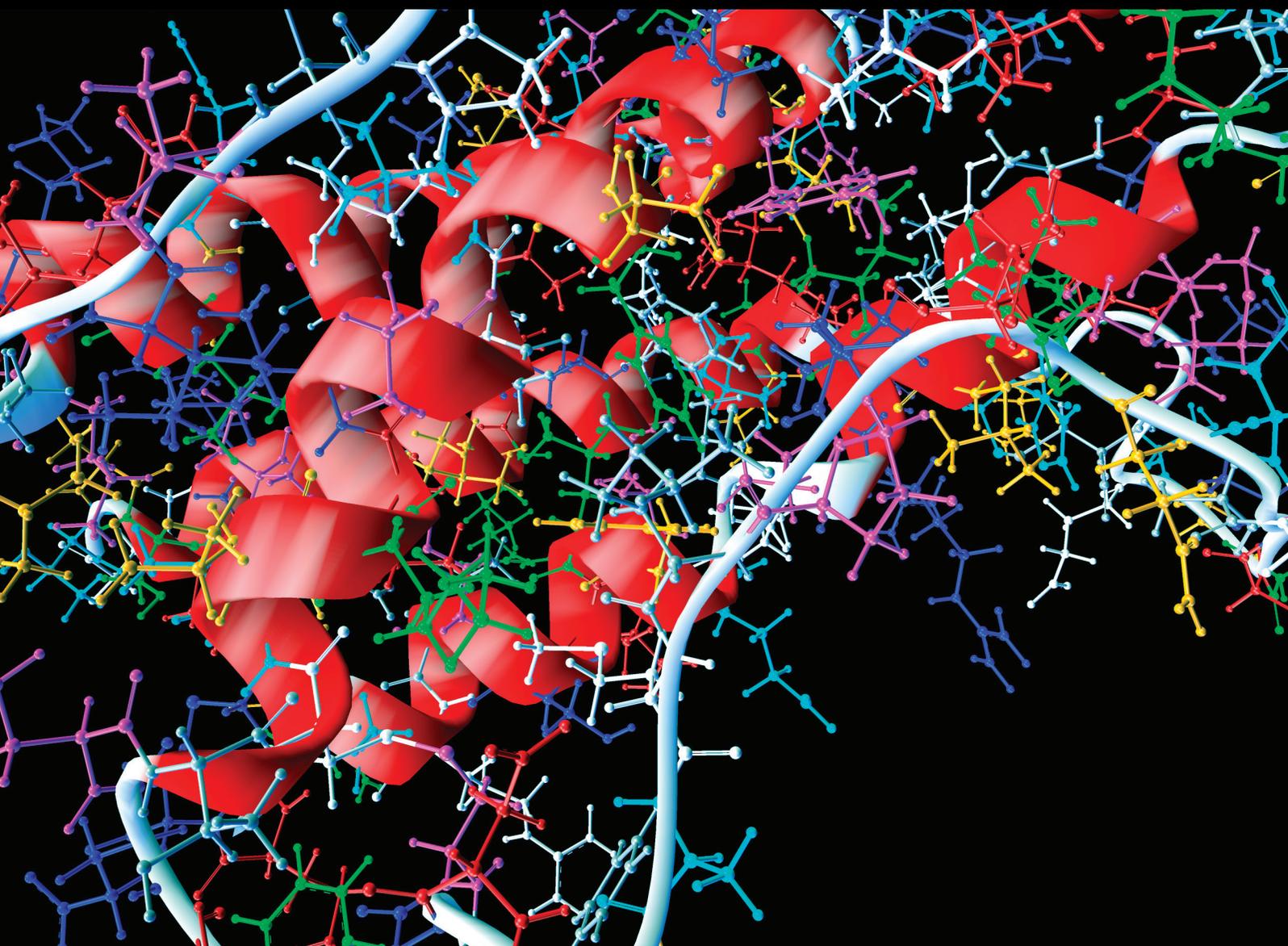


Computational and Mathematical Methods in Medicine

# Recent Advances in Statistical Data and Signal Analysis: Application to Real World Diagnostics from Medical and Biological Signals

Guest Editors: Dwarikanath Mahapatra, Krishna Agarwal, Reza Khosrowabadi, and Dilip K. Prasad





---

**Recent Advances in Statistical Data and Signal  
Analysis: Application to Real World Diagnostics  
from Medical and Biological Signals**

Computational and Mathematical Methods in Medicine

---

**Recent Advances in Statistical Data and Signal Analysis: Application to Real World Diagnostics from Medical and Biological Signals**

Guest Editors: Dwarikanath Mahapatra, Krishna Agarwal, Reza Khosrowabadi, and Dilip K. Prasad



Copyright © 2016 Hindawi Publishing Corporation. All rights reserved.

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

## Editorial Board

Emil Alexov, USA  
Elena Amato, Italy  
Konstantin G. Arbeev, USA  
Georgios Archontis, Cyprus  
Paolo Bagnaresi, Italy  
Enrique Berjano, Spain  
Elia Biganzoli, Italy  
Konstantin Blyuss, UK  
Hans A. Braun, Germany  
Thomas S. Buchanan, USA  
Zoran Bursac, USA  
Thierry Busso, France  
Xueyuan Cao, USA  
Carlos Castillo-Chavez, USA  
Prem Chapagain, USA  
Hsiu-Hsi Chen, Taiwan  
Ming-Huei Chen, USA  
Wai-Ki Ching, Hong Kong  
Nadia A. Chuzhanova, UK  
Maria N. D.S. Cordeiro, Portugal  
Irena Cosic, Australia  
Fabien Crauste, France  
Getachew Dagne, USA  
Qi Dai, China  
Chuangyin Dang, Hong Kong  
Justin Dauwels, Singapore  
Didier Delignières, France  
Jun Deng, USA  
Thomas Desaive, Belgium  
David Diller, USA  
Michel Dojat, France  
Irina Doytchinova, Bulgaria  
Esmaeil Ebrahimie, Australia  
Georges El Fakhri, USA  
Issam El Naqa, USA  
Angelo Facchiano, Italy  
Luca Faes, Italy  
Giancarlo Ferrigno, Italy  
Marc Thilo Figge, Germany  
Alfonso T. García-Sosa, Estonia  
Humberto González-Díaz, Spain  
Igor I. Goryanin, Japan  
Marko Gosak, Slovenia  
Damien Hall, Australia  
Volkhard Helms, Germany  
Roberto Hornero, Spain  
Tingjun Hou, China  
Seiya Imoto, Japan  
Sebastien Incerti, France  
Abdul Salam Jarrah, UAE  
Hsueh-Fen Juan, Taiwan  
Rafik Karaman, Palestine  
Lev Klebanov, Czech Republic  
Andrzej Kloczkowski, USA  
Xiang-Yin Kong, China  
Zuofeng Li, USA  
Chung-Min Liao, Taiwan  
Quan Long, UK  
Ezequiel López-Rubio, Spain  
Reinoud Maex, France  
Valeri Makarov, Spain  
Kostas Marias, Greece  
Richard J. Maude, Thailand  
Panagiotis Mavroidis, USA  
Georgia Melagraki, Greece  
Michele Migliore, Italy  
John Mitchell, UK  
Chee M. Ng, USA  
Michele Nichelatti, Italy  
Ernst Niebur, USA  
Kazuhisa Nishizawa, Japan  
Hugo Palmans, UK  
Francesco Pappalardo, Italy  
Matjaz Perc, Slovenia  
Jesús Picó, Spain  
Alberto Policriti, Italy  
Giuseppe Pontrelli, Italy  
Christopher Pretty, New Zealand  
Mihai V. Putz, Romania  
Ravi Radhakrishnan, USA  
Jose Joaquin Rieta, Spain  
Jan Rychtar, USA  
Moisés Santillán, Mexico  
Vinod Scaria, India  
Jörg Schaber, Germany  
Xu Shen, China  
Simon A. Sherman, USA  
Pengcheng Shi, USA  
Tieliu Shi, China  
Erik A. Siegbahn, Sweden  
Sivabal Sivaloganathan, Canada  
Dong Song, USA  
Xinyuan Song, Hong Kong  
Emiliano Spezi, UK  
Greg M. Thurber, USA  
Tianhai Tian, Australia  
Tianhai Tian, Australia  
Nestor V. Torres, Spain  
Nelson J. Trujillo-Barreto, UK  
Anna Tsantili-Kakoulidou, Greece  
Po-Hsiang Tsui, Taiwan  
Gabriel Turinici, France  
Edelmira Valero, Spain  
Raoul van Loon, UK  
Luigi Vitagliano, Italy  
Liangjiang Wang, USA  
Ruiqi Wang, China  
Ruisheng Wang, USA  
David A. Winkler, Australia  
Gabriel Wittum, Germany  
Yu Xue, China  
Yongqing Yang, China  
Chen Yanover, Israel  
Xiaojun Yao, China  
Kaan Yetilmezsoy, Turkey  
Hujun Yin, UK  
Hiro Yoshida, USA  
Henggui Zhang, UK  
Yuhai Zhao, China  
Xiaoqi Zheng, China  
Yunping Zhu, China

# Contents

---

**Recent Advances in Statistical Data and Signal Analysis: Application to Real World Diagnostics from Medical and Biological Signals**

Dwarikanath Mahapatra, Krishna Agarwal, Reza Khosrowabadi, and Dilip K. Prasad

Volume 2016, Article ID 1643687, 1 page

**Perfusion Angiography in Acute Ischemic Stroke**

Fabien Scalzo and David S. Liebeskind

Volume 2016, Article ID 2478324, 14 pages

**Detecting the Intention to Move Upper Limbs from Electroencephalographic Brain Signals**

Berenice Gudiño-Mendoza, Gildardo Sanchez-Ante, and Javier M. Antelis

Volume 2016, Article ID 3195373, 11 pages

**Frequency and Time Domain Analysis of Foetal Heart Rate Variability with Traditional Indexes: A Critical Survey**

Maria Romano, Luigi Iuppariello, Alfonso Maria Ponsiglione, Giovanni Improta, Paolo Bifulco, and Mario Cesarelli

Volume 2016, Article ID 9585431, 12 pages

**Centered Kernel Alignment Enhancing Neural Network Pretraining for MRI-Based Dementia Diagnosis**

David Cárdenas-Peña, Diego Collazos-Huertas, and German Castellanos-Dominguez

Volume 2016, Article ID 9523849, 10 pages

**Establishment of Relational Model of Congenital Heart Disease Markers and GO Functional Analysis of the Association between Its Serum Markers and Susceptibility Genes**

Min Liu, Luosha Zhao, and Jiaying Yuan

Volume 2016, Article ID 9506829, 14 pages

**Diagnosing Parkinson's Diseases Using Fuzzy Neural System**

Rahib H. Abiyev and Sanan Abizade

Volume 2016, Article ID 1267919, 9 pages

## *Editorial*

# **Recent Advances in Statistical Data and Signal Analysis: Application to Real World Diagnostics from Medical and Biological Signals**

**Dwarikanath Mahapatra,<sup>1</sup> Krishna Agarwal,<sup>2</sup> Reza Khosrowabadi,<sup>3</sup> and Dilip K. Prasad<sup>4</sup>**

<sup>1</sup>*IBM Research, Melbourne, VIC, Australia*

<sup>2</sup>*Singapore-MIT Alliance for Research and Technology, Singapore*

<sup>3</sup>*Shahid Beheshti University, Tehran, Iran*

<sup>4</sup>*Nanyang Technological University, Singapore*

Correspondence should be addressed to Dwarikanath Mahapatra; [dwarim@au.ibm.com](mailto:dwarim@au.ibm.com)

Received 19 June 2016; Accepted 19 June 2016

Copyright © 2016 Dwarikanath Mahapatra 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.

Medical and biological signals span almost the entire spectrum from EEG to X-rays and their sources range from molecular scales to large organs such as heart, brain, and muscles. Signal processing techniques (including image analysis) are constantly serving towards improving the state-of-the-art in medical and biological data analysis and interpretation. There is constant scientific endeavour to get better insight into the hidden information beneath the huge stack of medical data that we encounter. Consequently there has been a major shift towards quantitative analysis of medical data through various computational approaches. Computational approaches that have been hugely popular and found important applications include computational modeling, Bayesian and graphical models, machine learning, deep-learning, pattern recognition, optimization, spectral and pseudospectral analysis, stochastic modelling, iterative system model adaptation, and multiscale multiphysics analysis to name a few.

This special issue was an attempt to bring together interesting works that use advanced statistical techniques for cutting edge medical applications and biological signals for disease detection and diagnosis. We received submissions from a wide range of approaches and applications such as biological/medical image and signal processing, sensor and probe's signal analysis, imaging and microscopy techniques, human brain mapping, modeling and simulation of biological, biochemical, cellular, and subcellular processes,

sensor fusion, wearable devices based health informatics, histopathology image analysis, and brain computer interface in medicine.

We have selected papers which were particularly relevant to bringing forward methods and applications that we think will be interesting for a wide range of researchers. Out of 14 submissions we selected 6 manuscripts. We hope that this selective publication process would prove beneficial for those researchers who wish to advance the state-of-the-art in medical imaging research.

*Dwarikanath Mahapatra  
Krishna Agarwal  
Reza Khosrowabadi  
Dilip K. Prasad*

## Research Article

# Perfusion Angiography in Acute Ischemic Stroke

**Fabien Scalzo and David S. Liebeskind**

*Department of Neurology, University of California, Los Angeles, Los Angeles, CA 90024, USA*

Correspondence should be addressed to Fabien Scalzo; [fabien.scalzo@gmail.com](mailto:fabien.scalzo@gmail.com)

Received 7 December 2015; Revised 26 May 2016; Accepted 7 June 2016

Academic Editor: Dwarikanath Mahapatra

Copyright © 2016 F. Scalzo and D. S. Liebeskind. 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.

Visualization and quantification of blood flow are essential for the diagnosis and treatment evaluation of cerebrovascular diseases. For rapid imaging of the cerebrovasculature, digital subtraction angiography (DSA) remains the gold standard as it offers high spatial resolution. This paper lays out a methodological framework, named perfusion angiography, for the quantitative analysis and visualization of blood flow parameters from DSA images. The parameters, including cerebral blood flow (CBF) and cerebral blood volume (CBV), mean transit time (MTT), time-to-peak (TTP), and  $T_{\max}$ , are computed using a bolus tracking method based on the deconvolution of the time-density curve on a pixel-by-pixel basis. The method is tested on 66 acute ischemic stroke patients treated with thrombectomy and/or tissue plasminogen activator (tPA) and also evaluated on an estimation task with known ground truth. This novel imaging tool provides unique insights into flow mechanisms that cannot be observed directly in DSA sequences and might be used to evaluate the impact of endovascular interventions more precisely.

## 1. Introduction

Visualization of blood flow inside brain vessels is essential for the diagnosis and treatment evaluation of cerebrovascular disorders. First attempts date back to the early years of angiography [1], a technique that relies on X-ray imaging of iodinated radioopaque contrast agent previously injected into the blood stream. Over the years, the technique has improved and benefited from the appearance of digital cameras, leading to digital subtraction angiography (DSA) [2–4] which allows for the unwanted elements (e.g., skull) to be removed by image subtraction. Today, DSA remains a central and widely used imaging technique to assess blood flow during neurovascular interventions of stroke, for instance. In practice, several limitations hinder the use of DSA; images are qualitative; they are displayed in grayscale and need to be browsed frame by frame to observe temporal differences. This paper addresses these limitations by presenting a framework, perfusion angiography, for the quantitative analysis and visualization of perfusion and delay parameters from DSA.

The popularity of DSA can be attributed to its good spatiotemporal resolution which is not easily matched by other acquisition techniques such as magnetic resonance imaging (MRI) and computed tomography (CT). Vascular

abnormalities such as narrowing, blockage, or malformations can be visualized precisely in DSA. In addition, DSA is minimally invasive and is readily available in interventional suites of modern intensive care units (ICUs). Minimal cost, low risks, and rapid acquisition time are other features in favor of DSA. Although it may be argued that DSA will gradually be replaced in the future by CT angiography (CTA) during neurovascular interventions, DSA remains the gold standard worldwide.

Over the last three decades, numerous works have studied the role of DSA in both diagnosis and treatment evaluation of cardio- and cerebrovascular diseases. However, most of the existing studies were based on the visual review of image sequences by neurologists such that observations were collapsed to a simplified scale describing degree of reperfusion (Thrombolysis in Cerebral Infarction (TICI)) and recanalization (Arterial Occlusive Lesion (AOL)) after intervention. These dichotomizations are still a matter of ongoing debate [5] in the stroke community as their correlation with general outcome is limited and may also present interreader variability. There is a clear need to go beyond these manual scoring systems to obtain better evaluations for future clinical trials and endovascular devices. Although automatic TICI and AOL scores are still beyond the capabilities of current

methods, automated algorithms for quantitative blood flow estimation have been developed over the last 20 years (as reviewed in [6]). They have failed so far to be translated into meaningful tools that could improve clinical practice and treatment evaluation.

In addition to its high spatial resolution, DSA holds temporal information that can be used to track the contrast agent and compute parametric perfusion parameters, such as cerebral blood volume (CBV), cerebral blood flow (CBF), mean transit time (MTT), time-to-peak (TTP), and  $T_{\max}$ , thus providing a quantitative assessment of cerebral hemodynamics. Those parameters are very useful for assessment of cerebrovascular diseases as they can render the underlying functional information more easily recognizable. In contrast with MR and CT perfusion imaging that have been studied through clinical trials (e.g., DEFUSE [7]), the value of parametric imaging from DSA has been largely underappreciated. While the idea of extracting perfusion parameters from DSA has been introduced since the 1980s in sporadic studies for CBF [8], MTT [9], TTP [10], and  $T_{\max}$  [10, 11], it has not yet received the attention deserved by the acute stroke community. This overall lack in interest may be caused by the difficulty of real-time implementation of those algorithms on angiographic units, challenges due to the nature of the images (artifacts, vessel overlap), or perhaps failure to realize its full potential for quantitative decision support. Drawing from these observations, the overall goal of this paper is to describe how these important perfusion parameters can be extracted from the time-density curve and displayed in color-coded images that can be readily interpreted by neurologists and neurointerventionalists. After a brief historical review in Section 2.1 and a description of the dataset in Section 2.2, the paper summarizes the theory of densitometry in Section 2.3 and introduces the proposed framework in Section 2.4. The results of the experiments are presented in Section 3 and discussed in Section 4.

## 2. Methods

*2.1. Historical and Technical Overview.* Since its discovery [12], the application of X-ray for imaging purposes has greatly influenced medical diagnosis and interventions as it allows visualizing moving anatomical structures and endovascular devices. X-rays are produced by accelerating electrons emitted from a cathode towards a metal target anode using high voltage (50 kV). When directed towards the body and by passing through it, X-rays are partially absorbed and deflected, which causes attenuation of the incident beam. Various anatomical structures can be differentiated thanks to their specific level of absorption.

One of the decisive milestones of X-ray imaging was the introduction of angiography [13] which has made the visualization of blood flow within vessels possible. The acquisition of an angiogram relies on X-ray imaging of iodinated radioopaque contrast agent previously injected into the blood stream. The blood flow is observed thanks to the high level of absorption of the contrast agent.

The introduction of the image intensifier television (II-TV) that converts the incident X-rays into a visible image

was an ingredient to the success of angiography. With the modernization of computers in the 1980s, it became possible to record images digitally. This led to digital radiography [2] which allows for more flexible visualization of digitally enhanced images. Digital subtraction angiography (DSA) extends digital radiography [4, 14–17] by subtracting a background image (obtained before injection of the contrast agent) from subsequent images. The purpose is to eliminate the bone and soft tissue images that would otherwise be superposed on the vessels.

Despite excellent resolution characteristics, DSA has several inherent shortcomings. First, DSA images are subjected to two major types of noise: the quantum noise due to the random nature of X-ray distribution and the noise resulting from the electronic components. In addition, the image subtraction operation amplifies the noise already present in images. To overcome this problem, noise reduction techniques can be applied. This leads to a second weakness of DSA: noise reduction algorithms are generally coded in the hardware and algorithms cannot easily be accessed or customized. Another limitation is the possible motion of the patient during the image acquisition that creates spatial blur and artifacts as the background image is not aligned to subsequent frames. Finally, visualization of the temporal information from DSA sequences is challenged by the fact that images are typically displayed in a video mode with raw grayscale frames. Only recently have commercial systems started to introduce colormaps to better visualize the temporal information held in DSA.

In summary, DSA is technological evolution of digital radiography to remove unwanted bone and soft tissue from a set of successive images. Besides the technical limitations due to the nature of X-ray imaging, DSA is associated with a computational layer that may also introduce significant inaccuracies in the presence of even minor patient motion. Because internal parameters and source images used by DSA are generally not made available by scanner manufacturers, further postprocessing is particularly challenging. Despite these limiting factors, DSA remains the gold standard used during endovascular interventions. In this study, we proposed to extend DSA by introducing a computational framework for the computation of perfusion parameters.

*2.2. Patients Demographics and Data Acquisition.* The imaging dataset used in this study to evaluate our framework was collected from patients evaluated at a single, academic comprehensive stroke center and identified with symptoms of acute ischemic stroke. The use of this dataset was approved by the local Institutional Review Board (IRB). Inclusion criteria for this study included (1) presenting symptoms suggestive of acute stroke, (2) last known well time within six hours at admission, (3) digital subtraction angiography (DSA) of the brain performed at the end of a thrombectomy procedure, and (4) final diagnosis of ischemic stroke. A total of 66 patients (median age: 68 years (IQR 53, 79)), including 35 women, satisfied the above criteria. All patients underwent thrombectomy with various success in revascularization which was determined using the Thrombolysis in Cerebral Infarction (TICI) score. The distribution of TICI scores is

as follows: TICI 0 (4 patients), TICI 1 (1 patient), TICI 2a (17 patients), TICI 2b (35 patients), and TICI 3 (9 patients). Mechanical clot-retrieval devices include Trevo<sup>®</sup> (7 patients), MERCI<sup>®</sup> (17 patients), and Solitaire<sup>®</sup> (32 patients). The median NIH stroke scale (NIHSS) is 18, IQR 13, 21. The DSA scanning was performed on a Philips Allura Xper FD20<sup>®</sup> biplane using a routine timed contrast-bolus passage technique. Manual injection of Omnipaque 300 was performed at dilution of 70% (30% saline) such that 10 cc of contrast was administered intravenously at an approximate rate of 5 cm<sup>3</sup>/s. Image acquisition parameters vary across subjects. In the biplane acquisition setting, frames are acquired in an interleaved fashion at two standard viewpoints: anterior-posterior (AP) and lateral. The median number of frames acquired is 20 frames, IQR 17, 22, and the median peak voltage output is 95 kV, IQR 86, 104. Images sizes were all 1024 × 1024 but were acquired with different field of view.

**2.3. Video Densitometric Theory.** To derive perfusion parameters from DSA sequences by bolus tracking analysis, the concentration  $C$  of the contrast agent at any location must be known. It can be estimated through DSA as the intensity observed in the image is proportional to the contrast concentration [8, 18]:

$$I(t)_{(x,y)} = k\mu C(t)_{(x,y)} \rho_{(x,y)}, \quad (1)$$

where  $I(t)_{(x,y)}$  is the DSA image intensity value for a given pixel  $(x, y)$  at time  $t$ ,  $\mu$  is the mass attenuation coefficient of the contrast agent which is proportional to the X-ray energy,  $\rho_{(x,y)}$  is the thickness of the vessel,  $C_{(x,y)}$  is the contrast concentration, and  $k$  is a constant that accounts for the X-ray imaging system acquisition and amplification [19].

The vessel thickness  $\rho_{(x,y)}$  can be computed using one of the previously described frameworks (e.g., [20]) that first applies a vessel detector based on vesselness filtering and thresholding. Centerlines are then obtained via skeletonization. Finally, a perpendicular segment (computed along each point of the centerline) is used to measure the distance to the edges of the vessel and derive the thickness assuming cylindrical volume. The thickness is then applied on a cross-sectional basis to every point within the vessel using bicubic interpolation.

Animal studies [21, 22] of coronary circulation from DSA have demonstrated accurate estimation of the flow within blood vessels. Other studies [23] estimated flow related parameters from contrast time curves within the pulmonary parenchyma. In those cases, the vessel diameter within the parenchyma was too small to be measured on the image and had to be set to a constant value  $\rho_{(x,y)} = k_p$ .

We transpose these estimation methods of the concentration-time curve within blood vessels and the brain parenchyma to DSA imaging routinely acquired during endovascular treatment of acute ischemic strokes. The goal is to extract cerebral hemodynamic parameters to quantify degree of perfusion and delay, as described in the next section.

**2.4. Perfusion Parameters from DSA Using Bolus Tracking.** Bolus tracking algorithms [24–27] are well established methods to determine flow and timing parameters of a bolus

travelling from a source to a target location. This section describes the extraction process of hemodynamic indices that will provide a quantitative description of the tissue status from DSA.

From the contrast concentrations (see (1)), it is possible to estimate the CBV at any location  $u$  in the image by calculating the amount of contrast agent  $C_u$  that has passed through it with respect to the total amount of contrast measured at the feeding arterial vessel  $C_a$  (i.e., arterial input function (AIF)):

$$\text{CBV} = \frac{\int_{t=0}^{\infty} C_u(t) dt}{\int_{t=0}^{\infty} C_a(t) dt}. \quad (2)$$

Assuming no recirculation and therefore unimodality of the contrast curves, it is common to use the peak of the contrast curve as a temporal landmark. The time taken to reach that maximum is called time-to-peak (TTP).

It can be shown that the temporal relationship between the concentration at the feeding artery  $C_a$  and the target tissue  $C_u$  can be written as

$$C_u(t) = C_a(t) \otimes h(t), \quad (3)$$

where  $\otimes$  is the symbol for the convolution and  $h$  is the distribution of the transit times, as the contrast agent follows different paths through the vasculature. The transit times are related to the fraction of injected contrast agent still present in the vasculature at any given time  $t$ . This measure is described by the residue function  $R(t)$ :

$$R(t) = 1 - \int_{\tau=0}^t h(\tau) d\tau. \quad (4)$$

From  $R$ , we can establish the relation between the concentrations  $C_u$  and  $C_a$ :

$$C_u(t) = \text{CBF}(C_a \otimes R)(t) \quad (5)$$

which indicates that the contrast concentration  $C_u(t)$  in the target tissue at a given time  $t$  is proportional to the amount of blood passing through per unit time (i.e., CBF).

While the concentrations  $C_a$  and  $C_u$  can be estimated from DSA (see (1)), the residue function  $R$  and CBF require more complex computations. In practice, the concentration curves  $C_a$  and  $C_u$  are sampled at discrete time points,  $t_j \in [0, N - 1]$ :

$$C_u(t_j) = \Delta t \text{CBF} \sum_{i=0}^{N-1} C_a(t_i) R(t_j - t_i) \quad (6)$$

which can be rewritten as a matrix-vector notation:

$$C_u = \Delta t \text{CBF} C_A R, \quad (7)$$

where  $C_u, R \in \mathcal{R}^N$  and  $C_A$  is expanded to a Toeplitz matrix:

$$C_A = \begin{pmatrix} C_a(t_0) & 0 & \cdots & 0 \\ C_a(t_1) & C_a(t_0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ C_a(t_{n-1}) & C_a(t_{n-2}) & \cdots & C_a(t_0) \end{pmatrix}. \quad (8)$$

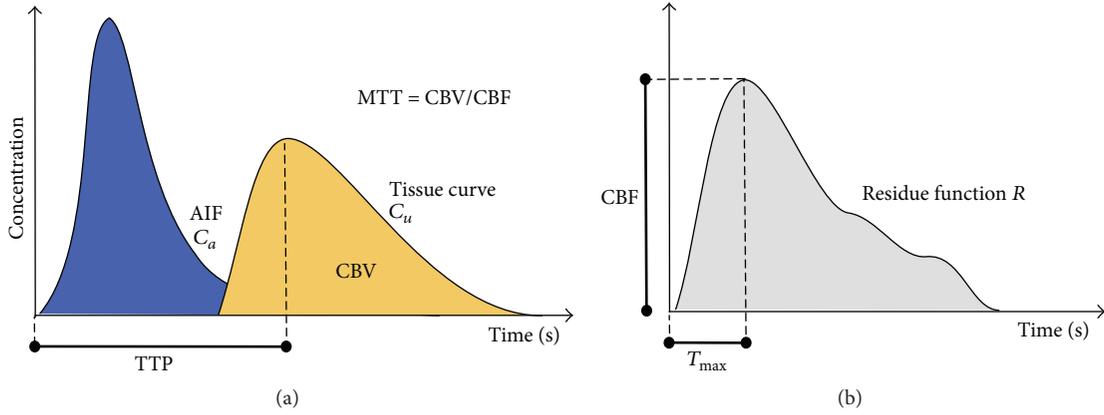


FIGURE 1: Illustration of a tissue concentration-time curve  $C_u$  (yellow) with respect to an arterial input function (AIF)  $C_a$  (blue). The deconvolution of the tissue curve  $C_u$  with  $C_a$  removes the dependence on the AIF and produces the residue function  $R$  (b). CBF is extracted at the maximum value reached at  $T_{\max}$ , while MTT is calculated as  $CBV/CBF$ , where  $CBV$  is determined as the area under the tissue curve (yellow). Because of the presence of arterial delays in stroke patients, the residue function is not always maximal at  $t = 0$  but might be maximal after a delay ( $T_{\max}$ ).

One way to recover  $R$  is to use singular value decomposition (SVD) of  $C_A$  into two orthogonal matrices,  $U$  and  $V^T$ , and a diagonal matrix,  $W$ , with singular values ordered descendingly in the diagonal,  $C_A = UWV^T$ . The solution is then given by

$$R = V\widehat{W}^{-1}U^T C_u, \quad (9)$$

where elements of  $\widehat{W}$  that are below a threshold are set to zero.

Given that  $\max(R) = 1$ , CBF is derived as the maximum of the estimated  $R$ , and  $T_{\max}$  is the time to reach this maximum. Once CBF has been estimated, MTT can be derived from the central volume theorem [28],  $MTT = CBV/CBF$ . The list of parameters extracted (CBF, CBV, MTT, TTP, and  $T_{\max}$ ) is illustrated in Figure 1.

**2.5. Solving Vessel Overlap with Gamma Mixture.** Overlap of the vessels may occur in biplane DSA and is one of the most challenging aspects of the estimation of perfusion parameters. This issue is illustrated in Figure 2 where a selected image location, shown as a yellow region, presents two contrast passages that lead to two peaks in the concentration-time curve. These two distributions correspond to the arterial and venous phase, respectively. The deconvolution method presented in Section 2.4 assumes unimodality of the concentration-time curve. Although it might be possible to use a previously acquired 3D model of the cerebrovasculature to delineate the vessels from the 2D projection, the direct processing of biplane DSA without any prior imaging is of great interest as other imaging modalities are not always available. To solve this problem, we suggest representing the concentration over time by a mixture of Gamma distributions that is automatically recovered at each point of the image using an expectation-maximization (EM) algorithm.

**2.5.1. Gamma-Variate Fitting.** The Gamma-variate function is the most commonly used prior distribution to represent

concentration-time curves as it has been shown to closely approximate the true contrast concentration. Drawing from the formulations present in the literature [29–31], we constrain the estimation of the concentration-time curves by assuming a minimum transit time  $\Delta_{\min}$  between the injection site and the brain which ensures that the maximum of the fitted distribution (which is also its inflection point) lies within the restricted domain. The density function  $\gamma_{\alpha,\beta}$  is written as

$$\gamma_{\alpha,\beta}(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} \exp^{-(x-\mu)\beta} (x-\mu)^{\alpha-1} & \text{if } x - \mu \geq \Delta_{\min} \\ 0 & \text{otherwise,} \end{cases} \quad (10)$$

where  $\alpha$ ,  $\beta$ , and  $\mu$  are the shape, scale, and location parameters, respectively. The Gamma function  $\Gamma(\alpha)$  is written as

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} \exp^{-t} dt. \quad (11)$$

The mean of the Gamma distribution is  $\alpha/\beta$ . The shape of the Gamma distribution is determined by the  $\alpha$  parameter, which intuitively relates to the contrast concentration variation. When  $\alpha > 1$ , the distribution is bell-shaped, suggesting low heterogeneity. In the case of  $\alpha < 1$ , the distribution is highly skewed which indicates high variation. This flexibility makes the distribution suitable for accommodating with different concentration-time curves as observed at different locations in the image.

**2.5.2. Mixture of Gamma-Variate Distributions.** To capture multiple contrast passages at a given image location, we propose to represent the concentration curve over time as a mixture of Gamma-variate distributions. This assumes that the overall distribution is generated from a few Gamma components, each with its own  $\alpha$  and  $\beta$  parameters. In our case,

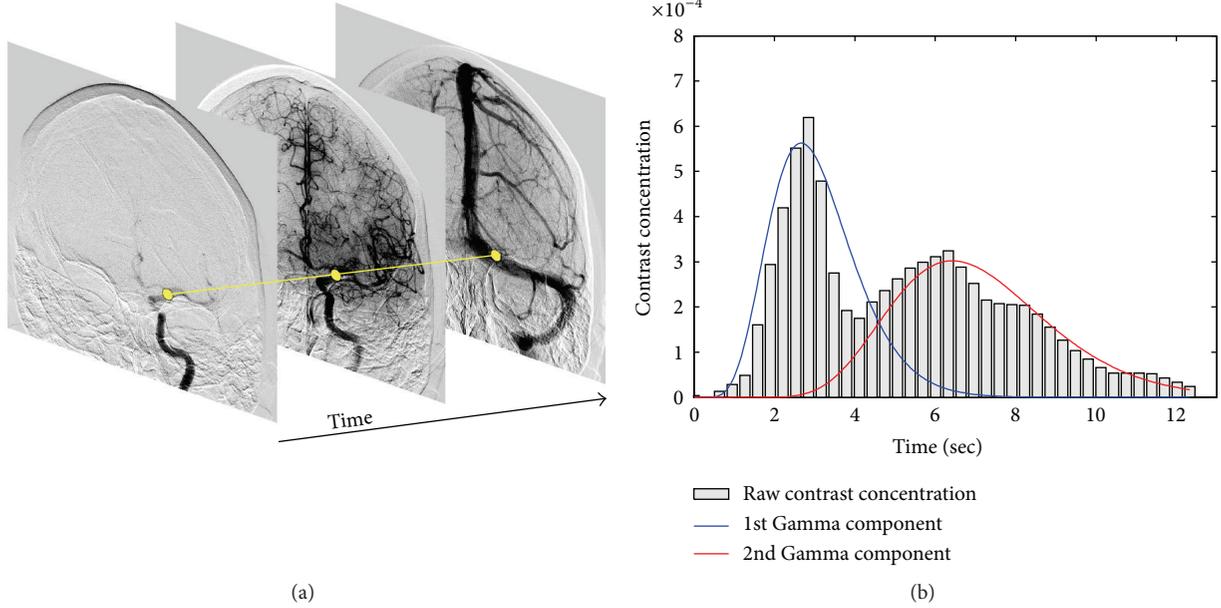


FIGURE 2: The bar graph of the contrast concentration-time curve (b) is shown for a specific location in a DSA sequence (shown in yellow on (a)). Two contrast passages can be observed in the concentration-time curve due to the overlap of the vessels. By applying the proposed method based on the EM algorithm, we are able to retrieve the individual components (represented by blue and red curves) using a Gamma mixture representation.

each component can be thought of as one contrast passage through one of the overlapped vessels at the current image location. Let  $K$  be the number of Gamma components in the mixture; the parameters of the  $j$ th component are denoted by  $\alpha_j$  and  $\beta_j$  and associated with the prior probability  $\tau_j$  that a measured concentration was drawn from the current component. Parameters of the overall distribution are summarized as  $\Theta = \{\alpha_j, \beta_j, \tau_j\}$ ,  $j = 1, \dots, K$ , with  $\sum_{j=1}^K \tau_j = 1$ , and the mixture is written as

$$\mathcal{M}(x, \Theta) = \sum_{j=1}^K \tau_j \gamma_{\alpha_j, \beta_j}(x), \quad (12)$$

where  $\gamma_{\alpha_j, \beta_j}(x)$  is the Gamma-variate distribution of the  $j$ th component (see (10)).

**2.5.3. Parameter Estimation.** The optimization of the parameters  $\Theta$  of the mixture is posed as a maximum likelihood estimation (MLE). The log-likelihood of parameter set  $\Theta$  is obtained by approximation using a weighted sum over discrete time:

$$\mathcal{L}(\Theta) = \sum_{i=1}^N \log \mathcal{M}(x_i, \Theta), \quad (13)$$

where  $i$  represents a discrete time point. The parameters  $\Theta$  of the model are unknown and are estimated using the expectation-maximization (EM) algorithm [32] which provides a convenient approximation in terms of an iterative maximization problem.

To be able to estimate the parameter set  $\Theta$  that maximizes  $\mathcal{L}$ , the EM algorithm introduces an unobservable matrix

$z \in \{0, 1\}^{N \times K}$  to specify which Gamma component the  $i$ th observation  $x_i$  comes from. In the original EM algorithm,  $z$  is defined as a binary variable that contains 1 for the component it comes from and 0 for all the others. Here, we use the soft EM definition where  $z$  is continuous and can take any value between 0 and 1, such that  $z \in [0, 1]^{N \times K}$ , and where the sum of the weights of each observed data point  $i$  is equal to 1,  $\sum_{j=1}^K z_{ij} = 1$ .

The complete discrete log-likelihood becomes

$$\mathcal{L}(\Theta) = \sum_{i=1}^N \sum_{j=1}^K z_{ij} \log \tau_j + C, \quad (14)$$

$$C = \sum_{i=1}^N \sum_{j=1}^K z_{ij} \log \gamma_{\alpha_j, \beta_j}(x_i).$$

EM uses the log-likelihood and iterates between the two following steps.

*E-Step.* Calculate the expected value  $Q(\Theta, \Theta^m)$  of the log-likelihood given current parameters  $\Theta^m$ , and

$$Q(\Theta, \Theta^m) = \sum_{i=1}^N \sum_{j=1}^K z_{ij}^m \log \tau_j + C, \quad (15)$$

where

$$z_{ij}^m = \frac{\tau_j^m \gamma_j(x_i; \alpha_j^m, \beta_j^m)}{\mathcal{M}(x_i, \Theta^m)}. \quad (16)$$

*M-Step.*  $Q(\Theta, \Theta^m)$  is maximized with respect to  $\Theta$  using numerical optimization

$$\Theta^{m+1} = \underset{\Theta}{\operatorname{argmax}} Q(\Theta, \Theta^m). \quad (17)$$

The iterative procedure is executed until the convergence criterion  $|\Theta^{m+1} - \Theta^m| < t_{EM}$  is satisfied or the maximum number of iterations reached (100). To avoid local maxima, it is repeated 5 times. The EM procedure can be performed for a different number of components  $K \in [1, 4]$ , for instance. The number  $K$  can be selected so that it minimizes the Bayesian Information Criterion (BIC) [33]. To allow for faster convergence and reduce the risk of falling into local maxima, the procedure is initialized with  $k$ -means algorithm.

*2.6. Experimental Setup.* This section describes the experimental protocol used in our study to evaluate the perfusion angiography framework. The proposed experiments will provide valuable insights about the following questions: Can the multivariate Gamma fitting method delineate individual contrast concentration curves in the presence of overlap and noise? How does it compare to a state-of-the-art fitting algorithm (RANSAC)? Is the computation of perfusion parameters from routinely acquired DSA feasible for assessment during endovascular interventions?

