhances the robustness of the model, and reduces the calibration error. These models may aid the clinicians in the decision-making process regarding clinical admission, early prevention, early clinical diagnosis, and application of clinical therapies. In this sense, this paper focuses on the development of PCA based soft computing approaches for prediction of BP reactivity to talking, which include conventional statistical method of PCA for data preprocessing. We developed PCA based ANN (PCA-ANN), PCA based ANFIS (PCA-ANFIS), and PCA-LS-SVM models for prediction of BP reactivity to talking in normotensive and hypertensive subjects. The prediction accuracy of developed models was assessed and compared using statistical indices including coefficient of determination (R^2), root mean square error (RMSE), and mean absolute percentage error (MAPE) to select the model that most accurately predicts the BP reactivity.

The rest of the paper is structured as follows. In Section 2, we present the details of data collection. Section 3 deals with the experimental approaches used for data analysis. Section 4 deals with the summary of results obtained. Section 5 describes the discussion and Section 6 concludes with future directions of work.

2. Data Collection

A total of 40 normotensive and 30 hypertensive female subjects among the students, staff, and faculty of Sant Longowal Institute of Engineering and Technology (Deemed

University), Longowal, District Sangrur, Punjab, India, volunteered for this study. Eligible participants had to be over 18 years of age. We excluded the subjects who were pregnant and who had arrhythmias. The institutional research committee approved the research protocol and all participants gave written informed consent before participation.

A standard questionnaire was administered to collect information on anthropometric characteristics of the participants. A specially separated room was used to conduct the study. This ensured minimal interference within the room while the tests were being carried out. The observers involved in the study were trained using the BHS’s BP measurement training materials [22].

To eliminate observer bias, BP was measured using a clinically validated (under standardized measurement conditions), newly purchased, and fully automated sphygmomanometer OMRON HEM-7203 (OMRON HEALTHCARE Co., Ltd., Kyoto, JAPAN) that uses the oscillometric method of measurement. The BP monitor is available with a small cuff (17–22 cm), medium cuff (22–32 cm), and large cuff (32–42 cm). The appropriate size of cuff was determined from the mid-arm circumference of the subject.

Subjects were advised to avoid alcohol, cigarette smoking, coffee/tea intake, and exercise for at least 30 minutes prior to their BP measurement. They were instructed to empty their bladder and sit upright with elbows on table, supported back, and feet flat on the ground, as they are the potential confounding factors. Moreover, they were asked not to talk and move during measurement [1].

After a rest period of 5 minutes [1], the measurements were performed four times repeatedly at an interval of one minute. First measurement was discarded and the average of last three measurements was taken into account. Subsequently, the same measurement protocol was repeated under talking phase during which the observer asked each subject to “tell me about your work in detail” [4]. During talking phase, the observer only talked to the subject to maintain the flow of conversation, making every possible effort to talk minimally. To improve the reliability of measurements, the subjects were examined for a week [3].

3. Experimental Methods

Data were expressed as mean \pm SD. A paired t -test was used to assess the difference between measurements of resting and talking conditions.

3.1. PCA. Firstly, Bartlett’s test of sphericity [23] and Kaiser Meyer Olkin (KMO) measure of sampling adequacy [24] were applied to check the suitability of data for application of PCA.

Bartlett’s test of sphericity tests the null hypothesis that the correlation matrix is an identity matrix or there is no relationship between predictor variables. Consider

$$\chi^2 = - \left\{ [N - 1] - \left[\frac{2K + 5}{6} \right] \right\} \log_e |R|, \quad (1)$$

where χ^2 is chi-square, N is sample size, K is number of predictor variables, \log_e is natural log, and $|R|$ is determinant of the correlation matrix.

KMO compares the magnitude of calculated correlation coefficients and partial correlation coefficients. The formula for KMO is given as follows:

$$\text{KMO} = \frac{\sum_{i \neq j} \sum r_{ij}^2}{\sum_{i \neq j} \sum r_{ij}^2 + \sum_{i \neq j} \sum a_{ij}^2}, \quad (2)$$

where $\sum_{i \neq j}$ is sum over all variables in the matrix when variable $i \neq j$, r_{ij} is Pearson correlation coefficient between variables i and j , and a_{ij} is partial correlation coefficient between variables i and j .

KMO index ranges, from 0 to 1, should be greater than 0.6 for the PCA to be considered appropriate.

PCA is a multivariable statistical analysis technique. The objective of PCA is to remove the multicollinearity problem and reduce the number of predictor variables and transform them into PCs which are independent linear combinations of the original data set and account for the maximum possible variance of the original data set so that adequate information from the original data set can be extracted [20, 21].

The eigenvalues of the standardized matrix are calculated from

$$|C - \lambda I| = 0, \quad (3)$$

where C is correlation matrix of the standardized data, λ is eigenvalues, and I is identity matrix. The weights of the variables in the PCs are then obtained by

$$|C - \lambda I| W = 0, \quad (4)$$

where W is matrix of weights.

To evaluate the influence of each predictor variable in the PCs, varimax rotation was used to obtain values of rotated factor loadings. These loadings represent the contribution of each predictor variable in a specific PC. The PCs used for the prediction of BP reactivity to unsupported back were obtained through multiplication of the standardized data matrix by weights (W) [25].

3.2. ANN. ANN's customary architecture was composed of an input layer, an output layer, and one or more intervening layers, also referred to as hidden layers to capture the nonlinearity in data.

Figure 1 shows an ANN model consisting of N nodes in input layer, one hidden layer with h hidden nodes, and an output layer with one node.

Network is trained by presenting one pair of input-output vector at a time. The weighted sum of inputs calculated at t th hidden node is

$$\text{NET}_t = \sum_{i=1}^N w_{ti} x_i + b_t, \quad (5)$$

where w_{ti} is weight on connection from the i th to the t th node, x_i is input data from input node, N is total number of input nodes, and b_t is bias on the t th hidden node.

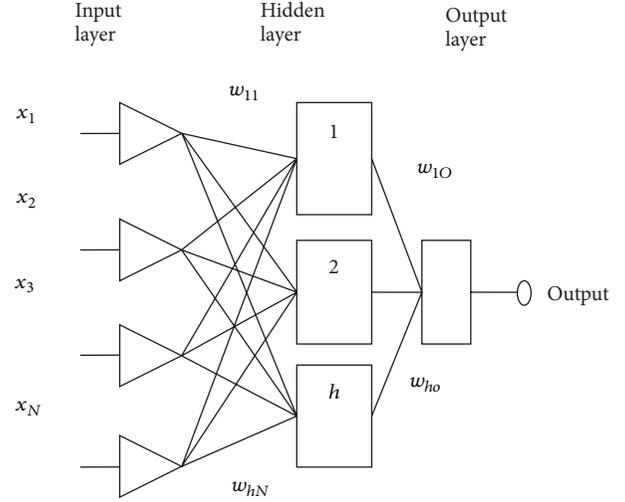


FIGURE 1: Architecture of ANN.

Each hidden node uses a tangent sigmoid transfer function to generate an output, say Z_t , between 0 and 1. The outputs from each hidden nodes, along with the bias b_0 on the output node, send to the output node and weighted sum becomes

$$\text{NET} = \sum_{t=1}^h v_t Z_t + b_0, \quad (6)$$

where h is total number of hidden nodes and v_t is weight from the t th hidden node to the output node.

The weighted sum NET becomes the input to the linear transfer function of the output node and the predicted output is

$$\hat{Y} = f(\text{NET}). \quad (7)$$

And then the second phase of the BP algorithm, adjustment of the connection weights, begins. The parameters of the ANN can be determined by minimizing the following objective function in the training process:

$$\text{SSE} = \sum_{j=1}^n (y_j - \hat{Y}_j)^2, \quad (8)$$

where \hat{Y}_j is output of the network from j th observation.

The sensitivity S_i of the outputs to each of the i th inputs, as partial derivatives of the output with respect to the input, under the assumption that relationship of Y and X is monotone [26], is given as

$$S_i = \frac{\partial \hat{Y}}{\partial X_i} = \sum_{t=1}^h \frac{\partial \hat{Y}}{\partial \text{NET}} \frac{\partial \text{NET}}{\partial Z_t} \frac{\partial Z_t}{\partial \text{NET}_t} \frac{\partial \text{NET}_t}{\partial X_i} \quad (9)$$

$$= \sum_{t=1}^h [f'(\text{NET}) v_t f'(\text{NET}_t) w_{ti}]$$

or

$$\hat{S}_i = \sum_{t=1}^h v_t w_{ti} \quad (10)$$

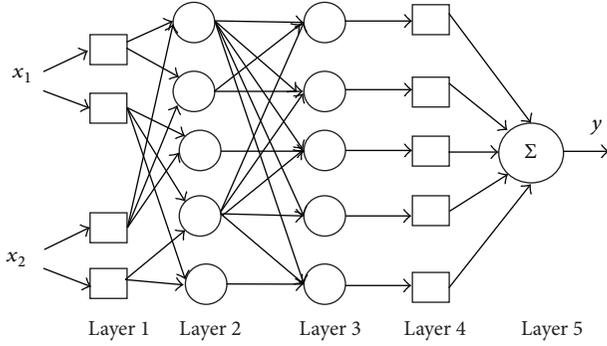


FIGURE 2: Architecture of ANFIS.

with the assumption that $f'(\text{NET})$ and $f'(\text{NET})_i$ are constants. The independent variable with higher relative positive or negative sensitivity has the higher positive or negative impact on dependent variable.

3.3. ANFIS. ANFIS, a multilayer feed forward network, uses neural network learning algorithms and fuzzy reasoning to map an input space to an output space, as shown in Figure 2. It has the ability to combine the verbal power of a fuzzy system with the numeric power of a neural network.

It can construct an input-output mapping based on human knowledge (if-then fuzzy rules) and stipulated input-output data pairs. The parameters of membership function, if-then rule exertion, and output parameters are calculated by training data set. The training algorithm is usually hybrid or back propagation. The ANFIS implements the rules of the form

$$R_r : \text{If } x_1 \text{ is } A_{j1}^{(1)} \dots \text{ and } x_n \text{ is } A_{jn}^{(n)}, \text{ then} \\ y = \alpha_0^{(r)} + \alpha_1^{(r)} \cdot x_1 + \dots + \alpha_n^{(r)} \cdot x_n, \quad (11)$$

where x_1 is independent variable, $A_{j1}^{(1)}$ is a fuzzy linguistic concept, and y is dependent variable.

Input Layer (L_1). Each unit of input layer stores parameters of membership functions to define a membership function that represents a linguistic term.

Input Membership Function Layer (L_2). Each unit of this layer represents a rule. The inputs to a unit $R_r \in L_2$ are degrees of membership which are multiplied to determine the degree of fulfilment τ_r for the rule represented by R_r .

Logical Nodes (L_3). This layer consists of a unit for each rule R_r that computes relative degree of fulfilment as follows:

$$\bar{\tau}_r = \frac{\tau_r}{\sum_{R_i \in L_2} \tau_i}. \quad (12)$$

Each unit of L_3 is connected to all the rule units in L_2 .

Output Membership Function Layer (L_4). Each unit of L_4 computes the output of a rule R_r as

$$O_r = \bar{\tau}_r (\alpha_0^{(r)} + \alpha_1^{(r)} \cdot x_1 + \dots + \alpha_n^{(r)} \cdot x_n). \quad (13)$$

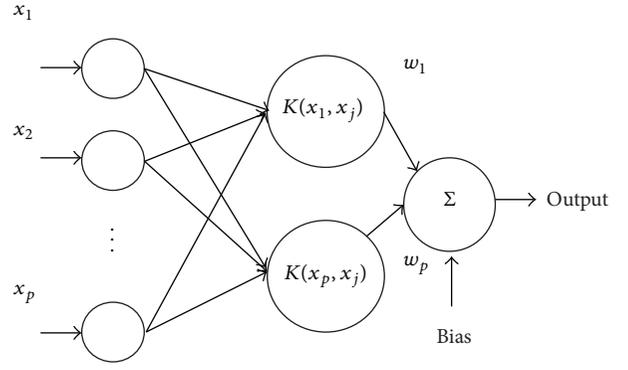


FIGURE 3: Architecture of LS-SVM.

The units L_4 are connected to all units of input layer and to exactly one unit in L_3 .

Output Layer (L_5). It computes the final output y by adding all the outputs from L_4 [27].

3.4. LS-SVM. LS-SVM is an extension of standard support vector machine. It converts the inequality constraints of SVM into equality ones which leads to solving a linear system instead of a quadratic problem, whose convergence speed is faster [28]. It has been widely used in estimation and approximation of function [29]. The architecture of LS-SVM is shown in Figure 3.

Given a set of training data set

$$(x_i, y_i), \quad i = 1, 2, \dots, l, \quad x_i \in R^2, \quad y_i \in R \quad (14)$$

with the input vector x_i and the output vector y_i , the regression function of least square-support vector machine, in feature space F , can be stated as

$$y(x) = w^T \theta(x) + b, \quad (15)$$

where w is weight vector and b is bias. $\theta(x)$ maps the input x into a vector in F .

The model is inferred from the training data set by minimizing the cost function given below

$$\frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{i=1}^l e_i^2, \quad (16)$$

subject to equality constraint

$$y_i = w^T \theta(x_i) + b + e_i, \quad (17)$$

where e_i is error, $i = 1, 2, 3, \dots, l$, and γ is regularization parameter.

Solving this optimization problem in dual space leads to finding the coefficients α_i and b in the following solution:

$$y(x) = \sum_{i=1}^l \alpha_i K(x_i, x) + b, \quad (18)$$

where $K(x_i, x)$ is kernel function.

3.5. Performance Indices Used for Model Comparison. For the comparison of developed models and selection of the optimal among them, performance of models was evaluated using R^2 and RMSE and MAPE.

R^2 is the square of the correlation coefficient between two variables x and y whose n pairs are available as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y_{di})^2}{\sum_{i=1}^n (y_{di} - y_m)^2}. \quad (19)$$

RMSE is the square root of mean square error, given by following equation:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_{di})^2}. \quad (20)$$

MAPE is defined as average of percentage errors, given by following equation:

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - y_{di}}{y_i} \right|, \quad (21)$$

where n is the number of samples, y_i is the predicted value obtained from the model, y_{di} is the actual value, and y_m is the average of the actual values.

The lower the RMSE and MAPE, the better the accuracy of the model in predicting the parameter. Also, the highest R^2 values indicated that the model performed the best [30].

4. Results

Descriptive statistics for each anthropometric characteristic is given as mean and SD in Table 1.

The results of paired t -test demonstrated statistically significant higher SBP, DBP, and mean arterial pressure (MAP), ($P < 0.001$) in talking condition. The mean rise was found to be higher in hypertensive individuals than normotensives, as shown in Table 2. These results are consistent with the recommendations of AHA for BP measurement in humans and experimental animals [1].

Table 3 presents the Pearson's correlation coefficients calculated for all anthropometric variables. High values of correlation coefficient (greater than 0.6) between pairs of anthropometric characteristics [31] revealed the existence of multicollinearity.

Before applying PCA, Bartlett's test of sphericity and Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy were applied to determine whether PCA was suitable for data studied. The results are shown in Table 4. High value of chi-square (χ^2) for Bartlett's test suggests that use of PCA is appropriate ($P < 0.001$) in normotensive and hypertensive subjects. The value of KMO is also greater than 0.6 which indicates that our sample size is enough to apply PCA [32].

The first four PCs (PC1–PC4), explaining more than 5% of total variation, as shown in Table 5, were retained for further analysis.

Rotated component loadings after varimax rotation represent the extent to which the original anthropometric

TABLE 1: Descriptive characteristics of anthropometric characteristics of study sample.

Anthropometric characteristic	Normotensive		Hypertensive	
	Mean	SD	Mean	SD
Age (year)	23.1	1.24	42.83	6.665
Height (m)	1.61	0.03	1.583	0.035
Weight (kg)	55.96	7.29	62.48	10.89
BMI (kg/m ²)	21.55	2.504	23.57	3.497
AC (cm)	26.56	2.45	26.72	2.4

characteristics are influential in forming PCs, as shown in Table 6.

The bold marked loads show the highest correlation between anthropometric characteristic and corresponding component. For both normotensive and hypertensive subjects, weight and BMI were positively highly correlated with PC1 and a negative high correlation between height and PC2 was observed.

Principal score values for assigned PCs were determined by using principal score coefficients.

Moreover, the value of Pearson's correlation (correlation coefficient < 0.6) between PCs, as shown in Table 7, indicates the elimination of multicollinearity effect presented in Table 3.

To develop PCA based soft computing prediction models 80% of data were used for training while entire data set was used for testing. Moreover, data must be normalized to achieve more accurate predictions [38]. The predicted BP reactivity values were denormalized for comparison with the actual values. MATLAB 7.5 version was used to develop the prediction models.

4.1. PCA-ANN. To achieve the best ANN structure for BP reactivity prediction, various structures of feed-forward neural network with different number of neurons in hidden layer were investigated. Finally, with consideration of statistical indices, a structure with two hidden layers, having six nodes in each hidden layer, was developed. There were four input nodes representing the four PCs and one output node representing the BP reactivity to talking. Tangent sigmoid and linear transfer functions were used as activation functions in the hidden and output layers. Back propagation learning algorithm based on Levenberg-Marquardt technique was used [39].

Figures 4 and 5 show the scatter plot between observed and predicted values of SBP, DBP, and MAP reactivity from PCA-ANN model in normotensive and hypertensive subjects, respectively.

4.2. PCA-ANFIS. PCA-ANFIS model was developed using genfis1 with grid partition on data. Different ANFIS parameters were tested in order to achieve the perfect training and maximum prediction accuracy.

Input membership functions "trapmf" and "gauss2mf" were used to predict SBP and DBP reactivities, respectively, in normotensive individuals, whereas membership function

TABLE 2: Results of paired t -test.

Subjects	BP, mm Hg	Mean difference \pm SD	t	P	95% CI of mean difference
Normotensives	SBP	6.31 ± 2.409	16.567	<0.001	5.540 to 7.081
	DBP	5.857 ± 1.584	23.388	<0.001	5.350 to 6.363
	MAP	6.008 ± 1.231	30.854	<0.001	5.164 to 6.402
Hypertensives	SBP	9.634 ± 1.283	41.117	<0.001	9.154 to 10.113
	DBP	7.816 ± 1.44	29.722	<0.001	7.278 to 8.354
	MAP	8.422 ± 1.105	41.757	<0.001	8.009 to 8.834

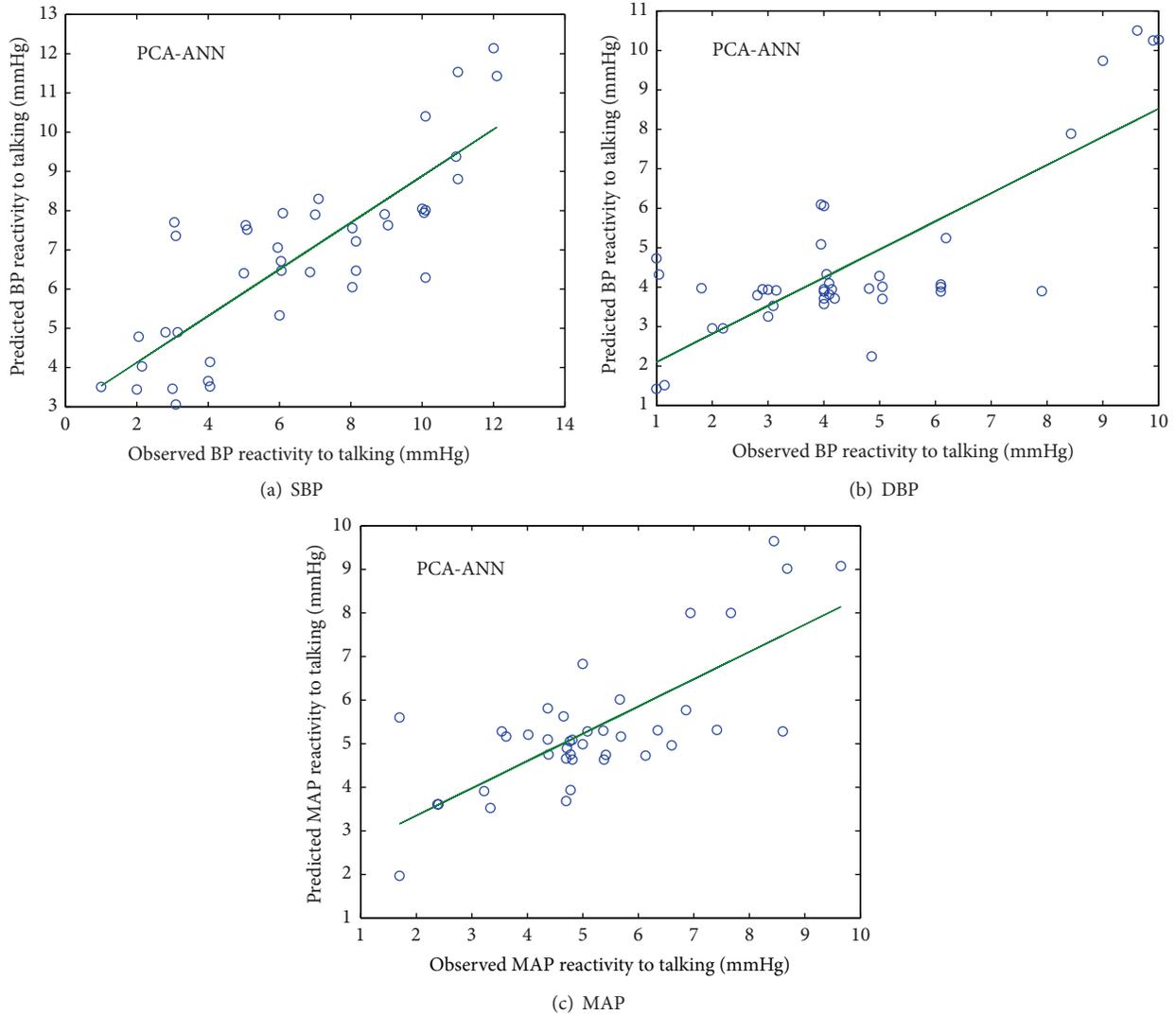


FIGURE 4: Scatter plot between observed and predicted values of SBP, DBP, and MAP reactivity of normotensive subjects using PCA-ANN model.

“psigmf” was used to predict SBP and DBP reactivity in hypertensive individuals. Output membership function “linear” was used.

Other parameters of trained PCA-ANFIS model were number of membership functions = 16, number of nodes = 55, number of linear parameters = 80, number of nonlinear parameters = 32, total number of parameters = 112, and number of fuzzy rules = 16.

The observed and predicted values of SBP, DBP, and MAP reactivity from PCA-ANFIS model for normotensive and hypertensive subjects were plotted in Figures 6 and 7.

4.3. PCA-LS-SVM. A PCA-LS-SVM model using RBF kernel and grid search optimization algorithm with 2-fold cross-validation was developed to obtain the optimal parameter

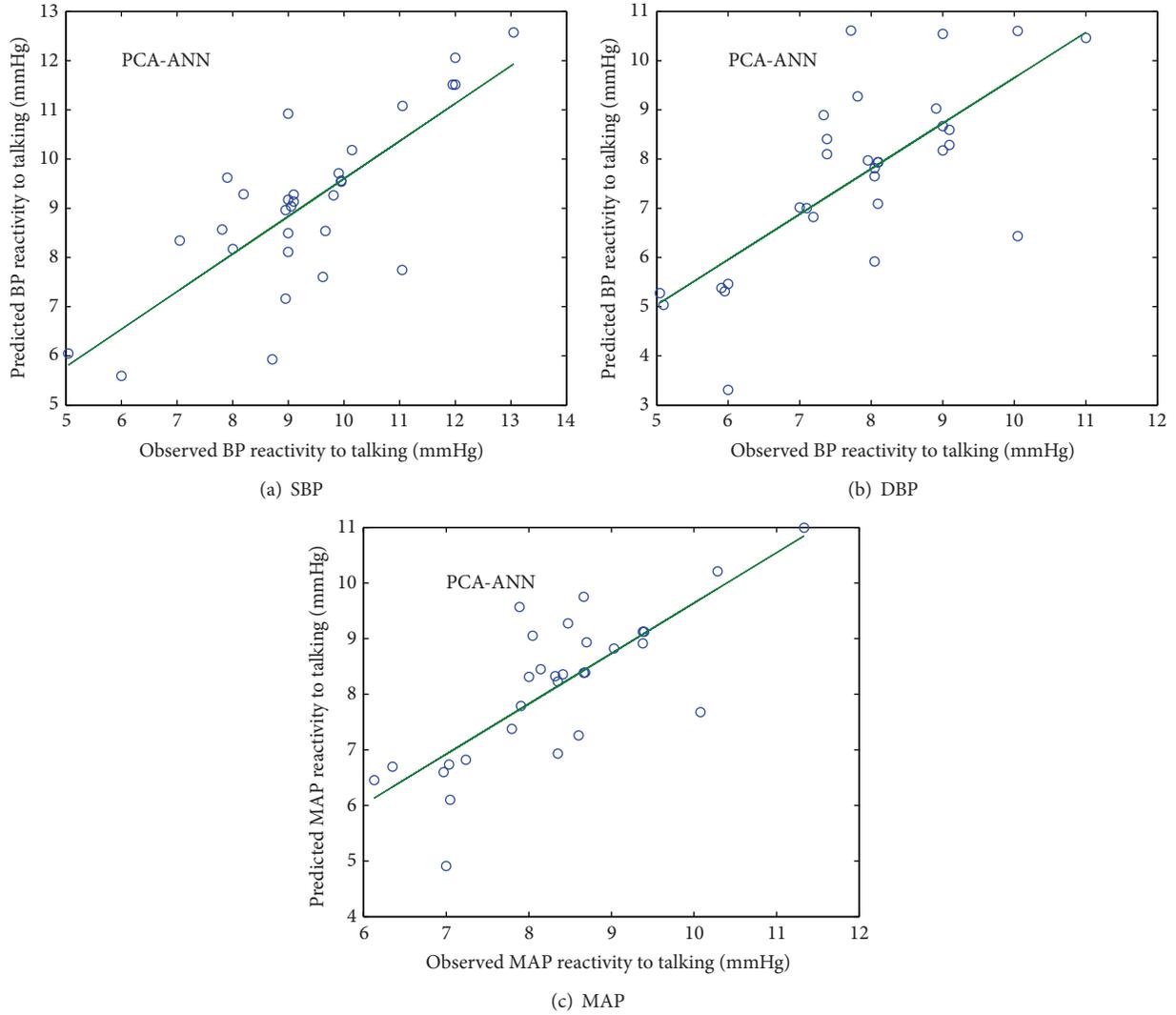


FIGURE 5: Scatter plot between observed and predicted values of SBP, DBP, and MAP reactivity of hypertensive subjects using PCA-ANN model.

TABLE 3: Pearson’s correlation coefficients between pairs of anthropometric characteristics in normotensive and hypertensive subjects.

Variable	Height	Weight	BMI	AC
Age (years)	0.535	0.784*	0.701*	0.668*
	0.113	0.598	0.509	0.585
Height (cms)		0.543	0.237	0.619*
		0.165	0.305	0.021
Weight (Kg)			0.934*	0.743*
			0.885*	0.767*
BMI (Kg/m ²)				0.617*
				0.691*

*It indicates $P < 0.001$; bold values indicate correlations in anthropometric characteristics of hypertensive subjects.

combination [40]. The optimal values of γ (regularization parameter) and σ^2 (squared bandwidth) for normotensive and hypertensive subjects were shown in Table 8.

Figures 8 and 9 show the scatter plot between observed and predicted values of SBP, DBP, and MAP reactivity from PCA-LS-SVM model in normotensive and hypertensive subjects, respectively.

Comparison of statistical indices for the models, as shown in Table 9, revealed that PCA-LS-SVM model has the highest value of R^2 and lowest value of RMSE for the prediction of BP reactivity to talking in normotensive and hypertensive subjects.

5. Discussion

For proper diagnosis and treatment of hypertension, accurate and reproducible BP measurements are essential.

This study confirms and extends previous studies [4, 6–13] by documenting a significant increase in BP with talking. This finding tends to support Weiner et al. [41] suggestion that there may be an association between verbal activity and

TABLE 4: Results of Bartlett’s test of sphericity and KMO.

Test		Normotensive subjects	Hypertensive subjects
Bartlett’s test of sphericity	Approx. χ^2	231.012	119.48
	DF	10	10
	P	<0.0001	<0.0001
KMO measure of sampling adequacy		0.63	0.75

DF: degree of freedom.

TABLE 5: Eigenvalues and % of variation explained by each PC in normotensive and hypertensive subjects.

PCs	Normotensive subjects			Hypertensive subjects		
	Eigenvalue	Individual%	Cumulative%	Eigenvalue	Individual%	Cumulative%
1	3.59	71.84	71.84	3.0550	61.10	61.10
2	0.83	16.58	88.42	1.1249	22.5	83.60
3	0.32	6.34	94.76	0.4393	8.78	92.38
4	0.25	5.04	99.8	0.2830	5.66	98.04
5	0.01	0.2	100.00	0.0978	1.96	100

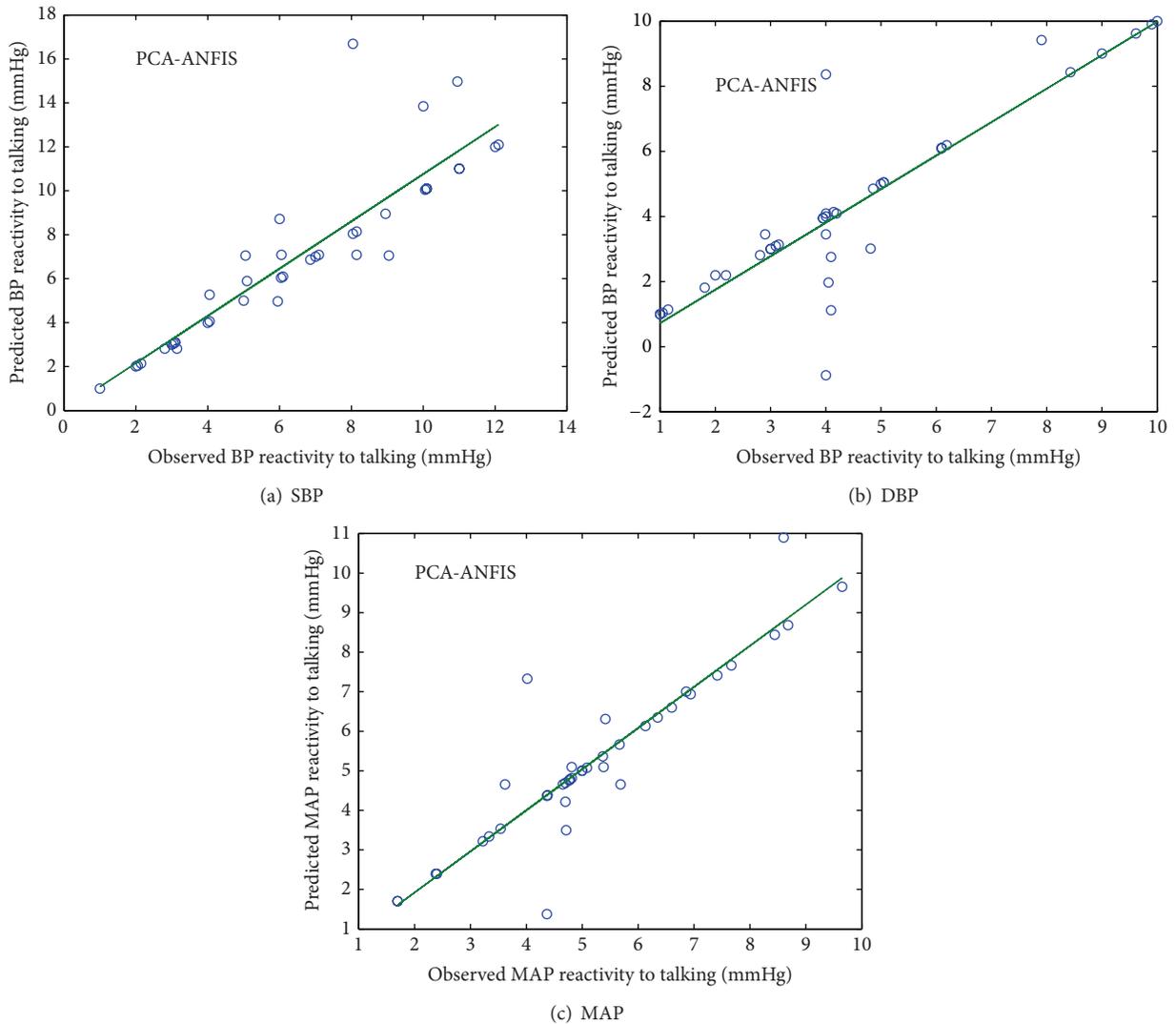


FIGURE 6: Scatter plot between observed and predicted values of SBP, DBP, and MAP reactivity of normotensive subjects using PCA-ANFIS model.

TABLE 6: Loadings of anthropometric characteristics in normotensive and hypertensive subjects.

Anthropometric characteristics	Loadings after varimax rotation							
	Normotensive subjects				Hypertensive subjects			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
Age	0.0004	-0.0006	-0.0000	-1.0000	-0.0036	0.0020	0.9988	-0.0026
Height	-0.0139	-0.9676	0.0002	-0.0008	0.0058	0.9968	0.0020	0.0043
Weight	-0.6569	-0.1812	-0.0008	0.0039	0.6576	0.0561	-0.0349	-0.0754
BMI	-0.7538	0.1757	0.0008	-0.0039	0.7533	-0.0566	0.0352	0.0760
AC	-0.0001	0.0001	-1.0000	-0.0000	0.0078	-0.0043	0.0027	-0.9942

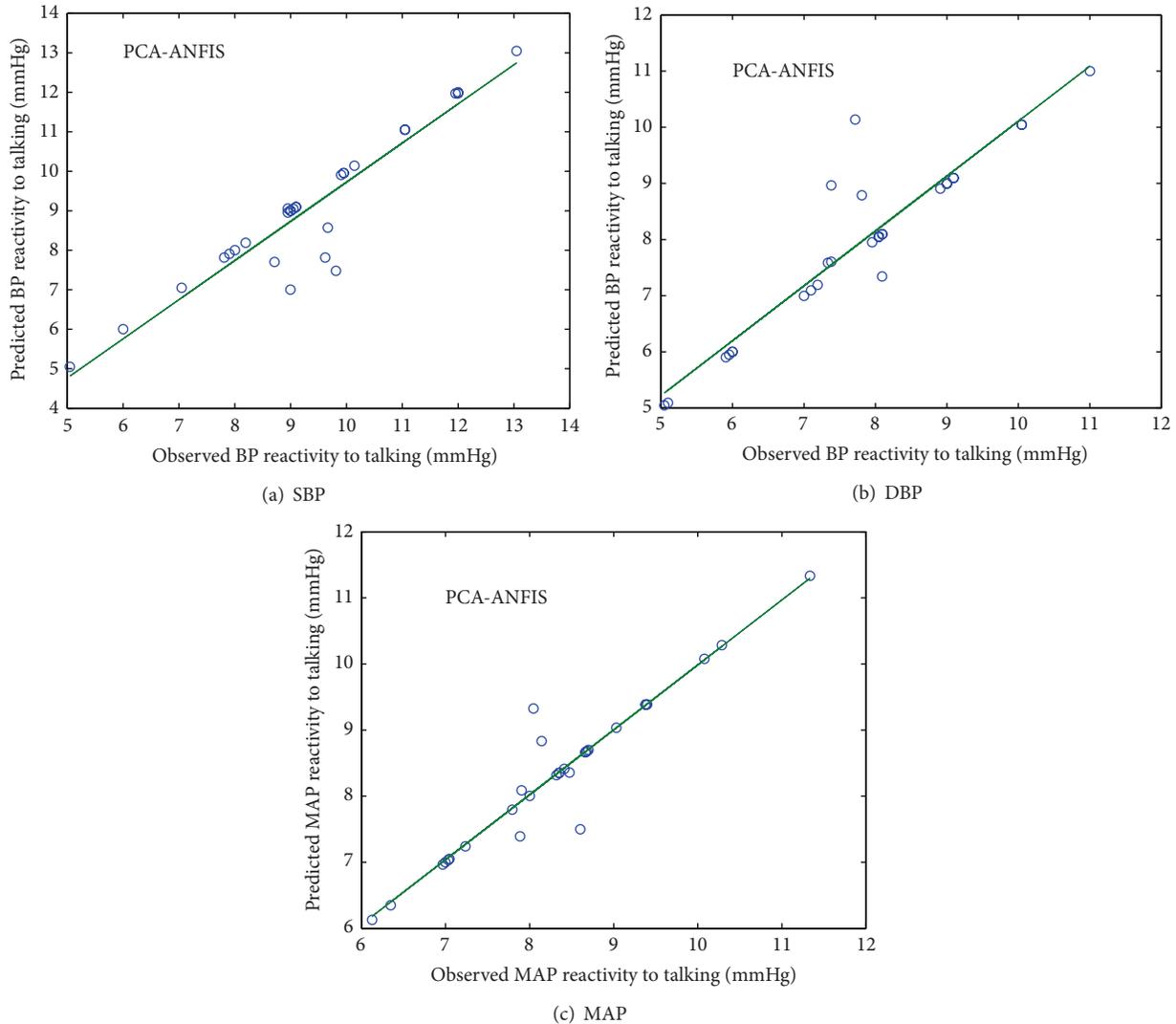


FIGURE 7: Scatter plot between observed and predicted values of SBP, DBP, and MAP reactivity of hypertensive subjects using PCA-ANFIS model.

BP elevations. And withdrawal from such verbal activity has important clinical implications for the cardiovascular system.

Furthermore, we illustrated an application of PCA based soft computing models in predicting the BP reactivity to talking. PCA corrects for confounding caused by anthropometric characteristics including age, height, weight, BMI, and AC

and, therefore, normotensive subjects were used to provide a basis for comparison.

As far as we know, this paper is the first study related to prediction of BP reactivity to talking using PCA based soft computing approaches. Therefore, the results were compared with indirectly related studies [33–37], as shown in Table 10.

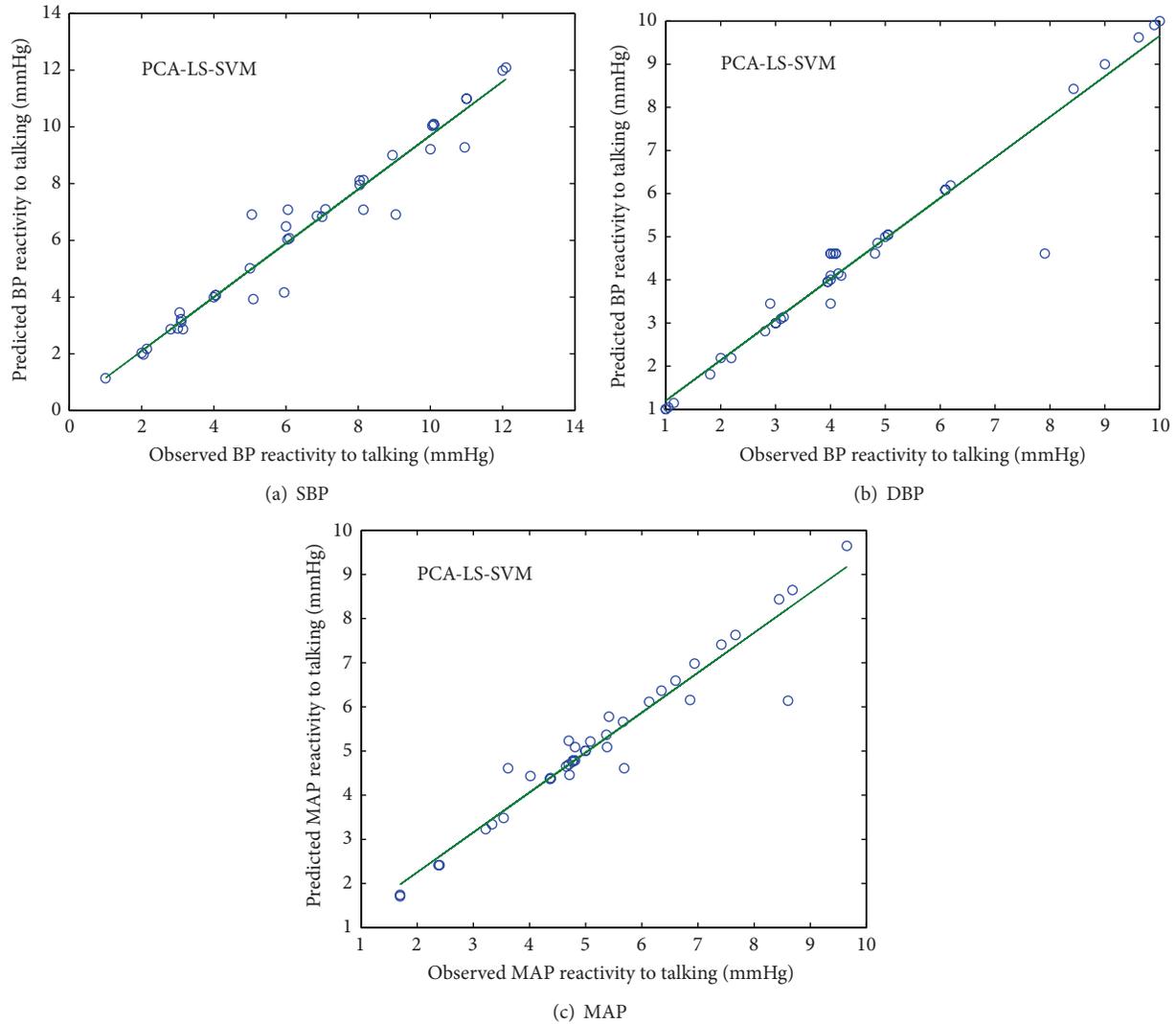


FIGURE 8: Scatter plot between observed and predicted values of SBP, DBP, and MAP reactivity of normotensive subjects using PCA-LS-SVM model.

TABLE 7: Pearson's correlation coefficient among all pairs of PCs in normotensive and hypertensive subjects.

PC	PC2	PC3	PC4
PC1	-0.00000225 0.00000878	0.0000000798 0.00000423	-0.0000167 0.00000659
PC2		-7.237e - 016 0.00000919	5.808e - 016 0.0000142
PC3			-7.557e - 017 0.0000175

Bold values indicate correlations in anthropometric characteristics of hypertensive subjects.

Promising results of soft computing techniques in all studies are due to their high degree of robustness and fault tolerance. In this work, specifically, the best performance of LS-SVM is sourced from its several advantages including global optimal solution ability, fast convergence rate, and good generalization with small size sample.

This study has a number of advantages. We used small, medium, and large size cuffs, which may have produced more accurate readings. And we took the mean of multiple readings to strengthen the accuracy of BP measurements.

However, any single comparison between the models might not reliably represent the true results. Validation of the computing models using larger database is essential to get an accurate measure of performance outside the development population.

6. Conclusion

The successful development of any prediction model depends largely on the quality and nature of data used for model development. To address the issue of multicollinearity within the anthropometric variables, PCA is incorporated. Furthermore, performance comparison of PCA-ANN, PCA-ANFIS, and PCA-LS-SVM models revealed the potential capability of PCA-LS-SVM model in predicting BP reactivity. This

TABLE 8: Optimal values of γ and σ^2 .

Parameters	Normotensive subjects		Hypertensive subjects	
	SBP	DBP	SBP	DBP
γ	369.9717	1.7430e + 004	1.7498e + 008	3.0844e + 006
σ^2	0.5882	7.0498e - 006	4.7587e - 004	2.3885e - 006

TABLE 9: Statistical indices for different models.

Model	Normotensive subjects						Hypertensive subjects					
	SBP		DBP		MAPE		SBP		DBP		MAPE	
	R^2 (%)	RMSE	MAPE	R^2 (%)	RMSE	MAPE	R^2 (%)	RMSE	MAPE	R^2 (%)	RMSE	MAPE
PCA-ANN	67.44	0.58	37.09	62.39	0.62	40.77	59.50	0.68	9.2	53.19	0.86	11.02
PCA-ANFIS	80.04	0.56	9.82	79.37	0.52	12.27	87.02	0.41	2.97	86.11	0.4	2.71
PCA-LS-SVM	95.42	0.21	5.88	94.22	0.24	4.05	98.76	0.11	0.88	98.78	0.11	0.84

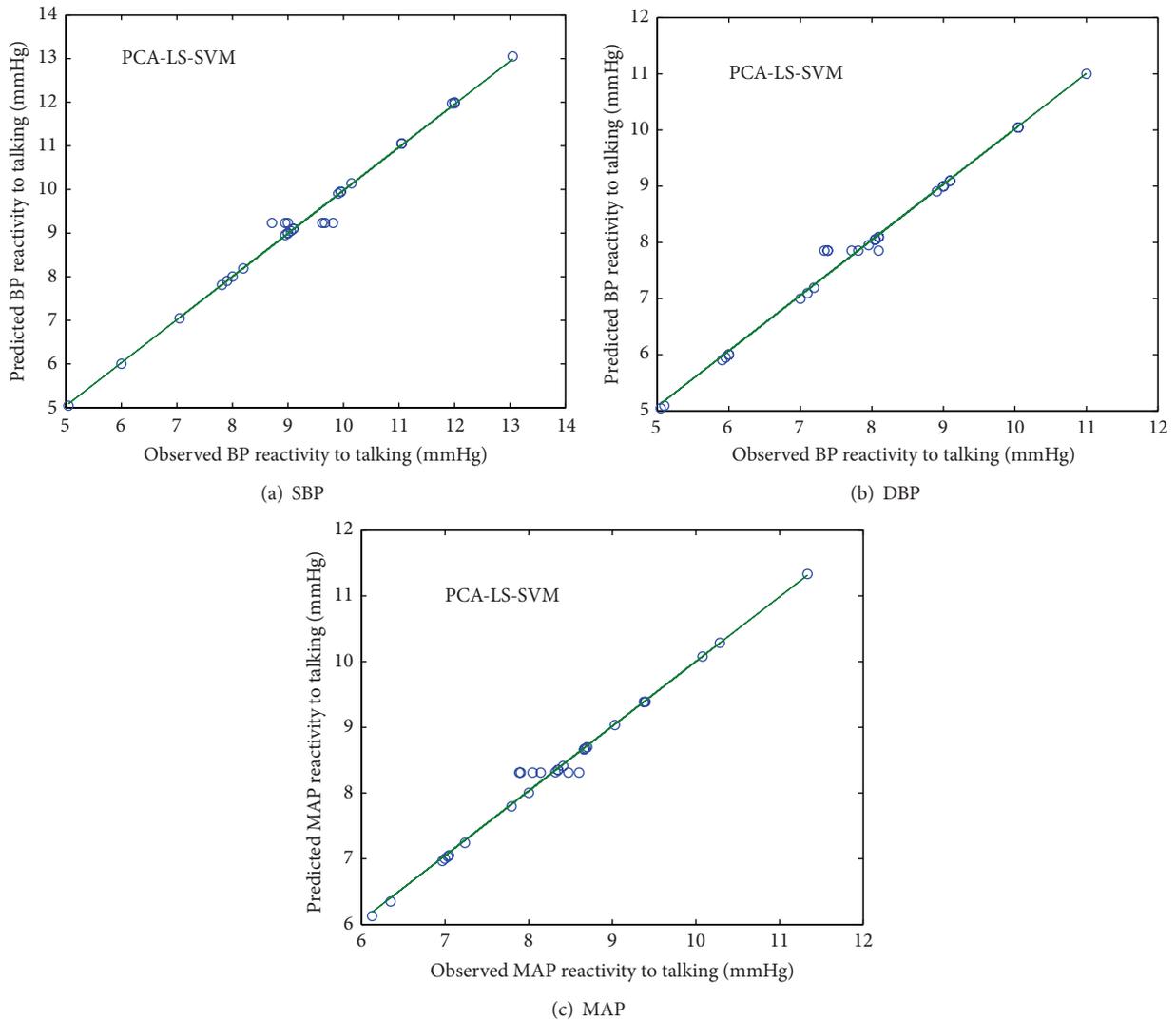


FIGURE 9: Scatter plot between observed and predicted values of SBP, DBP, and MAP reactivity of hypertensive subjects using PCA-LS-SVM model.

TABLE 10: Comparison of results with other studies.

Ref.	Model developed	Predicted parameter	Remarks
[33]	FIR and ANFIS	Mean BP and AEP during anaesthesia	No significant difference between the results of two models
[34]	ANN and multiple linear regression (MLR)	SBP	ANN outperformed MLR
[35]	RS-SVM	BP	Training rapidity and accuracy of the RS-SVM model are both evidently improved
[36]	PCA-ANFIS, conventional maximum amplitude algorithm	SBP and DBP	PCA-ANFIS outperformed
[37]	FIS	Effect of aerobic exercise on BP	Preliminary validation results of the performance of the FIS are promising
Our study	PCA-LS-SVM, PCA-ANN, and PCA-ANFIS	SBP, DBP, and MAP	PCA-LS-SVM outperformed PCA-ANN and PCA-ANFIS

work may provide a valuable reference for researchers and engineers who apply soft computing models for modeling biological variables. The results are helpful in physician's diagnosis for the prevention of hypertension in clinical medicine. Our future research is targeted to study an ensemble approach by combining the outputs of different hybrid techniques with more predictor variables and larger data sets to achieve wide clinical application of the soft computing.

Conflict of Interests

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

References

- [1] T. G. Pickering, J. E. Hall, L. J. Appel et al., "Recommendations for blood pressure measurement in humans and experimental animals: part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on high blood pressure research," *Circulation*, vol. 111, no. 5, pp. 697–716, 2005.
- [2] B. Williams, N. R. Poulter, M. J. Brown et al., "Guidelines for management of hypertension: report of the fourth working party of the British hypertension society, 2004—BHS IV," *Journal of Human Hypertension*, vol. 18, no. 3, pp. 139–185, 2004.
- [3] E. O'Brien, R. Asmar, L. Beilin et al., "European Society of Hypertension recommendations for conventional, ambulatory and home blood pressure measurement," *Journal of Hypertension*, vol. 21, no. 5, pp. 821–848, 2003.
- [4] J. J. Lynch, J. M. Long, S. A. Thomas, K. L. Malinow, and A. H. Katcher, "The effects of talking on the blood pressure of hypertensive and normotensive individuals," *Psychosomatic Medicine*, vol. 43, no. 1, pp. 25–33, 1981.
- [5] D. G. Vidt, R. S. Lang, R. J. Seballos, A. Misra-Hebert, J. Campbell, and J. F. Bena, "Taking blood pressure: too important to trust to humans?" *Cleveland Clinic Journal of Medicine*, vol. 77, no. 10, pp. 683–688, 2010.
- [6] D. Zheng, R. Giovannini, and A. Murray, "Effect of respiration, talking and small body movements on blood pressure measurement," *Journal of Human Hypertension*, vol. 26, no. 7, pp. 458–462, 2012.
- [7] C. Le Pailleur, G. Helft, P. Landais et al., "The effects of talking, reading, and silence on the "white coat" phenomenon in hypertensive patients," *The American Journal of Hypertension*, vol. 11, no. 2, pp. 203–207, 1998.
- [8] C. Le Pailleur, P. Montgermont, J. M. Feder, J. P. Metzger, and A. Vacheron, "Talking effect and "white coat" effect in hypertensive patients: physical effort or emotional content?" *Behavioral Medicine*, vol. 26, no. 4, pp. 149–157, 2001.
- [9] D. Zheng, R. Giovannini, and A. Murray, "Effect of talking on mean arterial blood pressure: agreement between manual auscultatory and automatic oscillometric techniques," *Computing in Cardiology*, vol. 38, pp. 841–844, 2011.
- [10] C. H. Tardy, W. R. Thompson, and M. T. Alen, "Cardiovascular responses during speech: does social support mediate the effects of talking on blood pressure?" *Journal of Language and Social Psychology*, vol. 8, pp. 271–285, 1989.
- [11] J. J. Lynch, K. E. Lynch, and E. Friedmann, "A cry unheard: sudden reductions in blood pressure while talking about feelings of hopelessness and helplessness," *Integrative Physiological and Behavioral Science*, vol. 27, no. 2, pp. 151–169, 1992.
- [12] J. M. Long, J. J. Lynch, N. M. Machiran, S. A. Thomas, and K. L. Malinow, "The effect of status on blood pressure during verbal communication," *Journal of Behavioral Medicine*, vol. 5, no. 2, pp. 165–172, 1982.
- [13] R. Hellmann and S. A. Grimm, "The influence of talking on diastolic blood pressure readings.," *Research in nursing & health*, vol. 7, no. 4, pp. 253–256, 1984.
- [14] F. D. Fuchs, M. Gus, L. B. Moreira et al., "Anthropometric indices and the incidence of hypertension: a comparative analysis," *Obesity Research*, vol. 13, no. 9, pp. 1515–1517, 2005.
- [15] M. Inoue, M. Minami, and E. Yano, "Body mass index, blood pressure, and glucose and lipid metabolism among permanent and fixed-term workers in the manufacturing industry: a cross-sectional study," *BMC Public Health*, vol. 14, article 207, no. 1, 2014.
- [16] D. C. Moser, I. De Carlos Back Giuliano, A. C. K. Titski, A. R. Gaya, M. J. Coelho-e-Silva, and N. Leite, "Anthropometric

- measures and blood pressure in school children,” *Jornal de Pediatria*, vol. 89, no. 3, pp. 243–249, 2013.
- [17] S. Doll, F. Paccaud, P. Bovet, M. Burnier, and V. Wietlisbach, “Body mass index, abdominal adiposity and blood pressure: consistency of their association across developing and developed countries,” *International Journal of Obesity*, vol. 26, no. 1, pp. 48–57, 2002.
- [18] K. Bose, A. Ghosh, S. Roy, and S. Gangopadhyay, “Blood pressure and waist circumference: An empirical study of the effects of waist circumference on blood pressure among Bengalee male jute mill workers of Belur, West Bengal, India,” *Journal of Physiological Anthropology and Applied Human Science*, vol. 22, no. 4, pp. 169–173, 2003.
- [19] R. Baba, M. Koketsu, M. Nagashima, H. Inasaka, M. Yoshinaga, and M. Yokota, “Adolescent obesity adversely affects blood pressure and resting heart rate,” *Circulation Journal*, vol. 71, no. 5, pp. 722–726, 2007.
- [20] T. Næs and B.-H. Mevik, “Understanding the collinearity problem in regression and discriminant analysis,” *Journal of Chemometrics*, vol. 15, no. 4, pp. 413–426, 2001.
- [21] I. T. Jolliffe, *Principal Component Analysis*, Springer, New York, NY, USA, 2002.
- [22] The British Hypertension Society, *Blood Pressure Measurement (CD-ROM)*, BMJ Books, London, UK, 1998.
- [23] S. Tobias and J. E. Carlsons, “Brief report: Bartlett’s test for sphericity and chance findings in factor analysis,” *Multivariate Behavioral Research*, vol. 4, pp. 375–377, 1969.
- [24] B. Williams, A. Onsmann, and T. Brown, “Exploratory factor analysis: a five-step guide for novices,” *Journal of Emergency Primary Health Care*, vol. 8, no. 3, pp. 1–13, 2010.
- [25] H. Camdevyren, N. Demyr, A. Kanik, and S. Keskin, “Use of principal component scores in multiple linear regression models for prediction of Chlorophyll-a in reservoirs,” *Ecological Modelling*, vol. 181, pp. 581–589, 2005.
- [26] J.-N. Hwang, J. J. Choi, S. Oh, and R. J. Marks II, “Query-based learning applied to partially trained multilayer perceptrons,” *IEEE Transactions on Neural Networks*, vol. 2, no. 1, pp. 131–136, 1991.
- [27] J.-S. R. Jang, “ANFIS: adaptive-network-based fuzzy inference system,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.
- [28] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least Squares Support Vector Machines*, World Scientific Publishing, London, UK, 2002.
- [29] H. Esen, F. Ozgen, M. Esen, and A. Sengur, “Modelling of a new solar air heater through least-squares support vector machines,” *Expert Systems with Applications*, vol. 36, no. 7, pp. 10673–10682, 2009.
- [30] D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith, “Model evaluation guidelines for systematic quantification of accuracy in watershed simulations,” *Transactions of the ASABE*, vol. 50, no. 3, pp. 885–900, 2007.
- [31] J. Benesty, J. Chen, and Y. Huang, “On the importance of the pearson correlation coefficient in noise reduction,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 16, no. 4, pp. 757–765, 2008.
- [32] H. F. Kaiser, “The application of electronic computers to factor analysis,” *Educational and Psychological Measurement*, vol. 20, pp. 141–151, 1960.
- [33] E. W. Jensen and A. Nebot, “Comparison of FIR and ANFIS methodologies for prediction of mean blood pressure and auditory evoked potentials index during anaesthesia,” in *Proceedings of the IEEE 20th Annual International Conference on Engineering in Medicine and Biology Society*, vol. 3, pp. 1385–1388, IEEE, Hong Kong, 1998.
- [34] J. Y. Kim, B. H. Cho, S. M. Lim, M. J. Jeon, I. Y. Kim, and S. Kim, “Comparative study on artificial neural network with multiple regressions for continuous estimation of blood pressure,” in *Proceedings of the 27th Annual International Conference on Engineering in Medicine and Biology Society of the IEEE*, Shanghai, China, 2006.
- [35] G. Zhang, “An improved hypertension prediction model based on RS and SVM in the three gorges area,” in *Proceedings of the International Conference on Computer Science and Software Engineering (CSSE ’08)*, pp. 810–814, Hubei, China, December 2008.
- [36] M. Forouzanfar, H. R. Dajani, V. Z. Groza, M. Bolic, and S. Rajan, “Adaptive neuro-fuzzy inference system for oscillometric blood pressure estimation,” in *Proceedings of the IEEE International Workshop on Medical Measurements and Applications*, pp. 125–129, Ontario, Canada, May 2010.
- [37] A. M. Honka, M. J. van Gils, and J. Parkka, “A personalized approach for predicting the effect of aerobic exercise on blood pressure using a fuzzy inference system,” in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC ’11)*, pp. 8299–8302, Boston, Mass, USA, August 2011.
- [38] D. Pyle, *Data Preparation for Data Mining*, Morgan Kaufmann, San Francisco, Calif, USA, 1999.
- [39] M. T. Hagan and M. B. Menhaj, “Training feedforward networks with the Marquardt algorithm,” *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989–993, 1994.
- [40] B. Scholkopf and A. J. Smola, *Learning with Kernels*, MIT Press, Cambridge, Mass, USA, 2002.
- [41] H. Weiner, M. T. Singert, and M. F. Reiser, “Cardiovascular responses and their psychological correlates. I. A. study in healthy young adult and patients with peptic ulcer and hypertension,” *Psychosomatic Medicine*, vol. 24, pp. 477–498, 1962.

Research Article

Detection and Measurement of the Intracellular Calcium Variation in Follicular Cells

Ana M. Herrera-Navarro,¹ Iván R. Terol-Villalobos,² Hugo Jiménez-Hernández,³
Hayde Peregrina-Barreto,⁴ and José-Joel Gonzalez-Barboza⁵

¹ Facultad de Informática, Universidad Autónoma de Querétaro, Campus Juriquilla, Avenida de las Ciencias s/n, 76230 Querétaro, QRO, Mexico

² Centro de Investigación y Desarrollo Tecnológico en Electroquímica S.C., Pedro Escobedo, 76703 Querétaro, QRO, Mexico

³ Centro de Investigación e Ingeniería Industrial. Avenida Playa Pie de la Cuesta No. 702, Desarrollo San Pablo, 76703 Querétaro, QRO, Mexico

⁴ Departamento de Ciencias Computacionales, Instituto Nacional de Astrofísica, Óptica y Electrónica, Luis Enrique Erro No. 1, CP 72840, Tonantzintla, PUE, Mexico

⁵ Centro de Investigación y Tecnología Aplicada, Cerro Blanco No. 141, Colonia Colinas del Cimatario, 76090 Querétaro, QRO, Mexico

Correspondence should be addressed to Ana M. Herrera-Navarro; anaherreranavarro@gmail.com

Received 13 June 2014; Revised 21 August 2014; Accepted 22 August 2014; Published 16 September 2014

Academic Editor: Ezequiel López-Rubio

Copyright © 2014 Ana M. Herrera-Navarro 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 work presents a new method for measuring the variation of intracellular calcium in follicular cells. The proposal consists in two stages: (i) the detection of the cell's nuclei and (ii) the analysis of the fluorescence variations. The first stage is performed via watershed modified transformation, where the process of labeling is controlled. The detection process uses the contours of the cells as descriptors, where they are enhanced with a morphological filter that homogenizes the luminance variation of the image. In the second stage, the fluorescence variations are modeled as an exponential decreasing function, where the fluorescence variations are highly correlated with the changes of intracellular free Ca^{2+} . Additionally, it is introduced a new morphological called medium reconstruction process, which helps to enhance the data for the modeling process. This filter exploits the undermodeling and overmodeling properties of reconstruction operators, such that it preserves the structure of the original signal. Finally, an experimental process shows evidence of the capabilities of the proposal.

1. Introduction

Calcium (Ca^{2+}) is an ubiquitous intracellular ion signaling responsible for controlling many cellular processes [1, 2]. Ca^{2+} acts as second messenger triggering pathological events, such as cells injury and death, as well as participating in pathological conditions, such as hypertension, cardiac arrhythmia, hematological problems, muscular diseases, and hormonal disorders, among others [1, 3, 4]. The role of Ca^{2+} in these diseases has just begun to be understood. Some Ca^{2+} that induce pathological conditions have been found with the help of substances that interfere with the movement or activation of Ca^{2+} [3, 5]. Due to the importance of Ca^{2+} , there are several numbers of optical and nonoptical techniques,

which have been developed for analyzing Ca^{2+} either dynamics or concentration. Fluorescent microscopy techniques are frequently used to observe the variation of intercellular Ca^{2+} concentration applying chemical fluorescent indicator as markers. Those indicators stimulate the cells causing a fluorescence effect [6]. The fluorescence is detected by microscopy and CCD sensors. To compute the cells with the greatest fluorescence variations, the user must select them manually and analyze all the images in the sequence in order to determine the changes of fluorescence over time. However, it consumes time and human resources being susceptible to error measuring. By such circumstance, the process for segmenting the cells and the study of dynamics of intracellular Ca^{2+} represent the main objective in this paper.

An important step in the study of intracellular Ca^{2+} consists of the segmentation of each individual cell. The image segmentation results are difficult because the environmental changing conditions are usually uncontrollable. In the literature typically depending on the properties of the cells, the segmentation method is proposed. For instance, some works use neural networks approaches [7], hierarchical threshold [8], or multiscale morphology [9], just to mention a few.

Watershed-plus-marker approach, on the other hand, is the traditional image segmentation method based on morphological mathematical [10, 11]. The success of this method depends mainly on the correct detection of the image's markers. The markers can be detected manually or automatically. Automatic approaches help the specialist to save time and resources. However, there are factors that affect the performance of an automatic detection of markers such as noise, cells occlusions, and abrupt changes in the images. Those factors can lead these algorithms to over- and undersegmentation, that is, create regions containing partial or multiple cells. In this sense, several approaches have been developed for improving cell segmentation. In this study, we introduce a method to analyze automatically the intracellular calcium variation. This approach consists in two stages: (1) the image enhancement and cell segmentation and (2) the calcium variation modeling. The image enhancement is carried out with a top-hat filter, which homogenizes the luminance conditions. The process of cells segmentation is performed using the marked-controlled watershed transform and filters by reconstruction, which is used to detect markers efficiently, after the homotopy of the gradient image was computed. To measure the calcium intracellular variations, the volume of each marked cell is computed. The procedure consists of estimating the volume using the luminance intensities for each marked cell along the time. After, least squares fitting (LSF) method is applied to create a model of the behavior of the variation of the fluorescence. The behavior model created is considered as an exponential decreasing function. To enhance the development of the model of the behavior of calcium, a novel morphological filter named as medium filter is introduced. This filter smoothes the fluorescence measures, exploiting the under- and overmodeling of reconstruction operator, preserving the information structure of original signal.