These questions are addressed by evaluating the perfusion angiography framework on two different experiments. The first experiment focuses on the estimation of the overlapped contrast concentration curves and identification of the individual components using the multivariate Gamma fitting technique (Section 2.5). To do so, we computed the average AIF concentration curves  $C_{a1}$  from 5 randomly selected patients from our dataset on which we selected a region of interest at a similar location on the intracerebral artery (ICA). The average concentration curve  $C_{a1}$  was smoothed using a Gaussian filter and interpolated to produce a set of  $N = 100$  values using bicubic interpolation. The overlap was simulated by duplicating the contrast curve  $C_{a1}$  to create a vector  $C_{a2}$ , shifting the duplicated vector  $C_{a2}$ , and merging them into a single vector  $C_{gt}$ , thus creating a simulated overlap between two similar contrast curves, as written as follows:

$$C_{gt}(i) = \begin{cases} C_{a1}(i) & \text{if } i - \text{shift} \leq 0, \forall i \in [1, N] \\ \max(C_{a1}(i), C_{a2}(i - \text{shift})) & \text{otherwise,} \end{cases} \quad (18)$$

where the ground truth  $C_{gt}$  corresponds to a multimodal contrast curve obtained from two contrast curves  $C_{a1}$  and  $C_{a2}$ , such that the latter is temporally shifted. In our experiments, we produced a set of merged concentration curves by varying the shifting amount from 5 to 100, ranging from almost full to no overlap. The objective of the experiment is then to measure how accurately it can fit and retrieve the two original contrast curves  $C_{a1}$  and  $C_{a2}$  using a Gamma-variate mixture  $\gamma_1, \gamma_2$  from the merged contrast curve  $C_{gt}$ . In addition to the evaluation of the robustness to the amount of overlap, various levels of white Gaussian noise are added to the signal, ranging from a SNR of 500 to 5.

Alternative methods to fitting Gamma distributions exist in the literature. Among them, the least squares fitting based on a discrete formulation would be possible but computationally costly. A more efficient technique is the random sample consensus (RANSAC) method [34] that has emerged as a versatile tool for robust parameter estimation in pattern recognition. It is typically used in computer vision to retrieve correspondence between images and estimate the geometric transformation matrix that relates them. The idea behind RANSAC is to estimate a large number of minimal-set fitting hypotheses. For each hypothesis, a robust score is calculated; this score is based on the alignment of the hypothesis with all points in the set. The best scoring minimal-set hypothesis is taken as the final estimate. In our experiments, a total of 300 fitting hypotheses were used and each hypothesis was made of 15 points. The accuracy of both the Gamma-variate and the RANSAC models is measured as the coefficient of determination or  $R$ -squared. For better estimation of the error, the process is repeated 10 times for each combination of error and overlap, and the average  $R$ -squared is reported.

For the second experiment, we ran the perfusion angiography on our dataset (Section 2.2) composed of DSA sequences following endovascular thrombectomy recorded on 66 acute ischemic stroke patients with MCA occlusion. The experiments are formulated such that the distribution of a given perfusion parameter across the MCA territory is averaged and studied with respect to the TIC1 score. Statistical measures of correlation and dispersion are extracted.

During our experiments, source DSA images of each patient are processed with perfusion angiography. The concentration-time curve of the arterial input function (AIF)  $C_a$  (see (2)) required for the computation of perfusion maps is obtained by extracting the average of the DSA values comprised within a region of interest (ROI) at each time point. This ROI was manually selected by a UCLA neurologist on the source DSA of each patient prior to the processing. In this study, it was set on the intracerebral artery (ICA) as an elliptical region fully included in the vessel. Note that, similar to perfusion MRI, it should be possible to detect or estimate the AIF automatically using constraints on early arrival time and maximum contrast values. However, to minimize possible source of error for the computation of perfusion parameters in this study, we chose to delineate the AIF manually. The perfusion angiography was ran using the BIC criterion to select among a maximum of two Gamma components to differentiate between the arterial and the venous phase. After processing, the following parameter maps are available; CBF,  $CBV_{full}$ , CBV, MTT, TTP, and  $T_{max}$ , where  $CBV_{full}$  is the cerebral blood flow computed over the entire cerebral cycle (including arterial and venous phases) and CBV is computed during the arterial phase only.

In order to evaluate the five perfusion parameters presented in Section 2.4 (map  $\in \{CBF, CBV_{full}, CBV, MTT, TTP, T_{max}\}$ ), the parameter maps need to be transformed into quantitative values  $x_{map}$  that can be used as input to the statistical analysis. As a preprocessing step, a UCLA neurologist (blinded to outcome and perfusion maps) delineated the MCA territory on each DSA using a template scaled down and rotated to cover the entire territory. Each perfusion

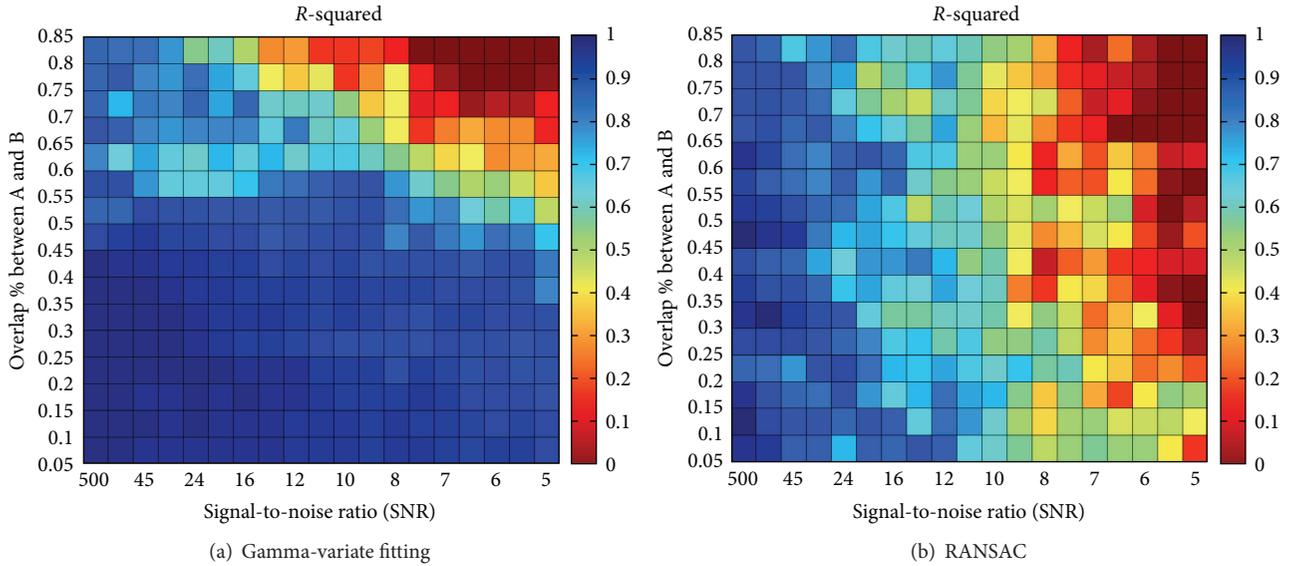


FIGURE 3: Illustration of the  $R$ -squared correlation coefficient between the estimated Gamma components and the ground truth for various levels of Gaussian white noise in terms of signal-to-noise ratio (SNR) and percentage of overlap between the two original components. The results are reported for the Gamma-variate method (a) and the RANSAC algorithm (b).

parameter is then characterized using the trimmed mean of the distribution of the values within the ROI. The trimmed mean computes the average of the values comprised between the 5th and 95th percentiles:

$$x_{\text{map}} = \frac{\sum_i \sum_j v(i, j)}{N_v}, \quad (19)$$

where  $N_v$  is the number of points included in the ROI and comprised between the 5th and 95th percentiles and  $v(i, j)$  is the value of the perfusion map at point  $[i, j]$ :

$$v(i, j) = \begin{cases} \text{map}(i, j); & \text{if } t_5 < \text{map}(i, j) < t_{95} \ \&\& \ \text{ROI}(i, j) == 1 \\ 0; & \text{otherwise.} \end{cases} \quad (20)$$

We evaluate the Pearson correlation between the following pairs of variables: (CBF, TICI), (CBV, TICI), (TTP, TICI), (MTT, TICI), and ( $T_{\text{max}}$ , TICI). To facilitate the statistical analysis, qualitative TICI scores (“0,” “1,” “2a,” “2b,” and “3”) are mapped to a continuous space, as follows: (“0,” 0); (“1,” 0.25); (“2a,” 0.5); (“2b,” 0.75); (“3,” 1).

### 3. Results

The results of the first experiment are reported in Figure 3 with the  $R$ -squared coefficient between the ground truth and the recovered mixture. It was computed for various levels of noise and degrees of overlap between two simulated contrast concentration curves within the ground truth. It can be observed that the Gamma-variate fitting framework is able to accurately retrieve the two components of the mixture in the presence of noise when the overlap is below 55%. When the overlap is greater than 55%, the accuracy

decreases significantly as the noise increases. As expected, the model fails to accurately recover the two components in the presence of very high levels of noise (SNR < 8) and high percentage of overlap (>70%). Fitting results are illustrated in Figure 4 for four different combinations of overlap amount and noise levels. RANSAC recovers the components with a decent accuracy regardless of overlap until a SNR of about 10, and then the error drastically increases in the presence of higher levels of noise. In comparison, the standard estimation of TTP (without multimodal fitting) taken at the maximum of the concentration-time curve would be misplaced in half of the cases depending on which component is the highest.

In the second experiment, the perfusion angiography framework processed successfully 89% (59 out of 66) of the DSA images included in our dataset. Seven cases failed during processing due to either patient motion, short acquisition time (i.e., the DSA acquisition did not cover the entire injection cycle), poor image quality, or low temporal resolution (i.e., insufficient number of frames).

As a first observation, we noted that most of the patients included in our dataset (93%; 55 out of 59) had poor outcomes (mRS greater than or equal to 3). We also noticed that a TICI score of 2b leads to a slightly better mRS outcome than 2a. However, patients that reached a TICI score of 3 (i.e., complete reperfusion of the MCA territory) were not associated with a better outcome than 2b patients. The phenomenon of futile recanalization is similar to what has been reported in other studies [35]. Possible explanation may include increased risk of hemorrhagic transformation. NIHSS at admission is linearly correlated with mRS outcome ( $r = 0.304$ ,  $p < 0.028$ ). In addition, low NIHSSs (i.e., not severe) are associated with larger variations in terms of outcome.

Linear regression analysis between CBF and CBV values revealed an overall strong correlation ( $r = 0.736$ ,  $p < 10^{-12}$ ).

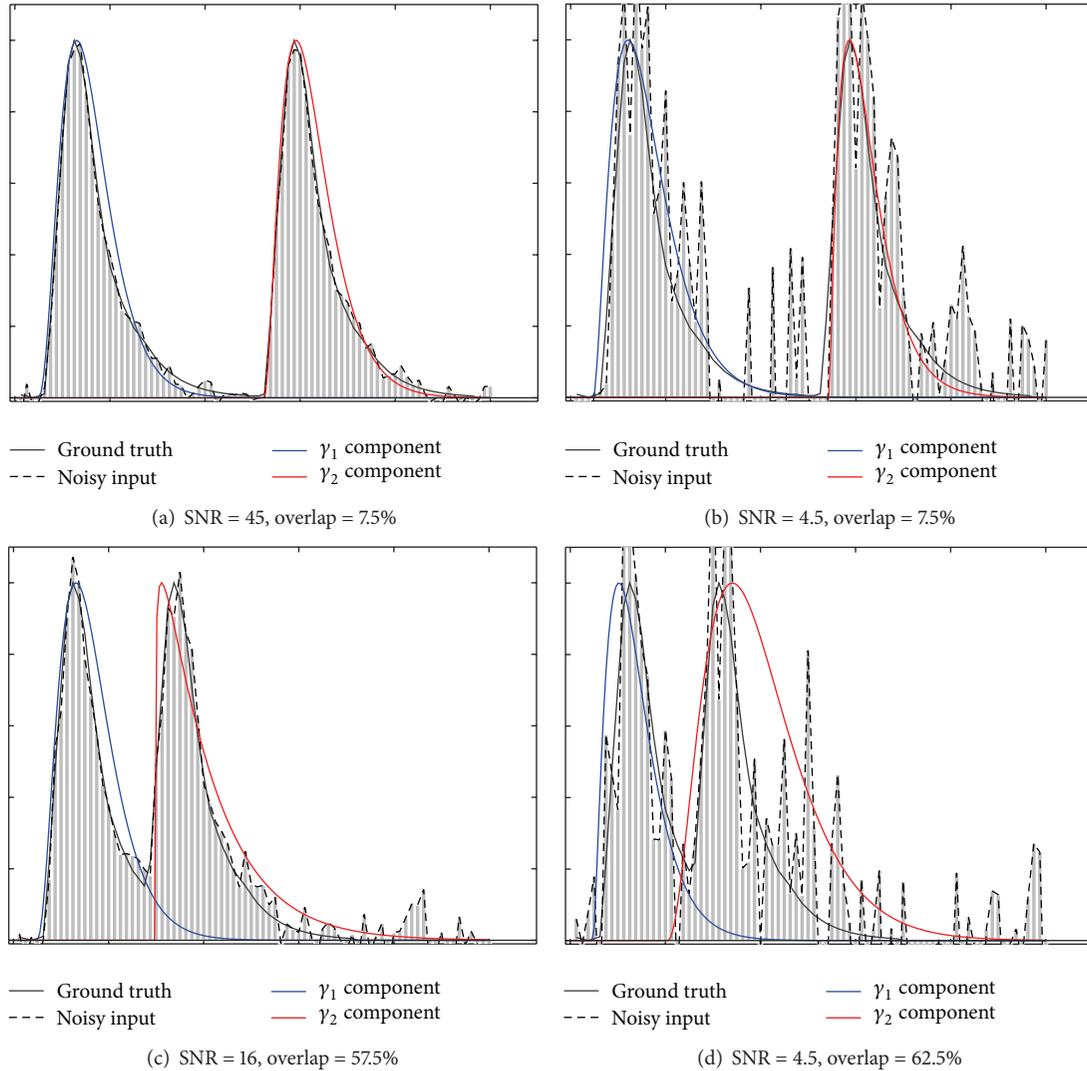


FIGURE 4: Illustration of the Gamma fitting process to recover two components for 4 different combinations of noise and overlap. Components  $\gamma_1$  and  $\gamma_2$  are shown in blue and red and were estimated using the EM-based Gamma-variate fitting (Section 2.5) based on the noisy input depicted by the dashed line.

Both values are estimated with perfusion angiography and averaged over the entire MCA territory. CBV was computed during the arterial phase of the cycle. This is an expected result that has been shown in previous MR and CT studies of perfusion [36] and could in principle be used to identify infarcted areas from penumbra [37].

Scatter plots representing the CBV and CBF perfusion angiography maps versus TICI score are illustrated in Figure 5 where each patient is depicted by a circle. The plots include CBF versus TICI (a) and CBV versus TICI (b). When plotted versus TICI, CBF shows a sign of positive correlation ( $r = 0.292$ ,  $p < 0.064$ ). However, low CBF is not always synonym of poor TICI score as slower flow might still lead to good revascularization and therefore a high TICI score. This may explain why larger TICI variations are observed for cases associated with low CBF. When CBV is studied with respect to TICI (Figure 5(b)), it shows weaker correlation ( $r = 0.218$ ,

$p < 0.170$ ). Significantly higher delays in terms of TTP were measured in the MCA territory for patients with no revascularization (TICI = 0). For other TICI grades, there was no correlation with TTP. Absence of equivalence between TICI and CBF/CBV estimated with perfusion angiography does not imply superiority of one measure to the other but rather it implies that they provide a different, perhaps complementary set of information.

The parametric maps computed for 8 patients are displayed in Figures 6(a) and 6(b). For each patient, the perfusion parameters are illustrated, including CBF,  $CBV_{full}$  (computed over the entire arteriovenous cycle), CBV (computed over the arterial phase), MTT, and TTP. For display purposes, each parametric map is normalized and color-coded to facilitate visualization. Red is used to show high value ( $\uparrow$  flow for CBF,  $\uparrow$  volume for CBV, and  $\uparrow$  delay for MTT and TTP), and blue is used to represent low values. In

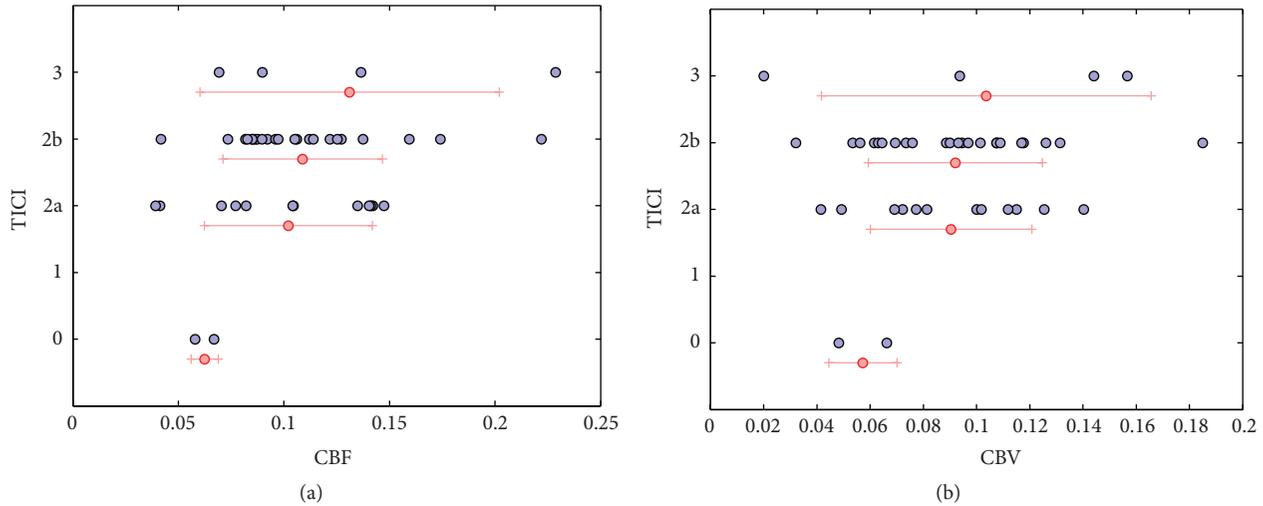


FIGURE 5: Scatter plots representing the perfusion angiography CBV and CBF versus TICI score. Average and standard deviation for specific TICI values are depicted by red lines.

addition, the source DSA on which perfusion angiography was performed is shown on the bottom row of each case. For matter of space, a subset of seven frames were sampled and displayed for each DSA sequence.

One of the particularities of the perfusion maps is to be bidimensional; therefore, a single image region may represent different anatomical structures that overlap across that region. Despite this limitation, these maps provide fine detail as they match the original spatial resolution of the DSA ( $1024 \times 1024$  in our dataset). When reviewed side by side, CBF, CBV, and TTP maps may help the expert eye to differentiate between antegrade and collateral flow and identify risk of hemorrhage, perfusion deficit, delay, and flow stagnation. The computation of the perfusion parameters for a single patient took 21 seconds. In principle, faster execution times can be obtained as the estimation of the perfusion parameters can be parallelized.

#### 4. Discussion

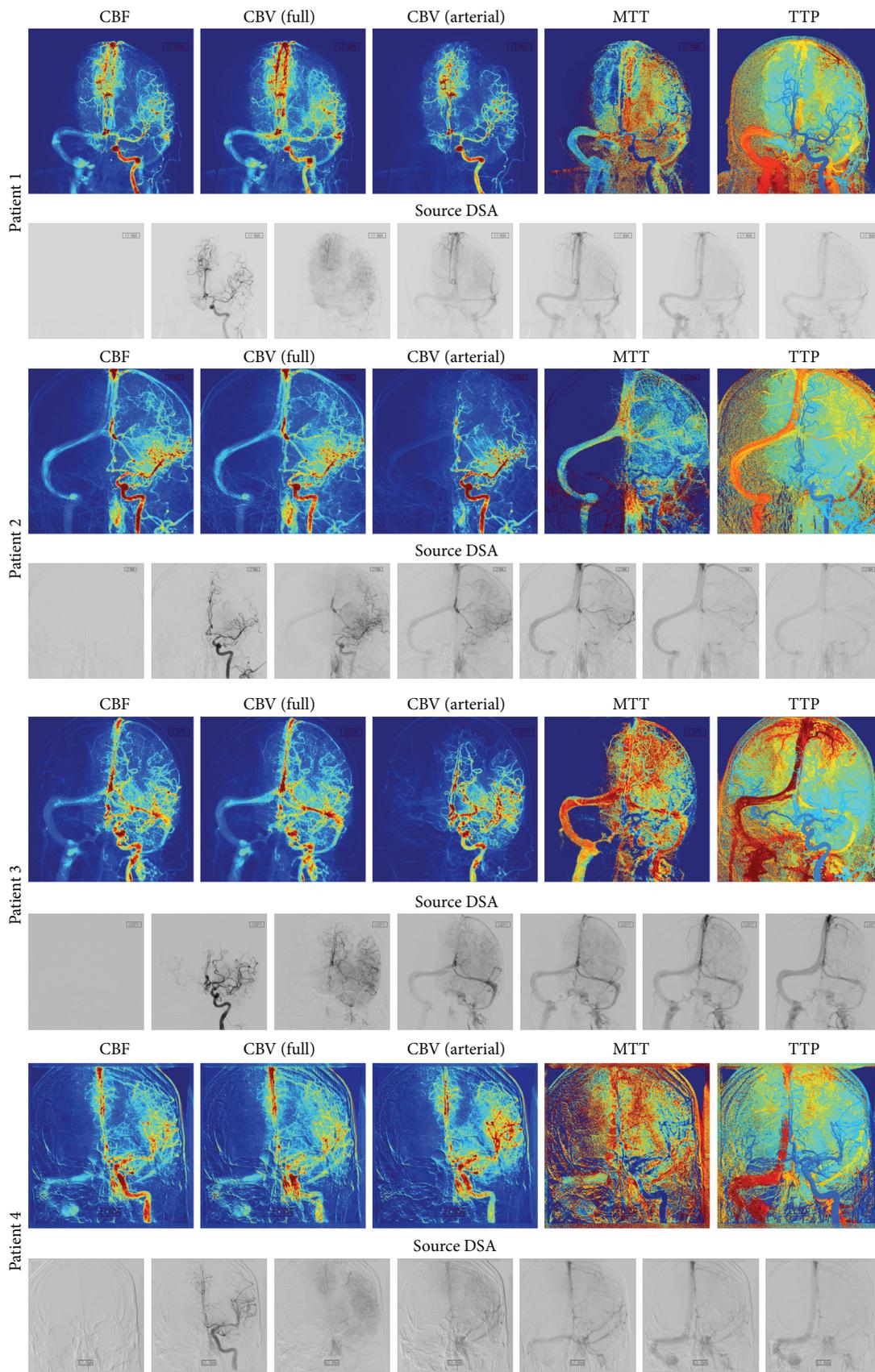
There is an overt need to provide imaging-based decision support to better guide and accelerate endovascular interventions in acute stroke. Among the available imaging techniques, DSA is a method of choice to visualize blood flow and guide endovascular interventions. Biplane DSA provides high-resolution spatiotemporal images that have mostly been used qualitatively through the manual review of raw grayscale video. The interpretation of DSA images could benefit from color-coded perfusion parameters that would enable the visualization of hemodynamic features that are not directly visible on source angiograms and allow for refined decisions without any delay in care, added X-ray exposure, or higher dose of contrast agent.

We have introduced in this paper a computational framework for the extraction of quantitative perfusion parameters from routine DSA. Similar to CT/MR perfusion, our

approach uses a deconvolution technique to derive CBF, CBV, MTT, TTP, and  $T_{\max}$ . A novel computational solution based on multimodal fitting was introduced to deal with overlap of the vessels. This study has demonstrated that routinely acquired DSA can be used to derive perfusion parameters that are similar in spirit to the ones obtained from CT/MR perfusion. However, the interpretation perfusion DSA is different due to the nature of the view (frontal or lateral) and the overlap of several brain structures within a given location.

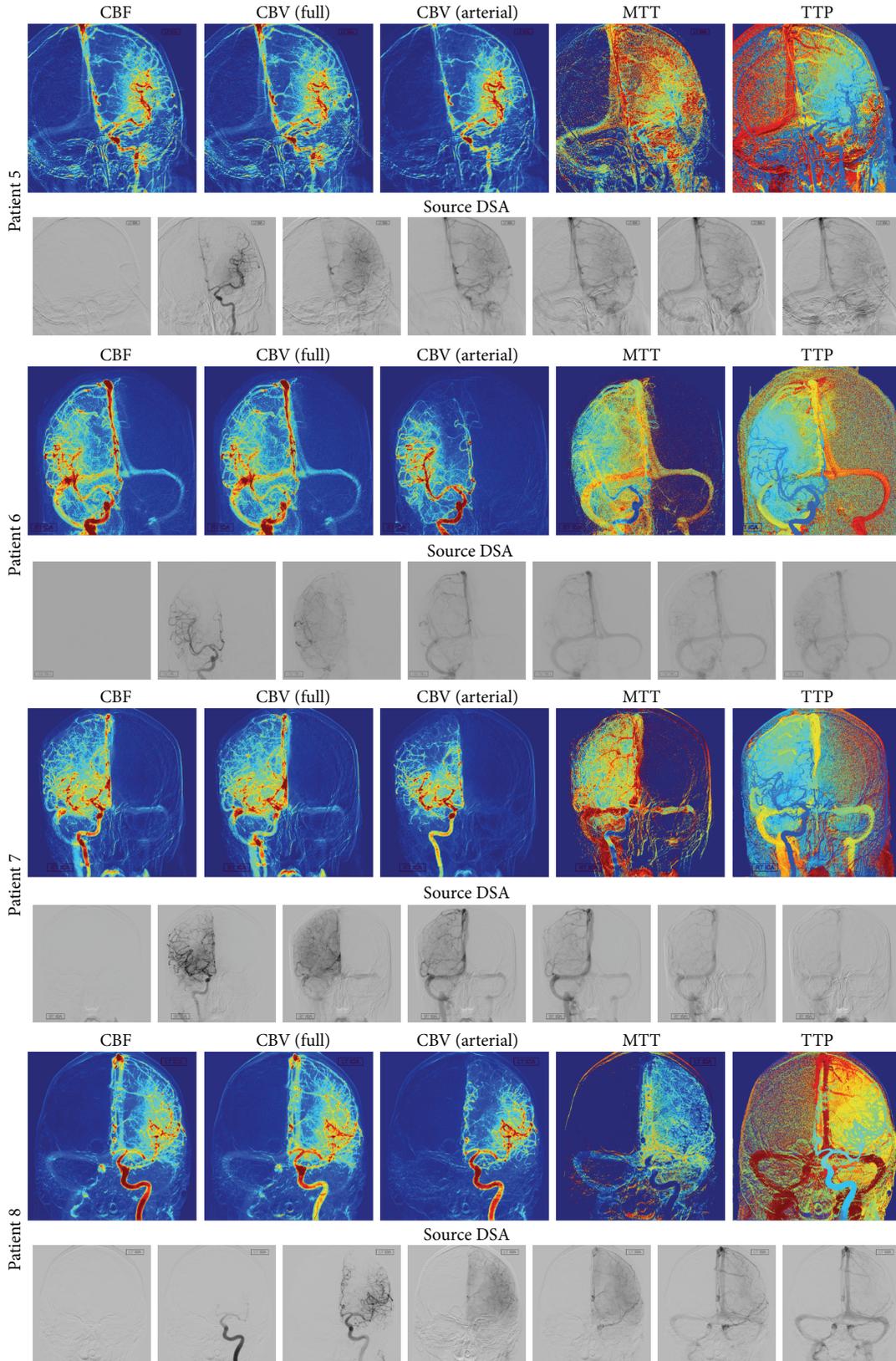
Taking a step back, it is clear that neuroimaging provides neurologists and neurointerventionalists with an immense source of information for guidance in clinical decision-making. Yet, perhaps because of the abundance of information held in those images, their use remains suboptimal. At UCLA, for example, the following modalities can be acquired: magnetic resonance imaging/angiography (MRI/MRA), diffusion/perfusion-weighted MRI, computed tomography/tomographic angiography (CT/CTA), perfusion CT, and digital subtraction angiography (DSA). Broadly speaking, these images offer different insights and mirror different steps of the therapy. Neuroimaging is used before treatment to classify the stroke using lesion size, tissue at risk, and involved vascular territory. This allows identifying stroke patients who can benefit the most from a specific treatment strategy and outweigh its potential risks. DSA images are acquired during therapy for decision-making. These iterative landmarks can be used to evaluate the degree of reperfusion and recanalization by visual scoring. Beyond the acute phase, neuroimaging is helpful in evaluating recovery and guiding other management strategies such as the augmentation of cerebral perfusion and reduction of mass effects from hemorrhage. Validation of perfusion angiography for estimation of hypoperfusion volume or degree of recanalization and reperfusion during endovascular interventions would be of great interest.

The need for neuroimaging insight is triggered by the complexity of personalized treatment and variability of stroke



(a)

FIGURE 6: Continued.



(b)

FIGURE 6: (a) Parametric maps computed for 4 patients. For each patient, the perfusion parameters are illustrated, including CBF, CBV (full) (computed over the entire arteriovenous cycle), CBV (arterial) (computed over the arterial phase), MTT, and TTP. The source DSA is shown on the bottom row of each patient. (b) Parametric maps computed for 4 patients. For each patient, the perfusion parameters are illustrated, including CBF, CBV (full) (computed over the entire arteriovenous cycle), CBV (arterial) (computed over the arterial phase), MTT, and TTP. The source DSA is shown on the bottom row of each patient.

outcomes. The patient population in acute ischemic stroke is incredibly heterogeneous; it presents a wide variety of outcomes and responses to treatment. For example, although the degree of recanalization correlates favorably with outcome, the risk of death remains stable. In addition, while the time from symptoms onset also correlates with outcome on average, it is not rare to observe that late recanalizers do better than early ones. These paradoxical observations can be linked to several factors such as blood pressure, NIH Stroke Scale (NIHSS), or age, but their individual predictive value is too weak for supporting prospective clinical decisions. The presence of collateral circulation beyond the site of occlusion may also be decisive as it could sustain tissue viability until recanalization occurs; however, its presence largely varies across patients. Therefore, careful patient selection for endovascular intervention based on collateral circulation and tissue status is key to tailor interventions and improve outcomes. Currently, collateral flow is evaluated on DSA but remains challenged by the lack of quantitative measure. Automatic evaluation of collateral flow and revascularization may be possible from perfusion angiography and should be considered for future studies.

**4.1. Challenges and Limitations.** The proposed framework holds several limitations due to the bias existing in the dataset studied, the nature of the source images, and other technical challenges related to the computation and acquisition. We discuss in this section these limitations and how they might affect the results obtained and could potentially be tackled in future studies.

A limiting factor of the study is that the dataset used in our experiments was rather small and not evenly distributed across degree of revascularization. Only five patients had poor TICl scores of 0 or 1. Conversely, mRS outcome was poor for most of the patients. Four patients fell within the range of mRS  $\in [0, 2]$ . Although the study of such a dataset can provide a proof of concept and applicability of the techniques, drawing conclusions or guidelines from the statistical analysis and the generalization power of specific perfusion parameters at the population level would require a larger, multicenter dataset. Such a study could provide more reliable estimates and possible relationship with outcome.

On the technical side, phantom calibration would in principle be required to obtain contrast curves that are generalizable across patients and hardware configuration. Because the study proposed in this paper was performed retrospectively on routinely acquired DSA, calibration values were not available. This lack of normalization has likely introduced errors in the estimated parameters. A prospective animal study with phantom calibration would be appropriate to test the accuracy of the parameter maps (especially CBF). In addition, vessel thickness was not considered during the computation of the perfusion maps. Because of this, it is likely that flow may have been incorrectly estimated in large vessel area. It should be possible to solve this problem by coupling the computation of the perfusion parameters with an automatic vessel detector that could extract the vessel diameter along the cerebrovasculature.

The time resolution of the DSA sequence (i.e., frame rate of the acquisition) has a great impact on the quality of the perfusion parameters. When the number of frames is too low, it becomes very difficult to delineate the different vessels when an overlap occurs. In addition, the estimation of CBV and TTP becomes approximative. Unlike CT and MRI where the time interval between each acquisition is kept constant, DSA sequences are acquired with a varying frame sampling rate. Therefore, the set of points of the contrast curve needs to be resampled and interpolated. In our framework, cubic spline interpolation was used and the interval was chosen as the minimum time interval observed between two successive frames in the current acquisition. Systematic tests should be performed to evaluate the sensitivity of the estimation of the perfusion parameters with respect to the frame rate.

There is a margin for improvement of the perfusion angiography framework by tackling these limitations. One of the most promising research directions would be to perform a comparative analysis to test equivalence of perfusion parameters estimated from DSA using perfusion angiography to the one obtained via MR or CT perfusion.

## 5. Conclusion

We have introduced in this paper perfusion angiography, a methodological framework for the quantitative analysis and visualization of blood flow parameters from DSA images. The parameters, including cerebral blood flow (CBF) and cerebral blood volume (CBV), mean transit time (MTT), time-to-peak (TTP), and  $T_{\max}$ , were reliably estimated using a bolus tracking method based on the deconvolution of the time-density curve on a pixel-by-pixel basis.

Although further study on a larger dataset would be necessary to establish statistical correspondence with outcome and TICl score and to provide comparative analysis with estimated parameter maps obtained with MR and CT perfusion, the proposed imaging tool provides unique insights into flow mechanisms that cannot be observed directly in DSA sequences and may be used to quantify perfusion impact of endovascular interventions.

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

This work was supported by a Beginning Grant-in-Aid award from the American Heart Association (AHA) no. 16BGIA27760152 to Professor F. Scalzo.

## References

- [1] A. Moniz, *Diagnostic des Tumeurs Cérébrales et Épreuve de L'encéphalographie Artérielle*, Masson et Cie, 1931.
- [2] M. M. Frost, H. D. Fisher, S. Nudelman, and H. Roehrig, "A digital video acquisition system for extraction of subvisual infor-

- mation in diagnostic medical imaging,” *Proceedings of SPIE*, vol. 127, pp. 208–215, 1977.
- [3] T. W. Ovitt, M. P. Capp, H. D. Fisher et al., “The development of a digital video subtraction system for intravenous angiography,” in *Proceedings of the Noninvasive Cardiovascular Measurements*, vol. 167 of *Proceedings of SPIE*, pp. 61–66.
  - [4] T. W. Ovitt, P. C. Christenson, H. D. Fisher III et al., “Intravenous angiography using digital video subtraction: X-ray imaging system,” *American Journal of Roentgenology*, vol. 135, no. 6, pp. 1141–1144, 1980.
  - [5] T. Tomsick, “TIMI, TIBI, TICl: I came, I saw, I got confused,” *American Journal of Neuroradiology*, vol. 28, no. 2, pp. 382–384, 2007.
  - [6] S. D. Shpilfoygel, R. A. Close, D. J. Valentino, and G. R. Duckwiler, “X-ray videodensitometric methods for blood flow and velocity measurement: a critical review of literature,” *Medical Physics*, vol. 27, no. 9, pp. 2008–2023, 2000.
  - [7] G. W. Albers, V. N. Thijs, L. Wechsler et al., “Magnetic resonance imaging profiles predict clinical response to early reperfusion: the diffusion and perfusion imaging evaluation for understanding stroke evolution (DEFUSE) study,” *Annals of Neurology*, vol. 60, no. 5, pp. 508–517, 2006.
  - [8] W. A. Bateman and R. A. Kruger, “Blood flow measurement using digital angiography and parametric imaging,” *Medical Physics*, vol. 11, no. 2, pp. 153–157, 1984.
  - [9] H. Touho, “Hemodynamic evaluation with dynamic DSA during the treatment of cerebral vasospasm: a retrospective study,” *Surgical Neurology*, vol. 44, no. 1, pp. 63–74, 1995.
  - [10] C. M. Strother, F. Bender, Y. Deuerling-Zheng et al., “Parametric color coding of digital subtraction angiography,” *American Journal of Neuroradiology*, vol. 31, no. 5, pp. 919–924, 2010.
  - [11] G. J. S. Hunter, J. V. Hunter, and N. J. G. Brown, “Parametric imaging using digital subtraction angiography,” *British Journal of Radiology*, vol. 59, no. 697, pp. 7–11, 1986.
  - [12] W. C. Röntgen, “On a new kind of rays,” *Nature*, vol. 53, no. 1369, pp. 274–276, 1896.
  - [13] A. E. Moniz, “L’encephalographie arterielle, son importance dans la localisation des tumeurs cerebrales,” *Revue Neurologique*, vol. 2, no. 34, pp. 72–90, 1927.
  - [14] T. F. Meaney, M. A. Weinstein, E. Buonocore et al., “Digital subtraction angiography of the human cardiovascular system,” *American Journal of Roentgenology*, vol. 135, no. 6, pp. 1153–1160, 1980.
  - [15] D. P. Harrington, L. M. Boxt, and P. D. Murray, “Digital subtraction angiography: overview of technical principles,” *American Journal of Roentgenology*, vol. 139, no. 4, pp. 781–786, 1982.
  - [16] S. Balter, D. Ergun, E. Tscholl, F. Buchmann, and L. Verhoeven, “Digital subtraction angiography: fundamental noise characteristics,” *Radiology*, vol. 152, no. 1, pp. 195–198, 1984.
  - [17] F. Buchmann, “Digital subtraction angiography (DSA): new perspectives in angiography,” *Neurosurgical Review*, vol. 8, no. 1, pp. 35–40, 1985.
  - [18] R. Le Goff and Y. Bizais, “Densitometric measurement of blood flow application to stenosis quantification,” in *Medical Imaging VI: Image Processing*, vol. 1652 of *Proceedings of SPIE*, pp. 378–385, February 1992.
  - [19] R. A. Kruger, C. A. Mistretta, and S. J. Riederer, “Physical and technical considerations of computerized fluoroscopy difference imaging,” *IEEE Transactions on Nuclear Science*, vol. 28, no. 1, pp. 205–212, 1981.
  - [20] A. V. Clough, G. S. Krenz, M. Owens, A. Al-Tinawi, C. A. Dawson, and J. H. Linehan, “An algorithm for angiographic estimation of blood vessel diameter,” *Journal of Applied Physiology*, vol. 71, no. 5, pp. 2050–2058, 1991.
  - [21] L. E. Fencil, K. Doi, K. G. Chua, and K. R. Hoffman, “Measurement of absolute flow rate in vessels using a stereoscopic DSA system,” *Physics in Medicine and Biology*, vol. 34, no. 6, pp. 659–671, 1989.
  - [22] M. Lin, C. T. Marshall, Y. Qi et al., “Quantitative blood flow measurements in the small animal cardiopulmonary system using digital subtraction angiography,” *Medical Physics*, vol. 36, no. 11, pp. 5347–5358, 2009.
  - [23] M. de Lin, L. Ning, C. T. Badea, N. N. Mistry, Y. Qi, and G. A. Johnson, “A high-precision contrast injector for small animal X-ray digital subtraction angiography,” *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 3, pp. 1082–1091, 2008.
  - [24] A. Gidlund, “Development of apparatus and methods for roentgen studies in haemodynamics,” *Acta radiologica*, no. 130, pp. 7–70, 1956.
  - [25] S. Yerushalmi and Y. Itzchak, “Angiographic methods for blood flow measurements,” *Medical Progress through Technology*, vol. 4, no. 3, pp. 107–115, 1976.
  - [26] L. Østergaard, R. M. Weisskoff, D. A. Chesler, G. Gyldensted, and B. R. Rosen, “High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: mathematical approach and statistical analysis,” *Magnetic Resonance in Medicine*, vol. 36, no. 5, pp. 715–725, 1996.
  - [27] L. Østergaard, “Principles of cerebral perfusion imaging by bolus tracking,” *Journal of Magnetic Resonance Imaging*, vol. 22, no. 6, pp. 710–717, 2005.
  - [28] G. N. Stewart, “Researches on the circulation time in organs and on the influences which affect it: parts I.-III,” *The Journal of Physiology*, vol. 15, no. 1-2, pp. 1–89, 1893.
  - [29] M. Rausch, K. Scheffler, M. Rudin, and E. W. Radü, “Analysis of input functions from different arterial branches with gamma variate functions and cluster analysis for quantitative blood volume measurements,” *Magnetic Resonance Imaging*, vol. 18, no. 10, pp. 1235–1243, 2000.
  - [30] P. Schramm, Y. Huang, G. Erb, E. Klotz, and S. Heiland, “How does the injection protocol influence the attenuation-time curve in CT perfusion measurements: comparison of measured and simulated data,” *Medical Physics*, vol. 36, no. 8, pp. 3487–3494, 2009.
  - [31] M. Wagner, Y. Deuerling-Zheng, M. Möhlenbruch, M. Bendzus, J. Boese, and S. Heiland, “A model based algorithm for perfusion estimation in interventional C-arm CT systems,” *Medical Physics*, vol. 40, no. 3, Article ID 031916, 2013.
  - [32] R. Sundberg, “Maximum likelihood theory for incomplete data from an exponential family,” *Scandinavian Journal of Statistics*, vol. 1, no. 2, pp. 49–58, 1974.
  - [33] G. Schwarz, “Estimating the dimension of a model,” *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.
  - [34] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
  - [35] M. P. Marks, M. G. Lansberg, M. Mlynash et al., “Correlation of AOL recanalization, TIMI reperfusion and TICl reperfusion with infarct growth and clinical outcome,” *Journal of NeuroInterventional Surgery*, vol. 6, no. 10, pp. 724–728, 2014.