The paper is organized as follows. In the next section, a review of some morphological filters is presented. In Section 3, a method based on marked controlled watershed transform to detect automatically the cells is shown. In Section 4, we show the procedure to estimate the volume of each marked cell starting from the fluorescence intensity along time; after applied least squares fitting and morphological medium filter, the model of fluorescence behavior is created. Finally, results and conclusions can be found in the last section.

2. Concepts of Morphological Filtering

2.1. Basic Notions of Morphological Filtering. Mathematical morphology is mainly based on the so-called increasing transformations [12–14]. A transformation T is increasing if

for all pairs of functions f and g , with $f \leq g \Rightarrow T(f) \leq T(g)$. In other words, increasing transformations preserve the order of the relation. A second property is the idempotence; that is, a transformation T is idempotent if and only if $T(T(f)) = T(f)$. The basic morphological filters are the morphological opening $\gamma_{\mu B}$ and the morphological closing $\varphi_{\mu B}$ with a given structuring element, where B represents the elementary structuring element (3×3 pixels, e.g.) containing its origin and μ is an homothetic parameter. Thus, the morphological opening and closing are given, respectively, by

$$\gamma_{\mu B}f(x) = \delta_{\mu B}(\varepsilon_{\mu B}(f)), \quad \varphi_{\mu B}f(x) = \varepsilon_{\mu B}(\delta_{\mu B}(f)), \quad (1)$$

where the morphological erosion $\varepsilon_{\mu B}(f(x))$ and dilation $\delta_{\mu B}(f(x))$ are expressed as $\varepsilon_{\mu B} = f \ominus \mu B : x \mapsto \inf_{y \in \mu B} f(x+y)$ and $\delta_{\mu B}(f) = f \oplus \mu B : x \mapsto \sup_{y \in \mu B} f(x-y)$.

Henceforth, the set B will be suppressed rendering, the expressions γ_{μ} and $\gamma_{\mu B}$ are equivalent ($\gamma_{\mu} = \gamma_{\mu B}$). When the parameter μ is equal to one, all parameters are suppressed ($\delta_B = \delta$).

2.2. Opening (Closing) by Reconstruction. The notion of reconstruction is a very useful concept provided by MM. Reconstruction transformations are built by means of the geodesic transformations. In a gray-level case, the geodesic dilation $\delta_f^1(g)$ (resp., the geodesic erosion $\varepsilon_f^1(g)$) with $g \leq f$ (resp., $g \geq f$) of size 1 given by $\delta_f^1(g) = f \wedge \delta_B(g)$ (resp., $\varepsilon_f^1(g) = f \vee \varepsilon_B(g)$) is iterated until idempotence. Consider two functions f and g , with $f \geq g$ ($f \leq g$). Reconstruction transformations of the marker function g in f , using geodesic dilations and erosions, expressed by $R(f, g)$ and $R^*(f, g)$, respectively, are defined by

$$R(f, g) = \lim_{n \rightarrow \infty} \varepsilon_f^n(g) = \varepsilon_f^1 \varepsilon_f^1 \cdots \varepsilon_f^1(g), \quad (2)$$

$$R^*(f, g) = \lim_{n \rightarrow \infty} \delta_f^n(g) = \delta_f^1 \delta_f^1 \cdots \delta_f^1(g).$$

When the marker function g is equal to the erosion or the dilation of the original function in (2), the opening and the closing by reconstruction are obtained:

$$\tilde{\gamma}_{\mu}(f) = \lim_{n \rightarrow \infty} \delta_f^n(\varepsilon_{\mu}(f)), \quad (3)$$

$$\tilde{\varphi}_{\mu}(f) = \lim_{n \rightarrow \infty} \varepsilon_f^n(\delta_{\mu}(f)).$$

3. Automatic Detection of Cells

The watershed-plus-marker approach transformation is a traditional image segmentation method based in mathematical morphology [10, 13]. However, this transformation makes use of an extensive set of morphological filters. This transform is used for segmenting images avoiding the oversegmentation [11]. The oversegmentation criterion consists of setting an upper limit in the number of minima regions detected. This process is performed with the minima impositions over the markers, exploiting homotopy property of the operators.

However, it needs to make some assumptions to use this approach. The most important assumption consists of the fact that the minima represent the center of the object (in the cell case, the core of it); a second assumption consists of the gradient estimation, such that the fact that it could be performed with morphological operators.

3.1. Marker Detection. Due to the features of the images, the nucleus of each cell is used as regional minimum. As matter of fact, a regional minimum M of a gray-scale image I is a connected component of pixels with uniform altitude without lower neighbors. Before computing the minima, the nucleus of the cells is mainly dark, surrounded by brighter region composed by the cytoplasm. However, the local luminance conditions of each cell differ singly to each other, affecting the detection of the nucleus. For the homogenization of luminance conditions the top-hat transform is used as a local contrast correction filter.

For the $\{I_i\}_{i \in S}$ be sequence of image. The top-hat transformation is defined as follows:

$$\text{Th}\omega_{\lambda B}(I) = (I_i)(x) - \gamma_{\lambda}(I_i)(x), \quad (4)$$

where the dimensions of structuring element are related with the luminance conditions of the scenario; in such a way, the luminance distribution in the image may be approximated with the morphological opening γ_{μ} . Whenever the dimension of structuring element becomes proportionally similar to the image dimension, it would represent global luminance sources affectations; on the other hand, small dimensions represent local luminance variations effects. This process is illustrated in Figure 1, where each process step is showed and in certain steps the image has been coded in pseudocolor to point out the effects involved at each step.

The direct appliance of the maxima transform detects all the maxima, including noise data, as it is appreciated in Figure 2(b). To avoid the extra minima detection, the image is enhanced with a closing by reconstruction operator. This operator allow grouping by all connected local minima, discarding the majority of noise effects. The closing by reconstruction uses a unitary structuring to approximate the original image. For illustration proposes in Figure 2(c) is appreciated the effect of apply the filter in the minima detection process. Observe that some cell nucleus are well detected, but others are omitted; due to the acquisition process the cytoplasm is not completely closed. This situation can be fixed using subsequent images, where additional minima are correctly detected. Thus, in order to obtain the maximum number of minima, a function that captures the occurrence of the minima in the sequence is constructed. Let $\{I_i\}_{i \in S}$ and $\{M_i\}_{i \in S}$ be the images of the sequence and the images containing the detected minima, respectively. $M_i(x)$ is a binary image such that it takes 1 value if the point x belongs to a regional minimum and 0 values otherwise. After using the subsequent frames, summation I_m is built as follows:

$$I_m(x) = \sum_{i \in S} M_i(x). \quad (5)$$

The surface I_m is drawn in Figure 2(d). Analyzing Figures 2(c) and 2(d), the majority of true minima are detected. Other

big areas are pointed out in the figure. They are discarded by a thresholding criterion. In the case of study, each cell is typically about four pixel of radius, which can be discovered using an opening by reconstruction filter. As complementary, a closing operator with a structure of 3 pixels of dimension is used to connect insolated regions. The process is illustrated in Figure 2. First, a morphological closing of size 3 is applied to fill the small holes; the results can be appreciated in Figure 2(c) (before closing filter) and Figure 2(e) (after closing filter). Next, applying the summation I_m (Figure 2(d)), the minima are found, and finally the regions with big and small areas are discarded, denoting the cells.

3.2. Gradient Operator. The watershed-plus-marker approach makes use of the gradient operator to impose the markers. In this sense, the morphological gradient can be used as contrast detector. Let $I(x)$ be a function defined in \mathbb{Z}^2 and B the basic structuring element of 3×3 dimensions, centered at point x . Then, the transformation is defined as follows for a discrete space:

$$\nabla_B I(x) = \delta_B I(x) - \varepsilon_B I(x). \quad (6)$$

There are other two versions of gradients in mathematical morphology, the internal and the external gradients defined, respectively, as follows:

$$\begin{aligned} \nabla_B I(x) &= I(x) - \varepsilon_B I(x), \\ \nabla_B I(x) &= \delta_B I(x) - I(x). \end{aligned} \quad (7)$$

Figures 3(b) and 3(c) show the internal and external gradients of the image in Figure 3(a). However, the indistinct use of any gradient approach has the consequence where the border should present double border. Typically they correspond to the cell generated among the nucleus of the cell and cytoplasm and the other one between the cytoplasm and the background of the image. The drawback to detect efficiently the true border in the image is an open task; then to deal with it, several tests have been applied to the images. Experimentally it is appreciated that external gradient offers smoother and thick borders (see Figure 3(c)), instead of internal gradient offering defined and clear borders as it is illustrated in Figure 3(c).

3.3. Imposed of Minima by Reconstruction. Once the markers cells signals are detected, these are imposed with their minima on the gradient image. To carry out this task the following procedure is performed. Let M and g be the set of markers computed as commented above and the gradient image, respectively. After, two news functions are built: the first one consists of a thresholding function $f(x)$, which is defined as $f(x) = 255$ if $x \notin M$ and $f(x) = 0$ if $x \in M$, while the second one is built through the gradient image as $g'(x) = g(x)$ if $x \notin M$ and $g(x) = 0$ if $x \in M$. Furthermore, the dual morphological reconstruction of $f(x)$ inside of $g'(x)$ is made, denoted by $R^*(g', f)$. The function $R^*(g', f)$ only has the minima of M , such that the watershed transformation is applied. Figure 3(d) illustrates the results getting after applying watershed segmentation.

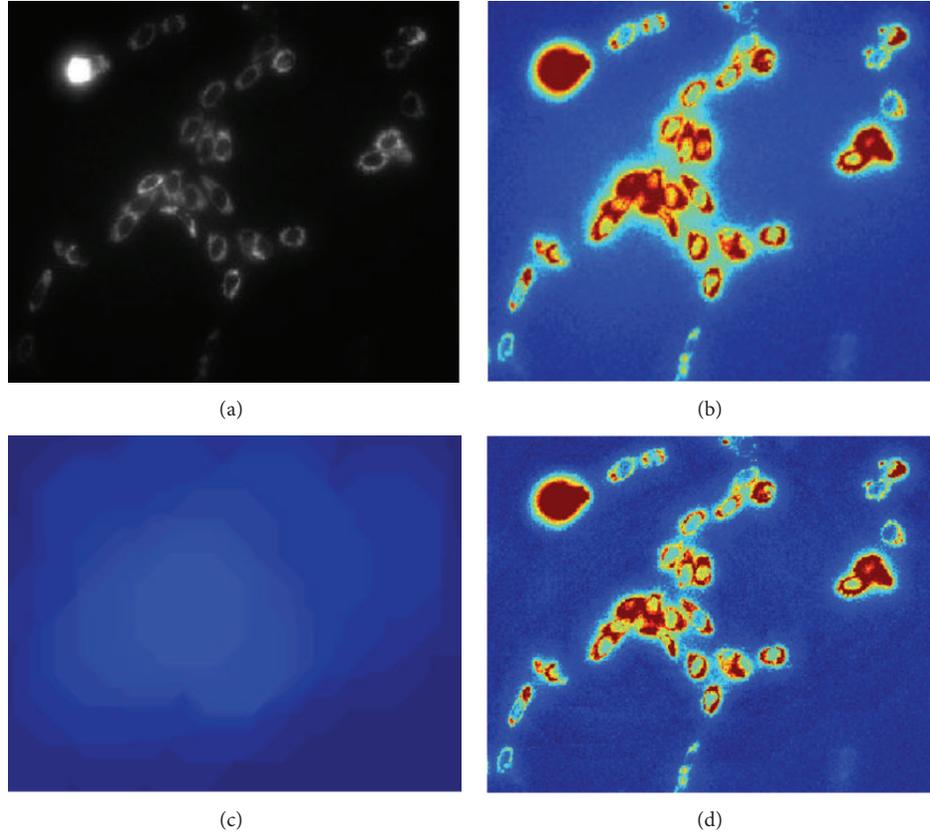


FIGURE 1: (a) Original image, (b) original image in pseudocolor before background correction, (c) opening morphological, and (d) image after background correction.

4. Modeling the Intracellular Calcium Dynamic

In this section, we deal with the problem of modeling the intracellular calcium dynamic. The procedure consists in three parts: the estimation of the calcium volume, the fitting of an exponential curve, and the calculus of the error.

4.1. Estimating Cell's Volume. The cell intensities are highly related to the amount of calcium contained in each cell. Then, the task of creating a model of the behavior of calcium in each cell is managed computing the volume of each cell in the image. The historical measures of time are used to represent the evolution of the dynamic of the variation of calcium for each particular cell. The historical measures of volumes are denoted by $\{V_n(i)\}_{i \in S}$, where the subindex n corresponds to a particular cell and i represent the particular volume for the time stamp i th. The volume is estimated with a discrete approximation of the integral as follows:

$$V = \int_{x_i}^{x_f} \int_{y_i}^{y_f} f(x, y) dy dx, \quad (8)$$

$$V \approx \sum_{x_i}^{x_f} \sum_{y_i}^{y_f} f(x, y) \Delta y \Delta x, \quad \text{for } \Delta x = \Delta y = 1.$$

4.2. Modeling Intracellular Calcium Variations. As it is appreciated in Figure 4, the dynamics of the calcium stimulus has an exponential behavior. Then, the purpose consists of creating a model of the decreasing behavior stimuli suffered by each cell, where the region of interest is located among the global maxima and the end of the signal. However, due to the noise, it is not possible to detect easily the maximum. To attenuate this inconvenience, an automatic process is performed detecting the maxima for the function $\{V_n(i)\}_{i \in S}$. The process consists of a sequential alternating filter in one-dimensional scenario. The alternating filter is constituted by a sequence of one closing by reconstruction followed by one opening by reconstruction $\hat{\varphi}_{\mu L}(\tilde{\gamma}_{\mu L}(V))(i)$ where the size of μ is varied into the interval $[0, k]$. The filter undermodels the original signal, smoothing the signal wave and allowing the detection of the global maxima efficiently.

Figure 5 illustrates the detection of a representative maximum detected that corresponds to a connected element in one dimension space. The center of the connected element represents the maxima location, such that it is estimated with the mean of the connected elements; that is, $c(\{x_i \mid x_i \in R(x_i, x_j)\}) = (1/n) \sum_{i=1}^n x_i$, such that $R(x_i, x_j)$ is an equivalent relation of the connectivity criterion. The behavior of the dynamic of the calcium for each particular cell should be

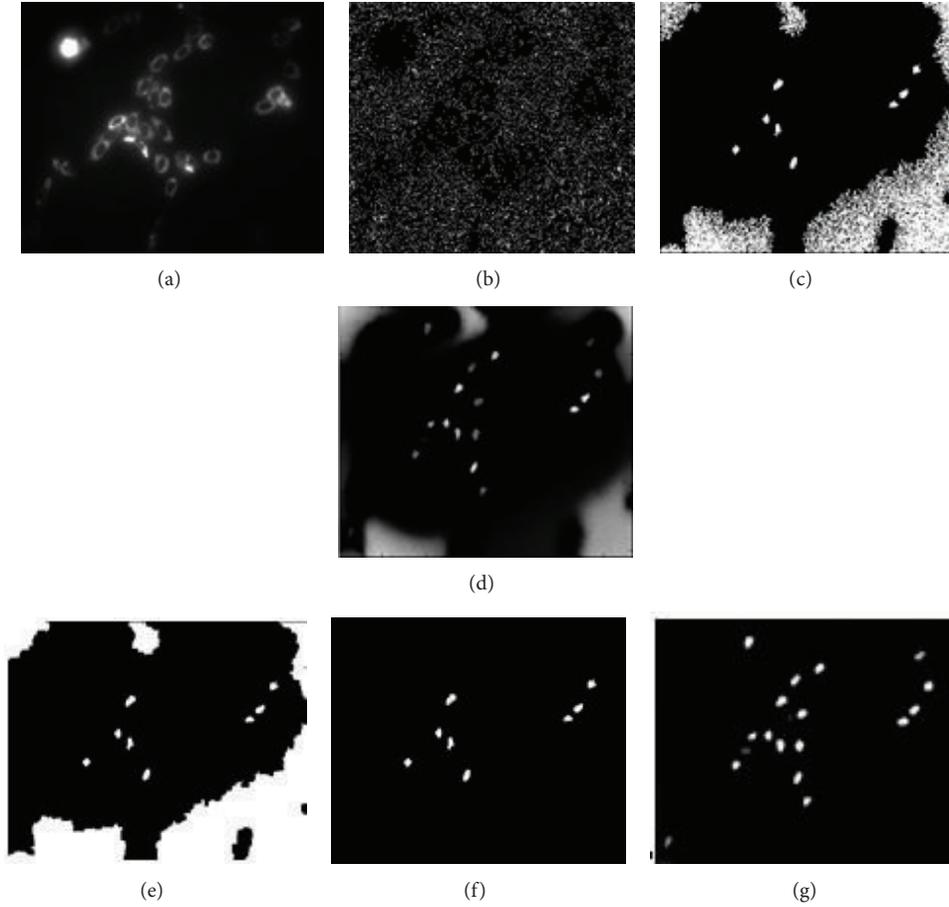


FIGURE 2: (a) Input image, (b) regional minima of original image, (c) minima obtained after applied closing by reconstruction, (d) function constructed from the minimum of the sequence of image, (e) morphological closing $\phi_{\lambda=3}$, (f) minima obtained by the difference: $M_i(x) = M_i(x) - \gamma_{\lambda=6} M_i(x)$, and (g) set of markers obtained by the function $I_m(x)$.

modeled as a polynomial decay time-decreasing function as follows:

$$y = \sum_{i=0}^n a_i x^i, \quad (9)$$

where α_i are the polynomial parameters, where the data used is taken from the maxima to the end of the data. The parameter estimation is performed by least squares the follow expression $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = (X^T X)^{-1} X^T \mathbf{y}$, such that $X = [\mathbf{x}^0 \ \dots \ \mathbf{x}^n]$, for \mathbf{x} data vertical vector. The order of polynomial is estimated from $(X^T X)^{-1}$ expression as follows: for a higher order n , the SVD decomposition of matrix $(X^T X)^{-1} = S \Sigma V^T$ is used. The rank of matrix is estimated when normalized information of eigenvalues represents 99.99% of information; this is $\sum_{j=1}^{n'} \sigma_j / \sum_{i=1}^n \sigma_i > 0.9999$, where each σ_i is taken from matrix $\Sigma = \begin{bmatrix} \sigma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_n \end{bmatrix}$ and n' represent the polynomial order for fitting. For illustration purposes in Figure 6, a fitting

sample is showed. The exponential help to model and analyze the decrease of the intensity registered in each cell.

4.3. Error Model Fitting. The correct construction of model over data is defined by introducing two measures of error: BIAS error and RMSE error. The first one is a measure of error modeling. The second measure is a precision error modeling criterion. The bias error provides information about how the model fits real data. Negative bias error means that model is undermodeling the data; that is, the model is a function under real data. Consequently, positive bias error represents overmodeling. Values near to zero mean that the model catches the dynamic of real data. Formally, bias error is defined as $\text{Bias}(x, x^*) = \sum_{i=0}^n x - x^*$, where x represent real data and x^* estimated data. Note that when BIAS is equal to zero it does not mean that the model is correct. It means that the same proportions of measures are below and under for real data. Then to quantify the precision error the RMSE is used. This error is the average of absolute differences among real and modeling data. RMSE is defined as follows: $\text{RMSE}(x, x^*) = (1/n) \sum_{i=1}^n (x^* - x)^2$ where x^* represent modeling function and x real data.

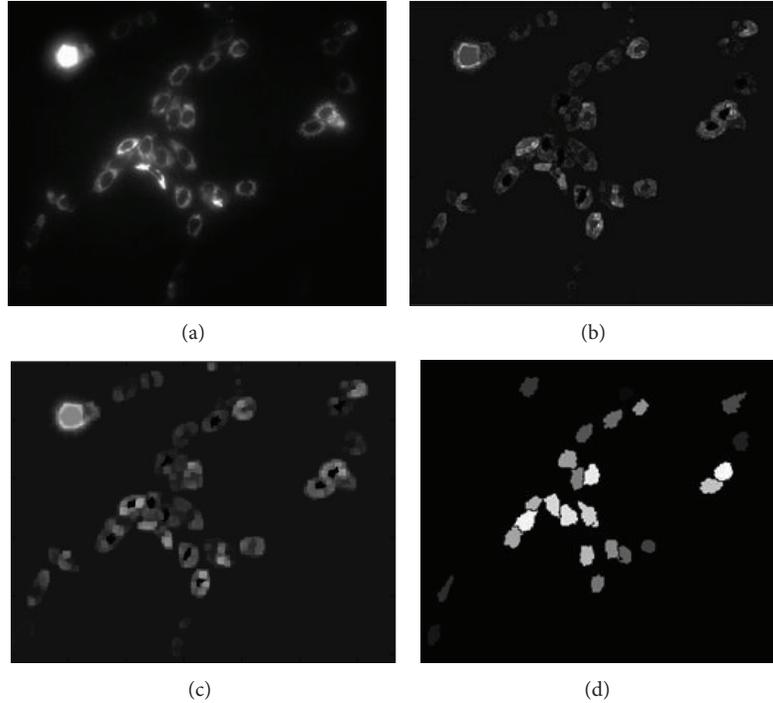


FIGURE 3: (a) Input image, (b) internal gradient, (c) external gradient, and (d) segmented cells.

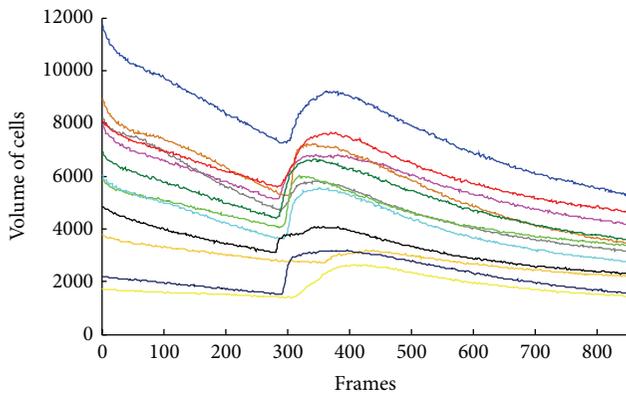


FIGURE 4: Cells volume curves over time showing an exponential decreasing behavior.

4.4. Enhancement of Data. Although the least squared method offers the optimal model, it depends on the data measurement having normal distribution. Then, by the nature of the model, it results hard in verifying that these measures have a normal distribution. As a consequence it is necessary to enhance the data in order to facilitate the convergence of the approach. For simplicity, it is assumed that any signal $V_n(i)$, resulted from the calculation of the volume of an n cell for an i time, is affected by additive noise with zero mean as follows:

$$V_n = V_n^* + N_n, \quad (10)$$

where V_n^* is the signal free noise and N_n is the additive noise with mean zero. In fact N_n has zero mean; the true signal data V_n^* is located into $\min\{\text{dom}(N_n)\}$ and $\max\{\text{dom}(N_n)\}$ values. However, given N_n is a random variable, it locally should not present a zero mean, making it difficult to estimate the V_n^* value. To figure it out, it is needed to analyze locally the information, inferring the trend, and make an estimation of the expected value. The proposal consists of exploiting certain properties of operators taken from morphology operators. The reconstruction operators are useful because they approximate a surface by iterating successively a marker, getting the other surface that has similar topological properties. The approximation does not keep the original level of detail of its shape, such that it depends on the form and properties of structural element used. It should be considered inconvenient, but, in practical terms, it is its major advantage in sense and it represents the main trend of the original data, eliminating variations less than the structural element (high frequencies) of the original signal and resulting in a new signal that under- or overmodels the original data.

Considering the basic operators by reconstruction (opening and closing), the property of extensive or antiextensive, respectively, cause the fact that the application of each one over a signal V_n results in $\tilde{\gamma}_{\mu L}(V)$ or $\tilde{\varphi}_{\mu L}(V)$ signals such as under- or overmodeling the original. Both of them remain as the global trend of the topological information of V_n . Consequently, the residual presents important topological information. However, the distribution of the data changes slightly: the shape of the derivative of the original and the approximated signal is different, changing the statistical properties of its PDFs. Figure 7 presents the probability density

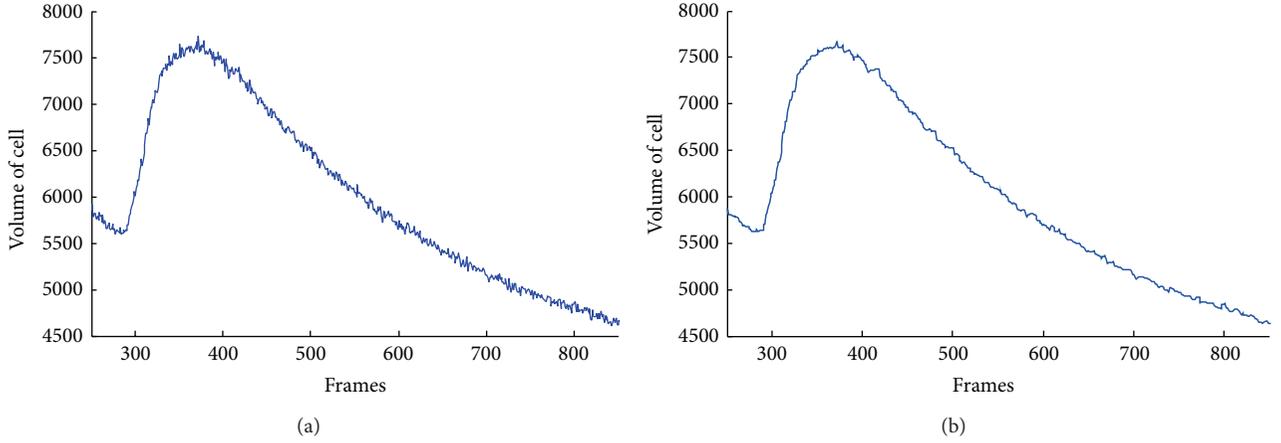


FIGURE 5: (a) Original signal that has multiple maxima caused by the noise interference. (b) Filtered signal presents a smoothing wave in which the global maximum is easy to detect.

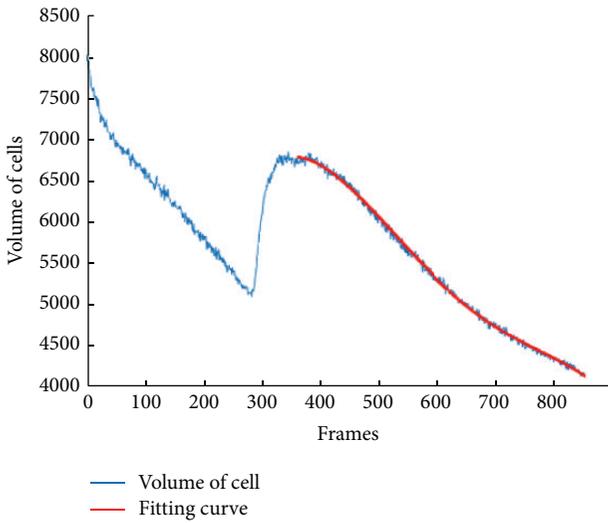


FIGURE 6: Fitting an exponential curve over the volume of cell behavior.

function (PDF) approximated via its histogram after application over a signal V_n . The histogram of opening operator presents a negative deviation, which means that the surface approximated is undermodeled. On the other hand, when we apply a closing operator, it overmodels the original signal and its histogram is deviated to the positive side of the range.

The proposal consists of mixing both filters, preserving the statistical information of the original signal. Noise effects are represented by the high frequencies. These frequencies must be discarded preserving the global trend of original signal V_n . The discarded frequencies are directly related to the size of the structural element and the sampling process; that is, given a structural element of size k , it represents a temporality of kf , where f is the mean frequency of acquisition of V_n . Then the process of filtering $f(V_n)$ is statistical consistent if and only if V_n^* , minus V_n preserve the following equality:

$$\rho(V_n - V_n^*) = G(0, \sigma). \quad (11)$$

This is, the density distribution function of the difference between filtered data and original data is a normal distribution centered at the origin. The development of correct statistical filter must satisfy (9), where it is appreciated that opening and closing reconstruction operators provide negative and positive bias information of the approximated surface. The original signal is enveloped by the opening reconstruction and closing reconstruction, respectively, such that $\tilde{\gamma}_{\mu L}(V) \leq V_n \leq \tilde{\varphi}_{\mu L}(V)$. Consequently, for estimating V_n , using $\tilde{\gamma}_{\mu L}(V)$ and $\tilde{\varphi}_{\mu L}(V)$ and considering that $E\{N_{1\dots n}\} = 0$, and approximation to V_n is

$$f_{\mu}^{\mu L}(V) = \alpha_1 \tilde{\gamma}_{\mu L}(V) + \alpha_2 \tilde{\varphi}_{\mu L}(V), \quad (12)$$

where α_1 and α_2 are values between $[0, 1]$ and its sum is the unit. In case that $\tilde{\gamma}_{\mu L}(V)$ and $\tilde{\varphi}_{\mu L}(V)$ use the same structural element, $\alpha_1 = \alpha_2 = 0.5$. In other cases these values would vary depending on effects of the geometry in the reconstruction process. The filter described above is denoted as a medium reconstruction filter. An extension of this filter implies a sequential form, where the properties of the structural element used in reconstruction stage should be varied as follows: let $p(\mu L, k)$ be a function that returns a structural element with particular properties for k instant; sequential version of medium reconstruction filter is defined as

$$f_{\mu}^{p(\mu L, k)}(V) = f_{\mu}^{p(\mu L, k)} \cdot f_{\mu}^{p(\mu L, k-1)}(V) \dots \dots f_{\mu}^{p(\mu L, 1)}(V). \quad (13)$$

Note that function $p(\mu L, k)$ would vary the size and the topology of the structural element. The topology and size will affect the model that fits the data. The effect of applying the medium filter by reconstruction is illustrated in Figure 8, where in Figure 8(a) are presented the original data (blue color) and the filtered data (red color). As is appreciated, filtered signal follows the main trend of the original signal, discarding the high frequencies, and always statistical properties are kept as it is appreciated in Figure 8(b). This figure shows the difference of filtered image and original. This property makes it ideal for filtering data, improving the results when raw

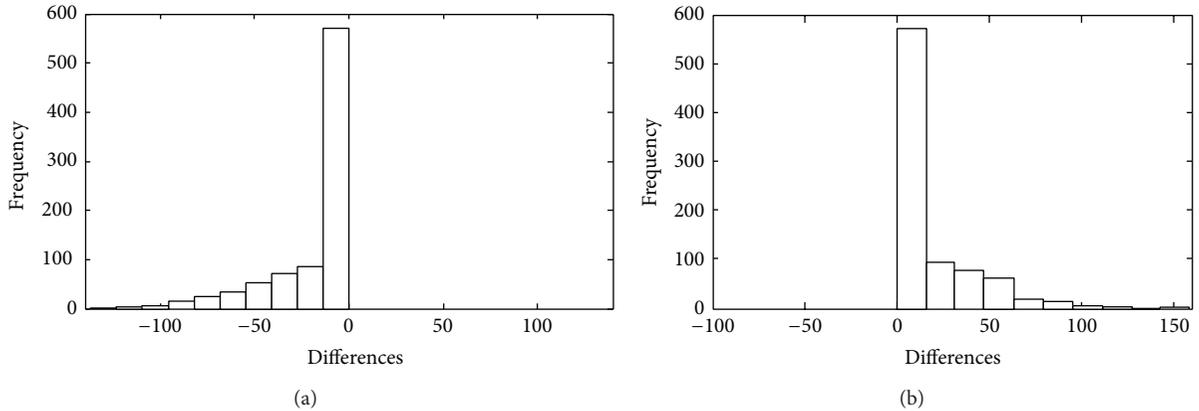


FIGURE 7: Histogram of differences from original signal and reconstructed signal (a) opening operator histogram, (b) closing operator histogram.

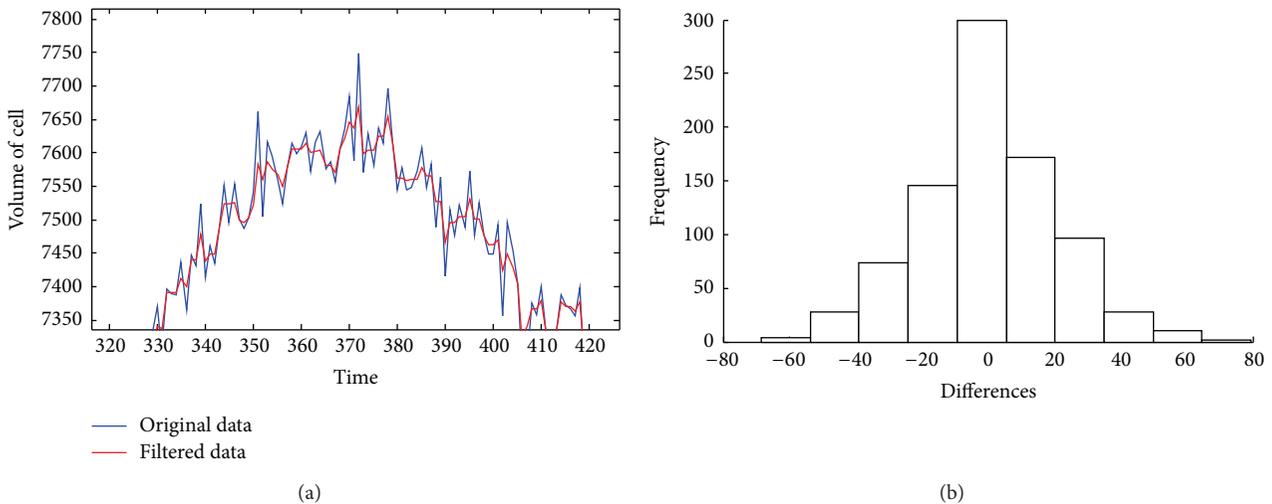


FIGURE 8: (a) Matching of original data and filtered data with the morphological medium filter. (b) Histogram of differences from original surface and reconstructed surface, as noted, the expected value is centered in zero and would be considered as a normal distribution.

data fit the exponential decreasing function. For a detailed analysis the bias and RMSE error are showed for the case of filter and nonfilter signal. Bias error behaves to close in both scenarios; but the RMSE is deeply reduced, which means that the fitting process results are better, after filter data (Table 1).

5. Results and Discuss

The proposal described above is tested under an experimental method that consists of analyzing a sequence of images that contains cells, which are exited applying Flour 4, in order to measure the effect over Ca^{2+} belonging to each cell and characterizing its behavior. The process is illustrated in Figure 9. The process diagram sums up the sequence of processing steps done over the sequence. The sequence of images was acquired from biological researchers of the Institute of Neurobiology, Campus UNAM-UAQ. The sequence was obtained from cells of *Xenopus laevis* frog. The calcium is

measured indirectly with its excitation via Fluo-4 (by Molecular Probes). The optical material consists of a microscopy of fluorescence, setting up in an Olympus camera sensor IX71 at 485 to 520 wavelength sensitivity nm (excitation-emission, respectably); finally the images were acquired with fast acquisition camera (Evolution Qe Media Cybernetics), at 30 frames per second (Fps) with a resolution of 320×240 pixels. Finally, for testing purposes, 1,000 images have been selected, which represents temporary a sequence of 33-second length. The cell detection is a tough task because there are many factors which inside directly in the analysis process, as the nonhomogeneity luminance conditions of the images and the conditions that present the cells of interest. After the image acquisition, the following task consists of finding out and segmenting each cell. This process is performed via watershed approach. However, starting from the first frame acquired does not warrant correct cell detection. To make more robust the cell process detection, for each image, the cells are detected, as described in the third th section. Once

TABLE 1: Errors of modeling with/without use of a median reconstruction filter.

Cell	Filtering		Without filter	
	BIAS	RMSE	BIAS	RMSE
1	0.0996	0.0415	0.0967	385.1000
2	0.5179	0.4742	0.1611	1304.4000
3	0.6913	1.7530	0.3772	7913.6000
4	1.1718	1.2087	0.3229	2906.9000
5	0.6538	0.7799	0.2144	2339.9000
6	0.7412	1.0012	0.1904	2306.6000
7	1.3467	1.2601	0.2079	1718.6000
8	1.1380	1.3794	0.2082	2282.1000
9	0.5010	0.4614	0.1823	1566.5000
10	1.4760	1.1212	0.1832	1198.6000
11	1.4346	0.9506	0.1091	620.1000
12	1.1730	1.6871	0.3167	4.1090
13	0.3320	0.2412	0.1132	0.5596
14	0.3510	0.1794	0.1601	703.2000
15	0.7043	0.7558	0.2293	2077.9000
16	1.6021	3.0539	0.2142	2721.2000
17	0.2594	0.1767	0.1605	999.4000
18	0.1688	0.0566	0.1144	343.6000

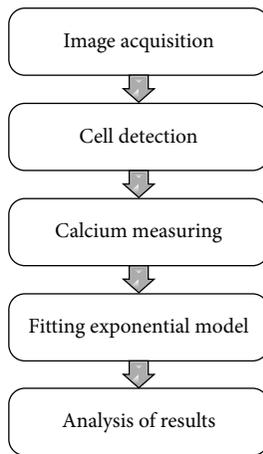
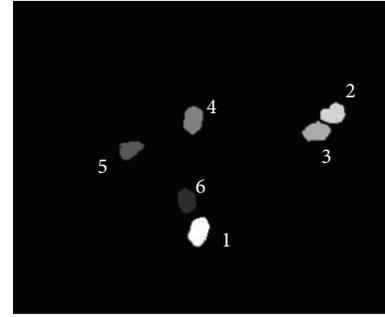
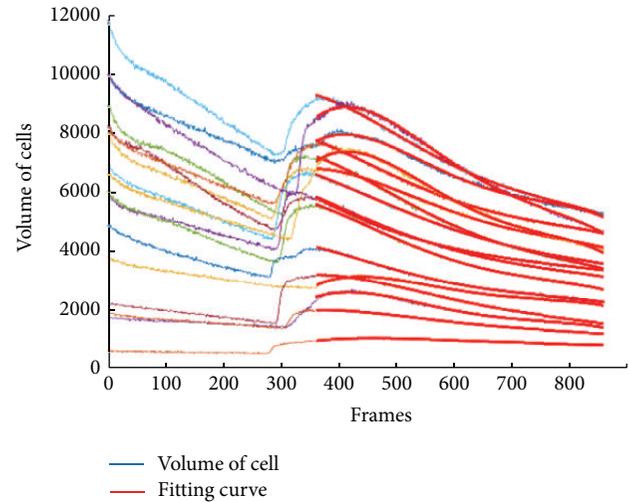


FIGURE 9: Block's diagram of the proposal.

the cells are detected, the neighborhood around the cell is considered to analyze the calcium concentration. The calcium concentration is performed by the measure of the luminance of each cell. The relation between the luminance intensity of the cell is highly correlated with the calcium concentration; that is, cells with high luminance have major calcium concentration. Next, the creation of the model of behavior results in a difficult task, because the behavior observed is not linear with the use of autoregression methods being inappropriate [15]. Then, discarding the times where the cell started to become excited, the dynamic of decreasing is modeled with an exponential function via least square method. The selected data includes the maxima location to the vanishing excited behavior. As recreated better the model, before applying least squared, the median filter by reconstruction is applied, which



(a)



(b)

FIGURE 10: (a) Cells segmented, (b) curves calculated by the least squares fitting.

improve the accuracy of the modeling. Finally the results are showed in Figure 10. Observe, in Figure 10(a), that the cells are detected and the dynamic modeled as exponential superposed over measured data is showed (Figure 10(b)). The use of filter dismisses the high frequencies smoothing the behavior of luminance variations. The dismissing of high frequencies adds extra accuracy warranting that the exponential fitting has more significance, although the data are affected by noise effects. Finally, the way of how the cells were segmented represents a framework to analyze the intracellular calcium, which segment automatically the set of cells. This process is convenient in the sense that many of microscopic dynamics could be analyzed efficiently providing better information to the biologists.

6. Conclusions

In this paper, an automatic method for the study of intracellular calcium based on a marked controlled watershed transform for segmenting stage is presented. A new filter based on reconstruction operators is introduced. Then, having a high precision of cell segmenting and efficient ways to discard the noise measurement result the base for an automatic frameworks analysis as the experimentation shows. Finally,

the reconstruction operators applied over one dimension data results usefully in the development of filters that help to create models of the dynamic of the calcium.

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 anonymous reviewers for their valuable comments since some ideas and precisions in this per come from them. The author I. Terol would like to thank Diego Rodrigo and Dario T. G. for their great encouragement. This work was funded by the government agency ASA-Conacyt (208466), México, 2014.

References

- [1] M. J. Berridge, P. Lipp, and M. D. Bootman, "The versatility and universality of calcium signalling," *Nature Reviews Molecular Cell Biology*, vol. 1, no. 1, pp. 11–21, 2000.
- [2] M. J. Berridge, M. D. Bootman, and H. L. Roderick, "Calcium signalling: dynamics, homeostasis and remodelling," *Nature Reviews Molecular Cell Biology*, vol. 4, no. 7, pp. 517–529, 2003.
- [3] B. F. Trump and I. K. Berezsky, "Role of sodium and calcium regulation in toxic cell injury," in *Drug Metabolism and Drug Toxicity*, Raven Press, New York, NY, USA, 1984.
- [4] S. A. Stricker, V. E. Centonze, S. W. Paddock, and G. Schatten, "Confocal microscopy of fertilization-induced calcium dynamics in sea urchin eggs," *Developmental Biology*, vol. 149, no. 2, pp. 370–380, 1992.
- [5] B. F. Trum, I. K. Berezsky, H. U. Laiho, A. R. Osornio, W. J. Mergner, and M. W. Smith, "The rol of calcium in cell injury. A review," *Scanning Electron Microscopy*, vol. 2, pp. 437–462, 1980.
- [6] M. Whitaker and R. Patel, "Calcium and cell cycle control," *Development*, vol. 108, no. 4, pp. 525–542, 1990.
- [7] T. W. Nattkemper, H. Wersing, W. Schubert, and H. Ritter, "A neural network architecture for automatic segmentation of fluorescence micrographs," *Neurocomputing*, vol. 48, no. 1–4, pp. 357–367, 2002.
- [8] C. Wählby, J. Lindblad, M. Vondrus, E. Bengtsson, and L. Björkesten, "Algorithms for cytoplasm segmentation of fluorescence labelled cells," *Analytical Cellular Pathology*, vol. 24, no. 2-3, pp. 101–111, 2002.
- [9] K. Wu, D. Gauthier, and M. Levine, *Live Cell Image Segmentation*, Center for Intelligent Machines, McGill University, Montreal, Canada, 1995.
- [10] V. Metzler, T. Lehmann, H. Bienert, K. Mottaghy, and K. Spitzer, "Scale-independent shape analysis for quantitative cytology using mathematical morphology," *Computers in Biology and Medicine*, vol. 30, no. 3, pp. 135–151, 2000.
- [11] F. Meyer and S. Beucher, "Morphological segmentation," *Journal of Visual Communication and Image Representation*, vol. 1, no. 1, pp. 21–46, 1990.
- [12] H. Heijmans, *Morphological Image Operators*, Academic Press, San Diego, Calif, USA, 1994.
- [13] P. Soille, *Morphological Image Analysis: Principles and Applications*, Springer, Berlin, Germany, 1999.
- [14] J. Serra, *Image Analysis and Mathematical Morphology: Theoretical Advances*, vol. 2, Academic Press, New York, NY, USA, 1988.
- [15] R. E. Walpole, R. H. Myers, and S. L. Myers, *Probability and Statistics for Engineers and Scientist*, Prentice-Hall, 9th edition, 2011.

Research Article

Feature Selection for Better Identification of Subtypes of Guillain-Barré Syndrome

José Hernández-Torruco,¹ Juana Canul-Reich,¹
Juan Frausto-Solís,² and Juan José Méndez-Castillo³

¹ *División Académica de Informática y Sistemas, Universidad Juárez Autónoma de Tabasco, Km. 1 Carretera Cunduacán-Jalpa de Méndez, Colonia La Esmeralda, 86690 Cunduacán, TAB, Mexico*

² *Universidad Politécnica del Estado de Morelos, Boulevard Cuauhnáhuac 566, Colonia Lomas del Texcal, 62574 Jiutepec, MOR, Mexico*

³ *Hospital General de Especialidades "Dr. Javier Buenfil Osorio", Avenida Lázaro Cárdenas 208, Colonia Las Flores, 24097 San Francisco de Campeche, CAM, Mexico*

Correspondence should be addressed to Juana Canul-Reich; juana.canul@ujat.mx

Received 6 June 2014; Revised 13 August 2014; Accepted 21 August 2014; Published 15 September 2014

Academic Editor: José M. Jerez

Copyright © 2014 José Hernández-Torruco 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.

Guillain-Barré syndrome (GBS) is a neurological disorder which has not been explored using clustering algorithms. Clustering algorithms perform more efficiently when they work only with relevant features. In this work, we applied correlation-based feature selection (CFS), chi-squared, information gain, symmetrical uncertainty, and consistency filter methods to select the most relevant features from a 156-feature real dataset. This dataset contains clinical, serological, and nerve conduction tests data obtained from GBS patients. The most relevant feature subsets, determined with each filter method, were used to identify four subtypes of GBS present in the dataset. We used partitions around medoids (PAM) clustering algorithm to form four clusters, corresponding to the GBS subtypes. We applied the purity of each cluster as evaluation measure. After experimentation, symmetrical uncertainty and information gain determined a feature subset of seven variables. These variables conformed as a dataset were used as input to PAM and reached a purity of 0.7984. This result leads to a first characterization of this syndrome using computational techniques.

1. Introduction

Guillain-Barré syndrome (GBS) is an autoimmune neurological disorder characterized by a fast evolution, generally from a few days up to four weeks [1]. GBS has an incidence of 1.3 to 2 per 100,000 people and a mortality rate from five to fifteen percent. The exact cause of GBS is unknown; however, it is frequently preceded by either a respiratory or a gastrointestinal infection. The diagnosis of GBS includes clinical, serological, and electrophysiological criteria [2]. The severity of GBS varies among subtypes, which can be mainly acute inflammatory demyelinating polyneuropathy (AIDP), acute motor axonal neuropathy (AMAN), acute motor sensory axonal neuropathy (AMSAN), and Miller-Fisher syndrome [1]. Electrodiagnostic criteria for distinguishing AIDP, AMAN, and AMSAN are well established in the literature [3], while the Miller-Fisher subtype is characterized by the clinical triad: ophthalmoplegia, ataxia, and areflexia [1].

A better understanding of the differences in the GBS subtypes is critical for the implementation of appropriate treatments for total recovery and in certain cases for the survival of patients. Hospitalization time and the cost of treatments vary according to the severity of the specific subtype. Finding a minimum feature subset to accurately identify GBS subtypes could lead to a simplified and cheaper process of diagnosis and treatment of the GBS case. The ultimate goal of a physician is to get patients to a full recovery. This can be more effectively achieved when an early diagnosis of the case is performed using a minimum number of medical features.

This work constitutes a first attempt to using machine learning techniques, specifically cluster analysis in combination with filter methods for feature selection. We aim at finding a small feature subset to identify four GBS subtypes. Machine learning techniques have been found in the literature to predict the prognoses of this syndrome [4, 5] as well as

to find predictors of respiratory failure and necessity of mechanical ventilation in GBS patients [6–8]. Nevertheless, no previous publications about specific subtypes identification of the syndrome using machine learning techniques were found in the literature.

Cluster analysis is a computational technique from the machine learning area that is shown to be useful to find different groups of objects in datasets [9–12]. However, datasets might contain a mixture of “bad” and “good” features. “Bad” features are redundant or noisy features and make algorithms slow and inaccurate. Feature selection techniques allow reducing the dimensionality of a dataset such that it only contains “good” features which would maximize the performance of the algorithms and thus enabling the possibility of reaching a higher accuracy [13]. For feature selection, several machine learning methods are available, which are usually classified as filter [14–19], wrapper [20–22], embedded [23–25], and hybrid [26–29]. From the machine learning point of view it is interesting to analyze the performance of feature selection methods in diverse scenarios with real data, as this case is.

In this work we use a real dataset consisting of 156 features and 129 cases of GBS patients; these are 20 AIDP cases, 37 AMAN, 59 AMSAN, and 13 Miller-Fisher cases. The dataset contains clinical, serological, and nerve conduction tests data.

We use PAM (Partitions Around Medoids) clustering algorithm to identify with the highest purity groups corresponding to four subtypes of GBS. A group with high purity contains the largest number of elements of the same type and the fewest number of elements of a different type. Purity is an external clustering validation metric that evaluates the quality of a clustering based on the grouping of objects into clusters and comparing this grouping with the ground truth. Although there are several clustering validation metrics, both internal and external [30], we selected purity since our interest was to find “pure” groups and to take advantage of the available prior knowledge of the true labels. The use of a prior knowledge to evaluate a clustering process is also known as supervised or semisupervised clustering; some examples can be found in [31–34].

In order to achieve the identification of the four groups with a high purity it is necessary to select the relevant features in the dataset; otherwise the purity magnitude would be compromised as stated in [13]. For this initial exploratory study, we chose filter methods as they are the simplest and lowest computational demanding methods available in the literature and as they work independently of the clustering algorithms. We focus on five filter methods: correlation-based feature selection (CFS), chi-squared, information gain, consistency, and symmetrical uncertainty methods.

The experimental results showed a good performance of the method and allowed us to obtain a first characterization of GBS using machine learning techniques.

2. Materials and Methods

2.1. Data. The dataset used in this work comprises 129 cases of patients seen at Instituto Nacional de Neurología y Neurocirugía located in Mexico City. Data were collected from 1993 through 2002. There are 20 AIDP cases, 37 AMAN,

59 AMSAN, and 13 Miller-Fisher cases. The identification of subtypes was made by a group of neurophysiologists based on the clinical and electrophysiological criteria established in the literature [1–3]. This dataset is not yet publicly available and this is the first time it is used in an experimental study. No public dataset was found to be used as a benchmark.

Originally, the dataset consisted of 365 attributes corresponding to epidemiological data, clinical data, results from two nerve conduction tests, and results from two cerebrospinal fluid (CSF) analyses. The second nerve conduction test was conducted in 22 patients and the second CSF analysis was conducted in 47 patients only. Therefore, data from these two tests were excluded from the dataset.

The diagnostic criteria for GBS are established in the literature [1–3]. These formal criteria were considered to determine which variables from the original dataset could be important in the characterization of the four subtypes of GBS. We made a preselection of variables based on these criteria. Originally, the dataset had 365 variables. After preselection, it was left with 156 variables: 121 variables from the nerve conduction test, 4 variables from the CSF analysis, and 31 clinical variables. As for the type of attributes, these are 28 categorical and 128 numeric attributes. The situation of dealing with mixed data types was solved using Gower’s similarity coefficient, as explained later.

2.2. Filter Methods. We selected filter methods for this initial exploratory study as they are in computational terms the fastest and simplest methods available in the literature for feature selection. Filters work independently from any clustering algorithm and base their decision solely on characteristics of data.

We chose these five particular methods based on their performance reported in the literature [15, 17, 35, 36]. Chosen filters apply diverse criteria to evaluate feature relevance. Filters investigated are CFS, chi-squared, information gain, symmetrical uncertainty, and consistency.

2.2.1. Correlation-Based Feature Selection (CFS). CFS [14] evaluates two aspects of a feature subset: its capacity to predict the class and the correlation between the features of the subset. This method seeks to maximize the first aspect and minimize the second one. This method results in a feature subset with the highest capacity to predict the class and the least correlation between features of the subset. Given a feature subset S containing k features, CFS finds the goodness of S denoted (M_S) as follows:

$$M_S = \frac{k\bar{r}_{cf}}{\sqrt{k + k(k-1)\bar{r}_{ff}}}, \quad (1)$$

where \bar{r}_{ff} is the average correlation of all feature-feature pairs, and $k\bar{r}_{cf}$ is the average correlation of all feature-class pairs.

2.2.2. Chi-Squared. This method evaluates the chi-square statistic of each feature taken individually with respect to the

- (1) k objects are arbitrarily selected as the initial medoids m .
- (2) **repeat**
 - (2.1) The distance is computed between each remaining object o and the medoids m .
 - (2.2) Each object o is assigned to the cluster with the nearest medoid m .
 - (2.3) An initial total cost E_{ini} is calculated.
 - (2.4) A random o is selected.
 - (2.5) A total cost E_{fin} is calculated as a result of swapping an arbitrary m with the o randomly selected.
 - (2.6) If $E_{\text{fin}} - E_{\text{ini}} < 0$ then m is replaced with o .
- until** $E_{\text{fin}} - E_{\text{ini}} = 0$.

ALGORITHM 1: Partitions around medoids (PAM).

class [15] and provides a feature ranking as a result. The chi-square test for a feature f and the class c is defined as follows:

$$X^2(f, c) = \frac{N \left[P(f, c) P(\bar{f}, \bar{c}) - P(f, \bar{c}) P(\bar{f}, c) \right]^2}{P(f) P(\bar{f}) P(c) P(\bar{c})}, \quad (2)$$

where N is the number of observations in the dataset, $P(x, y)$ is the joint probability of x and y , and $P(x)$ is the marginal probability of x .

2.2.3. Information Gain. Information gain measures the goodness of a feature to predict the class given that the presence or absence of the feature in the dataset is known. This method delivers a ranking according to the goodness of each feature.

Information gain [16] of a feature f_k and a class c_i is defined as follows:

$$IG(f_k, c_i) = \sum_{c \in \{c_i, \bar{c}_i\}} \sum_{f \in \{f_k, \bar{f}_k\}} P(f, c) \log_2 \frac{P(f, c)}{P(f) P(c)}, \quad (3)$$

where $c \in \{c_i, \bar{c}_i\}$ is the set of all classes, $f \in \{f_k, \bar{f}_k\}$ is the set of all features, $P(f, c)$ is the joint probability of feature f and class c , and $P(f)$ and $P(c)$ are the marginal probabilities of f and c , respectively.

2.2.4. Consistency. This method finds the smallest feature subset that presumably improves the discriminatory power of the original feature subset. This subset has the highest consistency. The consistency for a given feature subset S is computed as follows [17]:

$$\text{Consistency} = 1 - \text{inconsistency rate}. \quad (4)$$

Let us define a pattern as a set of values for S . An inconsistency arises when two patterns match exactly all attributes except for the class. The inconsistency count for a pattern is the number of times it appears in the dataset minus the number of times it appears in the majority class. The inconsistency rate is the sum of all the inconsistency counts for all possible patterns of S divided by the total number of patterns [18].

2.2.5. Symmetrical Uncertainty. This method measures the correlation between pairs of attributes using normalization of information gain. The normalization is performed to

compensate for the bias of information gain to benefit attributes with more values and to ensure that they are comparable [17]. This method results in a feature ranking.

Symmetrical uncertainty is computed as follows [19]:

$$U(A, B) = 2 * \frac{MI(A, B)}{\text{Entropy}(A) + \text{Entropy}(B)} \quad (5)$$

$$MI(A, B) = \sum P(A, B) \log_2 \frac{P(A, B)}{P(A) P(B)},$$

where $P(X)$ is the marginal probability of feature X , R_A is the range of feature A , and $P(A, B)$ is the joint probability of features A and B . Entropy is computed using the classical equation discussed in [17].

2.3. Clustering Algorithm: Partitions Around Medoids (PAM). As stated before, the dataset used in this work combines categorical and numeric data. PAM is a clustering algorithm capable of handling such situations. It receives a distance matrix between observations as input. The distance matrix was computed using Gower's coefficient, explained later.

PAM, introduced by Kaufman and Rousseeuw [37], aims to group data around the most central item of each group, known as medoid, which has the minimum sum of dissimilarities with respect to all data points. PAM forms clusters that minimize the total cost E of the configuration, defined as

$$E = \sum_{i=1}^k \sum_{o \in C_i} \text{dist}(o, m), \quad (6)$$

where k is the number of clusters, $o \in C_i$ is the set of objects in cluster C_i , and $\text{dist}(o, m)$ is the distance between an object o and a medoid m .

PAM works as shown in Algorithm 1 [38].

2.4. Gower's Similarity Coefficient. Distance metrics are used in clustering tasks to compute the distance between objects. The distance computed is used by clustering algorithms to determine how much similar or dissimilar the objects are and what cluster they belong to. There are many distance metrics. Some of them deal with numeric data, like Euclidean, Manhattan, and Minkowski [38]. To deal with binary data the Jaccard coefficient and Hamming are often used [38]. For categorical data, some distance metrics are Overlap, Goodall, and Gambaryan [39].

In this work we used for experimentation a dataset that contains mixed data, that is, both categorical and numeric data. To deal with this situation we selected Gower's coefficient. It is a robust and widely used distance metric for mixed data. We used this coefficient to obtain a matrix of distances between observations as PAM requires. It was introduced by Gower in 1971 [40]. Gower's coefficient is defined as follows [41]:

$$d_{ij}^2 = 1 - S_{ij} \quad (7)$$

$$S_{ij} = \frac{\sum_{h=1}^{p1} (1 - |x_{ih} - x_{jh}| / G_h) + a + \alpha}{p1 + p2 - d + p3},$$

where $p1$ is the number of quantitative variables, $p2$ is the number of binary variables, $p3$ is the number of qualitative variables, α is the number of coincidences for qualitative variables, a is the number of coincidences in 1 (feature presence) for binary variables, d is the number of coincidences in 0 (feature absence) for binary variables, and G_h is the range of the h th quantitative variable.

Gower's coefficient is within the range 0-1. A value near to 1 indicates strong similarity between items and a value near to 0 indicates weak similarity.

2.5. Metrics to Evaluate the Quality of a Clustering Process.

The quality of a clustering process can be evaluated using two types of metrics: internal and external. Internal metrics evaluate the quality of a clustering process based on some intrinsic characteristics, regularly, intra- and intercluster distances. Internal metrics assign high scores to clusters with largest distances among them (separability) and shortest distances among members of the same cluster (compactness). These metrics are very useful when the number of clusters is not known at all. Examples of internal metrics are Q-modularity [42], Davies-Bouldin index, Dunn index, and silhouette [43].

External metrics evaluate the quality of clusters based on data not used during the clustering process, such as the ground truth, that is, the real classes of the instances. The larger the number of instances correctly located according to the ground truth, the higher the index. Some examples of external metrics are Rand index, Folkes and Mallows index, Hubert's T statistic [30], and purity [44].

2.5.1. Purity. The dataset used in this work provides the ground truth. We know there are four classes in the dataset. The objective of this study was to find the features that identify with the highest accuracy possible four clusters, each corresponding to one class. To achieve this goal we selected purity as the metric to evaluate the quality of the clustering process.

Purity validates the quality of a clustering process based on the locations of data in each cluster with respect to the true classes. The more objects in each resultant cluster belong to the true class, the higher the purity. Formally [44],

$$\text{purity}(C, W) = \frac{1}{N} \sum_k \max_j (n_j^i), \quad (8)$$

where N is the number of samples, $C = \{c_1, c_2, \dots, c_k\}$ is the set of clusters found by the clustering algorithm,

TABLE 1: Purity of a clustering with three classes.

	Class A	Class B	Class C	
Cluster 1	0	14	1	15
Cluster 2	9	2	0	11
Cluster 3	3	1	21	25
	12	17	22	51

$W = \{w_1, w_2, \dots, w_k\}$ is the set of the classes of the objects, $n_j^i = |w_i \cap c_j|$ is the number of objects of cluster i being in class j , w_i is the set of objects in class k , and c_j is the set of objects in cluster k .

The value of purity ranges from 0 to 1. A purity value of 1 indicates that all the objects in each cluster belong to the same class. An example of purity calculation is shown in Table 1.

The number of objects of the majority class in each cluster is shown in bold. The purity of the clustering is computed as follows: $(9 + 14 + 21)/51 = 44/51 = 0.8627$.

3. Results and Discussion

3.1. Experimental Design. We used the 156-feature GBS dataset, described earlier, for experiments. This dataset contains a combination of categorical and numeric features. Gower's coefficient is able to deal with both types of features when present in the same dataset. We used this method to compute the distance matrix among instances, which is required as input to the PAM algorithm.

As we know beforehand, there are four GBS subtypes present in our dataset. This is why the number of clusters requested to PAM algorithm in our experiments was $k = 4$. We expected the clustering algorithm would identify each subtype as a cluster, with the highest purity possible. Five filter methods were used for feature selection, as clustering algorithms perform more efficiently when they work only with relevant attributes [13].

The class attribute was not used when the clustering algorithm was executed. We used it to compute the purity of the clusters obtained with PAM.

A baseline purity using all the 156 features included in the dataset was computed. This value was compared with the purity obtained using only the relevant features as determined by each filter method. Such comparison would allow for a clear view of the benefits of the feature selection process over using the entire dataset, in terms of purity.

Each of the five filter methods selected for experiments in this work was applied to the 156-feature dataset. Along with the features, the class attribute was included in the dataset during the filtering process.

As previously described, CFS and consistency methods include in their output the subset with the most relevant features found. In contrast, chi-squared, information gain, and symmetrical uncertainty methods output a feature ranking.

In all scenarios, new datasets were created with the best feature subsets. The distance matrix of the new datasets was calculated and used as input to the PAM algorithm. Finally, purity of clusters was computed.

TABLE 2: Results of filter methods ranked on purity.

Method	Number of features	Purity
Information gain	7	0.7984
Symmetrical uncertainty	7	0.7984
CFS	16	0.7984
Chi-squared	41	0.7829
Consistency	6	0.6589

TABLE 3: List of variables with the highest purity (0.7984) selected by information gain and symmetrical uncertainty.

Feature	Meaning
v105	Amplitude of left ulnar motor nerve
v106*	Area under the curve of left ulnar motor nerve
v116	Amplitude of right ulnar motor nerve
v172*	Amplitude of left median sensory nerve
v177*	Amplitude of right median sensory nerve
v182*	Amplitude of left ulnar sensory nerve
v187	Amplitude of right ulnar sensory nerve

In both CFS and consistency methods, the new datasets were created with the resultant most relevant features.

For chi-squared, information gain, and symmetrical uncertainty, feature rankings they produced were used to create the new datasets. Datasets with dimension from 2 through 156 were created, with the best two features, the best three features, and so on. The reason for a dataset of dimension 2 is that the calculation of the distance matrix requires at least 2 attributes. The best feature subset was the set of features conforming the dataset which led to the highest purity in the clustering process.

3.2. Results

3.2.1. *Identification of the Four GBS Subtypes.* The baseline purity of the four clusters obtained using all the 156 features included in the dataset was 0.6899. After experimentation, four filter methods found feature subsets which increased the baseline purity after the clustering process. Only the feature subset selected by the consistency method as the most relevant obtained a lower purity of 0.6589 than that of the baseline experiment.

Table 2 shows the results of purity of the five methods. Three methods tied with the highest purity (0.7984): information gain, symmetrical uncertainty, and CFS. Both information gain and symmetrical uncertainty selected seven relevant features while CFS selected 16 relevant features. Chi-squared method chose 41 nerve conduction test variables as the most relevant and reached 0.7829 of purity. The consistency method showed the worst performance, which reached a purity of 0.6589. The six relevant features selected by consistency method were two clinical and four corresponding to the nerve conduction test.

Table 3 shows the list of the variables selected by both information gain and symmetrical uncertainty. These variables conformed as a dataset were able to identify the four subtypes of GBS with a purity of 0.7984. All these variables are related to the nerve conduction test.

TABLE 4: List of variables with the highest purity (0.7984) selected by CFS.