- [36] D. G. Nabavi, A. Cenic, R. A. Craen et al., "CT assessment of cerebral perfusion: experimental validation and initial clinical experience," *Radiology*, vol. 213, no. 1, pp. 141–149, 1999.
- [37] B. D. Murphy, A. J. Fox, D. H. Lee et al., "White matter thresholds for ischemic penumbra and infarct core in patients with acute stroke: CT perfusion study," *Radiology*, vol. 247, no. 3, pp. 818–825, 2008.

## Research Article

# Detecting the Intention to Move Upper Limbs from Electroencephalographic Brain Signals

**Berenice Gudiño-Mendoza, Gildardo Sanchez-Ante, and Javier M. Antelis**

*Tecnologico de Monterrey, Campus Guadalajara, Avenida General Ramón Corona 2514, 45201 Zapopan, JAL, Mexico*

Correspondence should be addressed to Berenice Gudiño-Mendoza; [bereniceg@itesm.mx](mailto:bereniceg@itesm.mx)

Received 24 December 2015; Accepted 21 February 2016

Academic Editor: Reza Khosrowabadi

Copyright © 2016 Berenice Gudiño-Mendoza 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.

Early decoding of motor states directly from the brain activity is essential to develop brain-machine interfaces (BMI) for natural motor control of neuroprosthetic devices. Hence, this study aimed to investigate the detection of movement information before the actual movement occurs. This information piece could be useful to provide early control signals to drive BMI-based rehabilitation and motor assisted devices, thus providing a natural and active rehabilitation therapy. In this work, electroencephalographic (EEG) brain signals from six healthy right-handed participants were recorded during self-initiated reaching movements of the upper limbs. The analysis of these EEG traces showed that significant event-related desynchronization is present before and during the execution of the movements, predominantly in the motor-related  $\alpha$  and  $\beta$  frequency bands and in electrodes placed above the motor cortex. This oscillatory brain activity was used to continuously detect the intention to move the limbs, that is, to identify the motor phase prior to the actual execution of the reaching movement. The results showed, first, significant classification between relax and movement intention and, second, significant detection of movement intention prior to the onset of the executed movement. On the basis of these results, detection of movement intention could be used in BMI settings to reduce the gap between mental motor processes and the actual movement performed by an assisted or rehabilitation robotic device.

## 1. Introduction

Brain-machine interface (BMI) systems are emerging technologies that provide a novel communication channel for both healthy people and patients with limited communication or motor impairments [1, 2]. A BMI system decodes the mental tasks performed by the user using brain signals recorded with invasive or noninvasive techniques. This is in turn used to control an application or an external device such as the computer cursor, a robotic wheelchair, or an orthosis device [3, 4]. As the peripheral nervous system is not involved in this process, a BMI can be a promising assistive technology for people with partial or complete motor disabilities [5, 6]. The most important application of BMI systems is the control of motor assisted robotic devices, which are employed for motor restoration or motor rehabilitation [7, 8]. This also includes rehabilitation scenarios based on virtual reality environments [9, 10]. These applications may provide patients

who suffered stroke or spinal cord injury with the possibility of shortening the recovery period to improve their motor functioning.

The cue-based synchronous protocol is the traditional paradigm employed to decode mental tasks from the brain activity in BMI settings. In this paradigm, the mental task is first performed by the user and then the BMI technology applies decoding algorithms to identify the task [8]. Then, a control signal or command is provided to drive neuroprosthesis [11], for example, a real or virtual robotic arm. For instance, in a BMI based on motor imagery (MI), the user performs mental imagination of different limb movements and then the BMI technology identifies the moved limb which is used to provide a command of movement in the application [12]. Thus, the user's mental task is associated with the movement provided by the application. However, there exists an inherent delay between the time of the mental task and the time of the movement performed by the

application. In consequence, the movements performed by external devices, either real or virtual, are not found natural by the user.

Reducing the temporal gap between mental motor tasks and the actual movement performed by an assisted or rehabilitation robotic device might be useful to obtain fast and natural motor control. This can also promote motor recovery at the cortical level [13, 14]. To achieve this early detection of movement information, this work studies the decoding of natural movement information before a motor task is performed. Several previous works have studied this kind of decoding from the noninvasive electroencephalographic (EEG) brain signals. These studies are based on movement related cortical potentials (MRCP), spectral power (SP), and common spatial patterns (SCP) features of the EEG to detect movement information preceding actual movements. Table 1 presents a summary of the state of the art of the most relevant works devoted to the detection of movement intention. These studies have demonstrated the feasibility of detecting motor information before a movement is performed.

Nonetheless, more research is still required to achieve early detection of movement in more realistic situations. For this reason, this work proposed the continuous detection of movement intention in self-paced natural reaching movements of the upper limbs. This experimental task was chosen because it resembles the common daily-live-activity of grasping an object such as a glass of water or a bottle. It is important to detect the movement intention with sufficient preceding time in order to be able to use this information piece on time to provide natural movement control to users in BMI-based motor recovery and motor rehabilitation scenarios. Therefore, this study addresses the detection of the intention to move irrespective of the moved limb within a continuous decoding strategy. Six healthy right-handed subjects participated in the experimental sessions. The results revealed significant event-related desynchronization before and during the execution of the reaching movement task, and these cortical rhythms were used as features to continuously detect the intention to move the limbs. In addition, significant classification rate of movement intention was achieved before the onset of the executed reaching movements.

This paper is organized as follows: the description of the experiment, the data processing and analysis, and the evaluation process and metrics are detailed in Section 2; Section 3 describes the results, in particular the significant activity of event-related desynchronization/synchronization and the classification results oriented to detect movement intention; finally Section 4 presents the conclusions and future work.

## 2. Methods

*2.1. Design and Execution of the Experiment.* The experiment consisted of self-paced natural reaching movements of the upper right/left limb. This experiment was selected because it resembles the common daily-live-activity of grasping an object such as a glass of water. Participants were comfortably seated with both forearms resting on the chair's arm and a computer screen was located in front of them. The

experiment consisted of the execution of many repetitions or trials of reaching movements with either left or right arm and was guided by visual cues presented on the screen. Each trial consisted of three cues. The first cue showed the text "relax" for three seconds and indicated staying comfortably seated with the arms placed on the chair's arms in complete relaxation. Participants were requested not to execute or imagine any movement. The second cue showed for twelve seconds an image with an "arrow" pointing to the left/right and indicated moving naturally the corresponding arm towards the center of the screen. Participants were instructed not to initiate the movement immediately after the arrow was presented but to initiate it whenever they wish, waiting for at least five seconds while avoiding any mental count. Accordingly, the movement initiation varies across trials. Immediately after the reaching movement was completed, participants were instructed to return back the arm towards the chair's arms. The third cue showed the text "rest" and indicated resting, moving, or blinking for three seconds. Therefore, each trial lasted for eighteen seconds in total. Figure 1 displays the full temporal sequence of a trial during the experiment. Participants were asked to avoid any movement and to minimize blinking from the presentation of the first cue and up to the termination of the reaching movement.

The experiment was executed in four blocks of 24 trials (7.2 min per block) resulting in a total of 96 trials (28.8 min for all blocks). To avoid fatigue, patients could rest between blocks as long as they needed. To keep balance of the number of trials for the left and right arm, each block contained the same number of left and right movements, which were presented in a pseudorandom manner. This experiment was approved by the ethics committee of the university.

*2.2. Participants.* Six able-bodied right-handed subjects (two males and four females; age range 23–19 years; mean  $\pm$  std 20.33  $\pm$  1.51 years) without diagnosis of neurological or motor disease voluntarily participated in this study. All participants were students from the university and did not have experience with electroencephalogram (EEG) recording protocols or brain-machine interface (BMI) experiments. They were duly informed about the objective of the research and the experimental procedure and all of them signed informed consent forms. They were informed that they could leave the experiment when they wanted.

*2.3. Recording of EEG and EMG Signals.* EEG signals were recorded using monopolar electrodes at 21 scalp positions according to the 10/10 international electrode location system. EEG signals were recorded from scalp locations *Fp1*, *Fp2*, *F7*, *F3*, *Fz*, *F4*, *F8*, *T3*, *C3*, *Cz*, *C4*, *T4*, *T5*, *P3*, *Pz*, *P4*, *T6*, *O1*, *O2*, *A1*, and *A2*, with the ground at *Fpz* and the reference at the left earlobe. EMG signals were recorded with bipolar electrodes located above the biceps brachii muscle and the triceps brachii muscle. These EMG signals were recorded from both arms and they were used to establish the time of the movement initiation of each trial. EEG and EMG data were recorded at a sampling frequency of 2048 Hz and no filtering was applied. The electrode impedance was

TABLE 1: Description of the state of the art of works reporting decoding of motor information preceding actual movements from electroencephalographic brain signals. ME: motor execution; MI: motor imagery; MRCP: movement related cortical potentials; LDA: linear discriminant analysis.

State-of-the-art movement intention detection from EEG							
	Participants	Motor task	EEG features	Classifier	Decoding	Accuracy	Detection time
[27]	8: healthy	ME/MI	Spectral power	Bayesian	Static	Varied between subjects	—
[28]	4: stroke	ME	Spectral power	LDA	Static	—	From $-200$ to $-600$ ms
[29]	6: healthy	ME	Spectral power	Bayesian	Static	$62.9 \pm 7.5\%$	—
[30]	19: healthy 5: stroke	ME/MI	MRCPs	Template matching	Continuous	Healthy: ME: $82.5 \pm 7.8\%$ , MI: $64.5 \pm 5.33\%$ ; stroke: ME: $55.01 \pm 12.01\%$	Healthy: ME: $-66.6 \pm 121$ ms
[31]	20: healthy 5: stroke	ME/MI	MRCPs	Template matching	Continuous	Healthy: ME: $69 \pm 21\%$ , MI: $65 \pm 22\%$ ; stroke: ME: $58 \pm 11\%$	—
[32]	10: healthy 2: stroke	ME/MI	MRCPs	LDA	Continuous	—	$-500$ ms
[33]	10: healthy 2: stroke	ME	MRCPs	LDA	Continuous	Healthy: $76\%$ ; stroke: $47\%$	$-312.5$ ms
[34]	20: healthy	ME	Common spatial patterns	LDA	Static	$83\%$ in 12 subjects	—
[35]	6: healthy 3: SCI	ME	Spectral power and MRCPs	SDA	Continuous	From $75\%$ to $40\%$	From $-421$ ms to $-256$ ms

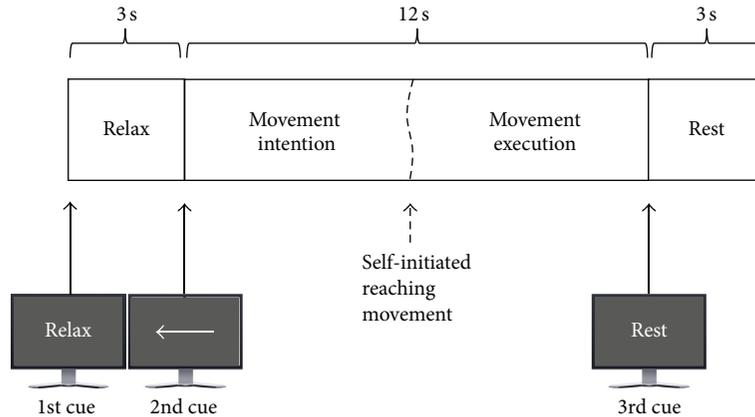


FIGURE 1: Temporal sequence of a trial with the three visual cues presented to the participants. Note that the second cue indicated self-initiating a reaching movement with the left arm.

kept below  $5\text{ k}\Omega$  for EEG and  $20\text{ k}\Omega$  for EMG. EEG and EMG signals were simultaneously recorded using a Nexus-32 electrophysiology monitoring system from Mind Media. BioTrace+ software was used to manage the presentation of the visual cues and the recording of the EEG and EMG signals and to store the data for offline processing.

2.4. *Data Preprocessing.* After the experimental sessions, recorded data were subjected to offline preprocessing and analysis. EEG and EMG data were trimmed from the presentation of the first cue up to the presentation of the third cue; thus, the resulting trials lasted for fifteen seconds. Then, the time latency of the movement onset of each trial was

computed with the EMG activity following this procedure: (i) the EMG signal from the moved arm was selected; (ii) this EMG signal was high-pass filtered with a cutoff frequency of 10 Hz using a sixth-order Butterworth-type infinite impulse response (IIR) filter; (iii) the Hilbert transform was computed from the filtered signal; (iv) the magnitude of the Hilbert transform was smoothed and normalized  $z$ -score; (v) the first value greater than zero in the resulting signal was defined as the EMG-based movement onset. Trials for which the movement onset was lower than 3 s (early arm movement initiation) and greater than 11 s (delayed arm movement initiation) relative to the presentation of the second visual cue were discharged. Then, the time axis of each trial was rereferenced to the EMG-based movement onset; that is,  $t = 0$  represents the initiation of the reaching movement. Finally, trials were trimmed from the initiation time  $t_{\text{ini}}$  (i.e.,  $t_{\text{ini}}$  is the time of the presentation of the first cue) up to 1 s relative to the EMG-based movement onset.

Frontal electrodes ( $Fp1$  and  $Fp2$ ), electrodes located near to the neap ( $O1$  and  $O2$ ), and other electrodes far away from the motor cortex ( $F7$ ,  $F8$ ,  $T3$ ,  $T4$ ,  $T5$ ,  $T6$ ,  $A1$ , and  $A2$ ) were removed from all participants as they are usually contaminated by eye blinks, muscle activity, and other artifacts; thus nine electrodes located on or surrounding the motor cortex ( $F3$ ,  $Fz$ ,  $F4$ ,  $C3$ ,  $Cz$ ,  $C4$ ,  $P3$ ,  $Pz$ , and  $P4$ ) were kept and used for the subsequent analysis. EEG data was resampled to 256 Hz, filtered from 0.1 Hz to 100 Hz using a zero-phase, four-order, bandpass Butterworth filter, and rereferenced using the common average reference (CAR) filter where the average across all channels is subtracted for each channel independently for each time sample.

**2.5. Event-Related Desynchronization/Synchronization.** To compute the significant event-related desynchronization/synchronization of each electrode, a bootstrap analysis of the time-frequency representation was performed. The goal in this analysis was to study the underlying task-related oscillatory brain activity during intention of motion [15]. All trials were trimmed from  $-6$  to  $1$  s relative to the EMG-based movement onset. This allows all trials to have the same length. For each trial and every channel, the time-frequency representation  $\text{TFR}(t, f)$  was computed in the frequency band  $[2, 40]$  Hz at the resolution of 1 Hz using Morlet wavelets [16]. For each channel individually, the event-related desynchronization/synchronization (i.e., power increase/decrease) relative to the baseline  $[-6, -3]$  s was computed for each time and frequency as  $\text{ERDS}(t, f) = 100 \times (\text{TFR}(t, f) - \text{TFR}_{\text{baseline}}(f)) / \text{TFR}_{\text{baseline}}(f)$ , where  $\text{TFR}_{\text{baseline}}(f)$  is the average of  $\text{TFR}(t, f)$  in the baseline interval for frequency  $f$ . The significant event-related desynchronization/synchronization of each channel was computed with a bootstrap analysis following [17] at the significant level of  $\alpha = 0.05$ .

**2.6. Detection of Movement Intention.** Detection of movement intention was based on spectral power features and on a support vector machine (SVM) used to distinguish between *relax* and *intention*.

**2.6.1. Features.** Spectral power features were computed with an autoregressive spectrum (ARS) model of order 16 [18, 19], where Burg's method was employed to estimate the model coefficients and the noise variance [20]. For each electrode, only values of the spectral power in the motor-related  $\alpha[8, 14]$  Hz and  $\beta[14, 25]$  Hz frequency bands were used. Spectral power values were computed at the resolution of 1 Hz. This resulted in 18 spectral power values per electrode. Thus, the feature vector is  $\mathbf{x} \in \mathbb{R}^D$  where  $D = 162$  (18 spectral power values  $\times$  9 electrodes), which is associated with a class label  $y \in \{\text{relax}, \text{intention}\}$ . For a given time instant  $t$  where  $t \in [t_{\text{ini}}, 1]$ , the spectral power features are computed from the EEG in the time window  $[t - T, t]$ , where  $T$  is the size of the window. Note that the time  $t$  corresponds to the endpoint of the used time window; therefore the computed features were causal.

**2.6.2. Classifier.** To discriminate between *relax* and *intention*, a SVM with a radial basis function (RBF) kernel [21] was employed. The implementation of the SVM relied on the LIBSVM library [22]. The hyperparameters of the RBF were  $C = 1$  for the regularization parameter and  $\sigma = 0.5$  for the width [23, 24]. To train this classifier, the features were extracted exclusively from the *relax* phase  $[t_{\text{ini}}, t_{\text{ini}} + 3]$  and the *movement intention* phase  $[-3, 0]$  and they were labeled as *relax* and *intention*, respectively. Figure 2(a) illustrates the segments of *relax* and *intention* used to extract features. Within these segments, the features were computed from nonoverlapping time windows of length  $T$ . In the *relax* phase features were computed at  $t_k = t_{\text{ini}} + 3 - kT$  for  $k = 0, 1, 2, \dots$  provided that  $t \in [t_{\text{ini}}, t_{\text{ini}} + 3]$ . In the *movement intention* phase features were computed at  $t_k = -kT$  for  $k = 0, 1, 2, \dots$  provided that  $t \in [-3, 0]$ . Prior to training, features were  $z$ -score normalized according to  $x_i = (x_i - \mu_i) / \sigma_i$ , ( $i = 1, 2, \dots, D$ ), where  $\mu_i$  and  $\sigma_i$  are the corresponding mean and standard deviation of the  $i$ th feature computed exclusively from training data.

**2.6.3. Evaluation Procedure and Metrics.** Detection of movement intention was assessed for each subject independently following this procedure:

- (i) Randomly select 80% of the trials as training set and use remaining 20% as test set.
- (ii) For each trial of the training set, extract the spectral power based features from the *relax* phase and the *movement intention* phase according to  $T$  and then train the classifier.
- (iii) Apply the classifier to each trial in the test set using sliding windows of size  $T$  in steps of 0.1 s (Figure 2(b) illustrates the process employed to perform classification in a test trial).
- (iv) Compute performance metrics using the entire test set.

The following metrics were considered: (i) classification accuracy (CA) (rate of correct classifications achieved within the *relax*  $[t_{\text{ini}}, t_{\text{ini}} + 3]$  and the *intention*  $[-3, 0]$  s phases),

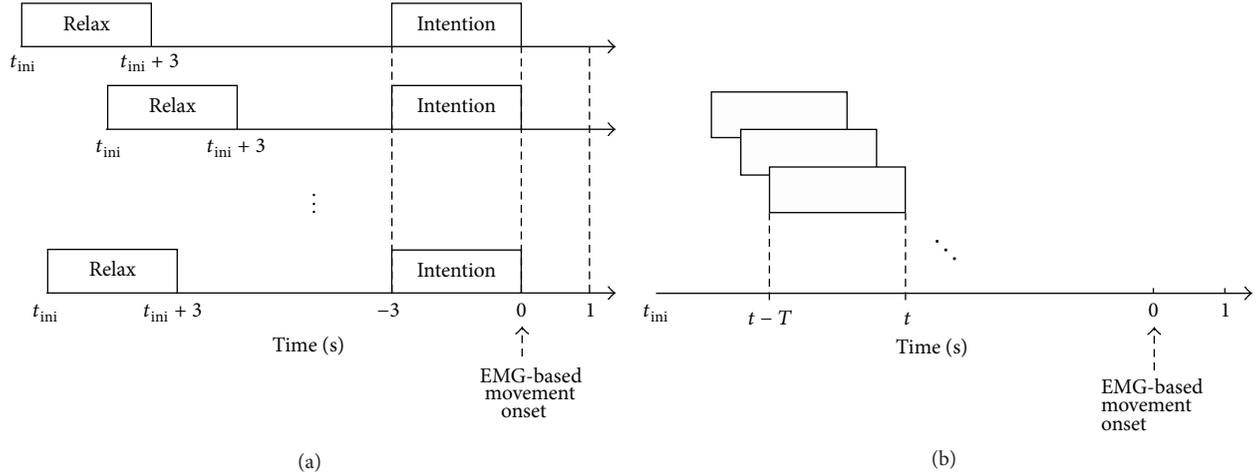


FIGURE 2: Illustration of the data segments used to train and to test the classification. (a) Data segments in the relax phase and the movement intention phase used to extract features to train the classifier. (b) Data segments used to carry out classification in a test trial.

true positive events (TPE) (movement intention detection rate obtained in the intention phase), and true negative events (TNE) (relax detection rate obtained in the relax phase); (ii) time-resolved movement intention detection accuracy or  $DA(t)$  (rate of movement intention detected at time  $t$ ); (iii) time instant of the movement intention onset or tMI (the lowest time instant for  $t < 0$  at which significant differences between  $DA(t)$  and the chance level are unequivocally achieved); and (iv) trials where movement intention is detected or  $NT_D$  (rate of trials where movement intention was unequivocally detected prior to movement initiation, i.e.,  $t < 0$ ).

This evaluation procedure was repeated 30 times and distributions and the mean  $\pm$  std of the performance metrics were computed for each participant and for all of them. The significant chance level of the detection accuracy or  $DA_{sig}(t)$  was computed empirically by randomly permuting the class labels during the training of the classifier. This procedure was conducted also 30 times for each subject using 80% of the trials for training (with random labels) and remaining 20% for evaluation. The significant chance level of the classification accuracy or  $CA_{sig}$  was computed as the maximum empirical chance level in the relax  $[t_{ini}, t_{ini} + 3]$  and intention  $[-3, 0]$  s phases. To examine significant differences between the distributions of CA and the significant chance level  $CA_{sig}$  the Wilcoxon signed rank test was used, while to examine significant differences between  $DA(t)$  and  $DA_{sig}(t)$  the Wilcoxon rank-sum test was employed. These statistical tests were performed at a confidence level of  $\alpha = 0.01$ .

### 3. Results

The time instant of the movement initiation computed with the EMG activity was estimated in all the trials of all subjects after the presentation of the second visual cue (i.e., the one instructed to self-initiate the reaching movement of the left/right arm). Movement onset was lower than 3 s in 3% of the trials while it was greater than 11 s in 1% of the trials.

TABLE 2: Summary of the estimated EEG-based movement onset for all subjects and the average across all of them. It only includes trials for which the movement onset was estimated from  $t \geq 3$  and  $t < 11$  with respect to the presentation of the second visual cue which instructed to self-initiate the reaching movement of the left/right arm.

	EMG-based movement onset (s)			
	Mean	Std	Min	Max
S1	6.19	1.37	3.18	9.35
S2	7.54	1.30	4.04	10.36
S3	7.52	1.39	4.27	10.82
S4	6.69	1.30	3.55	10.00
S5	6.93	1.29	3.77	9.66
S6	7.38	1.30	4.07	10.41
Avg	7.04	1.42	3.19	10.82

These trials were discharged and not used in the rest of the work. Then, the total number of trials across all subjects used in this study was on average  $92.83 \pm 2.79$  (minimum of 89 and maximum of 96). Table 2 shows a summary of the estimated EMG-based movement onset for all subjects and the average for all of them. The average movement onset across all subjects was  $7.04 \pm 1.42$  s (minimum of 3.18 s and a maximum 10.82 s).

The significant activity of event-related desynchronization/synchronization computed across all trials and subjects is presented in Figure 3. Significant desynchronization ( $p < 0.05$ ) is observed in all sensors and in the motor-related  $\alpha[8, 13]$  Hz and  $\beta[14, 30]$  Hz frequency bands around the movement onset  $t = 0$  s. This significant desynchronization starts in the movement intention phase roughly at 1 s prior to the movement onset and remains significant up to the movement execution interval  $t \geq 0$ . No significant desynchronization or synchronization ( $p > 0.05$ ) is observed before  $\approx -1$  s. Note that the significant desynchronization is uniformly

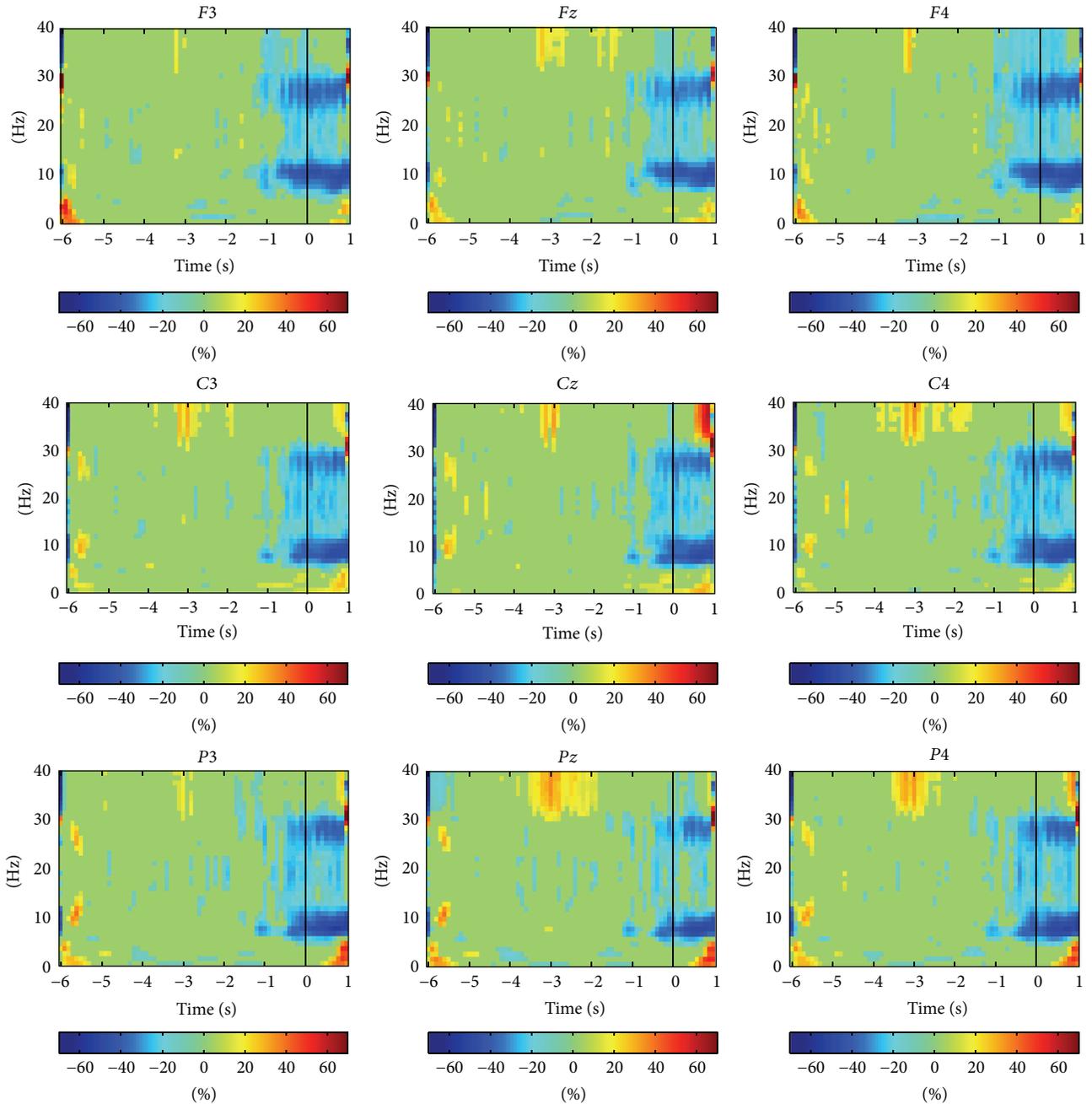


FIGURE 3: Significant event-related desynchronization/synchronization activity computed for all trials and subjects. Horizontal axis represents time (units of s) while vertical axis represents frequency (units of Hz). Solid black lines in all graphs represent  $t = 0$  s or the movement onset. Significant desynchronization ( $p < 0.05$ ) is observed in all sensors in the motor-related  $\alpha$ [8, 13] Hz and  $\beta$ [14, 30] Hz frequency bands from  $t \approx -1.5$  s, while no significant desynchronization or synchronization ( $p > 0.05$ ) is observed before  $t \approx -1.5$  s.

distributed in all sensors and in both hemispheres; that is, no spatial pattern of desynchronization/synchronization is observed across the motor cortex.

The average of significant event-related desynchronization/synchronization in the  $\alpha$ [8, 13] Hz and  $\beta$ [14, 30] Hz frequency bands of each electrode was computed over time windows along the entire duration of the trial. These results are presented in Table 3. In all electrodes, the significant

desynchronization is absent for time windows from  $[-6, -5)$  to  $[-4, -3)$  s but then begins to intensify gradually from  $[-3, -2)$  s up to  $[0, 1)$  s. Thus, the significant desynchronization starts prior to the movement onset, that is, at the movement intention phase, and remains significant up to the movement execution phase. Note that the event-related desynchronization/synchronization in the  $\alpha/\beta$  frequency bands averaged for all electrodes is  $-4.72/-3.84$ ,

TABLE 3: Across all subjects average of significant event-related desynchronization/synchronization in the motor-related  $\alpha$ [8, 13] Hz and  $\beta$ [14, 30] Hz frequency bands over several time windows from  $-6$  s to  $1$  s. Results are presented for each electrode and the last row presents the integrated values for all channels.

		Time window (s)						
		$[-6, -5)$	$[-5, -4)$	$[-4, -3)$	$[-3, -2)$	$[-2, -1)$	$[-1, 0)$	$[0, 1)$
$F3$	$\alpha$	2.12	0.00	0.00	-0.59	-8.09	-28.66	-40.04
	$\beta$	0.79	-2.23	0.26	-1.13	-0.83	-15.89	-18.23
$Fz$	$\alpha$	1.76	0.00	0.00	0.00	-3.99	-22.92	-37.74
	$\beta$	0.99	-1.05	0.20	-1.16	-1.34	-13.88	-16.39
$F4$	$\alpha$	2.07	-0.22	0.00	0.00	-5.82	-28.64	-41.14
	$\beta$	1.62	-0.70	-1.07	-0.17	-3.58	-16.18	-17.45
$C3$	$\alpha$	4.92	-0.39	0.00	0.00	-4.53	-23.90	-38.85
	$\beta$	0.31	-1.24	0.22	-1.50	-2.56	-14.74	-21.39
$Cz$	$\alpha$	4.43	-0.38	0.00	-0.37	-4.47	-23.44	-40.67
	$\beta$	1.36	0.10	0.00	-2.17	-3.97	-15.19	-19.10
$C4$	$\alpha$	3.36	-0.20	0.00	-0.77	-5.17	-26.23	-40.19
	$\beta$	0.60	0.42	0.00	-2.87	-5.01	-18.70	-20.82
$P3$	$\alpha$	4.70	-0.20	0.00	-0.44	-6.86	-21.79	-42.01
	$\beta$	1.01	-1.06	-1.14	-3.64	-2.41	-16.42	-16.79
$Pz$	$\alpha$	4.25	0.00	0.27	-0.21	-4.38	-16.06	-36.00
	$\beta$	0.57	-0.26	-2.00	-2.52	-1.71	-11.28	-9.26
$P4$	$\alpha$	4.09	0.00	0.00	0.00	-5.02	-21.56	-41.10
	$\beta$	0.73	0.04	-0.80	-0.94	-1.48	-16.26	-16.55
Avg	$\alpha$	4.24	-0.32	0.00	-0.38	-4.72	-24.53	-39.90
	$\beta$	0.76	-0.24	0.07	-2.18	-3.84	-16.21	-20.44

$-24.53/-16.21$ , and  $-39.90/-20.44$  for time windows  $[-2, -1)$ ,  $[-1, 0)$  and  $[0, 1)$  s, respectively. This shows that the significant desynchronization is more prominent during the movement execution phase than during the movement intention phase and that it is stronger in  $\alpha$ [8, 13] Hz than in the  $\beta$ [14, 30] Hz frequency band.

The first classification analysis explored the impact of the window size  $T$  used to compute the spectral power features to classify between relax and intention. Figure 4 shows, for each subject and for all of them, the distributions of the classification accuracy metric CA for window sizes of  $T = 0.5$ ,  $T = 0.75$ , and  $T = 1$  s and the significant chance level  $CA_{sig}$  (maximum chance level across all subjects achieved in the relax and movement intention phases). For subjects 1, 3, 4, and 6, the median of the distribution for all  $T$  is greater and significantly different than the chance level  $CA_{sig}$  ( $p < 0.01$ , Wilcoxon signed rank test). However, for subjects 2 and 5, no significant differences were found between the median of the distributions and the chance level  $CA_{sig}$  ( $p > 0.01$ , Wilcoxon signed rank test). The averages of CA for  $T = 0.5$  s were  $0.64 \pm 0.18$ ,  $0.55 \pm 0.16$ ,  $0.65 \pm 0.14$ ,  $0.65 \pm 0.17$ ,  $0.54 \pm 0.16$ , and  $0.68 \pm 0.14$ , for  $T = 0.75$  s were  $0.70 \pm 0.18$ ,  $0.57 \pm 0.19$ ,  $0.67 \pm 0.14$ ,  $0.66 \pm 0.18$ ,  $0.58 \pm 0.18$ , and  $0.68 \pm 0.14$ , and for  $T = 1$  s were  $0.72 \pm 0.19$ ,  $0.59 \pm 0.20$ ,  $0.70 \pm 0.16$ ,  $0.70 \pm 0.19$ ,  $0.59 \pm 0.20$ , and  $0.70 \pm 0.16$ , respectively, for subjects 1 to 6. The results across all subjects showed that the median of the distributions of CA for all  $T$  is also greater and significantly different than the chance level  $CA_{sig}$  ( $p < 0.01$ , Wilcoxon signed rank test). Table 4 summarizes the results of classification accuracy

TABLE 4: Summary of the classification accuracy, true positive events, and true negative events achieved for all subjects for different window sizes  $T = 0.5$ ,  $T = 0.75$ , and  $T = 1.0$  s.

	Classification accuracy				
	Mean	Min	Max	TPE	TNE
0.50	$0.62 \pm 0.06$	0.54	0.68	0.62	0.62
0.75	$0.64 \pm 0.05$	0.57	0.70	0.65	0.64
1.00	$0.67 \pm 0.06$	0.59	0.72	0.67	0.66

(CA), true positive events (TPE), and true negative events (TNE) obtained across all subjects for the three windows sizes. The averaged values of CA for  $T$  of 0.5, 0.75, and 1 were  $0.62 \pm 0.06$ ,  $0.64 \pm 0.05$ , and  $0.67 \pm 0.06$ , respectively, while TPE/TNE were 0.62/0.62, 0.65/0.64, and 0.67/0.66, respectively. These results show that the performance in the recognition between relax and intention increases as the time window size  $T$  increases. Therefore, a window size  $T = 1$  s was used in the rest of this work to study the detection of movement intention.

Figure 5 shows the time-resolved detection accuracy  $DA(t)$  and the significant chance level of the detection accuracy  $DA_{sig}(t)$ . Results are presented for each subject separately. In all subjects  $DA(t)$  is presented from  $t = -5$  s. This is due to the following: first, the trial's initiation time  $t_{ini}$  is different across all subjects and trials and the common initiation time across all of them is  $t = -6$  s and, second, the window size used to compute the causal features is  $T = 1$  s.

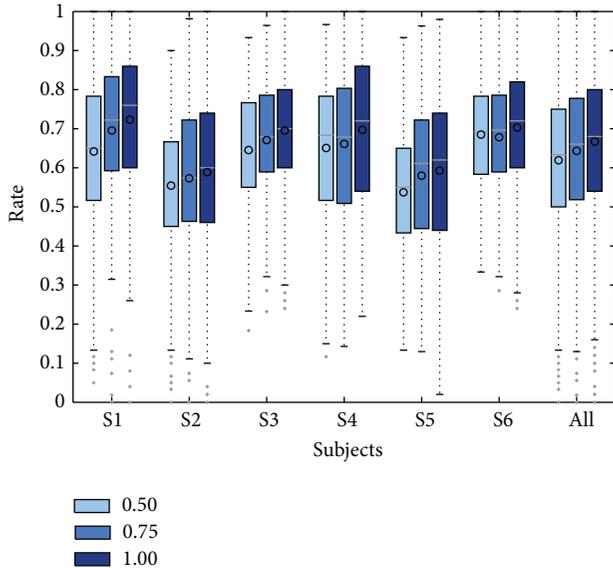


FIGURE 4: Distribution of the classification accuracy CA obtained for time windows  $T = 0.5, 0.75,$  and  $1$ . Results are clustered by subject and the last group of boxplots present the CA for all subjects. Horizontal line over each boxplot represents the median, meanwhile the circle represents the mean.

For all subjects (except number 5),  $DA(t)$  is initially at the chance level and starts to rise before the movement initiation at around  $t = -1$  s. In other words, no movement intention is detected from  $-6$  to  $\approx -1$  s while detection of movement intention is observed from  $\approx -1$  s. The maximum  $DA(t)$  is  $0.92, 0.73, 0.97, 0.86,$  and  $0.85$  for subjects 1 to 6, respectively, (excluding subject 5). These peaks of detection accuracy are achieved at  $t = 0.7, t = 0.9, t = 0.8,$  and  $t = 0.8,$  for subjects 1 to 4 and for subject 6 the maximum is reached in  $t = 0.2$  and  $t = 0.4$  (see vertical dotted blue lines in all plots of figures). Note that  $DA(t)$  always peaks at the movement execution phase  $t > 0$ . For subject 5,  $DA(t)$  is above chance level from  $-6$  to  $\approx 0$  s and suddenly drops at about  $t = 0$  s. This indicates that movement intention is always detected, even before the movement intention phase  $t < -3$  s (i.e., it is not possible to discriminate between movement intention and no movement intention) and that movement intention is at the chance level at the movement execution phase  $t > 0$  s. Thus, no movement intention information was detected for this participant. This result agrees with the distribution and average values of classification accuracy CA presented in Figure 4 for  $T = 1$  where subject 5 presented the lower performance.

The fraction of trials where movement intention was detected prior to movement initiation  $NT_D$  and the time of movement intention detection  $tMI$  are summarized in Table 5.  $tMI$  may also be observed in Figure 5. These metrics were not computed for subject 5 as no significant time-resolved detection accuracy  $DA(t)$  was achieved for this participant. On average, movement intention was detected in  $80 \pm 0.7\%$  of the trials across all subjects (minimum 69%,

TABLE 5: Rate of trials where movement intention was unequivocally detected ( $NT_D$ ) and the time instant of the movement intention detection ( $tMI$ ).

	Movement intention detection		
	$NT_D$	$1 - NT_D$	$tMI$
S1	0.84	0.16	-1.0
S2	0.69	0.31	-0.5
S3	0.87	0.13	-0.8
S4	0.79	0.21	-0.6
S5	—	—	—
S6	0.84	0.16	-1.0
Avg	$0.80 \pm 0.07$	0.20	$-0.78 \pm 0.23$

maximum 87%). These results reveal that movement intention is detected in the majority of the reaching movements performed by the participants. In addition, the time instant of movement intention detection is  $0.78 \pm 0.23$  s prior to movement initiation (minimum  $-0.5$  s, maximum  $-1.00$  s).

#### 4. Conclusions

This work proposed the continuous detection of movement intention from electroencephalographic (EEG) brain signals during natural self-paced reaching movements of the upper limbs. In the context of this work, movement intention was defined as the mental motor task (with no physical output) that occurs before the initiation of a movement, for example, motor planning. Six healthy subjects participated in this study and the EEG and electromyographic (EMG) activities were recorded.

The event-related synchronization/desynchronization of the EEG activity showed significant task-related cortical rhythms that started before the movement initiation, that is, at the movement intention phase, and remained during movement execution. Significant desynchronization ( $p < 0.05$ ) was observed in the motor-related  $\alpha[8, 13]$  Hz and  $\beta[14, 30]$  Hz frequency bands in all the selected sensors which were located above the motor cortex. This significant power decrease started about 1 s before the initiation of the movement and remains significant up to the movement execution, while no significant synchronization/desynchronization ( $p > 0.05$ ) was observed before this time. This significant desynchronization was detected in both hemispheres and was consistent with the experimental motor task which includes reaching movements with either the left or right arm.

These task related cortical rhythms were then used to investigate the feasibility of discriminating between the relax phase and the movement intention phase. Therefore, biclass classification between relax and intention was evaluated using the spectral power of the ongoing EEG activity and a support vector machine as classifier. This classification was evaluated using different window sizes  $T$  of EEG to compute the spectral power features. The results showed that classification accuracy between the relax and the intention phases increases as the time window size  $T$  increases. Thus, a window size

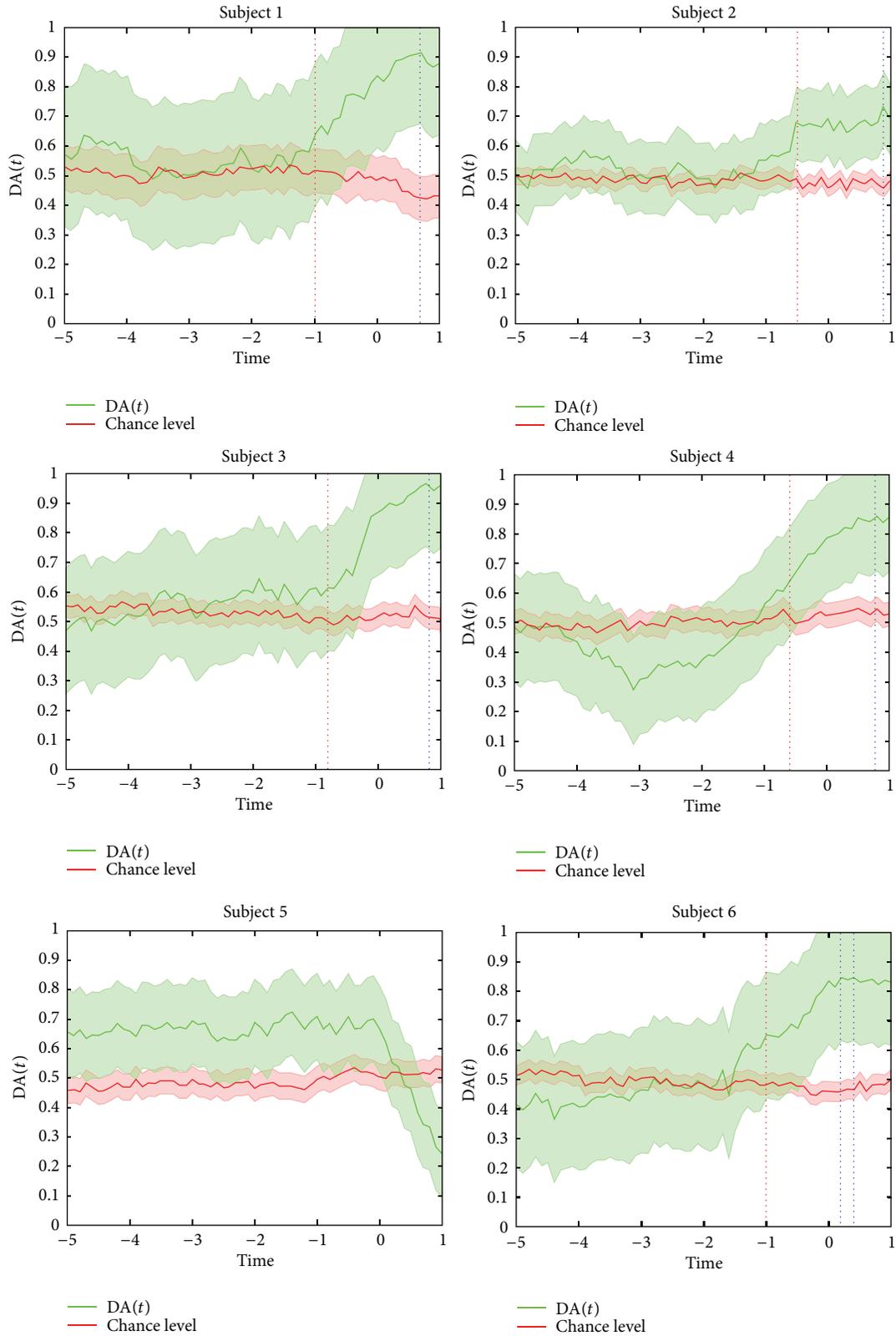


FIGURE 5: Time-resolved movement intention detection accuracy  $DA(t)$  (solid green line) and the empirical significant chance level of detection accuracy  $DA_{\text{sig}}(t)$  (solid red line) of each subject.  $t = 0$  refers to the initiation of the reaching movement. Shaded regions bounding the curves indicate the standard deviation. Vertical dotted blue lines represent the time of the maximum  $DA(t)$ . Vertical dotted red lines represent the tMI.

of  $T = 1$  s was selected to compute the spectral power features used to study the continuous detection of movement intention.

Finally, significant time-resolved detection accuracy was obtained in 5 out of the 6 participants before actual movement initiation. In one of the participants, it was not possible to distinguish between movement intention from relax. The significant detection of movement intention starts to rise at about 1 s before the onset of the movement and remained during the movement execution phase. The time-resolved detection accuracy reached the maximum during movement execution. This agrees with the observed significant desynchronization activity reported above. The initial time instant of movement intention was on average 0.78 s, that is, almost half a second before the actual movement, which was detected in 80% of the trials. The proposed detector of movement intention could be used in BMI-based robot-assisted rehabilitation scenarios. The advantage would be the reduction of the temporal delay between mental motor processes and the actual movement performed by the robotic devices. This could in principle provide fast, natural, and continuous motor control that enhances and promotes motor relearning at the cortical level.

The next steps for this research are (i) detection of the moved arm, (ii) determination of novel features based on the estimation of spikes, (iii) testing a novel classifier based on lattice neural networks with dendritic processing, and (iv) performing feature reduction and selection using Fast Correlation Based Filter (FCBF) [25, 26] or sequential forward selection (SFS).

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

The authors express their gratitude to CONACYT and COECYTJAL for the partial financial support through “Programa de Estímulos a la Innovación” and Project 3232-2015. The first author thanks CONACYT for the postdoctoral fellowship. And also, the authors thank Tecnológico de Monterrey, Campus Guadalajara.

## References

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, “Brain-computer interfaces for communication and control,” *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [2] A. Schlögl, C. Vidaurre, and K.-R. Müller, “Adaptive methods in BCI research—an introductory tutorial,” in *Brain-Computer Interfaces: Revolutionizing Human-Computer Interaction*, The Frontiers Collection, pp. 331–355, Springer, Berlin, Germany, 2010.
- [3] J. Becedas, “Brain-machine interfaces: basis and advances,” *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 42, no. 6, pp. 825–836, 2012.
- [4] B. Graimann, B. Z. Allison, and G. Pfurtscheller, *Brain-Computer Interfaces: Revolutionizing Human-Computer Interaction*, Springer, Berlin, Germany, 2010.
- [5] R. Leeb, L. Tonin, M. Rohm, L. Desideri, T. Carlson, and J. D. R. Millan, “Towards independence: a BCI telepresence robot for people with severe motor disabilities,” *Proceedings of the IEEE*, vol. 103, no. 6, pp. 969–982, 2015.
- [6] T. S. Rached, D. F. de S Santos, A. Perkusich, H. Almeida, and M. M. de Almeida Holanda, “BCI-aware pervasive multimedia for motor disabled people,” in *Proceedings of the International Conference on Information Society (i-Society '10)*, pp. 86–91, London, UK, June 2010.
- [7] J. D. R. Millán, R. Rupp, G. R. Müller-Putz et al., “Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges,” *Frontiers in Neuroscience*, vol. 4, article 161, 2010.
- [8] B. Z. Allison, S. Dunne, R. Leeb, J. D. R. Milln, and A. Nijholt, *Towards Practical Brain-Computer Interfaces: Bridging the Gap from Research to Real-World Applications*, Springer, Berlin, Germany, 2012.
- [9] R. Leeb, D. Friedman, G. R. Müller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller, “Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic,” *Computational Intelligence and Neuroscience*, vol. 2007, Article ID 79642, 8 pages, 2007.
- [10] N. Birbaumer and L. G. Cohen, “Brain-computer interfaces: communication and restoration of movement in paralysis,” *Journal of Physiology*, vol. 579, no. 3, pp. 621–636, 2007.
- [11] G. R. Müller-Putz, R. Scherer, G. Pfurtscheller, and R. Rupp, “EEG-based neuroprosthesis control: a step towards clinical practice,” *Neuroscience Letters*, vol. 382, no. 1-2, pp. 169–174, 2005.
- [12] G. Pfurtscheller and C. Neuper, “Motor imagery activates primary sensorimotor area in humans,” *Neuroscience Letters*, vol. 239, no. 2-3, pp. 65–68, 1997.
- [13] W. Wang, J. L. Collinger, M. A. Perez et al., “Neural interface technology for rehabilitation: exploiting and promoting neuroplasticity,” *Physical Medicine and Rehabilitation Clinics of North America*, vol. 21, no. 1, pp. 157–178, 2010.
- [14] M. Grosse-Wentrup, D. Mattia, and K. Oweiss, “Using brain-computer interfaces to induce neural plasticity and restore function,” *Journal of Neural Engineering*, vol. 8, no. 2, Article ID 025004, 2011.
- [15] G. Pfurtscheller and A. Aranibar, “Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement,” *Electroencephalography and Clinical Neurophysiology*, vol. 46, no. 2, pp. 138–146, 1979.
- [16] C. Tallon-Baudry, O. Bertrand, C. Delpuech, and J. Pernier, “Oscillatory  $\gamma$ -band (30–70 Hz) activity induced by a visual search task in humans,” *The Journal of Neuroscience*, vol. 17, no. 2, pp. 722–734, 1997.
- [17] B. Graimann and G. Pfurtscheller, “Quantification and visualization of event-related changes in oscillatory brain activity in the time frequency domain,” *Progress in Brain Research*, vol. 159, pp. 79–97, 2006.
- [18] D. J. McFarland and J. R. Wolpaw, “Sensorimotor rhythm-based brain-computer interface (BCI): model order selection for autoregressive spectral analysis,” *Journal of Neural Engineering*, vol. 5, no. 2, pp. 155–162, 2008.

- [19] D. J. McFarland, W. A. Sarnacki, and J. R. Wolpaw, "Electroencephalographic (EEG) control of three-dimensional movement," *Journal of Neural Engineering*, vol. 7, no. 3, Article ID 036007, 2010.
- [20] R. Bos, S. de Waele, and P. M. T. Broersen, "Autoregressive spectral estimation by application of the Burg algorithm to irregularly sampled data," *IEEE Transactions on Instrumentation and Measurement*, vol. 51, no. 6, pp. 1289–1294, 2002.
- [21] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [22] 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.
- [23] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain-computer interfaces," *Journal of Neural Engineering*, vol. 4, no. 2, pp. R1–R13, 2007.
- [24] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 141–144, 2003.
- [25] L. Yu and H. Liu, "Efficient feature selection via analysis of relevance and redundancy," *Journal of Machine Learning Research*, vol. 5, pp. 1205–1224, 2004.
- [26] U. M. Fayyad and K. B. Irani, "Multi-interval discretization of continuous-valued attributes for classification learning," in *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI '93)*, pp. 1022–1029, Chambéry, France, August 1993.
- [27] V. Morash, O. Bai, S. Furlani, P. Lin, and M. Hallett, "Classifying EEG signals preceding right hand, left hand, tongue, and right foot movements and motor imageries," *Clinical Neurophysiology*, vol. 119, no. 11, pp. 2570–2578, 2008.
- [28] A. Muralidharan, J. Chae, and D. M. Taylor, "Extracting attempted hand movements from eegs in people with complete hand paralysis following stroke," *Frontiers in Neuroscience*, vol. 5, article 39, 2011.
- [29] J. Ibáñez, J. I. Serrano, M. D. del Castillo, J. Minguez, and J. L. Pons, "Predictive classification of self-paced upper-limb analytical movements with EEG," *Medical and Biological Engineering & Computing*, vol. 53, no. 11, pp. 1201–1210, 2015.
- [30] I. K. Niazi, N. Jiang, O. Tiberghien, J. F. Nielsen, K. Dremstrup, and D. Farina, "Detection of movement intention from single-trial movement-related cortical potentials," *Journal of Neural Engineering*, vol. 8, no. 6, Article ID 066009, 2011.
- [31] I. K. Niazi, N. Jiang, M. Jochumsen, J. F. Nielsen, K. Dremstrup, and D. Farina, "Detection of movement-related cortical potentials based on subject-independent training," *Medical and Biological Engineering & Computing*, vol. 51, no. 5, pp. 507–512, 2013.
- [32] E. Lew, R. Chavarriaga, S. Silvoni, and J. D. R. Millán, "Detection of self-paced reaching movement intention from EEG signals," *Frontiers in Neuroengineering*, vol. 5, no. 13, 2012.
- [33] E. Y. Lew, R. Chavarriaga, S. Silvoni, and J. D. R. Millán, "Single trial prediction of self-paced reaching directions from EEG signals," *Frontiers in Neuroscience*, vol. 8, article 222, 2014.
- [34] M. Salvaris and P. Haggard, "Decoding intention at sensorimotor timescales," *PLoS ONE*, vol. 9, no. 2, Article ID e85100, 2014.
- [35] E. López-Larraz, L. Montesano, Á. Gil-Agudo, and J. Minguez, "Continuous decoding of movement intention of upper limb self-initiated analytic movements from pre-movement EEG correlates," *Journal of Neuroengineering and Rehabilitation*, vol. 11, no. 1, article 153, 2014.

## Review Article

# Frequency and Time Domain Analysis of Foetal Heart Rate Variability with Traditional Indexes: A Critical Survey

**Maria Romano,<sup>1</sup> Luigi Iuppariello,<sup>2</sup> Alfonso Maria Ponsiglione,<sup>2</sup> Giovanni Improta,<sup>3</sup> Paolo Bifulco,<sup>2</sup> and Mario Cesarelli<sup>2</sup>**

<sup>1</sup>*DMSC, University "Magna Graecia", Catanzaro, Italy*

<sup>2</sup>*DIETI, University of Naples "Federico II", Naples, Italy*

<sup>3</sup>*Department of Public Health, University of Naples "Federico II" Hospital, Naples, Italy*

Correspondence should be addressed to Mario Cesarelli; [cesarell@unina.it](mailto:cesarell@unina.it)

Received 22 December 2015; Revised 6 March 2016; Accepted 7 March 2016

Academic Editor: Krishna Agarwal

Copyright © 2016 Maria Romano 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.

Monitoring of foetal heart rate and its variability (FHRV) covers an important role in assessing health of foetus. Many analysis methods have been used to get quantitative measures of FHRV. FHRV has been studied in time and in frequency domain and interesting clinical results have been obtained. Nevertheless, a standardized definition of FHRV and a precise methodology to be used for its evaluation are lacking. We carried out a literature overview about both frequency domain analysis (FDA) and time domain analysis (TDA). Then, by using simulated FHR signals, we defined the methodology for FDA. Further, employing more than 400 real FHR signals, we analysed some of the most common indexes, Short Term Variability for TDA and power content of the spectrum bands and sympathovagal balance for FDA, and evaluated their ranges of values, which in many cases are a novelty. Finally, we verified the relationship between these indexes and two important parameters: week of gestation, indicator of foetal growth, and foetal state, classified as active or at rest. Our results indicate that, according to literature, it is necessary to standardize the procedure for FHRV evaluation and to consider week of gestation and foetal state before FHR analysis.

## 1. Introduction

Foetal heart rate (FHR) monitoring is of great importance to obtain information about foetal health during pregnancy and labour. In particular, electronic monitoring of the FHR (EFM), commonly named cardiotocography when an external Doppler probe is used, is the most employed method to detect foetal distress and prevent neurologic damage or even foetal death [1, 2]. However, despite its usefulness for obstetricians, the problem of EFM is its poor predictive value. It often lacks specificity leading to unnecessary interventions that increase caesarean delivery and operative vaginal delivery rates [1–4]. Moreover, there is often disagreement between obstetricians analysing FHR traces, since the interpretation is usually performed by means of a visual inspection, which obviously lacks objectivity and reproducibility [5, 6]. In recent years, hence, interest has grown in how to recognize changes in FHR that might predict more accurately foetal distress. For example, in order to overcome the subjective nature of FHR

interpretation, several attempts have been made to automate the diagnosis of the foetal status and many computerised algorithms have been developed to assess FHR parameters [1, 6–10].

Independently of the recording methods, the main FHR morphologic characteristics and parameters observed by physicians for foetal health evaluation are FHR mean value (which is related to week of gestation), baseline of the FHR, acceleration's rate and shape, deceleration's rate and shape, and FHR variability (FHRV) [11–15]. Among these, FHRV is probably the most important one, since it reflects the activity of autonomic nervous system (ANS) in the foetus who is growing and developing [1, 16–19], although the exact contributions of the two branches of the ANS are still object of investigation even in adult subjects [20]. The study of FHRV, often referred to as the beat-to-beat fluctuations of the FHR signal, could be, like for heart rate variability (HRV) of adult subjects, a base for a more powerful, detailed, and objective FHR analysis and for better knowledge of ANS reactions and

its development [11, 16, 21–24]. With respect to the first studies [25, 26], nowadays, the knowledge of FHRV is improved, so that different ranges of variability can be identified in order to classify FHR recordings [27] and its assessment employed to evaluate foetal reactivity and wellbeing in nonstress condition [28]. Even though the presence of good variability may not always be, by itself, a certain sign of reassuring FHR signal (corresponding to a well-oxygenated foetus), most clinicians agree that minimal or absent variability could be an indicator of foetal distress [28–30].

Due to its recognized importance, a large number of new analysis methods have been enforced to obtain more objective and quantitative measures of FHRV [31, 32]. Traditionally, as HRV of adult subjects, FHRV can be studied both in time (statistical indexes) and in frequency (spectral indexes) domain. In the time domain, Short Term Variability (STV) indexes and Long Term Variability (LTV) indexes are usually distinguished. In the frequency domain, different methods have been employed to estimate the power spectral density (PSD), which is widely considered the index that best covers all the information of the heart rate series. However, since FHRV signal shows a nonstationary behaviour, the time-frequency analysis of the FHRV is generally employed [1, 22, 33–39]. As far as time domain analysis (TDA) is concerned, it could present some limitations because it mainly relies on statistical measurements, so it can only describe the magnitude of the variability around an average value, without providing further information about the physiological mechanisms involved [40, 41]. Some limitations have been also shown for the frequency domain analysis (FDA) since it is generally sensitive to artifacts [42] and can provide information only about periodical fluctuations of the heart rate rhythm, without inspecting other possible nonperiodic trends embedded in the variability signal [40, 41]. In order to overcome these limitations and also to investigate and improve risk stratification, during the last decades, techniques analysing nonlinear dynamics, such as symbolic dynamics, approximate entropy, and fractal analysis, have been employed both in adults and in foetuses, even if, at the moment, none appears to be predominant and completely satisfying [26, 43, 44].

TDA and FDA hence remain the most used methods of HRV and FHRV analysis also due to their simplicity and higher acceptance in clinical environments. Nevertheless, a standardized definition of FHRV is still lacking and subject to changes and updates. Traditional analysis lack of standardized methods for computing time or frequency domain indexes and the relationship between FHRV and foetal growth is actually not completely clear and most of the studies inspecting this issue are by now rather old [52, 54, 71, 72].

This work aims to compare some common indexes of TDA and FDA, employing both real and artificial FHR signals.

To get this objective, we firstly introduce a brief report on the most relevant literature works about traditional FHRV analysis, focusing the attention on those studies based on the computation of the STV, as a time domain index, and on the power of the FHRV in different frequency bands.

Then, through the use of simulated FHR signals, we define the methodology to be employed for the estimation of the chosen indexes for FDA.

Finally, we apply the time and frequency indexes to the analysis of real FHR recordings in order to assess the capability of these indexes to correlate with foetal development during gestation and with foetal state and to provide an overview of their reference values.

## 2. Literature Report

*2.1. Spectral Analysis of FHRV: A Brief Literature Report.* The study of biologic signals in the frequency domain can offer deeper knowledge of their behaviour. In adults and foetuses, the spectral analysis permits estimating the power of the periodic HR fluctuations and it represents a noninvasive and powerful tool to understand ANS functional state and reactions [28].

After the study of Akselrod et al. [73], which introduced the power spectral analysis of short term heart rate fluctuations as a noninvasive quantitative probe of beat-to-beat cardiovascular autonomic control, FDA has been widely performed to point out the relation between the ANS activity and low frequency and high frequency bands, whose power content reflects changes of sympathetic and vagal activity [24, 26, 61]. Besides, changes in power distribution have been recognized as predictors of foetal distress, both in antepartum and in intrapartum periods [65, 74]. The work of Padhye et al. [19] investigated the correlation of power in LF (computed using the Lomb periodogram) and HF bands and noted an increasing trend of the power with gestational age. In their study, van Laar et al. [63] used fast Fourier transform to calculate the FHRV spectral power in LF and HF bands in order to compare spectral values between near term and postterm foetuses and found a sympathetic predominance in foetuses near term during active state, but an increased vagal modulation in postterm foetuses during rest state. According to Kwon et al. [4], changes in spectral power corresponding to a low pH are different between term and preterm foetuses, confirming a correlation between frequency indexes and gestational age.

Despite these interesting and important clinical results, problems in interpretation and comparison arise because literature works employ different frequency analysis methodologies, use discordant measure units for the PSD, and show disagreement about the frequency bands of the FHRV spectrum [36, 75], even if most of the literature agrees that, like the case for adult subjects, three bands can be detected in the FHR power spectrum: a very low frequency (VLF) band, which seems to be related to thermoregulation mechanisms [12, 67]; a low frequency (LF) band, which is mainly associated with the sympathetic branch activity and is an indicator of foetal development and wellbeing [54, 67]; a high frequency (HF) band, which reflects the respiratory activity and the vagal stimulation [38, 45, 47, 49, 61].

In order to clarify the definition of frequencies bands, a study of literature was conducted, involving about eight

TABLE I: Literature overview of FHRV spectral bands.

Reference	First author	Year	VLF <sub>l</sub>	VLF <sub>u</sub>	LF <sub>l</sub>	LF <sub>u</sub>	MF <sub>l</sub>	MF <sub>u</sub>	HF <sub>l</sub>	HF <sub>u</sub>
[45]	Divon	1985							0.70	0.95
[38]	Cerutti	1989	0.00	0.03	0.04	0.15			0.20	0.40
[46]	Ferrazzi	1989							0.70	0.90
[47]	Karin	1992							0.60	0.80
[48]	Metsälä	1993			0.025	0.069	0.07	0.129	0.13	1.00
[16]	Sibony	1994	0.02	0.05	0.05	0.15			0.15	0.50
[49]	Oppenheimer	1994			0.00	0.20			0.70	0.95
[50]	Groome	1994	0.00	0.04	0.04	0.20			0.20	2.50
[17]	Sibony	1995	0.02	0.04	0.04	0.16			0.16	0.30
[18]	Sibony	1995							0.20	0.50
[51]	Rassi	1995			0.07	0.12			0.56	1.10
[52]	Kimura	1996			0.00	0.30				
[53]	Moczko	1998	0.01	0.10					0.60	1.00
[54]	Ohta	1999			0.06	0.31			0.30	0.50
[55]	Rantonen	2000			0.03	0.07	0.07	0.13	0.13	1
[56]	Zhuravlev	2002		0.05	0.05	0.20			above 0.2	
[57]	Magenes	2002			0.04	0.15	0.15	0.50	0.50	1.00
[43]	Signorini	2003	0.00	0.30	0.03	0.15	0.15	0.50	0.50	1.00
[58]	Van Leeuwen	2003			0.04	0.15			0.15	0.40
[19]	Padhye	2004			0.05	0.25			0.25	1.00
[59]	Yum	2004			0.04	0.15			0.15	0.40
[60]	Siira	2005			0.04	0.15			0.15	1.00
[61]	David	2006	0.01	0.08	0.08	0.20			0.40	1.50
[62]	Tsoulos	2006	0.00	0.03	0.03	0.15	0.15	0.50	0.50	1.50
[36]	David	2007	0.02	0.08	0.08	0.20	0.20	0.40	0.40	1.70
[63]	van Laar	2009			0.04	0.15			0.40	1.50
[64]	Schneider	2009	0.02	0.08	0.08	0.20			0.40	1.70
[4]	Kwon	2012			0.04	0.15	0.15	0.50	0.50	1.00
[65]	Warrick	2012			0.03	0.15	0.15	0.50		
[66]	Reinhard	2012			0.04	0.15			0.15	0.40
[67]	Gonçalves	2013	0.00	0.03	0.03	0.15	0.15	0.50	0.50	1.00
[68]	Van Laar	2013			0.04	0.15			0.40	1.50
[69]	Van Leeuwen	2014			0.08	0.20			0.40	1.70

Results from the studied literature works (listed in chronological order) on FHRV spectral bands (l and u indicate, resp., the lower and upper limit of each band).

hundred literature works concerning foetal monitoring, published between 1983 and 2013. Among these, only a hundred works are directly related to frequency analysis and only about thirty works gave details about the three bands; the major disagreement is about the VLF band. Some researchers identifies the VLF band in the range from 0 to 0.03 Hz, while others consider VLF band ranging from 0 to 0.04 Hz or from 0 to 0.05 Hz [12, 50, 56, 67]. Furthermore, some authors introduce a middle frequency (MF) band in the range of 0.15–0.5 Hz or 0.2–0.4 Hz [36, 43].

A concise overview of different literature works focused on the computation of the FHRV spectral bands and corresponding power content is shown, respectively, in Tables I and 2. Empty cells are due to the absence of data in the original papers.

*2.2. STV in Foetal Monitoring: A Brief Literature Report.* As mentioned, in the time domain, Short Term Variability (STV)

indexes and Long Term Variability (LTV) indexes are usually distinguished. The former, also according to FIGO guidelines [76], refer to the continuous small changes in difference between successive interbeat intervals, which occur under physiological conditions. These minimal oscillations cannot be reliably interpreted by the naked eye; furthermore, a correspondent shared mathematical definition is lacking, so that this important parameter has lost part of its relevance and in some more recent guidelines it is not even considered but they speak broadly of variability, referring implicitly to the amplitude of FHR signal and without differentiating by LTV, since in practice they are visually determined as a unit [14, 77, 78].

When a computerised system is available, different indexes are used, many of which are borrowed from studies concerning adult heart rate. Among them there are Root Mean Square Successive Difference (RMSSD), that is, the

TABLE 2: Literature overview of FHRV power values.

Reference	First author	Year	VLF	LF	MF	HF	Measure unit	Foetal state
[38]	Cerutti	1989	27.30		9.30	44.20	%	Breathing
			69.20		25.30	2.60	%	Nonbreathing
			8.10		2.80	10.20	ms <sup>2</sup>	Breathing
			75.40	27.60	2.90		ms <sup>2</sup>	Nonbreathing
[50]	Groome	1994	35.6 ± 15.3	28.6 ± 10.7		35.8 ± 13.2	%	Breathing
			30.9 ± 11.6	28.6 ± 9.5		40.5 ± 13.9	%	Nonbreathing
			0.9 ± 0.67	0.62 ± 0.37		0.77 ± 0.29	ms <sup>2</sup>	Breathing
			0.4 ± 0.43	0.33 ± 0.24		0.42 ± 0.21	ms <sup>2</sup>	Nonbreathing
[43]	Signorini	2003		324 ± 174	28 ± 26		ms <sup>2</sup>	Active
				123 ± 95	16 ± 9		ms <sup>2</sup>	Quiet
			31.10	56.84	8.37	0.18	%	Active
			33.9 ± 15.7	48.3 ± 18.1	12.4 ± 5.6	1.27 ± 1.31	%	Quiet
[59]	Yum	2004		100.5 ± 6.3		15.5 ± 0.9	ms <sup>2</sup>	
[63]	van Laar	2009		0.8 ± 0.08		0.07 ± 0.03	%	Active
				0.69 ± 0.1		0.14 ± 0.06	%	Quiet
				429 ± 410		21.3 ± 7.3	ms <sup>2</sup>	Active
				92 ± 79.9		10.5 ± 5.3	ms <sup>2</sup>	Quiet
[70]	Ferrario	2009	274.82 ± 234.41	136.76 ± 84.21	19.13 ± 10.93	4.8 ± 3.61	ms <sup>2</sup>	
				83.82 ± 4.79	12.26 ± 2.51	3.93 ± 2.55	%	
[69]	Van Leeuwen	2014		45 ± 43		24 ± 12	ms <sup>2</sup>	

Results from the studied literature works (listed in chronological order) on FHRV power estimation.

square root of the mean squared differences of successive RR intervals, and pNN50, that is, the percentage of differences between following RR intervals greater than 50 ms [26, 79]. The LTV, instead, refer to fluctuations in the FHR over seconds, such as SDNN-Index, that is, the mean of the 5-minute standard deviation of the NN interval (normal to normal interval) calculated over 24 h, and SDANN, that is, the standard deviation of the average NN interval calculated over short periods [26].

In clinical practice, beat-to-beat indexes are often preferred since a good beat-to-beat variability is widely accepted as a significant index to assess foetal wellbeing, since a good beat-to-beat variability is a reliable indicator of a healthy foetal ANS [80]. Hence, many studies in time domain attempted to compute indexes for quantifying STV in fetuses by using very different techniques and methods (modification of the mean, standard deviation (SD), slope changes, and varying epoch lengths) [81]. The lack of a unique standardized methodology along with the fact that STV formulas, usually based on ECG, are often applied without any adaptation to the ultrasound technique makes the comprehension of the measure and the comparison between two or more indexes very difficult [79].

Despite the lack of standardization, STV is broadly employed. For example, Short Term Variability was found to be a good predictor of Apgar scores by Ayres-De-Campos et al. [82]. D'Elia et al. [83] analysed healthy term fetuses subjected to vibroacoustic stimulation by means of computerised CTG and found a statistically significant increase in foetal movements, acceleration rates, and STV with foetal activity. In their work, Serra et al. [84] examined the clinical value

of the STV in the timing of the delivery of severely growth-retarded fetuses and confirmed that the STV can assess the condition of fetuses with severe intrauterine growth restriction (IUGR) and that it is an important marker of perinatal outcome in severely growth-retarded fetuses. Also Galazios et al. [85] have rather recently observed that STV value is associated with foetal distress and, more recently, the study of Annunziata et al. [86] evaluated the impact of vibroacoustic stimulation on STV of CTG recordings in low and high risk pregnancies and noted that an increase in STV is significantly associated with good perinatal outcome.

Finally, Cesarelli et al. [79] proposed a comprehensive study on nine different mathematical indexes utilised to compute STV from CTG recordings, testing their robustness, sensitivity, and dependence on other parameters (FHR storage rate, FHR mean value, etc.) and demonstrated that the SD index, computed after floating line extraction, provides efficient information and is independent of the considered variables.

### 3. Methods

**3.1. Data Collection.** Real CTG traces were recorded by healthy pregnant women during the clinical practice, using commercially available cardiocographs (HP-135x or Sonicaid). Five hundred and eighty recordings, lasting on average more than 25 minutes, recorded from women between the 24th and the 42nd gestation weeks, were considered for the study. Gestational age was determined from the last menstruation date or from ultrasound measurements executed in the

first trimester of pregnancy. The database was completed with other pieces of clinical information of patients and newborns. All patients gave their informed consent to participate in the research concerning foetal monitoring.

CTG signals were processed by software previously developed by the authors [7, 84]. Firstly, they were preprocessed by means of a software [10, 84] which processes signals in output from the cardiocographs in order to recognize signal tracts having good and bad quality (these last including tracts of signal loss); for each segment of good quality, recover the real uneven FHR series when CTG output is evenly spaced (case of HP/Philips cardiocographs) [87] and detect and process outliers [88]; interpolate signal tracts of poor quality (according to an index provided by the equipment) or signal loss which last maximum 3 s, in order to avoid an excessive fragmentation of the signal.

In this way, all CTG signals available for further processing have the same characteristics and are unevenly sampled (in correspondence with the real heart beat) regardless of the equipment employed for their recording, hence regardless of the acquisition and sampling mode.

**3.2. CTG Simulation.** Like it is known, real FHR signals are affected by a considerable amount of variability and complexity, because of relationships, complicated and not yet fully known, among the different physiological mechanisms involved in heart rate regulation. Hence, in order to have available signals with characteristics known a priori, artificial CTG signals were employed for the analysis carried out in order to define an adequate methodology for FHRV evaluation.

According to previous works [79, 89, 90], an artificial uneven RR series with specific power spectrum characteristics, proper for FHR, was firstly generated. In particular, we set central frequencies of the spectral bands, the bandwidths, and the ratio between power of low and high frequency bands. Then, in line with the operations made by the cardiocographs which detect heart beats (by means of Doppler technique), compute interbeats distances series, and then reversing it get the FHR signal, this last one was computed as  $FHR = 60/RR$ , fixing FHR mean value and standard deviation of its peak-to-peak amplitude. After that, acceleration rates and deceleration rates with different amplitude and duration were simulated by using Gaussian-like signal tracts (for a more detailed explanation about how these synthetic signals were generated, please refer to previous works of the authors [79, 89, 90] and see Figure 1 for an example).

Our final simulated signals, hence, are in principle similar to those obtained with ultrasound technique [90].

The software for generating artificial CTG signals, as well as that for signal processing employed for his work, was developed using MATLAB R2011a [10].

**3.3. FHRV Estimation.** Previously, we proposed definition of FHRV as difference between FHR signal and its floating line [10, 12]. Of course, for FHRV assessment, the floating line has to be correctly estimated, so that we proposed also a procedure, using spline nonlinear filtering, developed to this

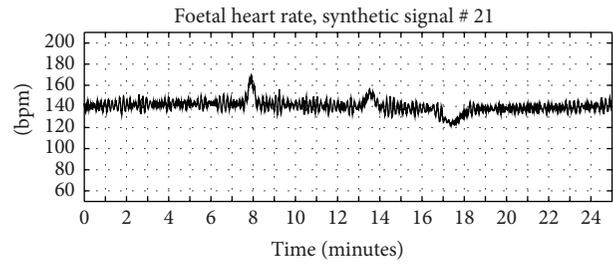


FIGURE 1: Synthetic FHR. Example of FHR artificially generated according to a procedure previously developed and published.

aim (the methodology has been recently updated, and results are submitted but not yet published).

Here, as further test, we compared the mean frequency spectrum computed on 30 simulated FHRV signals obtained with two different methodologies: by means of the application of our procedure for floating line estimation or simply as a result of the detrend operation, often employed in the literature [16, 55, 59, 63, 67]. (Let us remember that detrend is an operation which removes the linear trend from a signal; in Matlab it is a default function.)

**3.4. Frequency Domain Indexes.** We estimated FHRV as explained in Section 3.3 and considered for this signal LF (0.05–0.2 Hz) and HF bands (0.2–1 Hz); then, as VLF band, we computed the power spectral density (PSD) of the floating line.

Because of the nonstationarity of FHR and hence of FHRV, PSD was estimated by means of the Short Time Fourier Transformation [10], the methodology still more used for its simplicity, using a sliding Hamming window of 32 s [34, 47]. This window is shifted sample by sample and a new PSD is computed each time [35]. To be able to use the STFT, the FHR signal was previously interpolated (4 Hz sampling rate) by means of cubic interpolation, which has been demonstrated to reduce the error introduced [37]. Then, by means of a simple integral rule, we have computed the power values, absolute (called  $P_{LF}$ ,  $P_{HF}$ ), and percentage, with respect to the total power (called LF%, HF%), in the bands defined above. Further, the sympathovagal balance (SVB) index, which is an important index that reflects the relations between vagal and sympathetic branches of the ANS [26], was computed as ratio between  $P_{LF}$  and  $P_{HF}$ . Finally, using the same methodology, we computed  $P_{VLF}$  (and VLF%) as power of PSD of the floating line (using a commercial personal computer with a processor i5-3337U@1.8 GHz and RAM of 8 GB, the computation requires about 4 s for a CTG signal of more than 2 hours of length and about 10 s when also a 3D representation of the PSD is required).

**3.5. Time Domain Indexes.** As index of variability in time domain, we chose STV for its importance in monitoring of foetal health, since it is related to regulation mechanisms elicited by ANS activity [84, 86], as already mentioned in the previous sections. In agreement with a previous study of the authors, we assessed STV as standard deviation of the

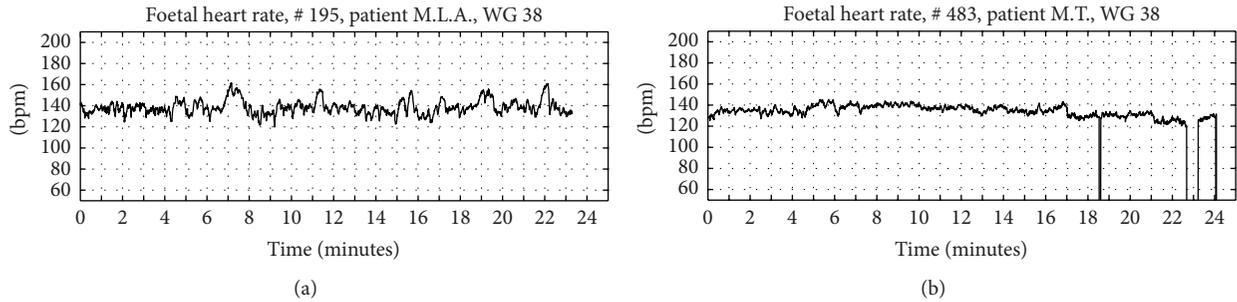


FIGURE 2: Foetal states. Example of FHR recorded from a foetus in an active state (a) and a foetus at rest (b) WG: week of gestation.

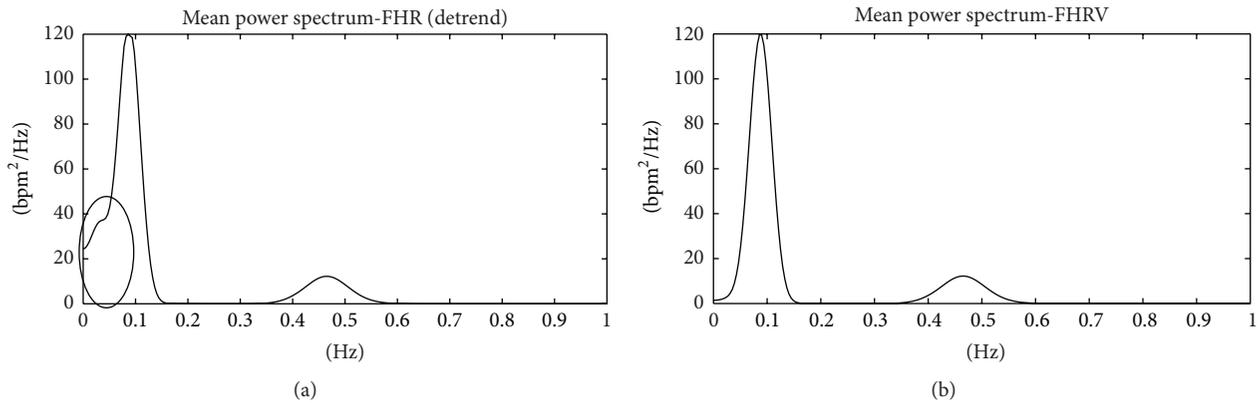


FIGURE 3: FHRV power spectrum. Mean power spectrum of FHRV computed using detrend (a) and mean power spectrum of FHRV computed after floating line subtraction (b). In the oval (a), the part of VLF band modified by a residual of VLF band.