Feature	Meaning
v29	Extraocular muscles involvement
v30	Ptosis
v40	Karnofsky at discharge
v105	Distal amplitude of left ulnar motor nerve
v106*	Area under the curve of left ulnar motor nerve
v108	Proximal amplitude of left ulnar motor nerve
v111	Average F-wave latency of left ulnar motor nerve
v116	Distal amplitude of right ulnar motor nerve
v134	F-wave amplitude of left tibial motor nerve
v172*	Amplitude of left median sensory nerve
v173	Area under the curve of left median sensory nerve
v177*	Amplitude of right median sensory nerve
v182*	Amplitude of left ulnar sensory nerve
v185	Conduction velocity of right ulnar sensory nerve
v187	Amplitude of right ulnar sensory nerve
v192	Amplitude of left sural sensory nerve

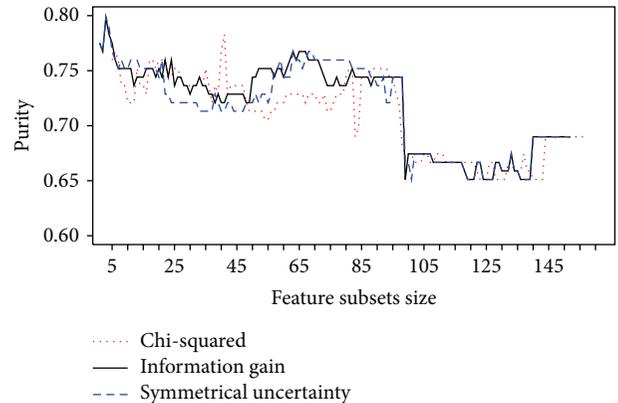


FIGURE 1: Purity reached by the best feature subsets as ranked by chi-squared, information gain, and symmetrical uncertainty methods.

CFS picked out 16 relevant variables, three of them clinical and 13 corresponding to the nerve conduction test, which reached a purity of 0.7984 as well. The list of 16 variables is shown in Table 4.

Four variables from Tables 3 and 4, denoted by (*), were selected by all methods.

Purity results of the clustering process using the datasets formed with the most relevant features as ranked by chi-squared, information gain, and symmetrical uncertainty, as described in methodology section, are shown in Figure 1. The three methods behave similarly. Both information gain and symmetrical uncertainty methods reached a maximum value with seven relevant variables, while chi-squared method reached its maximum with 41 variables. All three methods kept purity in the range of 0.7 and 0.8 for feature subsets of sizes between 2 and 102. For bigger subsets, purity lies in the range of 0.65 through 0.7.

TABLE 5: Purity of pairwise clustering of GBS subtypes.

GBS subtypes	IG	SU	CFS	Consistency	Chi-squared	All features
AIDP and AMAN	0.9649 (2)	0.9649 (2)	*	0.9824 (2)	0.9649 (2)	0.8771
AMAN and AMSAN	0.927 (3)	0.9375 (2)	0.875 (7)	0.927 (3)	0.9687 (4)	0.7395
AIDP and AMSAN	0.962 (3)	0.9367 (2)	0.9113 (9)	0.7468 (5)	0.962 (3)	0.8354
AIDP and MF	0.8787 (3)	0.8787 (4)	0.8787 (4)	0.8787 (3)	0.909 (3)	0.6666
AMAN and MF	0.98 (6)	0.96 (3)	0.98 (4)	0.74 (2)	0.98 (6)	0.96
AMSAN and MF	0.9305 (13)	0.9444 (13)	0.9583 (14)	0.8194 (5)	0.9444 (14)	0.8333

IG: information gain, SU: symmetrical uncertainty, and * one feature selected therefore purity was not computed. The number of features selected in each case is shown in parenthesis.

TABLE 6: Purity for different numbers of clusters using the four GBS subtypes.

k	IG	SU	CFS	Consistency	Chi-squared	All features
2	0.6434 (9)	0.6434 (10)	0.6511 (16)	0.4883 (6)	0.6124 (48)	0.5038
3	0.7906 (7)	0.7906 (7)	0.7286 (16)	0.6666 (6)	0.7829 (6)	0.5813
4	0.7984 (7)	0.7984 (7)	0.7984 (16)	0.6589 (6)	0.7829 (41)	0.6899
5	0.7984 (5)	0.7984 (5)	0.7906 (16)	0.7286 (6)	0.7829 (91)	0.6821
6	0.7751 (4)	0.7751 (7)	0.8139 (16)	0.7596 (6)	0.7751 (38)	0.6666
10	0.8139 (5)	0.8139 (5)	0.8217 (16)	0.7596 (6)	0.8062 (38)	0.6976
20	0.8449 (38)	0.8372 (31)	0.8527 (16)	0.8294 (6)	0.8294 (53)	0.7596

k : number of clusters, IG: information gain, and SU: symmetrical uncertainty. The number of features selected in each case is shown in parenthesis.

3.2.2. Pairwise Exploration of the GBS Subtypes. In order to investigate if any two pairs of GBS subtypes were distinguishable we conducted an additional experiment. We created six new datasets, each one containing instances of only two GBS subtypes. We calculated a baseline purity of each pair of GBS subtypes using all the 156 features. Our goal was to determine a feature subset capable of identifying each pair of GBS subtypes with a higher purity than that of the baseline. We used the five filter methods investigated all along this work to determine the most relevant features for each pair of GBS subtypes. For all scenarios we used $k = 2$, as there are only two GBS subtypes in each dataset. Finally, we applied PAM to form the clusters using only the relevant features determined with each filter method and calculated their purity.

Table 5 shows the results of this experiment. Each row represents a pair of GBS subtypes. Columns 2 to 6 represent a filter method. The right-most column indicates the purity achieved using all the features in the dataset, that is, doing no feature selection at all. Table entries indicate the purity obtained in each case. Numbers in bold show the highest purity obtained for each pair of GBS subtype. Based on the purity obtained, it was found that any filter method is better than using all the features. The highest purity for all pairs of GBS subtypes was superior to 0.9. This result demonstrates the effectiveness of filter methods and highlights the importance of feature selection.

3.2.3. Exploring Different Values of k . As explained at the beginning of Section 3.1, we performed the clustering process requesting $k = 4$ clusters as we know this is the number of existing GBS subtypes in the dataset. However, we wanted to explore the clustering process with different values of k . Purity results were analyzed and shown in Table 6.

The results of this experiment are shown in Table 6. The first column represents the different values of k analyzed. Each remaining column represents a filter method. The right-most column represents the purity obtained using all the features, that is, doing no feature selection at all. Each row represents the results obtained for each value of k . Table entries indicate the purity obtained in each case. The results indicate that, in general, purity keeps an ascending pattern as k increases. Purities for $k = 4$ and $k = 5$ are very close. In all cases, purity is low for $k = 2$ and very high for $k = 20$. The highest purity values were found for $k = 20$ in all cases; however, these numbers do not indicate that the real number of clusters in the dataset is 20; in fact this number of clusters does not correspond with the nature of GBS subtypes in real life. This result confirms what is reported in literature; higher values of purity are easily obtained for higher values of k [45]. Purity is a good evaluation metric for clustering when the number of clusters is known, as in this case.

3.3. Discussion. Our objective in this work was to find the best feature subset to identify four GBS subtypes with the highest purity. We did not find any similar work in the literature; therefore this one represents the first effort in this direction. In order to achieve our purpose, we applied machine learning techniques. We used five filter methods for feature selection and compared their performance.

3.3.1. Importance of Feature Selection to Identify GBS Subtypes. The clustering of the four GBS subtypes using all the 156 features in the dataset reached a purity of 0.6899. This means that many cases were mislocated in the clustering process. Table 2 shows that four of the five feature selection methods used in this work obtained a small feature subset that led to the identification of the four groups with a higher purity than that of the baseline.

The identification of GBS subtypes pairwise was achieved with a high purity. The initial baseline purity was improved in all cases (Table 5) when the algorithm used only the relevant features.

These results demonstrate that the clustering algorithm underperforms in the presence of redundant and irrelevant features and highlight the importance of feature selection methods.

3.3.2. Analysis of Different Numbers of Clusters. Purity is a good evaluation metric for clustering when the number of clusters is known, as in this case. Higher purity is easily achieved as the number of clusters increases [45] and that is demonstrated with the results shown in Table 6.

3.3.3. Identification of Four GBS Subtypes. The main contribution of this work is the identification of a subset of seven relevant features from a dataset of 156 variables which identified four GBS subtypes with a purity of 0.7984. Another contribution is the analysis of the performance of five filter methods for feature selection. Finally, this work contributes with the feature rankings produced by chi-squared, information gain, and symmetrical uncertainty methods.

A remarkable finding is that all five methods coincided in four variables. It is also noteworthy that only two of the five methods selected clinical variables. It is important to highlight the fact that the consistency method was not able to select a feature subset to improve the baseline purity (0.6899), but instead the six features selected by this method achieved a worse purity (0.6589).

Information gain, symmetrical uncertainty, and CFS were showed to be highly efficient as they could obtain a reduced subset of relevant features that allow identifying four subtypes of GBS with high purity (0.7984). The first two methods coincided in the same seven variables. CSF selected 16 variables. Further studies are needed to evaluate other methods of feature selection, such as wrapper, embedded, and hybrid methods.

4. Conclusions

In this work, we aimed to find a reduced feature subset for identifying four subtypes of GBS with the highest purity. This work represents the first effort on using cluster analysis to identify GBS subtypes. We used for experiments a real dataset of 156 features containing clinical, serological, and nerve conduction tests data. A clustering process was performed with PAM algorithm. In order to select the most relevant features from the dataset as input for PAM, we conducted experiments with five filter methods: CFS, chi-squared, information gain, symmetrical uncertainty, and consistency.

We succeeded as two filter methods were able to find a feature subset consisting of only seven variables that allowed us to obtain a purity of 0.7984. This result originated the first computational characterization of GBS subtypes. Besides, the reduced number of features found to identify the four GBS subtypes could guide physicians to design a faster, simpler, and cheaper diagnosis of the syndrome case.

Other filter methods like FCBF (Fast Correlation-Based Filter) [46] and INTERACT [47] could be used in further studies. Also, more sophisticated methods of feature selection are recommended for analysis, such as those listed in [48–50].

Finally, machine learning techniques such as neural networks or support vector machines could be used for clustering. Purity on their resultant clusters can be compared to that of PAM. This study is planned to further our research.

Conflict of Interests

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

Authors' Contribution

The first two authors contributed equally to this work.

References

- [1] S. Kuwabara, "Guillain-Barré syndrome," *Drugs*, vol. 64, no. 6, pp. 597–610, 2004.
- [2] S. I. Pascual Pascual, *Protocolos Diagnóstico Terapéuticos de la AEP: Neurología Pediátrica. Síndrome de Guillain-Barré*, Asociación Española de Pediatría, Madrid, Spain, 2008.
- [3] A. Uncini and S. Kuwabara, "Electrodiagnostic criteria for Guillain-Barré syndrome: a critical revision and the need for an update," *Clinical Neurophysiology*, vol. 123, no. 8, pp. 1487–1495, 2012.
- [4] R. van Koningsveld, E. W. Steyerberg, R. A. Hughes, A. V. Swan, P. A. van Doorn, and B. C. Jacobs, "A clinical prognostic scoring system for Guillain-Barré syndrome," *The Lancet Neurology*, vol. 6, no. 7, pp. 589–594, 2007.
- [5] C. Walgaard, H. F. Lingsma, L. Ruts, P. A. van Doorn, E. W. Steyerberg, and B. C. Jacobs, "Early recognition of poor prognosis in Guillain-Barré syndrome," *Neurology*, vol. 76, no. 11, pp. 968–975, 2011.
- [6] M.-C. Durand, R. Porcher, D. Orlikowski et al., "Clinical and electrophysiological predictors of respiratory failure in Guillain-Barré syndrome: a prospective study," *The Lancet Neurology*, vol. 5, no. 12, pp. 1021–1028, 2006.
- [7] B. S. Paul, R. Bhatia, K. Prasad, M. V. Padma, M. Tripathi, and M. B. Singh, "Clinical predictors of mechanical ventilation in Guillain-Barré syndrome," *Neurology India*, vol. 60, no. 2, pp. 150–153, 2012.
- [8] C. Walgaard, H. F. Lingsma, L. Ruts et al., "Prediction of respiratory insufficiency in Guillain-Barré syndrome," *Annals of Neurology*, vol. 67, no. 6, pp. 781–787, 2010.
- [9] A. V. Babkin, T. J. Kudryavtseva, and S. A. Utkina, "Identification and analysis of industrial cluster structure," *World Applied Sciences Journal*, vol. 28, no. 10, pp. 1408–1413, 2013.
- [10] J. L. Bravo Cabrera, E. Azpra Romero, V. Zarraluqui Such, C. Gay García, and F. Estrada Porrúa, "Cluster analysis for validated climatology stations using precipitation in Mexico," *Atmosfera*, vol. 25, no. 4, pp. 339–354, 2012.
- [11] P. R. Burgel, J. L. Paillasseur, D. Caillaud et al., "Clinical COPD phenotypes: a novel approach using principal component and cluster analyses," *European Respiratory Journal*, vol. 36, no. 3, pp. 531–539, 2010.

- [12] J. Angus Webb, N. R. Bond, S. R. Wealands, R. M. Nally, G. P. Quinn, and P. A. Vesik, "Bayesian clustering with autoclass explicitly recognizes uncertainties in landscape classification," *Ecography*, vol. 30, no. 4, pp. 526–536, 2007.
- [13] M. Dash and H. Liu, "Feature selection for clustering," in *Knowledge Discovery and Data Mining. Current Issues and New Applications*, T. Terano, H. Liu, and A. L. P. Chen, Eds., vol. 1805 of *Lecture Notes in Computer Science*, pp. 110–121, Springer, Berlin, Germany, 2000.
- [14] M. A. Hall, *Correlation-based feature selection for machine learning [Ph.D. thesis]*, University of Waikato, Hamilton, New Zealand, 1999.
- [15] Z. Zheng, X. Wu, and R. Srihari, "Feature selection for text categorization on imbalanced data," *ACM SIGKDD Explorations Newsletter*, vol. 6, no. 1, pp. 80–89, 2004.
- [16] F. Sebastiani, "Machine learning in automated text categorization," *ACM Computing Surveys*, vol. 34, no. 1, pp. 1–47, 2002.
- [17] Y. Liu and M. Schumann, "Data mining feature selection for credit scoring models," *Journal of the Operational Research Society*, vol. 56, no. 9, pp. 1099–1108, 2005.
- [18] M. Dash, H. Liu, and H. Motoda, "Consistency based feature selection," in *Knowledge Discovery and Data Mining. Current Issues and New Applications*, T. Terano, H. Liu, and A. L. P. Chen, Eds., vol. 1805, pp. 98–109, Springer, Berlin, Germany, 2000.
- [19] E. Sarhrouni, A. Hammouch, and D. Aboutajdine, "Application of symmetric uncertainty and mutual information to dimensionality reduction and classification of hyperspectral images," *International Journal of Engineering and Technology*, vol. 4, no. 5, pp. 268–276, 2012.
- [20] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial Intelligence*, vol. 97, no. 1-2, pp. 273–324, 1997.
- [21] D. J. Straczuzki and P. E. Utgoff, "Randomized variable elimination," *Journal of Machine Learning Research*, vol. 5, pp. 1331–1364, 2004.
- [22] I. Inza, P. Larrañaga, R. Etxeberria, and B. Sierra, "Feature Subset Selection by Bayesian network-based optimization," *Artificial Intelligence*, vol. 123, no. 1-2, pp. 157–184, 2000.
- [23] L. Brieman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*, Wadsworth Inc., 1984.
- [24] H. Fu, Z. Xiao, E. Dellandrea, W. Dou, and L. Chen, "Image categorization using ESFS: a new embedded feature selection method based on SFS," in *Advanced Concepts for Intelligent Vision Systems*, J. Blanc-Talon, W. Philips, D. Popescu, and P. Scheunders, Eds., vol. 5807 of *Lecture Notes in Computer Science*, pp. 288–299, 2009.
- [25] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [26] S. Das, "Filters, wrappers and a boosting-based hybrid for feature selection," in *Proceedings of the 8th International Conference on Machine Learning*, pp. 74–81, 2001.
- [27] Y. Couce, L. Franco, D. Urda, J. L. Subirats, and J. M. Jerez, "Hybrid (Generalization-Correlation) method for feature selection in high dimensional DNA microarray prediction problems," in *Advances in Computational Intelligence*, I. Cabestany, I. Rojas, and G. Joya, Eds., vol. 6692 of *Lecture Notes in Computer Science*, pp. 202–209, 2011.
- [28] S. Chebrolu, A. Abraham, and J. P. Thomas, "Hybrid feature selection for modeling intrusion detection systems," in *Neural Information Processing*, N. Pal, N. Kasabov, R. Mudi, S. Pal, and S. Parui, Eds., vol. 3316 of *Lecture Notes in Computer Science*, pp. 1020–1025, 2004.
- [29] Y. Shen, X. Qiu, C. Zhang et al., "Quad-PRE: a hybrid method to predict protein quaternary structure attributes," *Computational and Mathematical Methods in Medicine*, vol. 2014, Article ID 715494, 9 pages, 2014.
- [30] M. Halkidi, Y. Batistakis, and M. Vazirgiannis, "On clustering validation techniques," *Journal of Intelligent Information Systems*, vol. 17, no. 2-3, pp. 107–145, 2001.
- [31] K. Wagstaff, C. Cardie, S. Rogers, and S. Schroedl, "Constrained K-means clustering with background knowledge," in *Proceedings of the 18th International Conference on Machine Learning*, pp. 577–584, 2001.
- [32] G. Forestier, P. Gançarski, and C. Wemmert, "Collaborative clustering with background knowledge," *Data and Knowledge Engineering*, vol. 69, no. 2, pp. 211–228, 2010.
- [33] M. Leng, G. Cheng, J. Wang et al., "Active semisupervised clustering algorithm with label propagation for imbalanced and multidensity datasets," *Mathematical Problems in Engineering*, vol. 2013, Article ID 641927, 10 pages, 2013.
- [34] M. Zhu, F. Meng, and Y. Zhou, "Semisupervised clustering for networks based on fast affinity propagation," *Mathematical Problems in Engineering*, vol. 2013, Article ID 385265, 13 pages, 2013.
- [35] E. Pitt and R. Nayal, "The use of various data mining and feature selection methods in the analysis of a population survey dataset," in *Proceedings of the 2nd International Workshop on Integrating Artificial Intelligence and Data Mining*, vol. 84, pp. 83–93, 2007.
- [36] H. Liu, J. Li, and L. Wong, "A comparative study on feature selection and classification methods using gene expression profiles and proteomic patterns," *Genome Informatics*, vol. 13, pp. 51–60, 2002.
- [37] L. Kaufman and P. Rousseeuw, "Clustering by means of medoids," in *Statistical Data Analysis Based on the L_1 -Norm and Related Methods*, Y. Dodge, Ed., pp. 405–416, North-Holland, 1987.
- [38] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, San Francisco, Calif, USA, 2012.
- [39] S. Boriah, V. Chandola, and V. Kumar, "Similarity measures for categorical data: a comparative evaluation," in *Proceedings of the 8th SIAM International Conference on Data Mining*, pp. 243–254, April 2008.
- [40] J. Gower, "A general coefficient of similarity and some of its properties," *Biometrics*, vol. 27, no. 4, pp. 857–871, 1971.
- [41] D. Chávez Esponda, I. Miranda Cabrera, M. Varela Nualles, and L. Fernández, "Utilización del análisis de clusters con variables mixtas en la selección de genotipos de maíz," *Revista Investigación Operacional*, vol. 30, no. 3, pp. 209–216, 2010.
- [42] U. Brandes, D. Delling, M. Gaertler et al., "On modularity clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 2, pp. 172–188, 2008.
- [43] O. Arbelaitz, I. Gurrutxaga, J. Muguerza, J. M. Pérez, and I. Perona, "An extensive comparative study of cluster validity indices," *Pattern Recognition*, vol. 46, no. 1, pp. 243–256, 2013.
- [44] G. Forestier, C. Wemmert, and P. Ganarski, "Background knowledge integration in clustering using purity indexes," in *Knowledge Science, Engineering and Management*, Y. Bi and M. A. Williams, Eds., vol. 6291 of *Lecture Notes in Computer Science*, pp. 28–38, 2010.
- [45] C. D. Manning, P. Raghavan, and H. Schütze, *An Introduction to Information Retrieval*, Cambridge University Press, Cambridge, UK, 2009.

- [46] L. Yu and H. Liu, "Feature selection for high-dimensional data: a fast correlation-based filter solution," in *Proceedings of the 20th International Conference on Machine Learning*, pp. 856–863, August 2003.
- [47] Z. Zhao and H. Liu, "Searching for interacting features," in *Proceeding of the 20th International Joint Conference on Artificial Intelligence (IJCAI '07)*, pp. 1156–1161, San Francisco, Calif, USA, January 2007.
- [48] A. M. Taha, A. Mustapha, and S. D. Chen, "Naive bayes-guided bat algorithm for feature selection," *The Scientific World Journal*, vol. 2013, Article ID 325973, 9 pages, 2013.
- [49] M. S. Uzer, N. Yilmaz, and O. Inan, "Feature selection method based on artificial bee colony algorithm and support vector machines for medical datasets classification," *The Scientific World Journal*, vol. 2013, Article ID 419187, 10 pages, 2013.
- [50] K. Dai, H. Yu, and Q. Li, "A semisupervised feature selection with support vector machine," *Journal of Applied Mathematics*, vol. 2013, Article ID 416320, 11 pages, 2013.

Research Article

Validation in Principal Components Analysis Applied to EEG Data

**João Carlos G. D. Costa, Paulo José G. Da-Silva,
Renan Moritz V. R. Almeida, and Antonio Fernando C. Infantosi**

Biomedical Engineering Program, COPPE, Federal University of Rio de Janeiro, P.O. Box 68510, 21941-972 Rio de Janeiro, RJ, Brazil

Correspondence should be addressed to Antonio Fernando C. Infantosi; afci@peb.ufrj.br

Received 1 May 2014; Revised 13 August 2014; Accepted 14 August 2014; Published 8 September 2014

Academic Editor: Ezequiel López-Rubio

Copyright © 2014 João Carlos G. D. Costa 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.

The well-known multivariate technique Principal Components Analysis (PCA) is usually applied to a sample, and so component scores are subjected to sampling variability. However, few studies address their stability, an important topic when the sample size is small. This work presents three validation procedures applied to PCA, based on confidence regions generated by a variant of a nonparametric bootstrap called the partial bootstrap: (i) the assessment of PC scores variability by the spread and overlapping of “confidence regions” plotted around these scores; (ii) the use of the confidence regions centroids as a validation set; and (iii) the definition of the number of nontrivial axes to be retained for analysis. The methods were applied to EEG data collected during a postural control protocol with twenty-four volunteers. Two axes were retained for analysis, with 91.6% of explained variance. Results showed that the area of the confidence regions provided useful insights on the variability of scores and suggested that some subjects were not distinguishable from others, which was not evident from the principal planes. In addition, potential outliers, initially suggested by an analysis of the first principal plane, could not be confirmed by the confidence regions.

1. Introduction

A large number of variables is frequently required in many research fields and, especially, in the biomedical sciences. One of the most used methods for studying patterns in such large databases is the Principal Components Analysis (PCA) [1, 2]. PCA is suitable for dimensionality reduction and for exploratory purposes, allowing for the extraction of data features through variance maximization. However, as in any statistical model, a validation procedure must be employed if generalizability is required. Such procedures are even more important when only a small number of subjects/objects are available [3, 4]. Important statistics usually obtained in PCA are eigenvalues and principal component (PC) scores and, thus, nonparametric confidence intervals (C.I.) can be used to assess their variability. The latter can be, for example, generated by a resampling technique [4] and, then, computed as “confidence regions” around PC scores. Since the percent of explained variance is different according to each PCA dimension, corresponding C.I. are also different, helping outlier

identification (longer intervals suggest extreme observations) [4].

One of these resampling techniques is the nonparametric bootstrap, in which samples are drawn with replacement in order to mimic the empirical probability function of the data [5]. Although visual cluttering may result, the bootstrap (BST) can be employed for defining confidence regions in PCA, thus helping graphical display interpretation [4]. However, few texts address the subject of PCA confidence regions derived from BST. One of them is Linting et al. [6], in which 90% BST ellipses were drawn for a nonlinear PCA used to study interactions between children and caregivers in non-maternal child care. By comparing the results with those from a linear PCA, they suggested a guideline for users who wish to employ the BST procedure in linear and nonlinear PCA.

Classification of electroencephalographic (EEG) signals is an objective of many neurological studies, for example, for staging a neurologic disease or for brain-computer interface (BCI) systems. These systems, briefly, concern the transformation of human thoughts (through acquired EEG signals)

into a computer system, for instance, for helping people with motor or spelling impairments during specific tasks [7]. The electroencephalogram is the registry of a spatial-temporal cortical activity recorded from electrodes spatially placed on the scalp region and is mainly characterized by signals with different frequency bands, such as theta (4–8 Hz), alpha (8–13 Hz), and beta (13–30 Hz), and amplitude varying with pathological conditions and in specific behavior states (e.g., sleep or vigil, eyes open or closed) [8, 9]. Although online classification tasks are a prerequisite for practical BCI purposes, extensive offline studies are needed before establishing a trustworthy BCI device, hence indicating the importance of validation procedures.

The aim of this paper is to present three validation procedures for PCA using the nonparametric bootstrap, with an application to EEG data. These procedures allow for assessing the sampling variability of PC scores and the number of axes to be retained for analysis, especially if only a small number of subjects are available. The method concerns plotting “confidence regions” and constructing a “validation set” for PC scores (the centroids of the confidence regions). A variant of the ordinary nonparametric bootstrap called the partial bootstrap (PBST) was used to this end. Furthermore, a validation procedure was employed in order to confirm the number of nontrivial axes to be analyzed. The assessment of sampling variability of the PC scores was performed through the areas of the confidence regions, while the centroids were compared to the original scores through an unsupervised classification algorithm. An example with correlated attributes derived from time and frequency-domain EEG signals was used for introducing the proposed approach. The theory of PCA and nonparametric BST is introduced in Sections 2 and 3, respectively, while the validation methods are presented in Section 4. In Section 5, the method is applied to EEG data obtained from a postural control protocol.

2. Principal Components Analysis

Principal Components Analysis is comprehensively presented in many multivariate statistics textbooks, such as Jolliffe [2] and Lebart et al. [10], and only a brief introduction is given here. From p variables observed on n objects (an $n \times p$ matrix), that is, a raw data matrix \mathbf{X} , PCA derives new variables as linear combinations of the original ones, defined from a new orthogonal coordinate system onto which the original space is projected. This new system summarizes the total data variation in decreasing order so that the first new variable has the largest variation, the second has the second largest, and so on. These new variables are the principal components. The singular value decomposition (SVD) is used to estimate this new orthogonal space, by factoring \mathbf{X} as [11]

$$\mathbf{X} = \mathbf{U} \times \mathbf{D} \times \mathbf{V}^T, \quad (1)$$

where \mathbf{U} and \mathbf{V} are the left and right singular vectors matrices, respectively, $\mathbf{U}\mathbf{U}^T = \mathbf{I}_n (n \times n)$, $\mathbf{V}^T\mathbf{V} = \mathbf{I}_p (p \times p)$, and the superscript \mathbf{T} indicates the transpose of the matrix. \mathbf{D} is a diagonal matrix with singular values λ_i in decreasing order

$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, $i = 1, 2, \dots, p$. Squaring \mathbf{D} and dividing it by $(n - 1)$, one obtains

$$\text{cor}(\mathbf{X}) = \mathbf{V} \times \mathbf{\Sigma} \times \mathbf{V}^T, \quad (2)$$

where $\text{cor}(\mathbf{X})$ is the sample correlation matrix if \mathbf{X} has standardized variables. Matrix $\mathbf{\Sigma}$ is also a diagonal matrix, with elements related to the variance of \mathbf{X} , in which $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_p^2 \geq 0$, so that $(n - 1) \sigma_j^2 = \lambda_j^2$, $j = 1, 2, \dots, p$.

The PC scores (\mathbf{Z}) are obtained as follows:

$$\mathbf{Z} = \mathbf{X} \times \mathbf{V} = \mathbf{U} \times \mathbf{D}. \quad (3)$$

Therefore, elements of \mathbf{Z} are linear combinations of the elements of \mathbf{X} , with component coefficients given by the column-vector of \mathbf{V} (which is called the *loadings* matrix). The matrix \mathbf{V} defines, thereby, an orthonormal basis and its columns are linearly independent vectors. Indeed, \mathbf{Z} is the projection of \mathbf{X} onto the orthonormal basis \mathbf{V} .

As mentioned, of the most common uses of PCA is the dimensionality reduction of \mathbf{X} , keeping as much information (variance) as possible. If the reduced dimensionality is $m \leq p$, one may consider the model (in elementwise representation):

$$x_{ij} = \sum_{t=1}^m u_{it} \lambda_t v_{tj} + \varepsilon_{ij}, \quad (4)$$

where u_{it} and v_{tj} are the elements of matrices \mathbf{U} and \mathbf{V} , respectively, while ε_{ij} represents the residual terms or the noise present in the data, $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, p$. The proportion of variance explained by each dimension up to dimension m is given by

$$\text{var}\% = \frac{\sum_{j=1}^m \lambda_j}{\sum_{j=1}^p \lambda_j} \times 100\%. \quad (5)$$

The procedure for choosing the number of principal components m to be retained is not well-defined. One method is the Scree plot, based on a plot of eigenvalues against their order [12, 13]. Some authors suggest other empirical methods such as the retention of a number of dimensions corresponding to a fixed proportion of explained variance (usually 70–90%) and the Kaiser’s rule (retaining the eigenvalues of the correlation matrix higher than unity) [2, page 113–115] and [12, 13]. Statistical approaches have also been proposed, such as the Bartlett’s test or eigenvalues bootstrapping [13]. However, as Jolliffe [2, page 133] pointed out, there is still no clear advantage of a specific method over the others.

3. Nonparametric Bootstrap

The nonparametric BST is a computer-intensive technique, which attempts to replicate the probability distribution of a statistic of interest by resampling with replacement from the original sample (the observed data) a predefined (R) number of times [5]. Usually, this procedure generates new samples of the same size n of the original one, providing a mathematical framework for inferring the statistical accuracy of the desired

estimate [14]. Thus, in summary, the statistic of interest ($\tilde{\theta}$) is the observed value of some unknown population parameter θ , and the nonparametric BST generates R replicated samples of the original data (the BST samples), resulting in the set $\tilde{\theta}_{\text{set}}^* = \{\tilde{\theta}_1^*, \tilde{\theta}_2^*, \dots, \tilde{\theta}_R^*\}$. If the observed data is independent and identically distributed, the BST estimate of the observed value ($\tilde{\theta}^*$) can be calculated from $\tilde{\theta}_{\text{set}}^*$ as

$$\tilde{\theta}^* = \frac{\sum_{r=1}^R \theta_b^*}{R}, \quad (6)$$

implying that $\tilde{\theta}^*$ is an estimate of the true value θ . The accuracy of the BST estimates can be represented by confidence intervals (C.I.) calculated from $\tilde{\theta}_{\text{set}}^*$. The percentile method [14] is the simplest method for BST C.I. and is based on the percentiles of $\tilde{\theta}_{\text{set}}^*$, as

$$\text{C.I.} = \left[\tilde{\theta}_{\text{set}}^* \left(\frac{\alpha}{2} \right), \tilde{\theta}_{\text{set}}^* \left(1 - \frac{\alpha}{2} \right) \right], \quad (7)$$

where $\tilde{\theta}_{\text{set}}^*(\alpha/2)$ and $\tilde{\theta}_{\text{set}}^*(1-\alpha/2)$ are the $100\alpha/2\%$ and $100(1-\alpha/2)\%$ percentiles of the $\tilde{\theta}_{\text{set}}^*$ and α is the desired confidence level. For example, for $\alpha = 0.05$ and $R = 1000$, the C.I. inferior/superior limits are the 24th/976th elements of $\tilde{\theta}_{\text{set}}^*$. In general, it is advisable to have a large R [14], and, for PCA, Lebart [4] advocates $R \geq 30$, while Diaconis and Efron [15] employed 100 replications in a PCA study for grading college students.

4. Validation and Stability

The performance of a model is always better on data on which the model was estimated, and this rule applies for both exploratory and predictive methods [3, 16]. In order to evaluate the results obtained by the whole (or part of) observed sample (the training set), the model can be applied to a different data (the validation set), assessing its generalization ability [17, 18]. The procedure of applying the obtained model to new data is usually called *validation*. Therefore, modeling demands rigorous validation procedures, since a good model is supposed to have generalizability [16]. Basically, there are three kinds of validation: internal, external, and relative, the first being most commonly used due to its simplicity and lower costs. In the internal validation, the observed data can be split in two or more sets (such as cross-validation) or BST methods (one for training and the others for validating the model), while in the external validation, a new but plausible dataset is presented to the model. For the relative validation, a different model is applied to the available data. When the training set has a small number of subjects, BST becomes a good option for internal validation, since all subjects can be used for model development (no observation is discarded), and BST samples can be used for validation. Furthermore, the generalization concept described above can be connected to the concept of stability in PCA, because if the score coordinates do not change markedly, their positions onto principal planes can also be considered stable.

Applying nonparametric BST to, for instance, the above defined matrix $\mathbf{X}(n \times p)$, different matrices may be generated

by the replication of different rows, and their singular values and singular vectors will no longer be the same. The SVD applied to each of the R BST matrices is

$$\mathbf{X}_r^* = \mathbf{U}_r^* \times \mathbf{D}_r^* \times \mathbf{V}_r^{*\text{T}}, \quad (8)$$

where $*$ denotes a BST sample and $r = 1, 2, \dots, R$. \mathbf{D}_r^* has singular values in decreasing order $\lambda_{1r}^* \geq \lambda_{2r}^* \geq \dots \geq \lambda_{pr}^* \geq 0$, and $\Sigma_r^* = \mathbf{D}_r^{*2}/(n-1)$. Through these concepts, validation procedures using nonparametric BST can be applied to PCA.

4.1. Assessing the Number of Nontrivial Axes. If the chosen dimensionality is $m^* < p$, no overlapping between BST eigenvalues for ($\alpha = 0$) will occur if

$$\begin{aligned} \min(\sigma_1^*) - \max(\sigma_2^*) &> 0, \\ \min(\sigma_2^*) - \max(\sigma_3^*) &> 0, \\ &\vdots \\ \min(\sigma_{m^*-1}^*) - \max(\sigma_{m^*}^*) &> 0, \end{aligned} \quad (9)$$

where $\sigma_k^* = \{\sigma_{k1}^*, \sigma_{k2}^*, \dots, \sigma_{kR}^*\}$, for $k = 1, 2, \dots, m^*$; thus the number m^* obtained from BST can be compared to the number m obtained by Scree plot.

4.2. Assessing the Variability of PC Scores. After the application of BST to \mathbf{X} , R replicated matrices are obtained. Due to their different axes (defined by different eigenvectors), they cannot be directly compared to the original space (defined by the eigenvectors of the original correlation matrix), because of axes reflection or inversion [4, 19]. Since replicated samples do not have necessarily the same subjects compared to those in the original sample, different eigenvalues and eigenvectors can occur, and a correction procedure is needed, such as that provided by Procrustes Analysis [20]. To circumvent this problem, the PBST can be applied, consisting of projecting replicated components scores (as ‘‘supplementary’’ points) onto the orthonormal matrix \mathbf{V} :

$$\widehat{\mathbf{z}}_{ir} = \mathbf{x}_{ir}^* \times \mathbf{V}, \quad (10)$$

where $\widehat{\mathbf{z}}_{ir}$ is the i th component score of the r replicated, standardized object (\mathbf{x}_{ir}^*). Therefore, nR object scores can be visualized in the original space, generating n clouds of points. This approach has the advantage of maintaining the original PC planes, which is a better estimate than any of the replicated planes [4]. Thus, (10) can be expressed as

$$\widehat{\mathbf{Z}}_{\text{parc}} = \mathbf{X}^* \times \mathbf{V}, \quad (11)$$

where

$$\widehat{\mathbf{X}}^* = \begin{bmatrix} \mathbf{X}_1^* \\ \mathbf{X}_2^* \\ \vdots \\ \mathbf{X}_R^* \end{bmatrix}, \quad (12)$$

and \mathbf{V} is limited to m dimensions ($m \times m$) after the dimensionality reduction procedure is applied.

Since PCA displays are usually shown in a low-dimensional space, confidence regions are represented as polytopes [21] or, in a two-dimensional space, as polygons (or convex hulls) [22]. The interpretation of these polygons basically takes into account overlapping (which suggests similar objects) and spread (widespread polygons suggest unstable score coordinates, while narrow polygons suggest stability). Furthermore, these confidence regions allow for the estimation of new PC scores (through their centroids). Although any value of α can be used, Efron [23] states that $\alpha = 0.10$ is satisfactory in most cases, while Lebart [4] pointed out that when $\alpha = 0$ untypical values (e.g., outliers) can be easily identified (through the longer edges of the plotted convex hull).

4.3. Validation of PC Score Coordinates. The area (in square units) of a polygon can be calculated as

$$S = \frac{1}{2} \left[\left| \begin{array}{cc} x_1 & x_2 \\ y_1 & y_2 \end{array} \right| + \left| \begin{array}{cc} x_2 & x_3 \\ y_2 & y_3 \end{array} \right| + \dots + \left| \begin{array}{cc} x_l & x_1 \\ y_l & y_1 \end{array} \right| \right], \quad (13)$$

where $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)$ are the l vertices' coordinates of the polygon, in clockwise order, and $|\cdot|$ is the determinant of the matrix. Absolute value can be calculated, if necessary, and the centroid coordinates (\bar{x}_c, \bar{y}_c) calculated from any polygon are given by

$$\begin{aligned} \bar{x}_c &= \frac{1}{6S} \cdot \sum_{i=1}^l (x_{i-1} + x_i) (x_{i-1}y_i - x_iy_{i-1}), \\ \bar{y}_c &= \frac{1}{6S} \cdot \sum_{i=1}^l (y_{i-1} + y_i) (x_{i-1}y_i - x_iy_{i-1}). \end{aligned} \quad (14)$$

Thus, the centroid can be considered as the BST estimate (BST centroids) of the true component score.

The BST centroids are, therefore, estimates of the PC score coordinates, and the comparison of original scores and BST centroids allows for the comparison of both models, using, for example, an unsupervised classification method. These clustering methods concern procedures where the groups are not known a priori and the researcher must choose, based on previous knowledge or on some criteria, the number of clusters present in the data. This subjective procedure is mainly employed to visualize or suggest clusters, generating hypothesis for later investigation [24].

One kind of unsupervised classification method is the hierarchical algorithm, in which a nested-tree diagram (the dendrogram) is generated, suggesting, by inspection, the underlying clustering structure of the data. There are, basically, two kinds of hierarchical classification algorithms, the divisive and the agglomerative, which group objects according to some clustering rule [25]. Agglomerative Hierarchical Algorithms (AHA) are some of the most used classification algorithms and start by grouping two objects into a single cluster, and at each step of the algorithm, new objects are aggregated, forming a new cluster, and so on, until, in the last step, all objects are joined into a single cluster. "Cutting the tree" at some distance is one of the procedures for defining the

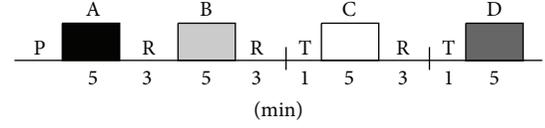


FIGURE 1: Complete experimental protocol sequence for the acquisition of EEG signals in 31 subjects (21 males and 10 females). P indicates the preparation procedure; R refers to the resting interval of three minutes; T indicates the transition between sat and upright standing positions.

cluster structure in AHA [24, Section 3.3.2], and the Average Linkage Algorithm is considered the most stable AHA [26]. Therefore, an AHA using this method can be used to compare groups generated by ordinary PC scores and BST centroids.

5. Application

5.1. Subjects. A data set from a postural control protocol was used in this study, including stabilometric and EEG signals. Thirty-one subjects (21 males and 10 females), ages 21 to 45 (31.0 ± 6.6) years, height 154 to 187 (172.7 ± 9.4) cm, and body weight 46 to 107 (73.3 ± 17.3) kg, participated in the initial study. All subjects presented no history of neurological pathologies, osseous, muscles or joints diseases, or equilibrium disorders. An anamnesis was performed to obtain information about headaches, illnesses, vertigo, eyestrain, and the use of contact lens or glasses. Subjects using lens or glasses were included when no problem with their use was reported. The study was approved by a Local Institutional Review Board (IESC/UFRJ - Ref. 100/2011). None of the authors participated as a volunteer.

5.2. Experimental Protocol. The EEG and stabilometric signals were acquired simultaneously, but only EEG signals were analyzed here. The experiments were performed in an electromagnetically shielded room, under controlled environmental conditions (23°C , attenuated sound and light control), with the subject barefooted on a force platform. The feet position (angle: 30° ; heels 2 cm apart) was previously delineated to standardize the same support base during the tests. The EEG signals were acquired during five minutes, with the subject in distinct postural conditions: (i) resting in a comfortable armchair with eyes closed (spontaneous EEG with room lights off, denoted as "A"); (ii) the same position as (i), but with eyes open ("B"); (iii) during stabilometric test in upright standing position with eyes open (denoted as "C"); and (iv) eyes closed ("D"). The trials with eyes open condition were conducted with room lights on and with the subject watching a white wall located 1 meter apart from the force platform. An interval of three minutes was taken between each condition, and the subject remained seated in the chair during this period. The stabilometric tests were performed one minute after the subject was standing on the force platform, in order to allow for the recovery of balance after rising from the chair. Figure 1 shows the experimental protocol sequence.

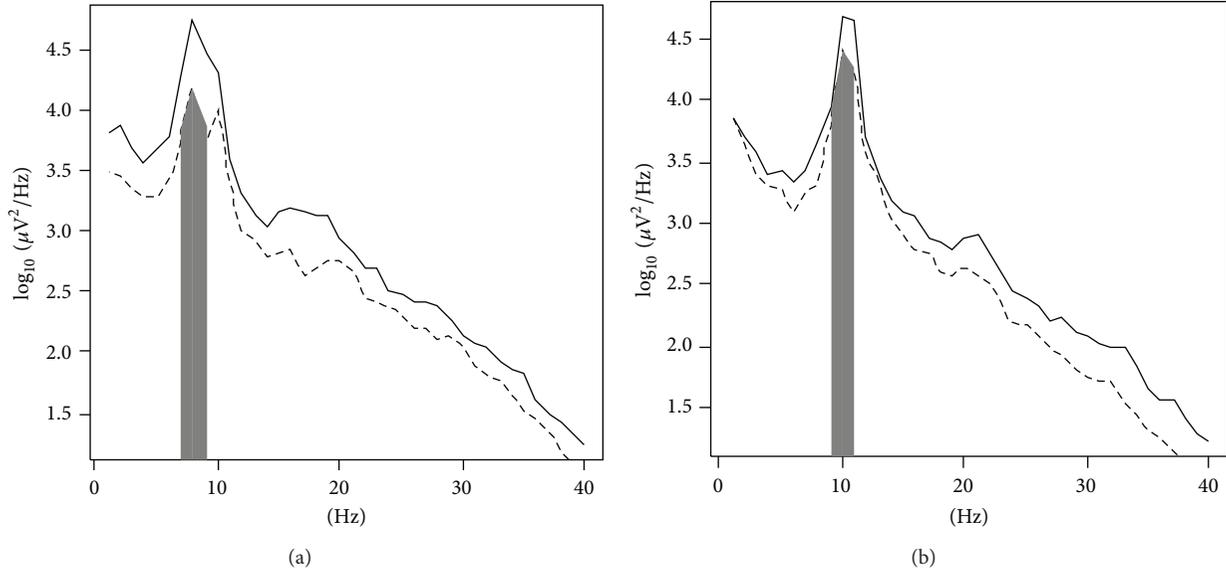


FIGURE 2: PSD of volunteers 1 (a) and 3 (b), solid line for condition A, and dashed line for condition B. (a) Maximum for condition A at 4.75 and maximum for condition B at 4.17. (b) Maximum for condition A at 4.68 and maximum for condition B at 4.40. Power for condition B was calculated as the area of the highlighted region in grey.

The EEG recordings were acquired using the BrainNet—BNT 36 (EMSA, Brazil, <http://www.emsamed.com.br>) device at a sampling frequency of 400 Hz and 16 bit A-to-D precision, with electrode position according to the International 10/20 System (monopolar derivations, averaged bilateral earlobe reference and ground in FPz). Scalp electrode impedances were below 5 k Ω throughout the session. The EEG recordings were analog-filtered by a fourth-order low-pass Butterworth with cutoff frequency at 100 Hz (antialiasing) and second-order high-pass Butterworth at 0.1 Hz and also by a digital notch filter in 60 Hz. The power spectral densities (PSDs) were determined by an additional offline digital filtering using a fourth-order, forward-reverse band-pass (1–40 Hz), Butterworth filter [27].

5.3. Data. The complete data set consisted of 5-minute EEG recordings (O1, O2, P3, P4, C3, C4, T3, T4, T5, T6, F3, F4, F7, F8, Fp1, Fp2, Fz, Cz, Pz, and Oz derivations) for the conditions A, B, C, and D. In order to reduce display cluttering, only one occipital derivation (O1) was used in this study. The O1 EEG signals were first segmented into 1-second zeroed-mean epochs of 400 samples. An artifacts rejection methodology proposed by Simpson et al. [28] was also applied, resulting in a different number of epochs for each volunteer and condition (min = 20, max = 300). To allow for better precision of estimates and for computational convenience (all epochs were stored in an array) only those volunteers with a minimum of 150 free artifacts epochs were retained in the study (24 subjects).

5.4. Variables in the Frequency and Time Domain. A rectangular window was subsequently applied to each epoch, and the averaged periodogram was calculated by the Bartlett method. Six variables were extracted from the periodogram:

maximum of the PSD magnitude of alpha (8–13 Hz), theta (4–8 Hz), and beta (13–30 Hz) bands in \log_{10} (micro V^2 /Hz) and an estimate of alpha, theta, and beta band power, defined as the trapezoidal area centered in the maximum peak of respective bands \log_{10} (micro V^2).

For each epoch, four statistics were estimated: the root mean square (RMS), the difference between maximal positive and minimal negative values (Mm); the standard deviation of the samples (SD); and the skewness coefficient [29]. Then, the median of each statistic for all 150 epochs was determined.

5.5. Variable Statistics. Mean \pm standard deviation for frequency domain variables was 6.75 ± 1.26 (alpha band power, range 4.28–9.12), 5.44 ± 0.73 (beta band power, range 4.11–6.82), 6.12 ± 0.96 (theta band power, range 4.54–9.12), 3.52 ± 0.68 (alpha band maximum, range 2.24–4.75), 2.72 ± 0.36 (beta band peak, range 2.03–3.37), and 3.07 ± 0.50 (theta band maximum, range 2.27–4.75). The Shapiro-Wilk test suggested that the alpha and beta band power (\log_{10}) and the alpha and beta band maximum (\log_{10}) variables were Gaussian. In Figure 2, PSDs of two volunteers (conditions A and B) are shown, with the areas corresponding to power at distinct bands highlighted. For volunteer 1, maximum peak for eyes open was achieved at 8 Hz, a transition frequency between theta and alpha bands (Figure 2(a)), while for volunteer 3, the maximum peak in the same condition occurred at 10 Hz (Figure 2(b)). For time domain variables, values were 12 ± 6 (RMS, range 4–30), 12 ± 6 (SD, range 4–30), 0.1 ± 0.1 (skewness, range -0.1 – $+0.3$), and 57 ± 26 (Mm, range 22–137). There were positive and significant ($P \ll 0.001$) correlations (rho's Spearman coefficient, range 0.62–1.00) between all variables, with perfect correlation (rho = 1.00) between alpha power and alpha maximum and between RMS and SD (redundant variables).

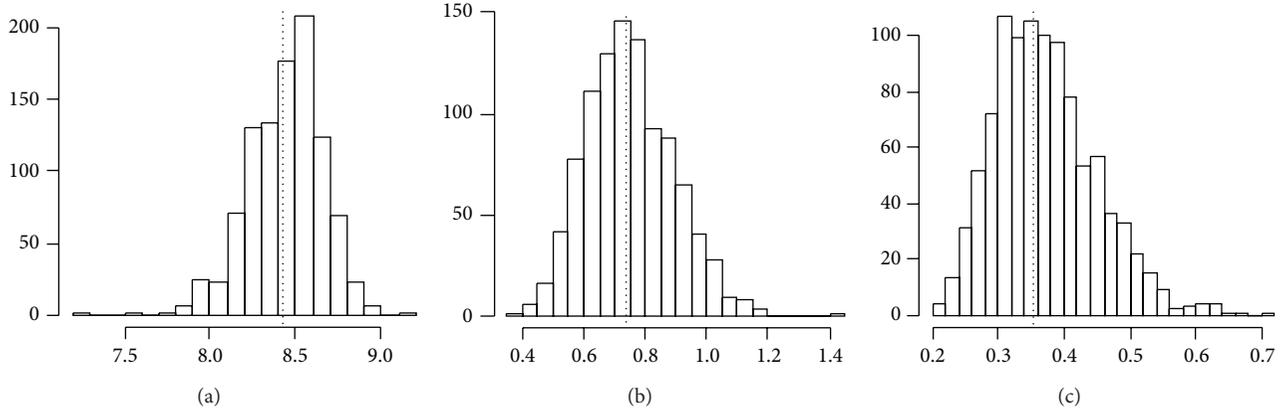


FIGURE 3: Histogram of replicated eigenvalues. (a) First eigenvalues; (b) second eigenvalues; (c) third eigenvalues. An overlapping between the 100% C.I. of second and third eigenvalues can be seen.

5.6. PCA and Nonparametric BST. Considering the 24 subjects to whom the experimental protocol was applied (conditions A and B) and the extracted variables from the EEG signal (six in frequency and four in time domains), the resulting data matrix (\mathbf{X}) with 48 rows and 10 columns (variables) was constructed. The SVD algorithm was applied to the zeroed-mean, standardized data matrix, and the \mathbf{U} , \mathbf{D} , \mathbf{V} , and \mathbf{Z} matrices were calculated, according to (1) and (3). The number of axes to be retained was assessed by the Scree plot and validated through an analysis of the nonoverlapping confidence intervals ($\alpha = 0$) of the replicated eigenvalues.

The nonparametric BST was performed according to the following steps.

- (1) Resampling the rows (of \mathbf{X}) with replacement, with $R = 1000$, which resulted in R replicated matrices 48×10 .
- (2) The R matrices were mean-centered and standardized.
- (3) The SVD algorithm was applied to all R standardized matrices.
- (4) $\mathbf{Z}_{\text{par}} = \mathbf{X}^* \times \mathbf{V}$ was obtained by (11).
- (5) The extreme points of all 48 clouds were determined.
- (6) The convex hulls (confidence polygons) were plotted around each object, with $\alpha = 0$ to analyse likely outliers.
- (7) The BST centroids and the areas of confidence polygons were calculated according to (14) and (13), respectively.

5.7. Validation. Validation was carried out as described in Section 4. To assess the variability of the PC scores, the areas of their corresponding convex hulls were compared, while BST centroids were compared to original scores by the dendrograms originated from an AHA, average method. The number of retained axes was assessed by the 100% C.I. obtained by a nonparametric BST.

The significance level adopted was 5% and data processing used the open access R statistical software [30], packages

TABLE 1: Loadings matrix showing the coefficients for the first two PCs in a PCA of 10 variables, relating to EEG signals from 24 volunteers.

Variables	1° PC	2° PC
Alpha power	-0.33	0.10
Beta power	-0.29	0.57
Theta power	-0.31	-0.38
Alpha max.	-0.32	0.10
Beta max.	-0.30	0.51
Theta max.	-0.30	-0.43
RMS	-0.33	-0.09
SD	-0.33	-0.09
Skewness	-0.29	-0.21
Mm	-0.34	-0.06

R.matlab [31], *signal* [32], *e1071* [33], and *pracma* [34]. Convex hulls and confidence polygons are terms interchangeably used in this text. Spearman correlation coefficient and Shapiro-Wilk tests were applied to verify correlation and Gaussianity, respectively.

5.8. Results. The dimensionality suggested by the Scree plot was two, corresponding to 91.6% of the explained variance (1st eigenvalue: 8.42; 2nd eigenvalue: 0.74). The coefficients for these two PCs are shown in Table 1. The dimensionality analysis was also confirmed by the C.I. of replicated eigenvalues (Figure 3).

Since the first PC is a linear combination with almost equal weights, none of these variables can be said to be “more influential.” Therefore, in this component, signal scores contrast only in relation to the origin. The second PC, however, shows a contrast between the beta and theta bands.

The histograms for the first three replicated eigenvalues are shown in Figure 3, with original eigenvalues in dashed lines. No overlapping between the first and two replicated eigenvalues occurred, since $\min(\sigma_1^* = 7.30) > \max(\sigma_2^* = 1.40)$. On the contrary, overlapping is present in the C.I. of the second and third replicated eigenvalues (Figures 3(b)

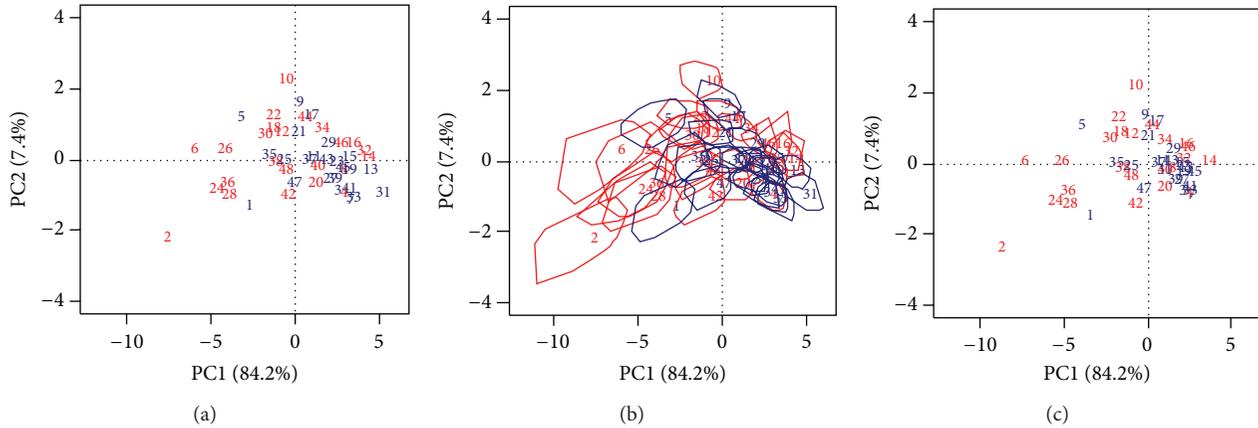


FIGURE 4: (a) Principal plane for signal scores, numbers from 1 to 48; (b) PC scores surrounded by their corresponding convex hulls; (c) BST-generated centroids. Signals corresponding to A and B conditions in red and blue, respectively.

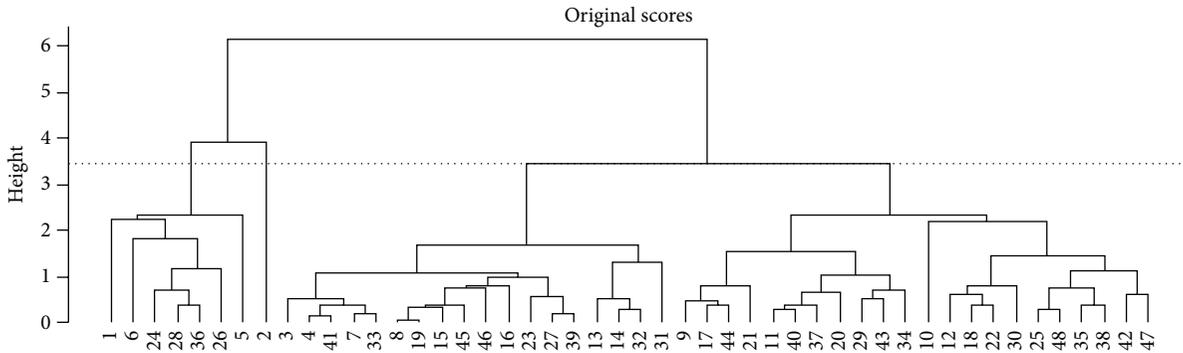


FIGURE 5: Dendrogram for average AHA; original PC scores as input. Two clusters and an outlier are suggested at height = 3.47.

and 3(c)). Explained variances in replicated samples, in two dimensions, varied from 86.1% to 95.0% (mean = 91.9%).

The projection of all replicated matrices in the original orthonormal basis provided 48000 points, which were synthesized by the convex hulls encompassing all 1000 replicated samples for each score. The number of replications varied between 931 (subject 17) and 1077 (subject 24), with mean = 1000. The principal plane is depicted in Figure 4(a), where the first PC has 84.4% of explained variance, while, in Figure 4(b), convex hulls are drawn around the original coordinates of the 48 scores. Extended overlapping convex hulls suggest signals with similar characteristics; therefore, their BST centroids are closer in the display. The area of the convex hulls varied between 0.96 and 7.79 (mean = 2.25). The two largest areas correspond to signals 6 (7.79) and 2 (7.37), while the smallest areas correspond to signals 29 (0.96) and 23 (1.02). Since these areas are located on opposite sides of the first PC, it can be said that this PC also discriminates between larger and smaller areas.

The areas of signals number 2 and number 6 also deserve attention. The convex hull corresponding to the latter is placed onto the second and third quadrants, while the former is placed on the third one. Signal 2 (female; Figure 3(a); solid line) had the highest and coincident measures for alpha (9.12) and theta (9.12) power and for alpha (4.75) and theta (4.75)

maximum (maximum peak was at 8 Hz in the transition frequency between alpha and theta bands). Also, this subject had the highest RMS (30.1), SD (30.1), skewness (0.31), and Mm (136.7). Signal 6 (male; Figure 3(b); solid line) had the highest values of beta power (6.82) and maximum beta (3.37). An analysis on overlapping polygons revealed a similarity for other signals, for each side of the first PC. The BST centroids are shown in Figure 4(c).

Scores from the original first principal plane were used as inputs for the AHA. According to the chosen separation distance (height), it was possible to identify (at least) two clusters, one with signals 1, 5, 6, 24, 26, 28, and 36 and another with all other signals (Figure 5). Signal number 2 was merged at the highest separated height, suggesting that this signal is an outlier. Figure 6 shows the Dendrogram obtained from the BST centroids, in which signal numbers 2 and 6 could be considered as another cluster.

6. Discussion

Validation is an important step in any statistical model and PCA is not an exception to this rule [3]. In PCA, distances between scores in a sample cannot be supposed to represent unbiasedly the true distances, especially if the sample size is small [35], and, in this context, an analysis of the sampling

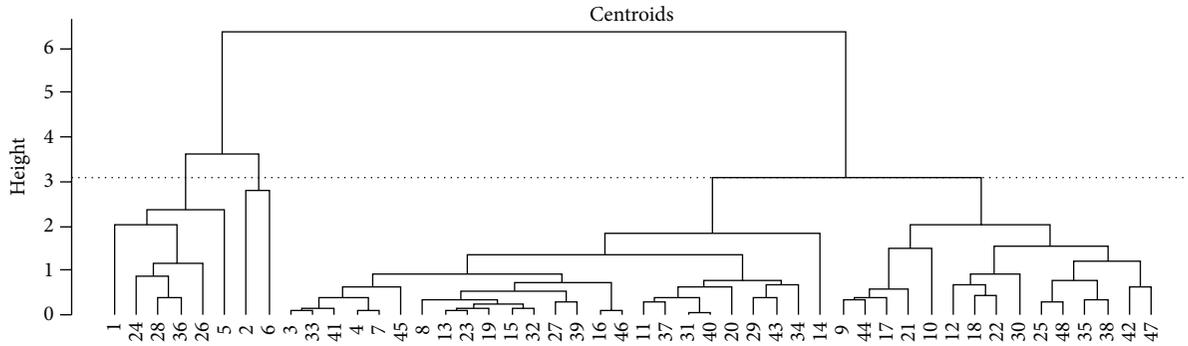


FIGURE 6: Dendrogram for average AHA; BST centroids as input. Three clusters are suggested at height = 3.07.

variability of the scores is important, allowing for smaller inferential errors.

As predicted by Efron and Tibshirani [14], increased processing power made the BST an important tool in statistical inference and model validation [36, 37]. However, when BST is applied to models that incorporate the SVD algorithm, care must be taken, due to reflection, stretching, or rotation of the principal components [4, 19]. To overcome these problems, this paper used “supplementary points” projected onto the space spanned by the principal components through a Partial BST, thus avoiding techniques such as the Procrustes Analysis simplifying the validation analysis [4]. Hence, the nonparametric BST is a reliable tool for result validation, as long as one takes into account the mentioned problems of different vector spaces generated by BST.

As mentioned, the C.I. of the replicated eigenvalues indicated that two PCs had to be retained. Indeed, an overlapping between the second up to the fifth replicated eigenvalues was present (not shown). The studied dataset had two highly positively correlated variables, a common feature when analyzing spectral power and time-domain EEG, and high correlation between variables usually results in a small number of axes to be retained (typically two). Thus, confidence regions were drawn in two dimensions, but, for dimensionalities larger than two, confidence polygons could be easily plotted using the described technique. Additionally, it should be noted that the first PC represented an overall average, a very common situation in biological data.

Agglomerative Hierarchical Algorithms are one of the unsupervised methods most used in classification studies [10]. Since PCA is sensitive to outliers, an initial analysis of the first principal plane together with the Dendrogram built from the original scores (Figures 4(a) and 5) would suggest that signal 2 was an outlier. However, this interpretation was not confirmed, since the confidence polygons of signals 2, 24, 28, and 36 had overlapping regions corresponding to condition A, thus suggesting a similarity among them (Figure 4(b)). Furthermore, when the BST centroids were used as input data, signal 2 was first merged with signal 6, suggesting a different group (Figure 6). This feature was not detected when the original scores were used as input data for the classifier.

As mentioned, PCA is widely used in biomedical signal analysis. For example, Casarotto et al. [38] employed PCA for reducing ocular artifacts in event-related potentials (in

39 children) by subtracting the principal component related to the electrooculogram (EOG) from the raw EEG. Since EOG is always present and has amplitude similar to EEG, reducing artifacts from this source is very important, and the authors concluded that the approach allowed for an efficient reduction of ocular artifacts. Kobayashi and Kuriki [39] employed PCA to increase the signal-to-noise ratio (SNR) in evoked neuromagnetic signals applied to four male subjects. The raw spontaneous neuromagnetic fields were recorded by a superconducting quantum interference device (SQUID) system and superposed to simulated evoked fields to mimic real signals. The authors retained three PCs for analysis and concluded that the suppression of the first PC improved the SNR compared to the common averaging method. Also Dafertshofer et al. [40], analyzing six electromyographic (EMG) signals from thoracic and lumbar muscles, obtained during a treadmill walking experiment, found that the first two PCs accounted for 88% of the data variance. The second PC suggested a contrast between right and left thoracic muscles, while the first PC represented an overall average. Analysis of gait kinematic data in stroke patients was performed in twenty-seven subjects by Milovanovic and Popović [41] who found differences between patients and healthy subjects by PCA. In that study, the authors retained the first two principal components and concluded that the first PC is related to severity of hemiplegia.

However, none of the studies above included a discussion about the generalization potential of their results. The methodology described here would be very useful to this end, owing to the small number of subjects in many of these studies. Furthermore, since PC scores can be used as input data in classification algorithms for BCI purposes, this assessment is especially important for avoiding inaccurate analysis in the training dataset.

The results suggested two and three main clusters for the analyzed dataset, mainly due to the importance of the first PC. As it is well-known, the occipital area (O1 and O2 derivations) is recognized as a visual area in the human cortex, while the parietal area (P3 and P4) is known to be part of the associative cortex, which corresponds to the sensory-motor integration within postural control. When individuals are in standing up position (orthostatic posture), especially in the eyes closed condition, other sensory (vestibular and proprioceptors) systems play an important role in balance,

increasing the activity in other cortex regions [42]. Therefore, we analyzed only stable postural conditions in which the volunteers were kept in “sat” position, to minimize the influence of other EEG derivations. Usually the cortical activity during balance perturbation is investigated in time domain by the coherence average method, to analyze the latency of the evoked potential after stimulation onset [43]. However, the evoked potential evaluation is not an automatic process of stimuli response identification, and objective response detection (ORD) techniques in frequency domain, such as the spectral F -test (SFT) and the event-related desynchronization/synchronization index (ERD/ERS), have been used to this end [44, 45].

In summary, this paper showed how BST methods can be applied to validate the most important PCA results, what is particularly relevant in small data sets, a common feature in EEG studies. One of the presented methods is a new procedure, which consists in estimating new PC scores as centroids of confidence regions calculated by a PBST of the original data (the BST centroids) and in using these centroids as a validation set. A comparison was performed on two Agglomerative Hierarchical Algorithms, one with the original and the other using the estimated component scores as inputs, and the estimated scores allowed for the detection of a cluster not discovered by the original scores. Furthermore, the confidence regions were able to help result interpretation, for instance, by the analysis of their overlap. As discussed, in this case, the area of a polygon increases together with the variability of the PC, providing additional insights about the data, for instance, concerning outliers and remote observations in the multidimensional space [46]. Studies using more complex classification algorithms and data with dimensionality larger than two would be useful for further developing this work.

Conflict of Interests

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

Acknowledgments

This work was partially funded by FAPERJ and the Brazilian Research Council (CNPq). The authors also thank the anonymous reviewer for thought-provoking comments and criticisms and for providing [23] which helped to improve the paper.

References

- [1] J. C. Gower, “Some distance properties of latent root and vector methods used in multivariate analysis,” *Biometrika*, vol. 53, no. 3-4, pp. 325–338, 1966.
- [2] I. T. Jolliffe, *Principal Component Analysis*, Springer Series in Statistics, Springer, New York, NY, USA, 2nd edition, 1986.
- [3] W. J. Krzanowski, “Sensitivity in metric scaling and analysis of distance,” *Biometrics*, vol. 62, no. 1, pp. 239–244, 2006.
- [4] L. Lebart, “Which bootstrap for principal axes methods?” in *Selected Contributions in Data Analysis and Classification*, Studies in Classification, Data Analysis and Knowledge Organization, pp. 581–588, Springer, New York, NY, USA, 2007.
- [5] B. Efron, “Bootstrap methods: another look at the jackknife,” *The Annals of Statistics*, vol. 7, no. 1, pp. 1–26, 1979.
- [6] M. Linting, J. J. Meulman, P. J. F. Groenen, and A. J. van der Kooij, “Stability of non-linear principal components analysis: an empirical study using the balanced bootstrap,” *Psychological Methods*, vol. 12, no. 3, pp. 359–379, 2007.
- [7] G. Pfurtscheller, D. Flotzinger, and J. Kalcher, “Brain-computer interface—a new communication device for handicapped persons,” *Journal of Microcomputer Applications*, vol. 16, no. 3, pp. 293–299, 1993.
- [8] C. M. Krause, L. Sillanmäki, M. Koivisto et al., “The effects of memory load on event-related EEG desynchronization and synchronization,” *Clinical Neurophysiology*, vol. 111, pp. 2071–2078, 2000.
- [9] W. Klimesch, “EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis,” *Brain Research Reviews*, vol. 29, no. 2-3, pp. 169–195, 1999.
- [10] L. Lebart, A. Morineau, and K. M. Warwick, *Multivariate Descriptive Statistical Analysis: Correspondence Analysis and Related techniques for Large Matrices*, John Wiley & Sons, New York, NY, USA, 1984.
- [11] G. H. Golub and C. F. van Loan, *Matrix Computations*, John Hopkins University Press, Baltimore, Md, USA, 3rd edition, 1996.
- [12] R. B. Cattell, “The Scree test for the number of factors,” *Multivariate Behavioural Research*, vol. 1, no. 1, pp. 245–276, 1966.
- [13] D. A. Jackson, “Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches,” *Ecology*, vol. 74, no. 8, pp. 2204–2214, 1993.
- [14] B. Efron and R. Tibshirani, “Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy,” *Statistical Science*, vol. 1, no. 1, pp. 54–75, 1986.
- [15] P. Diaconis and B. Efron, “Computer-intensive methods in statistics,” *Scientific American*, vol. 248, no. 5, pp. 116–131, 1983.
- [16] D. G. Altman and P. Royston, “What do we mean by validating a prognostic model?” *Statistics in Medicine*, vol. 19, pp. 453–473, 2000.
- [17] F. E. Harrel Jr., K. L. Lee, and D. B. Mark, “Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors,” *Statistics in Medicine*, vol. 15, pp. 361–387, 1996.
- [18] B. Efron, “Estimating the error rate of a prediction rule: improvement on cross-validation,” *Journal of the American Statistical Association*, vol. 78, no. 382, pp. 316–331, 1983.
- [19] L. Milan and J. Whittaker, “Application of the parametric bootstrap to models that incorporate a singular value decomposition,” *Applied Statistics*, vol. 44, no. 1, pp. 31–49, 1995.
- [20] R. Sibson, “Studies in the robustness of multidimensional scaling: procrustes statistics,” *Journal of the Royal Statistical Society B: Methodological*, vol. 40, no. 2, pp. 234–238, 1978.
- [21] H. S. M. Coxeter, *Regular Polytopes*, Dover, New York, NY, USA, 2nd edition, 1973.
- [22] W. F. Eddy, “A new convex hull algorithm for planar sets,” *ACM Transactions on Mathematical Software*, vol. 3, no. 4, pp. 398–403, 1977.
- [23] B. Efron, “Bootstrap confidence intervals: good or bad?” *Psychological Bulletin*, vol. 104, no. 2, pp. 293–296, 1988.

- [24] R. Dubes and A. K. Jain, "Validity studies in clustering methodologies," *Pattern Recognition*, vol. 11, no. 4, pp. 235–254, 1979.
- [25] A. D. Gordon, "A review of hierarchical classification," *Journal of the Royal Statistical Society A*, vol. 150, no. 2, pp. 119–137, 1987.
- [26] G. W. Milligan and P. D. Isaac, "The validation of four ultrametric clustering algorithms," *Pattern Recognition*, vol. 12, no. 2, pp. 41–50, 1980.
- [27] S. M. Kay and S. L. Marple Jr., "Spectrum analysis—a modern perspective," *Proceedings of the IEEE*, vol. 69, no. 11, pp. 1380–1419, 1981.
- [28] D. M. Simpson, C. J. Tierra-Criollo, R. T. Leite, E. J. B. Zayen, and A. F. C. Infantosi, "Objective response detection in an electroencephalogram during somatosensory stimulation," *Annals of Biomedical Engineering*, vol. 28, no. 6, pp. 691–698, 2000.
- [29] K. Paul, V. Krajča, Z. Roth, J. Melichar, and S. Petránek, "Quantitative topographic differentiation of the neonatal EEG," *Clinical Neurophysiology*, vol. 117, no. 9, pp. 2050–2058, 2006.
- [30] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012.
- [31] H. Bengtsson, "R.matlab: Read and write of MAT files together with R-to-Matlab connectivity," R package version 1.7.0, 2013, <http://CRAN.R-project.org/package=R.matlab>.
- [32] Signal Developers, "Signal: signal processing," 2011, <http://r-forge.r-project.org/projects/signal>.
- [33] E. Dimitriadou, K. Hornik, F. Leisch, D. Meyer, and A. Weingessel, "e1071: misc functions of the department of statistics (e1071)," 2011, <http://CRAN.R-project.org/package=e1071>.
- [34] H. W. Borchers, "Pracma: Practical Numerical Math Functions (pracma)," 2014, <http://CRAN.R-project.org/package=pracma>.
- [35] E. Guadagnoli and W. F. Velicer, "Relation of sample size to the stability of components patterns," *Psychological Bulletin*, vol. 103, no. 2, pp. 265–275, 1988.
- [36] E. W. Steyerberg, M. J. C. Eijkemans, F. E. Harrell Jr., and J. D. F. Habbema, "Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets," *Medical Decision Making*, vol. 21, no. 1, pp. 45–56, 2001.
- [37] R. M. V. R. Almeida, A. F. C. Infantosi, J. H. R. Suassuna, and J. C. G. D. Costa, "Multiple correspondence analysis in predictive logistic modelling: application to a living-donor kidney transplantation data," *Computer Methods and Programs in Biomedicine*, vol. 95, no. 2, pp. 116–128, 2009.
- [38] S. Casarotto, A. M. Bianchi, S. Cerutti, and G. A. Chiarenza, "Principal component analysis for reduction of ocular artefacts in event-related potentials of normal and dyslexic children," *Clinical Neurophysiology*, vol. 115, no. 3, pp. 609–619, 2004.
- [39] T. Kobayashi and S. Kuriki, "Principal component elimination method for the improvement of S/N in evoked neuromagnetic field measurements," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 8, pp. 951–958, 1999.
- [40] A. Daffertshofer, C. J. C. Lamoth, O. G. Meijer, and P. J. Beek, "PCA in studying coordination and variability: a tutorial," *Clinical Biomechanics*, vol. 19, no. 4, pp. 415–428, 2004.
- [41] I. Milovanović and D. B. Popović, "Principal component analysis of gait kinematics data in acute and chronic stroke patients," *Computational and Mathematical Methods in Medicine*, vol. 2012, Article ID 649743, 8 pages, 2012.
- [42] P. J. G. da Silva, J. Nadal, and A. F. C. Infantosi, "Investigating the center of pressure velocity Romberg's quotient for assessing the visual role on the body sway," *Revista Brasileira de Engenharia Biomedica*, vol. 28, no. 4, pp. 319–326, 2012.
- [43] S. Slobounov, M. Hallett, S. Stanhope, and H. Shibasaki, "Role of cerebral cortex in human postural control: an EEG study," *Clinical Neurophysiology*, vol. 116, no. 2, pp. 315–323, 2005.
- [44] A. F. C. Infantosi and A. M. F. L. Miranda de Sá, "A statistical test for evaluating the event-related synchronization/desynchronization and its potential use in brain-computer-interfaces," in *Proceedings of the IFMBE, Latin American Congress on Biomedical Engineering*, vol. 18, pp. 1122–1136, Margarita Island, Venezuela, 2007.
- [45] P. J. G. Da-Silva, A. M. F. L. Miranda de Sá, and A. F. C. Infantosi, "Dynamic visual stimulation effects on cortical response EEG desynchronization," in *IFMBE Proceedings, World Congress on Medical Physics and Biomedical Engineering*, vol. 99, pp. 1573–1576, Beijing, China, 2012.
- [46] N. R. Draper and J. A. John, "Influential observations and outliers in regression," *Technometrics*, vol. 23, no. 1, pp. 21–26, 1981.

Research Article

Log-Gabor Energy Based Multimodal Medical Image Fusion in NSCT Domain

Yong Yang,¹ Song Tong,¹ Shuying Huang,² and Pan Lin³

¹ School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China

² School of Software and Communication Engineering, Jiangxi University of Finance and Economics, Nanchang 330032, China

³ Institute of Biomedical Engineering, Xi'an Jiaotong University, Xi'an 710049, China

Correspondence should be addressed to Shuying Huang; shuyinghuang2010@126.com

Received 7 June 2014; Revised 5 August 2014; Accepted 6 August 2014; Published 24 August 2014

Academic Editor: Ezequiel López-Rubio

Copyright © 2014 Yong Yang 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.

Multimodal medical image fusion is a powerful tool in clinical applications such as noninvasive diagnosis, image-guided radiotherapy, and treatment planning. In this paper, a novel nonsubsampling Contourlet transform (NSCT) based method for multimodal medical image fusion is presented, which is approximately shift invariant and can effectively suppress the pseudo-Gibbs phenomena. The source medical images are initially transformed by NSCT followed by fusing low- and high-frequency components. The phase congruency that can provide a contrast and brightness-invariant representation is applied to fuse low-frequency coefficients, whereas the Log-Gabor energy that can efficiently determine the frequency coefficients from the clear and detail parts is employed to fuse the high-frequency coefficients. The proposed fusion method has been compared with the discrete wavelet transform (DWT), the fast discrete curvelet transform (FDCT), and the dual tree complex wavelet transform (DTCWT) based image fusion methods and other NSCT-based methods. Visually and quantitatively experimental results indicate that the proposed fusion method can obtain more effective and accurate fusion results of multimodal medical images than other algorithms. Further, the applicability of the proposed method has been testified by carrying out a clinical example on a woman affected with recurrent tumor images.

1. Introduction

Medical imaging has attracted increasing attention in the recent years due to its vital component in medical diagnostics and treatment [1]. However, each imaging modality reports on a restricted domain and provides information in limited domains that some are common, and some are unique [2]. For example, computed tomography (CT) image can provide dense structures like bones and hard tissues with less distortion whereas magnetic resonance imaging (MRI) image is better visualized in the case of soft tissues [3]. Similarly, T1-MRI image provides the details of an anatomical structure of tissues while T2-MRI image provides information about normal and pathological tissues [4]. As a result, multimodal medical images which have relevant and complementary information are necessary to be combined for a comprehensive figure [5]. The multimodal medical image fusion is

the possible way to integrate complementary information from multiple modality images [6]. The image fusion not only obtains a more accurate and complete description of the same target, but also reduces the randomness and redundancy to increase the clinical applicability of image-guided diagnosis and assessment of medical problems [7].

Generally image fusion techniques can be divided into spatial domain and frequency domain techniques [8]. Spatial domain techniques are carried out directly on the source images. Weighted average method is the simplest spatial domain approach. However, along with simplicity, this method leads to several undesirable side effects like reduced contrast [9]. Other spatial domain techniques have been developed, such as intensity-hue-saturation (IHS), principal component analysis (PCA), and the Brovey transform [10]. Although the fused images obtained by these methods have high spatial quality, they usually overlook the high quality

of spectral information and suffer from spectral degradation [10]. Li et al. [11] introduced the artificial neural network (ANN) to make image fusion. However, the performance of ANN depends on the sample images and this is not an appealing characteristic. Yang and Blum [12] used a statistical approach to fuse the images. In their method, the distortion is modeled as a mixture of Gaussian probability density functions which is a limiting assumption. Since the actual objects usually contain structures at many different scales or resolutions and multiscale techniques can provide the means to exploit this fact, the frequency domain techniques especially the multiscale techniques have attracted more and more interest in image fusion [13].

In frequency domain techniques, each source image is first decomposed into a sequence of multiscale coefficients. Various fusion rules are then employed in the selection of these coefficients, which are synthesized via inverse transforms to form the fused image. Recently, series of frequency domain methods have been explored by using multiscale transform, including Laplacian pyramid transform, gradient pyramid transform, filter-subtract-decimate pyramid transform, discrete wavelet transform (DWT), and complex wavelet transform (CWT) [14–20]. There is evidence that multiscale transform based signal decomposition is similar to the human visual system. As we know the wavelet analysis, with its upstanding localize peculiarity in both time and frequency domain, has become one of the most commonly used methods in the fields of multiscale transform based image fusion [16]. However, wavelet analysis cannot effectively represent the line singularities and plane singularities of the images and thus cannot represent the directions of the edges of images accurately. To overcome these shortcomings of the wavelet transform, Do and Vetterli [17] proposed Contourlet transform which can give the asymptotic optimal representation of contours and has been successfully used for image fusion. However, the up- and downsampling process of Contourlet decomposition and reconstruction results in the Contourlet transform lacking shift-invariance and having pseudo-Gibbs phenomena in the fused image [19]. Later, da Cunha et al. [20] proposed the nonsubsampling Contourlet transform (NSCT) based on Contourlet transform. This method inherits the advantages of Contourlet transform, while possessing shift-invariance and effectively suppressing pseudo-Gibbs phenomena.

Although quite good results have been reported by NSCT based method, there is still much room to improve the fusion performance in the coefficient selection as follows.

- (a) The low-frequency coefficients of the fused image can be simply acquired by averaging the low-frequency coefficients of the input images. This rule decreased contrast in the fused images [21] and cannot give the fused subimage of high quality for medical images.
- (b) The popularly used larger absolute rule is implemented in the value of a single pixel of the current high-frequency subband. The disadvantage of this method is that the coefficients only know the value of a single pixel but not any of the relationship between

the corresponding coefficients in high-frequency subbands [22].

- (c) Most fusion rules of the NSCT-based methods are implemented in multifocus images [23], remote sensing images [24], and infrared and visible images [25]. The results are not of the same quality as those of the multimodal medical images. For example, Chai et al. [22] proposed a NSCT method based on features contrast of multiscale products to fuse multifocus images. However, it has been proven that this algorithm is not able to utilize prominent information present in the subbands efficiently and results in the poor quality when it is used to fuse multimodal medical images [26].

In this paper, a novel fusion framework based on NSCT is proposed for multimodal medical images. The main contribution of the method lies in the proposed fusion rule, which can capture the best membership of source images' coefficients to the corresponding fused coefficient. The phase congruency and Log-Gabor energy are unified as the fusion rules for low- and high-frequency coefficients, respectively. The phase congruency provides a contrast and brightness-invariant representation of low-frequency coefficients whereas Log-Gabor energy efficiently determines the frequency coefficients from the clear and detail parts in the high frequency. The combinations of these two techniques can preserve more details from source images and thus improve the quality of the fused images. Experiments indicate that the proposed framework can provide a better fusion outcome when compared to series of traditional image fusion methods in terms of both subjective and objective evaluations.

The rest of the paper is organized as follows. NSCT and phase congruency are described in Section 2 followed by the proposed multimodal medical image fusion framework in Section 3. Experimental results and discussions are given in Section 4 and the concluding remarks are presented in Section 5.

2. Preliminaries

This section provides the description of concepts on which the proposed framework is based. These concepts, including NSCT and phase congruency, are described as follows.

2.1. Nonsubsampling Contourlet Transform (NSCT). Contourlet transform can be divided into two stages [19]: Laplacian pyramid (LP) and directional filter bank (DFB) and offers an efficient directional multiresolution image representation. Among them, LP is first utilized to capture the point singularities and then followed by the DFB to link the singular point into linear structures. LP is used to decompose the original images into low-frequency and high-frequency subimages, and DFB divides the high-frequency subbands into directional subbands. The Contourlet decomposed schematic diagram is shown in Figure 1.

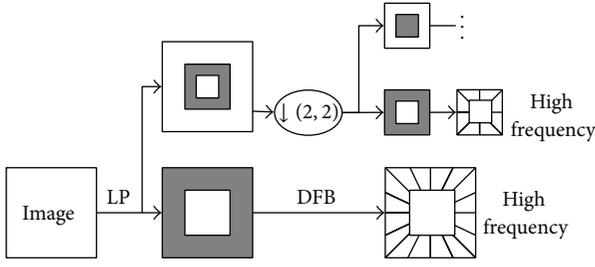


FIGURE 1: Contourlet decomposed schematic diagram.

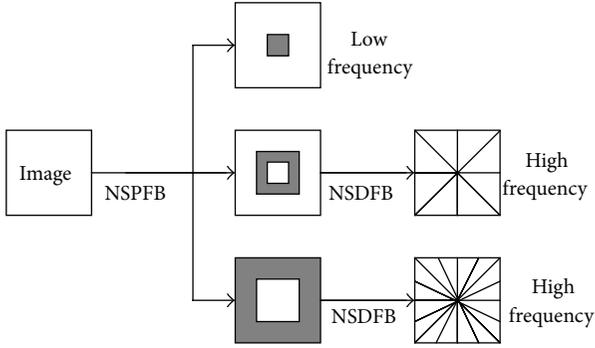


FIGURE 2: NSCT decomposed schematic diagram.

The NSCT is proposed based on the theory of Contourlet transform. NSCT inherits the advantage of Contourlet transform, enhances directional selectivity and shift-invariance, and effectively overcomes the pseudo-Gibbs phenomena. NSCT is built on nonsubsampled pyramid filter bank (NSPFB) and nonsubsampled directional filter bank (NSDFB) [21]. Figure 2 gives the NSCT decomposition framework with $k = 2$ levels.

The NSPFB ensures the multiscale performance by taking advantage of a two-channel nonsubsampled filter bank and one low-frequency subband and one high-frequency subband that can be produced at each decomposition level. The NSDFB is a two-channel nonsubsampled filter bank constructed by eliminating the downsamplers and upsamplers and combining the directional fan filter banks in the nonsubsampled directional filter [23]. NSDFB allows the direction decomposition with l levels in each high-frequency subbands from NSPFB and then produces 2^l directional subbands with the same size as the source images. Thus, the NSDFB provides the NSCT the multidirection performance and offers more precise directional detail information to get more accurate results [23]. Therefore, NSCT leads to better frequency selectivity and an essential property of the shift-invariance on account of nonsubsampled operation. The size of different subimages decomposed by NSCT is identical. Additionally, NSCT-based image fusion can effectively mitigate the effects of misregistration on the results [27]. Therefore, NSCT is more suitable for image fusion.

2.2. Phase Congruency. Phase congruency is a feature perception approach which provides information that is invariant to image illumination and contrast [28]. This model is built on

the Local Energy Model [29], which postulates that important features can be found at points where the Fourier components are maximally in phase. Furthermore, the angle at which the phase congruency occurs signifies the feature type. The phase congruency can be used for feature detection [30]. The model provides useful feature localization and noise compensation. The phase congruency at a point (i, j) can be defined as follows [31]:

$$PC^\theta(i, j) = \sum_n W^\theta(i, j) \left[A_n^\theta(i, j) \left(\cos(\phi_n^\theta(i, j) - \bar{\phi}_n^\theta(i, j)) \right) - \left| \sin(\phi_n^\theta(i, j) - \bar{\phi}_n^\theta(i, j)) \right| - T \right]_+ \times \left(\sum_n A_n^\theta(i, j) + \varepsilon \right)^{-1}, \quad (1)$$

where θ is the orientation, $W^\theta(i, j)$ is the weighting factor based on frequency spread, $A_n^\theta(i, j)$ and $\bar{\phi}_n^\theta(i, j)$ are the amplitude and phase for wavelet scale n , respectively, $\phi_n^\theta(i, j)$ is the weighted mean phase, T is a noise threshold constant, and ε is a small constant value to avoid division by zero. The notation $[\]_+$ denotes that the enclosed quantity is equal to itself when the value is positive, and zero otherwise. For details of phase congruency measure see [29].

As we know the multimodal medical images have the following characteristics:

- the images of different modal have significantly different pixel mappings;
- the capturing environment of different modalities varies and resulted in the change of illumination and contrast;
- the edges and corners in the images are identified by collecting frequency components of the image that is in phase.

According to the literature [26, 32], it is easy to find that phase congruency is not only invariant to different pixel intensity mappings, illumination, and contrast changes, but also gives the Fourier components that are maximally in phase. These all will lead to efficient fusion. That is why we use phase congruency for multimodal medical fusion.

3. Proposed Multimodal Medical Image Fusion Framework

The framework of the proposed multimodal medical image fusion algorithm is depicted in Figure 3, but before describing it, the definition of local Log-Gabor energy in NSCT domain is first described as follows.

3.1. Log-Gabor Energy in NSCT Domain. The high-frequency coefficients in NSCT domain represent the detailed components of the source images, such as the edges, textures, and region boundaries [21]. In general, the coefficients with larger absolute values correspond to the sharper brightness

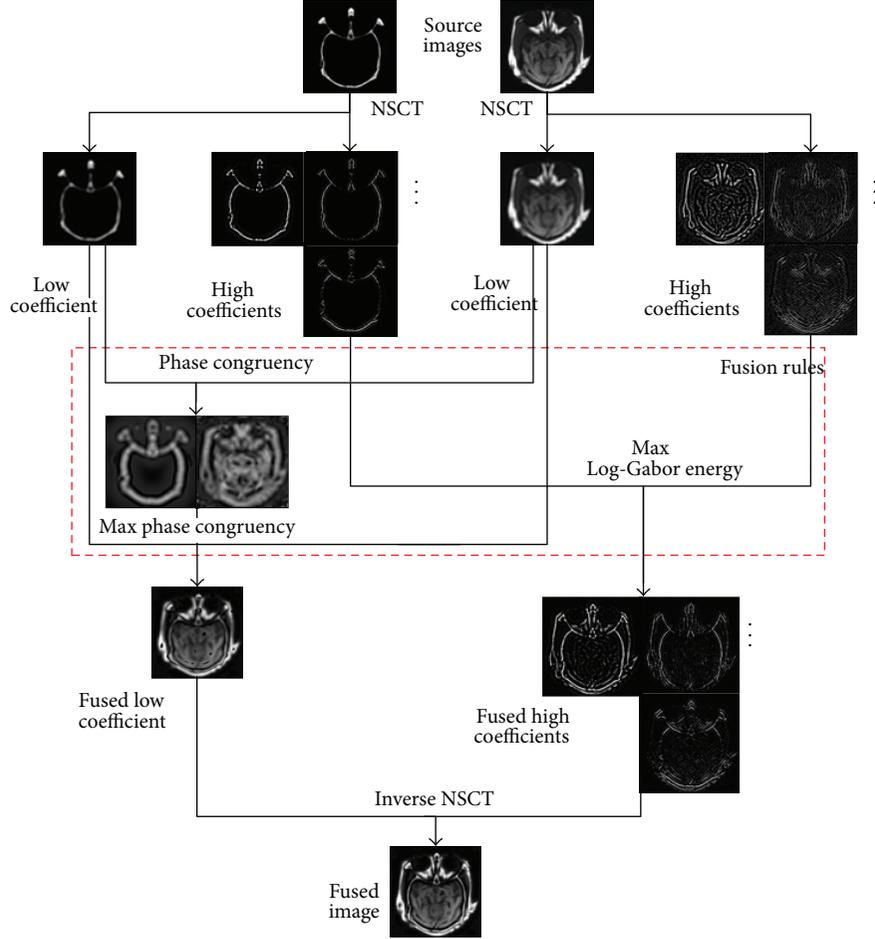


FIGURE 3: The framework of the proposed fusion algorithm.

in the image. It is to be noted that the noise is also related to high-frequency coefficients and may cause miscalculation of sharpness values and, therefore, affect the fusion performance [26]. Furthermore, human visual system is generally more sensitive to texture detail features than the value of a single pixel.

To overcome the defect mentioned above, a novel high-frequency fusion rule based on local Log-Gabor energy is designed in this paper. Gabor wavelet is a popular technique that has been extensively used to extract texture features [33]. Log-Gabor filters are proposed based on Gabor filters. Compared with Gabor filters, Log-Gabor filters cover the shortage of the high frequency of Gabor filter component expression and more in accord with human visual system [34]. Therefore, Log-Gabor wavelet can achieve optimal spatial orientation and wider spectrum information at the same time and thus more truly reflect the frequency response of the natural images and improve the performance in terms of the accuracy [35].

Under polar coordinates, the Log-Gabor wavelet is expressed as follows [36]:

$$g(f, \theta) = \exp \left\{ -\frac{[\ln(f/f_0)]^2}{2[\ln(\sigma/f_0)]^2} \right\} \times \exp \left\{ -\frac{(\theta - \theta_0)^2}{2\sigma_\theta^2} \right\}, \quad (2)$$

in which f_0 is the center frequency of the Log-Gabor filter, θ_0 is the direction of the filter, σ is used to determine the bandwidth of the radial filter, and σ_θ is used to determine the bandwidth of the orientation. If $g_{kl}^{uv}(i, j)$ correspond to Log-Gabor wavelets in scale u and direction v , the signal response is expressed as follows:

$$G_{kl}^{uv}(i, j) = H_{kl}(i, j) * g_{kl}^{uv}(i, j), \quad (3)$$

where $H_{kl}(i, j)$ is the coefficient located at (i, j) in high-frequency subimages of the source image A or B at the k th scale, l th direction, and $*$ denotes convolution operation. The Log-Gabor energy of high-frequency subimages at the k th scale, l th direction, is expressed as follows:

$$E_{kl}(i, j) = \sum_{u=1}^U \sum_{v=1}^V \sqrt{\text{real}(G_{kl}^{uv}(i, j))^2 + \text{imag}(G_{kl}^{uv}(i, j))^2}, \quad (4)$$

in which, $\text{real}(G_{kl}^{uv}(i, j))$ is the real part of $G_{kl}^{uv}(i, j)$ and $\text{imag}(G_{kl}^{uv}(i, j))$ is the imaginary part of $G_{kl}^{uv}(i, j)$. The Log-Gabor energy in NSCT domain at the local area around the pixel (i, j) is given as

$$\text{LE}_{kl}(i, j) = \frac{1}{(2M+1)(2N+1)} \sum_{m=-M}^M \sum_{n=-N}^N E_{kl}(i+m, j+n), \quad (5)$$

in which $(2M + 1)(2N + 1)$ is the window size. The proposed definition of the local Log-Gabor energy not only extracts more useful features from high-frequency coefficients, but also keeps a well performance in noisy environment.

3.2. Proposed Fusion Framework. The proposed NSCT-based image fusion framework is discussed in this subsection. Considering the input multimodal medical images (A and B) are perfectly registered. The framework of the proposed fusion method is shown Figure 3 and described as the following three steps.

Step 1. Perform θ -level NSCT on A and B to obtain one low-frequency subimage and series of high-frequency subimages at each level and direction l ; that is, $A : \{L^A, H_{k,l}^A\}$ and $B : \{L^B, H_{k,l}^B\}$, where L^A, L^B are the low-frequency subimages and $H_{k,l}^A, H_{k,l}^B$ represent the high-frequency subimages at level $k \in [1, \theta]$ in the orientation l .

Step 2. Fuse low- and high-frequency subbands via the following novel fusion rule to obtain composite low- and high-frequency subbands.

The low-frequency coefficients represent the approximation component of the source images. The popular widely used approach is to apply the averaging methods to produce the fused coefficients. However, this rule reduced contrast in the fused images and cannot give the fused subimage of high quality for medical image. Therefore, the criterion based on the phase congruency that is introduced in Section 2.2 is employed to fuse the low-frequency coefficients. The fusion rule for the low-frequency subbands is defined as

$$L^F(i, j) = \begin{cases} L^A(i, j), & \text{if } PC^{A\theta}(i, j) > PC^{B\theta}(i, j) \\ L^B(i, j), & \text{if } PC^{A\theta}(i, j) < PC^{B\theta}(i, j) \\ 0.5(L^A(i, j) + L^B(i, j)), & \text{if } PC^{A\theta}(i, j) = PC^{B\theta}(i, j), \end{cases} \quad (6)$$

where $PC^{A\theta}(i, j), PC^{B\theta}(i, j)$ is the phase congruency extracted from low-frequency subimages of the source images A and B , respectively.

For the high-frequency coefficients, the most common fusion rule is selecting the coefficient with larger absolute values. This rule does not take any consideration of the surrounding pixels and cannot give the fused components of high quality for medical image. Especially when the source images contain noise, the noise could be mistaken for fused coefficients and cause miscalculation of the sharpness value. Therefore, the criterion based on Log-Gabor energy is introduced to fuse high-frequency coefficients. The fusion rule for the high-frequency subbands is defined as

$$H_{kl}^F(i, j) = \begin{cases} H_{kl}^A(i, j), & \text{if } LE_{kl}^A(i, j) \geq LE_{kl}^B(i, j) \\ H_{kl}^B(i, j), & \text{otherwise,} \end{cases} \quad (7)$$

where $LE_{kl}^A(i, j)$ and $LE_{kl}^B(i, j)$ are the local Log-Gabor energy extracted from high-frequency subimages at the k th scale and l th direction of source images A and B , respectively.

Step 3. Perform θ -level by the inverse NSCT on the fused low- and high-frequency subimages. The fused image is obtained ultimately in this way.

4. The Experimental Results and Analysis

It is well known that different image quality metrics imply the visual quality of images from different aspects, but none of them can directly imply the quality. In this paper, we consider both the visual representation and quantitative assessment of the fused images. For evaluation of the proposed fusion method, we have considered five separate fusion performance metrics defined as below.

4.1. Evaluation Index System

4.1.1. Standard Deviation. The standard deviation (STD) of an image with size of $M_0 \times N_0$ is defined as [37]:

$$\text{Std} = \left(\frac{1}{M_0 \times N_0} \sum_{i=1}^{M_0} \sum_{j=1}^{N_0} (F(i, j) - \hat{\mu})^2 \right)^{1/2}, \quad (8)$$

where $F(i, j)$ is the pixel value of the fused image at the position (i, j) and $\hat{\mu}$ is the mean value of the image. The STD can be used to estimate how widely the gray values spread in an image. The larger the STD, the better the result.

4.1.2. Edge Based Similarity Measure. The edge based similarity measure ($Q^{AB/F}$) is proposed by Xydeas and Petrović [38]. The definition is given as

$$Q^{AB/F} = \frac{\sum_{i=1}^{M_0} \sum_{j=1}^{N_0} [Q^{AF}(i, j) w^A(i, j) + Q^{BF}(i, j) w^B(i, j)]}{\sum_{i=1}^{M_0} \sum_{j=1}^{N_0} [w^A(i, j) + w^B(i, j)]}, \quad (9)$$

where $w^A(i, j)$ and $w^B(i, j)$ are the corresponding gradient strength for images A and B , respectively. The definition of $Q^{AF}(i, j)$ and $Q^{BF}(i, j)$ is given as

$$\begin{aligned} Q^{AF}(i, j) &= Q_a^{AF}(i, j) Q_g^{AF}(i, j), \\ Q^{BF}(i, j) &= Q_a^{BF}(i, j) Q_g^{BF}(i, j), \end{aligned} \quad (10)$$

where $Q_a^{x^F}(i, j)$ and $Q_g^{x^F}(i, j)$ are the edge strength and orientation preservation values at location for image x (A or B), respectively. The edge based similarity measure gives the similarity between the edges transferred from the input images to the fused image [26]. The larger the value, the better the fusion result.

4.1.3. *Mutual Information.* Mutual information (MI) [39] between the fusion image F and the source images A and B is defined as follows:

$$\begin{aligned} MI &= MI^{AF} + MI^{BF}, \\ MI^{AF} &= \sum_{f=0}^L \sum_{a=0}^L p^{AF}(a, f) \log_2 \left(\frac{p^{AF}(a, f)}{p^A(a) p^F(f)} \right), \\ MI^{BF} &= \sum_{f=0}^L \sum_{b=0}^L p^{BF}(b, f) \log_2 \left(\frac{p^{BF}(b, f)}{p^B(b) p^F(f)} \right), \end{aligned} \quad (11)$$

where MI^{x^F} denotes the normalized mutual information between the fused image and the input image A, B ; a, b and $f \in [0, L]$, and L is the number of bins. $p^A(a)$, $p^B(b)$, and $p^F(f)$ are the normalized gray level histograms of source images and fused image. $p^{x^F}(a, f)$ is the joint gray level histograms between fused image and each source image.

MI can indicate how much information the fused image F conveys about the source images A and B [22]. Therefore, the greater the value of MI, the better the fusion effect.

4.1.4. *Cross Entropy.* The cross entropy is defined as [8]:

$$CE = \sum_{l=0}^{L-1} P_l \log_2 \frac{P_l}{Q_l}, \quad (12)$$

where P_l and Q_l denote the gray level histogram of the source image and the fused image, respectively. The cross entropy is used to evaluate the difference between the source images and the fused image. The lower value corresponds to the better fusion result.

4.1.5. *Spatial Frequency.* Spatial frequency is defined as [40]:

$$SF = \sqrt{RF^2 + CF^2}, \quad (13)$$

where RF and CF are the row frequency and column frequency, respectively, and are defined as

$$\begin{aligned} RF &= \sqrt{\frac{1}{M_0 N_0} \sum_{i=0}^{M_0-1} \sum_{j=0}^{N_0-1} [F(i, j) - F(i, j-1)]^2}, \\ CF &= \sqrt{\frac{1}{M_0 N_0} \sum_{i=0}^{M_0-1} \sum_{j=0}^{N_0-1} [F(i, j) - F(i-1, j)]^2}. \end{aligned} \quad (14)$$

The spatial frequency reflects the edge information of the fused image. Larger spatial frequency values indicate better image quality.

4.2. *Experiments on Multimodal Medical Image Fusion.* To evaluate the performance of the proposed image fusion approach, the experiments are performed on three groups of multimodal medical images. These images are characterized in two distinct pairs: (1) CT and MRI; (2) MR-T1 and MR-T2. The images in Figures 4(a1)-4(b1) and 4(a2)-4(b2) are

CT and MRI images, whereas Figures 4(a3)-4(b3) are T1-weighted MR image (MR-T1) and T2-weighted MR image (MR-T2). All images have the same size of 256×256 pixel, with 256-level gray scale. For all these image groups, the results of the proposed fusion framework are compared with those of the traditional discrete wavelet transform (DWT) [13, 16], the second generation curvelet transform (fast discrete curvelet transform, FDCT) [41, 42], the dual tree complex wavelet transform (DTCWT) [4], and the nonsubsampled Contourlet transform (NSCT-1 and NSCT-2) based methods. The high-frequency coefficients and low-frequency coefficients of DWT, FDCT, DTCWT, and NSCT-1 based methods are merged by the popular widely used fusion rule of selecting the coefficient with larger absolute values and the averaging rule (average-maximum rule), respectively. NSCT-2 based method is merged by the fusion rules proposed by Bhatnagar, et al. in [26]. In order to perform a fair comparison, the source images are all decomposed into the same levels with 3 for those methods except FDCT method. For DWT method, the images are decomposed using the DBSS (2, 2) wavelet. For implementing NSCT, "9-7" filters and "pkva" filters (how to set the filters can be seen in [43]) are used as the pyramidal and directional filters, respectively.

4.2.1. *Subjective Evaluation.* The first pair of medical images are two groups of brain CT and MRI images on different aspects, shown in Figures 4(a1), 4(b1) and 4(a2), 4(b2), respectively. It can be easily seen that the CT image shows the dense structure while MRI provides information about soft tissues. The obtained fused images from DWT, FDCT, DTCWT, NSCT-1, and NSCT-2 are shown in Figures 4(c1)-4(g1) and 4(c2)-4(g2), respectively. The results for the proposed fusion method have been shown in Figures 4(h1) and 4(h2). On comparing these results, it can be easily observed that the proposed method outperforms those fusion methods and has good visual representation of fused image.

The second pair of medical images are MR-T1 and MR-T2 images, shown in Figures 4(a3) and 4(b3). The comparison of DWT, FDCT, DTCWT, NSCT-1, NSCT-2, and proposed method, shown in Figures 4(c3)-4(h3), clearly implies that the fusion result of the proposed method has better quality and contrast in comparison to other methods.

Similarly, on observing the noticeable improvement has been emphasized in Figure 4 by the red arrows and the analysis above, one can easily verify the fact that again the proposed method has been found superior in terms of visual representation over DWT, FDCT, DTCWT, NSCT-1, and NSCT-2 fusion methods.

4.2.2. *Objective Evaluation.* For objective evaluation of the fusion results, shown in Figure 4, we have used five fusion metrics: cross entropy, spatial frequency, STD, $Q^{AB/F}$, and MI. The quantitative comparison of cross entropy and spatial frequency for these fused images is visually given by Figures 5-6 and other metrics are given by Table 1.

On observing Figure 5, one can easily observe all the three results of the proposed scheme have lower values of cross entropy than any of the DWT, FDCT, DTCWT, NSCT-1, and NSCT-2 fusion methods. The cross entropy is used to

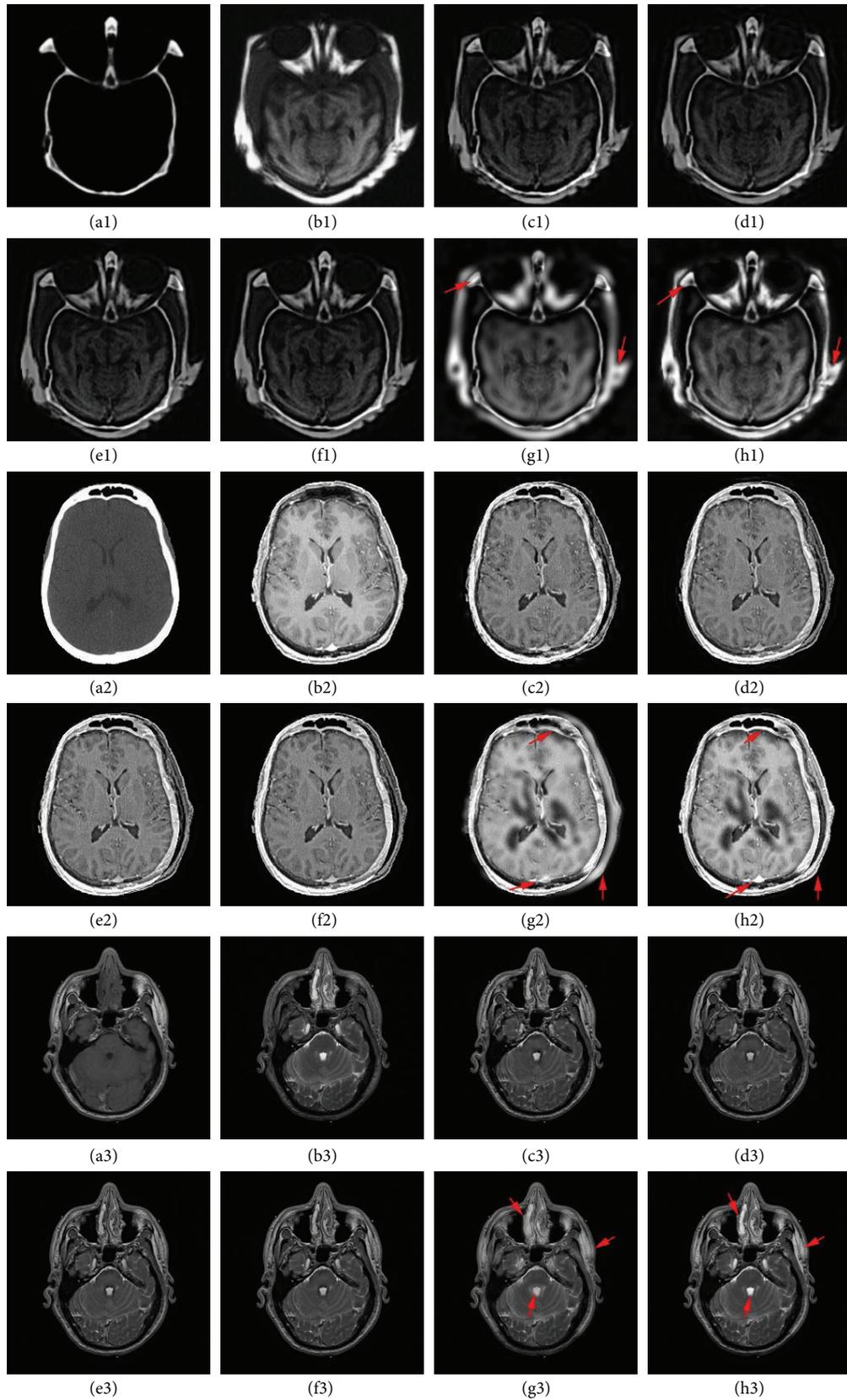


FIGURE 4: Source multimodal medical images: (a1), (b1) image group 1 (CT and MRI); (a2), (b2) image group 2 (CT and MRI); (a3), (b3) image group 3 (MR-T1 and MR-T2); fused images from (c1), (c2), (c3) DWT based method; (d1), (d2), (d3) FDCT based method; (e1), (e2), (e3) DTCWT based method; (f1), (f2), (f3) NSCT-1 based method; (g1), (g2), (g3) NSCT-2 based method; (h1), (h2), (h3) our proposed method.

TABLE 1: Comparison on quantitative evaluation of different methods for the first set medical images.

Images	Evaluation	DWT	FDCT	DTCWT	NSCT-1	NSCT-2	Proposed method
Image group 1 (CT and MRI)	STD	41.160	39.170	40.132	40.581	54.641	58.476
	$Q^{AB/F}$	0.618	0.592	0.652	0.683	0.427	0.716
	MI	2.371	2.029	2.393	2.438	1.886	2.580
Image group 2 (CT and MRI)	STD	72.260	70.572	71.407	71.733	80.378	83.334
	$Q^{AB/F}$	0.567	0.584	0.606	0.622	0.450	0.618
	MI	3.137	3.138	3.198	3.235	2.915	3.302
Image group 3 (MR-T1 and MR-T2)	STD	37.358	36.856	37.358	37.277	39.023	40.630
	$Q^{AB/F}$	0.635	0.647	0.668	0.682	0.589	0.689
	MI	3.209	3.142	3.302	3.347	3.349	3.469

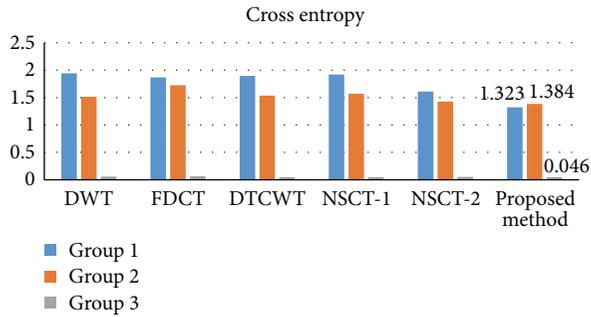


FIGURE 5: Comparison on cross entropy of different methods and images.

evaluate the difference between the source images and the fused image. Therefore, the lower value corresponds to the better fusion result.

On observing Figure 6, two values of the spatial frequency of the fused image obtained by the proposed method are the highest, and the other one is 6.447 which is close to the highest value 6.581. Observation of Table 1 yields that all the three results of the proposed fusion scheme have higher values of STD, $Q^{AB/F}$, and MI than any of other methods except one value of $Q^{AB/F}$ (image group 2) is the second best. However, an overall comparison shows the superiority of the proposed fusion scheme.

4.2.3. Combined Evaluation. Since the subjective and objective evaluations separately are not able to examine fusion results, we have combined them. From these figures (Figures 4–6) and table (Table 1), it is clearly to find that the proposed method not only preserves most of the source images characteristics and information, but also improves the definition and the spatial quality better than the existing methods, which can be justified by the optimum values of objective criteria except one value of spatial frequency (image group 3) and one value of $Q^{AB/F}$ (image group 2). Consider the example of the first set of images: the five criteria values of the proposed method are 1.323 (cross entropy), 7.050 (spatial frequency), 58.476 (STD), 0.716 ($Q^{AB/F}$), and 2.580 (MI), respectively. Each of them is the optimal one in the first set of experiments.

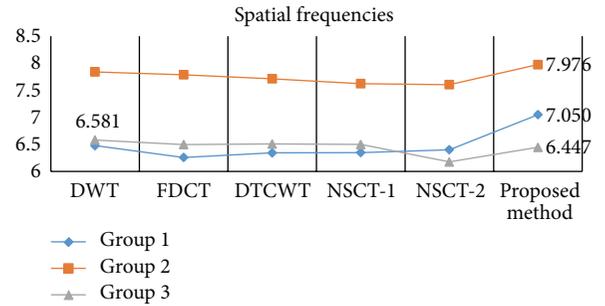


FIGURE 6: Comparison on spatial frequencies of different methods and images.

Among these methods, the result of NSCT-2 based method also gives poor results when comparing to the proposed NSCT-based method. This stems from the fact that high-frequency fusion rule of NSCT-2 based method is not able to extract the detail information in the high frequency effectively. Also, by carefully looking at the outputs of the proposed NSCT-based method (Figures 4(h1), 4(h2), and 4(h3)), we can find that they get more contrast and more spatial resolution than the outputs of NSCT-2 based method (highlighted by the red arrows) and other methods. The main reason behind the better performance is that the proposed fusion rules for low- and high-frequency coefficients can effectively extract prominent and detail information from the source images. Therefore, it can be possible to conclude that the proposed method is better than the existing methods.

4.3. Fusion of Multimodal Medical Noisy Images and a Clinical Example. To evaluate the performance of the proposed method in noisy environment, the input image group 1 has been additionally corrupted with Gaussian noise, with a standard deviation of 5% (shown in Figures 7(a) and 7(b)). In addition, a clinical applicability on noninvasive diagnosis of neoplastic disease is given in the last subsection.

4.3.1. Fusion of Multimodal Medical Noisy Images. For comparison, apart from visual observation, objective criteria on STD, MI, and $Q^{AB/F}$ are used to evaluate how much clear or detail information of the source images is transferred to the fused images. However, maybe these criteria cannot

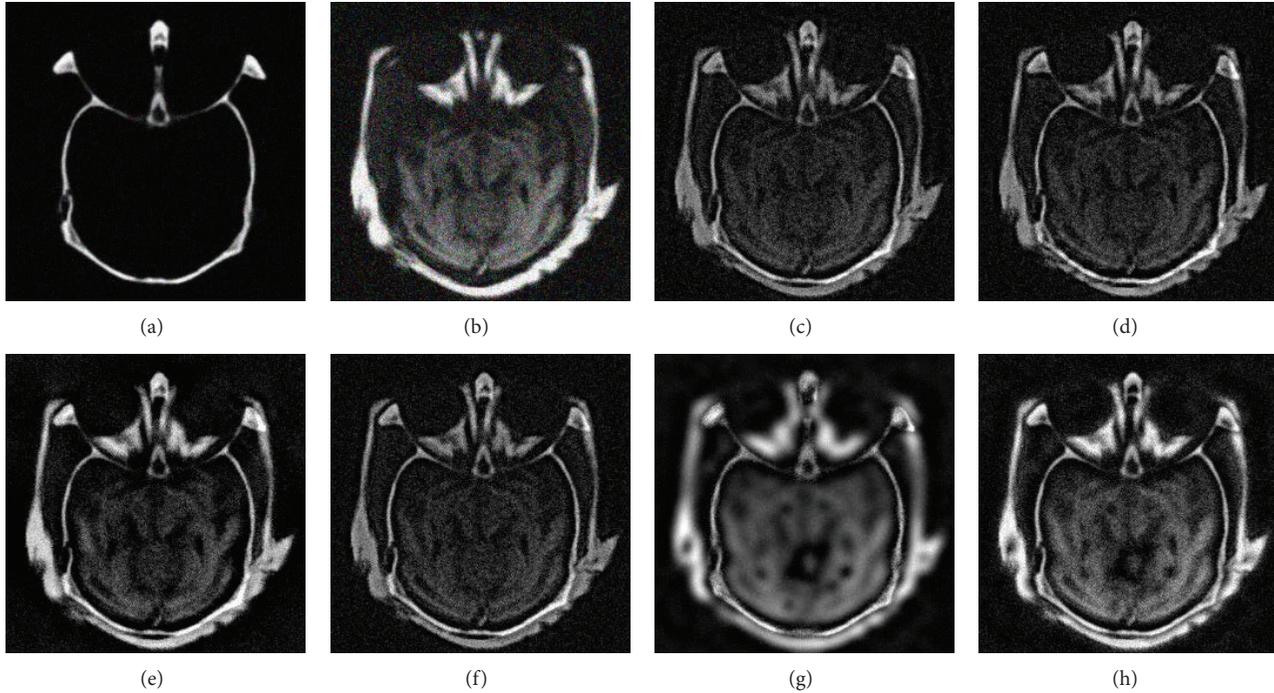


FIGURE 7: The multimodal medical images with noise: (a) CT image with 5% noise, (b) MRI image with 5% noise; fused images by (c) DWT, (d) FDCT, (e) DTCWT, (f) NSCT-1, (g) NSCT-2, and (h) proposed method.

TABLE 2: Comparison on quantitative evaluation of different methods for the noise medical images.

Evaluation	DWT	FDCT	DTCWT	NSCT-1	NSCT-2	Proposed method
STD	41.893	40.044	40.995	41.338	52.533	57.787
$Q^{AB/F}$	0.592	0.579	0.619	0.637	0.269	0.640
MI	2.785	2.539	2.826	2.882	1.816	2.914

effectively evaluate the performance of the fusion methods in terms of the noise transmission. For further comparison, Peak Signal to Noise Ratio (PSNR), a ratio between the maximum possible power of a signal and the power of noise that affects the fidelity [44], is used. The larger the value of PSNR, the less the image distortion [45]. PSNR is formulated as

$$\text{PSNR} = 10 \lg \left| \frac{255^2}{\text{RMSE}^2} \right|, \quad (15)$$

where RMSE denotes the Root Mean Square Error between the fused image and the reference image. The reference image in the following experiment is selecting from Figure 4(h1), which is proven to be the best performance compared to other images.

Figure 7 illustrates the fusion results obtained by the different methods. The comparison of the images fused by DWT, FDCT, DTCWT, NSCT-1, NSCT-2, and proposed method, shown in Figures 7(c)–7(h), clearly implies that the fused image by proposed method has better quality and contrast than other methods. Figure 8 shows the values of PSNR of different methods in fusing noisy images. One can observe that the proposed method has higher values of

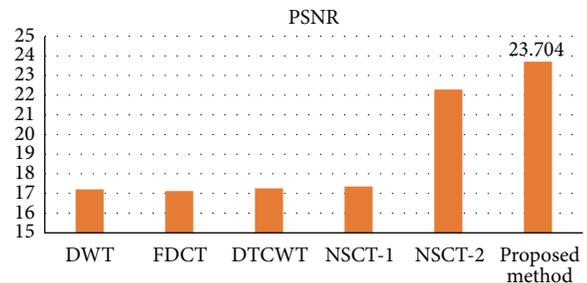


FIGURE 8: Comparison on PSNR of different methods for the noise medical images.

PSNR than any of the DWT, FDCT, DTCWT, NSCT-1, and NSCT-2 fusion methods. Table 2 gives the quantitative results of fused images and shows that the values of STD, $Q^{AB/F}$, and MI are also the highest of all the six methods. From the analysis above, we can also observe that the proposed scheme provides the best performance and outperforms the other algorithms. In addition, compared with the result of the NSCT-1 method using the average-maximum rule, it

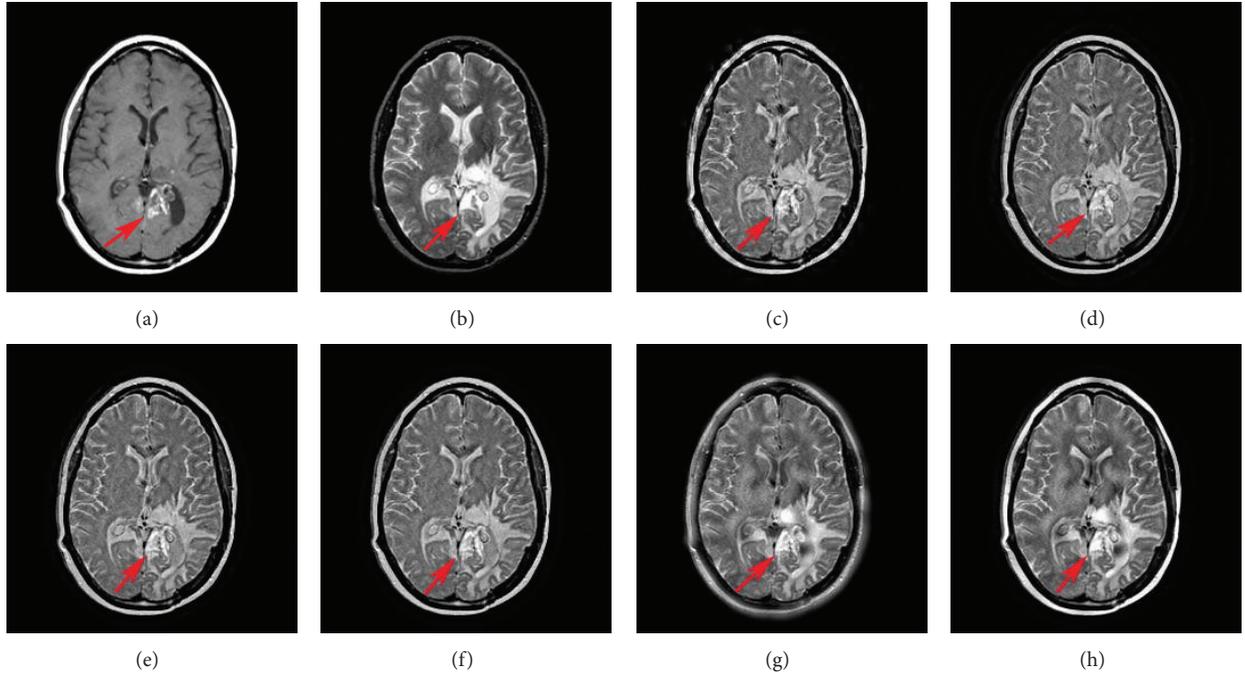


FIGURE 9: Brain images of the man with recurrent tumor: (a) MR-T1 image, (b) MR-T2 image; fused images by (c) DWT, (d) FDCT, (e) DTCWT, (f) NSCT-1, (g) NSCT-2, and (h) proposed method.

TABLE 3: Comparison on quantitative evaluation of different methods for the clinical medical images.

Evaluation	DWT	FDCT	DTCWT	NSCT-1	NSCT-2	Proposed method
STD	64.433	62.487	63.194	63.652	60.722	67.155
$Q^{AB/F}$	0.541	0.565	0.578	0.597	0.411	0.624
MI	2.507	2.455	2.540	2.584	2.392	2.624

demonstrated the validities of the proposed fusion rule in noisy environment.

4.3.2. A Clinical Example on Noninvasive Diagnosis. In order to demonstrate the practical value of the proposed scheme in medical imaging, one clinical case on neoplastic diagnosis is considered where MR-T1/MR-T2 medical modalities are used. The images have been downloaded from the Harvard University site (<http://www.med.harvard.edu/AANLIB/home.html>). Figures 9(a)-9(b) show the recurrent tumor case of a 51-year-old woman who sought medical attention because of gradually increasing right hemiparesis (weakness) and hemianopia (visual loss). At craniotomy, left parietal anaplastic astrocytoma was found. A right frontal lesion was biopsied. A large region of mixed signal on MR-T1 and MR-T2 images gives the signs of the possibility of active tumor (highlighted by the red arrows).

Figures 9(c)-9(h) show the fused images by DWT, FDCT, DTCWT, NSCT-1, NSCT-2, and proposed method. It is obvious that the fused image by proposed method has better contrast and sharpness of active tumor (highlighted by the red arrows) than other methods. Table 3 shows the quantitative evaluation of different methods for the clinical medical images. The values of the proposed method are optimum in

terms of STD, MI, and $Q^{AB/F}$. From Figure 9 and Table 3, we can obtain the same conclusion that the proposed scheme provides the best performance and outperforms the other algorithms.

5. Conclusions

Multimodal medical image fusion plays an important role in clinical applications. But the real challenge is to obtain a visually enhanced image through fusion process. In this paper, a novel and effective image fusion framework based on NSCT and Log-Gabor energy is proposed. The potential advantages include (1) NSCT is more suitable for image fusion because of its advantages such as multiresolution, multidirection, and shift-invariance; (2) a new couple of fusion rules based on phase congruency and Log-Gabor energy are used to preserve more useful information in the fused image to improve the quality of the fused images and overcome the limitations of the traditional fusion rules; and (3) the proposed method can provide a better performance than the current fusion methods whatever the source images are clean or noisy. In the experiments, five groups of multimodal medical images, including one group with noise and one group clinical example of a woman affected with recurrent tumor, are

fused by using traditional fusion methods and the proposed framework. The subjective and objective comparisons clearly demonstrate that the proposed algorithm can enhance the details of the fused image and can improve the visual effect with less information distortion than other fusion methods. In the future, we plan to design a pure C++ platform to reduce the time cost and extend our method for 3D or 4D medical image fusion.

Conflict of Interests

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

Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 61262034, no. 61462031 and no. 61473221), by the Key Project of Chinese Ministry of Education (no. 211087), by the Doctoral Fund of Ministry of Education of China (no. 20120201120071), by the Natural Science Foundation of Jiangxi Province (no. 20114BAB211020 and no. 20132BAB201025), by the Young Scientist Foundation of Jiangxi Province (no. 20122BCB23017), by the Science and Technology Application Project of Jiangxi Province (no. KJLD14031), by the Science and Technology Research Project of the Education Department of Jiangxi Province (no. GJJ14334), and by the Fundamental Research Funds for the Central Universities of China.

References

- [1] R. Shen, I. Cheng, and A. Basu, "Cross-scale coefficient selection for volumetric medical image fusion," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 4, pp. 1069–1079, 2013.
- [2] V. D. Calhoun and T. Adali, "Feature-based fusion of medical imaging data," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 711–720, 2009.
- [3] Y. Yang, D. S. Park, S. Huang, and J. Yang, "Fusion of CT and MR images using an improved wavelet based method," *Journal of X-Ray Science and Technology*, vol. 18, no. 2, pp. 157–170, 2010.
- [4] R. Singh, R. Srivastava, O. Prakash, and A. Khare, "Multimodal medical image fusion in dual tree complex wavelet transform domain using maximum and average fusion rules," *Journal of Medical Imaging and Health Informatics*, vol. 2, no. 2, pp. 168–173, 2012.
- [5] Z. Wang and Y. Ma, "Medical image fusion using m -PCNN," *Information Fusion*, vol. 9, no. 2, pp. 176–185, 2008.
- [6] V. Barra and J. Boire, "A general framework for the fusion of anatomical and functional medical images," *NeuroImage*, vol. 13, no. 3, pp. 410–424, 2001.
- [7] A. P. James and B. V. Dasarathy, "Medical image fusion: a survey of the state of the art," *Information Fusion*, vol. 19, pp. 4–19, 2014.
- [8] P. Balasubramaniam and V. P. Ananthi, "Image fusion using intuitionistic fuzzy sets," *Information Fusion*, vol. 20, pp. 21–30, 2014.
- [9] Y. Yang, S. Y. Huang, F. J. Gao et al., "Multi-focus image fusion using an effective discrete wavelet transform based algorithm," *Measurement Science Review*, vol. 14, no. 2, pp. 102–108, 2014.
- [10] P. S. Pradhan, R. L. King, N. H. Younan, and D. W. Holcomb, "Estimation of the number of decomposition levels for a wavelet-based multiresolution multisensor image fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 12, pp. 3674–3686, 2006.
- [11] S. Li, J. T. Kwok, and Y. Wang, "Multifocus image fusion using artificial neural networks," *Pattern Recognition Letters*, vol. 23, no. 8, pp. 985–997, 2002.
- [12] J. Yang and R. S. Blum, "A statistical signal processing approach to image fusion for concealed weapon detection," in *Proceedings of the International Conference on Image Processing (ICIP '02)*, vol. 1, pp. 513–516, September 2002.
- [13] R. Singh and A. Khare, "Multiscale medical image fusion in wavelet domain," *The Scientific World Journal*, vol. 2013, Article ID 521034, 10 pages, 2013.
- [14] B. Aiazzi, L. Alparone, A. Barducci, S. Baronti, and I. Pippi, "Multispectral fusion of multisensor image data by the generalized Laplacian pyramid," in *Proceedings of the 1999 IEEE International Geoscience and Remote Sensing Symposium (IGARSS '99)*, vol. 2, pp. 1183–1185, Hamburg, Germany, July 1999.
- [15] V. S. Petrović and C. S. Xydeas, "Gradient-based multiresolution image fusion," *IEEE Transactions on Image Processing*, vol. 13, no. 2, pp. 228–237, 2004.
- [16] Y. Yang, D. S. Park, S. Huang, and N. Rao, "Medical image fusion via an effective wavelet-based approach," *Eurasip Journal on Advances in Signal Processing*, vol. 2010, Article ID 579341, 2010.
- [17] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091–2106, 2005.
- [18] R. Singh and A. Khare, "Fusion of multimodal medical images using Daubechies complex wavelet transform—a multiresolution approach," *Information Fusion*, vol. 19, pp. 49–60, 2014.
- [19] X. B. Qu, J. W. Yan, and G. D. Yang, "Multifocus image fusion method of sharp frequency localized Contourlet transform domain based on sum-modified-Laplacian," *Optics and Precision Engineering*, vol. 17, no. 5, pp. 1203–1212, 2009.
- [20] A. L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsampling contourlet transform: theory, design, and applications," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3089–3101, 2006.
- [21] Q. Zhang and B. Guo, "Multifocus image fusion using the nonsubsampling contourlet transform," *Signal Processing*, vol. 89, no. 7, pp. 1334–1346, 2009.
- [22] Y. Chai, H. Li, and X. Zhang, "Multifocus image fusion based on features contrast of multiscale products in nonsubsampling contourlet transform domain," *Optik*, vol. 123, no. 7, pp. 569–581, 2012.
- [23] X. B. Qu, J. W. Yan, H. Z. Xiao, and Z. Zhu, "Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampling contourlet transform domain," *Acta Automatica Sinica*, vol. 34, no. 12, pp. 1508–1514, 2008.
- [24] Y. Wu, C. Wu, and S. Wu, "Fusion of remote sensing images based on nonsubsampling contourlet transform and region segmentation," *Journal of Shanghai Jiaotong University (Science)*, vol. 16, no. 6, pp. 722–727, 2011.
- [25] J. Adu, J. Gan, and Y. Wang, "Image fusion based on nonsubsampling contourlet transform for infrared and visible light image," *Infrared Physics & Technology*, vol. 61, pp. 94–100, 2013.