FHRV, obtained after subtraction of the floating line from the FHR signal [79]. In order to enrich the available information, without worsening the computational complexity, the calculation is carried out on sliding windows of length  $M$  (with  $M$  covering 30 s), with an overlap of  $M - 1$  samples [10]; then, to provide an overall STV index of the signal, the mean of all STV values is computed (the computation and the graphical representation require about 8 s for a CTG signal of more than 2 hours of length, using the same personal computer utilised for FDA).

**3.6. Week of Gestation and Foetal State.** To test time and frequency indexes, we chose to verify their relationship with the week of gestation, the most simple indicator of foetal growth, and the foetal state.

It is known that foetal behavioural states have a great importance in influencing FHR patterns. However, the exact recognition of the foetal state is not a simple task, since respiratory acts, eyes closure and opening, just to name some aspect, should be simultaneously detected by ultrasound imaging. Hence, according to the literature [36, 63], for CTG recorded from 30th week of gestation onward, we defined active or resting foetal state. In particular, we classified a FHR signal as recorded from a foetus in active state if, at visual inspection, it showed a good variability (at least equal to 5 bpm) and at least two acceleration rates in 20 minutes of recording. If the signal showed low variability and absence of

acceleration rates, the foetus was classified as at rest. Doubt cases were excluded from the analysis.

In Figure 2, examples of FHR signals classified as active and rest are shown ((a) and (b), resp.).

**3.7. Statistical Analysis.** For all indexes studied, we analysed the correlation with week of gestation by means of regression lines (a polynomial quadratic curve was considered).

Besides, we tested the ability of these indexes to differentiate active from rest foetal state by means of the Mann-Whitney test, because of the non-Gaussian distribution of data.

## 4. Results

**4.1. FHRV Mean Spectrum.** In order to verify which is the more appropriate operation for FHRV estimation, detrend or floating line subtraction, in Figures 3(a) and 3(b), we show results of the frequency analysis conducted on simulated FHR signals.

It is clear that when the detrend operation is used (Figure 3(a)), a residual of VLF component appears superimposed on LF component. Of course, that can alter power computation and results of sympathovagal balance estimation.

**4.2. Foetal State.** In Table 3, the results of the Mann-Whitney test carried out for all tested indexes are shown; they concern

TABLE 3: Results of the Mann-Whitney test.

		Rest	Activity	Mann-Whitney $U$	$p$
		# CTG: 55	# CTG: 384		
		Median value			
DT	STV	1.74	2.80	1850	****
	$P_{VLF}$	11.3	44.7	1366	****
	$P_{LF}$	3.22	7.87	2123	****
	$P_{HF}$	0.41	0.99	2526	****
DF	$P_{tot}$	15.8	55.5	958	****
	SVB	8.20	8.54	11124	ns
	VLF%	76.5	83.4	7196	****
	LF%	20.2	14.4	7291	****
	HF%	2.56	1.73	7348	***

\*\*\* for  $p < 0.001$  that is statistically highly significant; \*\*\*\* for  $p \leq 0.0001$ ; ns for not significant.

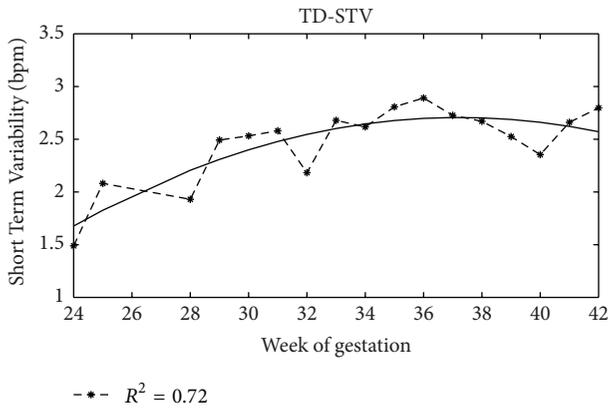


FIGURE 4: STV trend. Short Term Variability values and regression curve with week of gestation.

the comparison between foetuses at rest (55 CTG recordings) and active foetuses (384 CTG).

4.3. *FHRV Indexes.* Since all indexes, except SVB, resulted significantly differently in active state with respect to rest state, in Tables 4 and 5 we report ranges (minimum and maximum value) of all indexes here analysed and their mean and SD, computed on real FHR, separately for the two foetal states.

4.4. *Foetal Development.* In Figures 4–7, as examples, the trends with the pregnancy course of average values of some analysed indexes are depicted and in Table 6 values of coefficient of determination ( $R^2$ ) are shown for all indexes here analysed. In this case, the analysis is carried out without distinguishing foetal states that are not defined early in pregnancy.

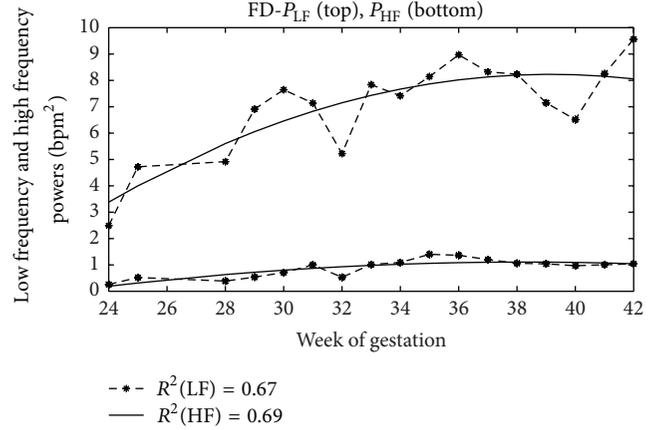


FIGURE 5:  $P_{LF}$  and  $P_{HF}$  trends. Power values in LF (on the top) and HF (on the bottom) bands and regression curves with week of gestation.

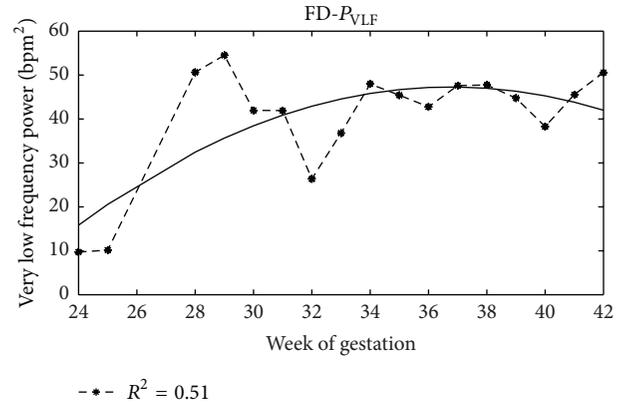


FIGURE 6:  $P_{VLF}$  trend. Power values of VLF band and regression curve with week of gestation.

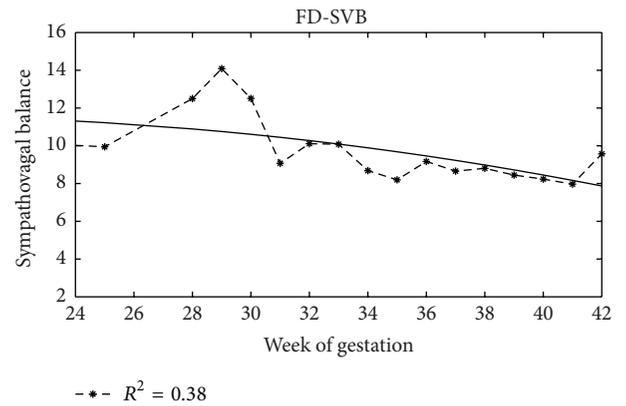


FIGURE 7: SVB trend. Sympathovagal balance values (dimensionless) and regression curve with week of gestation.

## 5. Conclusions

In this paper, firstly we provided an overview of the literature concerning the use of some traditional indexes. It is a concise overview but, at the best of our knowledge, it is the first

TABLE 4: Values ranges of all indexes.

	STV [bpm]	$P_{VLF}$ [bpm <sup>2</sup> ]	$P_{LF}$ [bpm <sup>2</sup> ]	$P_{HF}$ [bpm <sup>2</sup> ]	$P_{tot}$ [bpm <sup>2</sup> ]	SVB	VLF%	LF%	HF%
Mean	<b>1.80</b>	<b>14.72</b>	<b>3.57</b>	<b>0.47</b>	<b>18.76</b>	<b>8.37</b>	<b>74.04</b>	<b>22.40</b>	<b>3.23</b>
SD	<b>0.42</b>	<b>10.32</b>	<b>1.93</b>	<b>0.24</b>	<b>10.78</b>	<b>3.85</b>	<b>14.24</b>	<b>12.13</b>	<b>2.93</b>
Minimum	1.13	1.79	0.85	0.16	4.31	1.91	31.10	5.40	0.43
Maximum	2.88	46.46	9.49	1.44	50.89	19.21	94.18	64.87	17.02

Rest foetal state [55 CTG].

TABLE 5: Values ranges of all indexes.

	STV [bpm]	$P_{VLF}$ [bpm <sup>2</sup> ]	$P_{LF}$ [bpm <sup>2</sup> ]	$P_{HF}$ [bpm <sup>2</sup> ]	$P_{tot}$ [bpm <sup>2</sup> ]	SVB	VLF%	LF%	HF%
Mean	<b>2.90</b>	<b>52.39</b>	<b>9.04</b>	<b>1.24</b>	<b>62.50</b>	<b>8.53</b>	<b>81.31</b>	<b>16.19</b>	<b>2.41</b>
SD	<b>0.78</b>	<b>30.12</b>	<b>4.64</b>	<b>0.94</b>	<b>31.36</b>	<b>3.62</b>	<b>10.06</b>	<b>8.31</b>	<b>2.28</b>
Minimum	1.53	11.20	2.05	0.20	15.65	1.92	38.67	2.68	0.26
Maximum	5.77	179.64	27.09	5.61	189.59	20.64	97.06	47.88	16.42

Active foetal state [384 CTG].

TABLE 6: Values of the coefficients of determination ( $R^2$ ).

STV	$P_{VLF}$	$P_{LF}$	$P_{HF}$	$P_{tot}$	SVB	VLF%	LF%	HF%
0.72	0.51	0.67	0.69	0.57	0.38	0.12	0.16	0.03

time that so many quantitative indications and results are compared (Tables 1 and 2).

In the review section, we reported just some clinical results (being this aspect out of the main aim of the work and already treated in the literature [75]) and we neglected the choice of the methodology to compute PSD, since in a previous work we compared three methodologies and obtained not so different results. So, we focused on the definition of FHRV (since a FHRV definition shared and mathematically translatable is still missing) and frequency bands.

As here shown, about the frequency spectrum, although there is agreement in considering three main bands, VLF, LF and HF, and a MF when a more detailed analysis is necessary, even the same author or authors of the same research group employ different bands limits in their works (Table 1). Besides, frequency analysis is often carried out without providing power values and, in case, a large variability is present among different papers (Table 2), so that a comparison is quite difficult.

Furthermore, it is worth underlying that the analyses are very often carried out starting from RR intervals, maybe for historical reasons or for “continuity” with studies involving adult subjects. Nevertheless, ECG is not yet so diffused in clinical routine for difficulty in signal processing (in case of recording through maternal abdomen) or ethical reasons (in case of direct fECG), whereas clinicians are used to analyse FHR signal, measured in beats per minute (bpm).

For these reasons, we prefer to process FHR signal (without converting it into RR signal) and propose definition of FHRV as the difference between FHR and floating line. Preliminary results here shown (Figure 3) confirmed the usefulness of this methodology. About PSD computation, in

previous works [34, 37], we used the Lomb method which can be applied directly on uneven series; however, it has too long computational time by losing an advantage of the frequency analysis that can be used, if desired, also for real-time analysis; therefore, for successive as well as for this work we employed STFT. Furthermore, comparing different papers and taking into account experimental results obtained on our database, we decided to define for FHR signals the following bands limits: 0–0.05 Hz for VLF; 0.05–0.2 Hz for LF; and 0.2–1 Hz for HF.

Once defined details of the procedure are to be employed, through a retrospective study on a very large amount of CTG data recorded during physiological pregnancies (439 signals whereas most papers report results computed on a much smaller number of signals), we provided range of values of all frequency indexes here considered (differentiated on the basis of the foetal state, Tables 4 and 5). Some of them show high values of SD that put in evidence the high intersubject variability, even in healthy foetuses.

About STV index, the literature is even in more disagreement and we did not find reference values to report in our brief summary. However, we limited our analysis to a literature overview of the main results obtained in its clinical use since previously we already analysed different mathematical formula employed for its assessment and proposed a new evaluation methodology. By means of this methodology, we computed its value in healthy foetuses and for week of gestation. It, on average, resulted in increase with gestational age (Figure 4), according to the literature [91], and is very different in rest or active foetal state.

Then, we carried out a statistical analysis to test the correlation between the different indexes employed and the foetal growth (week of gestation).

Regression analysis showed that, on average, only some indexes can represent foetal growth in a satisfying way ( $R^2$  about 0.7, Table 6). Moreover, although a reliable comparison is rather difficult, because of the significant differences in employed methodologies and in a general lack of details provided, as proved here and in the literature [75], they are

substantially consistent with others. Absolute LF power and HF power, for example, (Figure 5), increase with gestational age but decrease their growth rate towards the end of pregnancy [58, 61]. With regard to the VLF, we do not have clear literature results with which we make a comparison; however, we may observe that the power in this band increases with week of gestation (Figure 6) and this is coherent with the increase of foetal movements and acceleration rates [61, 92] (let us remember that we computed  $P_{VLF}$  as power of the floating line which, in turn, includes acceleration rates and deceleration rates). Besides, its value represents the most of the total power so that the two indexes have an equivalent behaviour both with gestational age and with regard to foetal state. We can observe also that in the last weeks of gestation all powers change really little (both in absolute and percentage values), confirming the almost ended development of the SNA (further modifications will concern the adaptation to postnatal life).

The trend of SVB (Figure 7) is an exception but it is not a surprising result. Its value decreases with gestational age, coherently with the literature [36, 56, 61], and with the increase of HF band, related to development of the vagal branch.

Finally, about the foetal state, almost all indexes appeared to be able to separate the two groups (Table 3).

The findings we get in this work lead us to say, in accordance with the literature [63, 75], that standardization of FHRV assessment is necessary and that, since foetal state and gestational age can strongly affect results, it is not possible to process FHR signals regardless of these conditions.

## Competing Interests

The authors declare that they have no competing interests regarding the publication of this paper.

## Acknowledgments

This study was partially supported by the project “Sviluppo di Sistemi per il Monitoraggio Fetale Domiciliare” funded by “Legge 5” of Campania Region and by DRIVE IN2 project funded by the Italian National Program Piano Operativo Nazionale Ricerca e Competitività 2007/13.

## References

- [1] E. Salamalekis, P. Thomopoulos, D. Giannaris et al., “Computerised intrapartum diagnosis of fetal hypoxia based on fetal heart rate monitoring and fetal pulse oximetry recordings utilising wavelet analysis and neural networks,” *BJOG: An International Journal of Obstetrics & Gynaecology*, vol. 109, no. 10, pp. 1137–1142, 2002.
- [2] A. M. Vinitzleas, A. Antsaklis, I. Varvarigos, C. Papas, I. Sofatzis, and J. T. Montgomery, “A randomized trial of intrapartum electronic fetal monitoring versus intermittent auscultation,” *Obstetrics & Gynecology*, vol. 81, no. 6, pp. 899–907, 1993.
- [3] S. L. Clark and G. D. V. Hankins, “Temporal and demographic trends in cerebral palsy—fact and fiction,” *American Journal of Obstetrics and Gynecology*, vol. 188, no. 3, pp. 628–633, 2003.
- [4] J. Y. Kwon, I. Y. Park, J. C. Shin, J. Song, R. Tafreshi, and J. Lim, “Specific change in spectral power of fetal heart rate variability related to fetal acidemia during labor: comparison between preterm and term fetuses,” *Early Human Development*, vol. 88, no. 4, pp. 203–207, 2012.
- [5] R. Mantel, H. P. Van Geijn, I. A. P. Ververs, G. J. Colenbrander, and P. J. Kostense, “Automated analysis of antepartum fetal heart rate in relation to fetal rest-activity states: a longitudinal study of uncomplicated pregnancies using the Sonicaid System 8000,” *European Journal of Obstetrics Gynecology & Reproductive Biology*, vol. 71, no. 1, pp. 41–51, 1997.
- [6] J. Bernardes, A. Costa-Pereira, D. Ayres-De-Campos, H. P. Van Geijn, and L. Pereira-Leite, “Evaluation of interobserver agreement of cardiotocograms,” *International Journal of Gynecology and Obstetrics*, vol. 57, no. 1, pp. 33–37, 1997.
- [7] G. Improta, M. Romano, F. Amato, M. Sansone, and M. Cesarelli, “Development of a software for automatic analysis of CTG recordings,” in *Proceedings of the 3rd Congress Gruppo Nazionale Bioingegneria (GNB '12)*, Paper no. 69, Rome, Italy, 2012.
- [8] J. Pardey, M. Moulden, and C. W. G. Redman, “A computer system for the numerical analysis of nonstress tests,” *American Journal of Obstetrics and Gynecology*, vol. 186, no. 5, pp. 1095–1103, 2002.
- [9] R. M. Grivell, Z. Alfirevic, G. M. L. Gyte, and D. Devane, “Antenatal cardiotocography for fetal assessment,” *Cochrane Database of Systematic Reviews*, no. 1, Article ID CD007863, 2010.
- [10] M. Romano, P. Bifulco, M. Ruffo, G. Improta, F. Clemente, and M. Cesarelli, “Software for computerised analysis of cardiotocographic traces,” *Computer Methods and Programs in Biomedicine*, vol. 124, pp. 121–137, 2016.
- [11] M. Romano, P. Bifulco, M. Cesarelli, M. Sansone, and M. Bracale, “Foetal heart rate power spectrum response to uterine contraction,” *Medical and Biological Engineering & Computing*, vol. 44, no. 3, pp. 188–201, 2006.
- [12] M. Cesarelli, M. Romano, G. D’Addio et al., “Floatingline estimation in FHR signal analysis,” in *5th European IFMBE Conference, 14, Budapest, Hungary*, vol. 37 of *IFMBE Proceedings*, pp. 179–182, 2011.
- [13] A. Sweha, T. W. Hacker, and J. Nuovo, “Interpretation of the electronic fetal heart rate during labor,” *American Family Physician*, vol. 59, no. 9, pp. 2487–2500, 1999.
- [14] National Institute of Child Health and Human Development Research Planning Workshop, “Electronic fetal heart rate monitoring: research guidelines for interpretation,” *American Journal of Obstetrics & Gynecology*, vol. 177, no. 6, pp. 1385–1390, 1997.
- [15] Royal College of Obstetricians and Gynaecologists, *The Use of Electronic Fetal Monitoring. The Use and Interpretation of Cardiotocography in Intrapartum Fetal Surveillance*, vol. 8 of *Evidence-Based Clinical Guideline*, Royal College of Obstetricians and Gynaecologists, 2001.
- [16] O. Sibony, J.-P. Fouillot, M. Benaoudia et al., “Quantification of the fetal heart rate variability by spectral analysis of fetal well-being and fetal distress,” *European Journal of Obstetrics and Gynecology and Reproductive Biology*, vol. 54, no. 2, pp. 103–108, 1994.
- [17] O. Sibony, J. P. Fouillot, D. Luton, J. F. Oury, and P. Blot, “Effects of neuromuscular blockade on fetal heart rate variability: a power spectrum analysis,” *Journal of Applied Physiology*, vol. 79, no. 1, pp. 63–65, 1995.

- [18] O. Sibony, J.-P. Fouillot, M. Bennaoudia, D. Luton, P. Blot, and C. Sureau, "Spectral analysis of fetal heart rate in flat recordings," *Early Human Development*, vol. 41, no. 3, pp. 215–220, 1995.
- [19] N. S. Padhye, A. Brazdeikis, and M. T. Verklan, "Monitoring fetal development with magnetocardiography," in *Proceedings of the 26th Annual International Conference of the IEEE EMBS*, San Francisco, Calif, USA, 2004.
- [20] G. E. Billman, "Heart rate variability—a historical perspective," *Frontiers in Physiology*, vol. 2, article 86, 2011.
- [21] D. Hoyer, E. Heinicke, S. Jaekel et al., "Indices of fetal development derived from heart rate patterns," *Early Human Development*, vol. 85, no. 6, pp. 379–386, 2009.
- [22] M. Cesarelli, M. Romano, M. Ruffo, P. Bifulco, and G. Pasquariello, "Foetal heart rate variability frequency characteristics with respect to uterine contractions," *Journal of Biomedical Science and Engineering*, vol. 3, no. 10, pp. 1014–1021, 2010.
- [23] M. Romano, M. Cesarelli, P. Bifulco, M. Sansone, and M. Bracale, "Development of an algorithm for homogeneous FHR signals identification," in *Proceedings of the 2nd European Medical and Biological Engineering Conference (EMBEC '02)*, vol. 2, p. 1542, Vienna, Austria, December 2002.
- [24] C. M. A. Van Ravenswaaij-Arts, L. A. A. Kollee, J. C. W. Hopman, G. B. A. Stoelinga, and H. P. Van Geijn, "Heart rate variability," *Annals of Internal Medicine*, vol. 118, no. 6, pp. 436–447, 1993.
- [25] E. H. Hon and S. T. Lee, "Electronic evaluations of the fetal heart rate patterns preceding fetal death, further observations," *American Journal of Obstetrics & Gynecology*, vol. 87, pp. 814–826, 1965.
- [26] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, "Guidelines, heart rate variability, standards of measurement, physiological interpretation, and clinical use," *European Heart Journal*, vol. 17, pp. 354–381, 1996.
- [27] K. P. Williams and F. Galerneau, "Fetal heart rate parameters predictive of neonatal outcome in the presence of a prolonged deceleration," *The American College of Obstetricians and Gynecologists*, vol. 100, no. 5, part 1, pp. 951–954, 2002.
- [28] M. Romano, M. Bracale, M. Cesarelli et al., "Antepartum cardiotocography: a study of fetal reactivity in frequency domain," *Computers in Biology and Medicine*, vol. 36, no. 6, pp. 619–633, 2006.
- [29] A. M. Ponsiglione, M. Romano, G. Improta, P. Bifulco, D. G. Addio, and M. Cesarelli, "Symbolic dynamics analysis as a tool for assessing foetal heart rate variability," in *Congresso del Gruppo Nazionale di Bioingegneria (GNB '14)*, Paper M-36, 2014.
- [30] G. S. Dawes, M. Moulden, and C. W. G. Redman, "System 8000: computerized antenatal FHR analysis," *Journal of Perinatal Medicine*, vol. 19, no. 1-2, pp. 47–51, 1991.
- [31] H. Cao, D. E. Lake, J. E. Ferguson, C. A. Chisholm, M. P. Griffin, and J. R. Moorman, "Toward quantitative fetal heart rate monitoring," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 1, pp. 111–118, 2006.
- [32] D. Cysarz, P. Van Leeuwen, and H. Bettermann, "Irregularities and nonlinearities in fetal heart period time series in the course of pregnancy," *Herzschrittmachertherapie und Elektrophysiologie*, vol. 11, no. 3, pp. 179–183, 2000.
- [33] P. Laguna, G. B. Moody, and R. G. Mark, "Power spectral density of unevenly sampled data by least-square analysis: performance and application to heart rate signals," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 6, pp. 698–715, 1998.
- [34] M. Romano, M. Cesarelli, P. Bifulco, M. Ruffo, A. Fratini, and G. Pasquariello, "Time-frequency analysis of CTG signals," *Current Development in Theory and Applications of Wavelets*, vol. 3, no. 2, pp. 169–192, 2009.
- [35] J. L. A. Carvalho, A. F. Rocha, L. F. Junqueira, J. S. Neto, I. Santos, and F. A. O. Nascimento, "A tool for time-frequency analysis of heart rate variability," in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '03)*, September 2003.
- [36] M. David, M. Hirsch, J. Karin, E. Toledo, and S. Akselrod, "An estimate of fetal autonomic state by time-frequency analysis of fetal heart rate variability," *Journal of Applied Physiology*, vol. 102, no. 3, pp. 1057–1064, 2007.
- [37] M. Cesarelli, M. Romano, M. Ruffo, P. Bifulco, G. Pasquariello, and A. Fratini, "PSD modifications of FHRV due to interpolation and CTG storage rate," *Biomedical Signal Processing and Control*, vol. 6, no. 3, pp. 225–230, 2011.
- [38] S. Cerutti, S. Civardi, M. G. Signorini, E. Ferrazzi, G. Pardi, and A. Bianchi, "Spectral analysis of antepartum heart rate variability," *Clinical Physics and Physiological Measurement*, vol. 10, no. 4B, pp. 27–31, 1989.
- [39] K. K. Spyridou and L. J. Hadjileontiadis, "Analysis of fetal heart rate in healthy and pathological pregnancies using wavelet-based features," in *Proceedings of the 29th Annual International Conference of the IEEE EMBS. Cité Internationale*, Lyon, France, 2007.
- [40] A. Voss, J. Kurths, H. J. Kleiner et al., "The application of methods of non-linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death," *Cardiovascular Research*, vol. 31, no. 3, pp. 419–433, 1996.
- [41] H. V. Huikuri, T. H. Mäkitallio, and J. Perkiömäki, "Measurement of heart rate variability by methods based on nonlinear dynamics," *Journal of Electrocardiology*, vol. 36, pp. 95–99, 2003.
- [42] C. Peters, R. Vullings, J. Bergmans, G. Oei, and P. Wijn, "The effect of artifact correction on spectral estimates of heart rate variability," in *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '08)*, pp. 2669–2672, IEEE, August 2008.
- [43] M. G. Signorini, G. Magenes, S. Cerutti, and D. Arduini, "Linear and nonlinear parameters for the analysis of fetal heart rate signal from cardiotocographic recordings," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 3, pp. 365–374, 2003.
- [44] M. Cesarelli, M. Romano, P. Bifulco, G. Improta, DG. Addio, and M. Cesarelli, "An application of symbolic dynamics for FHRV assessment," *Studies in Health Technology and Informatics*, vol. 180, pp. 123–127, 2012.
- [45] M. Y. Divon, S.-Y. Yeh, E. Z. Zimmer, L. D. Platt, E. Paldi, and R. H. Paul, "Respiratory sinus arrhythmia in the human fetus," *American Journal of Obstetrics & Gynecology*, vol. 151, no. 4, pp. 425–428, 1985.
- [46] E. Ferrazzi, G. Pardi, P. Levi Setti, M. Rodolfi, S. Civardi, and S. Cerutti, "Power spectral analysis of the heart rate of the human fetus at 26 and 36 weeks of gestation," *Clinical Physics and Physiological Measurement B*, vol. 10, pp. 57–60, 1989.
- [47] J. Karin, M. Hirsch, C. Sagiv, and S. Akselrod, "Fetal autonomic nervous system activity monitoring by spectral analysis of heart rate variations," in *Proceedings of the Computers in Cardiology*, pp. 479–482, IEEE, Durham, NC, USA, October 1992.
- [48] T. H. Metsälä, J. P. Pirhonen, J. O. Jalonen, R. U. Erkkola, and I. A. T. Välimäki, "Association of abnormal flow velocity waveforms in the uterine artery with frequency-specific fetal

- heart rate variability,” *Early Human Development*, vol. 34, no. 3, pp. 217–225, 1993.
- [49] L. W. Oppenheimer and R. M. Lewinsky, “Power spectral analysis of fetal heart rate,” *Baillière’s Clinical Obstetrics and Gynaecology*, vol. 8, no. 3, pp. 643–661, 1994.
- [50] L. J. Groome, D. M. Mooney, L. S. Bentz, and K. P. Singh, “Spectral analysis of heart rate variability during quiet sleep in normal human fetuses between 36 and 40 weeks of gestation,” *Early Human Development*, vol. 38, no. 1, pp. 1–9, 1994.
- [51] D. Rassi and M. J. Lewis, “Power spectral analysis of the foetal magnetocardiogram,” *Physiological Measurement*, vol. 16, no. 2, pp. 111–120, 1995.
- [52] Y. Kimura, K. Okamura, and A. Yajima, “Spectral analysis of beat-to-beat intervals of the fetal heart obtained by Doppler ultrasound,” *Gynecologic and Obstetric Investigation*, vol. 41, no. 1, pp. 5–9, 1996.
- [53] J. Moczko, J. Jezewski, and A. Gacek, “Detection of fetal breathing movements with joint time-frequency analysis of cardiocogram records,” in *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 20, no. 3, pp. 1501–1504, Hong Kong, 1998.
- [54] T. Ohta, K. Okamura, Y. Kimura et al., “Alteration in the low-frequency domain in power spectral analysis of fetal heart beat fluctuations,” *Fetal Diagnosis and Therapy*, vol. 14, no. 2, pp. 92–97, 1999.
- [55] T. Rantonen, E. Ekholm, S. Siira et al., “Periodic spectral components of fetal heart rate variability reflect the changes in cord arterial base deficit values: a preliminary report,” *Early Human Development*, vol. 60, no. 3, pp. 233–238, 2001.
- [56] Y. E. Zhuravlev, D. Rassi, A. A. Mishin, and S. J. Emery, “Dynamic analysis of beat-to-beat fetal heart rate variability recorded by squid magnetometer: quantification of sympatho-vagal balance,” *Early Human Development*, vol. 66, no. 1, pp. 1–10, 2002.
- [57] G. Magenes, M. G. Signorini, C. Pizzicardi, D. Arduini, and S. Cerutti, “Fetal heart rate variability due to vibroacoustic stimulation: linear and nonlinear contribution,” in *Proceedings of the 4th International Workshop on Biosignal Interpretation (BSI ’02)*, p. 139, Como, Italy, 2002.
- [58] P. Van Leeuwen, D. Geue, S. Lange, W. Hatzmann, and D. Grönemeyer, “Changes in the frequency power spectrum of fetal heart rate in the course of pregnancy,” *Prenatal Diagnosis*, vol. 23, no. 11, pp. 909–916, 2003.
- [59] M.-K. Yum, C.-R. Kim, E.-Y. Park, and J.-H. Kim, “Instability and frequency-domain variability of heart rates in fetuses with or without growth restriction affected by severe preeclampsia,” *Physiological Measurement*, vol. 25, no. 5, pp. 1105–1113, 2004.
- [60] S. M. Siira, T. H. Ojala, T. J. Vahlberg et al., “Marked fetal acidosis and specific changes in power spectrum analysis of fetal heart rate variability recorded during the last hour of labour,” *BJOG: An International Journal of Obstetrics and Gynaecology*, vol. 112, no. 4, pp. 418–423, 2005.
- [61] M. David, M. Hirsch, and S. Akselrod, “Maturation of fetal cardiac autonomic control as expressed by Fetal Heart Rate variability,” in *Proceedings of the Computers in Cardiology*, pp. 901–904, IEEE, Valencia, Spain, September 2006.
- [62] J. Tsoulos, G. Georgoulas, D. Gavrilis, C. Stylios, J. Bemardes, and P. Groumpos, “Introducing grammatical evolution in fetal heart rate analysis and classification,” in *Proceedings of the 3rd International IEEE Conference on Intelligent Systems*, pp. 560–565, London, UK, September 2006.
- [63] J. O. E. H. van Laar, C. H. L. Peters, R. Vullings, S. Houterman, and S. G. Oei, “Power spectrum analysis of fetal heart rate variability at near term and post term gestation during active sleep and quiet sleep,” *Early Human Development*, vol. 85, no. 12, pp. 795–798, 2009.
- [64] U. Schneider, E. Schleussner, A. Fiedler et al., “Fetal heart rate variability reveals differential dynamics in the intrauterine development of the sympathetic and parasympathetic branches of the autonomic nervous system,” *Physiological Measurement*, vol. 30, no. 2, pp. 215–226, 2009.
- [65] P. A. Warrick and E. F. Hamilton, “Fetal heart-rate variability response to uterine contractions during labour and delivery,” *Computing in Cardiology*, vol. 39, pp. 471–420, 2012.
- [66] J. Reinhard, B. R. Hayes-Gill, S. Schiermeier et al., “Change of spectral analysis of fetal heart rate during clinical hypnosis: a prospective randomised trial from the 20th week of gestation till term,” *Geburtshilfe und Frauenheilkunde*, vol. 72, no. 4, pp. 316–321, 2012.
- [67] H. Gonçalves, A. Costa, D. Ayres-de-Campos, C. Costa-Santos, A. P. Rocha, and J. Bernardes, “Comparison of real beat-to-beat signals with commercially available 4 Hz sampling on the evaluation of foetal heart rate variability,” *Medical & Biological Engineering & Computing*, vol. 51, no. 6, pp. 665–676, 2013.
- [68] J. O. E. H. Van Laar, G. J. J. Warmerdam, K. M. J. Verdurmen et al., “Fetal heart rate variability during pregnancy, obtained from non-invasive electrocardiogram recordings,” *Acta Obstetrica et Gynecologica Scandinavica*, vol. 93, no. 1, pp. 93–101, 2014.
- [69] P. Van Leeuwen, L. Werner, Z. Hilal, S. Schiermeier, W. Hatzmann, and D. Grönemeyer, “Fetal electrocardiographic measurements in the assessment of fetal heart rate variability in the antepartum period,” *Physiological Measurement*, vol. 35, no. 3, pp. 441–454, 2014.
- [70] M. Ferrario, M. G. Signorini, and G. Magenes, “Complexity analysis of the fetal heart rate variability: early identification of severe intrauterine growth-restricted fetuses,” *Medical and Biological Engineering and Computing*, vol. 47, no. 9, pp. 911–919, 2009.
- [71] M. Romano, G. D’Addio, F. Clemente, A. M. Ponsiglione, G. Improta, and M. Cesarelli, “Symbolic dynamic and frequency analysis in foetal monitoring,” in *Proceedings of the 9th IEEE International Symposium on Medical Measurements and Applications (MeMeA ’14)*, pp. 1–5, IEEE, Lisboa, Portugal, June 2014.
- [72] G. Improta, G. Faiella, M. Romano, P. Bifulco, F. Clemente, and M. Cesarelli, “Power spectral analysis of foetal heart rate variability,” in *Proceedings of the Gruppo Nazionale Bioingegneria (GNB ’14)*, Paper M-16, Pavia, Italy, June 2014.
- [73] S. Akselrod, D. Gordon, F. A. Ubel, D. C. Shannon, A. C. Berger, and R. J. Cohen, “Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control,” *Science*, vol. 213, no. 4504, pp. 220–222, 1981.
- [74] D. Y. Chung, Y. B. Sim, K. T. Park, S. H. Yi, J. C. Shin, and S. P. Kim, “Spectral analysis of fetal heart rate variability as a predictor of intrapartum fetal distress,” *International Journal of Gynecology and Obstetrics*, vol. 73, no. 2, pp. 109–116, 2001.
- [75] J. O. E. H. Van Laar, M. M. Porath, C. H. L. Peters, and S. G. Oei, “Spectral analysis of fetal heart rate variability for fetal surveillance: review of the literature,” *Acta Obstetrica et Gynecologica Scandinavica*, vol. 87, no. 3, pp. 300–306, 2008.
- [76] FIGO News, “Guidelines for the use of fetal monitoring,” *International Federation of Gynaecology & Obstetrics*, vol. 25, pp. 159–167, 1987.

- [77] R. Liston, D. Sawchuck, and D. Young, "Fetal health surveillance: antepartum and intrapartum consensus guideline," *Journal of Obstetrics and Gynaecology Canada*, vol. 29, supplement 4, no. 9, p. 56, 2007.
- [78] NCC. National Certification Corporation, "NICHD definitions and classifications: application to electronic fetal monitoring interpretation," *NCC Monograph*, vol. 3, no. 1, 2010.
- [79] M. Cesarelli, M. Romano, and P. Bifulco, "Comparison of short term variability indexes in cardiotocographic foetal monitoring," *Computers in Biology and Medicine*, vol. 39, no. 2, pp. 106–118, 2009.
- [80] D. G. Chaffin, C. C. Goldberg, and K. L. Reed, "The dimension of chaos in the fetal heart rate," *American Journal of Obstetrics and Gynecology*, vol. 165, no. 5, pp. 1425–1429, 1991.
- [81] W. J. Parer, J. T. Parer, R. H. Holbrook, and B. S. B. Block, "Validity of mathematical methods of quantitating fetal heart rate variability," *American Journal of Obstetrics & Gynecology*, vol. 153, no. 4, pp. 402–409, 1985.
- [82] D. Ayres-De-Campos, C. Costa-Santos, and J. Bernardes, "Prediction of neonatal state by computer analysis of fetal heart rate tracings: the antepartum arm of the SisPorto® multicentre validation study," *European Journal of Obstetrics Gynecology & Reproductive Biology*, vol. 118, no. 1, pp. 52–60, 2005.
- [83] A. D'Elia, M. Pighetti, F. Vanacore, G. Fabbrocini, and L. Arpaia, "Vibroacoustic stimulation in normal term human pregnancy," *Early Human Development*, vol. 81, no. 5, pp. 449–453, 2005.
- [84] V. Serra, M. Moulden, J. Bellver, and C. W. G. Redman, "The value of the short-term fetal heart rate variation for timing the delivery of growth-retarded fetuses," *BJOG: An International Journal of Obstetrics & Gynaecology*, vol. 115, no. 9, pp. 1101–1107, 2008.
- [85] G. Galazios, G. Tripsianis, P. Tsikouras, N. Koutlaki, and V. Liberis, "Fetal distress evaluation using and analyzing the variables of antepartum computerized cardiotocography," *Archives of Gynecology and Obstetrics*, vol. 281, no. 2, pp. 229–233, 2010.
- [86] M. L. Annunziata, M. Scala, N. Giuliano et al., "Fetal vibroacoustic stimulation in computerized cardiotocographic analysis: the role of short-term variability and approximate entropy," *Journal of Pregnancy*, vol. 2012, Article ID 814987, 7 pages, 2012.
- [87] "HP Series 50 XM. Monitor Fetale/Materno," Informazioni Tecniche, 1995.
- [88] M. Romano, G. Faiella, P. Bifulco, G. D'Addio, F. Clemente, and M. Cesarelli, "Outliers detection and processing in CTG monitoring," in *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013: MEDICON 2013, 25–28 September 2013, Seville, Spain*, vol. 41 of *IFMBE Proceedings*, pp. 25–28, Springer, Berlin, Germany, 2014.
- [89] M. Cesarelli, M. Romano, P. Bifulco, F. Fedele, and M. Bracale, "An algorithm for the recovery of fetal heart rate series from CTG data," *Computers in Biology and Medicine*, vol. 37, no. 5, pp. 663–669, 2007.
- [90] G. Improta, M. Romano, and A. M. Ponsiglione, "Computerized cardiotocography: a software to generate synthetic signals," *Journal of Health & Medical Informatics*, vol. 5, no. 4, p. 162, 2014.
- [91] S. Lange, P. Van Leeuwen, D. Geue, W. Hatzmann, and D. Grönemeyer, "Influence of gestational age, heart rate, gender and time of day on fetal heart rate variability," *Medical and Biological Engineering and Computing*, vol. 43, no. 4, pp. 481–486, 2005.
- [92] G. S. Dawes, Y. J. Meir, and G. P. Mandruzzato, "Computerized evaluation of fetal heart-rate patterns," *Journal of Perinatal Medicine*, vol. 22, no. 6, pp. 491–499, 1994.

## Research Article

# Centered Kernel Alignment Enhancing Neural Network Pretraining for MRI-Based Dementia Diagnosis

David Cárdenas-Peña, Diego Collazos-Huertas, and German Castellanos-Dominguez

*Signal Processing and Recognition Group, Universidad Nacional de Colombia, Manizales, Colombia*

Correspondence should be addressed to David Cárdenas-Peña; [dcardenasp@unal.edu.co](mailto:dcardenasp@unal.edu.co)

Received 24 December 2015; Accepted 3 March 2016

Academic Editor: Dwarikanath Mahapatra

Copyright © 2016 David Cárdenas-Peña 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.