- [26] G. Bhatnagar, Q. M. J. Wu, and Z. Liu, "Directive contrast based multimodal medical image fusion in NSCT domain," *IEEE Transactions on Multimedia*, vol. 15, no. 5, pp. 1014–1024, 2013.
- [27] S. Das and M. K. Kundu, "NSCT-based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency," *Medical and Biological Engineering and Computing*, vol. 50, no. 10, pp. 1105–1114, 2012.
- [28] P. Kovési, "Image features from phase congruency," *Videre*, vol. 1, no. 3, pp. 1–26, 1999.
- [29] M. C. Morrone and R. A. Owens, "Feature detection from local energy," *Pattern Recognition Letters*, vol. 6, no. 5, pp. 303–313, 1987.
- [30] P. Kovési, "Phase congruency: a low-level image invariant," *Psychological Research*, vol. 64, no. 2, pp. 136–148, 2000.
- [31] L. Zhang, X. Mou, and D. Zhang, "FSIM: a feature similarity index for image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.
- [32] A. Wong and W. Bishop, "Efficient least squares fusion of MRI and CT images using a phase congruency model," *Pattern Recognition Letters*, vol. 29, no. 3, pp. 173–180, 2008.
- [33] L. Yu, Z. He, and Q. Cao, "Gabor texture representation method for face recognition using the Gamma and generalized Gaussian models," *Image and Vision Computing*, vol. 28, no. 1, pp. 177–187, 2010.
- [34] P. Yao, J. Li, X. Ye, Z. Zhuang, and B. Li, "Iris recognition algorithm using modified Log-Gabor filters," in *Proceedings of the 18th International Conference on Pattern Recognition (ICPR '06)*, pp. 461–464, August 2006.
- [35] R. Raghavendra and C. Busch, "Novel image fusion scheme based on dependency measure for robust multispectral palm-print recognition," *Pattern Recognition*, vol. 47, no. 6, pp. 2205–2221, 2014.
- [36] R. Redondo, F. Šroubek, S. Fischer, and G. Cristóbal, "Multifocus image fusion using the log-Gabor transform and a Multisize Windows technique," *Information Fusion*, vol. 10, no. 2, pp. 163–171, 2009.
- [37] X. Zhang, X. Li, and Y. Feng, "A new image fusion performance measure using Riesz transforms," *Optik*, vol. 125, no. 3, pp. 1427–1433, 2014.
- [38] C. S. Xydeas and V. Petrović, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308–309, 2000.
- [39] B. Yang and S. Li, "Pixel-level image fusion with simultaneous orthogonal matching pursuit," *Information Fusion*, vol. 13, no. 1, pp. 10–19, 2012.
- [40] Q. Miao, C. Shi, P. Xu, M. Yang, and Y. Shi, "A novel algorithm of image fusion using shearlets," *Optics Communications*, vol. 284, no. 6, pp. 1540–1547, 2011.
- [41] E. Candès, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Modeling and Simulation*, vol. 5, no. 3, pp. 861–899, 2006.
- [42] M. Fu and C. Zhao, "Fusion of infrared and visible images based on the second generation curvelet transform," *Journal of Infrared and Millimeter Waves*, vol. 28, no. 4, pp. 254–258, 2009.
- [43] S. Li, B. Yang, and J. Hu, "Performance comparison of different multi-resolution transforms for image fusion," *Information Fusion*, vol. 12, no. 2, pp. 74–84, 2011.
- [44] Y. Phamila and R. Amutha, "Discrete Cosine Transform based fusion of multi-focus images for visual sensor networks," *Signal Processing*, vol. 95, pp. 161–170, 2014.
- [45] L. Wang, B. Li, and L. Tian, "Multi-modal medical volumetric data fusion using 3D discrete shearlet transform and global-to-local rule," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 1, pp. 197–206, 2014.

Research Article

Image Tracking for the High Similarity Drug Tablets Based on Light Intensity Reflective Energy and Artificial Neural Network

Zhongwei Liang,¹ Liang Zhou,² Xiaochu Liu,¹ and Xiaogang Wang¹

¹ School of Mechanical & Electrical Engineering, Guangzhou University, Guangzhou 510006, China

² School of Mathematics & Information Science, Shanghai Lixin University of Commerce, Shanghai 201620, China

Correspondence should be addressed to Zhongwei Liang; lzwstalin@126.com

Received 24 April 2014; Revised 6 June 2014; Accepted 22 June 2014; Published 17 July 2014

Academic Editor: Rafael M. Luque-Baena

Copyright © 2014 Zhongwei Liang 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.

It is obvious that tablet image tracking exerts a notable influence on the efficiency and reliability of high-speed drug mass production, and, simultaneously, it also emerges as a big difficult problem and targeted focus during production monitoring in recent years, due to the high similarity shape and random position distribution of those objectives to be searched for. For the purpose of tracking tablets accurately in random distribution, through using surface fitting approach and transitional vector determination, the calibrated surface of light intensity reflective energy can be established, describing the shape topology and topography details of objective tablet. On this basis, the mathematical properties of these established surfaces have been proposed, and thereafter artificial neural network (ANN) has been employed for classifying those moving targeted tablets by recognizing their different surface properties; therefore, the instantaneous coordinate positions of those drug tablets on one image frame can then be determined. By repeating identical pattern recognition on the next image frame, the real-time movements of objective tablet templates were successfully tracked in sequence. This paper provides reliable references and new research ideas for the real-time objective tracking in the case of drug production practices.

1. Introduction

Nowadays, the automation level of drug mass production and quality inspection still remains undeveloped; most drug tablets in random position distribution were manually inspected for picking out those unqualified ones and then hand-encapsulated after amount counting or bottle packaging by ocular estimation. All drug tablets were highly similar or identical to each other in shape, size, surface topography, and color, which imposes colossal operation overload and time exhaustion on the working operators. In the interests of dealing with these difficulties, tablet tracking by computer video was greatly hampered by their identical size and similar shape, keeping a far distance from the traditional applications of image recognition, such as face recognition or vehicle video monitoring. Nevertheless, light intensity reflective energy and ANN recognition have remarkable superiority in locating those scattered tablet objectives, which undoubtedly confirms its reliable results and high efficient performance. Based on this theoretical hypothesis, the labor intensities will

be greatly reduced, and simultaneously the high precision of tracking processes can also be ensured.

According to the recently published literatures, it can be learned that image tracking or video tracking were generally employed in such conditions as moving vehicle monitoring or people tracking. Wang and Wang [1] have proposed an approach based on computer version and image processing technique for automatic grain tracking; through an efficient box-tracking-based method, presented by Fukuda et al. [2], mammalian cells (fibroblast) on microcarrier have been clearly recognized. At the same time, Dill et al. [3] studied the tracking cells supported by digital image processing and obtained a rather satisfactory result for multiobjective positioning during their high-speed movement. Kuba et al. [4] also made their contributions on automatic particle detection by one-class SVM from microscope image. Similar performances have been obtained by González et al. and so forth [5] on tracking olive trees in high-resolution satellite images.

On the other hand, Guarino et al. [6] have optimized the traditional image analysis approach. But their optimization

process needs a much more computation storage and time interval than that of other traditional ones, which restricts its direct usages in real-time drug mass production; Marçal [7] presented new methods based on mathematical morphology and were suitable for grains of circular shape; Kim et al. [8] have successfully proposed a real-time approach for tracking the number of passing people by using one single camera. Similar research founding proposed by Professor Luque-Baena et al. can also be learned from [9, 10] as well. In the area of moving-objective image detection and its automatic classification, Wong et al. [11] used an example of outdoor people tracking for tourist-flow estimation in a constrained environment. According to the research result of Buczkowski et al. [12], the box-tracking method (BCM) shows a very high percentage of error in experimental practice, due to the identical shape and high-speed movement of those drug tablets to be studied. In order to cover this deficiency, Wang et al. [13] have proposed a new and efficient framework for pedestrian analysis and tracking, which consists of rule induction classifier and linear regression model. Agustin and Oh [14] preprocessed the detected foreground objects to eliminate pixel noise and small artifacts by performing opening morphology operation. Besides, Lien et al. [15] also paid high attention on a novel vehicle detection method without background modeling simultaneously. Zhan and Luo [16] made their contribution on system design of real-time vehicle recognition based on video for Windows (AVI) files; the development of a block-based real-time tracking system has been concluded by Park et al. [17] and Głowacz et al. [18] and Park et al. [19] have paid their attentions to the optical flow calculation and the area-based decision rule, respectively, which helps in acquiring reliable image tracking results. When discussing the cross-correlation in pattern recognition, Hong [20] used correlation coefficient of fuzzy numbers under arithmetic operations. Park et al. [21] have studied the correlation coefficient of interval-valued intuitionistic probability sets to multiple attribute group decisions. Ye [22] and Son [23], respectively, proposed probability decision-making method and fuzzy entropy determination for investigating the mutual-relationships among different factors. Other related research results can be learned from [24–26]. Furthermore, literature [27–29] also presented their latest progresses in cross-correlation mechanisms while such evaluation approaches as intuitionistic probability, interval-valued probability sets, and gamma rank correlation coefficient were, respectively, described in [30–32].

Since these commonly used approaches were found being influenced by traditional limitations, the real-time tablet tracking cannot be ensured easily; more importantly, due to the fact that most existing image tracking technologies focus on human tracking or vehicle monitoring, which can be distinguished easily by the different shapes and the widely divergent sizes of those targeted objects in practice [33–35], the high similarity objective tracking method suitable for high-speed drug mass production still remains unstudied and undeveloped, that becomes a research margin of advanced video inspection in these years.

This paper is structured as follows. Section 1 outlines the importance and necessity of objective tablet tracking.

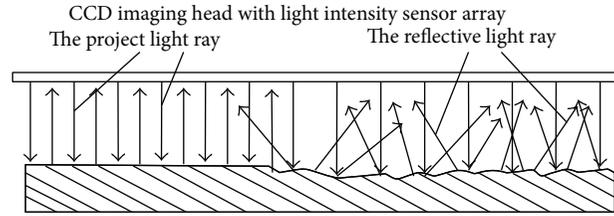


FIGURE 1: The light reflective effect caused by different surface topography areas, which composes an original property demonstration for calibrating the inflective-energy surface properties of drug tablets.

Section 2 describes the theoretical foundation of light intensity reflective energy surface. Section 3 presents mathematical properties for describing the reflective energy surfaces of drug tablets, Section 4 discusses the detailed tracking processes, with tracking discussions and performance comparisons mentioned in Section 5, and finally Section 6 concludes this paper as respected.

2. The Theoretical Foundation of Light Intensity Reflective Energy Surface

During the imaging process, the light was projected vertically thereafter a clear image can be obtained for showing the surface topography and geometrical shape of objective tablet. But, at the same time, the unsmooth surface topography on those drug tablets can also be identified due to their different light reflective effects. Figure 1 shows the light reflective effect according to different topography areas, which composes an original property demonstration for calibrating the inflective-energy of objective tablet in this experiment. Since different topographic areas on objective tablet present different reflective effects of light, an average-distributed reflective energy condition of light intensity can be obtained on the smooth part of tablet surface. When the lighting ray casts on the objective surface to be observed, it will be reflected to CCD unit in the form of roughly parallel light routes; thus, an image with uniform distribution of light intensity will be gotten and a low concentration of lightness energy can also be shown. On the contrary, the light illumination on the rugged areas of tablet surface brings about the diffuse reflection of lighting rays, which results in the ununiformed concentration of lightness energy. Based on this theoretical presumption, the topography properties of objective drug tablet can be clearly classified with the reflective light intensity measurement, the computation of illumination energy, and the characteristic modeling of energy distribution.

3. Surface Mathematical Properties

Since the reflective energy of light intensity demonstrated by one given image pixel can be regarded as the height value in the z axis, thereafter the reflective energy distribution of the whole tablet image can be described in the form of free-form surface, from the perspective of spatial surface fitting domain. Based on this theoretical foundation the following physical

properties have been proposed, for the sake of calibrating the established reflective energy surface of objective tablet in a quantitative way [36, 37].

Property 1. As it was well known that surface elasticity demonstrates the transformation resisting capability of one objective spatial surface shape under the force impact caused by external force loading, it becomes a typical index to calibrate the shape properties of spatial surfaces in a mathematical coordinate system, and simultaneously surface elasticity shows the concentricity level of characteristic distribution in the form of elasticity value as well. In this experiment, Property 1 demonstrates the elasticity distribution concentrating on one certain topography section, which makes it being conveniently used to describe topography properties from the perspective of elasticity variance distribution:

$$\begin{aligned} \varphi = & \alpha_1 \sum_{i=1}^m W_{ui}^2 + \beta_1 \sum_{i=1}^m W_{uui}^2 + \alpha_2 \sum_{j=1}^n W_{vj}^2 + \beta_2 \sum_{j=1}^n W_{vjj}^2 \\ & + \alpha_1 \alpha_2 \beta_1 \beta_2 \sum_{i=1}^m \sum_{j=1}^n W_{uivj}^2 - 2f(u, v)W. \end{aligned} \quad (1)$$

Here, W denotes an objective topography surface in the form of B-spline basis function; W_u , W_v , W_{uu} , W_{vv} , and W_{uv} are the partial derivatives of W in the first order, second order, and hybrid state of u, v axes, respectively; α_1 , α_2 , β_1 , and β_2 are the given coefficients, $f(u, v)$ denotes a given function of surface vector, and finally m, n denote the order amounts of surface vector in u, v axes.

Property 2. Energy of surface plays a prominent role as fairness function in the occasions of geometric modeling or surface microproperty analysis. Therefore, energy distribution can be computed and quantified to denote its belonged experimental conditions and geometric properties. This property defines the scattering level of surface energy in a mathematical sense, for a high value shows a more decentralized scattering of surface energy:

$$\begin{aligned} \zeta = & \sum_{i=1}^m \int_{\Omega} S_u(u_i)^2 du + \sum_{i=1}^m \sum_{i=1}^m \iint_{i \in \Omega} S_{uu}(u_{ii})^2 du du \\ & + \sum_{j=1}^n \int_{\Omega} S_v(v_j)^2 dv + \sum_{j=1}^n \sum_{j=1}^n \iint_{j \in \Omega} S_{vv}(v_{jj})^2 dv dv \quad (2) \\ & + \sum_{i=1}^m \sum_{j=1}^n \iint_{\Omega} S_{uv}(u_i v_j)^2 du dv. \end{aligned}$$

Here, $S_u(u_i)$, $S_{uu}(u_{ii})$, $S_v(v_j)$, $S_{vv}(v_{jj})$, and $S_{uv}(u_i v_j)$ denote the first order, second order, and hybrid derivatives of objective surface $f(u, v)$ in u, v axes.

Property 3. Since surface construction highly depends on external loading effect, the fitted result will be deformed in an obvious scale. In the purpose of quantifying the influential

effects caused by external loading, this property was newly proposed to calibrate the difference deviation and variation principle between the forced and original surfaces, as high amendment quantity demonstrates a relative critical deformation of resultant topography, with the definition shown as follows:

$$\begin{aligned} \rho = & -2 \sum_{i=m_u}^u \sum_{j=m_v}^v V_{i,j} \iint_{i,j \in \Omega} N_{i,s_u}(u) N_{j,s_v}(v) N_{i,j}(uv) \\ & \times f(u, v) du dv. \end{aligned} \quad (3)$$

Here, $N_{i,s_u}(u)$, $N_{j,s_v}(v)$, and $N_{i,j}(uv)$ denote the B-spline boundary control surfaces in u, v axes, respectively, with $V_{i,j}$ denoting the transitional vector between two adjacent control vertexes impacted by external loading.

Property 4. The computation of the radial polynomial of one coordinate point (x, y, z) on free surface consists of three steps: computations of radial polynomials, radial basis functions, and radial polynomial moments, through projecting the coordinates of one control point on its belonged basis functions. $R_{nm}(\gamma)$ denotes the radial polynomial of one control point (x, y, z) :

$$R_{nm}(\gamma) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s [(n-s)!] \gamma^{n-2s}}{s! ((n+|m|)/2-s)! ((n-|m|)/2-s)!}. \quad (4)$$

Here, n denotes a positive integer or zero and m denotes an integer number and $n - |m| = \text{even number}$, with $|m| \leq n$; r denotes the vector length spaced from the origin point to one control point on surface $f(x, y, z)$, defined as $\gamma = \sqrt{x^2 + y^2 + z^2}$, $-1 < x, y, z < 1$.

Property 5. As the traditional Zernike moments have been extensively used, they receive much research attention in a number of fields: object recognition, terrain reconstruction, roughness segmentation, edge detection, and biomedical measurement. They have already become one of the most widely used families of orthogonal mathematical moments, owing to their extraordinary properties of being invariant or insensitive to any arbitrary rotation of the objective surface in the three-dimensional (3D) coordinate system. Based on the radial polynomial computation of one coordinate point (x, y, z) mentioned before, an improved Zernike moment was presented by focusing on their magnitude values as follows:

$$Z_{nm} = \frac{n+1}{\pi} \int_0^1 \int_0^{2\pi} R_{nm}(\gamma) e^{jm\theta} f(\gamma, \theta) \gamma d\gamma d\theta. \quad (5)$$

Property 6. $P(u, v)$ ($0 < u, v < 1$) was supposed to be the objective surface labeled by m control vertexes in u direction, with the node vectors being $(u_0, u_0, u_0, u_0, \dots, u_{m-4}, u_{m-3}, u_{m-3}, u_{m-3}, u_{m-3})$, $u_0 = 0$, $u_{m-3} = 1$, and n control vertexes in v direction, with the node vectors being $(v_0, v_0, v_0, v_0, \dots, v_{n-4}, v_{n-3}, v_{n-3}, v_{n-3}, v_{n-3})$, $v_0 = 0$, $v_{n-3} = 1$. Two signless integral numbers n_1, n_2 were employed for

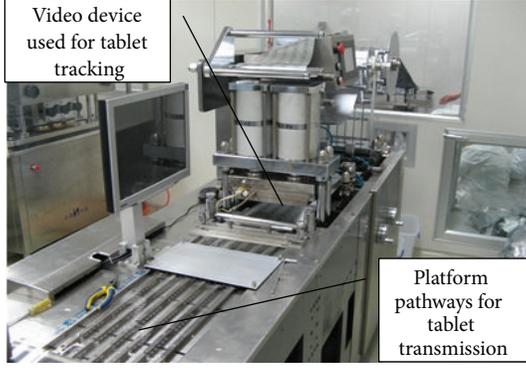


FIGURE 2: The video inspection platform employed in this experiment for tracking drug tablets.

$m \leq 2^{n_1} + 3$, $n \leq 2^{n_2} + 3$, and V represents the congregation of $(m-2)(n-2)/2$ parametric points in $P(u, v)$: (u_k, v_l) ($k = 0, 1, \dots, m-3$; $l = 0, 1, \dots, n-3$). Then, the fairing error of objective surface was defined as

$$\varepsilon_{n_1, n_2} = \frac{\max_{(u,v) \in V} |P(u, v) - P_{n_1, n_2}(u, v)|}{|P(u, v)|} \times 100\%; \quad (6)$$

$$\overline{\varepsilon}_{n_1, n_2} = 1 - \frac{\varepsilon - \varepsilon_{n_1, n_2}}{\varepsilon_{n_1-1, n_2-1} - \varepsilon_{n_1, n_2}}.$$

Here, $P_{n_1, n_2}(u, v)$ denotes the approximated surface of $P(u, v)$ with $n_1 \times n_2$ control vertexes, which can be obtained from the frequency decomposition of $P(u, v)$ in its 3D coordinate lattice.

4. Experiment for Drug Tablet Tracking

4.1. The Arrangement of Experimental Device. Stainless steel material was employed to manufacture the experimental platform, with the purpose of avoiding the high light reflective index and mirror effect of background material, causing these two physical phenomena to reduce the contrast ratio and blur the objective clearance in video image remarkably [38, 39]. On the stainless steel pathways, the drug tablets were scattered into the monitoring field for high-speed video detection. Figure 2 illustrates the video inspection platform employed, including the video device for tablet tracing and the platform pathways for tablet transmission. For the convenience of tracing these tablet objectives with high accuracy and efficiency, we subdivided the whole imaging field into several blocks by the size of $100 \text{ mm} \times 100 \text{ mm}$ and then used the DMIRM CCD high-speed imaging system manufactured by German LEICA to capture the moving processes of drug tablets at the rate of 24 frames per second (FPS). It is noteworthy that this video imaging system consists of three essential components: a precise high-speed CCD head equipped with light intensity sensor array, a PC equipped with video processing board, and a set of multidimensional triangulation processing software for ensuring high clearances of video image frames. Since the light source being used is a visible coaxial LED (220 V; 20 W;

700 nm) combined with diffraction grating that generates a fixed number of stripes, which enables the system to capture more than 450,000 pixels in one single image frame. The range measurements were computed by using triangulation operation between the binary digital images and the calibrated imaging plane locations. The high-speed CCD microscopic camera was able to capture image pixel data in 0.01 ms (enabling image acquisition from a continuously moving manipulator) and process data array in 10 ms and has a measurement deviation of $1/20000$ in the field of view. This CCD imaging head has been attached to a DMIRM 3-axis directly driven configuration manipulator for self-adaptively video imaging; thus, it eliminates image distortion caused by inhomogeneous light and keeps a rapid reaction and stable tracing performance simultaneously. For the purpose of improving the detection precision or dynamic tracing of drug tablets to the precise scale (1~5 mm), we predetermined the imaging depth as 0.3 m; therefore, the planar platform can be focused. On the other hand, the position overlap of drug tablets can be effectively avoided due to the continual high-frequency vibration of the experimental platform pathways, which reduces the probability of erroneous judgment to the maximum extent. In this experiment, the exposure time for digital imaging was supposed to be 0.5 s whereas the moving speed of drug tablet is about 3.5 m/s; therefore, one given tablet moves 1.75 m during the whole imaging process. In order to reduce the blur effect on the captured images, image sharpening was used to keep the high resolution of those studied objective tablets and thus accurate pattern recognition can be ensured.

4.2. The Computation of Transitional Vectors between Surface Control Points. In order to establish surface model for describing the light intensity reflective energy of drug tablets, the obtained energy value at the specific coordinate position of one pixel can be demonstrated as the z -axis value in the vertical direction; therefore, it was considered as the spatial coordinate of one control point. Under usual illumination, the external surface of drug tablet presents a constant situation of light intensity and uniform distribution of grey level. Based on this precondition, the three-dimensional (3D) transitional vectors can be computed.

During the capturing of digital images, the observer sight line keeps orthogonal intersecting with the imaging plane called XOY . The light intensity of one surface point (x, y) can be denoted by $I(x, y)$, its project vector was determined as $[pi, qi, -1]$, and then the normal vector of its reflective energy surface can be described as $[p, q, -1]$. In this paper, θ denotes the included angle that lies between the project vector of lighting and the normal vector of reflective energy surface. The reflective intensity of lighting energy in the direction of normal vector can be computed as [40, 41]

$$E(x, y) = I(x, y) \rho \cos \theta. \quad (7)$$

Here, $E(x, y)$ denotes the reflective energy intensity of the image pixel (x, y) , whereas $I(x, y)$ denotes the incident illumination intensity and is supposed as a constant number;

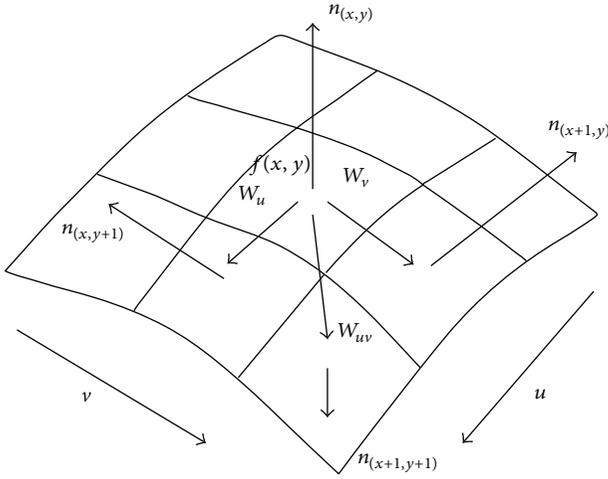


FIGURE 3: The transitional vectors and their spatial distribution on the fitted reflective energy surface.

ρ denotes the light reflectivity level of objective tablet surface (constant number).

Then, the cosine value of this included angle θ can be expressed as [42]

$$\cos \theta = \frac{pp_i + qq_i + rr_i}{\sqrt{p^2 + q^2 + r^2} \sqrt{p_i^2 + q_i^2 + r_i^2}}. \quad (8)$$

The energy intensity of one image pixel (or called as the reflective light intensity at the objective pixel) can be identified as

$$E(x, y) = I(x, y) \rho \frac{pp_i + qq_i + rr_i}{\sqrt{p^2 + q^2 + r^2} \sqrt{p_i^2 + q_i^2 + r_i^2}}. \quad (9)$$

Figure 3 shows the transitional vectors between surface control points and their spatial distribution. Through presenting a specific image pixel at (x, y) , this research obtained the normal vector of energy surface: $n_{(x,y)} = [p_1, q_1, r_1]$, together with other normal ones at its neighboring: $n_{(x+1,y)} = [p_2, q_2, r_2]$, $n_{(x,y+1)} = [p_3, q_3, r_3]$, and $n_{(x+1,y+1)} = [p_4, q_4, r_4]$. On this basis, the tangent vectors W_u , W_v , and W_{uv} at the positions of $(x+1, y)$, $(x, y+1)$, and $(x+1, y+1)$ can be determined, respectively:

$$W_u = n_{(x,y)} \times n_{(x+1,y)} = \begin{vmatrix} i & j & k \\ p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \end{vmatrix} \\ = \left[\begin{vmatrix} q_1 & r_1 \\ q_2 & r_2 \end{vmatrix}, \begin{vmatrix} r_1 & p_1 \\ r_2 & p_2 \end{vmatrix}, \begin{vmatrix} p_1 & q_2 \\ p_2 & q_1 \end{vmatrix} \right] = (u_1, u_2, u_3),$$

$$W_v = n_{(x,y)} \times n_{(x,y+1)} = \begin{vmatrix} i & j & k \\ p_1 & q_1 & r_1 \\ p_3 & q_3 & r_3 \end{vmatrix} \\ = \left[\begin{vmatrix} q_1 & r_1 \\ q_3 & r_3 \end{vmatrix}, \begin{vmatrix} r_1 & p_1 \\ r_3 & p_3 \end{vmatrix}, \begin{vmatrix} p_1 & q_2 \\ p_3 & q_3 \end{vmatrix} \right] = (v_1, v_2, v_3), \\ W_{uv} = n_{(x+1,y)} \times n_{(x,y+1)} = \begin{vmatrix} i & j & k \\ p_2 & q_2 & r_2 \\ p_3 & q_3 & r_3 \end{vmatrix} \\ = \left[\begin{vmatrix} q_2 & r_2 \\ q_3 & r_3 \end{vmatrix}, \begin{vmatrix} r_2 & p_2 \\ r_3 & p_3 \end{vmatrix}, \begin{vmatrix} p_2 & q_2 \\ p_3 & q_3 \end{vmatrix} \right] = (uv_1, uv_2, uv_3). \quad (10)$$

The secondary partial derivative vectors were identified as $W_{uu} = (u_{11}, u_{22}, u_{33})$ and $W_{vv} = (v_{11}, v_{22}, v_{33})$, which demonstrates the variance ratio of those tangent vectors in u and v directions, respectively. Therefore, the finite difference method can be used to determine these two items:

$$W_{uu1} = (uu_1, uu_2, uu_3)_{(x,y)} = W_{u2} - W_{u1} \\ = (u_1, u_2, u_3)_{(x+1,y)} - (u_1, u_2, u_3)_{(x,y)} \\ W_{uu2} = (uu_1, uu_2, uu_3)_{(x+1,y)} = W_{u3} - W_{u2} \\ = (u_1, u_2, u_3)_{(x+2,y)} - (u_1, u_2, u_3)_{(x+1,y)} \\ \vdots \\ W_{uuk} = (uu_1, uu_2, uu_3)_{(x+k,y)} = W_{u(k+1)} - W_{uk} \\ = (u_1, u_2, u_3)_{(x+k+1,y)} - (u_1, u_2, u_3)_{(x+k,y)} \\ W_{vv1} = (vv_1, vv_2, vv_3)_{(x,y)} = W_{v2} - W_{v1} \\ = (v_1, v_2, v_3)_{(x,y+1)} - (v_1, v_2, v_3)_{(x,y)} \\ W_{vv2} = (vv_1, vv_2, vv_3)_{(x,y+1)} = W_{v3} - W_{v2} \\ = (v_1, v_2, v_3)_{(x,y+2)} - (v_1, v_2, v_3)_{(x,y+1)} \\ \vdots \\ W_{vvl} = (vv_1, vv_2, vv_3)_{(x,y+l)} = W_{v(l+1)} - W_{vl} \\ = (v_1, v_2, v_3)_{(x,y+l+1)} - (v_1, v_2, v_3)_{(x,y+l)}. \quad (11)$$

Based on the above-mentioned process, the normal vector $n_{(x,y)} = [p_1, q_1, r_1]$ of one specific surface point can be determined the same as its partial derivative vectors at the 1st and 2nd order, respectively. Consider

$$W_u = (u_1, u_2, u_3); \quad W_{uu} = (u_{11}, u_{22}, u_{33}); \\ W_v = (v_1, v_2, v_3); \quad W_{vv} = (v_{11}, v_{22}, v_{33}); \quad (12) \\ W_{uv} = (uv_1, uv_2, uv_3).$$

4.3. The Establishment of Light Intensity Reflective Energy Surface. Figure 4 shows some typical image frames for objective

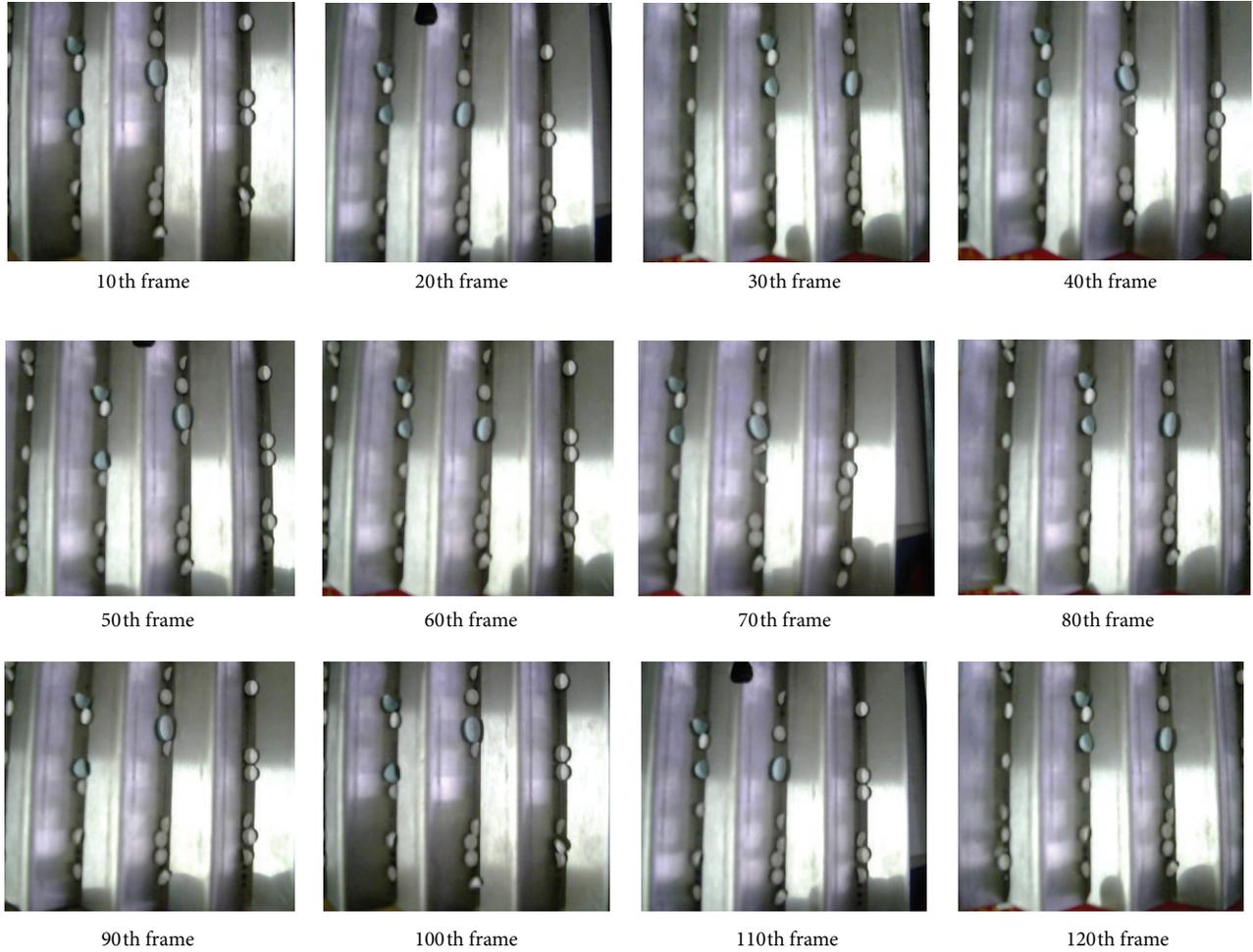


FIGURE 4: Typical image frames obtained from the tablet video for objective tracking in high similarity.

tracking, the amount of drug tablets being seen in the monitoring area was strictly limited under 20 by predetermined setting, in the interest of keeping clear shape description and instantaneous tracking of those objective tablets. Through characteristic statistical analysis of drug tablets in geometric shapes, 12 tablet templates in typical geometric shapes and different laying angles were selected out for the following pattern recognition, with all possible shape morphologies and variant position arrangements during monitoring process being fully included, as Figure 5 demonstrates. Compared with other mathematical criterions, using tablet template facilitates the following determination of their geometrical characteristics in a clearer way. Since energy optimization fitting becomes a frequently used modeling method supported by mathematical programming and structure optimization, it regards the studied reflective energy surface providing the minimum physical distorting energy as its ultimate objective; therefore, various means of energy restriction and employed loadings were applied to achieve this goal in practice [43–45].

Energy optimization method regarding the laminose elastic deformation equation of elastic mechanics as its

mathematical references:

$$E_{\text{surface}} = \iint \left[\begin{aligned} & \left[\alpha_{11}W_u^2 + 2\alpha_{12}W_uW_v + \alpha_{22}W_v^2 \right] \\ & \left[+\beta_{11}W_{uu}^2 + 2\beta_{12}W_{uv}^2 + \beta_{22}W_{vv}^2 \right] \\ & - 2Wf(u, v) \end{aligned} \right] du dv. \quad (13)$$

Here, W denotes the objective surface represented in u, v axes; $W_u, W_v, W_{uu},$ and W_{vv} denote the first-order and second-order derivative vectors of W in u and v axes, respectively. W_{uv} denotes a mixing derivative vector, α and β were given parameters, and f denotes a given vector function to be approximated.

In practical modeling, α and β often denote the referential parameters of material characteristic; f denotes the external employed loading used for energy surface control. Under normal condition, the influence results exerted by $\alpha, \beta,$ and f

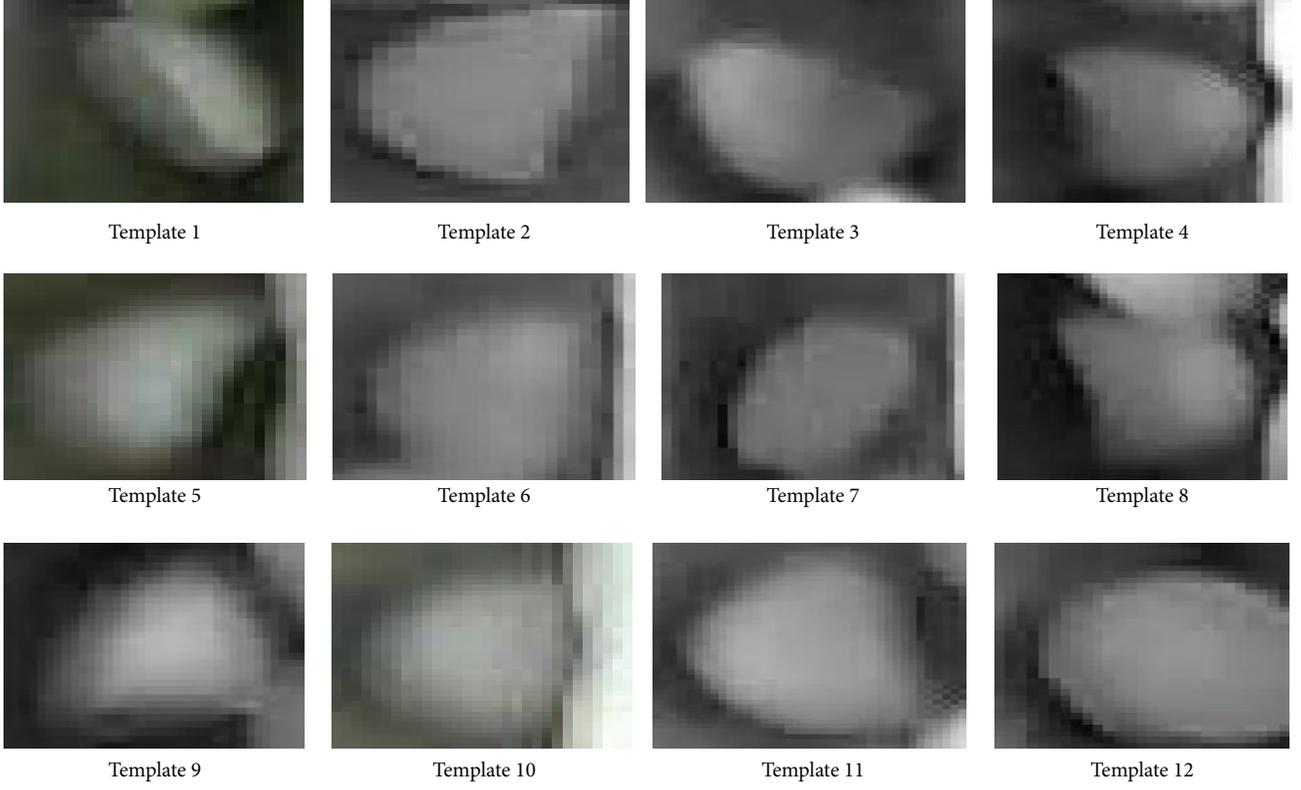


FIGURE 5: Objective tablet templates for image tracking in this experiment.

will be generally neglected and thereafter valuated as 1, 1, and 0, respectively. Consider

$$E_{\text{surface}} = \iint \left[\begin{array}{c} W_u^2 + 2W_u W_v + W_v^2 \\ + W_{uu}^2 + 2W_{vv}^2 + W_{vv}^2 \end{array} \right] du dv. \quad (14)$$

By using the basis function of B-spline surface, the above-mentioned function was transferred into

$$w(u, v) = \sum_{\substack{i=0, mu \\ j=0, mv}} V_{i,j} N_{i,su}(u) N_{j,sv}(v). \quad (15)$$

Here, V denotes a control point of surface $w(u, v)$; $mu + 1$, $mv + 1$ denote the number of control points in two axes; su and sv denote the power of studied surface; $N_{i,su}(u)N_{j,sv}(v)$ are the B-spline basis function defined by su , sv , together with the knot vector denoted by K_U, K_V ($U = [u_0, u_1, u_2, \dots, u_{m+s+1}]$, $V = [v_0, v_1, v_2, \dots, v_{m+s+1}]$) as well. All variables in these equations were defined in u or v directions.

Equation (16) defines the first-order and second-order derivative vectors $w_u(u)$ and $w_{uu}(u)$:

$$w_u(u, v) = \sum_{\substack{i=0, mu \\ j=0, mv}} V_{i,j} N'_{i,su}(u) N_{j,sv}(v);$$

$$w_v(u, v) = \sum_{\substack{i=0, mu \\ j=0, mv}} V_{i,j} N_{i,su}(u) N'_{j,sv}(v);$$

$$w_{uu}(u, v) = \sum_{\substack{i=0, mu \\ j=0, mv}} V_{i,j} N''_{i,su}(u) N_{j,sv}(v);$$

$$w_{vv}(u, v) = \sum_{\substack{i=0, mu \\ j=0, mv}} V_{i,j} N_{i,su}(u) N''_{j,sv}(v);$$

$$w_{uv}(u, v) = \sum_{\substack{i=0, mu \\ j=0, mv}} V_{i,j} B'_{i,su}(u) B'_{j,sv}(v).$$

(16)

Substituting (16) into (14) and letting $f = 0$, then

$$E = \iint_0^1 \sum_{\substack{i=0,mu \\ j=0,mv}} V_{i,j} \sum_{\substack{k=0,mu \\ l=0,mv}} V_{k,l} * \begin{bmatrix} \alpha_{11} N'_{i,su}(u) N_{j,sv}(v) N'_{k,su}(u) N_{l,sv}(v) \\ + 2\alpha_{12} N'_{i,su}(u) N_{j,sv}(v) N'_{k,su}(u) N_{l,sv}(v) \\ + \alpha_{22} N'_{i,su}(u) N_{j,sv}(v) N'_{k,su}(u) N_{l,sv}(v) \\ + \beta_{11} N''_{i,su}(u) N_{j,sv}(v) N''_{k,su}(u) N_{l,sv}(v) \\ + 2\beta_{12} N''_{i,su}(u) N_{j,sv}(v) N''_{k,su}(u) N_{l,sv}(v) \\ + \beta_{22} N''_{i,su}(u) N_{j,sv}(v) N''_{k,su}(u) N_{l,sv}(v) \end{bmatrix} du dv. \quad (17)$$

Then,

$$S_{i,j,k,l} = \iint_0^1 \begin{bmatrix} \alpha_{11} N'_{i,su}(u) N_{j,sv}(v) N'_{k,su}(u) N_{l,sv}(v) \\ + 2\alpha_{12} N'_{i,su}(u) N_{j,sv}(v) N'_{k,su}(u) N_{l,sv}(v) \\ + \alpha_{22} N'_{i,su}(u) N_{j,sv}(v) N'_{k,su}(u) N_{l,sv}(v) \\ + \beta_{11} N''_{i,su}(u) N_{j,sv}(v) N''_{k,su}(u) N_{l,sv}(v) \\ + 2\beta_{12} N''_{i,su}(u) N_{j,sv}(v) N''_{k,su}(u) N_{l,sv}(v) \\ + \beta_{22} N''_{i,su}(u) N_{j,sv}(v) N''_{k,su}(u) N_{l,sv}(v) \end{bmatrix} du dv. \quad (18)$$

Thus, (17) will be simplified into

$$E = \sum_{\substack{i=0,mu \\ j=0,mv}} V_{i,j} \sum_{\substack{k=0,mu \\ l=0,mv}} V_{k,l} * S_{i,j,k,l}. \quad (19)$$

Here, $S_{i,j,k,l}$ denotes the integrating operation of one-known basis function; thus, the mathematical model of energy optimization surface can be transferred into the quadratic function of control vertexes V [46]. Based on this surface fitting method and the tablet templates shown in Figure 5, their corresponding reflective energy surfaces of light intensity can then be established with the resultant details being demonstrated in Figure 6.

4.4. The Application of ANN for Recognizing Surface Properties. In this study, back propagation (BP) learning algorithm, which has a unique learning principle, generally called delta rule, is used. As we all know, the back propagation learning algorithm is a common method for training artificial neural networks. Firstly, forward propagation of input vectors was realized to generate the propagation output activations, and then ANN uses the pattern target to generate the deviation of all output and hidden neurons. Based on this precondition, multiplying the calculation between the output deviation and the input activation was operated for obtaining the gradient vectors of those weigh values, which results into the percentage subtraction of their gradient vectors. Repeat these steps again until the performance of ANN network reaches a satisfied level. The three layers of the network architecture include the input layer, middle layer (hidden layer), and output layer. Layers include several processing units known as

neurons. They connected with each other by variable weights to be determined. In this network, the input layer receives information from external source and passes this information to the network for data processing. The middle layer receives data from the input layer and does all information analysis. The output layer receives the processed information from the middle layer and sends the results to an external receptor [47].

A neuron in ANN network produces its input by processing the net input through nonlinear activation (transfer) function. The sigmoidal activation function is the most utilized one, which updates the weight and derivative values of ANN according to the resilient back propagation algorithm; therefore, it is usually trained for updating ANN networks. In this training algorithm, the update value for ANN weight is increased whenever the derivative value of its performance function has the same sign for two successive calculate iterations; on the other hand it is decreased according to the derivative value with respect to that weight changes sign from the previous iteration. If the derivative value is zero, then the update value for ANN weight remains the same. If the objective weight value continues to change in the same direction for several training iterations, then the magnitude of the weight change will be increased accordingly. In this experiment, we supposed that the input layer has NI neurons, with the middle layer and output layer having NJ and NK neurons, respectively, whereas the input elements for the j th neuron at the middle layer can then be described as

$$\text{net}_j = \sum_{i=1}^{NI} w_{ij} O_i; \quad j = 1, 2, \dots, NJ. \quad (20)$$

Here, w_{ij} denotes the weight calibrating from the i th neuron on the input layer to the j th neuron on the middle layer, with O_i demonstrating the output result of the i th neuron on the input layer. Meanwhile, the resultant input of the k th neuron on the output layer can also be identified as

$$\text{net}_k = \sum_{j=1}^{NJ} w_{jk} O_j; \quad k = 1, 2, \dots, NK. \quad (21)$$

Here, w_{jk} denotes the weight calibrating from the j th neuron on the middle layer to the k th neuron on the output layer, with O_j demonstrating the output result of the k th neuron on the middle layer. Through substituting the input neurons of the middle and output layer, as denoted by net_j

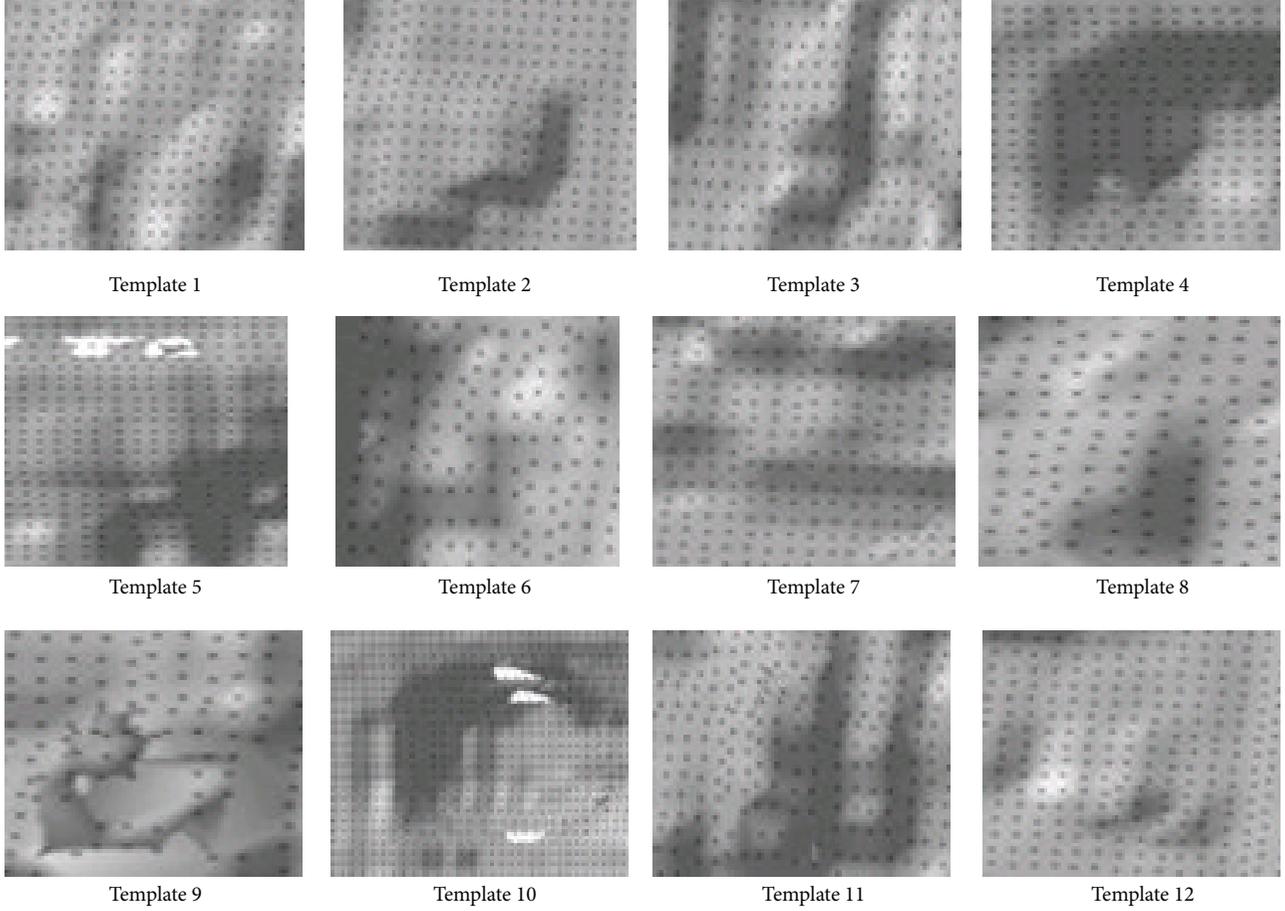


FIGURE 6: The established reflective energy surfaces of objective tablet templates.

and net_k , into (20), the output results of these three layers can be illustrated as follows:

$$\begin{aligned}
 O_i &= net_i = x_i, \\
 O_j &= f_j(net_j, \theta_j) = \frac{1}{1 + e^{-(net_j - \theta_j)}}, \\
 O_k &= f_k(net_k, \theta_k) = \frac{1}{1 + e^{-(net_k - \theta_k)}}.
 \end{aligned} \tag{22}$$

Here, θ_j and θ_k denote the output threshold value used for assessing the computation results of the j th neuron on the middle layer and the k th neuron on the output layer, respectively. In this experiment, the gradient-based learning rule was used to train this established network, with the minimum mean square error (MSE) between the computation results caused by the network output and tablet template being predetermined as the ultimate train objective. The template amount for training was supposed to be P , and the input elements can then be denoted by x_1, x_2, \dots, x_p , with the output elements being denoted by y_1, y_2, \dots, y_p . Here, this experiment used t_1, t_2, \dots, t_p to denote the data vector

consisted by training samples. Thus, MSE of the p th sample can be obtained as follows:

$$E^p = \frac{1}{2} \sum_{k=1}^{NK} (t_k^p - y_k^p)^2. \tag{23}$$

Here, t_k^p denotes the reference value of the p th sample for the k th output neuron, while y_k^p denotes the practical output in the same condition. Therefore, the weight on the output layer can then be adjusted to

$$\begin{aligned}
 \Delta w_{jk}(n+1) &= \eta \varepsilon_k^p O_j^p + \alpha \Delta w_{jk}(n); \\
 \varepsilon_k^p &= (t_k^p - y_k^p) f'_k(net_k^p).
 \end{aligned} \tag{24}$$

Here, η represents the learning rate, α denotes the momentum factor, and δ_{pk} denotes the deviation value between t_k^p and y_k^p , which shows the difference between the reference value and the practical output of the p th sample

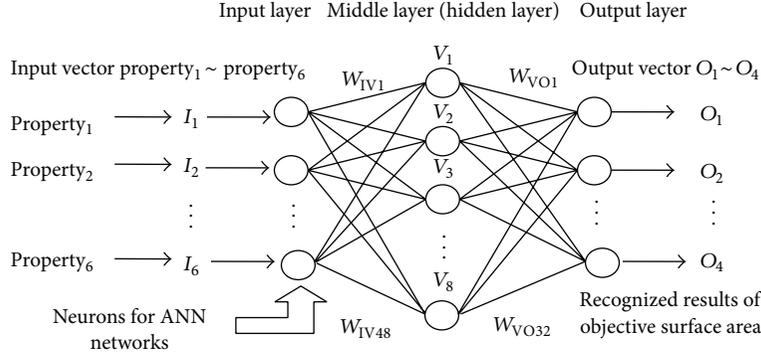


FIGURE 7: Schematic illustration of ANN used for surface property recognition and tablet tracking.

caused by the k th output neuron [48]. Therefore, the weights of middle layer can then be adjusted into

$$\Delta w_{ij}(n+1) = \eta \varepsilon_j^p O_i^p + \alpha \Delta w_{ij}(n); \quad (25)$$

$$\varepsilon_j^p = f_j'(\text{net}_j^p) \sum_{k=1}^{NK} \delta_{pk} w_{jk}.$$

When focusing on the ANN Testing process, it should be noted that the idea which we have used for testing neural networks and generally applying test vectors is that we use a multicase test vector to determine the decision strength of this neural network. For the final evaluation, we captured 100 images describing the random movement cases of drug tablets. Such a large testing case amount was used to minimize the uncertainty of the objective classifier's performance estimation in this experiment. Then, according to the reflective energy surfaces of those drug tablets, we input their corresponding property vectors to the established ANN and check the obtained output vectors in sequence. Based on the testing operations, this neural network reaches the desired performance; thus, a high improvement in the video detection of drug tablets can be observed.

The weight values in ANN network markedly affect the classification of tablet templates, which keep a close correlation with the computation or tracking errors in the proposed experimental conditions; they were also easily impacted by the arrangement of neural neutrons and input/output vectors as well as weight values provide correlation among network layers, which ensures the feasibility of recognizing the geometrical shapes of tablet template and the property distribution of its reflective energy surface with the self-adaptive control of ANN. As Figure 7 shows, the detailed setup of ANN network is used, and it was observed that, for the purpose of realizing the improvements in network recognition and objective classification, the weights between input and middle layers should be kept in a relatively stable state, which provides a useful tool to markedly eliminate the external error-training interference caused by noisy sample vectors or error distribution gradient of input data. Simultaneously, the weights between middle and output layers should be kept in a radically distinction state, which ensures a clearer

classification of tablet properties, especially when the self-adaptive learning rule was being employed. Based on these preconditions, ANN network was gradually optimized for improving its recognition performance [49, 50].

5. Tracking Process Discussions and Performance Comparisons

With the trained neural network and the surface properties computed by (1)–(6), they were used as the input vector for recognizing the tablet templates:

$$\begin{aligned} \text{Input} &= [I_1, I_2, I_3, I_4, I_5, I_6] \\ &= [\text{property}_1, \text{property}_2, \text{property}_3, \\ &\quad \text{property}_4, \text{property}_5, \text{property}_6]. \end{aligned} \quad (26)$$

Then, the output vector can be predetermined as $[O_1, O_2, O_3, O_4]$ by using forward propagation algorithm; the maximum possible reiteration times were determined as 15000, in order to prevent the case of training threshold does not be met; the node number of the hidden layer was 2, the learning rate coefficient was supposed to be 0.45, the momentum factor was supposed to be 0.065, the train step was supposed to be 0.15, and the interval illustration factor for recognition process was predetermined as 30. The error function was determined by (27), and it can also be employed to describe the ANN performance capability:

$$E_{\text{all}} = \sum_{p=1}^P E^p. \quad (27)$$

This index examines if the prerequisite was satisfied or not: $E_{\text{all}} < \varepsilon$? If it was confirmed, then the training processes of ANN can be terminated or it will be iterated for the next time until this prerequisite is satisfied.

When discussing the determination of ε , through establishing the weights between network layers, then

$$\varepsilon = 1 - \xi \sqrt{\sum_{j=1}^{12} |\Delta w_{kj} (1 - \delta_{pk})|^p}. \quad (28)$$

TABLE 1: The original and corrected weight values of the established ANN in this experiment.

The original weight value of ANN before training								
$W_{IV1} \sim W_{IV8}$	2.358	1.298	-3.114	5.558	-6.025	-0.553	3.339	6.554
$W_{IV9} \sim W_{IV16}$	-5.159	0.587	1.265	-3.558	-5.554	7.663	2.663	-2.226
$W_{IV17} \sim W_{IV24}$	0.668	-5.487	-6.587	9.558	3.214	3.271	1.098	-3.554
$W_{IV25} \sim W_{IV32}$	6.558	1.245	-6.584	-5.145	9.552	1.118	2.337	6.524
$W_{IV33} \sim W_{IV40}$	-3.224	-4.859	1.548	4.887	3.225	-3.876	8.098	-6.554
$W_{IV41} \sim W_{IV48}$	2.335	2.982	3.112	-3.442	7.078	1.987	9.003	2.338
$W_{VO1} \sim W_{VO8}$	3.265	2.554	-6.254	9.551	3.025	2.336	-8.476	-2.114
$W_{VO9} \sim W_{VO16}$	1.487	6.584	4.559	-6.002	5.554	4.098	-4.442	6.335
$W_{VO17} \sim W_{VO24}$	-6.954	8.885	3.654	8.201	-6.002	2.221	8.098	-6.598
$W_{VO25} \sim W_{VO32}$	6.547	-6.548	6.554	8.887	-6.254	-2.189	3.228	0.621
$W_{VO29} \sim W_{VO32}$	2.337	-3.072	4.976	-0.887	7.765	2.037	3.047	5.887
The corrected weight value of ANN after training								
$W_{IV1} \sim W_{IV8}$	3.334	4.228	2.998	2.882	2.332	3.728	2.391	3.311
$W_{IV9} \sim W_{IV16}$	-0.887	6.447	-8.472	8.472	8.482	5.772	5.339	4.591
$W_{IV17} \sim W_{IV24}$	2.337	6.992	-1.993	6.472	3.442	6.102	5.281	5.993
$W_{IV25} \sim W_{IV32}$	5.823	9.073	4.378	3.448	5.312	9.082	-0.998	-2.921
$W_{IV33} \sim W_{IV40}$	9.008	2.774	7.461	5.729	7.422	5.698	-9.003	3.549
$W_{IV41} \sim W_{IV48}$	1.894	-8.372	8.472	5.773	9.082	4.332	-5.211	5.281
$W_{VO1} \sim W_{VO8}$	-8.774	0.987	4.228	8.728	9.228	6.937	7.381	-6.391
$W_{VO9} \sim W_{VO16}$	5.662	8.443	-0.879	8.227	-8.492	2.184	1.391	-5.112
$W_{VO17} \sim W_{VO24}$	8.447	9.622	2.398	2.173	7.472	-7.471	4.291	5.295
$W_{VO25} \sim W_{VO32}$	0.182	8.446	1.094	7.422	4.442	4.552	5.281	-3.591
$W_{VO29} \sim W_{VO32}$	-4.372	6.364	4.553	-0.921	7.311	6.082	8.391	-8.422

TABLE 2: The reflective energy surface properties of representative templates on the 35th image frame, which were employed as the input vector of ANN for pattern recognition, the same as the other surface property groups.

The input values of the surface properties						
	<i>Property 1</i>	<i>Property 2</i>	<i>Property 3</i>	<i>Property 4</i>	<i>Property 5</i>	<i>Property 6</i>
1	3587.23	548.52	1587.9	10155.4	956.6	0.2669
2	1035.69	669.52	2036.5	12336.5	875.2	0.3022
3	2598.32	716.55	3544.5	10269.5	665.2	0.3654
4	4782.61	215.84	3055.8	9854.6	365.4	0.1485
5	5548.32	985.36	6248.2	9965.5	302.1	0.3954
6	3699.14	888.36	5598.7	7825.5	954.6	0.3022
7	6685.26	441.56	6694.8	6359.4	865.7	0.2035
8	1334.69	332.58	6025.8	9806.5	1012.5	0.3144
9	4026.54	698.22	4785.9	11025.4	1124.5	0.6022
10	4026.57	725.36	6698.5	12004.5	1325.9	0.3598
11	9833.21	889.65	4755.2	7985.6	968.8	0.4752
12	5476.32	102.54	6321.5	9586.6	1024.5	0.3301

Here, ξ denotes a distinguishing coefficient located in $[0, 1]$. Δw_{kj} denotes the self-adaptive deviation value between the weight values calibrated from the j th neuron on the middle layer to the k th neuron on the output; δ_{pk} denotes the deviation value between t_k^p and y_k^p . By using this threshold screening mechanism, the accuracy of pattern recognition can be ensured, and thereafter the overtraining phenomenon in ANN network could also be prevented. The greater this threshold value is, clearer the objective recognition result would be.

Table 1 demonstrates the original and corrected weights of ANN. It is noteworthy that the original weights were tentatively selected with reference to (20)-(21), in the hope of reducing the computation deviation between the desired and practical output vectors to the minimum scale [51, 52]. Table 2 demonstrates the energy surface properties of tablet templates on the 35th image frame. In the interest of clearly describing the detailed process of tablet template recognition and its position justification, a program flow schematic was provided, with the details shown in Figure 8. Table 3

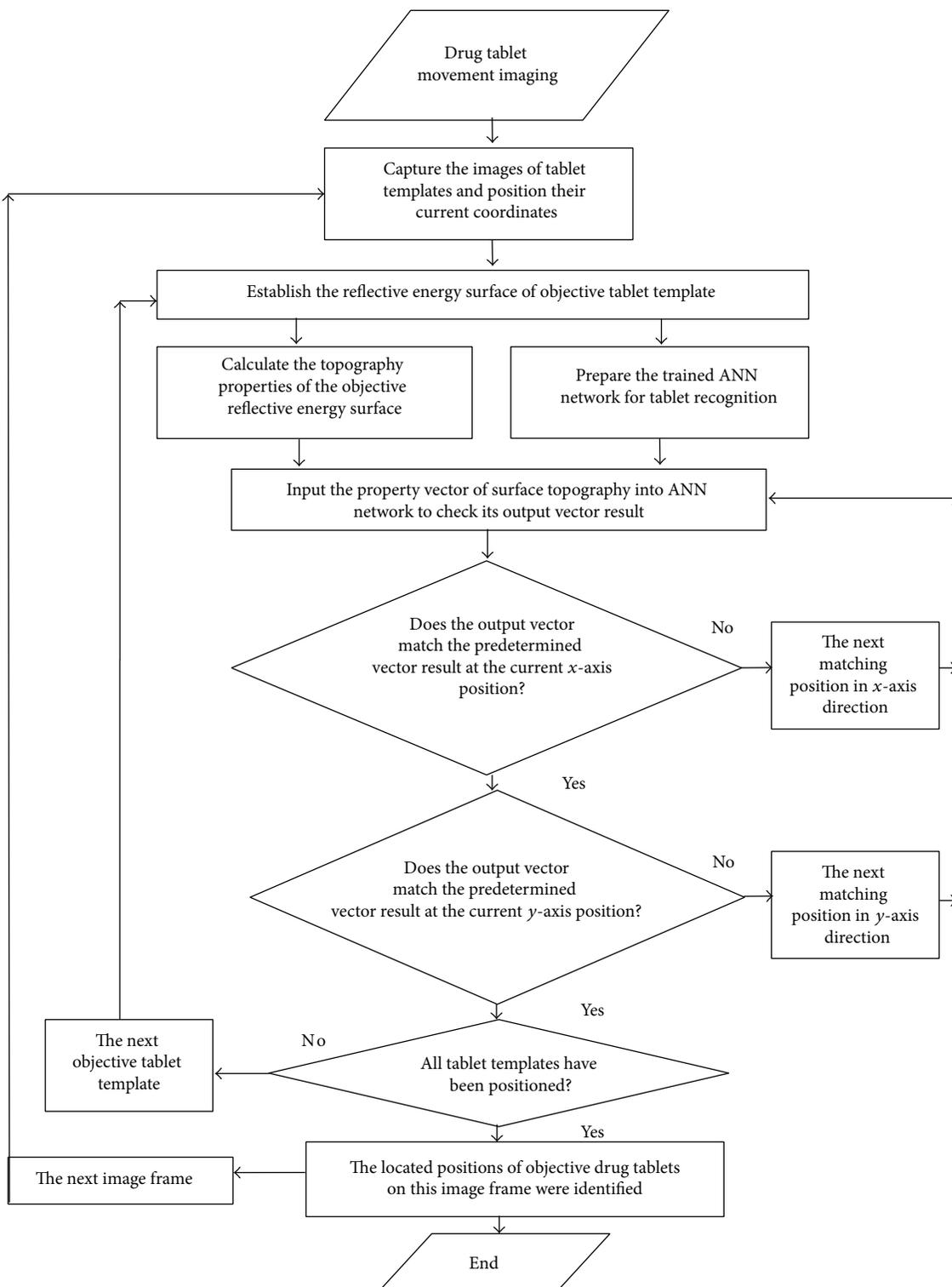


FIGURE 8: The program flow schematic of tablet template recognition and position justification.

describes the recognition results of surface properties; similar operations were repeatedly operated for times. Output vector composed by four digits is used to denote the specific objective template. For example, $[0\ 0\ 0\ 1]$ represents template 1,

$[0\ 0\ 1\ 0]$ represents template 2, $[0\ 0\ 1\ 1]$ represents template 3, $[0\ 1\ 0\ 0]$ represents template 4, ..., and $[1\ 1\ 0\ 0]$ represents template 12, with the mean values of recognition results based on repeatedly experiments being used. Through

TABLE 3: Representative recognition results of reflective energy surface properties for tracking those tablet templates on the 35th image frame.