Dementia is a growing problem that affects elderly people worldwide. More accurate evaluation of dementia diagnosis can help during the medical examination. Several methods for computer-aided dementia diagnosis have been proposed using resonance imaging scans to discriminate between patients with Alzheimer's disease (AD) or mild cognitive impairment (MCI) and healthy controls (NC). Nonetheless, the computer-aided diagnosis is especially challenging because of the heterogeneous and intermediate nature of MCI. We address the automated dementia diagnosis by introducing a novel supervised pretraining approach that takes advantage of the artificial neural network (ANN) for complex classification tasks. The proposal initializes an ANN based on linear projections to achieve more discriminating spaces. Such projections are estimated by maximizing the centered kernel alignment criterion that assesses the affinity between the resonance imaging data kernel matrix and the label target matrix. As a result, the performed linear embedding allows accounting for features that contribute the most to the MCI class discrimination. We compare the supervised pretraining approach to two unsupervised initialization methods (autoencoders and Principal Component Analysis) and against the best four performing classification methods of the 2014 *CADDementia* challenge. As a result, our proposal outperforms all the baselines (7% of classification accuracy and area under the receiver-operating-characteristic curve) at the time it reduces the class biasing.

## 1. Introduction

In 2010, the number of people aged over 60 years with dementia was estimated at 35.6 million worldwide and this figure had been expected to double over the next two decades [1]. Actually, World Health Organization and the Alzheimer's Disease International had declared dementia as a public health priority, encouraging articulating government policies and promoting actions at international and national levels [2]. Alzheimer's disease (AD) is the most diagnosed dementia-related chronic illness that demands very expensive costs of care, living arrangements, and therapies. Thus, efforts are underway to improve treatment which may delay, at least, one year the AD onset and development, leading to decreasing the number of cases by nine millions [3]. AD can be early diagnosed by predicting the conversion to dementia from a state of mild cognitive impairment (MCI) that especially increases the AD risk [4].

In this regard, early diagnosis is directly related to the effectiveness of interventions [5]. Along with clinical history, neuropsychological tests, and laboratory assessment, the joint clinical diagnosis of AD also includes neuroimaging techniques like positron emission tomography (PET) and magnetic resonance imaging (MRI). These techniques are usually incorporated in the routine workup for excluding secondary pathology causes (e.g., tumors) [6, 7]. However, factors related to image quality and radiologist experience may limit their use [8]. For dealing with this issue, the imaging-based automatic assessment of quantitative biomarkers has been proven to enhance the performance for dementia diagnosis. In the particular case of AD, there are two groups of widely studied biomarkers: (i) patterns of brain amyloid-beta, such as low cerebrospinal fluid (CSF)  $A\beta_{42}$  and amyloid PET imaging, and (ii) measures of neuronal injury or degeneration like CSF tau measurement, fluorodeoxyglucose PET, and atrophy on structural MRI [9]. Thus, structural MRI

has become valuable for biomarker assessment since this noninvasive technique explains structural changes at the onset of cognitive impairment [10].

For the purpose of automated diagnosis, the first stage to implement is the structure-wise feature extraction from available MRI data, including voxel-based morphometry, volume, thickness, shape, and intensity relation. Nonetheless, more emphasis usually focuses on the classification approach due to its strong influence on the entire diagnosis system. With regard to neurodegenerative diseases, the reported classifiers range from straightforward approaches ( $k$ -Nearest Neighbors [11], Linear Discriminant Analysis [12], Support Vector Machines [13], Random Forests [14], and Regressions [15]) to the combination of classifiers [16]. Most of the above approaches had been evaluated for the *2014 CADDementia challenge* which aimed to reproduce the clinical diagnosis of 354 subjects in a multiclass classification problem of three diagnostic groups [17], Alzheimer’s diagnosed patients, subjects with MCI, and healthy controls (NC), given their T1-weighted MRI scans. As a result, the best-performing algorithm yielded an accuracy of 63.0% and an area under the receiver-operating-characteristic (ROC) curve of 78.8%. Nonetheless, reported true positive rates are 96.9% and 28.7% for NC and MCI, respectively, resulting in class biasing.

Generally speaking, dementia diagnosis from MRI still remains a challenging task, mainly, because of the nature of mild cognitive impairment; that is, there is a heterogeneous and intermediate category between the NC and AD diagnostic groups, from which subjects may convert to AD or return to the normal cognition [4]. For overcoming this shortcoming, machine learning tools as the artificial neural networks (ANN) have been developed to enhance dementia diagnosis, presenting the following advantages [18, 19]: (i) ability to process a large amount of data, (ii) reduced likelihood of overlooking relevant information, and (iii) reduction of diagnosis time.

Nonetheless, an essential procedure for ANN implementation is initializing deep architecture (termed pretraining) which can be carried out by training a deep network to optimize directly only the supervised objective of interest, starting from a set of randomly initialized parameters. However, this strategy performs poorly in practice [20]. With the aim to improve each initial-random guess, a local unsupervised criterion is considered to pretrain each layer stepwise, trying to produce a useful higher-level description based on the adjacent low-level representation output of the previous layer. Particular examples that use unsupervised learning are the following: Restricted Boltzmann Machines [21], autoencoders [22], sparse autoencoders [23], and the greedy layer-wise unsupervised learning which is the most common approach that learns one layer of a deep architecture at a time [24]. Although the unsupervised pretraining generates hidden representations that are more useful than the input space, many of the resulting features may be irrelevant for the discrimination task [25, 26].

In this paper, we benefit from the ANN advantages for complex classification tasks to introduce a novel supervised ANN initialization approach devoted to the automated dementia diagnosis. The proposed pretraining approach

searches for a linear projection into a more discriminating space so that the resulting embedding features and labels become as much as possible associated. Consequently, the obtained ANN architecture should match better the nature of supervised training data. Taking into account the fact that the ANN straightforward hybridization with other approaches yields stronger paradigms for solving complex and computationally expensive problems [27, 28], we also incorporate kernel theory for assessing the affinity between projected data and available labels. The use of kernel approaches offers an elegant, functional analysis framework for tasks, gathering multiple information sources (e.g., features and labels) as the minimum variance unbiased estimation of regression coefficients and least squares estimation of random variables [29]. Moreover, we consider the centered kernel alignment criterion as the affinity measure between a data kernel matrix and a target label matrix [30, 31]. As a result, the linear embedding allows accounting for features that contribute the most to the class discrimination.

The present paper is organized as follows: Section 2 firstly describes the mathematical background on learning projections using CKA and ANN for classification. Section 3 introduces all the carried out experiments for tuning the algorithm parameters and the evaluation scheme with blinded data. Then, achieved results are discussed in Section 4. Finally, Section 5 presents the concluding remarks and future research directions.

## 2. Materials and Methods

*2.1. Classification Using Artificial Neural Networks.* Within the classification framework, an  $L$ -layered ANN is assumed to predict the needed class label set through a battery of feedforward deterministic transformations, which are implemented by the hidden layers  $\mathbf{h}^l$ , which map the input space  $\mathbf{x}$  to the network output  $\mathbf{h}^L$  as follows [27]:

$$\begin{aligned} \mathbf{h}^l &= \phi(\mathbf{b}^l + \mathbf{W}^l \mathbf{h}^{l-1}), \quad \forall l = 1, \dots, L-1, \\ \mathbf{h}^0 &= \mathbf{x}, \end{aligned} \quad (1)$$

where  $\mathbf{b}^l \in \mathbb{R}^{m_l+1}$  is the  $l$ th offset vector,  $\mathbf{W}^l \in \mathbb{R}^{m_l+1 \times m_l}$  is the  $l$ th linear projection, and  $m_l \in \mathbb{Z}^+$  is the size of the  $l$ th layer. The function  $\phi(\cdot) \in \mathbb{R}$  applies saturating, nonlinear, element-wise operations. Here, we choose the standard sigmoid,  $\phi(z) = \text{sigmoid}(z)$ , expressed as follows:

$$\text{sigmoid}(z) = \frac{\tanh(z) + 1}{2}. \quad (2)$$

The first layer in (1) (i.e.,  $\mathbf{h}^0 \in \mathbb{R}^D$ ) is conventionally adjusted to the input feature vector. In turn, the output layer  $\mathbf{h}^L \in [0, 1]^C$  predicts the class when combined with a provided target  $t \in \{1, \dots, C\}$  into a loss function  $\mathcal{L}(\mathbf{h}^L, t)$ . In practice, the output layer can be carried out by the nonlinear softmax function described as follows:

$$h_c^L = \frac{\exp(b_c^L + \mathbf{w}_c^L \mathbf{h}^{L-1})}{\sum_j \exp(b_j^L + \mathbf{w}_j^L \mathbf{h}^{L-1})}, \quad (3)$$

where  $b_c^L$  is the  $c$ th element of  $\mathbf{b}^L$ ,  $\mathbf{w}_c^L$  is the  $c$ th row of  $\mathbf{W}^L$ ,  $\mathbf{h}^L$  is positive, and  $\sum_c h_c^L = 1$ .

The rationale behind the choice of softmax function is that each yielded output  $h_c^L$  can be used as an estimator of  $P(t_i = c | \mathbf{x}_i)$ , so that the interpretation of  $t_i$  relates to the class associated with input pattern  $\mathbf{x}_i$ . In this case, the softmax loss function corresponds often to the negative conditional log-likelihood:

$$\mathcal{L}(\mathbf{h}^L, t) = -\log \sum_c P(t = c | \mathbf{x}). \quad (4)$$

Therefore, the expected value over  $(\mathbf{x}, t)$  pairs is minimized with respect to the biases and weighting matrices.

**2.2. ANN Pretraining Using Centered Kernel Alignment.** Let  $\mathbf{X} \in \{\mathbf{x}_i \in \mathbb{R}^D : i \in N\}$  be the input feature matrix with size  $\mathbb{R}^{D \times N}$  which holds  $N$  trajectories and let  $\mathbf{x}_i \subset \mathcal{X}$  be a  $D$ -dimensional random process. In order to encode the affinity between a couple of trajectories,  $\{\mathbf{x}_i, \mathbf{x}_j\}$ , we determine the following kernel function:

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \langle \varphi(\mathbf{x}_i), \varphi(\mathbf{x}_j) \rangle, \quad \forall i, j \in N. \quad (5)$$

$\langle \cdot, \cdot \rangle$  stands for the inner product and  $\varphi(\cdot) : \mathbb{R}^D \rightarrow \mathcal{H}$  maps from the original domain,  $\mathbb{R}^D$ , into a Reproduced Kernel Hilbert Space (RKHS),  $\mathcal{H}$ . As a rule, it holds that  $|\mathcal{H}| \rightarrow \infty$ , so that  $|\mathbb{R}^D| \ll |\mathcal{H}|$  can be assumed. Nevertheless, there is no need for computing  $\varphi(\cdot)$  directly. Instead, the well-known *kernel trick* is employed for computing (5) through the positive definite and infinitely divisible kernel function as follows:

$$k_{ij} = \kappa(d(\mathbf{x}_i, \mathbf{x}_j)), \quad (6)$$

where  $d : \mathbb{R}^D \times \mathbb{R}^D \mapsto \mathbb{R}^+$  is a distance operator implementing the positive definite kernel function  $\kappa(\cdot)$ . A kernel matrix  $\mathbf{K} \in \mathbb{R}^{N \times N}$  that results from the application of  $\kappa$  over each sample pair in  $\mathbf{X}$  is assumed as the covariance estimator of the random process  $\mathcal{X}$  over the RKHS.

With the purpose of improving the system performance in terms of learning speed and classification accuracy, we introduce the prior label knowledge into the initialization process. Thus, we compute the pairwise relations between the feature vectors through the introduced feature similarity kernel matrix  $\mathbf{K} \in \mathbb{R}^{N \times N}$  which has elements as follows:

$$k_{ij} = \kappa_{\mathbf{x}}(d_W(\mathbf{x}_i, \mathbf{x}_j)), \quad \forall i, j \in \{1, \dots, N\}, \quad (7)$$

with  $d_W : \mathbb{R}^D \times \mathbb{R}^D \mapsto \mathbb{R}^+$  being a distance operator that implements the positive definite kernel function  $\kappa_{\mathbf{x}}(\cdot)$ , and  $\{(\mathbf{x}_i, t_i) : i = 1, \dots, N\}$  is a set of input-label pairs with  $\mathbf{x}_i \in \mathbb{R}^D$  and  $t_i \in \{1, C\}$ , with  $C$  being the number of classes to identify.

Since we look for a suitable weighting matrix for initializing the ANN optimization, we rely on the Mahalanobis distance that is defined on a  $D$ -dimensional space by the following inverse covariance matrix  $\mathbf{W}^T \mathbf{W}$ :

$$d_W(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{W}^T \mathbf{W} (\mathbf{x}_i - \mathbf{x}_j), \quad (8)$$

where matrix  $\mathbf{W} \in \mathbb{R}^{m_1 \times D}$  holds the linear projection  $\mathbf{y}_i = \mathbf{W} \mathbf{x}_i$ , with  $\mathbf{y}_i \in \mathbb{R}^{m_1}$ ,  $m_1 \leq D$ .

Based on the already estimated feature similarities, we propose further to learn the matrix  $\mathbf{W}$  by adding the prior knowledge about the feasible sample membership (e.g., healthy or diseased groups) enclosed in a matrix  $\mathbf{B} \in \mathbb{R}^{N \times N}$  with elements  $b_{ij} = \delta(t_i - t_j)$ . Thus, we measure the similarity between the matrices  $\mathbf{K}$  and  $\mathbf{B}$  through the following function of centered kernel alignment (CKA) [32]:

$$\rho(\mathbf{K}, \mathbf{B}) = \frac{\langle \mathbf{H} \mathbf{K} \mathbf{H}, \mathbf{H} \mathbf{B} \mathbf{H} \rangle_F}{\|\mathbf{H} \mathbf{K} \mathbf{H}\|_F \|\mathbf{H} \mathbf{B} \mathbf{H}\|_F}, \quad \rho \in [0, 1], \quad (9)$$

where  $\mathbf{H} = \mathbf{I} - N^{-1} \mathbf{1} \mathbf{1}^T$ , with  $\mathbf{H} \in \mathbb{R}^{N \times N}$ , is a centering matrix,  $\mathbf{1} \in \mathbb{R}^N$  is an all-ones vector, and  $\langle \cdot, \cdot \rangle_F$  and  $\|\cdot\|_F$  stand for the Frobenius inner product and norm, respectively.

Therefore, the centered version of the alignment coefficient leads to better correlation estimation compared to its uncentered version [31]. Therefore, the CKA cost function, described in (9), highlights relevant features by learning the matrix  $\mathbf{W}$  that best matches all relations between the resulting feature vectors and provided target classes. Consequently, we state the following optimization problem to compute the projection matrix:

$$\mathbf{W}^* = \arg \max_{\mathbf{W}} \rho(\mathbf{K}_{\mathbf{W}}, \mathbf{B}), \quad (10)$$

and we thus initialize the first layer of the ANN with  $\mathbf{W}^*$ .

Additionally, the weighting matrix allows analyzing the contribution of the input feature set for building the projection matrix by computing the feature relevance vector  $\boldsymbol{\rho} \in \mathbb{R}^D$  in the following form:

$$\rho_d = \mathcal{E} \{w_{ud}^2 : \forall u \in [1, m_1]\}, \quad (11)$$

where  $w_{ud} \in \mathbb{R}$  is the weight that associates each  $d$ th feature to  $u$ th hidden neuron.  $\mathcal{E}\{\cdot\}$  stands for the averaging operator. The main assumption behind the introduced relevance in (11) is that the larger the values of  $\rho_d$  the larger the dependency of the estimated embedding on the input attribute.

### 3. Experimental Setup

An automated, computer-aided diagnosis system based on artificial neural networks is introduced to classify structural magnetic resonance imaging (MRI) scans in accordance with the following three neurological classes: normal control (NC), mild cognitive impairment (MCI), and Alzheimer's disease (AD). Figure 1 illustrates the methodological development of the proposed approach.

**3.1. ADNI Data.** Data used in the preparation of this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu/>) which was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies, and nonprofit

TABLE 1: Demographic and clinical details of the selected ADNI cohort.

	“best” quality			“partial” quality		
	NC	MCI	AD	NC	MCI	AD
$N$	655	825	513	465	130	34
Age	$74.9 \pm 5.0$	$74.4 \pm 7.4$	$74.0 \pm 7.4$	$76.6 \pm 6.4$	$76.0 \pm 6.3$	$74.3 \pm 6.5$
Male	47.5%	39.5%	47.6%	70.1%	62.3%	70.6%
MMSE	$29.0 \pm 1.0$	$27.1 \pm 2.5$	$21.9 \pm 4.4$	$27.5 \pm 2.0$	$21.2 \pm 1.6$	$14.4 \pm 2.8$

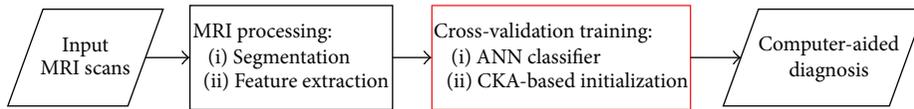


FIGURE 1: General processing pipeline: FreeSurfer independently segments and extracts features from given MRIs. Centered kernel alignment is proposed to learn a projection matrix initializing the NN training in a 5-fold cross-validation scheme. Tuned model is used for classification task.

organizations. The primary goal of ADNI is to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer’s disease (AD). From the ADNI 1, ADNI 2, and ADNI GO phases, we selected a subset of 633 subjects with scans that had been noted with the “best” quality mark. As a result, the selected subset holds  $N = 1993$  images with three class labels described above;  $C = 3$ . Besides, a random subset of 70% data was chosen for tuning and training stages, while the remaining 30% is for the test purpose. In addition, 629 images with a “partial” quality mark were selected in order to assess the performance under more complicated imaging conditions. Table 1 briefly describes the demographic information for the ADNI selected cohort.

**3.2. Processing of MRI Data.** We used FreeSurfer, version 5.1 (a free available (<http://surfer.nmr.mgh.harvard.edu/>), widely used and extensively validated brain MRI analysis software package), to process the structural brain MRI scans and compute the morphological measurements [33]. FreeSurfer morphometric procedures have been demonstrated to show good test-retest reliability across scanner manufacturers and across field strengths [34]. The FreeSurfer pipeline is fully automatic and includes the next procedures: a watershed-based skull stripping [35], a transformation to the Talairach, an intensity normalization and bias field correction [36], tessellation of the gray/white matter boundary, topology correction [37], and a surface deformation [38]. Consequently, a representation of the cortical surface between white and gray matters, of the pial surface, and segmentation of white matter from the rest of the brain are obtained. FreeSurfer computes structure-specific volume, area, and thickness measurements. Cortical Volumes and Subcortical Volumes are normalized to each subject’s Total Intracranial Volume (eTIV) [39]. Table 2 summarizes the five feature sets extracted for each subject, which are concatenated into the feature matrix  $\mathbf{X}$  with dimensions  $N = 1993$  and  $D = 324$ .

TABLE 2: FreeSurfer extracted features. # stands for the number of features.

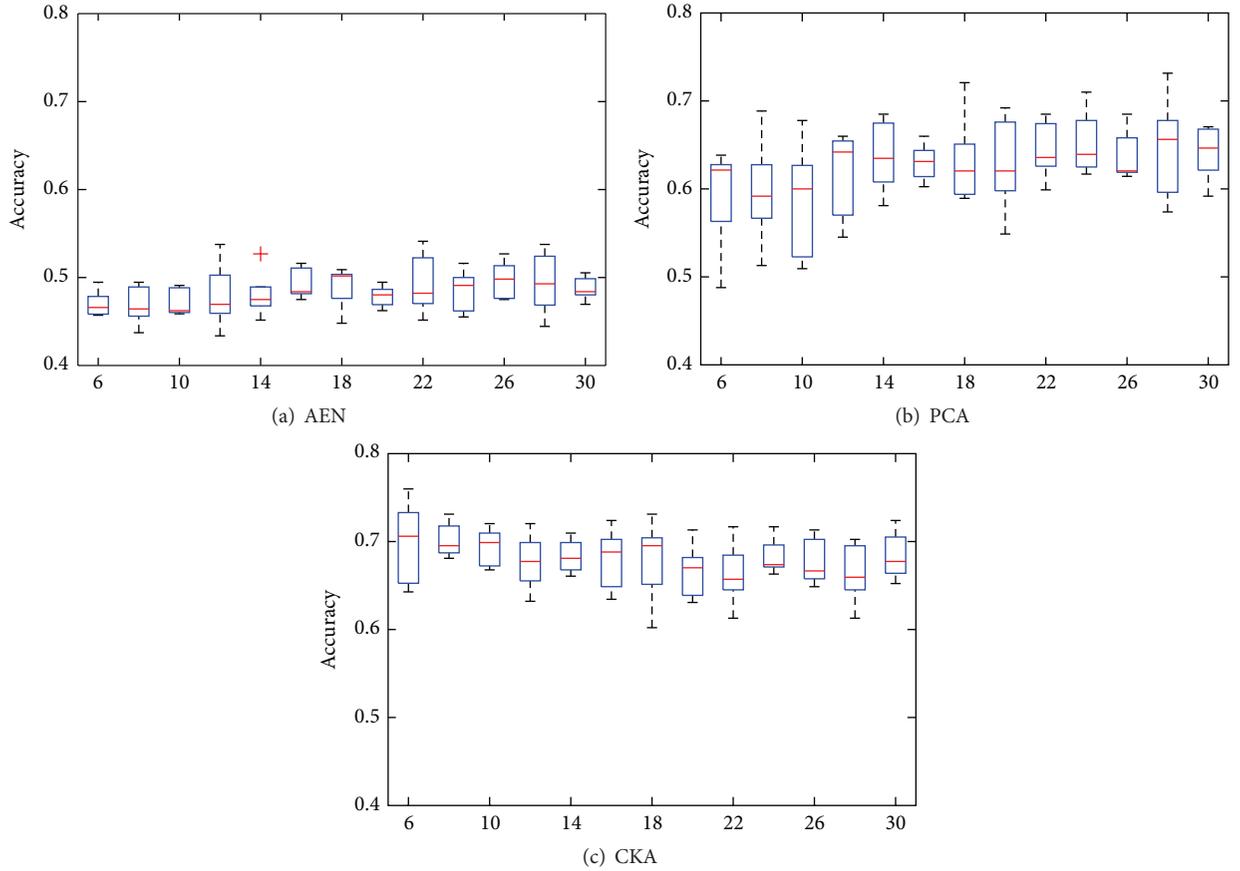
Type	# features	Units
Cortical Volumes (CV)	70	$\text{mm}^3$
Subcortical Volumes (SV)	42	$\text{mm}^3$
Surface Area (SA)	72	$\text{mm}^2$
Thickness Average (TA)	70	mm
Thickness Std. (TS)	70	mm
<b>Total</b>	324	

**3.3. Tuning of ANN Model Parameter.** Given input  $D = 324$  MRI features for classification of the 3 neurological classes, we use the feedforward ANNs with one hidden layer: 324-input and 3-output neurons. An exhaustive search is carried out for tuning the single free parameter, namely, the number of neurons in the hidden layer ( $m_1$ ). We also compare our proposal against autoencoders (AEN) [20] and the well-known Principal Components Analysis (PCA) for the initialization stage. All of these approaches (AEN, PCA, and CKA) provide a projection matrix with an output dimension that, in this case, equates the hidden layer size. Thus, resulting projections are used as the initial weights for the first layer. Also, biases and output layer weights are randomly initialized. For a different number of neurons, Figure 2 shows the accuracy results obtained by each considered strategy of initialization using 5-fold cross-validation scheme. Since we look for the most accurate and stable network configuration, we chose the optimal net as the one with the highest mean-to-deviation ratio. The resulting search indicates that the best number of hidden neurons is accomplished at  $m_1 = 20$ ,  $m_1 = 16$ , and  $m_1 = 14$  for AEN, PCA, and CKA approaches, respectively.

We further analyze the influence of each feature to the initialization process regarding the relevance criterion introduced in (11). Obtained results of relevance in Figure 3 show that the proposed CKA approach enhances the Subcortical

TABLE 3: Best performing algorithms in the 2014 CADDementia challenge [17].

Algorithm	Features	Classifier
Abdulkadir	Voxel-based morphometry	Support Vector Machine
Ledig	Volume and intensity relations	Random Forest classifier
Sørensen	Volume, thickness, shape, and intensity relations	Regularized Linear Discriminant Analysis
Wachinger	Volume, thickness, and shape	Generalized Linear Model


 FIGURE 2: Artificial neural network performance along the number of nodes in the hidden layer ( $m_1$ ) for the three initialization approaches: autoencoder, PCA-based projection, and CKA-based projection. Results are computed under 5-fold cross-validation scheme.

Volume features at the time it diminishes the influence of most Cortical Volumes and Thickness Averages. The relevance of each feature set provided by AEN and PCA is practically the same. Hence, CKA allows the selection of relevant biomarkers from MRI.

**3.4. Classifier Performance of Neurological Classes.** As shown in Table 3, the ANN models that have been tuned for the three initialization strategies are contrasted with the best four performing approaches of the 2014 CADDementia challenge [17]. The compared algorithms are evaluated in terms of their classification performance, accuracy ( $\alpha$ ), area under the receiver-operating-characteristic curve ( $\beta$ ), and class-wise

true positive rate ( $\tau_p^c$ ) criteria, respectively, which are defined as

$$\begin{aligned}
 \alpha &= \frac{\sum_c (t_p^c + t_n^c)}{\sum_c N^c}, \\
 \tau^c &= \frac{t_p^c}{N^c}, \\
 \beta &= \frac{\sum_c \beta^c \cdot N^c}{\sum_c N^c},
 \end{aligned} \tag{12}$$

where  $c \in \{\text{NC, MCI, AD}\}$  indexes each class and  $N^c$ ,  $t_p^c$ , and  $t_n^c$  are the number of samples, true positives, and true negatives for the  $c$ th class, respectively. The area under the

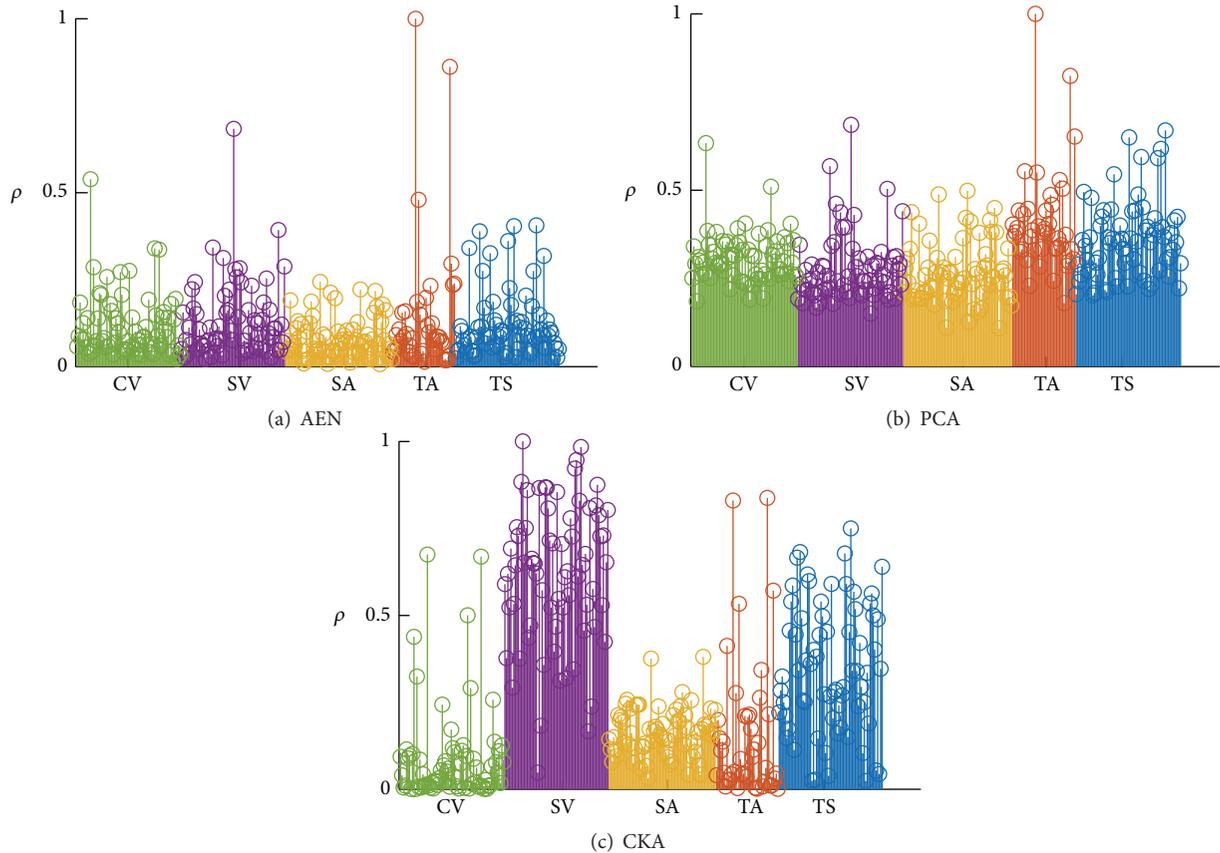


FIGURE 3: Relevance indexes grouped by feature type: Cortical Volume (CV), Subcortical Volume (SV), Surface Area (SA), Thickness Average (TA), and Thickness Std. (TS).

curve  $\beta$  is the weighted average of the area under the ROC curve of each class  $\beta^c$ . Presented results for the baseline approaches are the ones reported on the challenge for 354 images. Although the testing groups on the challenge and on this paper are not exactly the same, the amount of data, their characteristics, and the blind setup make those two groups equivalent for evaluation purposes.

As seen in Table 4 which compares the classification performance on the 30% “best” quality test set for considered algorithms, the proposed approach, besides outperforming other compared approaches of initialization, also performs better than other computer-aided diagnosis methods as a whole. For the “partial” quality images, as expected, the general performance diminishes in all ANN approaches. Nonetheless, the overall accuracy and AUC are still competitive with respect to the challenge winner. Based on the displayed ROC curves and confusion matrices for the ANN-based classifiers with the optimum parameter set (see Figure 4), we also infer that the proposed approach improves MCI discrimination.

#### 4. Discussion

From the validation carried out above for MRI-based dementia diagnosis, the following aspects emerge as relevant for the developed proposal of ANN pretraining:

TABLE 4: Classification performance on the testing groups for considered algorithms under evaluation criteria. Top: baseline approaches. Bottom: ANN pretrainings.

Algorithm	$\alpha$	$\tau^{\text{NC}}$	$\tau^{\text{MCI}}$	$\tau^{\text{AD}}$	$\beta$	$\beta^{\text{NC}}$	$\beta^{\text{MCI}}$	$\beta^{\text{AD}}$
2014 CADDementia								
Sørensen	<b>63.0</b>	<b>96.9</b>	28.7	<b>61.2</b>	<b>78.8</b>	86.3	<b>63.1</b>	87.5
Wachinger	59.0	72.1	51.6	51.5	77.0	83.3	59.4	<b>88.2</b>
Ledig	57.9	89.1	41.0	38.8	76.7	<b>86.6</b>	59.7	84.9
Abdulkadir	53.7	45.7	<b>65.6</b>	49.5	77.7	85.6	59.9	86.7
“best” quality testing								
NN-AEN	47.6	73.4	33.1	38.1	64.9	71.4	53.4	75.1
NN-PCA	63.8	70.4	56.7	66.9	80.0	87.2	70.0	87.0
NN-CKA	<b>70.9</b>	78.4	<b>66.6</b>	<b>68.3</b>	<b>85.3</b>	<b>91.7</b>	<b>78.4</b>	<b>88.3</b>
“partial” quality								
NN-AEN	62.9	64.6	46.4	32.0	77.0	82.5	65.6	72.5
NN-PCA	64.4	67.6	<b>49.3</b>	26.0	78.4	82.3	67.5	79.2
NN-CKA	<b>65.2</b>	68.6	38.6	42.0	<b>81.6</b>	85.7	<b>70.1</b>	82.4

- (i) As commonly implemented by the state-of-the-art ANN algorithms, the proposed initialization approach also has one free model parameter which is the number of hidden neurons. Tuning of this parameter is proposed to be carried out heuristically

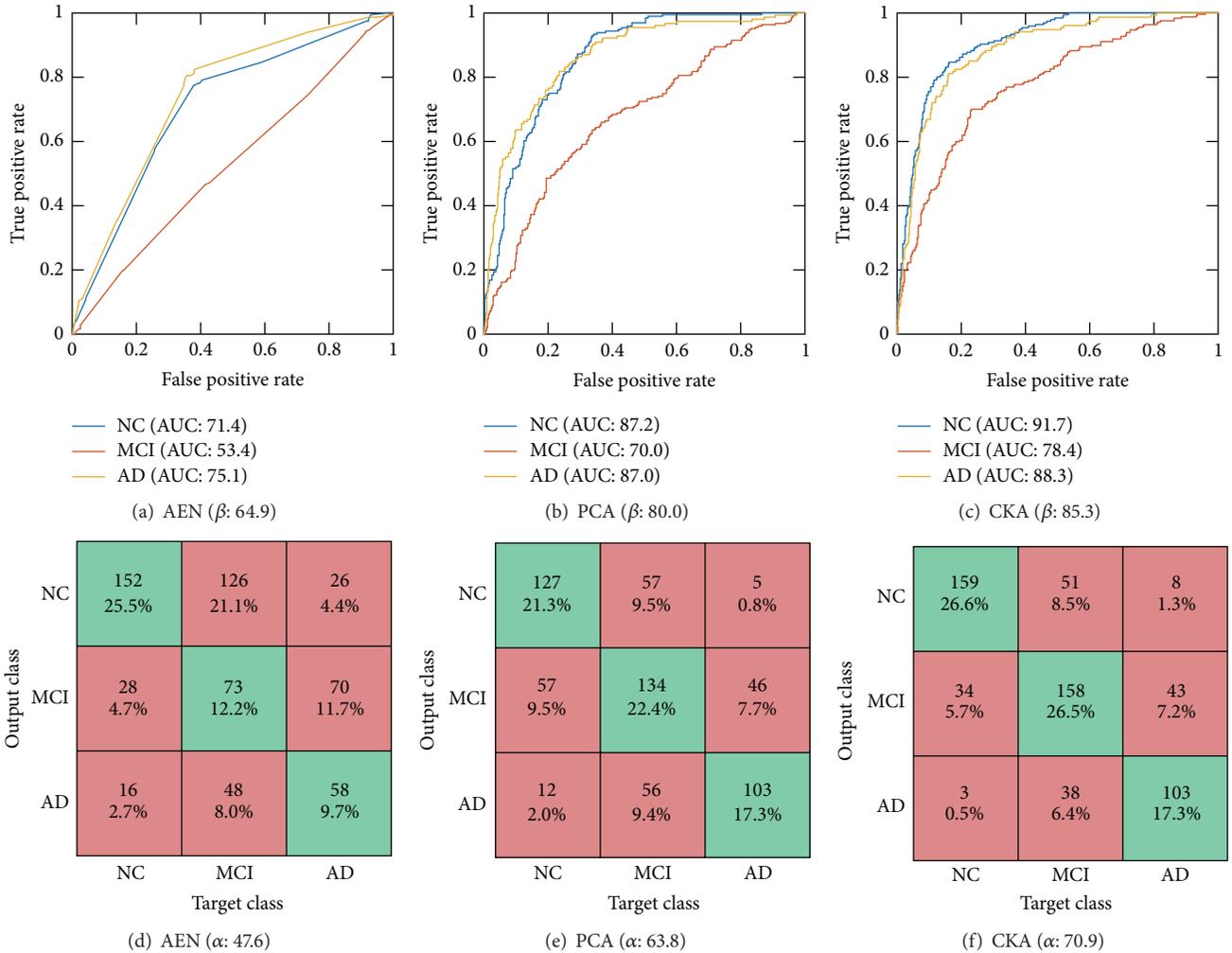


FIGURE 4: Receiver-operating-characteristic curve ((a), (b), and (c)) and confusion matrix ((d), (e), and (f)) on the 30% test data for AEN ((a) and (d)), PCA ((b) and (e)), and CKA ((c) and (f)) initialization approaches at the best parameter set of the ANN classifier.

by an exhaustive search so as to reach the highest accuracy on a 5-fold cross-validation (see Figure 2). Thus, 24, 20, and 16 hidden neurons are selected for CKA, AEN, and PCA, respectively. As a result, the suggested CKA approach improves other pretraining ANN approaches (in about 10%) with the additional benefit of decreasing the performed parameter sensitivity.

- (ii) We assess the influence of each MRI feature at the pretraining procedure regarding the relevance criterion introduced in (11). As follows from Figure 3, AEN and PCA ponder every feature evenly, restraining their ability to extract biomarkers. By contrast, CKA enhances the influence of Subcortical Volumes and Thickness Standard deviations at the time it diminishes the contribution of Cortical Volumes and Thickness Averages. Consequently, the proposed approach is also suitable for feature selection tasks.
- (iii) In the interest of comparing, we contrast the developed ANN pretraining approach with the best four

classification strategies of the 2014 CADDementia, devoted especially to dementia classification. From the obtained results, summarized in Table 4, it follows that proposed CKA outperforms other algorithms in most of the evaluation criteria and imaging conditions, providing the most balanced performance over all classes. Particularly for the 30% testing images, CKA increases by 7%-points the classification accuracy and average area under the ROC curve. It is worth noting that although Sørensen’s approach accomplishes a  $\tau^{NC}$  value that is 18.5%-points higher than the proposal, its performance turns out to be biased towards the NC, yielding a worse value of MCI. That is, CKA carries out unbiased class performance of the dementia classification. In the case of “partial” quality images, in spite of the general performance reduction, CKA remains as the best ANN initialization approach. Moreover, the overall measures are still competitive with the results provided by the CADDementia challenge.

- (iv) Figure 4 shows the per-class ROC curves and confusion matrices obtained by the contrasted approaches. In all cases, the area under the curve and accuracy for NC and AD classes are higher than the ones achieved by the MCI class (Figures 4(a)–4(c)). Hence, MCI classification from the incorporated MRI features remains a challenging task due to the following facts: the widely known MCI heterogeneity, the MCI being an intermediate class between healthy individuals and those diagnosed with Alzheimer’s disease, and the possibility of MCI subjects eventually converting to AD or NC. Moreover, confusion matrices displayed in Figures 4(d)–4(f) confirm that NC and AD are suitable for distinction in most of the cases. Nevertheless, the MCI class introduces the most errors when considered as both target and output class. Therefore, particular studies on the mild cognitive impairment should improve the diagnosis [5, 40].

## 5. Conclusion and Future Work

In this paper, we propose a supervised method for initializing the training of artificial neural networks, aiming to improve the computer-aided diagnosis of dementia. Given a set of volume, area, surface, and thickness features extracted from the subject’s brain MRI, the examined dementia diagnosis task consists of assigning subjects to the next neurological groups: normal control, mild cognitive impairment (MCI), or Alzheimer’s disease. This dementia classification task is particularly challenging because MCI is a heterogeneous and intermediate category between NC and AD. Also, MCI subjects may convert to AD or come back to NC.

To improve the classification performance, we incorporate a matrix projecting the samples into a more discriminating feature space so that the affinity between projected features and class labels is maximized. Such a criterion is implemented by the centered kernel alignment (CKA) between the feature and target label kernels, providing two key benefits: (i) the only free parameter is the hidden dimension; (ii) a relevance analysis can be introduced to find biomarkers. As a result, our proposal of ANN pretraining outperforms the contrasted algorithms (7% of classification accuracy and area under the ROC curve) and reduces the class biasing, resulting in better MCI discrimination.

Nonetheless, the use of CKA implies a couple of restrictions. Firstly, the number of samples should be larger than input and output dimensions to avoid overfitted linear projections. We cope with this drawback by considering a large enough subset of samples for training purposes (about 1300). Secondly, attained projections must always be of lower dimension compared to the original feature space. In this case, the enhancement on class discrimination is due to the affinity between labels and features, not due to an increase of the dimension.

As future work, we plan to evaluate the CKA discriminative capabilities in other neuropathological tasks from MRI as predicting Alzheimer’s conversion from MCI and attention deficit hyperactivity disorder classification. We also expect to

develop a neural network training scheme using CKA as the cost function.

## Appendix

### Gradient Descend-Based Optimization of CKA Approach

The explicit objective function of the empirical CKA in (9) yields [32]

$$\begin{aligned} \hat{\rho}_{\text{CKA}}(\mathbf{K}_W, \mathbf{B}) &= \log(\text{tr}(\mathbf{K}_W \mathbf{H} \mathbf{B} \mathbf{H})) \\ &\quad - \frac{1}{2} \log(\text{tr}(\mathbf{K}_W \mathbf{H} \mathbf{K}_W \mathbf{H})) + \rho_0, \end{aligned} \quad (\text{A.1})$$

with  $\rho_0 \in \mathbb{R}$  being a constant independent of  $\mathbf{W}$ . We then consider the gradient descent approach to iteratively solve the optimization problem. To this end, we compute the gradient of the explicit function in (A.1) with respect to  $\mathbf{W}$  as

$$\begin{aligned} \nabla_W(\hat{\rho}_{\text{CKA}}(\mathbf{K}_W, \mathbf{B})) \\ = -4\mathbf{W}((\mathbf{G} \circ \mathbf{K}_W) - \text{diag}(\mathbf{1}^\top (\mathbf{G} \circ \mathbf{K}_W)))(\mathbf{X}\mathbf{W})^\top, \end{aligned} \quad (\text{A.2})$$

where  $\text{diag}(\cdot)$  and  $\circ$  denote the diagonal operator and the Hadamard product, respectively.  $\mathbf{G} \in \mathbb{R}^{N \times N}$  is the gradient of the objective function with respect to the kernel matrix  $\mathbf{K}_W$ :

$$\begin{aligned} \mathbf{G} &= \nabla_{\mathbf{K}_W}(\hat{\rho}_{\text{CKA}}(\mathbf{K}_A, \mathbf{B})) \\ &= \frac{\mathbf{H}\mathbf{B}\mathbf{H}}{\text{tr}(\mathbf{K}_W \mathbf{H} \mathbf{B} \mathbf{H})} - \frac{\mathbf{H}\mathbf{K}_W \mathbf{H}}{\text{tr}(\mathbf{K}_W \mathbf{H} \mathbf{K}_W \mathbf{H})}. \end{aligned} \quad (\text{A.3})$$

As a result, the updating rule for  $\mathbf{W}$ , given the initial guess  $\mathbf{W}^0$ , becomes

$$\mathbf{W}^{t+1} = \mathbf{W}^t - \mu_{W^t}^t \nabla_{\mathbf{W}^t}(\hat{\rho}_{\text{CKA}}(\mathbf{K}_W, \mathbf{B})), \quad (\text{A.4})$$

with  $\mu_{W^t}^t \in \mathbb{R}^+$  being the step size of the updating rule and  $\mathbf{W}^t$  being the estimated projection matrix at iteration  $t$ .

## Competing Interests

The authors declare that there are no competing financial, professional, or personal interests influencing the performance or presentation of the work described in this paper.

## Acknowledgments

This work was supported by *Programa Nacional de Formación de Investigadores “Generación del Bicentenario” 2011* and the research Project no. 111956933522, both funded by COLCIENCIAS. Besides, this research would not have been possible without the funding of the E-health project *“Plataforma tecnológica para los servicios de teleasistencia, emergencias médicas, seguimiento y monitoreo permanente de pacientes y apoyo a los programas de prevención” Eje 3-ARTICA*.

## References

- [1] M. Prince, R. Bryce, E. Albanese, A. Wimo, W. Ribeiro, and C. P. Ferri, "The global prevalence of dementia: a systematic review and metaanalysis," *Alzheimer's & Dementia*, vol. 9, no. 1, pp. 63.e2–75.e2, 2013.
- [2] M. Wortmann, "Dementia: a global health priority—highlights from an ADI and World Health Organization report," *Alzheimer's Research & Therapy*, vol. 4, no. 5, article 40, 2012.
- [3] R. Brookmeyer, E. Johnson, K. Ziegler-Graham, and H. M. Arrighi, "Forecasting the global burden of Alzheimer's disease," *Alzheimer's & Dementia*, vol. 3, no. 3, pp. 186–191, 2007.
- [4] S. Klöppel, J. Peter, A. Ludl et al., "Applying automated MR-based diagnostic methods to the memory clinic: a prospective study," *Journal of Alzheimer's Disease*, vol. 47, no. 4, pp. 939–954, 2015.
- [5] R. Wolz, V. Julkunen, J. Koikkalainen et al., "Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease," *PLoS ONE*, vol. 6, no. 10, Article ID e25446, pp. 1–9, 2011.
- [6] B. M. Tijms, A. M. Wink, W. de Haan et al., "Alzheimer's disease: connecting findings from graph theoretical studies of brain networks," *Neurobiology of Aging*, vol. 34, no. 8, pp. 2023–2036, 2013.
- [7] S. Lithfous, A. Dufour, and O. Després, "Spatial navigation in normal aging and the prodromal stage of Alzheimer's disease: insights from imaging and behavioral studies," *Ageing Research Reviews*, vol. 12, no. 1, pp. 201–213, 2013.
- [8] B. Dubois, H. H. Feldman, C. Jacova et al., "Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria," *The Lancet Neurology*, vol. 13, no. 6, pp. 614–629, 2014.
- [9] 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.
- [10] C. R. Jack, D. S. Knopman, W. J. Jagust et al., "Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers," *The Lancet Neurology*, vol. 12, no. 2, pp. 207–216, 2013.
- [11] G. A. Papakostas, A. Savio, M. Graña, and V. G. Kaburlasos, "A lattice computing approach to Alzheimer's disease computer assisted diagnosis based on MRI data," *Neurocomputing*, vol. 150, pp. 37–42, 2015.
- [12] L. Sørensen, A. Pai, C. Igel, and M. Nielsen, "Hippocampal texture predicts conversion from MCI to Alzheimer's disease," *Alzheimer's & Dementia*, vol. 9, no. 4, p. P581, 2013.
- [13] S. Klöppel, A. Abdulkadir, C. R. Jack, N. Koutsouleris, J. Mourão-Miranda, and P. Vemuri, "Diagnostic neuroimaging across diseases," *NeuroImage*, vol. 61, no. 2, pp. 457–463, 2012.
- [14] E. Moradi, A. Pepe, C. Gaser, H. Huttunen, and J. Tohka, "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects," *NeuroImage*, vol. 104, pp. 398–412, 2015.
- [15] S. F. Eskildsen, P. Coupé, V. S. Fonov, J. C. Pruessner, and D. L. Collins, "Structural imaging biomarkers of Alzheimer's disease: predicting disease progression," *Neurobiology of Aging*, vol. 36, supplement 1, pp. S23–S31, 2015.
- [16] S. Farhan, M. A. Fahiem, and H. Tauseef, "An ensemble-of-classifiers based approach for early diagnosis of Alzheimer's disease: classification using structural features of brain images," *Computational and Mathematical Methods in Medicine*, vol. 2014, Article ID 862307, 11 pages, 2014.
- [17] E. E. Bron, M. Smits, W. M. van der Flier et al., "Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge," *NeuroImage*, vol. 111, pp. 562–579, 2015.
- [18] F. Amato, A. López, E. M. Peña-Méndez, P. Vañhara, A. Hampl, and J. Havel, "Artificial neural networks in medical diagnosis," *Journal of Applied Biomedicine*, vol. 11, no. 2, pp. 47–58, 2013.
- [19] D. Chyzyk, A. Savio, and M. Graña, "Evolutionary ELM wrapper feature selection for Alzheimer's disease CAD on anatomical brain MRI," *Neurocomputing*, vol. 128, pp. 73–80, 2014.
- [20] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion," *Journal of Machine Learning Research*, vol. 11, no. 3, pp. 3371–3408, 2010.
- [21] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.
- [22] Y. Bengio and P. Lamblin, "Greedy layer-wise training of deep networks," in *Advances in Neural Information Processing Systems 19*, pp. 153–160, MIT Press, 2007.
- [23] M. Ranzato, C. Poultney, S. Chopra, and Y. L. Cun, "Efficient learning of sparse representations with an energy-based model," in *Proceedings of the Advances in Neural Information Processing Systems (NIPS '07)*, pp. 1137–1144, 2007.
- [24] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," in *Neural Networks: Tricks of the Trade*, vol. 7700 of *Lecture Notes in Computer Science*, pp. 437–478, Springer, Berlin, Germany, 2012.
- [25] J. Weston, F. Ratle, H. Mobahi, and R. Collobert, "Deep learning via semi-supervised embedding," in *Neural Networks: Tricks of the Trade*, G. Montavon, G. B. Orr, and K.-R. Müller, Eds., vol. 7700 of *Lecture Notes in Computer Science*, pp. 639–655, Springer, Berlin, Germany, 2012.
- [26] A.-R. Mohamed, T. N. Sainath, G. Dahl, B. Ramabhadran, G. E. Hinton, and M. A. Picheny, "Deep belief networks using discriminative features for phone recognition," in *Proceedings of the 36th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '11)*, pp. 5060–5063, Prague, Czech Republic, May 2011.
- [27] Y. Bengio, "Learning deep architectures for AI," *Foundations and Trends in Machine Learning*, vol. 2, no. 1, pp. 1–127, 2009.
- [28] I. A. Basheer and M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application," *Journal of Microbiological Methods*, vol. 43, no. 1, pp. 3–31, 2000.
- [29] J.-W. Xu, A. R. Paiva, I. Park, and J. C. Principe, "A reproducing kernel Hilbert space framework for information-theoretic learning," *IEEE Transactions on Signal Processing*, vol. 56, no. 12, pp. 5891–5902, 2008.
- [30] M. Orbes-Arteaga, D. Cárdenas-Peña, M. A. Álvarez, A. A. Orozco, and G. Castellanos-Domínguez, "Kernel centered alignment supervised metric for multi-atlas segmentation," in *Image Analysis and Processing—ICIAP 2015: 18th International Conference, Genoa, Italy, September 7–11, 2015, Proceedings, Part I*, vol. 9279 of *Lecture Notes in Computer Science*, pp. 658–667, Springer, Berlin, Germany, 2015.
- [31] C. Cortes, M. Mohri, and A. Rostamizadeh, "Algorithms for learning kernels based on centered alignment," *Journal of Machine Learning Research*, vol. 13, pp. 795–828, 2012.

- [32] A. J. Brockmeier, J. S. Choi, E. G. Kriminger, J. T. Francis, and J. C. Principe, "Neural decoding with kernel-based metric learning," *Neural Computation*, vol. 26, no. 6, pp. 1080–1107, 2014.
- [33] B. Fischl, "FreeSurfer," *NeuroImage*, vol. 62, no. 2, pp. 774–781, 2012.
- [34] X. Han, J. Jovicich, D. Salat et al., "Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer," *NeuroImage*, vol. 32, no. 1, pp. 180–194, 2006.
- [35] F. Ségonne, A. M. Dale, E. Busa et al., "A hybrid approach to the skull stripping problem in MRI," *NeuroImage*, vol. 22, no. 3, pp. 1060–1075, 2004.
- [36] 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.
- [37] F. Ségonne, J. Pacheco, and B. Fischl, "Geometrically accurate topology-correction of cortical surfaces using nonseparating loops," *IEEE Transactions on Medical Imaging*, vol. 26, no. 4, pp. 518–529, 2007.
- [38] B. Fischl, A. van der Kouwe, C. Destrieux et al., "Automatically parcellating the human cerebral cortex," *Cerebral Cortex*, vol. 14, no. 1, pp. 11–22, 2004.
- [39] R. L. Buckner, D. Head, J. Parker et al., "A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume," *NeuroImage*, vol. 23, no. 2, pp. 724–738, 2004.
- [40] J. Ramírez, J. M. Górriz, A. Ortiz, P. Padilla, and F. J. Martínez-Murcia, "Ensemble tree learning techniques for magnetic resonance image analysis," in *Innovation in Medicine and Healthcare 2015*, vol. 45 of *Smart Innovation, Systems and Technologies*, pp. 395–404, Springer, Berlin, Germany, 2016.

## Research Article

# Establishment of Relational Model of Congenital Heart Disease Markers and GO Functional Analysis of the Association between Its Serum Markers and Susceptibility Genes

Min Liu,<sup>1,2</sup> Luosha Zhao,<sup>1</sup> and Jiaying Yuan<sup>3</sup>

<sup>1</sup>Department of Cardiovascular Medicine, The First Affiliated Hospital of Zhengzhou University, No. 1 East Jianshe Road, Zhengzhou 450052, China

<sup>2</sup>Department of Cardiovascular Medicine, Zhengzhou Central Hospital, Zhengzhou University, No. 195 Tongbai Road, Zhengzhou 450007, China

<sup>3</sup>Department of Ultrasound Diagnosis, Directly under Hospital of Henan Military Region, No. 18 Jinshui Road, Zhengzhou 450000, China

Correspondence should be addressed to Luosha Zhao; zhaols12@163.com

Received 3 August 2015; Revised 24 September 2015; Accepted 1 October 2015

Academic Editor: Krishna Agarwal

Copyright © 2016 Min Liu 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.

**Purpose.** The purpose of present study was to construct the best screening model of congenital heart disease serum markers and to provide reference for further prevention and treatment of the disease. **Methods.** Documents from 2006 to 2014 were collected and meta-analysis was used for screening susceptibility genes and serum markers closely related to the diagnosis of congenital heart disease. Data of serum markers were extracted from 80 congenital heart disease patients and 80 healthy controls, respectively, and then logistic regression analysis and support vector machine were utilized to establish prediction models of serum markers and Gene Ontology (GO) functional annotation. **Results.** Results showed that NKX2.5, GATA4, and FOG2 were susceptibility genes of congenital heart disease. CRP, BNP, and cTnI were risk factors of congenital heart disease ( $p < 0.05$ ); cTnI, hs-CRP, BNP, and Lp(a) were significantly close to congenital heart disease ( $p < 0.01$ ). ROC curve indicated that the accuracy rate of Lp(a) and cTnI, Lp(a) and BNP, and BNP and cTnI joint prediction was 93.4%, 87.1%, and 97.2%, respectively. But the detection accuracy rate of the markers' relational model established by support vector machine was only 85%. GO analysis suggested that NKX2.5, GATA4, and FOG2 were functionally related to Lp(a) and BNP. **Conclusions.** The combined markers model of BNP and cTnI had the highest accuracy rate, providing a theoretical basis for the diagnosis of congenital heart disease.

## 1. Introduction

Congenital heart disease (CHD) indicates the presence of abnormality in heart and vascular structure and function at birth, the pathogenesis of which is complex. It is the interaction results of multiple factors like heredity and environment. The known risk factors include mental stimulation during pregnancy [1], harmful substances exposure [2], smoking and drinking [3], viral infections at early stage of pregnancy [4], diabetes mellitus [5], history of unhealthy pregnancy [6], and too high maternal age [7]. Its clinical consequences are extremely serious. It is the important cause of miscarriage, stillbirth, neonatal death, and children, adolescents, and

adults with disabilities. The incidence of fetal CHD reaches as much as 6% to 10% [8] and continues to show a significant upward trend in China [9].

Currently, CHD is still cured by surgery. Many scholars believe that a number of indicators such as the level of serum C-reactive protein (CRP), brain natriuretic peptide (BNP), cardiac troponin I (cTnI), and Lipoprotein(a) (Lp(a)) can better reflect the functional status of the heart in patients with CHD and have good potential in clinical analysis. These proteins may serve as indicators in prognosis evaluation.

Since the United States has announced precision medicine plan, countries around the world have increased the support for precision medicine. With the enrichment and

improvement of clinical big data and biological networks, it has become a general trend to complete interdisciplinary collaboration in disease prediction, diagnosis, and etiology analysis. In daily life, clinicians commonly use Logistic regression analysis to analyze the prognostic factors of the disease and estimate the probability of occurrence of variables [10]. Support vector machine (SVM) is a new machine learning method based on statistical theory. SVM is good at coping with linearly nonseparable sample data, which is achieved mainly through the slack variables (which are also called punishment variables) and kernel technology. It provides a unified framework in solving learning problems of finite samples [11].

Increasing studies show that the pathogenesis of congenital heart disease is related to certain transcription factors, while the relationship between the susceptibility genes and serological markers of congenital heart disease is not yet reported. With the rapid application of bioinformatics, Gene Ontology (GO) has become important tool and method in the field of bioinformatics. In terms of gene function annotation, GO plays a huge role. It can analyze the location of gene or protein in the cell, molecular functions, and biological processes involved; thus it simplifies the annotation of genes and their products as standardized vocabularies.

In this study, data of the susceptibility genes and clinical serology risk factors literatures of CHD were performed Meta-analysis to systematically evaluate them. By detecting levels of serum markers in patients with CHD, Logistic regression analysis, receiver operating characteristic (ROC) curve, and SVM approaches were used to evaluate the value of each serum marker in clinical diagnosis of CHD. The detection model of serum markers of this disease was then established. The functional relationship between susceptibility genes and serum markers was established by GO analysis. As a result, this study provides a theoretical basis for clinical practice and personalized treatment of cardiovascular disease.

## 2. Materials and Methods

### 2.1. Meta-Analysis

**2.1.1. Subjects.** Clinical research documents on susceptibility genes and serological markers of CHD published in China and foreign countries from January 2006 to October 2014 were selected.

**2.1.2. Document Retrieval.** Google Scholar was a major source of Chinese documents; PubMed, EMBASE, MEDLINE, and MD consult were main sources of English documents and the Chinese or English key words were “congenital heart disease”, “gene”, and “mutation” as well as “congenital heart disease”, “serum markers”, and “diagnosis”. The years of publication were from January 1, 2000, to October 31, 2014.

**2.2. Statistical Analysis.** RevMan5.1 was used for meta-analysis of the included literature.  $p \geq 0.05$  showed that the merge statistics of multiple studies had no statistical

significance;  $p < 0.05$  indicated that the combined statistics were statistically significant.

### 2.3. Establishing Relational Model of CHD Markers Group

**2.3.1. Research Data.** In this study, 80 CHD patients (33 with atrial septal defect, 36 with ventricular septal defect, 3 with patent ductus arteriosus, and 8 with tetralogy of Fallot) received treatment in the Department of Cardiac Surgery at our hospital from December 2009 to September 2014 (54 males and 26 females, aged from 7 days to 59 years) and 80 healthy outpatients as determined by a physical examination given at the hospital (38 males and 42 females, aged 3.6 months to 51 years) were selected as the subjects. Patients in case group were confirmed by echocardiography and (or) surgery, and the following cases were excluded: (1) renal insufficiency, chronic liver disease, and acute and chronic infectious diseases; (2) systemic lupus erythematosus, rheumatoid, and other immune system diseases; and (3) infectious endocarditis, rheumatic heart disease, cardiac tumors, myocarditis, and other types of heart disease. Healthy control group denied a family history of CHD. They were confirmed to have no cardiac dysfunction and organic diseases by physical examination and echocardiography. Infection, trauma, autoimmune diseases, cancer, and so on were also excluded.

10 mL of venous blood was collected from all study subjects in the morning after 12 h overnight fasting and put into the EDTA anticoagulant tube. Samples were centrifuged within 2 h at 3000 r/min for 10 min, and then the supernatants were collected.

**2.3.2. Sample Testing.** Serum BNP level was detected using enzyme-linked immunosorbent assay (ELISA). Serum hs-CRP was examined using immune rate nephelometry. Immunofluorescence method was used to determine serum cTnI level. ELISA double-antibody sandwich assay was adopted to test serum Lp(a) level. Detection methods were carried out in strict accordance with the kit instructions. Each sample received parallel testing twice and the average value was regarded as final test results.

**2.3.3. Establishing Relational Model of CHD Markers Group Based on Logistic Regression Analysis.** Serum markers BNP, hs-CRP, cTnI, and Lp(a) levels of CHD patients and healthy control group undergone Logistic regression analysis with the new variables of Logistic regression model as test variables and the pathological diagnosis results as state variables; the ROC curve was drawn. According to the value of the area under the curve (AUC) of ROC and diagnostic accuracy, its application value in early diagnosis of CHD was evaluated.

**2.3.4. Establishing Relational Model of CHD Markers Group Based on SVM.** Data of the 80 CHD patients were treated with normalization processing. The establishment, training, and validation of SVM model were achieved through MATLAB programming.

**2.3.5. Statistical Analysis.** The data obtained undergone significance of difference analysis using statistical software SPSS19.0 and the data were expressed by the following: mean  $\pm$  standard deviation.  $p < 0.05$  indicated that the difference was statistically significant.

### 2.4. Bioinformatics Functional Analysis of Serum Markers

#### *Lp(a) and BNP and Susceptibility Genes of CHD*

**2.4.1. GO Retrieval.** Congenital heart disease-related susceptibility genes NKX2.5, GATA4, and FOG2 and serological markers hs-CRP, Lp(a), BNP, and cTnI undergone GO functional annotation using AmiGO platform.

**2.4.2. RT-PCR.** RNA kit from TAKARA (Takara Bio Inc., Shiga, Japan) was used to extract serum RNA, and Thermo Scientific RevertAid First Strand cDNA Synthesis Kit was used for reverse transcription experiments. With the synthesized cDNA template and GAPDH as template, we performed fluorescence quantitative PCR reactions. Fluorescent dye SYBR and quantitative real-time PCR instrument CFX96 were applied in this experiment. Primers are shown in Table 1 (primers were synthesized by Shanghai Sangon Biotech Co., Ltd., Shanghai, China). 20  $\mu$ L system of PCR reaction was as follows: 10  $\mu$ LSYBG Mix + 8  $\mu$ L H<sub>2</sub>O + 0.5  $\mu$ L upstream primer + 0.5  $\mu$ L downstream primer + 1  $\mu$ L cDNA; reaction conditions were as follows: denaturation at 95°C for 30 s, PCR reaction at 95°C for 5 s, and collecting fluorescence at 55°C for 30 s, with a total of 40 cycles, repeated three times.

## 3. Results

**3.1. Meta-Analysis of Susceptibility Genes and Serum Markers.** There were 176 documents about susceptibility and 216 documents about serum markers for initial survey after screening, there were 19 documents about susceptibility [12–31], and 20 documents about serum markers [32–51] were eventually included for meta-analysis.

Meta-analysis results of susceptibility genes and serum markers are shown in Tables 2 and 3. The heterogeneity test result of susceptibility genes NKX2.5 and FOG2 was  $p > 0.05$ , indicating the consistency of the literatures was well, so fixed effect model was used to pool the data. The heterogeneity test result of GATA4 was  $p < 0.05$ , suggesting that heterogeneity existed between the literatures, so the random effect model was adopted. The upper and lower limit of pooled SMD and 95% CI were greater than 1, indicating that the correlation between the mutation of three genes and congenital heart disease was statistical significance. The heterogeneity test result of three serum markers was  $p < 0.05$ , indicating that heterogeneity existed between literatures, so the random effect model was adopted. The upper and lower limit of pooled WMD and 95% CI were all greater than 0. Additionally, 95% CI transverse lines of three serum markers fell to the left side of the invalid vertical lines, suggesting that the incidence rate of the experimental group was bigger than that of the control group. Specific meta-analysis results are shown in Additional Files 1–6 (see Supplementary Material available online at <http://dx.doi.org/10.1155/2016/9506829>).

TABLE 1: Name, sequence, and product length of each primer.

Primer name	Primer sequence	Product length
NKX2.5-F	AGAAGACAGAGGCGGACAAC	175 bp
NKX2.5-R	CGTGGACGTGAGTTTCAGCA	
GATA4-F	GTGTCCCAGACGTTCTCAGT	226 bp
GATA4-R	TCCGTGCAGGAATTTGAGGA	
FOG2-F	TTAATCAACGGAAGCAAATG	466 bp
FOG2-R	CCACTCAAATACAGGGTTAGG	
GAPDH-F	AGAAGGCTGGGGCTCATTTC	258 bp
GAPDH-R	AGGGGCCATCCACAGTCTTC	

**3.2. Test Results of Serum Markers.** The test results of serum markers cTnI, hs-CRP, BNP, and Lp(a) of 80 patients with CHD and 80 healthy persons are shown in Figure 1. As can be seen from the figure, the levels of cTnI, hs-CRP, BNP, and Lp(a) in the case group were significantly higher than those in the controls ( $p < 0.05$ ).

**3.3. Logistic Regression Analysis Results.** With cTnI, hs-CRP, BNP, and Lp(a) as independent variables and sick or not as the dependent variable, SPSS19.0 was used for dichotomy Logistic regression analysis. Univariate regression analysis results are presented in Table 4, which suggested that the relationship between Lp(a), BNP, and cTnI with CHD was statistically significant ( $p < 0.05$ ). These three factors were then used for multivariate Logistic regression analysis. The results showed that the combination of these three factors was unfavorable for accurate diagnosis of CHD ( $p > 0.05$ , Table 5). Pairwise combinations of three factors were conducted for multivariate Logistic regression analysis and the results are presented in Table 6. It was indicated that the relationship between Lp(a), BNP, and cTnI with CHD had statistical significance ( $p < 0.05$ ). The accuracy rates of combined predication of Lp(a) and cTnI, Lp(a) and BNP, and BNP and cTnI were 93.4%, 87.1%, and 97.2%, respectively.

**3.4. Application Value Evaluation of Serum Markers on the Detection of CHD.** SPSS19.0 software was adopted to evaluate the application value of Lp(a), BNP, and cTnI combined detection of CHD. ROC curves are shown in Figure 2. The AUC of Lp(a) and cTnI, Lp(a) and BNP, and BNP and cTnI joint detection were 0.994, 0.981, and 0.999, respectively, showing a high application value.

**3.5. Establishing Relational Model of CHD Serum Markers Group Based on SVM.** Serum markers cTnI, hs-CRP, BNP, and Lp (a) levels of 80 CHD patients and 80 healthy controls undergone attributive analysis. It was indicated that attributive analysis had significant classification and the data were consistent with the basic calculation requirements of SVM (Figure 3).

The relational model of CHD serum markers group based on SVM was established. Then, the test data of 20 CHD patients and 20 healthy controls were input into it. The test results are shown in Figure 4. The hollow circles represent the

TABLE 2: Meta-analysis results of susceptibility genes.

Susceptibility gene	Number of documents	Heterogeneity test		OR	95% CI
		$I^2$	$p$		
NKX2.5	7	36%	0.16	2.02	1.42–2.86
GATA4	11	58%	0.01	2.08	1.50–2.88
FOG2	4	0%	0.85	19.43	4.52–83.63

TABLE 3: Meta-analysis of serum markers.

Serum markers	Number of documents	Heterogeneity test		WMD	95% CI
		$I^2$	$p$		
cTnI	4	99.2%	0.000	0.33	0.12~0.55
hs-CRP	8	87.5%	0.000	1.84	1.36~2.32
BNP	19	99.9%	0.000	321.33	279.03–364.29

TABLE 4: Univariate regression analysis of serum markers associated with CHD.

Serum markers	$B$	SE	Wals	Sig.	Exp( $B$ )
Lp(a)	0.036	0.006	40.422	0.000	1.036
BNP	0.478	0.077	38.954	0.000	1.612
hs-CRP	83.858	400.679	0.044	0.834	2.625E36
cTnI	172.096	37.484	21.079	0.000	5.502E74

TABLE 5: Three-factor multivariate regression analysis of serum markers associated with CHD.

Serum markers	$B$	SE	Wals	Sig.	Exp( $B$ )
Lp(a)	0.046	52.695	0.018	0.869	1.047
BNP	156.787	72890.460	0.008	0.924	1.235
cTnI	0.665	152.952	0.000	0.997	1.944

TABLE 6: Two-factor multivariate regression analysis of serum markers associated with CHD.

	Serum markers	$B$	SE	Wals	Sig.	Exp( $B$ )
1	Lp(a)	0.045	0.011	15.672	0.000	1.046
	BNP	0.508	0.114	19.774	0.000	1.662
2	Lp(a)	0.043	0.012	11.876	0.001	1.044
	cTnI	211.657	61.321	11.914	0.001	8.348E91
3	BNP	0.682	0.247	7.619	0.006	1.978
	cTnI	263.631	111.283	5.612	0.018	3.115E114

target output; “\*” is the actual simulation output of SVM. As can be seen from the figure, the diagnostic accuracy of the model was  $34/40 = 85\%$ .

**3.6. GO Functional Annotation Results Comparison between Susceptibility Genes and Serum Markers of CHD.** After comparing the GO functional annotation results of susceptibility genes NKX2.5, GATA4, and FOG2 and serological indicators hs-CRP, Lp(a), BNP, and cTnI, it was found that NKX2.5,

GATA4, and FOG2 had same GO functional annotation with Lp(a) and BNP. The functional relations between three susceptibility genes and BNP were mainly in gene expression and metabolic process. The internal connections between Lp(a) and NKX2.5, GATA4, and FOG2 were mainly in function, especially in the aspects of Lipoprotein transmembrane transport and blood circulation. The same GO functional annotations of them are shown in Tables 7–9.

**3.7. Relative Expression Contents of Susceptibility Genes in mRNA Level.** Real-time fluorogenic quantitative PCR was used to detect the expression levels of susceptibility genes NKX2.5, GATA4, and FOG2 in mRNA.  $2^{-\Delta\Delta C_t}$  was used to calculate the relative expression levels, and the results were  $0.59 \pm 0.18$ ,  $0.47 \pm 0.14$ , and  $0.33 \pm 0.09$ , respectively. If the content of the control group was 1, the relative expression levels of NKX2.5, GATA4, and FOG2 in the case group were  $0.59 \pm 0.18$ ,  $0.47 \pm 0.14$ , and  $0.33 \pm 0.09$ , respectively (Figure 5). The expression levels of susceptibility genes NKX2.5, GATA4, and FOG2 in the case group were obviously lower than those in the controls. The results of serum indexes detection showed that Lp(a) and BNP levels in the case group were significantly higher than those in the controls (Figure 1). Thus it can be inferred that the unusual increase of serum Lp(a) and BNP levels may be related to the abnormal expression of NKX2.5, GATA4, and FOG2 genes.

## 4. Discussion

CHD is the most common congenital malformation at present and also the leading cause of infant death. Many factors interact with each other temporally and spatially in the development of heart. The combined actions of hereditary and environmental factors in embryonic phase will lead to the dysplasia of heart. Due to the complex genetic mechanism of CHD, the reason resulting in the malformation of heart is still unclear. The type of CHD is diverse, which has become a big problem in the treatment and prevention of CHD.

In this study, meta-analysis found that the mutation of NKX2.5, GATA4, and FOG2 genes played an important role in the development of CHD. The mutation of NKX2.5

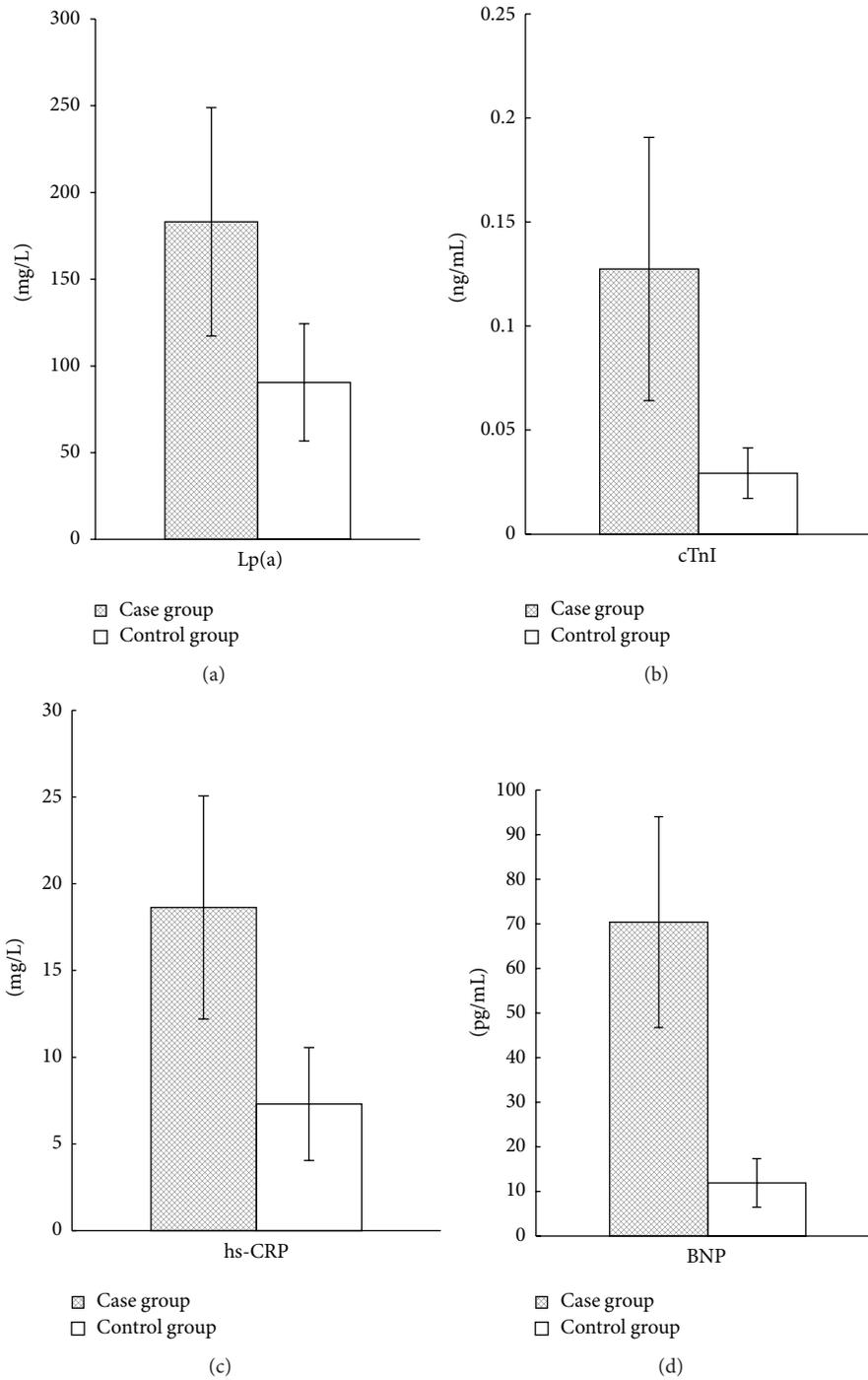


FIGURE 1: Detection results of four serum markers of CHD patients. \* indicates  $p < 0.05$ .

gene occurred mainly in homeodomain structural domain. McElhinney et al. [52] reported that the mutation of exon 1 of NKX2.5 gene existed in various CHD. The pathological and physiological effects of GATA4 gene related to heart development have been extensively researched. Garg et al. [53] have verified that GATA4 gene mutation is one of the causes of CHD for the first time by the molecular genetics research on two independent and simple CHD families. FOG2 gene is a transcription factor with early expression in

the process of heart development. Its interaction with GATA4 runs through the entire process of heart development. FOG2 plays an essential role in the development process of heart [3]. Both Tan and De Luca found a mutation in FOG2 gene exon from patients with double-outlet right ventricle combined ventricular septal defect [30, 31]. This paper found that serum markers cTnI, hs-CRP, and BNP were related to CHD and they can predict the occurrence of the disease. Guo [32] believed that the changes in serum levels of cTnI were of

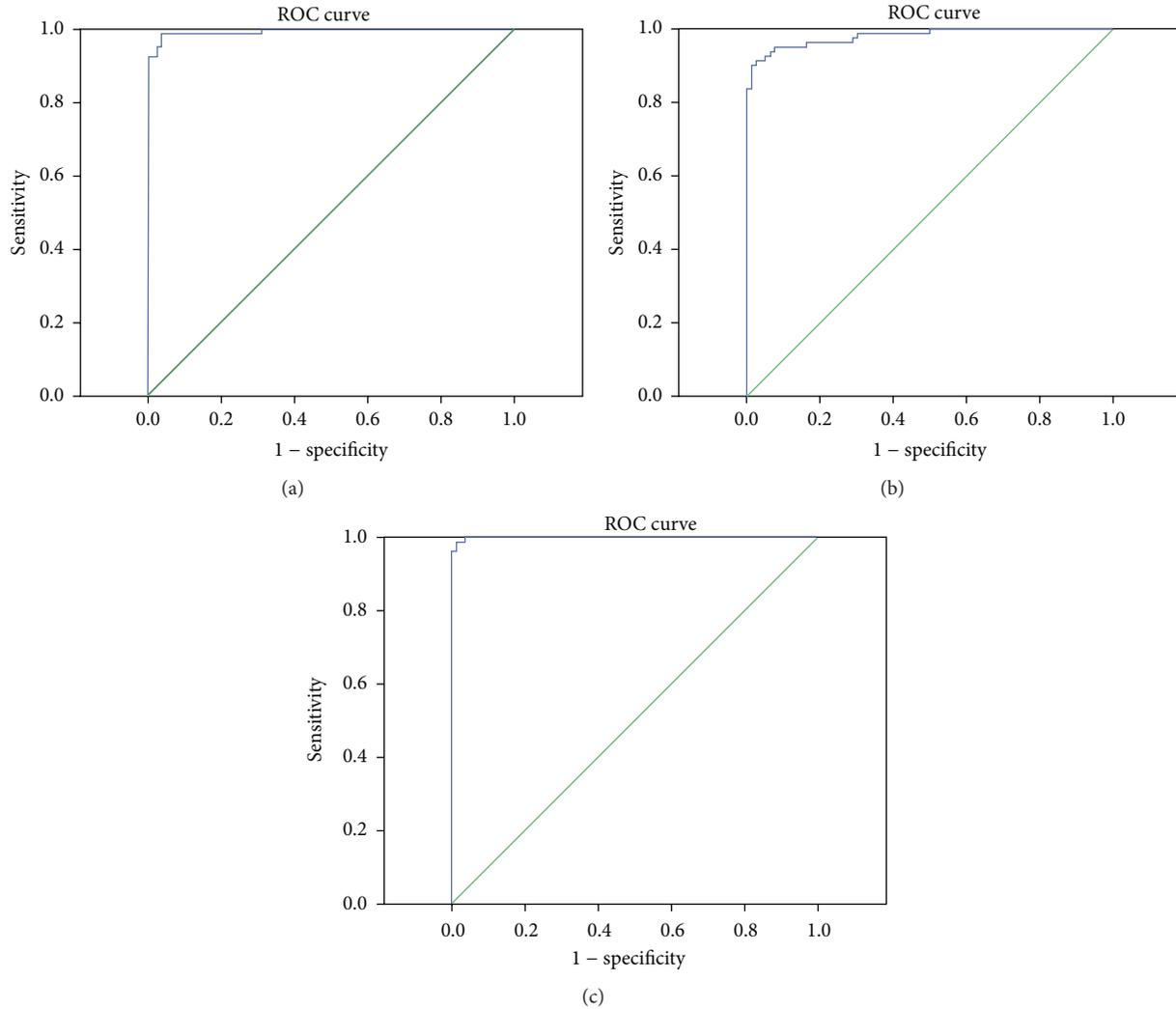


FIGURE 2: ROC curve of combined detection of CHD. (a) Lp(a) and cTnI; (b) Lp(a) and BNP; (c) BNP and cTnI.

great value in understanding the state and prognosis of CHD. However, researches on the relationship between Lp(a) and CHD were much rare, and Lp(a) did not meet the condition of meta-analysis, so we could not perform analysis of this factor.