Objective tablets	The recognized results (the obtained number vectors were the mean values based on 10 times of repeated experiments)				Recognition ratio	
Template 1	[0.0053 0.0024 0.0066 1.0035]				Practical result vector	95.6%
	[0 0 0 1]				Predetermined vector	
Template 2	[0.0014 0.0012 9.9971 0.0032]				Practical result vector	94.6%
	[0 0 1 0]				Predetermined vector	
Template 3	[0.0032 -0.0011 1.0036 1.0042]				Practical result vector	95.2%
	[0 0 1 1]				Predetermined vector	
Template 4	[-0.0018 1.0051 0.0041 0.0017]				Practical result vector	92.9%
	[0 1 0 0]				Predetermined vector	
Template 5	[-0.0017 9.9981 -0.0056 1.0088]				Practical result vector	93.3%
	[0 1 0 1]				Predetermined vector	
Template 6	[0.0055 1.0006 9.9968 0.0016]				Practical result vector	95.1%
	[0 1 1 0]				Predetermined vector	
Template 7	[-0.0055 1.0041 0.9988 1.0044]				Practical result vector	98.2%
	[0 1 1 1]				Predetermined vector	
Template 8	[1.0058 0.0066 0.0014 0.0058]				Practical result vector	93.9%
	[1 0 0 0]				Predetermined vector	
Template 9	[1.0055 0.0041 0.0088 0.9944]				Practical result vector	95.6%
	[1 0 0 1]				Predetermined vector	
Template 10	[1.0043 0.0022 0.9997 -0.0035]				Practical result vector	93.9%
	[1 0 1 0]				Predetermined vector	
Template 11	[1.0013 0.0025 0.9927 -0.0089]				Practical result vector	94.4%
	[1 0 1 1]				Predetermined vector	
Template 12	[1.0036 1.0022 0.0088 -0.0071]				Practical result vector	97.5%
	[1 1 0 0]				Predetermined vector	

data comparison, the obtained results keep a close difference with those predetermined one, which undoubtedly confirms their effectiveness and accuracy. As one typical pattern recognition, Figure 9 shows the coordinate positions of surface areas best matching their corresponding tablet templates on the 35th image frame, with the black boxes showing the probable positions of the targeted tablets to be studied. In this figure, the most possible matching positions have been magnified and displayed in surrounding images for clearer description; the coordinate positions of template centers were highlighted by the focus marks. For better illustration of the tablet tracking result, 10 key tablets have been labeled with numbers, especially those close ones that were highly focused, as Figure 10 shows. We can observe that the practical operation with this algorithm can be smoothly identified, and those tablets in close position can also be distinguished clearly. Besides, as the interval distance between two adjacent dashed lines was predetermined as 30 mm, the instantaneous moving speeds of objective tablets can be calculated.

In order to prove the accuracy and validity of this newly proposed method, several commonly used tracking methods were investigated with reference to the published literatures, including the maximum likelihood estimation, the Bayesian

estimation, the prior probability-density statistics, the non-prior information estimation, the invariant prior probability-density statistics, the Jeffreys statistics, the box-tracing-based method, and the sensor tracing method. Typical evaluation indexes were employed when an identical experimental condition has been considered: such as the computation time (the accurate time consumption in which the whole experiment uses the identical tracing platform; this research measures the computation time by means of the automated time counting instrument), the distributed Hash degree (the decentralized distributed complexity level that provides a lookup assessment index for algorithm evaluation similar to traditional Hash table, with a small value associated with the high-efficiently structure of tracing algorithm, and vice versa. This research calculates them through the performance process of one given algorithm and then makes a definite assessment on its performance capabilities as respected), the tracking error (the tracking ratio between the amount of error-tracking objectives to total ones, by which the tracing accuracy can be quantitatively identified. It can be calibrated by counting the amount of wrong-located tablets and total ones, resp., and then the ratio between them can be used as the tracking error), and the computation storage (the internal memory

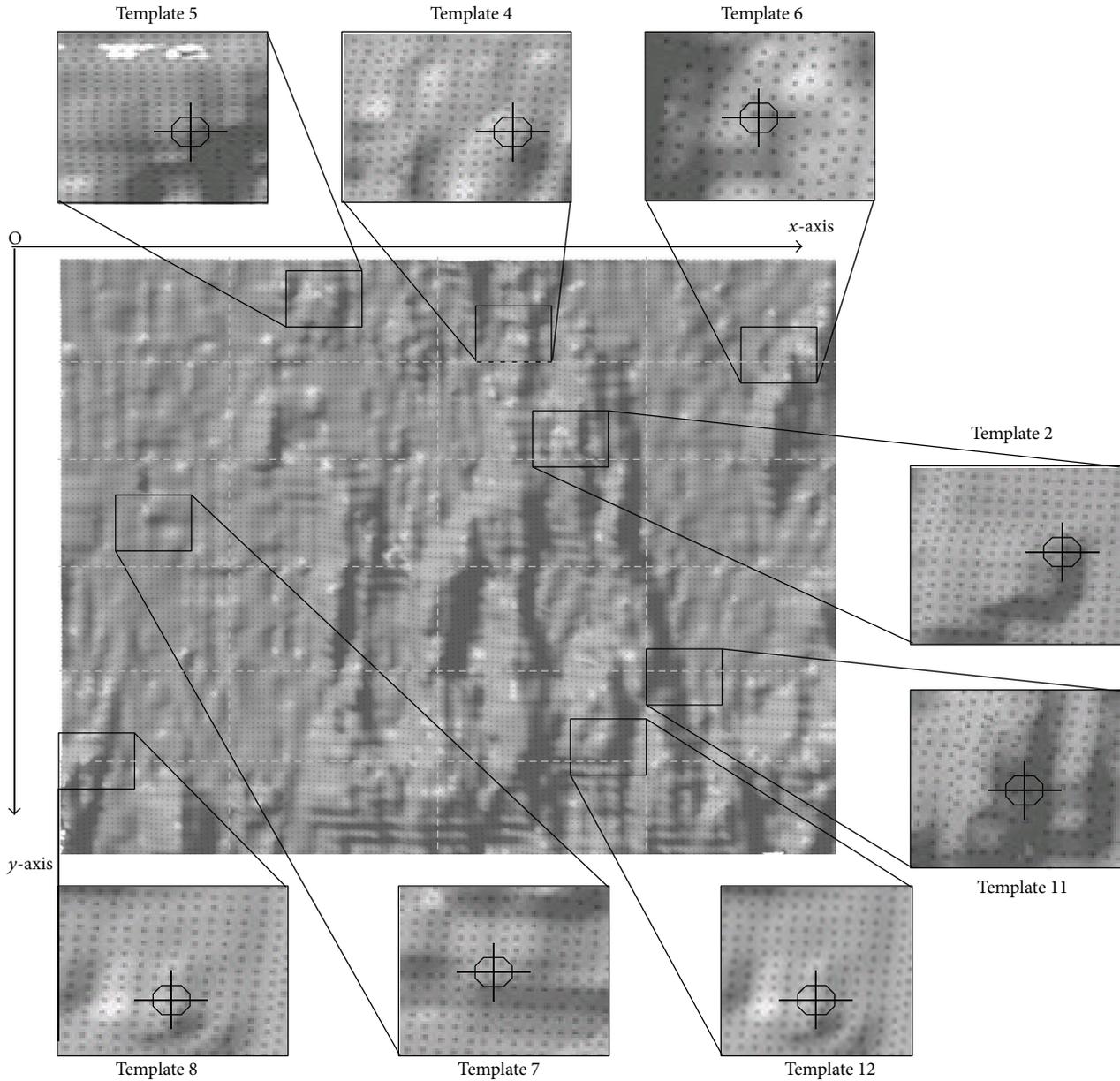


FIGURE 9: The position distributions of surface areas to best match their corresponding tablet templates on the 35th image frame.

capacity used for the whole experimental computations; we calibrate them from the real-time performance indicator of the video tracking system that is being involved during the whole process of tablet tracking) [44–48]. After determining these evaluation indexes, their average performance capabilities were compared to each other, with the results shown in Table 4 after data equalization and value regularization. Since this experiment is repeatable, we take a detailed notes about the practical performances for 10, 20, 30, 40, and 50 times, and thereafter a statistical evaluation on their mean index values can be made by employing those mentioned algorithms, respectively. It can be seen from Figure 11 that, with the increment of experiment times, most of these tracking algorithms show a growth tendency in the average

computation times. But it is noteworthy that a relative-stable performance capability can be ensured by using our newly proposed algorithm, which is kept at about 0.443 s throughout the whole process rather than changing radically. Similar performance evaluations can also be observed in the variation tendency schematics of Figure 12 with the average Hash degree (from 59.8% to 5.66%), the average tracking error (from 2.3% to 2.445%) in Figure 13, and the average computation storage (from 797.5 kb to 804.5 kb) in Figure 14.

From Table 4, it can be observed that the maximum likelihood estimation or the sensor tracing method has a relative superiority in the tracking error and distributed Hash degree; they can be widely used to investigate a stable statistical condition, by concerning the dynamic moving characteristics

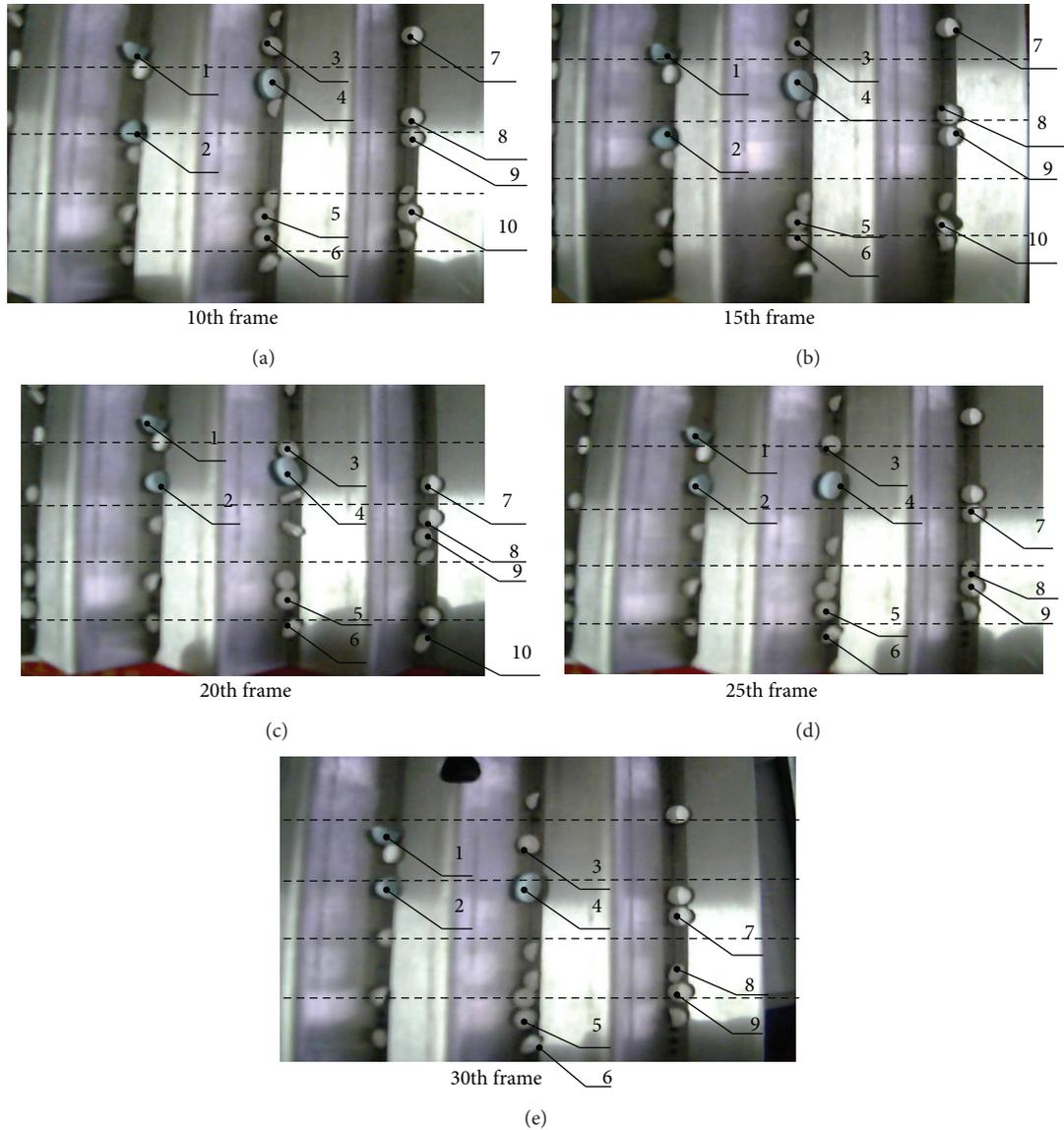


FIGURE 10: Representative video tracking results on different image frames, with the key objective tablets that have been labelled by numbers to show their moving processes in details, respectively.

of identical tablets in huge cluster; the Bayesian estimation obtains a good performance result in the computation time; therefore, it will be more suitable to fast assess a simpler mutual relationship mechanism for the approximate shape properties of those drug tablets; the prior probability-density statistics method or the Jeffreys statistics approach have excellent computation performances when the tracking error or distributed Hash degree were highly emphasized, which ensure the accurate demonstrations of data relationship analysis and thereafter make a series of remarkable progresses when compared with other traditional methods; the nonprior information estimation method and the invariant prior probability-density statistics approach have extraordinary capabilities in the distributed Hash degree, which show a more precise algorithm complexity of data distribution in multidimensional statistical domain. Finally a

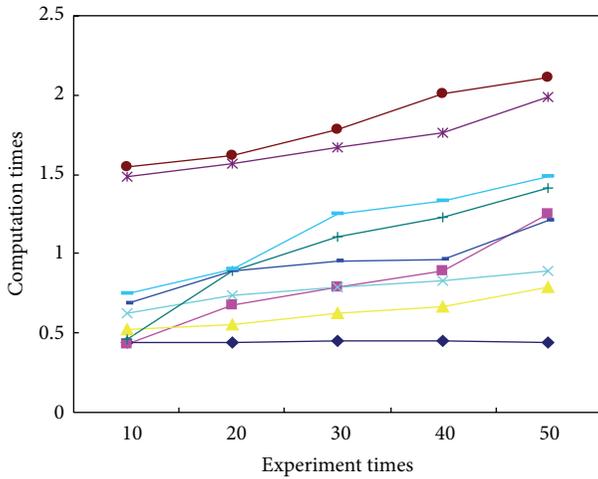
good performance evaluation result can be obtained by using this newly proposed algorithm, especially in the computation time, computation storage, or distributed Hash degree. Performance comparison proves its validation and efficiency when similar shape or identical size of objective tablets should be taken into account, and the working estimation on how well these alternative methods perform on average can be facilitated as well.

6. Conclusions

A new image tracking approach for high similarity drug tablets based on light intensity reflective energy and ANN recognition has been investigated. Considering the practical condition and precision requirement, the reflective energy

TABLE 4: Performance comparisons during drug tablet tracing by using different methods.

Method	Average computation time	Average distributed Hash degree	Average tracking error	Average computation storage
New algorithm	0.443 s	5.66%	2.445%	797.5 kb
Maximum likelihood estimation	0.678 s	4.90%	2.147%	883.1 kb
Bayesian estimation	0.558 s	5.03%	2.798%	1566.6 kb
Prior probability-density statistics	0.734 s	7.65%	3.446%	1251.4 kb
Nonprior information estimation	1.565 s	7.35%	3.709%	785.2 kb
Jeffreys statistics	1.623 s	6.44%	2.366%	975.1 kb
Invariant prior probability-density statistics	0.889 s	6.72%	2.477%	1032.5 kb
Box-tracing-based method	0.894 s	7.91%	3.674%	1377.3 kb
Senor tracing method	0.905 s	8.60%	2.668%	558.5 kb



◆ New algorithm
 ■ Maximum likelihood estimation
 ▲ Bayesian estimation
 ✕ Prior probability-density statistics
 * Nonprior information estimation
 ● Jeffreys statistics
 + Invariant prior probability-density statistics
 ▬ Box-tracing-based method
 — Sensor tracing method

FIGURE 11: The variation tendency of average computation times for different experimental times with different objective tracking algorithms denoted by colored lines, the same as follows.

surfaces were established for modeling the high similarity geometric characteristics and surface topography of those targeted drug tablets, which results into the maximum deletion of external optical signal interferences or calibration error in practice, and simultaneously an accurate description of tablet objectives in the light intensity domain can also be ensured. Thereafter, ANN was used for recognizing those studied objective tablets with the computed surface properties, and then their instantaneous positions on one image frame can be identified clearly. Through repeating these steps on the sequential image frames, a series of tablet tracking results can be obtained. After being compared

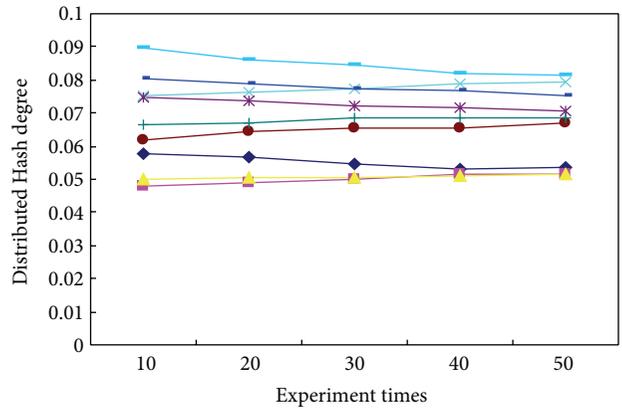


FIGURE 12: The variation tendency of average distributed Hash degrees for different experimental times.

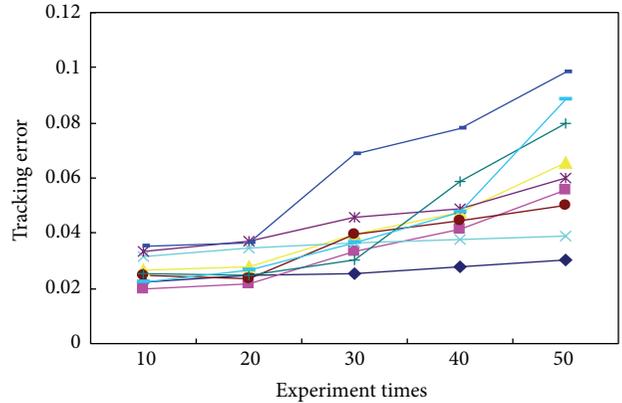


FIGURE 13: The variation tendency of average tracking errors for different experimental times.

with other tracking methods in performance indexes and determined results, it can be learned that this newly proposed method solves numerous difficult problems characterized by monitoring precision, complicated calculations or tracking errors, and other signal interferences caused by mathematical calibration or pattern recognition. Therefore, the quantitative

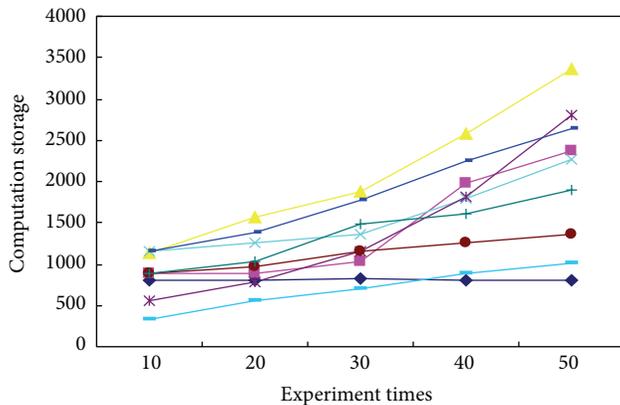


FIGURE 14: The variation tendency of average computation storages for different experimental times.

analysis of the objective tracking can be simplified, and the searching precision or computation efficiency can also be greatly improved, accordingly. Experimental computation and result analysis confirmed the validation and accuracy of this new method; new research ideas for real-time objective recognition or tablet tracking can be provided as well.

Conflict of Interests

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

Acknowledgments

The author acknowledges the funding of the following science foundations: National Natural Science Foundation of China (51205073), China National Spark Program (2013GA780063), China Postdoctoral Science Foundation Funded Project (2013T60797, 2012M510197), the Water Resource Science and Technology Program of Guangdong Province of China (2012-11), the Project of Department of Education of Guangdong Province (2013KJ CX0142), the Science and Technology Project of Guangzhou City (2012J4100053), the Foundation Projects of the State Key Laboratory of Fluid Power Transmission and Control (GZKF-201201), the Traction Power State Key Laboratory (TPL1311), and the National Engineering Research Centre of Near-Net-Shape Forming for Metallic Materials (2012007). They are also appreciated for supporting this work.

References

- [1] W. Wang and L. Wang, "An image processing method for grains counting," *Communications in Computer and Information Science*, vol. 326, no. 1, pp. 181–185, 2012.
- [2] T. Fukuda, H. Ishigami, S. Shiotani et al., "Recognition and counting method of mammalian cells on micro—carrier using image processing and artificial neural network," *Animal Cell Technology: Basic & Applied Aspects*, vol. 4, pp. 185–191, 1992.
- [3] D. Dill, A. Scholz, M. Gül, and B. Wolf, "Methods for counting cells supported by digital image processing," in *14th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics*, vol. 20 of *IFMBE Proceedings*, pp. 493–496, 2008.
- [4] H. Kuba, K. Hotta, and H. Takahashi, "Automatic particle detection and counting by one-class SVM from microscope image," *Lecture Notes in Computer Science*, vol. 5507, no. 2, pp. 361–368, 2009.
- [5] J. González, C. Galindo, V. Arevalo, and G. Ambrosio, "Applying image analysis and probabilistic techniques for counting olive trees in high-resolution satellite images," *Lecture Notes in Computer Science*, vol. 4678, pp. 920–931, 2007.
- [6] V. Guarino, A. Guaccio, P. A. Netti, and L. Ambrosio, "Image processing and fractal box counting: user-assisted method for multi-scale porous scaffold characterization," *Journal of Materials Science: Materials in Medicine*, vol. 21, no. 12, pp. 3109–3118, 2010.
- [7] A. R. S. Marçal, "Alternative methods for counting overlapping grains in digital images," *Lecture Notes in Computer Science*, vol. 5112, pp. 1051–1060, 2008.
- [8] J.-W. Kim, K.-S. Choi, B.-D. Choi, J.-Y. Lee, and S.-J. Ko, "Real-time system for counting the number of passing people using a single camera," in *Pattern Recognition*, vol. 2781 of *Lecture Notes in Computer Science*, pp. 466–473, Springer, Berlin, Germany, 2003.
- [9] R. M. Luque-Baena, J. M. Ortiz-De-Lazcano-Lobato, E. López-Rubio, E. Domínguez, and E. J. Palomo, "A competitive neural network for multiple object tracking in video sequence analysis," *Neural Processing Letters*, vol. 37, no. 1, pp. 47–67, 2013.
- [10] R. M. Luque, J. M. Ortiz-De-Lazcano-Lobato, E. Lopez-Rubio, and E. J. Palomo, "Object tracking in video sequences by unsupervised learning," *Lecture Notes in Computer Science*, vol. 5702, pp. 1070–1077, 2009.
- [11] Y. S. Wong, C. W. Tam, S. M. Lee, C. P. Chan, and H. Fu, "Video-base people counting and gender recognition," *Lecture Notes in Computer Science*, vol. 7332, no. 2, pp. 228–235, 2012.
- [12] S. Buczkowski, S. Kyriacos, F. Nekka, and L. Cartilier, "The modified box-counting method: analysis of some characteristic parameters," *Pattern Recognition*, vol. 31, no. 4, pp. 411–418, 1998.
- [13] Z.-B. Wang, H.-W. Hao, Y. Li, X.-C. Yin, and S. Tian, "Pedestrian analysis and counting system with videos," in *Neural Information Processing*, vol. 7667 of *Lecture Notes in Computer Science*, pp. 91–99, Springer, Berlin, Germany, 2012.
- [14] O. C. Agustin and B.-J. Oh, "People counting using object detection and grid size estimation," *Communications in Computer and Information Science*, vol. 265, no. 1, pp. 244–253, 2011.
- [15] C. C. Lien, Y. T. Tsai, M. H. Tsai, and L. G. Jang, "Vehicle counting without background modeling," *Lecture Notes in Computer Science*, vol. 6523, pp. 446–456, 2011.
- [16] W. Zhan and Z. Luo, "System design of real time vehicle type recognition based on video for windows (AVI) files," *Communications in Computer and Information Science*, vol. 135, no. 2, pp. 681–686, 2011.
- [17] H. H. Park, H. G. Lee, S. Noh, and J. Kim, "Development of a block-based real-time people counting system," *Lecture Notes in Computer Science*, vol. 4109, pp. 366–374, 2006.
- [18] A. Głowacz, Z. Mikrut, and P. Pawlik, "Video detection algorithm using an optical flow calculation method," *Communications in Computer and Information Science*, vol. 287, pp. 118–129, 2012.

- [19] H. H. Park, H. G. Lee, S.-I. Noh, and J. Kim, "An area-based decision rule for people-counting systems," in *Multimedia Content Representation, Classification and Security*, vol. 4105 of *Lecture Notes in Computer Science*, pp. 450–457, Springer, Berlin, Germany, 2006.
- [20] D. H. Hong, "Fuzzy measures for a correlation coefficient of fuzzy numbers under T_w (the weakest t -norm)-based fuzzy arithmetic operations," *Information Sciences*, vol. 176, no. 2, pp. 150–160, 2006.
- [21] D. G. Park, Y. C. Kwun, and J. H. Park, "Correlation coefficient of interval-valued intuitionistic fuzzy sets and its application to multiple attribute group decision making problems," *Mathematical and Computer Modelling*, vol. 50, no. 9-10, pp. 1279–1293, 2009.
- [22] J. Ye, "Fuzzy decision-making method based on the weighted correlation coefficient under intuitionistic fuzzy environment," *European Journal of Operational Research*, vol. 205, no. 1, pp. 202–204, 2010.
- [23] C. Son, "Correlation between learning (probability of success) and fuzzy entropy in control of intelligent robot's part macro-assembly tasks with sensor fusion techniques," *Robotics and Computer-Integrated Manufacturing*, vol. 23, no. 1, pp. 47–62, 2007.
- [24] R. M. Luque-Baena, D. Elizondo, E. López-Rubio, E. J. Palomo, and T. Watson, "Assessment of geometric features for individual identification and verification in biometric hand systems," *Expert Systems with Applications*, vol. 40, no. 9, pp. 3580–3594, 2013.
- [25] Z. Liang, X. Liu, B. Ye, and Y. Wang, "Performance investigation of fitting algorithms in surface micro-topography grinding processes based on multi-dimensional fuzzy relation set," *International Journal of Advanced Manufacturing Technology*, vol. 67, no. 9, pp. 2779–2798, 2013.
- [26] Z. Liang, B. Ye, Y. Wang, and R. K. Brauwer, "Three-dimensional fuzzy influence analysis of fitting algorithms on integrated chip topographic modeling," *Journal of Mechanical Science and Technology*, vol. 26, no. 10, pp. 3177–3191, 2012.
- [27] C. Guan, C. Juang, and G. Chen, "Face localization using fuzzy classifier with wavelet-localized focus color features and shape features," *Digital Signal Processing*, vol. 22, no. 6, pp. 961–970, 2012.
- [28] N. Chen, Z. Xu, and M. Xia, "Correlation coefficients of hesitant fuzzy sets and their applications to clustering analysis," *Applied Mathematical Modelling*, vol. 37, no. 4, pp. 2197–2211, 2013.
- [29] A. Dutta, A. Kar, and B. N. Chatterji, "A new approach to corner matching from image sequence using fuzzy similarity index," *Pattern Recognition Letters*, vol. 32, no. 5, pp. 712–720, 2011.
- [30] C. G. da Costa, B. Bedregal, and A. D. Dória Neto, "Atanassov's intuitionistic fuzzy probability and Markov chains," *Knowledge-Based Systems*, vol. 43, pp. 52–62, 2013.
- [31] J. E. B. Maia, G. A. Barreto, and A. L. V. Coelho, "Visual object tracking by an evolutionary self-organizing neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 22, no. 2-3, pp. 69–81, 2011.
- [32] M. D. Ruiz and E. Hüllermeier, "A formal and empirical analysis of the fuzzy gamma rank correlation coefficient," *Information Sciences*, vol. 206, pp. 1–17, 2012.
- [33] C. Zhong, G. Yanli, H. Huang, L. Tan, Y. Wu, and W. Wenting, "Three-dimensional reconstruction of coronary arteries and its application in localization of coronary artery segments corresponding to myocardial segments identified by transthoracic echocardiography," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 783939, 8 pages, 2013.
- [34] J. Lia, Q. Dub, and C. Suna, "An improved box-tracking method for video fractal dimension estimation," *Pattern Recognition*, vol. 42, no. 11, pp. 2460–2469, 2009.
- [35] Y. Chen, W. Xu, F. Kuang, and S. Gao, "The research and application of visual saliency and adaptive support vector machine in target tracking field," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 925341, 8 pages, 2013.
- [36] Y. Yi, A. Mouloud, and B. Salah, "Automatic segmentation and measurement of vasculature in retinal fundus images using probabilistic formulation," *Computational and Mathematical Methods in Medicine*, pp. 260410–260421, 2013.
- [37] C. M. B. Nobre, R. A. Braga Jr., A. G. Costa, R. R. Cardoso, W. S. da Silva, and T. Sáfiadi, "Biospeckle laser spectral analysis under Inertia Moment, Entropy and Cross-Spectrum methods," *Optics Communications*, vol. 282, no. 11, pp. 2236–2242, 2009.
- [38] Z. Liang, X. Liu, B. Ye, and R. K. Brauwer, "Fuzzy performance between surface fitting and energy distribution in turbulence runner," *The Scientific World Journal*, vol. 2012, Article ID 408949, 14 pages, 2012.
- [39] J. Hua, L. Meng, Z. Xu, and G. Li, "An adaptive signal-to-noise ratio estimator in mobile communication channels," *Digital Signal Processing*, vol. 20, no. 3, pp. 692–698, 2010.
- [40] R. Xiao, J. Yang, M. Goyal, Y. Liu, and Y. Wang, "Automatic vasculature identification in coronary angiograms by adaptive geometrical tracking," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 796342, 11 pages, 2013.
- [41] J. A. Revie, D. Stevenson, J. G. Chase et al., "Evaluation of a model-based hemodynamic monitoring method in a porcine study of septic shock," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 505417, 17 pages, 2013.
- [42] J. L. C. Mission and H. Kim, "Design charts for elastic pile shortening in the equivalent top-down load-settlement curve from a bidirectional load test," *Computers and Geotechnics*, vol. 38, no. 2, pp. 167–177, 2011.
- [43] Z. Liang, X. Liu, and B. Ye, "Four-dimensional fuzzy relation investigation in turbulence kinetic energy distribution, surface cluster modeling," *Arabian Journal for Science and Engineering*, vol. 39, no. 3, pp. 2339–2351, 2014.
- [44] T. P. Vo and J. Lee, "Interaction curves for vibration and buckling of thin-walled composite box beams under axial loads and end moments," *Applied Mathematical Modelling. Simulation and Computation for Engineering and Environmental Systems*, vol. 34, no. 10, pp. 3142–3157, 2010.
- [45] G. A. Papakostas, Y. S. Boutalis, and D. A. a. Karras, "Pattern classification by using improved wavelet compressed Zernike moments," *Applied Mathematics and Computation*, vol. 212, no. 1, pp. 162–176, 2009.
- [46] T.-W. Sung and C.-S. Yang, "A Voronoi-based sensor handover protocol for target tracking in distributed visual sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2014, Article ID 586210, 14 pages, 2014.
- [47] J. Pradeep, E. Srinivasan, and S. Himavathi, "An investigation on the performance of hybrid features for feed forward neural network based English handwritten character recognition system," *WSEAS Transactions on Signal Processing*, vol. 10, no. 1, pp. 21–29, 2014.
- [48] D. B. Sinha, N. M. Ledbetter, and D. L. Barbour, "Spike-timing computation properties of a feed-forward neural network model," *Frontiers in Computational Neuroscience*, vol. 8, no. 5, pp. 1–17, 2014.

- [49] P. Fuangkhon, "An incremental learning preprocessor for feed-forward neural network," *Artificial Intelligence Review*, vol. 41, no. 2, pp. 183–210, 2014.
- [50] G. Rajabzadeh, S. Salehi, A. Nemati, R. Tavakoli, and M. Solati Hashjin, "Enhancing glass ionomer cement features by using the HA/YSZ nanocomposite: a feed forward neural network modelling," *Journal of the Mechanical Behavior of Biomedical Materials*, vol. 29, pp. 317–327, 2014.
- [51] F. N. Sibai, A. Nuaimi, A. Maamari, and R. Kuwair, "Ear recognition with feed-forward artificial neural networks," *Neural Computing and Applications*, vol. 23, no. 5, pp. 1265–1273, 2013.
- [52] X. Zhang and Z. Zou, "Black-box modeling of ship manoeuvring motion based on feed-forward neural network with Chebyshev orthogonal basis function," *Journal of Marine Science and Technology*, vol. 18, no. 1, pp. 42–49, 2013.

Research Article

Screening for Prediabetes Using Machine Learning Models

**Soo Beom Choi,^{1,2} Won Jae Kim,³ Tae Keun Yoo,^{1,3} Jee Soo Park,^{1,3} Jai Won Chung,^{1,4}
Yong-ho Lee,⁵ Eun Seok Kang,⁵ and Deok Won Kim^{1,4}**

¹ Department of Medical Engineering, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 120-752, Republic of Korea

² Brain Korea 21 PLUS Project for Medical Science, Yonsei University, Republic of Korea

³ Department of Medicine, Yonsei University College of Medicine, Republic of Korea

⁴ Graduate Program in Biomedical Engineering, Yonsei University, Seoul, Republic of Korea

⁵ Department of Internal Medicine, Yonsei University Health System, Republic of Korea

Correspondence should be addressed to Deok Won Kim; kdw@yuhs.ac

Received 28 May 2014; Accepted 8 July 2014; Published 16 July 2014

Academic Editor: David A. Elizondo

Copyright © 2014 Soo Beom Choi 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.

The global prevalence of diabetes is rapidly increasing. Studies support the necessity of screening and interventions for prediabetes, which could result in serious complications and diabetes. This study aimed at developing an intelligence-based screening model for prediabetes. Data from the Korean National Health and Nutrition Examination Survey (KNHANES) were used, excluding subjects with diabetes. The KNHANES 2010 data ($n = 4685$) were used for training and internal validation, while data from KNHANES 2011 ($n = 4566$) were used for external validation. We developed two models to screen for prediabetes using an artificial neural network (ANN) and support vector machine (SVM) and performed a systematic evaluation of the models using internal and external validation. We compared the performance of our models with that of a screening score model based on logistic regression analysis for prediabetes that had been developed previously. The SVM model showed the areas under the curve of 0.731 in the external datasets, which is higher than those of the ANN model (0.729) and the screening score model (0.712), respectively. The prescreening methods developed in this study performed better than the screening score model that had been developed previously and may be more effective method for prediabetes screening.

1. Introduction

The prevalence of type 2 diabetes is dramatically increasing, resulting in a global public health issue [1]. The prevalence of diabetes was estimated at 285 million or 6.4% of adults in the world in 2010 [2], and this prevalence is expected to rise to 552 million by 2030 [3]. The increasing rates of obesity are expected to result in a faster increase in the prevalence of type 2 diabetes in the future [4]. However, owing to the absence of symptoms and/or disease-related knowledge, diabetes often goes undetected, and approximately one-third of people with diabetes are not aware of their status [5–7]. Therefore, development of a simple accurate screening method is needed. Historically, the majority of the clinical screening methods consisted of surveys developed using logistic regression analyses to predict diabetes [8–13].

Prediabetes was first recognized as an intermediate diagnosis and indication of a relatively high risk for the future development of diabetes by the Expert Committee on Diagnosis and Classification of Diabetes Mellitus in 1997 [14], and it has been reported that approximately 5–10% of patients with untreated prediabetes subsequently develop diabetes [15, 16]. This is significant considering that prediabetes based on impaired fasting glucose (IFG) was estimated to affect 4.9 million people, accounting for 17.4% of Korean adults in 2005 [6], with a further 35% of adults in the US with prediabetes in 2008 [17]. The definition of prediabetes includes a fasting plasma glucose (FPG) level in the range of 100–125 mg/dL (5.6–6.9 mmol/L), impaired glucose tolerance (IGT) (oral glucose tolerance test (OGTT) 2 h measurement in the range of 140–199 mg/dL (7.8–11.0 mmol/L)), or HbA1c level in the range of 5.7–6.4% (39–46 mmol/mol). Similar to diabetes,

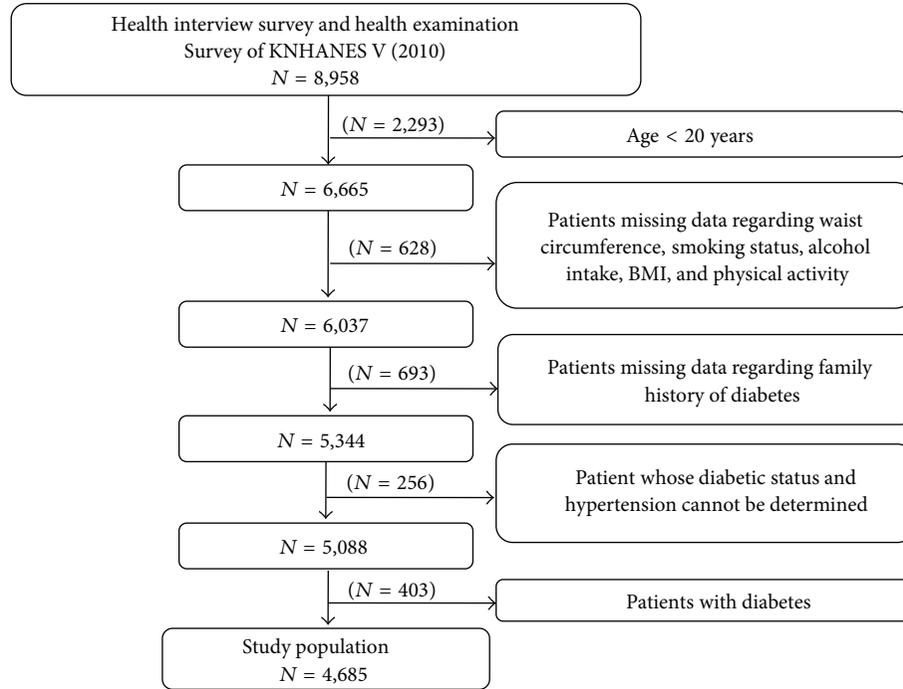


FIGURE 1: Flow chart of excluding subjects for the KNHANES 2010.

the risk of microvascular complications is increased with prediabetes [18], and the risk for cardiovascular disease and total mortality is almost twice as high in individuals with prediabetes [19, 20]. Early diagnosis and intervention for prediabetes could prevent these complications, prevent delay, or prevent the transition to diabetes [21] and be cost-effective [22].

Machine learning is an area of artificial intelligence research, which uses statistical methods for data classification. Several machine learning techniques have been applied in clinical settings to predict disease and have shown higher accuracy for diagnosis than classical methods [23]. Support vector machines (SVM) and artificial neural networks (ANN) have been widely used approaches in machine learning. They are the most frequently used supervised learning methods for analyzing complex medical data [24].

In this study, we aimed to develop and validate models to predict prediabetes using artificial neural network (ANN) and support vector machine (SVM) methods, which could be effective as simple and accurate screening tools. The model performance was compared to that of the screening score model that we modified for prediabetes based on the screening score for diabetes by Lee et al. [8], with respect to accuracy and area under the curve (AUC) of the receiver operating characteristic (ROC).

2. Materials and Methods

2.1. Data Source and Subjects. Data from the Korean National Health and Nutrition Examination Survey (KNHANES) 2010 [25] and 2011 [26] were used to develop and validate,

respectively, the ANN and SVM models for prediabetes. The KNHANES is a cross-sectional survey that includes approximately 800 questions; it is conducted by the Division of Chronic Disease Surveillance, Korea Centers for Disease Control and Prevention. The survey represents the entire nation using rolling sampling survey method. The following exclusion criteria applied to the subjects in both datasets: <20 years of age, missing data for waist circumference, smoking status, alcohol intake, body mass index (BMI), physical activity, family history of diabetes, undetermined diabetic status or hypertension status, and diagnosed diabetes or undiagnosed diabetes. Diagnosed and undiagnosed patients were excluded in order to focusing on predicting prediabetes. Undiagnosed diabetes was defined as a FPG \geq 126 mg/dL without diagnosis by clinician. Of the 8,958 subjects in the KNHANES 2010, 4,685 were included as shown in Figure 1. Of the 8,518 subjects in the KNHANES 2011, 4,566 subjects were included using the same flow chart as Figure 1. The subjects of the KNHANES 2010 and 2011 data sets were not overlapped.

Figure 2 illustrates the study flow. The development dataset from KNHANES 2010 was randomly divided into training and internal validation sets using a 2:1 ratio. The training set ($n = 3,134$) was used to construct the ANN and SVM models. The internal validation set ($n = 1,551$) was used to assess the ability to predict prediabetes. Additionally, data from KNHANES 2011 were used as an external validation set ($n = 4,566$). All individuals in the surveys participated voluntarily and informed consent was obtained from all participants. The survey protocol was approved by the Institutional Review Board of the Korean Centers for Disease Control and Prevention.

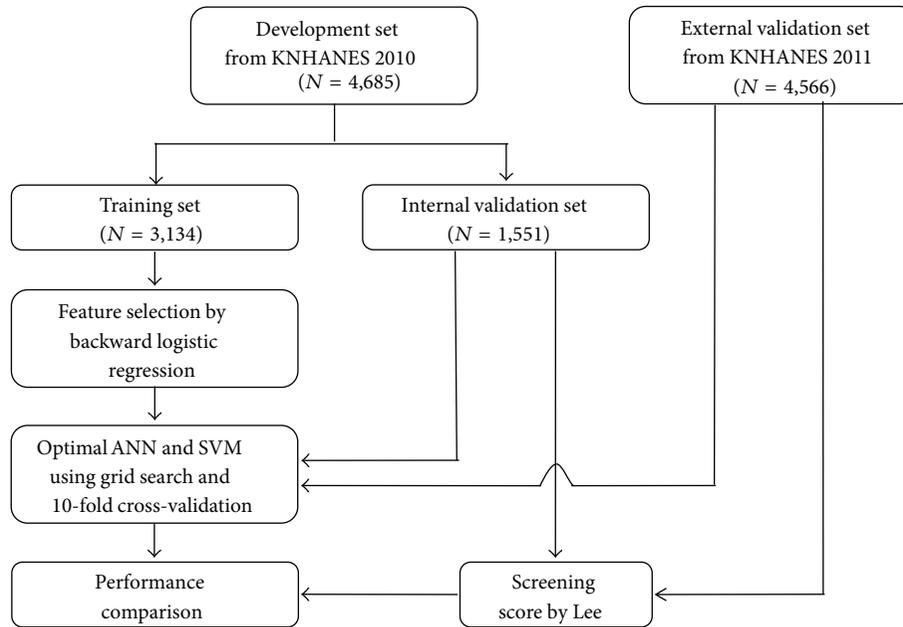


FIGURE 2: Chart depicting the flow of data from the Korean National Health and Nutrition Examination Survey (KNHANES) 2010 and 2011 to develop and validate a prediabetes model. KNHANES: Korean National Health and Nutrition Examination Survey; ANN: artificial neural network; SVM: support vector machine.

2.2. Risk Factors. We adopted the most frequently used nine variables from previous studies regarding diabetes prediction models: age, gender, family history of diabetes, hypertension, alcohol intake, BMI, smoking status, waist circumference, and physical activity [8–13]. FPG was determined using glucose levels that were collected following ≥ 8 hours of fasting. We considered only FPG although there are three methods to diagnose prediabetes. In KNHANES 2010, FPG was obtained from every subject, but OGTT was not tested, and HbA1c was tested only for subjects with diabetes. A family history of diabetes was limited to parents and siblings. Hypertension was defined as systolic blood pressure (SBP) > 140 mmHg, diastolic blood pressure (DBP) > 90 mmHg, or use of medication for blood pressure control [8]. Alcohol intake was calculated using 2 questions: (1) alcohol consumption frequency during the previous 12 months and (2) average number of drinks on those days. The amount of alcohol was calculated based on the number of glasses, regardless of the kind of beverage, assuming that the amount of alcohol was approximately the same in each glass (approximately 8 g alcohol per glass). Smoking status was divided into “currently smoking regularly” and “others,” with the latter group including subjects who had never smoked or had quit smoking. The subjects who answered more than “moderate” to the question “how intense is your everyday activity?” were considered as physically active.

2.3. Artificial Neural Network. ANN is an artificial intelligence technology, inspired by the architecture of biological neurons such as that in the human brain [27]. The technology is specialized for classification, and it is mostly used to identify underlying patterns for risk factors in medicine. When trained properly, neural networks are known to have

more accurate predictive abilities than conventional methods such as logistic regression. There have been a number of recent advances in ANN methodology that enable automatic detection of an optimal predictive model [28, 29]. Unlike logistic regression, ANNs are able to detect complex nonlinear relationships between multiple predictors and diseases, which make them useful in support systems for medical decisions [30, 31].

The ANN models were constructed using NeuroSolution version 6.0 (NeuroDimension, Gainesville, FL), which is professional software that simplifies the construction of ANN [32]. This software allows simultaneous testing of various types of neural networks, including generalized regression neural network, multilayer perceptron, probabilistic neural network, radial basis neural network, feedforward neural network, and support vector machine. To avoid overfitting, the prediction models were internally validated using cross-validation. Performance of the prediction models was monitored during training and cross-validation to obtain optimal algorithm and parameters, such as learning rate, momentum, and number of hidden nodes. The ANN was trained with 7 predictors including age, gender, waist circumference, BMI, family history of diabetes, hypertension, and alcohol intake, which were selected using backward logistic regression. The model chosen for prediabetes prediction was a multilayer perceptron model with 1 hidden layer, batch training, and momentum learning (MLP-1-BM) of backpropagation feed-forward algorithm, which demonstrated the best performance as a desired ANN.

2.4. Support Vector Machine. SVM maps data to a higher dimensional space through a kernel function that linearly separates data patterns. The data are divided into two groups

TABLE 1: The weighted characteristics of the data from the Korean National Health and Nutrition Examination Survey (KNHANES) 2010.

	Normal ($n = 3,681$)	Prediabetes ($n = 1,004$)	P^*
Age (years)	41.9 \pm 0.5 [41.0–42.8]	52.5 \pm 0.6 [51.3–53.7]	<0.001
Gender (% men)	46.9 (0.9) [45.2–48.7]	58.8 (1.9) [55.0–62.5]	<0.001
Family history of diabetes (%)	18.3 (0.9) [16.6–20.2]	22.9 (1.7) [19.8–26.4]	0.007
Current smoker (%)	27.5 (1.0) [25.5–29.6]	26.9 (1.9) [23.4–30.8]	0.799
Alcohol intake (drinks/day)	0.8 \pm 0.0 [0.7–0.9]	1.0 \pm 0.1 [0.9–1.2]	<0.001
Physically active (%)	50.6 \pm 1.1 [48.4–52.9]	52.1 \pm 2.1 [48.0–56.3]	0.535
BMI (kg/m ²)	23.2 \pm 0.1 [23.1–23.3]	25.1 \pm 0.1 [24.8–25.3]	<0.001
Waist circumference (cm)	79.1 \pm 0.2 [78.7–79.6]	85.8 \pm 0.4 [85.1–86.6]	<0.001
FPG (mg/dL)	89.0 \pm 0.1 [88.7–89.3]	107.4 \pm 0.3 [106.9–108.0]	<0.001
Systolic blood pressure (mmHg)	116.8 \pm 0.4 [116.0–117.5]	127.7 \pm 0.7 [126.4–129.1]	<0.001
Diastolic blood pressure (mmHg)	76.4 \pm 0.3 [75.8–77.0]	81.5 \pm 0.5 [80.6–82.4]	<0.001
Hypertension (%)	16.4 (0.8) [14.9–18.0]	41.1 (2.2) [36.8–45.5]	<0.001

BMI: body mass index; FPG: fasting plasma glucose.

Table values are given as mean \pm standard error or % (standard error) [95% confidence interval] unless otherwise indicated. P^* were obtained by t -test or chi-square test.

Impaired fasting glucose was considered with values ≥ 100 mg/dL and <126 mg/dL.

by the training data referred to as a support vector. SVM models are determined by choosing the maximum-margin hyperplane with the nearest support vector of the two groups [33]. SVM improves the accuracy of a model through the optimization of separating space using the kernel function, but one of the disadvantages of SVM is that it requires many trials to construct an optimal SVM model in comparison with other machine learning techniques [34].

The same seven risk factors as those in the ANN model were employed for the SVM. To obtain the optimal model, we adopted a grid search in which a range of parameter values (penalty parameter [C] of 0.01, 0.1, 1, 10, and 100 and scaling factor [σ] of 0.001, 0.01, 0.1, 1, 10, and 100) was tested using the 10-fold cross-validation strategy. The optimal parameter values with a C of 10 and σ of 10 for SVM using the Gaussian kernel function were obtained. The SVM models were constructed using MATLAB Version 2012a (Mathworks Inc., Natick, MA).

2.5. Screening Score of Our Models for Prediabetes. The models constructed by ANN and SVM were compared with a previously developed screening survey to illustrate performance of our models and the possibility of their use in real situations. For this purpose, we used a screening score model for the Korean population constructed by Lee et al. [8]; we felt this was appropriate because both studies constructed models for the Korean population. Lee et al. used data from KNHANES 2001 and 2005 for training and data from KNHANES 2007 and 2008 for external validation. In addition, the screening score model by Lee et al. used very similar risk factors to ours, with the exception of current smoking status. Those 6 variables independently associated with undiagnosed diabetes were chosen for their model: age, family history of diabetes, hypertension, waist circumference, smoking, and alcohol intake.

The risk score was assigned according to the odds ratio for each risk factor in the logistic regression model defined by Lee

et al. [8]. Within the total score range of 0–11 points, a cut-off score of ≥ 5 points was selected to indicate an individual at high risk for undiagnosed diabetes; this cut-off resulted in the highest value for the Youden index. The 6 risk factors jointly yielded an AUC of 0.730 for both the internal and external validation sets [8]. To compare with our models for prediabetes, we constructed a new screening score model for prediabetes by adjusting the cut-off point value based on our definition of prediabetes ($100 \text{ mg/dL} \leq \text{FPG} < 126 \text{ mg/dL}$), given that the screening score for diabetes used by Lee et al. was based on $\text{FPG} \geq 126 \text{ mg/dL}$ [8]. The screening score for prediabetes was designed with the same risk score model of the 6 risk factors using our training set for prediabetes (KNHANES 2010) and the Youden index; as a result, a cut-off score of ≥ 5 points was identified to indicate an individual with prediabetes.

2.6. Statistical Analyses. The weighted characteristics of the data from the KNHANES 2010 to represent the entire normal and prediabetes people in Korea are summarized by descriptive statistics in Table 1. For comparison of the factors between normal and prediabetes, the continuous and categorical characteristics were tested using t -test and chi-square test, respectively.

To obtain the optimal variables for the prediction model, backward logistic regression was performed with the training set. Each step of the backward regression excluded the variables without a statistically meaningful correlation with the outcome, prediabetes. Three steps of backward regression were executed, and the selected 7 variables were age, BMI, hypertension, gender, alcohol intake, waist circumference, and family history. The Hosmer-Lemeshow test resulted in a P value of 0.132, indicating that the chosen variables were well fitted.

ROC curve analysis is the most commonly used method in clinical analysis to establish an optimal cut-off point [35].

TABLE 2: Performance of the ANN, SVM, and screening score (Lee et al. [8]) models using the internal and external validation sets for predicting prediabetes.

		AUC	Accuracy (%)	Sensitivity (%)	Specificity (%)
Internal validation set ($n = 1,551$)	ANN*	0.768	69.0	74.1	67.5
	SVM†	0.761	64.9	78.9	61.2
	Screening score‡	0.734	63.4	76.1	60.0
External validation set ($n = 4,566$)	ANN*	0.729	60.7	77.2	56.7
	SVM†	0.731	66.1	69.4	65.3
	Screening score‡	0.712	59.9	74.3	56.4

AUC: area under the curve; ANN: artificial neural network; SVM: support vector machine.

The internal validation set was comprised of data from the Korean National Health and Nutrition Examination Survey (KNHANES) 2010, and the external validation set included data from KNHANES 2011. *The chosen model was a multilayer perceptron model with 1 hidden layer, batch training, and momentum learning (MLP-1-B-M) of backpropagation feedforward algorithm. †The optimal model was found using Gaussian kernel function with a penalty parameter (C) of 10 and scaling factor (σ) of 10. ‡The performance was calculated by applying the screening score model for prediabetes based on that of Lee et al. [8] to the data from KNHANES 2010 and 2011.

TABLE 3: Performance of the screening score model (Lee et al. [8]) in predicting prediabetes and undiagnosed diabetes using the data from the Korean National Health and Nutrition Examination Survey (KNHANES) 2010 and 2011.

		AUC	Accuracy (%)	Sensitivity (%)	Specificity (%)
Prediabetes	KNHANES 2010* (internal validation)	0.734	63.4	76.1	60.0
	KNHANES 2011* (external validation)	0.712	59.9	74.3	56.4
Undiagnosed diabetes	KNHANES 2010† (internal validation)	0.772	66.6	76.5	66.4
	KNHANES 2011† (external validation)	0.751	64.6	74.4	64.3

AUC: area under the curve; KNHANES: Korean National Health and Nutrition Examination Survey.

Prediabetes was defined as fasting plasma glucose, with values ≥ 100 mg/dL and < 126 mg/dL. *Internal and external validation sets to evaluate the screening score for prediabetes ($n = 1,551$ for KNHANES 2010 and $n = 4,566$ for KNHANES 2011). †Internal and external validation sets to evaluate the screening score for undiagnosed diabetes ($n = 1,585$ for KNHANES 2010 and $n = 4,683$ for KNHANES 2011).

Therefore, we generated ROC curves and the selected cut-off points that maximized the Youden index [36] to compare the performance of our optimal machine learning models with that of the screening score model for prediabetes based on the screening score by Lee et al. [8], using our internal and external validation sets. Following the ROC analysis, the AUC, accuracy, sensitivity, and specificity of our models and screening score model for prediabetes were calculated. The classification accuracy measured the proportion of cases correctly classified. Sensitivity measured the fraction of positive cases that were classified as positive. Specificity measured the fraction of negative cases that were classified as negative. We used SPSS 20.0 (IBM Corp, Armonk, NY) for statistical analysis and MedCalc 12.4 (MedCalc Inc., Mariakerke, Belgium) for ROC analysis. Statistical significance was set at $P < 0.05$.

3. Results

The weighted characteristics of the KNHANES 2010 data are summarized in Table 1. The factors that were significantly related to prediabetes were age, gender, family history of diabetes, alcohol intake, BMI, waist circumference, FPG, systolic and diastolic blood pressures, and hypertension.

When the prediction performance of the ANN model using 10-fold cross-validation was assessed for the training set, the final model showed an AUC of 0.706 and an accuracy of 65.6% for prediabetes. Cross-validation of the optimal

SVM parameters with the training set resulted in an AUC of 0.742 and accuracy of 69.9%. These results are not included in Table 2. The similar performance observed between the training and validation sets in Table 2 indicates that the trained models were not overfitting.

With both the internal and the external validation sets, our ANN and SVM models showed better performance than the existing screening score model using logistic regression, especially in terms of AUC, which is known as a better measure than accuracy in evaluating learning algorithms [37] (Table 2). In the external validation set, the accuracy of the SVM model was 5.4% and 6.2% higher than that of the ANN and screening score models, respectively. The ROC curves of the ANN, SVM, and screening score models are depicted for the internal and external validation sets in Figure 3.

Table 3 shows the performance obtained by applying the screening score model by Lee et al. [8] to the data from KNHANES 2010 and 2011 for predicting prediabetes and undiagnosed diabetes. For all performance parameters (AUC, accuracy, sensitivity, and specificity) in both datasets, the ability to predict prediabetes was inferior to that for diabetes. In particular, the AUC and accuracy for prediabetes in the external validation set were lower than those for undiagnosed diabetes by 0.039 and 4.7%, respectively. AUC and accuracy of the SVM model for external validation are higher than those of the screen score model for prediabetes by 0.019 and 6.2%, respectively.

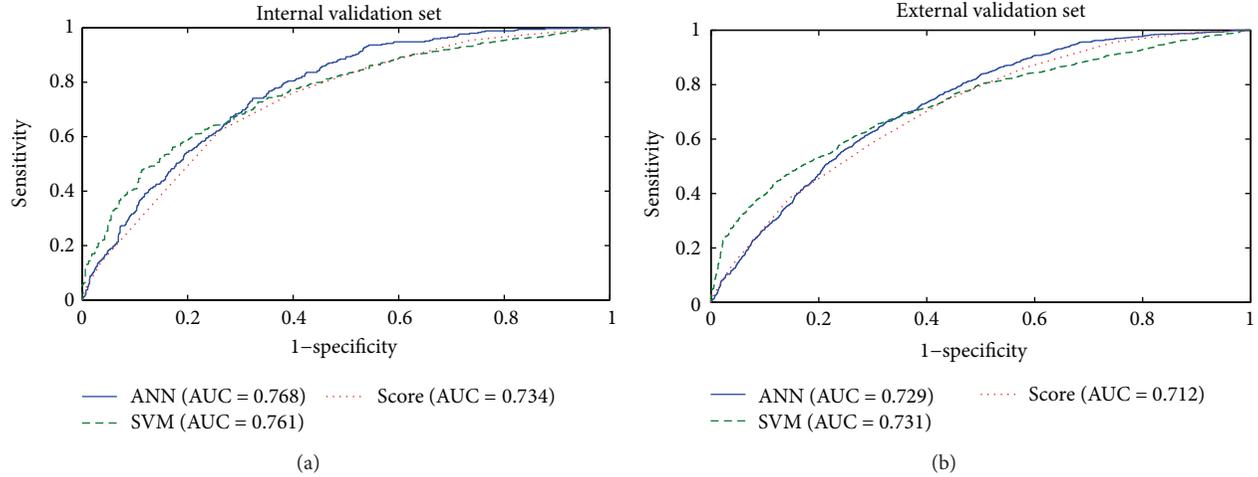


FIGURE 3: Receiver operating characteristic curves (ROC) of artificial neural network (ANN), support vector machine (SVM), and screening score in predicting prediabetes for internal validation set (a) and external validation set (b).

4. Discussion

The results of the present study indicate that the ANN and SVM models that we developed to predict prediabetes, defined as IFG, performed better than the existing clinical screening score model, as indicated by the AUC and accuracy measures (Table 2). Although logistic regression analysis and ANN share common roots in statistical pattern recognition, the latter is a generalization of the former [38], which might explain why our ANN model performed better than the screening score model, which was based on logistic regression analysis. The SVM model performed particularly well due to the ability of SVM to efficiently find a unique optimal solution, incorporate multiple types of data with a degree of flexibility, and model nonlinear patterns [39]. We also investigated SVM models with different numbers of risk factors to find optimal parameters. The best performance of SVM model with six risk factors including age, body mass index, hypertension, gender, daily alcohol intake, and waist circumference was an AUC of 0.743 and accuracy of 70.2% in training set, which was almost the same as the performance of SVM model with seven risk factors in our paper, resulting in an AUC of 0.742 and accuracy of 69.9%.

Meng et al. [40] compared the performance of logistic regression, ANNs, and decision tree models for predicting diabetes or prediabetes using common risk factors in China population. In Meng et al. study, the ANNs model was the poorest of the three models, with 73.23% accuracy. This result is consistent with ours that the performance of ANN model was lower than SVM model. However, the performance of ANN can be case-by-case depending on characteristics of data or developers.

Although similar statistical analyses were conducted (i.e., backward regression models), there were slight differences in the variables included in the present study and those in the study by Lee et al. [8]. Lee et al. included current smoking status as a risk factor in the training set based on the data from KNHANES 2001 and 2005; however, current

smoking status was not included in our training set using data from KNHANES 2010. This may have resulted from lifestyle changes in the Korean population between those years, including a decline in the overall smoking rate and stronger antismoking laws [41]. Although several screening score models have been developed and used clinically, our prediction model is unique in several ways.

First, owing to the similarity between our machine learning models and the existing screening score models, we were able to compare the performance of our machine learning models with the existing models. Second, to the best of our knowledge, there are few studies investigating prediabetes; instead, the majority of the other models have been developed to predict undiagnosed diabetes. However, prediabetes is increasingly becoming a significant public health issue. Using our model to screen patients for prediabetes would enable interventions at an earlier stage, which would be easier to implement and more successful than interventions implemented following diabetes screening.

Prediabetes was more difficult to predict than diabetes using any of the parameters across all of the models, which is not unexpected. AUC and accuracy of the SVM model for external validation are higher than those of the screen score model for prediabetes by 0.019 and 6.2%, respectively (Table 2). Therefore, we demonstrated that the machine learning methods could help to overcome the difficulty in predicting prediabetes.

This study has certain limitations. First, FPG level was the only measurement that we used to define prediabetes and diabetes; OGTT and HbA1c were not taken into consideration. Data were lacking for these measurements; however, the use of FPG level was consistent with the model developed by Lee et al. [8], with which we compared our models.

Second, the screening score model for diabetes developed by Lee et al. [8] did not correspond perfectly with our model for prediabetes. For a more precise comparison in future studies, a screening score model for prediabetes should be constructed using the new regression equation with different

risk factors for prediabetes. In spite of this limitation, the suggested model with the new cut-off point is considered a good model for predicting prediabetes with AUCs of 0.734 and 0.712 in the internal and external validation sets, respectively.

Last, the new models that we developed are limited in terms of convenience and potential widespread use. Although the screening score model is not the most effective one for disease prediction, it is simple and accessible. However, machine learning models could also become more accessible through the use of calculator software, particularly with the widespread use of devices such as computers, smart phones, and tablet PCs. Future studies could develop a calculator in which the values are entered via a website or application and the results are immediately delivered to the end user. The decision tree method is also warranted for easy interpreting tree-like plot in the future.

5. Conclusion

Our study constructed a reasonably good model to predict prediabetes in the Korean population. By applying similar methods in other countries, researchers could develop country-specific machine learning models for nationwide use. The creation of a user-friendly calculator program would enable access to screening by the general population, in addition to medical professionals. This widespread use could result in early diagnosis and treatment for people with prediabetes and diabetes, helping to relieve the public health diabetes burden and reducing the number of people who remain undiagnosed.

Conflict of Interests

The authors have no potential conflict of interests relevant to this paper.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MEST) (NRF-2012R1A2A2A03045612).

References

- [1] B. C. Zyriax, R. Salazar, W. Hoepfner, E. Vettorazzi, C. Herder, and E. Windler, "The association of genetic markers for type 2 diabetes with prediabetic status - cross-sectional data of a diabetes prevention trial," *PLoS ONE*, vol. 8, Article ID e75807, 2013.
- [2] J. E. Shaw, R. A. Sicree, and P. Z. Zimmet, "Global estimates of the prevalence of diabetes for 2010 and 2030," *Diabetes Research and Clinical Practice*, vol. 87, no. 1, pp. 4–14, 2010.
- [3] D. R. Whiting, L. Guariguata, C. Weil, and J. Shaw, "IDF Diabetes Atlas: Global estimates of the prevalence of diabetes for 2011 and 2030," *Diabetes Research and Clinical Practice*, vol. 94, no. 3, pp. 311–321, 2011.
- [4] P. Zimmet, K. G. M. M. Alberti, and J. Shaw, "Global and societal implications of the diabetes epidemic," *Nature*, vol. 414, no. 6865, pp. 782–787, 2001.
- [5] S. M. Kim, J. S. Lee, J. Lee et al., "Prevalence of diabetes and impaired fasting glucose in Korea: Korean National Health and Nutrition Survey 2001," *Diabetes Care*, vol. 29, no. 2, pp. 226–231, 2006.
- [6] J. C. Yong, C. K. Hyeon, M. K. Hee, W. P. Seok, K. Jongoh, and J. K. Dae, "Prevalence and management of diabetes in Korean adults: Korea National Health and Nutrition Examination Surveys 1998–2005," *Diabetes Care*, vol. 32, no. 11, pp. 2016–2020, 2009.
- [7] C. C. Cowie, K. F. Rust, D. D. Byrd-Holt et al., "Prevalence of diabetes and impaired fasting glucose in adults in the U.S. population: national Health and Nutrition Examination Survey 1999–2002," *Diabetes Care*, vol. 29, no. 6, pp. 1263–1268, 2006.
- [8] Y. Lee, H. Bang, H. C. Kim, H. M. Kim, S. W. Park, and D. J. Kim, "A simple screening score for diabetes for the Korean population: development, validation, and comparison with other scores," *Diabetes Care*, vol. 35, no. 8, pp. 1723–1730, 2012.
- [9] J. Lindström and J. Tuomilehto, "The diabetes risk score: a practical tool to predict type 2 diabetes risk," *Diabetes Care*, vol. 26, no. 3, pp. 725–731, 2003.
- [10] M. B. Schulze, K. Hoffmann, H. Boeing et al., "An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes," *Diabetes Care*, vol. 30, no. 3, pp. 510–515, 2007.
- [11] H. Bang, A. M. Edwards, A. S. Bombardieri et al., "Development and validation of a patient self-assessment score for diabetes risk," *Annals of Internal Medicine*, vol. 151, no. 11, pp. 775–783, 2009.
- [12] C. A. Baan, J. B. Ruige, R. P. Stolk et al., "Performance of a predictive model to identify undiagnosed diabetes in a health care setting," *Diabetes Care*, vol. 22, no. 2, pp. 213–219, 1999.
- [13] S. Griffin, P. S. Little, C. N. Hales, A. L. Kinmonth, and N. J. Wareham, "Diabetes risk score: towards earlier detection of type 2 diabetes in general practice," *Diabetes Metabolic Research Review*, vol. 16, no. 3, pp. 164–171, 2000.
- [14] American Diabetes Association, "Diagnosis and classification of diabetes mellitus (Position Statement)," *Diabetes Care*, vol. 35, supplement 1, pp. S64–S71, 2012.
- [15] F. De Vegt, J. M. Dekker, A. Jager et al., "Relation of impaired fasting and postload glucose with incident type 2 diabetes in a Dutch population: the Hoorn study," *Journal of the American Medical Association*, vol. 285, no. 16, pp. 2109–2113, 2001.
- [16] The Diabetes Prevention Program (DPP) Research Group, "The diabetes prevention program (DPP): description of lifestyle intervention," *Diabetes Care*, vol. 25, no. 12, pp. 2165–2171, 2002.
- [17] T. L. Sentell, G. He, E. W. Gregg, and D. Schillinger, "Racial/ethnic variation in prevalence estimates for United States prediabetes under alternative 2010 American diabetes association criteria: 1988–2008," *Ethnicity & Disease*, vol. 22, no. 4, pp. 451–458, 2012.
- [18] A. G. Tabák, C. Herder, W. Rathmann, E. J. Brunner, and M. Kivimäki, "Prediabetes: a high-risk state for diabetes development," *The Lancet*, vol. 379, no. 9833, pp. 2279–2290, 2012.
- [19] M. Coutinho, H. C. Gerstein, Y. Wang, and S. Yusuf, "The relationship between glucose and incident cardiovascular events: a metaregression analysis of published data from 20 studies of 95,783 individuals followed for 12.4 years," *Diabetes Care*, vol. 22, no. 2, pp. 233–240, 1999.

- [20] S. C. Port, M. O. Goodarzi, N. G. Boyle, and R. I. Jennrich, "Blood glucose: a strong risk factor for mortality in nondiabetic patients with cardiovascular disease," *American Heart Journal*, vol. 150, no. 2, pp. 209–214, 2005.
- [21] L. Tian, J. Zhu, L. Liu, Y. Liang, J. Li, and Y. Yang, "Prediabetes and short-term outcomes in nondiabetic patients after acute ST-elevation myocardial infarction," *Cardiology*, vol. 127, no. 1, pp. 55–61, 2014.
- [22] W. H. Herman, T. J. Hoerger, M. Brandle et al., "The cost-effectiveness of lifestyle modification or metformin in preventing type 2 diabetes in adults with impaired glucose tolerance," *Annals of Internal Medicine*, vol. 142, no. 5, pp. 323–332, 2005.
- [23] T. K. Yoo, S. K. Kim, D. W. Kim et al., "Osteoporosis risk prediction for bone mineral density assessment of postmenopausal women using machine learning," *Yonsei Medical Journal*, vol. 54, no. 6, pp. 1321–1330, 2013.
- [24] C. Hsieh, R. Lu, N. Lee, W. Chiu, M. Hsu, and Y. Li, "Novel solutions for an old disease: diagnosis of acute appendicitis with random forest, support vector machines, and artificial neural networks," *Surgery*, vol. 149, no. 1, pp. 87–93, 2011.
- [25] Korea Centers for Disease Control and Prevention, *The Fifth Korea National Health and Nutrition Examination Survey (KNHANES V)*, Ministry of Health and Welfare, Seoul, South Korea, 2010.
- [26] Korea Centers for Disease Control and Prevention, *The Fifth Korea National Health and Nutrition Examination Survey (KNHANES V-2)*, Ministry of Health and Welfare, Seoul, Republic of Korea, 2011.
- [27] J. Zou, Y. Han, and S. So, "Overview of artificial neural networks," *Methods in Molecular Biology*, vol. 458, pp. 15–23, 2008.
- [28] A. A. Motsinger-Reif and M. D. Ritchie, "Neural networks for genetic epidemiology: past, present, and future," *BioData Mining*, vol. 1, p. 3, 2008.
- [29] W. Yao, X. Chen, Y. Zhao, and M. van Tooren, "Concurrent subspace width optimization method for RBF neural network modeling," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 2, pp. 247–259, 2012.
- [30] W. G. Baxt, "Application of artificial neural networks to clinical medicine," *The Lancet*, vol. 346, no. 8983, pp. 1135–1138, 1995.
- [31] C. Eller-Vainicher, I. Chiodini, I. Santi et al., "Recognition of morphometric vertebral fractures by artificial neural networks: Analysis from gismo Lombardia database," *PLoS ONE*, vol. 6, no. 11, Article ID e27277, 2011.
- [32] A. Heiat, "Comparison of artificial neural network and regression models for estimating software development effort," *Information and Software Technology*, vol. 44, no. 15, pp. 911–922, 2002.
- [33] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [34] K. Kim, J. Y. Choi, T. K. Yoo, S. K. Kim, K. Chung, and D. W. Kim, "Mortality prediction of rats in acute hemorrhagic shock using machine learning techniques," *Medical and Biological Engineering and Computing*, vol. 51, no. 9, pp. 1059–1067, 2013.
- [35] A. K. Akobeng, "Understanding diagnostic tests 3: receiver operating characteristic curves," *Acta Paediatrica*, vol. 96, no. 5, pp. 644–647, 2007.
- [36] R. Fluss, D. Faraggi, and B. Reiser, "Estimation of the Youden index and its associated cutoff point," *Biometrical Journal*, vol. 47, no. 4, pp. 458–472, 2005.
- [37] J. Huang and C. X. Ling, "Using AUC and accuracy in evaluating learning algorithms," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 3, pp. 299–310, 2005.
- [38] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," *Journal of Biomedical Informatics*, vol. 35, no. 5-6, pp. 352–359, 2002.
- [39] R. C. Thurston, K. A. Matthews, J. Hernandez, and F. de la Torre, "Improving the performance of physiologic hot flash measures with support vector machines," *Psychophysiology*, vol. 46, no. 2, pp. 285–292, 2009.
- [40] X. H. Meng, Y. X. Huang, D. P. Rao, Q. Zhang, and Q. Liu, "Comparison of three data mining models for predicting diabetes or prediabetes by risk factors," *Kaohsiung Journal of Medical Sciences*, vol. 29, no. 2, pp. 93–99, 2013.
- [41] Y. Khang, S. Yun, H. Cho, and K. Jung-Choi, "The impact of governmental antismoking policy on socioeconomic disparities in cigarette smoking in South Korea," *Nicotine and Tobacco Research*, vol. 11, no. 3, pp. 262–269, 2009.

Research Article

Privacy-Preserving Self-Helped Medical Diagnosis Scheme Based on Secure Two-Party Computation in Wireless Sensor Networks

Yi Sun,¹ Qiaoyan Wen,¹ Yudong Zhang,² and Wenmin Li¹

¹ State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

² Brain Image Processing, Columbia University, New York, NY 10032, USA

Correspondence should be addressed to Yi Sun; sybupt@bupt.edu.cn

Received 10 April 2014; Accepted 23 June 2014; Published 14 July 2014

Academic Editor: David A. Elizondo

Copyright © 2014 Yi Sun 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.

With the continuing growth of wireless sensor networks in pervasive medical care, people pay more and more attention to privacy in medical monitoring, diagnosis, treatment, and patient care. On one hand, we expect the public health institutions to provide us with better service. On the other hand, we would not like to leak our personal health information to them. In order to balance this contradiction, in this paper we design a privacy-preserving self-helped medical diagnosis scheme based on secure two-party computation in wireless sensor networks so that patients can privately diagnose themselves by inputting a health card into a self-helped medical diagnosis ATM to obtain a diagnostic report just like drawing money from a bank ATM without revealing patients' health information and doctors' diagnostic skill. It makes secure self-helped disease diagnosis feasible and greatly benefits patients as well as relieving the heavy pressure of public health institutions.

1. Introduction

With the rapid development of science, more and more advanced technologies such as the internet of things and cloud computing are utilized in the area of modern medicine and this trend further pushes healthcare into the digital era [1–3]. Currently, numerous healthcare devices such as heart rate monitor, blood pressure monitor, and electrocardiogram are already popular in people's normal life. It makes it convenient for people to be aware of their health situation by viewing the reports of these devices. Especially, by the growing use of sensor technology in telecare, the new field known as wireless body area networks (WBAN) [1, 4] has designed various sensor devices that can be used to supervise critical body parameters and activities anytime and anywhere. People can easily and conveniently get the health data by these advanced sensor devices [5] such as temperature measurement, respiration monitor, heart rate monitor, pulse oximeter SpO₂, blood pressure monitor, pH monitor, glucose sensor, cardiac arrhythmia monitor/recorder, brain liquid pressure sensor, and endoscope capsule. What is more, these

devices are becoming more functional and portable. More and more mobile medical monitors have already been used to serve us [2].

Therefore, people no longer worry about how to obtain the health data but are concerned about how to securely deal with these sensitive data to have disease diagnosis with a medical institution. Traditionally, the issue of privacy of medical data has been dealt with primarily as a policy problem [6, 7]. Many related laws have been issued to protect the privacy of patients. However, it is still far away from satisfactory and people still fear the leakage of their private data. Hence, the most efficient solution to this problem is to protect patients' privacy in technology rather than in policy alone. In this aspect, most of previous literatures have introduced homomorphic encryption (HE) [8–10] to protect patients' privacy in some privacy-preserving medical applications [11]. However, HE will inevitably introduce tremendous cost and is not applicable to practical large-scale applications. Therefore, in this paper, we focus on building a secure and practical privacy-preserving medical diagnosis system that can serve us in our daily life. Starting from the aspiration of

the patient, the most secure and plausible diagnostic method is to apply the processed data rather than the original data to interact with the hospital which owns a disease database to diagnose the health status privately. Moreover, it requires that after diagnosis, the hospital gets nothing about the patient's health data and the patient has no idea of the hospital's disease database.

Inspired by daily used bank automated teller machine (ATM), we introduce the privacy-preserving self-helped medical diagnosis ATM (MD-ATM) so that after obtaining a healthcare card that stores some information about the health data which is collected by various sensor medical devices, patients can privately diagnose himself by inserting the health card into the MD-ATM to obtain diagnostic report just like drawing money from a bank ATM without revealing patient's health information and the disease database or doctors' diagnostic skill. When needing local computing, storing, or inputting some information, the patient uses his own portable device, called portable medical diagnostic device (PMDD).

In this paper, we will show how to realize this modern diagnosis system without HE. The main idea and technology we used in this scheme are secure two-party computation (STC) and oblivious transfer (OT). Firstly, we assume that patients themselves collect related data by various wireless sensor medical devices and further process and store them in their own health cards using PMDD. When diagnosing, the patient firstly transforms the original data locally and then inserts the card into the MD-ATM of the hospital to check up his health. Operating following the instructions of the MD-ATM, the patient will finally obtain a diagnostic report through OT and the patient then completes the self-helped diagnosis. In brief, our main contributions can be summarized as follows.

Our Contributions.

- (i) We build a new "patient-centered" medical diagnosis model in wireless sensor networks where patients themselves collect health data by various sensor medical devices while the hospital provides a disease database to help patients to complete disease diagnosis by themselves. Compared with traditional "doctor-centered" medical diagnosis model where patients have to depend on the doctor, our system is more appropriate especially when people pay more and more attention to privacy in wireless sensor networks.
- (ii) We firstly propose the privacy-preserving self-helped MD-ATM to construct a secure medical diagnosis scheme following the idea of STC. It makes secure self-helped medical diagnosis feasible and convenient just like drawing money from a bank ATM. It will greatly benefit patients as well as relieving the heavy pressure of public health institutions.
- (iii) We construct the self-helped medical diagnosis system based on OT without expensive HE. It provides us with another perspective to consider the problem of secure medical diagnosis for patients.

The rest of this paper is organized as follows. In Section 2, we briefly give an overview of secure two-party computation and oblivious transfer, and then we present our medical diagnosis system model in Section 3. In Section 4, we propose our privacy-preserving self-helped medical diagnosis scheme in detail and give a strict proof based on real-ideal simulation paradigm in Section 5. Finally, we summarize our work of this paper in the last section.

2. Preliminaries

2.1. Secure Two-Party Computation. Secure multiparty computation (SMC) is dedicated to deal with the problem of secure computation among distrustful participants. It was first introduced by Yao in 1982 [12] and then was extended by Goldreich et al. [13] and many other researchers [14–19]. Generally speaking, SMC is a method to implement cooperative computation with participants' private data, ensuring the correctness of the computation as well as not disclosing additional information except the necessary results. It has become a research focus in the international cryptographic community due to its wide applications in various areas and a mass of research results have been published one after another. Secure two-party computation (STC) [20] is a special case in SMC where there are only two participants. The well-known millionaires' problem [12] put forward by Yao is the representative problem of STC. In our discussing, we will consider the two-party case.

Generally speaking, STC is dedicated to computing a certain function between two mutually distrusted participants on their private inputs without revealing their private information. Informally, assuming that there are 2 participants, P_1, P_2 , each of them has a private number, x_1, x_2 , respectively. They want to cooperate to compute the function $y = f(x_1, x_2)$. A STC protocol is dubbed secure if no participant can learn more from the description of the public function and the result of the global calculation than what he can learn from his own information.

Formally, we usually analyze the security of a STC protocol using the real-ideal paradigm in the semihonest model where both of the two parties act semihonestly, following the protocol but making effort to gain more information about other parties' inputs, intermediate results, or overall outputs by the transcripts of the protocol [15]. We can overview the real-ideal paradigm as follows.

Firstly, in the ideal world, we assume that the computation of the functionality F on users private inputs is conducted by an additional trusted party, who receives x_i from user P_i , $i = 1, 2$, and returns the result $f(x_1, x_2)$ to P_i , $i = 1, 2$. However, there is no trusted party in the real world and so the two parties have to run a protocol Π to get the desired result. During executing protocol Π , both parties act semihonestly. Herein, the view of the i th party during an execution of Π on x_1, x_2 is denoted as $\text{VIEW}_i^\Pi(x_1, x_2)$, which contains P_i 's input, random tape, and the messages received from the other party. For a deterministic private function f , we say that Π privately computes f if there exist probabilistic polynomial-time algorithms S_1, S_2 , such that the simulated distribution

OT_1^m	
Inputs:	S inputs a set of m messages: (r_1, r_2, \dots, r_m) ; R inputs an index: z ;
Outputs:	S obtains λ , which means <i>Null</i> . R obtains r_z .

ALGORITHM 1

$\{S_i(x_i; f(x_1, x_2))\}$ is indistinguishable to $\{VIEW_i^\Pi(x_1, x_2)\}$, $i = 1, 2$. That is,

$$\begin{aligned} \{S_1(x_1; f(x_1, x_2))\} &\equiv \{VIEW_1^\Pi(x_1, x_2)\}, \\ \{S_2(x_2; f(x_1, x_2))\} &\equiv \{VIEW_1^\Pi(x_1, x_2)\}. \end{aligned} \quad (1)$$

2.2. Oblivious Transfer. In cryptography, OT is a type of protocol in which a sender transfers one of potentially many pieces of information to a receiver but remains oblivious as to which piece has been transferred. It was firstly introduced by Rabin [21] in 1981. Therein, the sender sends a message to the receiver with probability 1/2, while the sender remains oblivious as to whether or not the receiver received the message. Rabin's oblivious transfer scheme is based on the RSA cryptosystem. In 1985, Even et al. [22] proposed a more useful OT called 1-out-of-2 OT (OT_1^2) to build protocols for secure multiparty computation.

Afterwards, it has been generalized to 1-out-of- m OT (OT_1^m) [23] where the receiver gets exactly one message without the sender getting to know which message was queried and the receiver getting to know anything about the other messages that were not retrieved. OT_1^m has become a fundamental tool in cryptography and is usually used as a black-box when constructing protocols.

Formally, we can describe an OT_1^m protocol as follows. There are 2 participants called the sender S and the receiver R . Specifically, S has m messages, and R has an index z . R wishes to receive the z th message of the sender's m messages without leaking z to S , while knowing nothing about the rest $m - 1$ messages. A simplified OT_1^m protocol can be presented as in Algorithm 1.

3. System Model

In this section, we present the system model including the goals we aim to achieve in detail.

In this paper, we consider the privacy-preserving medical diagnosis system with two participants: the patient and the hospital. We assume that each patient can collect his own health data such as heart beat and blood pressure, in the form of a vector, called query vector, easily by various advanced medical devices. Herein, we call the heart beat, blood pressure, and so forth, as parameter items and the health data corresponding to heart beat, blood pressure, and so forth, as parameter values. For example, $q_P = (q_{P1}, \dots, q_{Pn})$ is the query vector of the patient P , where all $\{q_{Pw}\}_{w=1, \dots, n}$ are

the necessary parameters the hospital needs for diagnosis, and q_{Pw} is the parameter value of the parameter item heart beat. Each patient has a health card to store related data and a portable device PMDD to read the data stored in the card and to do some related computations after inserting the card. The hospital has a disease database $DB = \{d_i\}_{i=1, \dots, m}$, which in fact is the standard to determine which disease the patient has got. Each record of the disease database is presented as a triple $d_i = (i, t_i, r_i)$, $i = 1, \dots, m$, where m is the capacity of the disease database; i is the index of a disease; t_i , called the trait vector of the disease d_i , is a vector that covers all necessary parameters the hospital needs for diagnosis; and r_i is the disease diagnostic report including the disease name, doctors' advices, and prescriptions corresponding to the i th disease d_i . Concerning these parameters, we have some illustrations as follows.

- (i) q_P : it includes all necessary parameter items the hospital needs for diagnosis such as heart beat and blood pressure. The query vectors of the same patient P are different if P goes to different hospitals since their medical levels are different. The query vectors of the same patient P may be also different if P goes to the same hospital at different time since the medical level of the hospital has been always keeping improving. The dimension n and every parameter item $\{q_{Pw}\}_{w=1, \dots, n}$ of the query vector are determined by the trait vector t_i of the hospital. After registering to the hospital, patient can know what parameter items are needed in this diagnosis by reading the health card.
- (ii) $DB = \{d_i\}_{i=1, \dots, m}$: it concludes all diseases a hospital can diagnose. Different hospitals have different disease databases and the same hospital has different disease databases at different time since its medical level has been keeping improving. The dimension m is determined by the hospital's medical level.
- (iii) t_i : it includes all necessary parameter items the hospital needs for diagnosis such as heart beat and blood pressure. Different hospitals have different trait vectors and the trait vectors of the same hospital at different time may be different since the medical level of the hospital has been always keeping improving. The dimension n and every parameter item $\{t_{ij}\}_{j=1, \dots, n}$ of the trait vector are determined by the hospital's medical level. In order to improve the precision of our diagnosis system, the hospital can consider as many



FIGURE 1: Self-helped medical diagnosis model of our scheme.

factors as possible such as adding more personal feelings, symptoms, and previous medical features from the patient as parameter items. Although we only can diagnose some simple diseases currently, it is believed that it will be feasible for more complicated diseases in the future by extending the dimension n of the parameter items.

- (iv) r_i : it includes the disease name, doctors' advices, and prescriptions corresponding to the i th disease d_i . Each report may conclude many doctors' advices and prescriptions. Herein, we assume that every report obtained from the MD-ATM following the self-helped medical diagnosis is authorized by the hospital and all advices and prescriptions of a report are signed by corresponding doctors. After receiving the diagnostic report, patient can choose one doctor's advice and prescription to treat himself.

In this paper, the system makes medical diagnosis according to the Euclidean distances of two vectors. Specifically, taking the query vector of the patient and a trait vector of the database as an example, given a patient's query vector $q_P = (q_{P1}, \dots, q_{Pn})$ and a disease trait vector $t_i = (t_{i1}, \dots, t_{in})$, $i \in \{1, \dots, m\}$, their Euclidean distance [3] denoted by $\text{dist}_{q_P \leftrightarrow t_i}$ is

$$\text{dist}_{q_P \leftrightarrow t_i} = \sqrt{\sum_{w=1}^n (q_{Pw} - t_{iw})^2}. \quad (2)$$

Herein, we compare the squares of the Euclidean distances,

$$\text{dist}_{q_P \leftrightarrow t_i}^2 - \text{dist}_{q_P \leftrightarrow t_j}^2 = \sum_{w=1}^n (q_{Pw} - t_{iw})^2 - \sum_{w=1}^n (q_{Pw} - t_{jw})^2. \quad (3)$$

It is obvious that we can figure out which one has smaller distance with patient's query vector just by checking the sign of (3) without exact result of $\text{dist}_{q_P \leftrightarrow t_i}$ or $\text{dist}_{q_P \leftrightarrow t_j}$. Assuming that the report r_z corresponding to the trait vector

t_z , $z \in \{1, \dots, m\}$, is the diagnosed disease report, we have the following result, for all $j = 1, \dots, m$, $j \neq z$:

$$\text{dist}_{q_P \leftrightarrow t_z}^2 - \text{dist}_{q_P \leftrightarrow t_j}^2 < 0. \quad (4)$$

In our scheme, we will compare the squares of the Euclidean distances of the query vector and the trait vectors to find the diagnostic report that satisfies (4).

In real application, the hospital provides a MD-ATM, which is connected with the disease database and can read the data of the card, to direct patients to complete self-helped disease diagnosis. Specifically, we assume that each patient registers to the hospital for the first time and gets a health card. The hospital provides a self-helped MD-ATM in public just like a bank ATM. Whenever P wants to have a diagnosis, inserting his health card into the MD-ATM and following the instructions, P can complete the self-helped diagnosis by himself. The basic model can be illustrated in Figure 1.

Apart from the above, to enable a privacy-preserving medical diagnosis system, our scheme should simultaneously fulfill the following two security goals.

- (i) Confidentiality of disease database should be protected during the self-helped diagnosis process.
- (ii) Confidentiality of patient's private health data should be protected during the self-helped diagnosis process.

4. Our Scheme

In this section, we propose our privacy-preserving self-helped medical diagnosis scheme (PP-SH-MDS) in detail to show how a patient can diagnose by himself using his PMDD and the self-helped MD-ATM. The core of our construction can be summarized in Figure 2.

Specifically, the patient P executes as follows to make a self-helped diagnosis using his PMDD and the MD-ATM.

In the setup phase, P registers to a hospital as traditional medical diagnosis and gets a health card.

In the diagnosis phase, there are three subphases.

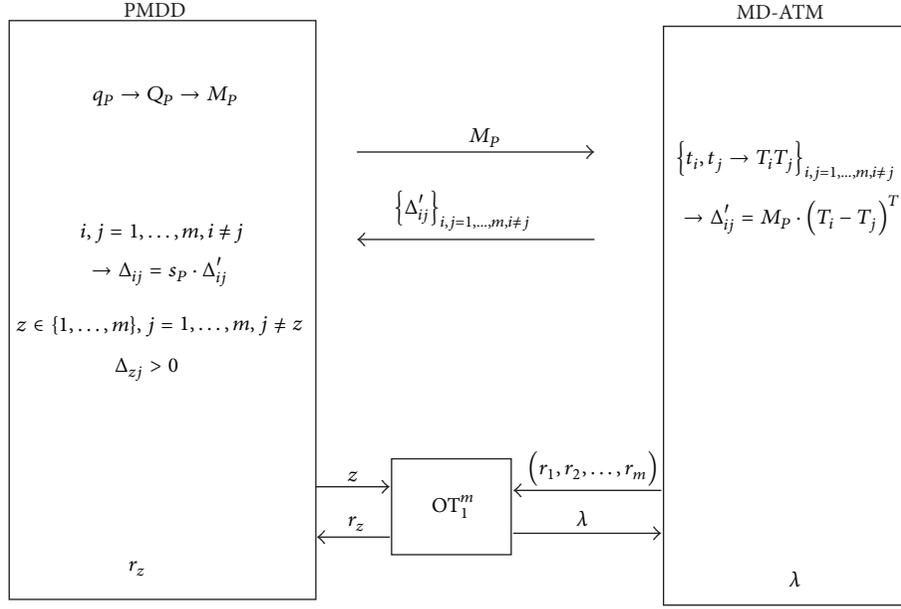


FIGURE 2: PP-SH-MDS.

(1) *Local Preprocessing.* Whenever P wants to have a diagnosis, he firstly conducts the following two transformations on PMDD locally.

(i) Vector-to-Vector.

(a) P firstly extends his original health data from an n -vector $q_P = (q_{P1}, \dots, q_{Pn})$ to an $(n+2)$ -vector $Q_P = (q_{P1}, \dots, q_{Pn}, q_{P,n+1}, q_{P,n+2})$, where $q_{P,n+1} = -(1/2) \sum_{w=1}^n q_{Pw}^2$ and $q_{P,n+2} = 1$.

(ii) Vector-to-Matrix.

(a) P randomly chooses a password $s_P = (s_{P1}, \dots, s_{Pk})$ and then generates a $k \times (n+2)$ matrix

$$B_P = \begin{bmatrix} b_{11} & \cdots & b_{1,n+2} \\ \vdots & & \vdots \\ b_{k1} & \cdots & b_{k,n+2} \end{bmatrix}, \quad (5)$$

where $\sum_{u=1}^k s_{Pu} \cdot b_{uw} = 1$, $w = 1, \dots, n+2$.

(b) By blinding Q_P using the matrix B_P , P further extends Q_P to a matrix

$$M_P = \begin{bmatrix} b_{11} \cdot q_{P1} & \cdots & b_{1,n+2} \cdot q_{P,n+2} \\ \vdots & & \vdots \\ b_{k1} \cdot q_{P1} & \cdots & b_{k,n+2} \cdot q_{P,n+2} \end{bmatrix}. \quad (6)$$

After completing the above steps, P stores the matrix M_P in the health card.

(2) *Diagnosis.*

(i) After local preprocessing, P inserts his health card into the MD-ATM and then the MD-ATM reads the card to get the matrix M_P and randomly chooses two trait vectors $t_i = (t_{i1}, \dots, t_{in})$ and $t_j = (t_{j1}, \dots, t_{jn})$ and, respectively, extends them to $(n+2)$ -vectors $T_i = (t_{i1}, \dots, t_{in}, t_{i,n+1}, t_{i,n+2})$ and $T_j = (t_{j1}, \dots, t_{jn}, t_{j,n+1}, t_{j,n+2})$, where $t_{i,n+1} = 1$, $t_{i,n+2} = -(1/2) \sum_{w=1}^n t_{iw}^2$ and $t_{j,n+1} = 1$, $t_{j,n+2} = -(1/2) \sum_{w=1}^n t_{jw}^2$. Then the MD-ATM computes $\Delta'_{ij} = M_P \cdot (T_i - T_j)^T$, $i, j = 1, \dots, m, i \neq j$, and writes (Δ'_{ij}, i, j) in the card and indicates the patient to get back his card.

(ii) Following the instructions of the MD-ATM, the patient gets the card back and inserts it into PMDD. After inputting his password, PMDD begins to compute $\Delta_{ij} = s_P \cdot \Delta'_{ij}$, $i, j = 1, \dots, m, i \neq j$, and finds the index z so that for all $j = 1, \dots, m, j \neq z, \Delta_{ij} > 0$. z is the input of the following OT₁^m protocol.

(3) *1-out-of- m OT Protocol.*

(i) P inserts his card into the MD-ATM and invokes an OT₁^m protocol, where P 's input is the index z and the MD-ATM's input is the diagnostic report set (r_1, r_2, \dots, r_m) of the database.

After executing the OT₁^m protocol, P gets the diagnostic report r_z corresponding to the disease d_z according to the index z , while the MD-ATM gets *Null*, denoted by λ .

5. Analysis

In this section, we analyze our scheme in detail. We firstly have a look at the correctness and then give a strict security proof following the real-ideal simulation paradigm of STC in the scenarios of semihonest adversaries.

5.1. Correctness. In this aspect, we follow the steps of our scheme and make sure that the patient indeed finds out the most possible disease from the disease database of the hospital using his health data by comparing Euclidean distances.

Following the scheme, we can see that the patient transforms the health data in two steps. Firstly, he extends his original health data from an n -vector $q_P = (q_{P1}, \dots, q_{Pn})$ to an $(n+2)$ -vector $Q_P = (q_{P1}, \dots, q_{Pn}, q_{P,n+1}, q_{P,n+2})$, where $q_{P,n+1} = -(1/2) \sum_{w=1}^n q_{Pw}^2$ and $q_{P,n+2} = 1$ and then blinds and extends Q_P using the matrix B_P to a matrix

$$M_P = \begin{bmatrix} b_{11} \cdot q_{P1} & \cdots & b_{1,n+2} \cdot q_{P,n+2} \\ \vdots & & \vdots \\ b_{k1} \cdot q_{P1} & \cdots & b_{k,n+2} \cdot q_{P,n+2} \end{bmatrix}. \quad (7)$$

On the other hand, the MD-ATM randomly selects two trait vectors $t_i = (t_{i1}, \dots, t_{in})$ and $t_j = (t_{j1}, \dots, t_{jn})$ and, respectively, extends them to $(n+2)$ -vectors $T_i = (t_{i1}, \dots, t_{in}, t_{i,n+1}, t_{i,n+2})$ and $T_j = (t_{j1}, \dots, t_{jn}, t_{j,n+1}, t_{j,n+2})$, where $t_{i,n+1} = 1$, $t_{i,n+2} = -(1/2) \sum_{w=1}^n t_{iw}^2$ and $t_{j,n+1} = 1$, $t_{j,n+2} = -(1/2) \sum_{w=1}^n t_{jw}^2$. After receiving M_P , for $i, j = 1, \dots, m$, $i \neq j$, the MD-ATM computes

$$\begin{aligned} \Delta'_{ij} &= M_P \cdot (T_i - T_j)^T \\ &= \begin{bmatrix} b_{11} \cdot q_{P1} & \cdots & b_{1,n+2} \cdot q_{P,n+2} \\ \vdots & & \vdots \\ b_{k1} \cdot q_{P1} & \cdots & b_{k,n+2} \cdot q_{P,n+2} \end{bmatrix} \cdot \begin{bmatrix} t_{i1} - t_{j1} \\ \vdots \\ t_{i,n+2} - t_{j,n+2} \end{bmatrix}. \end{aligned} \quad (8)$$

After receiving the returned message Δ'_{ij} , for $i, j = 1, \dots, m$, $i \neq j$, P uses PMDD to compute

$$\begin{aligned} \Delta_{ij} &= s_P \cdot \Delta'_{ij} = s_P \cdot M_P \cdot (T_i - T_j)^T \\ &= (s_{P1}, \dots, s_{Pk}) \cdot \begin{bmatrix} b_{11} \cdot q_{P1} & \cdots & b_{1,n+2} \cdot q_{P,n+2} \\ \vdots & & \vdots \\ b_{k1} \cdot q_{P1} & \cdots & b_{k,n+2} \cdot q_{P,n+2} \end{bmatrix} \\ &\quad \cdot \begin{bmatrix} t_{i1} - t_{j1} \\ \vdots \\ t_{i,n+2} - t_{j,n+2} \end{bmatrix} \end{aligned}$$

$$\begin{aligned} &= \sum_{u=1}^k s_{Pu} \cdot b_{u1} \cdot q_{P1} \cdot (t_{i1} - t_{j1}) + \cdots \\ &\quad + \sum_{u=1}^k s_{Pu} \cdot b_{u,n+2} \cdot q_{P,n+2} \cdot (t_{i,n+2} - t_{j,n+2}) \\ &= \sum_{w=1}^{n+2} q_{Pw} \cdot (t_{iw} - t_{jw}). \end{aligned} \quad (9)$$

Thus, we have

$$\begin{aligned} \Delta_{ij} &= \sum_{w=1}^{n+2} q_{Pw} \cdot t_{iw} - \sum_{w=1}^{n+2} q_{Pw} \cdot t_{jw} \\ &= \left(\sum_{w=1}^n q_{Pw} \cdot t_{iw} + q_{P,n+1} \cdot t_{i,n+1} + q_{P,n+2} \cdot t_{i,n+2} \right) \\ &\quad - \left(\sum_{w=1}^n q_{Pw} \cdot t_{jw} + q_{P,n+1} \cdot t_{j,n+1} + q_{P,n+2} \cdot t_{j,n+2} \right) \\ &= \left(\sum_{w=1}^n q_{Pw} \cdot t_{iw} + q_{P,n+1} + t_{i,n+2} \right) \\ &\quad - \left(\sum_{w=1}^n q_{Pw} \cdot t_{jw} + q_{P,n+1} + t_{j,n+2} \right) \\ &= \left(\sum_{w=1}^n q_{Pw} \cdot t_{iw} - \frac{1}{2} \sum_{w=1}^n q_{Pw}^2 - \frac{1}{2} \sum_{w=1}^n t_{iw}^2 \right) \\ &\quad - \left(\sum_{w=1}^n q_{Pw} \cdot t_{jw} - \frac{1}{2} \sum_{w=1}^n q_{Pw}^2 - \frac{1}{2} \sum_{w=1}^n t_{jw}^2 \right) \\ &= \frac{1}{2} \left[\sum_{w=1}^n (q_{Pw} - t_{jw})^2 - \sum_{w=1}^n (q_{Pw} - t_{iw})^2 \right] \\ &= \frac{1}{2} \left(\text{dist}_{q_P \leftrightarrow t_j}^2 - \text{dist}_{q_P \leftrightarrow t_i}^2 \right). \end{aligned} \quad (10)$$

Obviously, if for all $j = 1, \dots, m$, $j \neq z$, $\Delta_{zj} = (1/2) (\text{dist}_{q_P \leftrightarrow t_j}^2 - \text{dist}_{q_P \leftrightarrow t_z}^2) > 0$, then t_z , $z \in \{1, \dots, m\}$, is the trait vector of the diagnosed disease. The report r_z corresponding to the trait vector t_z is the diagnosed report. Taking the index z as the input of the following OT_1^m protocol, the patient can finally get the disease report r_z from the set (r_1, r_2, \dots, r_m) of the database.

Therefore, our scheme is correct.

5.2. Security. In this subsection, we strictly prove the security of our scheme. From the whole process, we can specify that the two parties in our system are the patient P and the hospital. They cooperate to compute the function $f(q_P, (d_1, d_2, \dots, d_m)) = r_z$, where r_z is the disease diagnostic report corresponding to the disease d_z and the distance $\text{dist}_{q_P \leftrightarrow t_z}$ satisfies the condition $\text{dist}_{q_P \leftrightarrow t_z}^2 = \min\{\text{dist}_{q_P \leftrightarrow t_j}^2\}_{j=1, \dots, m}$. As mentioned in Section 3, we should

achieve two security goals, that is, keeping both parties' inputs private. We apply the real-ideal simulation paradigm to prove that our scheme has achieved the two goals in the scenarios of semihonest adversaries assuming the OT_1^m protocol we used is secure.

Theory 1. *Our privacy-preserving self-helped medical diagnosis scheme is secure against semihonest adversaries if the OT_1^m protocol is secure.*

Proof. Notice that the view of P_i , $\{\text{VIEW}_i(q_p, \{d_1, \dots, d_m\})\}_{i=1,2}$, in the real execution consists of three parts, the private input, random tape, and the messages received from the other party including the output. Therefore, we can get the views of P_1 and P_2 , respectively, in the real execution as follows:

$$\begin{aligned} & \{\text{VIEW}_1(q_p, \{d_i\}_{i=1, \dots, m})\} \\ &= \left\{ q_p, s_p, B_p, \{\Delta'_{ij}\}_{i,j=1, \dots, m, i \neq j} \right\} \end{aligned} \quad (11)$$

$$\begin{aligned} & \{\text{VIEW}_1^{\text{OT}_1^m}(z, \{r_i\}_{i=1, \dots, m})\}, r_z \}; \\ & \{\text{VIEW}_2(q_p, \{d_i\}_{i=1, \dots, m})\} \\ &= \left\{ \{d_i\}_{i=1, \dots, m}, M_p, \{\text{VIEW}_2^{\text{OT}_1^m}(z, \{r_i\}_{i=1, \dots, m})\}, \lambda \right\}, \end{aligned} \quad (12)$$

where $\{\text{VIEW}_1^{\text{OT}_1^m}(z, \{r_i\}_{i=1, \dots, m})\}$ and $\{\text{VIEW}_2^{\text{OT}_1^m}(z, \{r_i\}_{i=1, \dots, m})\}$ are the views produced in the execution of OT_1^m protocol.

From the definition of security, we need to construct the probabilistic polynomial-time algorithm S_1/S_2 so that given the input and output of the patient P_1 /the hospital P_2 , $(q_p; r_z)/(\{d_i\}_{i=1, \dots, m}; \lambda)$, it can output a simulated view $\{S_1(q_p; r_z)\}/\{S_2(\{d_i\}_{i=1, \dots, m}; \lambda)\}$, which is indistinguishable to the view $\{\text{VIEW}_1(q_p, \{d_i\}_{i=1, \dots, m})\}/\{\text{VIEW}_2(q_p, \{d_i\}_{i=1, \dots, m})\}$ in the real execution of the scheme; that is,

$$\begin{aligned} & \{S_1(q_p; r_z)\} \cong \{\text{VIEW}_1(q_p, \{d_i\}_{i=1, \dots, m})\}, \\ & \{S_2(\{d_i\}_{i=1, \dots, m}; \lambda)\} \cong \{\text{VIEW}_2(q_p, \{d_i\}_{i=1, \dots, m})\}. \end{aligned} \quad (13)$$

In the following discussion, we follow the real-ideal simulation paradigm to construct such probabilistic polynomial-time algorithms S_1, S_2 . We separately prove the case when P_2 is semihonest and when P_1 is semihonest.

Case 1 (P_2 is semihonest). In this case, we only need to construct a simulator S_2 so that, given P_2 's input $\{d_i\}_{i=1, \dots, m}$ and output λ , S_2 can simulate P_2 's view in the real execution presented above as (12).

Firstly, since we assume that the OT_1^m protocol used in our scheme is secure and can be taken as a black-box, there exists an algorithm $S_2^{\text{OT}_1^m}$: given the input $\{r_i\}_{i=1, \dots, m}$ and the output λ , it can simulate P_2 's view of the OT_1^m execution and output $\{S_2^{\text{OT}_1^m}(\{r_i\}_{i=1, \dots, m}; \lambda)\}$ so that

$$\{S_2^{\text{OT}_1^m}(\{r_i\}_{i=1, \dots, m}; \lambda)\} \cong \{\text{VIEW}_2^{\text{OT}_1^m}(z, \{r_i\}_{i=1, \dots, m})\}. \quad (14)$$

Next, notice that S_2 is given $(\{d_i\}_{i=1, \dots, m}; \lambda)$; it can easily simulate the remaining parts of (12) by randomly choosing a $k \times (n+2)$ matrix M which is indistinguishable to the blinded matrix M_p . Then, S_2 outputs the simulated view,

$$\begin{aligned} & \{S_2(\{d_i\}_{i=1, \dots, m}; \lambda)\} \\ &= \left\{ \{d_i\}_{i=1, \dots, m}, M, \{S_2^{\text{OT}_1^m}(\{r_i\}_{i=1, \dots, m}; \lambda)\}, \lambda \right\}. \end{aligned} \quad (15)$$

Obviously, we can conclude that

$$\{S_2(\{d_i\}_{i=1, \dots, m}; \lambda)\} \cong \{\text{VIEW}_2(q_p, \{d_i\}_{i=1, \dots, m})\}. \quad (16)$$

Case 2 (P_1 is semihonest). Similar to Case 1, we only need to construct a simulator S_1 so that given P_1 's input q_p and output r_z , S_1 can simulate P_1 's view in the real execution presented above as (11).

As discussed above, since the OT_1^m protocol is secure, there exists an algorithm $S_1^{\text{OT}_1^m}$: given the input z and the output r_z , it can simulate P_1 's view of the OT_1^m execution and output $\{S_1^{\text{OT}_1^m}(z; r_z)\}$ so that

$$\{S_1^{\text{OT}_1^m}(z; r_z)\} \cong \{\text{VIEW}_1^{\text{OT}_1^m}(z, \{r_i\}_{i=1, \dots, m})\}. \quad (17)$$

Next, given $(q_p; r_z)$, S_1 then simulates the remaining parts of P_1 's view in the real execution as follows.

Firstly, as in the real execution, S_1 extends the original health data from an n -vector $q_p = (q_{p1}, \dots, q_{pn})$ to an $(n+2)$ -vector $Q_p = (q_{p1}, \dots, q_{pn}, q_{p,n+1}, q_{p,n+2})$, where $q_{p,n+1} = -(1/2) \sum_{w=1}^n q_{pw}^2$ and $q_{p,n+2} = 1$. Then, S_1 randomly chooses a password $s'_p = (s'_{p1}, \dots, s'_{pk})$ and then generates a $k \times (n+2)$ matrix

$$B'_p = \begin{bmatrix} b'_{11} & \cdots & b'_{1,n+2} \\ \vdots & & \vdots \\ b'_{k1} & \cdots & b'_{k,n+2} \end{bmatrix}, \quad (18)$$

where $\sum_{u=1}^k s'_{pu} \cdot b'_{uw} = 1$, $w = 1, \dots, n+2$. By blinding Q_p using the matrix B'_p , S_1 further extends Q_p to a matrix

$$M'_p = \begin{bmatrix} b'_{11} \cdot q_{p1} & \cdots & b'_{1,n+2} \cdot q_{p,n+2} \\ \vdots & & \vdots \\ b'_{k1} \cdot q_{p1} & \cdots & b'_{k,n+2} \cdot q_{p,n+2} \end{bmatrix}. \quad (19)$$

Then S_1 randomly selects m vectors, $T'_j = (t'_{j1}, \dots, t'_{jn}, t'_{j,n+1}, t'_{j,n+2})$, $t'_{j,n+1} = 1$, $t'_{j,n+2} = -(1/2) \sum_{w=1}^n t'_{jw}{}^2$, $j = 1, \dots, m$, and for all $j = 1, \dots, m$, $j \neq z$, $\sum_{w=1}^n (q_{pw} - t'_{zw})^2 < \sum_{w=1}^n (q_{pw} - t'_{jw})^2$; otherwise, S_1 reselects T'_j , $j = 1, \dots, m$. Afterwards, S_1 computes $\Lambda'_{ij} = M'_p \cdot (T'_i - T'_j)^T$, $i, j = 1, \dots, m$, $i \neq j$. Thus, for $i, j = 1, \dots, m$, $i \neq j$, $\Lambda_{ij} = s'_p \cdot \Lambda'_{ij} = \sum_{w=1}^{n+2} q_{pw} \cdot (t'_{iw} - t'_{jw})$.

Therefore, we have

$$\begin{aligned}
\Lambda_{ij} &= \sum_{w=1}^{n+2} q_{Pw} \cdot t'_{iw} - \sum_{w=1}^{n+2} q_{Pw} \cdot t'_{jw} \\
&= \left(\sum_{w=1}^n q_{Pw} \cdot t'_{iw} + q_{P,n+1} \cdot t'_{i,n+1} + q_{P,n+2} \cdot t'_{i,n+2} \right) \\
&\quad - \left(\sum_{w=1}^n q_{Pw} \cdot t'_{jw} + q_{P,n+1} \cdot t'_{j,n+1} + q_{P,n+2} \cdot t'_{j,n+2} \right) \\
&= \left(\sum_{w=1}^n q_{Pw} \cdot t'_{iw} + q_{P,n+1} \cdot t'_{i,n+2} \right) \\
&\quad - \left(\sum_{w=1}^n q_{Pw} \cdot t'_{jw} + q_{P,n+1} \cdot t'_{j,n+2} \right) \quad (20) \\
&= \left(\sum_{w=1}^n q_{Pw} \cdot t'_{iw} - \frac{1}{2} \sum_{w=1}^n q_{Pw}^2 - \frac{1}{2} \sum_{w=1}^n t'^2_{iw} \right) \\
&\quad - \left(\sum_{w=1}^n q_{Pw} \cdot t'_{jw} - \frac{1}{2} \sum_{w=1}^n q_{Pw}^2 - \frac{1}{2} \sum_{w=1}^n t'^2_{jw} \right) \\
&= \frac{1}{2} \left[\sum_{w=1}^n (q_{Pw} - t'_{jw})^2 - \sum_{w=1}^n (q_{Pw} - t'_{iw})^2 \right] \\
&= \frac{1}{2} \left(\text{dist}_{q_p \leftrightarrow t'_j}^2 - \text{dist}_{q_p \leftrightarrow t'_i}^2 \right).
\end{aligned}$$

Since for all $j = 1, \dots, m$, $j \neq z$, we have $\sum_{w=1}^n (q_{Pw} - t'_{zw})^2 < \sum_{w=1}^n (q_{Pw} - t'_{jw})^2$. Obviously, for all $j = 1, \dots, m$, $j \neq z$, $\Lambda_{zj} = (1/2)(\text{dist}_{q_p \leftrightarrow t'_j}^2 - \text{dist}_{q_p \leftrightarrow t'_z}^2) > 0$ and t_z , $z \in \{1, \dots, m\}$, is the trait vector of the diagnosed disease. The report r_z corresponding to the trait vector t_z is the diagnosed report, which matches the relationship in the real execution.

Now, S_1 can output the simulated view,

$$\begin{aligned}
\{S_1(q_p; r_z)\} \\
= \left\{ q_p, s'_p, B'_p, \{\Lambda'_{ij}\}_{i,j=1,\dots,m,i \neq j}, \{S_1^{\text{OT}_1^m}(z; r_z)\}, r_z \right\}. \quad (21)
\end{aligned}$$

Since s'_p and B'_p are randomly chosen and $\sum_{u=1}^k s'_{Pu} \cdot b'_{uw} = 1$, $w = 1, \dots, n+2$, as s_p and B_p in the real view $\{\text{VIEW}_1(q_p, \{d_i\}_{i=1,\dots,m})\}$. Due to the randomness and relationship, it is easy to find that (s'_p, B'_p) is indistinguishable to (s_p, B_p) . From the construction process of $\{\Lambda'_{ij}\}_{i,j=1,\dots,m,i \neq j}$, it is obvious to conclude that $\{\Lambda'_{ij}\}_{i,j=1,\dots,m,i \neq j}$ is indistinguishable to the set $\{\Lambda'_{ij}\}_{i,j=1,\dots,m,i \neq j}$. Combined with (17), we have

$$\{S_1(q_p; r_z)\} \cong \{\text{VIEW}_1(q_p, \{d_i\}_{i=1,\dots,m})\}. \quad (22)$$

□

6. Conclusions

In this paper, we consider the problem of how to securely make diagnosis without leaking patient's health data, diagnosed result, and hospital's disease database in wireless sensor networks. By applying the idea of secure two-party computation and the technology of oblivious transfer, we propose a privacy-preserving self-helped medical diagnosis scheme so that patients can privately diagnose themselves by inserting a health card into a self-helped MD-ATM to obtain the diagnostic report just like drawing money from a bank ATM. We also have a detailed analysis about the correctness and further strictly prove the security following the real-idea simulation paradigm. We expect to provide people another perspective on future medical care.

Conflict of Interests

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

Acknowledgments

This work is supported by NSFC (Grant nos. 61300181, 61272057, 61202434, 61170270, 61100203, and 61121061) and the Fundamental Research Funds for the Central Universities (Grant nos. 2012RC0612 and 2011YB01).

References

- [1] J. Wan, S. Ullah, C. Lai et al., "Cloud-enabled wireless body area networks for pervasive healthcare," *IEEE Network*, vol. 27, no. 5, pp. 56–61, 2013.
- [2] C. Doukas, T. Pliakas, and I. Maglogiannis, "Mobile healthcare information management utilizing Cloud Computing and Android OS," in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10)*, pp. 1037–1040, September 2010.
- [3] J. Yuan and S. Yu, "Efficient privacy-preserving biometric identification in cloud computing," in *Proceedings of the 32nd IEEE International Conference on Computer Communications (INFOCOM '13)*, pp. 2652–2660, 2013.
- [4] J. Liu, Q. Wang, J. Wan, J. Xiong, and B. Zeng, "Towards key issues of disaster aid based on wireless body area networks," *KSII Transactions on Internet and Information Systems*, vol. 7, no. 5, pp. 1014–1035, 2013.
- [5] M. Al Ameen, J. Liu, and K. Kwak, "Security and privacy issues in wireless sensor networks for healthcare applications," *Journal of Medical Systems*, vol. 36, no. 1, pp. 93–101, 2012.
- [6] P. M. Schwartz and J. R. Reidenberg, *Data Privacy Law: A Study of United States Data Protection*, LEXIS Law, Dayton, Ohio, USA, 1996.
- [7] H. Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Stanford University, Stanford, Calif, USA, 2010.
- [8] C. Gentry, *A fully homomorphic encryption scheme [doctoral dissertation]*, Stanford University.
- [9] Z. Brakerski and V. Vaikuntanathan, "Efficient fully homomorphic encryption from (standard) LWE," in *Proceedings of the IEEE 52nd Annual Symposium on Foundations of Computer*

- Science (FOCS '11)*, pp. 97–106, Palm Springs, Calif, USA, October 2011.
- [10] A. López-Alt, E. Tromer, and V. Vaikuntanathan, “On-the-fly multiparty computation on the cloud via multikey fully homomorphic encryption,” in *Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC '12)*, pp. 1219–1234, ACM, May 2012.
- [11] J. Bringer, H. Chabanne, and A. Patey, “Privacy-preserving biometric identification using secure multiparty computation: an overview and recent trends,” *IEEE Signal Processing Magazine*, vol. 30, no. 2, pp. 42–52, 2013.
- [12] A. C. Yao, “Protocols for secure computations,” in *Proceedings of the 23rd Annual IEEE Symposium on Foundations of Computer Science*, pp. 160–164, Chicago, Ill, USA, 1982.
- [13] O. S. Goldreich, S. Micali, and A. Wigderson, “How to play any mental game,” in *Proceedings of the 19th Annual ACM Symposium on Theory of Computing (STOC '87)*, pp. 218–229, ACM, New York, NY, USA, 1987.
- [14] O. S. Goldreich, “Secure multiparty computation,” Manuscript, Preliminary version, 1998.
- [15] O. Goldreich, *Foundations of Cryptography: Volume 2, Basic Applications*, Cambridge University Press, Cambridge, UK, 2004.
- [16] M. M. Prabhakaran and A. Sahai, Eds., *Secure Multiparty Computation*, IOS Press, 2013.
- [17] D. Chaum, C. Crepeau, and I. Damgard, “Multi-party unconditionally secure protocols (extended abstract),” in *Proceedings of the STOC*, pp. 11–19, ACM, 1988.
- [18] I. Damgard, V. Pastro, N. P. Smart, and S. Zakarias, “Multiparty computation from somewhat homomorphic encryption,” in *Advances in Cryptology—Crypto 2012*, vol. 7417 of *Lecture Notes in Computer Science*, pp. 643–662, Springer, 2012.
- [19] Y. Lindell and B. Pinkas, “An efficient protocol for secure two-party computation in the presence of malicious adversaries,” in *Advances in Cryptology—EUROCRYPT 2007*, vol. 4515 of *Lecture Notes in Computer Science*, pp. 52–78, Springer, Berlin, Germany, 2007.
- [20] B. Pinkas, T. Schneider, N. P. Smart, and S. C. Williams, “Secure two-party computation is practical,” in *Advances in Cryptology—ASIACRYPT 2009*, vol. 5912 of *Lecture Notes in Computer Science*, pp. 250–267, Springer, Berlin, Germany, 2009.
- [21] M. O. Rabin, “How to exchange secrets by oblivious transfer,” Tech. Rep. TR-81, Aiken Computation Laboratory, Harvard University, 1981.
- [22] S. Even, O. Goldreich, and A. Lempel, “A randomized protocol for signing contracts,” *Communications of the ACM*, vol. 28, no. 6, pp. 637–647, 1985.
- [23] M. Naor and B. Pinkas, “Oblivious transfer with adaptive queries,” in *Advances in Cryptology—CRYPTO'99*, vol. 1666 of *Lecture Notes in Computer Science*, pp. 573–590, Springer, 1999.

Research Article

Mixed-Norm Regularization for Brain Decoding

R. Flamary,¹ N. Jrad,² R. Phlypo,² M. Congedo,² and A. Rakotomamonjy³

¹ *Laboratoire Lagrange, UMR7293, Université de Nice, 00006 Nice, France*

² *Gipsa Lab, Domaine Universitaire BP 46, 38402 Saint Martin d'Hères, France*

³ *LITIS, EA 4108-INSA, Université de Rouen, 76000 Rouen, France*

Correspondence should be addressed to R. Flamary; remi.flamary@unice.fr and A. Rakotomamonjy; alain.rakoto@insa-rouen.fr

Received 9 January 2014; Accepted 13 March 2014; Published 17 April 2014

Academic Editor: David A. Elizondo

Copyright © 2014 R. Flamary 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 work investigates the use of mixed-norm regularization for sensor selection in event-related potential (ERP) based brain-computer interfaces (BCI). The classification problem is cast as a discriminative optimization framework where sensor selection is induced through the use of mixed-norms. This framework is extended to the multitask learning situation where several similar classification tasks related to different subjects are learned simultaneously. In this case, multitask learning helps in leveraging data scarcity issue yielding to more robust classifiers. For this purpose, we have introduced a regularizer that induces both sensor selection and classifier similarities. The different regularization approaches are compared on three ERP datasets showing the interest of mixed-norm regularization in terms of sensor selection. The multitask approaches are evaluated when a small number of learning examples are available yielding to significant performance improvements especially for subjects performing poorly.

1. Introduction

Brain computer interfaces (BCI) are systems that help disabled people communicate with their environment through the use of brain signals [1]. At the present time, one of the most prominent BCI is based on electroencephalography (EEG) because of its low-cost, portability, and its noninvasiveness. Among EEG based BCI, a paradigm of interest is the one based on event-related potentials (ERP) which are responses of the brain to some external stimuli. In this context, the innermost part of a BCI is the pattern recognition stage which has to correctly recognize presence of these ERPs. However, EEG signals are blurred due to the diffusion of the skull and the skin [2]. Furthermore, EEG recordings are highly contaminated by noise of biological, instrumental, and environmental origins. For addressing these issues, advanced signal processing and machine learning techniques have been employed to learn ERP patterns from training EEG signals leading to robust systems able to recognize the presence of these events [3–8]. Note that while some ERPs are used for generating BCI commands, some others can be used for improving BCI efficiency. Indeed, recent studies have also

tried to develop algorithms for automated recognition of error-related potentials [9]. These potentials are responses elicited when a subject commits an error in a BCI task or observes an error [10, 11] and thus they can help in correcting errors or in providing feedbacks to BCI users.

In this context of automated recognition of event-related potentials for BCI systems, reducing the number of EEG sensors is of primary importance since it reduces the implementation cost of the BCI by minimizing the number of EEG sensor and speeding up experimental setup and calibration time. For this purpose, some studies have proposed to choose relevant sensors according to prior knowledge of brain functions. For instance, sensors located above the motor cortex region are preferred for motor imagery tasks, while for visual event-related potential (ERP), sensors located on the visual cortex are favored [12]. Recent works have focused on automatic sensor selection adapted to the specificity of a subject [4, 13–17]. For instance, Rakotomamonjy and Guigue [18] performed a recursive backward sensor selection using cross-validation classification performances as an elimination criterion. Another approach for exploring subset sensors has been proposed by [15]; it consists in using a genetic

algorithm for sensor selection coupled with artificial neural networks for prediction. Those methods have been proven efficient but computationally demanding. A quicker way is to estimate the relevance of the sensors in terms of signal to noise ratio (SNR) [4] and to keep the most relevant ones. Note that this approach does not optimize a discrimination criterion, although the final aim is a classification task. Recently, van Gerven et al. [19] proposed a graceful approach for embedding sensor selection into a discriminative framework. They performed sensor selection and learn a decision function by solving a unique optimization problem. In their framework, a logistic regression classifier is learned and the group-lasso regularization, also known as $\ell_1 - \ell_2$ mixed-norm, is used to promote sensor selection. They have also investigated the use of this groupwise regularization for frequency band selection and their applications to transfer learning. The same idea has been explored by Tomioka and Müller [20] which also considered groupwise regularization for classifying EEG signals. In this work, we go beyond these studies by providing an in-depth study of the use of mixed-norms for sensor selection in a single subject setting and by discussing the utility of mixed-norms when learning decision functions for multiple subjects simultaneously.

Our first contribution addresses the problem of robust sensor selection embedded into a discriminative framework. We broaden the analysis of van Gerven et al. [19] by considering regularizers whose forms are $\ell_1 - \ell_q$ mixed-norms, with ($1 \leq q \leq 2$), as well as adaptive mixed-norms, so as to promote sparsity among group of features or sensors. In addition to providing a sparse and accurate sensor selection, mixed-norm regularization has several advantages. First, sensor selection is cast into an elegant discriminative framework, using for instance a large margin paradigm, which does not require any additional hyperparameter to be optimized. Secondly, since sensor selection is jointly learned with the classifier by optimizing an “all-in-one” problem, selected sensors are directed to the goal of discriminating relevant EEG patterns. Hence, mixed-norm regularization helps locating sensors which are relevant for an optimal classification performance.

A common drawback of all the aforementioned sensor selection techniques is that selected set of sensors may vary, more or less substantially, from subject to subject. This variability is due partly to subject specific differences and partly to acquisition noise and limited number of training examples. In such a case, selecting a robust subset of sensors may become a complex problem. Addressing this issue is the point of our second contribution. We propose a multitask learning (MTL) framework that helps in learning robust classifiers able to cope with the scarcity of learning examples. MTL is one way of achieving inductive transfer between tasks. The goal of inductive transfer is to leverage additional sources of information to improve the performance of learning on the current task. The main hypothesis underlying MTL is that tasks are related in some ways. In most cases, this relatedness is translated into a prior knowledge, for example, a regularization term, that a machine learning algorithm can take advantage of. For instance, regularization terms may promote similarity between all the tasks [21] or enforce

classifier parameters to lie in a low dimensional linear subspace [22] or to jointly select the relevant features [23]. MTL has been proven efficient for motor imagery in [24] where several classifiers were learned simultaneously from several BCI subject datasets. Our second contribution is thus focused on the problem of performing sensor selection and learning robust classifiers through the use of an MTL mixed-norm regularization framework. We propose a novel regularizer promoting sensor selection and similarity between classifiers. By doing so, our goal is then to yield sensor selection and robust classifiers that are able to overcome the data scarcity problem by sharing information between the different classifiers to be learned.

The paper is organized as follows. The first part of the paper presents the discriminative framework and the different regularization terms we have considered for channel selection and multitask learning. The second part is devoted to the description of the datasets, the preprocessing steps applied to each of them, and the results achieved in terms of performances and sensor selection. In order to promote reproducible research, the code needed for generating the results in this paper is available on the author’s website (URL: <http://remi.flamary.com/soft/soft-gsvm.html>).

2. Learning Framework

In this section, we introduce our mixed-norm regularization framework that can be used to perform sensor selection in a single task or in a transfer learning setting.

2.1. Channel Selection in a Single Task Learning Setting. Typically in BCI problems, one wants to learn a classifier that is able to predict the class of some EEG trials, from a set of learning examples. We denoted as $\{\mathbf{x}_i, y_i\}_{i \in \{1, \dots, n\}}$ the learning set such that $\mathbf{x}_i \in \mathbb{R}^d$ is a trial and $y_i \in \{-1, 1\}$ is its corresponding class, usually related to the absence or presence of an event-related potential. In most cases, a trial \mathbf{x}_i is extracted from a multidimensional signal and thus is characterized by r features for each of the p sensors, leading to a dimensionality $d = r \times p$. Our aim is to learn, for a single subject, a linear classifier f that will predict the class of a trial $\mathbf{x} \in \mathbb{R}^d$, by looking at the sign of the function $f(\cdot)$ defined as

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{w} + b \quad (1)$$

with $\mathbf{w} \in \mathbb{R}^d$ the normal vector to the separating hyperplane and $b \in \mathbb{R}$ a bias term. Parameters of this function are learned by solving the optimization problem:

$$\min_{\mathbf{w}, b} \sum_i^n L_o(\mathbf{y}_i, \mathbf{x}_i^T \mathbf{w} + b) + \lambda \Omega(\mathbf{w}), \quad (2)$$

where L_o is a loss function that measures the discrepancy between actual and predicted labels, $\Omega(\cdot)$ is a regularization term that expresses some prior knowledge about the learning problem, and λ is a parameter that balances both terms. In this work, we choose L_o to be the squared hinge loss $L_o(y, \hat{y}) = \max(0, 1 - y\hat{y})^2$, thus promoting a large margin classifier.

2.1.1. Regularization Terms. We now discuss different regularization terms that may be used for single task learning along with their significances in terms of channel selection.

ℓ_2 Norm. The first regularization term that comes to mind is the standard squared ℓ_2 norm regularization:

$$\Omega_2(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|_2^2, \quad (3)$$

where $\|\cdot\|_2$ is the Euclidean norm. This is the common regularization term used for SVMs and it will be considered in our experiments as the baseline approach. Intuitively, this regularizer tends to downweigh the amplitude of each component of \mathbf{w} leading to a better control of the margin width of our large-margin classifier and thus it helps in reducing overfitting.

ℓ_1 Norm. When only few of the features are discriminative for a classification task, a common way to select the relevant ones is to use an ℓ_1 norm of the form

$$\Omega_1(\mathbf{w}) = \sum_{i=1}^d |w_i| \quad (4)$$

as a regularizer [25]. Owing to its mathematical properties (nondifferentiability at 0), unlike the ℓ_2 norm, this regularization term promotes sparsity, which means that at optimality of problem (2), some components of \mathbf{w} are exactly 0. In a Bayesian framework, the ℓ_1 norm is related to the use of prior on \mathbf{w} that forces its component to vanish [19]. This is typically obtained by means of Laplacian prior over the weight. However, ℓ_1 norm ignores the structure of the features (which may be grouped by sensors) since each component of w is considered independently to the others. As such, this norm precludes grouped feature selection and allows only for feature selection.

$\ell_1 - \ell_q$ Mixed-Norm. A way to take into account the fact that features are structured is to use a mixed-norm that will group them and regularize them together. Here, we consider mixed-norm of the form

$$\Omega_{1-q}(\mathbf{w}) = \sum_{g \in \mathcal{G}} \|\mathbf{w}_g\|_q \quad (5)$$

with $1 \leq q \leq 2$ and \mathcal{G} being a partition of the set $\{1, \dots, d\}$. Intuitively, this $\ell_1 - \ell_q$ mixed-norm can be interpreted as an ℓ_1 norm applied to the vector containing the ℓ_q norm of each group of features. It promotes sparsity on each \mathbf{w}_g norm and consequently on the \mathbf{w}_g components as well. For our BCI problem, a natural choice for \mathcal{G} is to group the features by sensors yielding thus to p groups (one per sensor) of r features as reported in Figure 1. Note that unlike the $\ell_1 - \ell_2$ norm as used by van Gerven et al. [19] and Tomioka and Müller [20], the use of an inner ℓ_q norm leads to more flexibility as it spans from the $\ell_1 - \ell_1$ (equivalent to the ℓ_1 -norm and leading thus to unstructured feature selection) to the $\ell_1 - \ell_2$ which strongly ties together the components of a group. Examples of the use of ℓ_q norm and mixed-norm regularizations in other biomedical contexts can be found for instance in [26, 27].

Adaptive $\ell_1 - \ell_q$. The ℓ_1 and $\ell_1 - \ell_q$ norms described above are well known to lead to grouped feature selection. However, they are also known to lead to poor statistical properties (at least when used with a square loss function) [28]. For instance, they are known to have consistency issue in the sense that, even with an arbitrarily large number of training examples, these norms may be unable to select the true subset of features. In practice, this means that when used in (2), the optimal weight vector \mathbf{w} will tend to overestimate the number of relevant sensors. These issues can be addressed by considering an adaptive $\ell_1 - \ell_q$ mixed-norm of the form [28, 29]

$$\Omega_{a:1-q}(\mathbf{w}) = \sum_{g \in \mathcal{G}} \beta_g \|\mathbf{w}_g\|_q, \quad (6)$$

where the weights β_g are selected so as to enhance the sparsity pattern of \mathbf{w} . In our experiments, we obtain them by first solving the $\ell_1 - \ell_q$ problem with $\beta_g = 1$, which outputs an optimal parameter \mathbf{w}^* , and by finally defining $\beta_g = 1/\|\mathbf{w}_g^*\|_q$. Then, solving the weighted $\ell_1 - \ell_q$ problem yields an optimal solution with increased sparsity pattern compared to \mathbf{w}^* since the β_g augments the penalization of groups with norm $\|\mathbf{w}_g^*\|_q$ smaller than 1.

2.1.2. Algorithms. Let us now discuss how problem (2) is solved when one of these regularizers is in play.

Using the ℓ_2 norm regularization makes the problem differentiable. Hence a first- or second-order descent based algorithm can be considered [30].

Because the other regularizers are not differentiable, we have deployed an algorithm [31] tailored for minimizing objective function of the form $f_1(\mathbf{w}) + f_2(\mathbf{w})$ with f_1 a smooth and differentiable convex function with Lipschitz constant L and f_2 a continuous and convex nondifferentiable function having a simple proximal operator, that is, a closed-form or an easy-to-compute solution of the problem

$$\text{prox}_{f_2}(\mathbf{v}) := \underset{\mathbf{u}}{\text{argmin}} \frac{1}{2} \|\mathbf{v} - \mathbf{u}\|_2^2 + f_2(\mathbf{u}). \quad (7)$$

Such an algorithm, known as forward-backward splitting [31], is simply based on the following iterative approach:

$$\mathbf{w}^{k+1} = \text{prox}_{(1/\gamma)f_2}(\mathbf{w}^k - \gamma \nabla_{\mathbf{w}} f_1(\mathbf{w}^k)) \quad (8)$$

with γ being a stepsize in the gradient descent. This algorithm can be easily derived by considering, instead of directly minimizing $f_1(\mathbf{w}) + f_2(\mathbf{w})$, an iterative scheme which at each iteration replaces f_1 with a quadratic approximation of $f_1(\cdot)$ in the neighborhood of \mathbf{w}^k . Hence, \mathbf{w}^{k+1} is the minimizer of

$$f_1(\mathbf{w}^k) + \langle \nabla_{\mathbf{w}} f_1(\mathbf{w}^k), \mathbf{w} - \mathbf{w}^k \rangle + \frac{\gamma}{2} \|\mathbf{w} - \mathbf{w}^k\|_2^2 + f_2(\mathbf{w}) \quad (9)$$

whose closed-form is given in (8). This algorithm is known to converge towards a minimizer of $f_1(\mathbf{w}) + f_2(\mathbf{w})$ under some weak conditions on the stepsize [31], which is satisfied by choosing for instance $\gamma = 1/L$. We can note that the algorithm defined in (8) has the same flavor as a projected

gradient algorithm which first takes a gradient step and then “projects” back the solution owing to the proximal operator. More details can also be found in [32].

For our problem (2), we choose $f_1(\mathbf{w})$ to be the squared hinge loss and $f_2(\mathbf{w})$ the nonsmooth regularizer. The square hinge loss is indeed gradient Lipschitz with a constant L being $2 \sum_{i=1}^n \|\mathbf{x}_i\|_2^2$. Proof of this statement is available in Appendix A. Proximal operators of the ℓ_1 and the $\ell_1 - \ell_2$ regularization term can be easily shown to be the soft-thresholding and the block-soft thresholding operator [25]. The general $\ell_1 - \ell_q$ norm does not admit a closed-form solution, but its proximal operator can be simply computed by means of an iterative algorithm [23]. More details on these proximal operators are also available in Appendix C.

2.2. Channel Selection and Transfer Learning in Multiple Task Setting. We now address the problem of channel selection in cases where training examples for several subjects are at our disposal. We have claimed that in such a situation, it would be beneficial to learn the decision functions related to all subjects simultaneously, while inducing selected channels to be alike for all subjects, as well as inducing decision function parameters to be related in some sense. These two hypotheses make reasonable sense since brain regions related to the appearance of a given ERP are expected to be somewhat location-invariant across subjects. For solving this problem, we apply a machine learning paradigm, known as multitask learning, where in our case, each task is related to the decision function of a given subject and where the regularizer should reflect the above-described prior knowledge on the problem. Given m subjects, the resulting optimization problem boils down to be

$$\min_{\mathbf{W}, \mathbf{b}} \sum_t \sum_{i=1}^{n_t} L(y_{i,t}, \mathbf{x}_{i,t}^T \mathbf{w}_t + \mathbf{b}_t) + \Omega_{\text{mtl}}(\mathbf{W}) \quad (10)$$

with $\{\mathbf{x}_{i,t}, y_{i,t}\}_{i \in \{1, \dots, n_t\}}$ being the training examples related to each task, $t \in 1, \dots, m$, $(\mathbf{w}_t, \mathbf{b}_t)$ being the classifier parameters for task t , and $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_m] \in \mathbb{R}^{d \times m}$ being a matrix concatenating all vectors $\{\mathbf{w}_t\}$. Note that the multitask learning framework applied to single EEG trial classification has already been investigated by van Gerven et al. [19]. The main contribution we bring compared to their works is the use of regularizer that explicitly induces all subject classifiers to be similar to an average one, in addition to a regularizer that enforces selected channels to be the same for all subjects. The intuition behind this point is that we believe that since the classification tasks we are dealing with are similar for all subjects and all related to the same BCI paradigm, selected channels and classifier parameters should not differ that much from subject to subject. We also think that inducing task parameters to be similar may be more important than enforcing selected channels to be similar when the number of training examples is small since it helps in reducing overfitting. For this purpose, we have proposed a novel regularization term of the form

$$\Omega_{\text{mtl}}(\mathbf{W}) = \lambda_r \sum_{g \in \mathcal{G}'} \|\mathbf{W}_g\|_2 + \lambda_s \sum_{t=1}^m \|\mathbf{w}_t - \widehat{\mathbf{w}}\|_2^2, \quad (11)$$

where $\widehat{\mathbf{w}} = (1/m) \sum_t \mathbf{w}_t$ is the average classifier across tasks and \mathcal{G}' contains nonoverlapping groups of components from matrix \mathbf{W} . The first term in (11) is a mixed-norm term that promotes group regularization. In this work, we defined groups in \mathcal{G}' based on the sensors, which means that all the features across subject related to a given sensor are in the same group g , leading to p groups of $r \times m$ feature, as depicted in Figure 1. The second term is a similarity promoting term as introduced in Evgeniou and Pontil [21]. It can be interpreted as a term enforcing the minimization of the classifier’s parameter variance. In other words, it promotes classifiers to be similar to the average one, and it helps improving performances when the number of learning examples for each task is limited, by reducing overfitting. Note that λ_r and λ_s , respectively, control the sparsity induced by the first term and the similarity induced by the second one. Hence, when setting $\lambda_s = 0$, the regularizer given in (11) boils down to be similar to the one used by van Gerven et al. [19]. Note that in practice λ_r and λ_s are selected by means of a nested cross-validation which aims at classification accuracy. Thus, it may occur that classifier similarity is preferred over sensor selection leading to robust classifiers which still use most of the sensors.

Similar to the single task optimization framework given in (2), the objective function for problem (10) can be expressed as a sum of gradient Lipschitz continuous term $f_1(\mathbf{W}) = \sum_{t,i} L(\cdot) + \lambda_s \sum_{t=1}^m \|\mathbf{w}_t - \widehat{\mathbf{w}}\|_2^2$ and a nondifferentiable term $f_2(\mathbf{W}) = \lambda_r \sum_{g \in \mathcal{G}'} \|\mathbf{W}_g\|_2$ having a closed-form proximal operator (see Appendix B). Hence, we have again considered a forward-backward splitting algorithm whose iterates are given in (8).

3. Numerical Experiments

We now present how these novel approaches perform on different BCI problems. Before delving into the details of the results, we introduce the simulated and real datasets.

3.1. Experimental Data. We have first evaluated the proposed approaches on a simple simulated P300 dataset generated as follows. A P300 wave is extracted using the grand average of a single subject data from the EPFL dataset described in the following. We generate 11000 simulated examples with 8 discriminative channels containing the P300 out of 16 channels for positive examples. A Gaussian noise of standard deviation 0.2 is added to all signals making the dataset more realistic. 1000 of these examples have been used for training.

The first real P300 dataset we used is the EPFL dataset, based on eight subjects performing P300 related tasks [33]. The subjects were asked to focus on one of the $3 \times 2 = 6$ images on the screen while one of the images is flashed at random. The EEG signals were acquired from 32 channels, sampled at 1024 Hz, and 4 recording sessions per subject have been realized. Signals are preprocessed exactly according to the steps described in [33]: a [1, 8] Hz bandpass Butterworth filter of order 3 is applied to all signals followed by a downsampling. Hence, for each trial (training example), we have 8 time-sample features per channel corresponding to a 1000 ms time-window after stimulus, which leads to 256 features for all

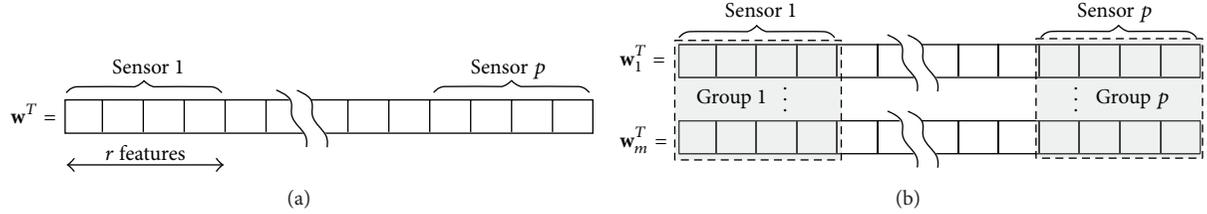


FIGURE 1: Examples of feature grouping for (a) single task and (b) multiple task learning.

channels ($32 \times 8 = 256$ features). Overall, the training set of a given subject is composed of about 3000 trials.

Another P300 dataset, recorded by the Neuroimaging Laboratory of Universidad Autónoma Metropolitana (UAM, Mexico) [34], has also been utilized. The data have been obtained from 30 subjects performing P300 spelling tasks on a 6×6 virtual keyboard. Signals are recorded over 10 channels leading thus to a very challenging dataset for sensor selection, as there are just few sensors left to select. For this dataset, we only use the first 3 sessions in order to have the same number of trials for all subjects (≈ 4000 samples). The EEG signals have been preprocessed according to the following steps: a $[2, 20]$ Hz Chebyshev bandpass filter of order 5 is first applied followed by a decimation, resulting in a poststimulus time-window of 31 samples per channels. Hence, each trial is composed of 310 (10×31) features.

We have also studied the effectiveness of our methods on an error-related potential (ErrP) dataset that has been recorded in the GIPSA Lab. The subjects were asked to memorize the position of 2 to 9 digits and to remind the position of one of these digits; operation has been repeated 72 times for each subject. The signal following the visualization of the result (correct/error on the memorized position) was recorded from 31 electrodes and sampled at 512 Hz. Similar to Jrad et al. [17], a $[1, 10]$ Hz Butterworth filter of order 4 and a downsampling has been applied to all channel signals. Finally, a time window of 1000 ms is considered as a trial (training example) with a dimensionality of $16 \times 31 = 496$.

3.2. Evaluation Criterion, Methods, and Experimental Protocol. We have compared several regularizers that induce feature/channel selection embedded in the learning algorithm, in a single subject learning setting as defined in (2). The performance measure commonly used in BCI competitions [3] is the area under the Roc curve (AUC). This measure is an estimate of the probability for a positive class to have a higher score than a negative class. It makes particularly sense to use AUC when evaluating a P300 speller as the letter in the keyboard is usually chosen by comparing score returned by the classifier for every column or line. In addition, AUC does not depend on the proportion of positive/negative examples in the data which makes it more robust than classification error rate. Our baseline algorithm is an SVM, which uses an ℓ_2 regularizer and thus does not perform any selection. Using an ℓ_1 regularizer yields a classifier which embeds feature selection, denoted as SVM-1 in the sequel. Three mixed-norm regularizers inducing sensor selection have also been considered: an $\ell_1 - \ell_2$ denoted as GSVM-2, and $\ell_1 - \ell_q$ referred

as GSVM- q , with q being selected in the set $\{1, 1.2, \dots, 1.8, 2\}$ by a nested cross-validation stage, and adaptive $\ell_1 - \ell_q$ norm, with $q = 2$ denoted as GSVM-a.

For the multitask learning setting, two MTL methods were compared to two baseline approaches which use all features, namely, a method that treats each tasks separately by learning one SVM per task (SVM) and a method denoted as SVM-Full, which on the contrary learns a unique SVM from all subject datasets. The two MTL methods are, respectively, a MTL as described in (10), denoted as MGSVM-2s and the same MTL but without similarity promoting regularization term, which actually means that we set $\lambda_s = 0$, indicated as MGSVM-2. For these approaches, performances are evaluated as the average AUC of the decision functions over all the subjects.

The experimental setup is described in the following. For each subject, the dataset is randomly split into a training set of $n = 1000$ trials and a test set containing the rest of the trials. The regularization parameter λ has been selected from a log-spaced grid ($[10^{-3}, 10^1]$) according to a nested 3-fold cross-validation step on the training set. When necessary, the selection of q is also included in this CV procedure. Finally, the selected value of λ is used to learn a classifier on the training examples and performances are evaluated on the independent test set. We run this procedure 10 times for every subject and report average performances. A Wilcoxon signed-rank test, which takes ties into account, is used to evaluate the statistical difference of the mean performances of all methods compared to the baseline SVM. We believe that such a test is more appropriate for comparing methods than merely looking at the standard deviation due to the high intersubject variability in BCI problems.

3.3. Results and Discussions. We now present the results we achieved on the above-described datasets.

3.3.1. Simulated Dataset. Average (over 10 runs) performance of the different regularizers on the simulated dataset is reported in Table 1 through AUC, sensor selection rate, and F -measure. This latter criterion measures the relevance of the selected channels compared to the true relevant ones. F -measure is formally defined as

$$F\text{-measure} = 2 \frac{|\mathcal{C} \cap \mathcal{C}^*|}{|\mathcal{C}^*| + |\mathcal{C}|}, \quad (12)$$

where \mathcal{C} and \mathcal{C}^* are, respectively, the set of selected channels and true relevant channels and $|\cdot|$ here denotes the cardinality

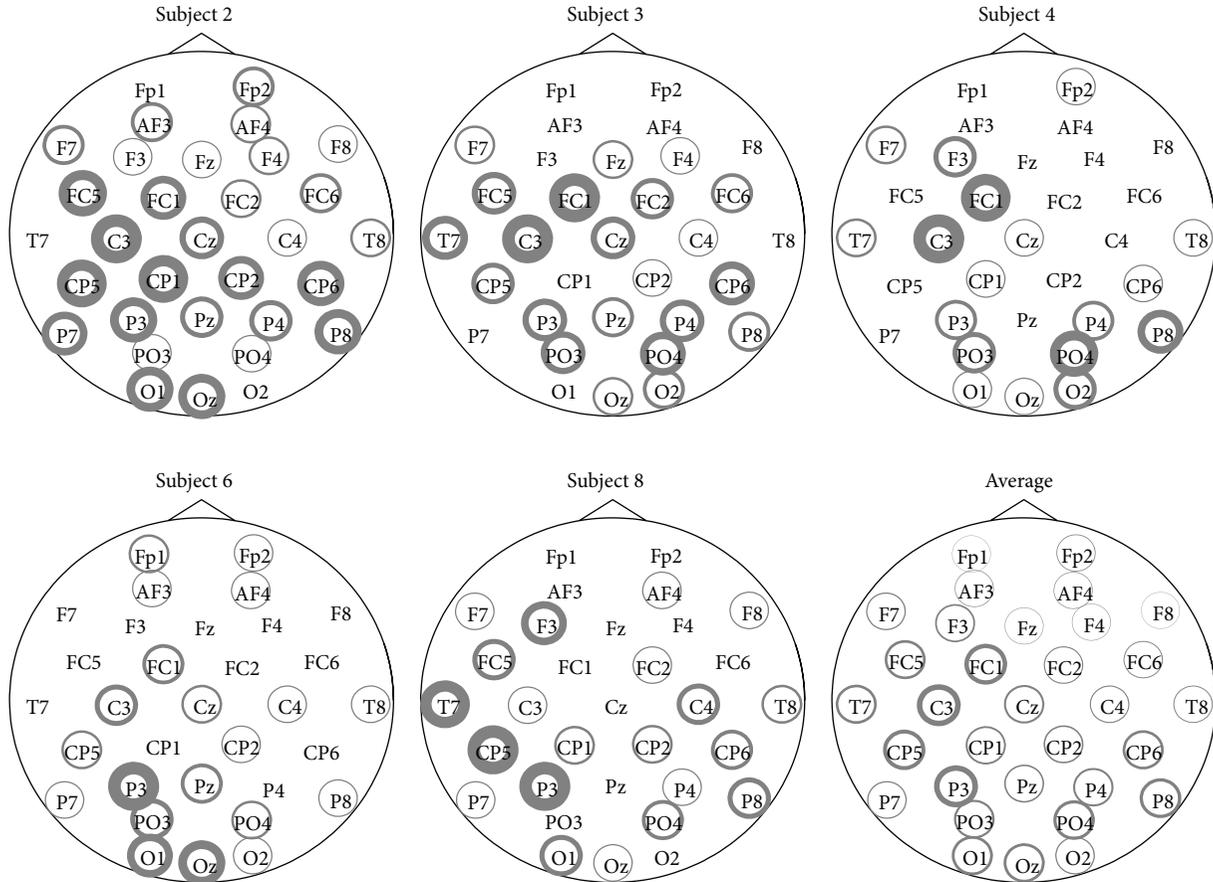


FIGURE 2: Selected sensors for the EPFL dataset. The line width of the circle is proportional to the number of times the sensor is selected for different splits. No circle means that the sensor has never been selected.

of a set. Note that if the selected channels are all the relevant ones, then the F -measure is equal to one. Most of the approaches provide similar AUC performances. We can although highlight that group-regularization approaches (GSVM-2, GSVM-p, GSVM-a) drastically reduce the number of selected channels since only 62% and 45% of the sensors are selected. A clear advantage goes to the adaptive regularization that is both sparser and is more capable of retrieving the true relevant channels.

3.3.2. P300 Datasets. Results for these datasets are reported in Table 2. For the EPFL dataset, all methods achieve performances that are not statistically different. However, we note that GSVM-2 leads to sensor selection (80% of sensor selected) while GSVM-a yields to classifiers that, on average, use 26% of the sensors at the cost of a slight loss in performances (1.5% AUC).

Results for the UAM dataset follow the same trend in terms of sensor selection but we also observe that the mixed-norm regularizers yield to increased performances. GSVM-2 performs statistically better than SVM although most of the sensors (9 out of 10) have been kept in the model. This shows that even if few channels have been removed, the group-regularization improves performances by bringing sensor prior knowledge to the problem. We also notice

that GSVM-a performance is statistically equivalent to the baseline SVM one while using only half of the sensors and GSVM-p consistently gives similar results to GSVM-2.

To summarize, concerning the performances of the different mixed-norm regularization, we outline that on one hand, GSVM-2 is at worst equivalent to the baseline SVM while achieving sensor selection and on the other hand GSVM-a yields to a substantial channel selection at the expense of a slight loss of performances.

A visualization of the electrodes selected by GSVM-a can be seen in Figure 2 for the EPFL dataset and in Figure 3 for the UAM dataset. Interestingly, we observe that for the EPFL dataset, the selected channels are highly dependent on the subject. The most recurring ones are the following: FC1 C3 T7 CP5 P3 PO3 PO4 Pz and the electrodes located above visual cortex O1, Oz, and O2. We see sensors from the occipital area that are known to be relevant [12] for P300 recognition, but sensors such as T7 and C3, from other brain regions, are also frequently selected. These results are however consistent with those presented in the recent literature [4, 18].

The UAM dataset uses only 10 electrodes that are already known to perform well in P300 recognition problem, but we can see from Figure 3 that the adaptive mixed-norm regularizer further selects some sensors that are essentially located in the occipital region. Note that despite the good

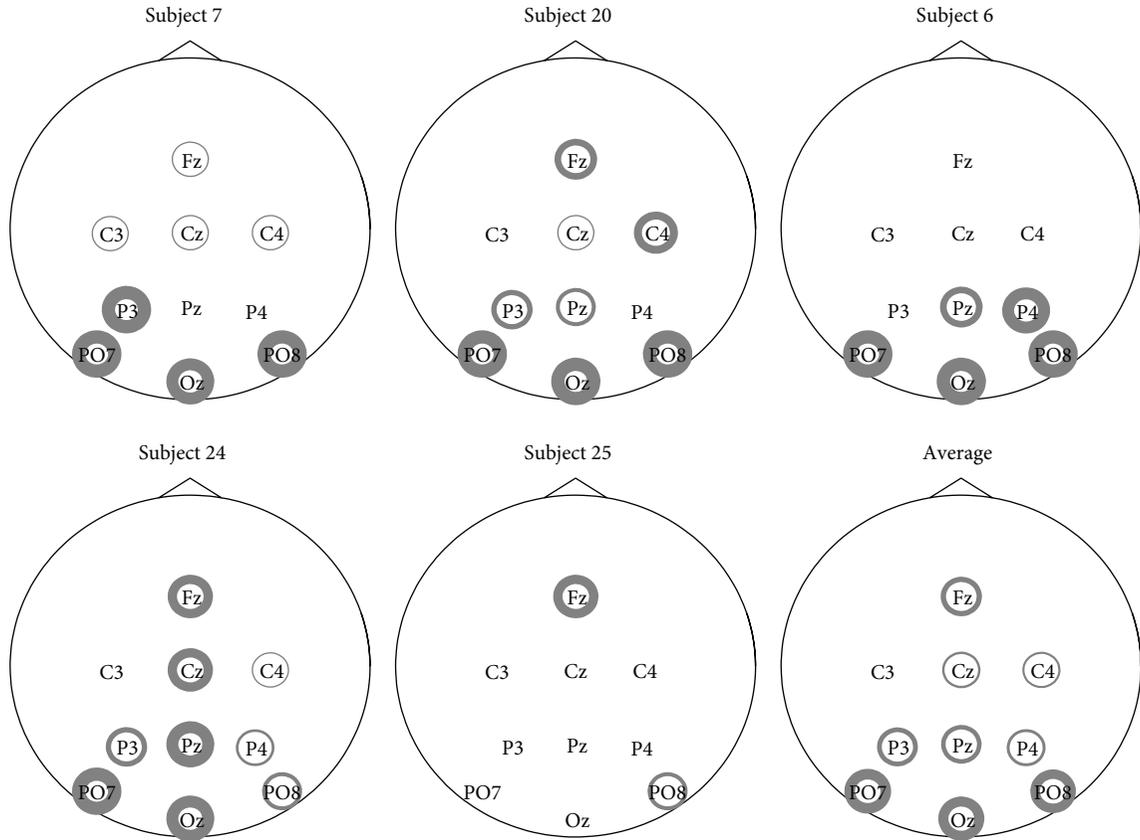


FIGURE 3: Selected sensors for the UAM dataset. The line width of the circle is proportional to the number of times the sensor is selected for different splits. No circle means that the sensor has never been selected.

TABLE 1: Performance results on the simulated datasets: the average performance in AUC (in %), the average percent of selected sensors (Sel), and the F -measure of the selected channels (in %).

Methods	Avg. AUC	AUC P -val	Avg. Sel	F -measure
SVM	79.79	—	100.00	66.67
GSVM-1	79.32	0.027	98.75	67.25
GSVM-2	80.96	0.004	62.50	89.72
GSVM-p	80.74	0.020	63.12	89.40
GSVM-a	80.51	0.014	45.62	93.98

Best results for each performance measure are in bold.

The P value refers to the one of a Wilcoxon signed-rank test with SVM as a baseline.

TABLE 2: Performance results for the 3 datasets: the average performance (over subjects) in AUC (in %), the average percent of selected sensors (Sel), and the P value of the Wilcoxon signed-rank test for the AUC when compared to the baseline SVM's one.

Methods	Datasets								
	EPFL dataset (8 Sub., 32 Ch.)			UAM dataset (30 Sub., 10 Ch.)			ErrP dataset (8 Sub., 32 Ch)		
	Avg. AUC	Avg. Sel	P value	Avg. AUC	Avg. Sel	P value	Avg. AUC	Avg. Sel	P value
SVM	80.35	100.00	—	84.47	100.00	—	76.96	100.00	—
SVM-1	79.88	87.66	0.15	84.45	96.27	0.5577	68.84	45.85	0.3125
GSVM-2	80.53	78.24	0.31	84.94	88.77	0.0001	77.29	29.84	0.5469
GSVM-p	80.38	77.81	0.74	84.94	90.80	0.0001	76.84	37.18	0.7422
GSVM-a	79.01	26.60	0.01	84.12	45.07	0.1109	67.25	7.14	0.1484

Best performing algorithms for each performance measure are in bold.

The P value refers to the one of a Wilcoxon signed-rank test with SVM as a baseline.

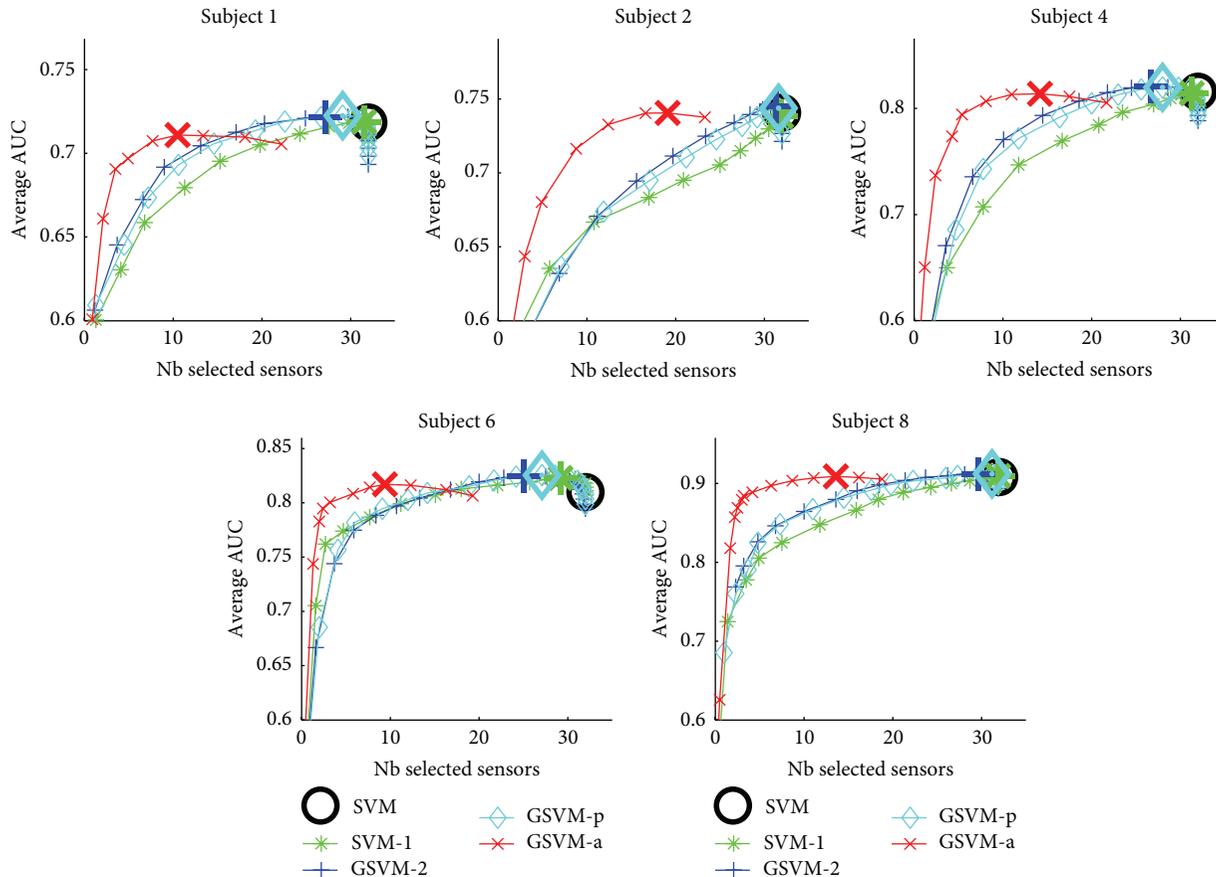


FIGURE 4: Performance versus sensor selection visualisation for the EPFL dataset. The large marker corresponds to the best model along the regularization path.

average performances reported in Table 2, some subjects in this dataset achieve very poor performances, of about 50% of AUC, regardless of the considered method. Selected channels for one of these subjects (Subject 25) are depicted in Figure 3 and, interestingly, they strongly differ from those of other subjects providing rationales for the poor AUC.

We have also investigated the impact of sparsity on the overall performance of the classifiers. To this aim, we have plotted the average performance of the different classifiers as a function of the number of selected sensors. These plots are depicted in Figure 4 for the EPFL dataset and on Figure 5 for the UAM dataset. For both datasets, GSVM-a frequently achieves a better AUC for a given level of sparsity. For most of the subjects, GSVM-a performs as well as SVM but using far less sensors. A rationale may be that in addition to selecting the relevant sensors, GSVM-a may provide a better estimation of the classifier parameters leading to better performances for a fixed number of sensors. As a summary, we suggest thus the use of an adaptive mixed-norm regularizer instead of an $\ell_1 - \ell_2$ mixed-norm as in van Gerven et al. [19] when sparsity and channel selection are of primary importance.

3.3.3. ErrP Dataset. The ErrP dataset differs from the others as its number of examples is small (72 examples per subject). The same experimental protocol as above has been used

for evaluating the methods but only 57 examples out of 72 have been retained for validation/training. Classification performances are reported on Table 2. For this dataset, the best performance is achieved by GSVM-2 but the Wilcoxon test shows that all methods are actually statistically equivalent. Interestingly, many channels of this dataset seem to be irrelevant for the classification task. Indeed, GSVM-2 selects only 30% of them while GSVM-a uses only 7% of the channels at the cost of 10% AUC loss. We believe that this loss is essentially caused by the aggressive regularization of GSVM-a and the difficulty to select the regularization parameter λ using only a subset of the 57 training examples. Channels selected by GSVM-2 can be visualized on Figure 6. Despite the high variance in terms of selected sensors, probably due to the small number of examples, sensors in the central area seem to be the most selected one, which is consistent with previous results in ErrP [35].

3.3.4. Multitask Learning. We now evaluate the impact of the approach we proposed in (10) and (11) on the P300 datasets. We expect that since multitask learning allows transferring some information between the different classification tasks, it will help in leveraging classification performances especially when the number of available training examples is small. Note that the ErrP dataset has not been tested in this MTL framework, because the above-described results suggest an

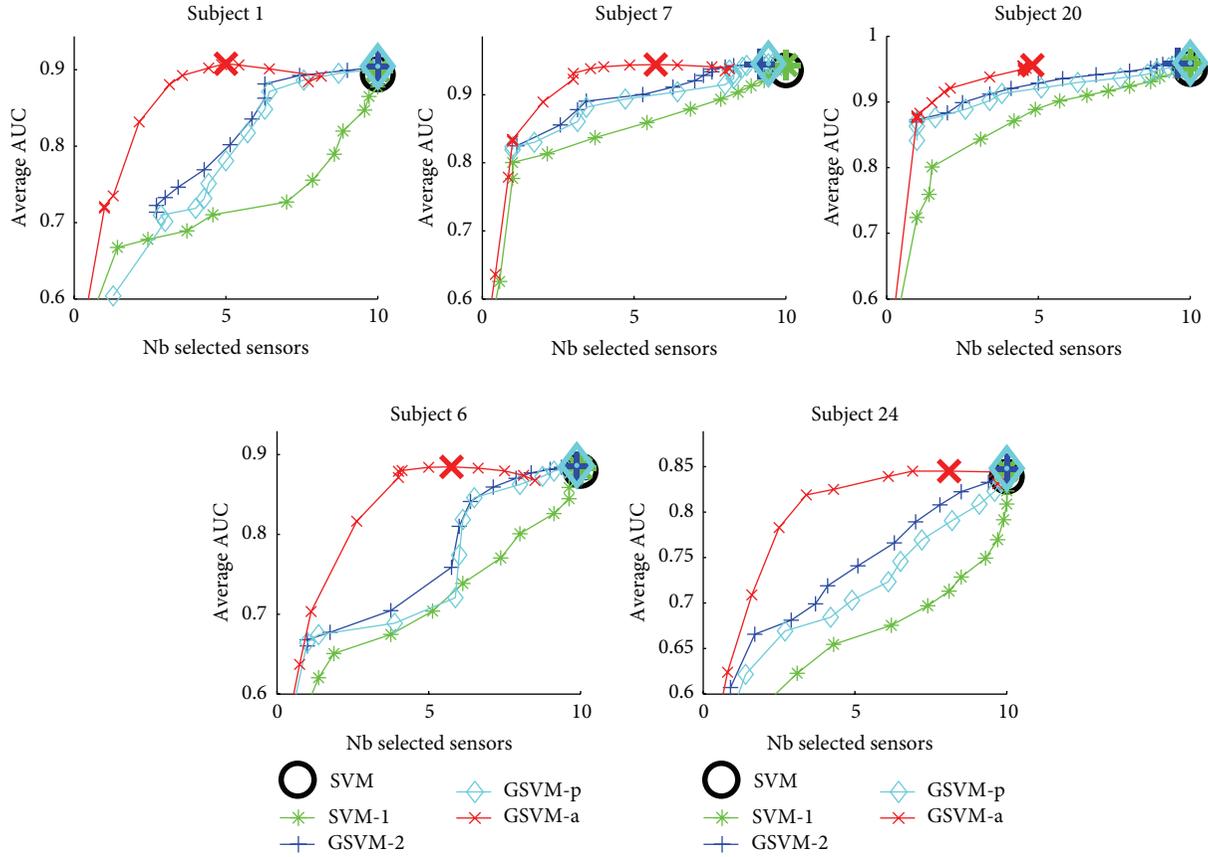


FIGURE 5: Performance versus sensor selection visualisation for the UAM dataset. The large marker corresponds to the best model along the regularization path.

important variance in the selected channels for all subjects. Hence, we believe that this learning problem does not fit into the prior knowledge considered through (11).

We have followed the same experimental protocol as for the single task learning except that training and test sets have been formed as follows. We first create training and test examples for a given subject by randomly splitting all examples of that subject and then gather all subject's training/test sets to form the multitask learning training/test sets. Hence, all the subjects are equally represented in these sets. A 3-fold nested cross-validation method is performed in order to automatically select the regularization terms (λ_r and λ_s).

Performances of the different methods have been evaluated for increasing number of training examples per subject and are reported in Figure 7. We can first see that for the EPFL dataset, MGSVM-2 and MGSVM-2s yield a slight but consistent improvement over the single-task classifiers (SVM-Full being a single classifier trained on all subject's examples and SVM being the average performances of subject-specific classifiers). The poor performances of the SVM-Full approach are probably due to the high intersubject variability in this dataset, which includes impaired patients.

For the UAM dataset, results are quite different since the SVM-Full and MGSVM-2s show a significant improvement over the single-task learning. We also note that when only

the joint channel selection regularizer is in play (MGSVM-2), multitask learning leads to poorer performance than the SVM-Full for a number of trials lower than 500. We justify this by the difficulty of achieving appropriate channel selection based only on few training examples, as confirmed by the performance of GSVM-2. From Figure 8, we can see that the good performance of MGSVM-2s is the outcome of performance improvement of about 10% AUC over SVM, achieved on some subjects that perform poorly. More importantly, while performances of these subjects are significantly increased, those that perform well still achieve good AUC scores. In addition, we emphasize that these improvements are essentially due to the similarity-inducing regularizer.

For both datasets, the MTL approach MGSVM-2s is consistently better than those of other single-task approaches thanks to the regularization parameters λ_r and λ_s that can adapt to the intersubject similarity (weak similarity for EPFL and strong similarity for UAM). These are interesting results showing that multitask learning can be a way to handle the problem related to some subjects that achieve poor performances. Moreover, results also indicate that multitask learning is useful for drastically shortening the calibration time. For instance, for the UAM dataset, 80% AUC was achieved using only 100 training examples (less than 1 minute of training example recordings). Note that the validation procedure tends to maximize performances and does not lead

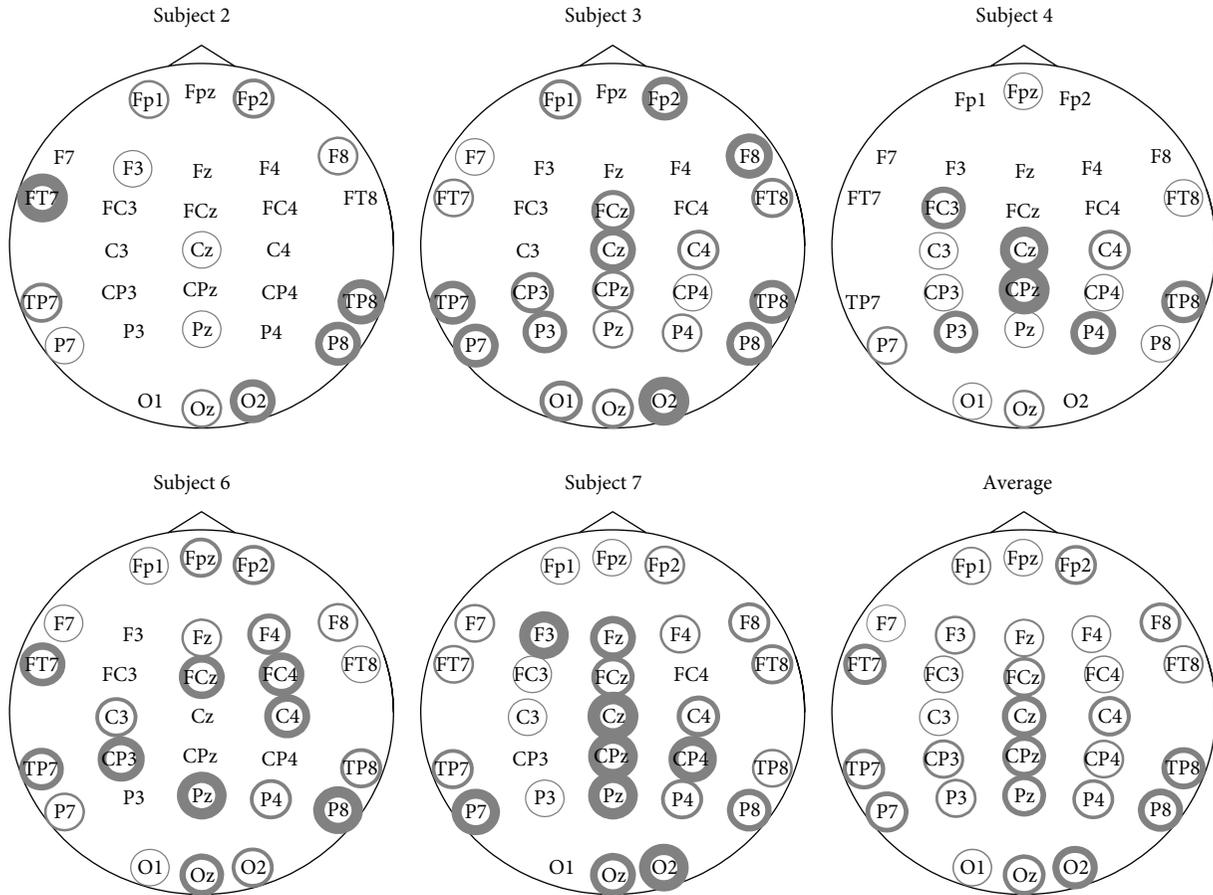


FIGURE 6: Selected sensors for the ERP dataset. The line width of the circle is proportional to the number of times the sensor is selected. No circle means that the sensor has never been selected.

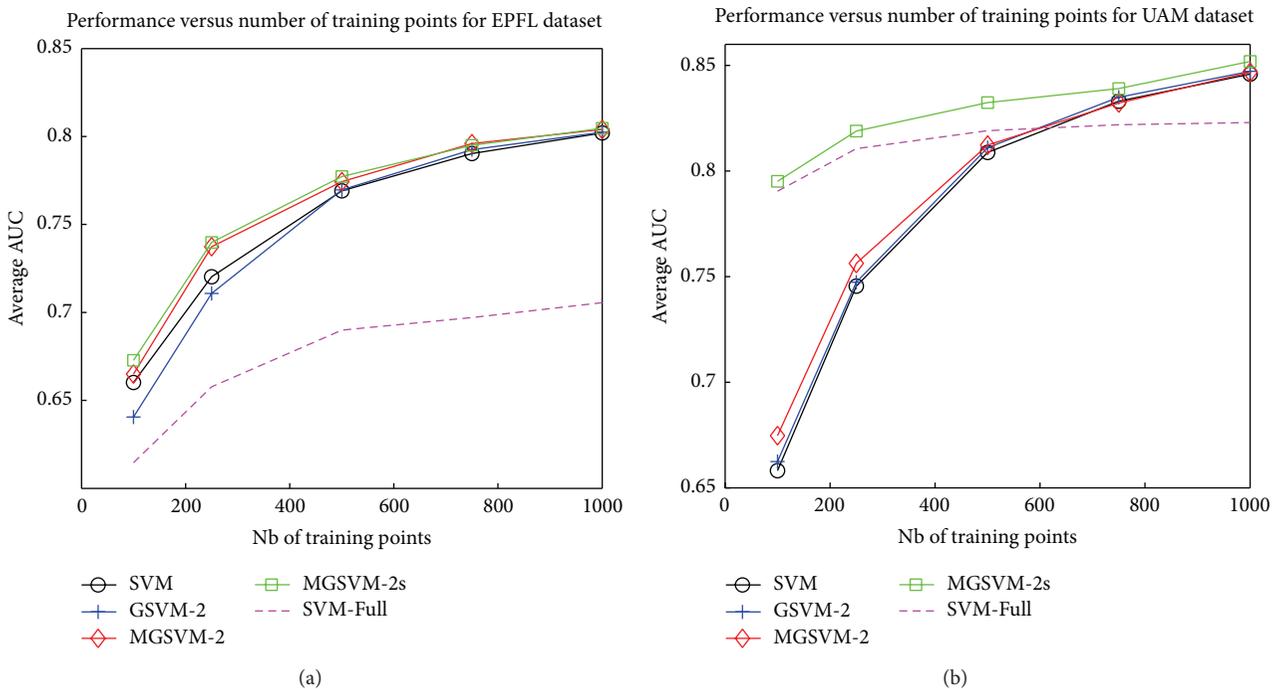


FIGURE 7: Multitask learning performances (AUC) for the EPFL (a) and UAM (b) datasets for different number of training examples per subject.

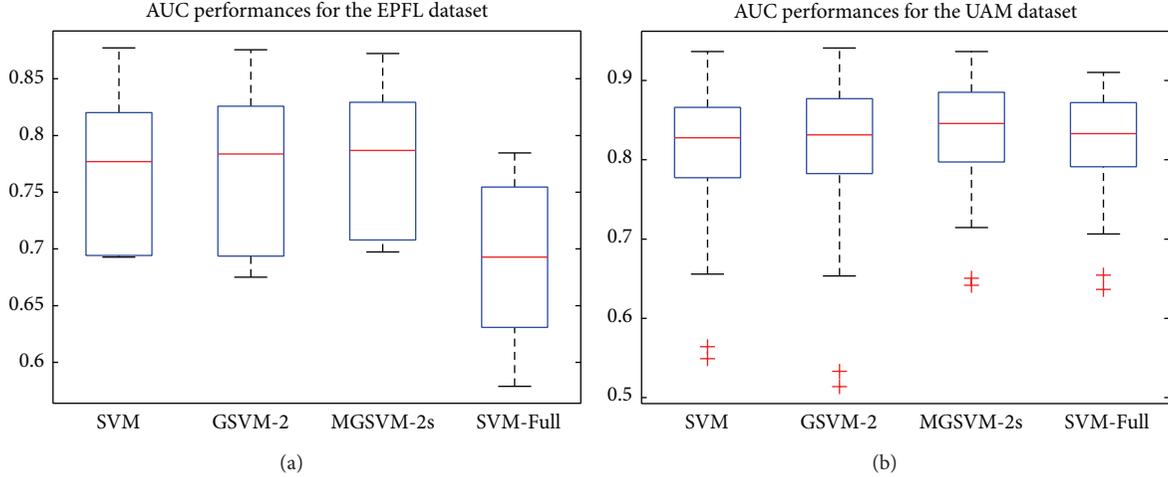


FIGURE 8: AUC performances comparison with EPFL (a) and UAM (b) for 500 training examples per subject.

to sparse classifiers for MTL approaches. As shown in Figures 2 and 3, the relevant sensors are quite different between subjects thus a joint sensor selection can lead to a slight loss of performances, hence the tendency of the cross-validation procedure to select nonsparse classifiers.

4. Conclusion

In this work, we have investigated the use of mixed-norm regularizers for discriminating event-related potentials in BCI. We have extended the discriminative framework of van Gerven et al. [19] by studying general mixed-norms and proposed the use of the adaptive mixed-norms as sparsity-inducing regularizers. This discriminative framework has been broadened to the multitask learning framework where classifiers related to different subjects are jointly trained. For this framework, we have introduced a novel regularizer that induces channel selection and classifier similarities. The different proposed approaches were tested on three different datasets involving a substantial number of subjects. Results from these experiments have highlighted that the $\ell_1 - \ell_2$ regularizer has been proven interesting for improving classification performance whereas adaptive mixed-norm is the regularizer to be considered when sensor selection is the primary objective. Regarding the multitask learning framework, our most interesting finding is that this learning framework allows, by learning more robust classifiers, significant performance improvement on some subjects that perform poorly in a single-task learning context.

In future work, we plan to investigate a different grouping of the features, such as temporal groups. This kind of group regularization could be for instance used in conjunction with the sensors group in order to promote both feature selection and temporal selection in the classifier. While the resulting problem is still convex, its resolution poses some issues so that a dedicated solver would be necessary.

Another research direction would be to investigate the use of asymmetrical MTL. This could prove handy when a poorly performing subject will negatively influence the

other subject performances in MTL while improving his own performances. In this case one would like subject classifier to be similar to the other's classifier without impacting their classifiers.

Appendices

A. Proof of Lipschitz Gradient of the Squared Hinge Loss

Given the training examples $\{\mathbf{x}_i, y_i\}$, the squared Hinge loss is written as

$$J = \sum_{i=1}^n \max(0, 1 - y_i \mathbf{x}_i^T \mathbf{w})^2 \quad (\text{A.1})$$

and its gradient is

$$\nabla_{\mathbf{w}} J = -2 \sum_i \mathbf{x}_i y_i \max(0, 1 - y_i \mathbf{x}_i^T \mathbf{w}). \quad (\text{A.2})$$

The squared Hinge loss is gradient Lipschitz if there exists a constant L such that

$$\|\nabla J(\mathbf{w}_1) - \nabla J(\mathbf{w}_2)\|_2 \leq L \|\mathbf{w}_1 - \mathbf{w}_2\|_2 \quad (\text{A.3})$$

$$\forall \mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^d.$$

The proof essentially relies on showing that $\mathbf{x}_i y_i \max(0, 1 - y_i \mathbf{x}_i^T \mathbf{w})$ is Lipschitz itself; that is, there exists $L' \in \mathbb{R}$ such that

$$\|\mathbf{x}_i y_i \max(0, 1 - y_i \mathbf{x}_i^T \mathbf{w}_1) - \mathbf{x}_i y_i \max(0, 1 - y_i \mathbf{x}_i^T \mathbf{w}_2)\| \leq L' \|\mathbf{w}_1 - \mathbf{w}_2\|. \quad (\text{A.4})$$

Now let us consider different situations. For a given \mathbf{w}_1 and \mathbf{w}_2 , if $1 - \mathbf{x}_i^T \mathbf{w}_1 \leq 0$ and $1 - \mathbf{x}_i^T \mathbf{w}_2 \leq 0$, then the left-hand side is equal to 0 and any L' would satisfy the inequality. If

$1 - \mathbf{x}_i^T \mathbf{w}_1 \leq 0$ and $1 - \mathbf{x}_i^T \mathbf{w}_2 \geq 0$, then the left-hand side (lhs) is

$$\begin{aligned} \text{lhs} &= \|\mathbf{x}_i\|_2 (1 - \mathbf{x}_i^T \mathbf{w}_2) \\ &\leq \|\mathbf{x}_i\|_2 (\mathbf{x}_i^T \mathbf{w}_1 - \mathbf{x}_i^T \mathbf{w}_2) \\ &\leq \|\mathbf{x}_i\|_2^2 \|\mathbf{w}_1 - \mathbf{w}_2\|_2. \end{aligned} \quad (\text{A.5})$$

A similar reasoning yields to the same bound when $1 - \mathbf{x}_i^T \mathbf{w}_1 \geq 0$ $1 - \mathbf{x}_i^T \mathbf{w}_2 \leq 0$ and $1 - \mathbf{x}_i^T \mathbf{w}_2 \geq 0$ and $1 - \mathbf{x}_i^T \mathbf{w}_1 \leq 0$. Thus, $\mathbf{x}_i y_i \max(0, 1 - y_i \mathbf{x}_i^T \mathbf{w})$ is Lipschitz with a constant $\|\mathbf{x}_i\|^2$. Now, we can conclude the proof by stating that $\nabla_{\mathbf{w}} J$ is Lipschitz as it is a sum of Lipschitz function and the related constant is $\sum_{i=1}^n \|\mathbf{x}_i\|_2^2$.

B. Lipschitz Gradient for the Multitask Learning Problem

For the multitask learning problem, we want to prove that the function

$$\begin{aligned} &\sum_{t=1}^m \sum_{i=1}^n L(y_{i,t}, \mathbf{x}_{i,t}^T \mathbf{w}_t + \mathbf{b}_t) \\ &+ \lambda_s \sum_{t=1}^m \left\| \mathbf{w}_t - \frac{1}{m} \sum_{j=1}^m \mathbf{w}_j \right\|_2^2 \end{aligned} \quad (\text{B.1})$$

is gradient Lipschitz, $L(\cdot, \cdot)$ being the square Hinge loss. From the above results, it is easy to show that the first term is gradient Lipschitz as the sum of gradient Lipschitz functions.

Now, we also show that the similarity term

$$\sum_t \left\| \mathbf{w}_t - \frac{1}{m} \sum_{j=1}^m \mathbf{w}_j \right\|_2^2 \quad (\text{B.2})$$

is also gradient Lipschitz.

This term can be expressed as

$$\begin{aligned} \left\| \mathbf{w}_t - \frac{1}{m} \sum_{j=1}^m \mathbf{w}_j \right\|_2^2 &= \sum_t \langle \mathbf{w}_t, \mathbf{w}_t \rangle - \frac{1}{m} \sum_{i,j=1}^m \langle \mathbf{w}_i, \mathbf{w}_j \rangle \\ &= \mathbf{w}^T \mathbf{M} \mathbf{w}, \end{aligned} \quad (\text{B.3})$$

where $\mathbf{w}^T = [\mathbf{w}_1^T, \dots, \mathbf{w}_m^T]$ is the vector of all classifier parameters and $\mathbf{M} \in \mathbb{R}^{md \times md}$ is the Hessian matrix of the similarity regularizer of the form

$$\mathbf{M} = \mathbf{I} - \frac{1}{m} \sum_{t=1}^m \mathbf{D}_t \quad (\text{B.4})$$

with \mathbf{I} the identity matrix and \mathbf{D}_t a block matrix with \mathbf{D}_t a $(t-1)$ -diagonal matrix where each block is an identity matrix \mathbf{I} with appropriate circular shift. \mathbf{D}_t is thus a $(t-1)$ row-shifted version of \mathbf{I} .

Once we have this formulation, we can use the fact that a function f is gradient Lipschitz of constant L if the largest

eigenvalue of its Hessian is bounded by L on its domain [36]. Hence, since we have

$$\|\mathbf{M}\|_2 \leq \|\mathbf{I}\|_2 + \frac{1}{m} \sum_{t=1}^m \|\mathbf{D}_t\|_2 = 2 \quad (\text{B.5})$$

the Hessian matrix of the similarity term $2 \cdot \mathbf{M}$ has consequently bounded eigenvalues. This concludes the proof that the function $\mathbf{w}^T \mathbf{M} \mathbf{w}$ is gradient Lipschitz continuous.

C. Proximal Operators

C.1. ℓ_1 Norm. The proximal operator of the ℓ_1 norm is defined as

$$\text{prox}_{\lambda \|\mathbf{x}\|_1}(\mathbf{u}) = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{x} - \mathbf{u}\|_2^2 + \lambda \|\mathbf{x}\|_1 \quad (\text{C.1})$$

and has the following closed-form solution for which each component is

$$[\text{prox}_{\lambda \|\mathbf{x}\|_1}(\mathbf{u})]_i = \text{sign}(u_i) (|u_i| - \lambda)_+ \quad (\text{C.2})$$

C.2. $\ell_1 - \ell_2$ Norm. The proximal operator of the $\ell_1 - \ell_2$ norm is defined as

$$\begin{aligned} &\text{prox}_{\lambda \sum_{g \in \mathcal{G}} \|\mathbf{x}_g\|_2}(\mathbf{u}) \\ &= \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{x} - \mathbf{u}\|_2^2 + \lambda \sum_{g \in \mathcal{G}} \|\mathbf{x}_g\|_2. \end{aligned} \quad (\text{C.3})$$

The minimization problem can be decomposed into several ones since the indices g are separable. Hence, we can just focus on the problem

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{x} - \mathbf{u}\|_2^2 + \lambda \|\mathbf{x}\|_2 \quad (\text{C.4})$$

whose minimizer is

$$\begin{aligned} &0 \quad \text{if } \|\mathbf{u}\|_2 \leq \lambda \\ &\left(1 - \frac{\lambda}{\|\mathbf{u}\|_2}\right) \mathbf{u} \quad \text{otherwise.} \end{aligned} \quad (\text{C.5})$$

Conflict of Interests

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

Acknowledgment

This work was partially supported by the French Agence Nationale de la Recherche (Grant ANR-11-JS02-10).

References

- [1] G. Dornhege, J. Millán, T. Hinterberger, D. McFarland, and K. Müller, *Toward Brain-Computer Interfacing*, vol. 74, Mit Press, Cambridge, Mass, USA, 2007.

- [2] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain*, Oxford University Press, New York, NY, USA, 2nd edition, 2006.
- [3] B. Blankertz, K.-R. Müller, D. J. Krusienski et al., “The BCI competition III: validating alternative approaches to actual BCI problems,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153–159, 2006.
- [4] B. Rivet, H. Cecotti, R. Phlypo, O. Bertrand, E. Maby, and J. Matout, “EEG sensor selection by sparse spatial filtering in P300 speller Brain-Computer Interface,” in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10)*, pp. 5379–5382, IEEE, September 2010.
- [5] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, “Single-trial analysis and classification of ERP components—a tutorial,” *NeuroImage*, vol. 56, no. 2, pp. 814–825, 2011.
- [6] J. Müller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, “Designing optimal spatial filters for single-trial EEG classification in a movement task,” *Clinical Neurophysiology*, vol. 110, no. 5, pp. 787–798, 1999.
- [7] C. Gouy-Pailler, M. Congedo, C. Brunner, C. Jutten, and G. Pfurtscheller, “Nonstationary brain source separation for multiclass motor imagery,” *IEEE Transactions on Bio-Medical Engineering*, vol. 57, no. 2, pp. 469–478, 2010.
- [8] G. Salimi-Khorshidi, A. M. Nasrabadi, and M. H. Golpayegani, “Fusion of classic P300 detection methods’ inferences in a framework of fuzzy labels,” *Artificial Intelligence in Medicine*, vol. 44, no. 3, pp. 247–259, 2008.
- [9] M. Falkenstein, J. Hohnsbein, J. Hoormann, and L. Blanke, “Effects of crossmodal divided attention on late ERP components—II. Error processing in choice reaction tasks,” *Electroencephalography and Clinical Neurophysiology*, vol. 78, no. 6, pp. 447–455, 1991.
- [10] P. Ferrez and J. Millán, *Error-Related Eeg Potentials in Brain-Computer Interfaces. Towards Brain-Computer Interfacing*, MIT Press, Cambridge, Mass, USA, 2007.
- [11] A. Buttfeld, P. W. Ferrez, and J. D. R. Millán, “Towards a robust BCI: error potentials and online learning,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 164–168, 2006.
- [12] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, “Toward enhanced P300 speller performance,” *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 15–21, 2008.
- [13] U. Hoffman, A. Yazdani, J. Vesin, and T. Ebrahimi, “Bayesian feature selection applied in a P300 brain—computer interface,” in *Proceedings of the 16th European Signal Processing Conference*, 2008.
- [14] T. N. Lal, M. Schröder, T. Hinterberger et al., “Support vector channel selection in BCI,” *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1003–1010, 2004.
- [15] J. Yang, H. Singh, E. L. Hines et al., “Channel selection and classification of electroencephalogram signals: an artificial neural network and genetic algorithm-based approach,” *Artificial Intelligence in Medicine*, vol. 55, no. 2, pp. 117–126, 2012.
- [16] H. Cecotti, B. Rivet, M. Congedo et al., “A robust sensor-selection method for P300 brain-computer interfaces,” *Journal of Neural Engineering*, vol. 8, no. 1, Article ID 016001, 2011.
- [17] N. Jrad, M. Congedo, R. Phlypo et al., “Sw-SVM: sensor weighting support vector machines for EEG-based brain-computer interfaces,” *Journal of Neural Engineering*, vol. 8, no. 5, Article ID 056004, 2011.
- [18] A. Rakotomamonjy and V. Guigue, “BCI competition III: dataset II-ensemble of SVMs for BCI P300 speller,” *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 3, pp. 1147–1154, 2008.
- [19] M. van Gerven, C. Hesse, O. Jensen, and T. Heskes, “Interpreting single trial data using groupwise regularisation,” *NeuroImage*, vol. 46, no. 3, pp. 665–676, 2009.
- [20] R. Tomioka and K.-R. Müller, “A regularized discriminative framework for EEG analysis with application to brain-computer interface,” *NeuroImage*, vol. 49, no. 1, pp. 415–432, 2010.
- [21] T. Evgeniou and M. Pontil, “Regularized multi-task learning,” in *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 109–117, August 2004.
- [22] A. Argyriou, T. Evgeniou, and M. Pontil, “Convex multi-task feature learning,” *Machine Learning*, vol. 73, no. 3, pp. 243–272, 2008.
- [23] A. Rakotomamonjy, R. Flamary, G. Gasso, and S. Canu, “ $\ell_{p/q}$ penalty for sparse linear and sparse multiple kernel multitask learning,” *IEEE Transactions on Neural Networks*, vol. 22, no. 8, pp. 1307–1320, 2011.
- [24] M. Alamgir, M. Grosse-Wentrup, and Y. Altun, “Multi-task learning for brain-computer interfaces,” in *AI & Statistics*, 2010.
- [25] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, “Convex optimization with sparsity-inducing norms,” in *Optimization for Machine Learning*, S. Sra, S. Nowozin, and S. J. Wright, Eds., pp. 19–53, MIT Press, 2011.
- [26] A. Rahimi, J. Xu, and L. Wang, “ L_p -norm regularization in volumetric imaging of cardiac current sources,” *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 276478, 10 pages, 2013.
- [27] A. Liu, T. Hao, Z. Gao, Y. Su, and Z. Yang, “Non-negative mixed-norm convex optimization for mitotic cell detection in phase contrast microscopy,” *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 176272, 10 pages, 2013.
- [28] F. R. Bach, “Consistency of the group lasso and multiple kernel learning,” *Journal of Machine Learning Research*, vol. 9, pp. 1179–1225, 2008.
- [29] H. Zou, “The adaptive lasso and its oracle properties,” *Journal of the American Statistical Association*, vol. 101, no. 476, pp. 1418–1429, 2006.
- [30] O. Chapelle, “Training a support vector machine in the primal,” *Neural Computation*, vol. 19, no. 5, pp. 1155–1178, 2007.
- [31] P. Combettes and J. Pesquet, “Proximal splitting methods in signal processing,” in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pp. 185–212, 2011.
- [32] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” *SIAM Journal on Imaging Sciences*, vol. 2, pp. 183–202, 2009.
- [33] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, “An efficient P300-based brain-computer interface for disabled subjects,” *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 115–125, 2008.
- [34] C. Ledesma-Ramirez, E. Bojorges Valdez, O. Yáñez Suarez, C. Saavedra, L. Bougrain, and G. G. Gentiletti, “An open-access P300 speller database,” in *Proceedings of the 4th International Brain-Computer Interface Meeting*, 2010.
- [35] S. Dehaene, M. Posner, and D. Tucker, “Localization of a neural system for error detection and compensation,” *Psychological Science*, vol. 5, no. 5, pp. 303–305, 1994.
- [36] D. Bertsekas, *Nonlinear Programming*, Athena Scientific, 1999.