By examining the levels of cTnI, hs-CRP, BNP, and Lp(a) of 80 CHD patients and 80 healthy control subjects, this study showed that the levels of cTnI, hs-CRP, BNP, and Lp(a) in the case group were significantly higher than those in the controls, and the difference was statistically significant. Geiger et al. [54] found that, compared to the non-CHD subjects, BNP level of CHD children was obviously increased. Similarly, Akhabue et al. [55] also believe that the difference of BNP concentration between CHD children patients and non-CHD children was significant. A number of studies show that the relationship between LP(a) and atherosclerotic disease was close, and the increased LP(a) is an independent risk factor of cardiovascular events [56–59]. Guo [32] has shown that serum cTnI level in patients with CHD was significantly higher than that in normal people. Logistic regression analysis showed that there existed

significant correlations between cTnI, BNP, Lp(a), and CHD. When performing combined diagnosis, cTnI, BNP, and Lp(a) pairwise binding were associated with CHD. According to the joint detection ROC curve, it was found that the pairwise combination AUC of cTnI, BNP, and Lp(a) were greater than 0.9, and the accuracy rates were higher than 87%. The bigger the data is, the better the effect is when using Logistic regression model. SVM in contrast has a higher accuracy rate as to small sample size.

Recent studies showed that GATA4 and GATA6 can collaborate and regulate the expression of brain natriuretic peptide (BNP). The deletion of any factor of GATA will lead to the downregulation of BNP level [60]. Other studies indicated that NKX2.5 and FOG2 could cooperate with GATA4, all of which play an important role in the normal process of heart development [61, 62]. As an independent protein molecule having a specific antigenicity, the metabolic pathways of Lp(a) is completely different from other apolipoproteins. It can interfere with lipid metabolism and the fibrinolytic system and then play an important role in cardiovascular diseases

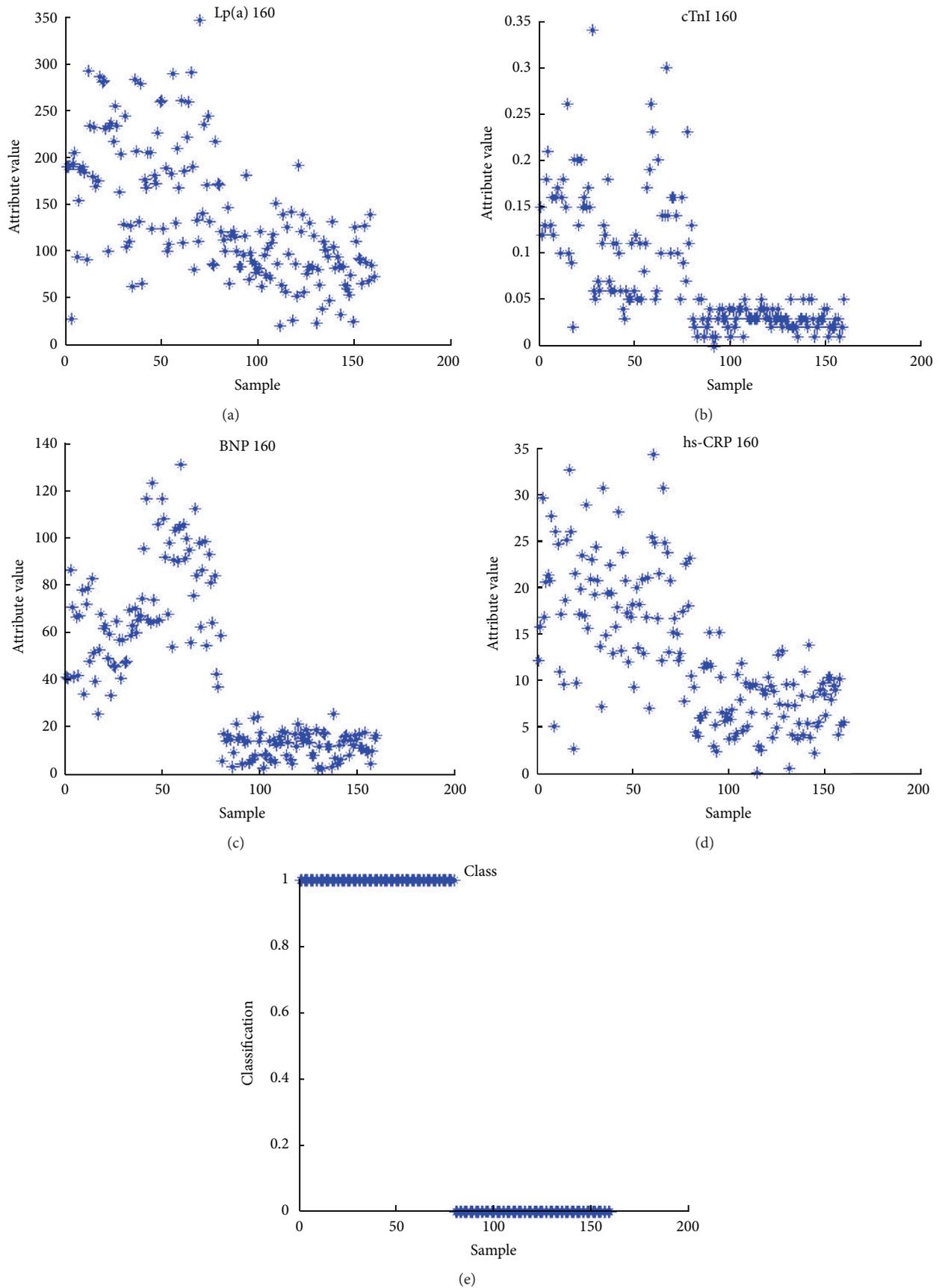


FIGURE 3: Attribution analyses of four serum markers of CHD.

TABLE 7: The same GO semantic annotation of NKX2.5 gene and serum markers Lp(a) and BNP.

Gene	GO number	Annotation	Ontology	Evidence	Reference	Serum marker
NKX2.5	0010467	Gene expression	Biological process	TAS	Reactome: REACT_71	Brain natriuretic peptide (BNP)
	0006367	Transcription initiation from RNA polymerase II promoter	Biological process	TAS	Reactome: REACT_118713 Reactome: REACT_12627	
	0007166	Cell surface receptor signaling pathway	Biological process	NAS	PMID: 12727915	
	0008015	Blood circulation	Biological process	TAS	PMID: 8047165	Lipoprotein(a) (Lp(a))
	0055085	Transmembrane transport	Biological process	TAS	Reactome: REACT_15480	
	0007399	Nervous system development	Biological process	IEA: with Ensembl ENSMUSP00000036044	GO REF: 0000019	

TABLE 8: The same GO semantic annotation of FOG2 gene and serum markers Lp(a) and BNP.

Gene	GO number	Annotation	Ontology	Evidence	Reference	Serum marker
FOG2	0001701	In utero embryonic development	Biological process	IEA: with Ensemb ENSMUSP00000036044	GO REF: 0000019	Brain natriuretic peptide (BNP)
	0007596	Blood coagulation	Biological process	TAS	Reactome: REACT_604	Lipoprotein(a)
	0010467	Gene expression	Biological process	TAS	Reactome: REACT_71	(Lp(a))
	0044702	Single organism reproductive process	Biological process	IEA: with Ensemb ENSMUSP000000274491	GO REF: 0000019	

TABLE 9: The same GO semantic annotation of GATA4 gene and serum markers Lp(a) and BNP.

Gene	GO number	Annotation	Ontology	Evidence	Reference	Serum marker	
GATA4	0010467	Gene expression	Biological process	TAS	Reactome: REACT_71	Brain natriuretic peptide (BNP)	
	0007166	Cell surface receptor signaling pathway	Biological process	NAS	PMID: 12727915		
	0014898	Cardiac muscle hypertrophy in response to stress	Biological process	IEA: with Ensemb ENSMUSP00000099520	GO REF: 0000019		
	0044702	Single organism reproductive process	Biological process	IEA: with Ensemb ENSMUSP000000274491	GO REF: 0000019		
	0008015	Blood circulation	Biological process	TAS	PMID: 8047165		
	0001701	In utero embryonic development	Biological process	IEA: with Ensemb ENSMUSP00000036044	GO REF: 0000019		
	0007283	Spermatogenesis	Biological process	IEA: with Ensemb ENSMUSP00000036044	GO REF: 0000019		
	0009743	Response to carbohydrate	Biological process	IEA: with Ensemb ENSRNOP00000039779	GO REF: 0000019		

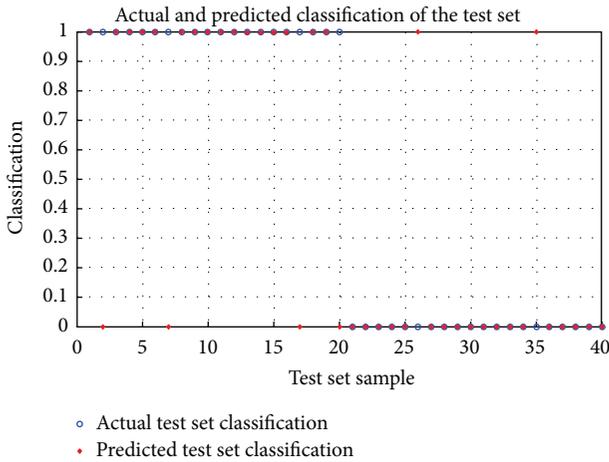


FIGURE 4: Relational model of CHD markers group based on SVM.

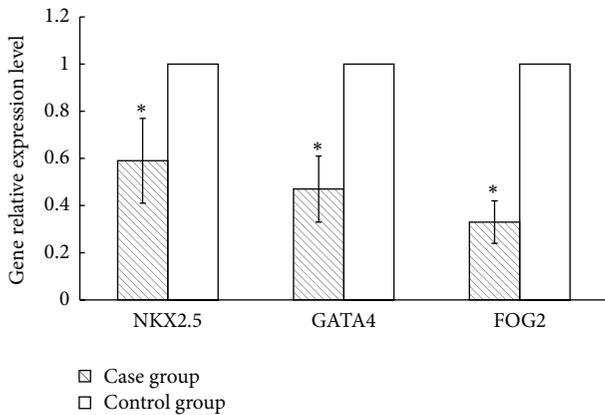


FIGURE 5: The relative expression levels of CHD susceptibility genes. \* indicates  $p < 0.05$ .

like thrombosis and atherosclerosis [63, 64]. Studies have shown that Lp(a) is an independent risk factor for myocardial infarction, coronary heart disease, and other cardiovascular diseases [65–68], but few researches are conducted on the relationship between Lp(a) and CHD. At present, it is not reported which transcription factor Lp(a) is regulated by. By bioinformatics analysis, this study showed that there were the same GO functional annotations between susceptibility gene NKX2.5, GATA4, and FOG2 and Lp(a) and BNP. The links between susceptibility genes and BNP existed mainly in gene expression and metabolism. Lp(a), especially in Lipoprotein membrane transport and blood circulation, was intrinsically linked to NKX2.5, GATA4, and FOG2. This paper conducted a study on the mRNA relative expression levels of susceptibility genes, Lp(a) and BNP. It was indicated that the levels of NKX2.5, GATA4, and FOG2 of the case group were significantly lower than those of the controls. The contents of Lp(a) and BNP of the case group were significantly higher than those of the controls, suggesting that the abnormal expression of susceptibility genes may lead to the increase of BNP level. However, the mechanism which causes the abnormal expression of Lp(a) is still not clear, so

further study is required. This also gives us a direction on the in-depth study of CHD.

### 5. Conclusions

In conclusion, as risk factors associated with CHD, cTnI, CRP, BNP and Lp(a) also have functional relation with susceptibility genes; therefore, they may provide a basis for the clinical detection of CHD, but its specific application still requires a lot of clinical cases data to train and optimize, thus making it more accurate. Clinical auxiliary testing model is only as an auxiliary tool at the early stage and cannot completely replace an experienced clinician’s diagnosis. The clinical diagnosis of CHD still needs to integrate all aspects of judgments.

### Ethical Approval

We certify that this study has followed the Declaration of Helsinki (1964).

### Consent

All subjects have given their written informed consent.

### Conflict of Interests

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

### Acknowledgments

The authors of this paper thank all authors who have contributed to it for advice and comments and thank Dr. Zhang for technical expertise. This study was supported by Funds 20120900, 20130459, and 132102310332 and the National Natural Science Foundation of China (Grant 81303150).

### References

- [1] S. Malik, M. A. Cleves, M. A. Honein et al., “Maternal smoking and congenital heart defects,” *Pediatrics*, vol. 121, no. 4, pp. e810–e816, 2008.
- [2] P. Dadvand, J. Rankin, S. Rushton, and T. Pless-Mulloli, “Association between maternal exposure to ambient air pollution and congenital heart disease: a register-based spatiotemporal analysis,” *American Journal of Epidemiology*, vol. 173, no. 2, pp. 171–182, 2011.
- [3] R. Kučienė and V. Dulskienė, “Parental cigarette smoking and the risk of congenital heart septal defects,” *Medicina*, vol. 46, no. 9, pp. 635–641, 2010.
- [4] A. Miller, T. Riehle-Colarusso, C. Siffel, J. L. Frias, and A. Correa, “Maternal age and prevalence of isolated congenital heart defects in an urban area of the United States,” *American Journal of Medical Genetics A*, vol. 155, no. 9, pp. 2137–2145, 2011.
- [5] L. A. Lisowski, P. M. Verheijen, J. A. Copel et al., “Congenital heart disease in pregnancies complicated by maternal diabetes mellitus: an international clinical collaboration, literature review, and meta-analysis,” *Herz*, vol. 35, no. 1, pp. 19–26, 2010.

- [6] L. L. Lipscombe, A. T. Banerjee, S. McTavish et al., "Readiness for diabetes prevention and barriers to lifestyle change in women with a history of gestational diabetes mellitus: rationale and study design," *Diabetes Research and Clinical Practice*, vol. 106, no. 1, pp. 57–66, 2014.
- [7] B. Pejtsik, J. Pintér, M. Horváth, and J. Hadnagy, "Relationship between congenital heart disease and various factors affecting pregnancy," *Orvosi Hetilap*, vol. 133, no. 3, pp. 155–158, 1992.
- [8] P. Pradat, C. Francannet, J. A. Harris, and E. Robert, "The epidemiology of cardiovascular defects, part I: a study based on data from three large registries of congenital malformations," *Pediatric Cardiology*, vol. 24, no. 3, pp. 195–221, 2003.
- [9] Z. Hu, X. Yuan, K. Rao, Z. Zheng, and S. Hu, "National trend in congenital heart disease mortality in China during 2003 to 2010: a population-based study," *Journal of Thoracic and Cardiovascular Surgery*, vol. 148, no. 2, pp. 596–602, 2014.
- [10] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, Berlin, Germany, 2nd edition, 1999.
- [11] X. Song and N.-K. Chen, "A SVM-based quantitative fMRI method for resting-state functional network detection," *Magnetic Resonance Imaging*, vol. 32, no. 7, pp. 819–831, 2014.
- [12] S. Pikkarainen, H. Tokola, R. Kerkelä, and H. Ruskoaho, "GATA transcription factors in the developing and adult heart," *Cardiovascular Research*, vol. 63, no. 2, pp. 196–207, 2004.
- [13] M. Zhao, X. Y. Zhu, X. L. Tian et al., "Detection of mutation in ZFPM2/FOG2 gene coding region in patients of Tetralogy of Fallot," *Chinese Journal of Cardiovascular Research*, vol. 10, pp. 422–424, 2012.
- [14] W. Huang, H. Meng, Y. Qiao, S. Pang, D. Chen, and B. Yan, "Two novel and functional DNA sequence variants within an upstream enhancer of the human NKX2-5 gene in ventricular septal defects," *Gene*, vol. 524, no. 2, pp. 152–155, 2013.
- [15] F. Xiong, Q. Li, C. Zhang et al., "Analyses of GATA4, NKX2.5, and TFAP2B genes in subjects from southern China with sporadic congenital heart disease," *Cardiovascular Pathology*, vol. 22, no. 2, pp. 141–145, 2013.
- [16] S. Pang, J. Shan, Y. Qiao et al., "Genetic and functional analysis of the NKX2-5 gene promoter in patients with ventricular septal defects," *Pediatric Cardiology*, vol. 33, no. 8, pp. 1355–1361, 2012.
- [17] X. Fang, *Exploration of the Transcription Factor NKX2.5 and GATA4 in Patients with Congenital Heart Disease*, Southeast University, 2011.
- [18] T. Peng, L. Wang, S.-F. Zhou, and X. Li, "Mutations of the GATA4 and NKX2.5 genes in Chinese pediatric patients with non-familial congenital heart disease," *Genetica*, vol. 138, no. 11, pp. 1231–1240, 2010.
- [19] J.-D. Ding, K.-R. Li, X.-L. Zhang et al., "Preliminary exploration of transcription factor Nkx2.5 mutations and congenital heart diseases," *National Medical Journal of China*, vol. 89, no. 16, pp. 1114–1116, 2009.
- [20] K. R. Li, J. D. Ding, X. L. Zhang, Y. Y. Yao, L. Q. Ren, and G. S. Ma, "Observation of transcription factor Nkx2.5 mutations of congenital heart disease patients," *Shandong Medical Journal*, vol. 48, pp. 83–84, 2008.
- [21] C. Cheng, Y. Lin, F. Yang et al., "Mutational screening of affected cardiac tissues and peripheral blood cells identified novel somatic mutations in GATA4 in patients with ventricular septal defect," *Journal of Biomedical Research*, vol. 25, no. 6, pp. 425–430, 2011.
- [22] W.-M. Zhang, X.-F. Li, Z.-Y. Ma et al., "GATA4 and NKX2.5 gene analysis in Chinese Uygur patients with congenital heart disease," *Chinese Medical Journal*, vol. 122, no. 4, pp. 416–419, 2009.
- [23] B. B. Dong, G. F. Zheng, H. Zhao, D. Wei, and X. Y. Liu, "Mutational analysis of the GATA4 gene in patients with congenital heart disease," *International Journal of Cardiovascular Disease*, vol. 39, pp. 48–52, 2012.
- [24] Y. Q. Yang, Y. Q. Tang, and X. Y. Liu, "A novel GATA4 mutation leading to congenital heart disease ventricular septal defect," *Chinese Journal of Medical Genetics*, vol. 27, pp. 512–516, 2010.
- [25] X. Y. Liu, Y. Q. Yang, J. Ma, X. P. Lin, and Y. H. Chen, "Novel GATA4 mutations in patients with congenital ventricular septal defects," *Chinese Journal of Laboratory Medicine*, vol. 33, pp. 343–347, 2010.
- [26] X. P. Lin and Y. H. Chen, "GATA4 mutations and sporadic congenital heart disease," *Journal of Tongji University*, vol. 30, pp. 9–13, 2009.
- [27] M. W. Chen, *Screening and Functional Analysis of GATA-4 Mutations in Congenital Cardiac Septal Defects*, Guangxi Medical University, 2009.
- [28] Z.-H. Tang, L. Xia, W. Chang et al., "Two novel missense mutations of GATA4 gene in Chinese patients with sporadic congenital heart defects," *Zhonghua Yi Xue Yi Chuan Xue Za Zhi*, vol. 23, no. 2, pp. 134–137, 2006.
- [29] W. Zhang, L. Shen, Z. Deng et al., "Novel missense variants of ZFPM2/FOG2 identified in conotruncal heart defect patients do not impair interaction with GATA4," *PLoS ONE*, vol. 9, no. 7, Article ID e102379, 2014.
- [30] Z.-P. Tan, C. Huang, Z.-B. Xu, J.-F. Yang, and Y.-F. Yang, "Novel ZFPM2/FOG2 variants in patients with double outlet right ventricle," *Clinical Genetics*, vol. 82, no. 5, pp. 466–471, 2012.
- [31] A. De Luca, A. Sarkozy, R. Ferese et al., "New mutations in ZFPM2/FOG2 gene in tetralogy of Fallot and double outlet right ventricle," *Clinical Genetics*, vol. 80, no. 2, pp. 184–190, 2011.
- [32] Y. Guo, "Clinical value of heart color-ultrasonography combined measurement of serum cTnI, IGF-I and IGFBP-3 levels in patients with congenital heart diseases," *Journal of Radioimmunology*, vol. 24, no. 5, pp. 500–502, 2011.
- [33] H. Sha, Q. Li, W. Qian, H. Li, and G. Q. Chen, "Evaluation of cardiac function by H-FABP, NT-proBNP, cTnI in children with congenital heart diseases and pneumonia," *Chinese Journal of Pediatrics*, vol. 29, pp. 939–941, 2011.
- [34] A. Uner, M. Doğan, M. Ay, and Ç. Acar, "The evaluation of serum N-terminal prohormone brain-type natriuretic peptide, troponin-I, and high-sensitivity C-reactive protein levels in children with congenital heart disease," *Human and Experimental Toxicology*, vol. 33, no. 11, pp. 1158–1166, 2014.
- [35] F.-J. Zhou, C.-Y. Zhou, Y.-J. Tian et al., "Diagnostic value of analysis of H-FABP, NT-proBNP, and cTnI in heart function in children with congenital heart disease and pneumonia," *European Review for Medical and Pharmacological Sciences*, vol. 18, no. 10, pp. 1513–1516, 2014.
- [36] H. Tomita, M. Takamuro, W. Soda, K. Hatakeyama, and H. Tsutsumi, "Increased serum high-sensitivity C-reactive protein is related to hypoxia and brain natriuretic peptide in congenital heart disease," *Pediatrics International*, vol. 50, no. 4, pp. 436–440, 2008.
- [37] X. F. Guang, H. L. Dai, C. H. Zhang et al., "Study of the changes of serum BNP, VEGF, and hs-CRP in patients with congenital heart disease combined with pulmonary arterial hypertension," *Chinese Journal of Clinicians*, vol. 6, pp. 891–893, 2012.

- [38] S. Yang, *Clinical Research of the Value of Measurement of Plasma Brain Natriuretic Peptide and High Sensitive C-Reactive Protein in Congenital Heart Disease*, Hebei Medical University, 2012.
- [39] X. J. Zhao, Y. L. Geng, Z. L. Chen, G. X. Wang, J. Zhao, and X. B. Li, "Clinical analysis of serum hs-CRP and BNP levels of children with blue type congenital heart disease," *Hebei Medical Journal*, vol. 36, pp. 1790–1792, 2014.
- [40] W. Li, *Clinical significance and changes in copeptin, brain natriuretic peptide and C-reactive protein concentration on peripheral blood in children with congenital heart disease combined with heart failure [M.S. thesis]*, Qingdao University, Qingdao, China, 2014.
- [41] X. N. Wang, S. B. Yang, X. Y. Tian et al., "Clinical value of C-reactive protein combined with B-type natriuretic peptide in evaluation of cardiac function of congenital heart disease children with pulmonary infections," *Chinese Journal of Nosocomiology*, vol. 24, pp. 1263–1264, 2014.
- [42] X. L. Lin and H. P. Liu, "Clinical analysis of C-reactive protein and B-type natriuretic peptide assessment of cardiac function in congenital heart disease children with lung infections," *The Journal of Medical Theory & Practice*, vol. 27, pp. 2001–2002, 2014.
- [43] Y. Qiu, X. L. Huang, J. Wu, H. Zheng, W. S. Wu, and Z. Lin, "Plasma brain natriuretic peptide and Tei index assessment of heart function in children with left to right congenital heart defects," *Maternal and Child Health Care of China*, vol. 22, pp. 2209–2211, 2007.
- [44] B. Y. Meng, T. Wang, Q. Zhang et al., "Brain natriuretic peptide and Tei-index for assessing preoperative left heart function in infants with congenital heart diseases," *South China Journal of Cardiovascular Diseases*, vol. 16, pp. 378–381, 2010.
- [45] J. Hu, H. Sha, and Q. Li, "The change of serum levels of B-type natriuretic peptides and the correlation to children with congenital heart disease and serious pneumonia," *Chinese Medicine*, vol. 6, pp. 1402–1403, 2011.
- [46] X. Li, M. Y. Yang, J. T. Qin, Y. Lin, D. Wang, and C. Y. Wang, "Relationship of serum N-terminal pro-brain natriuretic peptide and pulmonary artery pressure in pediatric patients with congenital heart disease," *Hainan Medical Journal*, vol. 23, no. 1, pp. 84–85, 2012.
- [47] X. Li, M. Y. Yang, J. T. Qin et al., "Clinical significance of changes of plasma N-terminal pro-brain natriuretic peptide in children with congenital heart disease combined with heart failure," *Journal of Applied Clinical Pediatrics*, vol. 27, pp. 19–20, 2012.
- [48] H. L. Wang and X. R. A, "Clinical significance of plasma BNP testing of children with congenital heart disease in plateau area," *Laboratory Medicine*, vol. 28, pp. 865–866, 2013.
- [49] M. Kavga, G. Varlamis, A. Giannopoulos et al., "Correlation of plasma B-type natriuretic peptide with shunt volume in children with congenital heart disease involving left-to-right shunt," *Hellenic Journal of Cardiology*, vol. 54, no. 3, pp. 192–198, 2013.
- [50] J. M. Fang, W. Guo, J. Jiang, L. F. Yu, and Y. Y. Guo, "Symptoms and clinical significance of N-terminal pro-brain natriuretic peptide in children suffering from congenital heart disease combined with heart failure," *China Medical Herald*, vol. 10, pp. 43–45, 2013.
- [51] C.-W. Lin, X.-L. Zeng, X.-H. Meng, S.-H. Jiang, and H. Ou-Yang, "Clinical importance of preoperative measurement of plasma amino-terminal pro-B-type natriuretic peptide in infants with congenital heart disease," *Chinese Journal of Contemporary Pediatrics*, vol. 16, no. 1, pp. 40–43, 2014.
- [52] D. B. McElhinney, E. Geiger, J. Blinder, D. W. Benson, and E. Goldmuntz, "NKX2.5 mutations in patients with congenital heart disease," *Journal of the American College of Cardiology*, vol. 42, no. 9, pp. 1650–1655, 2003.
- [53] V. Garg, I. S. Kathiriya, R. Barnes et al., "GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5," *Nature*, vol. 424, no. 6947, pp. 443–447, 2003.
- [54] R. Geiger, A. Hammerer-Lercher, C. Url et al., "NT-proBNP concentrations indicate cardiac disease in pediatric patients," *International Journal of Cardiology*, vol. 123, no. 1, pp. 63–65, 2007.
- [55] E. Akhabue, J. Thiboutot, J.-W. Cheng et al., "New and emerging risk factors for coronary heart disease," *The American Journal of the Medical Sciences*, vol. 347, no. 2, pp. 151–158, 2014.
- [56] A. Bennet, E. Di Angelantonio, S. Erqou et al., "Lipoprotein(a) levels and risk of future coronary heart disease: large-scale prospective data," *Archives of Internal Medicine*, vol. 168, no. 6, pp. 598–608, 2008.
- [57] P. R. Kamstrup, M. Benn, A. Tybjaerg-Hansen, and B. G. Nordestgaard, "Extreme lipoprotein(a) levels and risk of myocardial infarction in the general population: the Copenhagen City Heart Study," *Circulation*, vol. 117, no. 2, pp. 176–184, 2008.
- [58] Y. Momiyama, R. Ohmori, Z. A. Fayad et al., "Associations between serum lipoprotein(a) levels and the severity of coronary and aortic atherosclerosis," *Atherosclerosis*, vol. 222, no. 1, pp. 241–244, 2012.
- [59] Y. Jiang, K. Guo, M. Chen, J. Bao, C. Shen, and Y. Li, "Serum lipoprotein(a) positively correlates with coronary artery calcification in low-risk Chinese Han patients: a study from a single center," *PLoS ONE*, vol. 8, no. 8, Article ID e71673, 2013.
- [60] F. Charron, P. Paradis, O. Bronchain, G. Nemer, and M. Nemer, "Cooperative interaction between GATA-4 and GATA-6 regulates myocardial gene expression," *Molecular and Cellular Biology*, vol. 19, no. 6, pp. 4355–4365, 1999.
- [61] M. Salazar, F. Consoli, V. Villegas et al., "Search of somatic GATA4 and NKX2.5 gene mutations in sporadic septal heart defects," *European Journal of Medical Genetics*, vol. 54, no. 3, pp. 306–309, 2011.
- [62] T. M. Chlon and J. D. Crispino, "Combinatorial regulation of tissue specification by GATA and FOG factors," *Development*, vol. 139, no. 21, pp. 3905–3916, 2012.
- [63] F. Malek, J. Dvorak, J. Svitil et al., "Correlation of lipoprotein(a) concentration with the extent of coronary artery disease in patients on lipid lowering therapy," *Neuro Endocrinology Letters*, vol. 2, pp. 55–59, 2012.
- [64] M. Greif, T. Arnoldt, F. von Ziegler et al., "Lipoprotein(a) is independently correlated with coronary artery calcification," *European Journal of Internal Medicine*, vol. 24, no. 1, pp. 75–79, 2013.
- [65] Z.-G. Li, G. Li, Y.-L. Zhou et al., "Lack of association between lipoprotein(a) genetic variants and subsequent cardiovascular events in Chinese Han patients with coronary artery disease after percutaneous coronary intervention," *Lipids in Health and Disease*, vol. 12, article 127, 2013.
- [66] H. Konishi, K. Miyauchi, T. Kasai et al., "Impact of lipoprotein(a) as residual risk on long-term outcomes in patients after percutaneous coronary intervention," *American Journal of Cardiology*, vol. 115, no. 2, pp. 157–160, 2015.

- [67] V. Gudnason, “Lipoprotein (a): a causal independent risk factor for coronary heart disease?” *Current Opinion in Cardiology*, vol. 24, no. 5, pp. 490–495, 2009.
- [68] A. von Zychlinski, T. Kleffmann, M. J. A. Williams, and S. P. McCormick, “Proteomics of Lipoprotein(a) identifies a protein complement associated with response to wounding,” *Journal of Proteomics*, vol. 74, no. 12, pp. 2881–2891, 2011.

## Research Article

# Diagnosing Parkinson's Diseases Using Fuzzy Neural System

**Rahib H. Abiyev<sup>1</sup> and Sanan Abizade<sup>2</sup>**

<sup>1</sup>*Department of Computer Engineering, Applied Artificial Intelligence Research Centre, Near East University, Lefkosa, Northern Cyprus, Mersin 10, Turkey*

<sup>2</sup>*Department of Electrical and Electronic Engineering, Applied Artificial Intelligence Research Centre, Near East University, Lefkosa, Northern Cyprus, Mersin 10, Turkey*

Correspondence should be addressed to Rahib H. Abiyev; [rahib.abiyev@neu.edu.tr](mailto:rahib.abiyev@neu.edu.tr)

Received 30 September 2015; Accepted 14 December 2015

Academic Editor: Reza Khosrowabadi

Copyright © 2016 R. H. Abiyev and S. Abizade. 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 study presents the design of the recognition system that will discriminate between healthy people and people with Parkinson's disease. A diagnosing of Parkinson's diseases is performed using fusion of the fuzzy system and neural networks. The structure and learning algorithms of the proposed fuzzy neural system (FNS) are presented. The approach described in this paper allows enhancing the capability of the designed system and efficiently distinguishing healthy individuals. It was proved through simulation of the system that has been performed using data obtained from UCI machine learning repository. A comparative study was carried out and the simulation results demonstrated that the proposed fuzzy neural system improves the recognition rate of the designed system.

## 1. Introduction

In the world, many people suffer from Parkinson's disease (PD). The disease more often appeared after the age of 60 [1]. Parkinson's disease is a chronic disorder of central nervous system which causes the death of the nervous cell in the brain. Parkinson's disease is progressive and the number of people suffering from the disease is expected to rise. The disease usually happens slowly and persists over a long period of time.

The symptoms of the PD continue and worsen over time. The basic symptoms of PD are movement related symptoms. These are tremor, rigidity or stiffness of the limbs and trunk, bradykinesia or slow movement, and problems with balance or walking [2, 3]. Tremor is a basic symptom which may affect shaking or trembling of legs, arms, hands, jaw, or face. The patients may have difficulty talking, walking, or completing some other simple tasks as these symptoms become more pronounced. Other symptoms are related to the behavioural problems, depression, thinking, sleep, and emotional problems. A person with Parkinson's may have a trouble in speaking and swallowing and chewing problems. Especially in advanced stages of the disease nonmotor features, such as dementia and dysautonomia, occur frequently.

The diagnosis and timely treatments are important in order to manage its symptoms. The diagnosis is based on neurological examination and medical history of patients. The diagnosis of the disease in the early stages is difficult [3]. Diagnosis of PD depends on the presence of two or more of the above symptoms.

Vocal symptoms that include impairment of vocal sounds (*dysphonia*) and problems with the normal articulation of speech (*dysarthria*) are important in diagnosis of PD [4]. The research paper [5] shows that the most important symptom of PD is dysphonia. The dysphonia is the disorder of voice. Dysphonic symptoms typically include reduced loudness and roughness and breathiness and decreased energy in the higher parts of the harmonic spectrum and exaggerated vocal tremor. The treatment of these symptoms is difficult for the people having Parkinson's disease. In [4–6] it was shown that approximately 90% of people with Parkinson's disease have dysphonia. Dysphonia includes any pathological or functional problem with voice [6]. The voice will sound hoarse, strained, or effortful. It may be difficult to understand the voice of people having PD. The used method for diagnosis of Parkinson's disease (PD) is basically based on speech measurement for general voice disorders [4, 7–9].

Specialists doctors need to make an analysis of many factors for accurate diagnose of PD. Usually, decisions made are based on evaluating the current test results of patients. The problem becomes too difficult if the number of attributes that the specialist wants to evaluate is high. Recently various computational tools have been developed in order to improve the accuracy of diagnosis of PD. These tools have provided excellent help to the doctors and medical specialists in making decisions about the patients. Different Artificial Intelligence (AI) techniques, expert systems, and decision making systems are designed for diagnosis or classification of diseases. They were potential and good supportive tools for the expert/doctor. The development of efficient recognition systems in medical diagnosis is becoming more important. Nowadays various Artificial Intelligence techniques such as expert systems, fuzzy systems, and neural networks are actively applied for diagnosis of Parkinson's diseases using voice signals. Reference [4] introduces a new measure of dysphonia, pitch period entropy (PPE), which is robust to many uncontrollable confounding effects including noisy acoustic environments and separates healthy people from the people having PD. Nonlinear dynamical systems theory [4, 10] and statistical learning theory, such as *linear discriminant analysis* (LDA) and *support vector machines* (SVMs) [5, 11], are preferred for classification of healthy people or those with PD and discriminate the healthy people on the basis of measures of dysphonia. Different techniques, such as SVM [12], SVM with RBF (radial based function) kernel [13], SVM with Multiple Layer Perceptron (MLP), and a Radial Basis Function Network (RBFN) [14], are used for diagnosis of PD. In [15] integration of Kohonen self-organizing map (KSOM) and least squares support vector machine (LS-SVM), and in [3, 16] nonlinear time series analysis tools are applied for diagnosing of PD. Reference [17] uses fuzzy *c*-means algorithm, [18] uses four independent classification schemas, neural networks, DMneural, regression, and decision tree for classification purpose, and a comparative study was carried out.

The above methods are used in order to increase classification accuracy of PD. Classification systems can help in increasing the accuracy and reliability of diagnoses and minimizing possible errors, as well as making the diagnoses more time efficient. Success in the discovery of knowledge depends on the ability to explore different classes of specific data and to apply appropriate methods in order to extract the main features. This paper deals with the application of fusion of fuzzy systems and neural networks for designing of the recognition system of PD.

Fuzzy systems can handle uncertainties associated with information or data in the knowledge bases [19] and are widely used to solve different real world problems. Fuzzy system uses data and knowledge specific to chaotic dynamics of the process and increases the performance of the system. In the literature, different neural and fuzzy structures are proposed for solving various problems [20–26]. In [22, 23] clustering algorithm and gradient descent algorithm are applied for the design of multi-input and single output FNS. Well known ANFIS (adaptive neurofuzzy inference system) structure is used for solving cervical cancer recognition [27],

for optimizing the chiller loading [28], and for distinguishing ESES (electrical status epilepticus) and normal EEG (electroencephalography) signals [29]. The use of multiple ANFIS structures, in [27], leads to the increase of the number of parameters of the network. In these papers the used systems are designed for special purpose and most of them are basically based on Mamdani type rules. Performances of these systems are determined by measuring classification rate. In this paper, in order to improve the performance of classification system, a multi-input and multioutput fuzzy neural system (FNS) based on TSK rules is proposed for identification of the PD.

The paper is organized as follows. Section 2 describes the structure of proposed fuzzy neural system used for recognition of PD. The parameter update rule of the proposed system is presented in Section 3. Section 4 describes the simulation results. The conclusions are given in Section 5.

## 2. FNS Based Recognition

The fuzzy neural system combines the learning capabilities of neural networks with the linguistic rule interpretation of fuzzy inference systems. The design of FNS includes the development of the fuzzy rules that have if-then form. This can be achieved by dint of optimal definition of the premise and consequent parts of fuzzy if-then rules for the classification system through the training capability of neural networks. The two basic types of if-then rules used in fuzzy systems are Mamdani and Takagi-Sugeno-Kang (TSK) type fuzzy rules. The first type consists of rules, whose antecedent and consequent parts utilize fuzzy values. The second one uses the fuzzy rules that have fuzzy antecedent and crisp consequent parts. In the paper we use TSK type fuzzy rules for system design. The second type of fuzzy system approximates nonlinear system with linear systems and has the following form:

$$\begin{aligned} &\text{If } x_1 \text{ is } A_{1j} \text{ and } x_2 \text{ is } A_{2j} \text{ and } \dots \text{ and } x_m \text{ is } A_{mj} \\ &\text{Then } y_j \text{ is } \sum_{i=1}^m a_{ij}x_i + b_j, \end{aligned} \quad (1)$$

where  $x_i$  and  $y_j$  are input and output signals of the system, respectively,  $i = 1, \dots, m$  is the number of input signals, and  $j = 1, \dots, r$  is the number of rules.  $A_{ij}$  are input fuzzy sets;  $b_j$  and  $a_{ij}$  are coefficients.

The structure of fuzzy neural networks used for the classification of PDs is based on TSK type fuzzy rules and is given in Figure 1. The FNS includes six layers. In the first layer,  $x_i$  ( $i = 1, \dots, m$ ) input signals are distributed. The second layer includes membership functions. Here each node corresponds to one linguistic term. Here, for each input signal entering the system, the membership degree to which input value belongs to a fuzzy set is calculated. The Gaussian membership function is used in order to describe linguistic terms:

$$\mu_{1j}(x_i) = e^{-(x_i - c_{ij})^2 / \sigma_{ij}^2}, \quad i = 1, \dots, m, \quad j = 1, \dots, r, \quad (2)$$

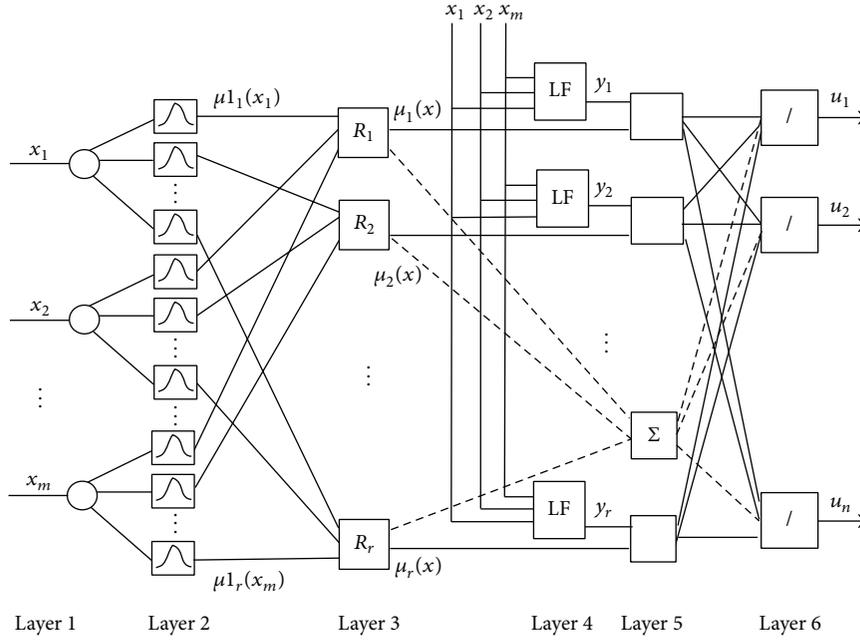


FIGURE 1: Classifier based on FNS.

where  $c_{ij}$  and  $\sigma_{ij}$  are center and width of the Gaussian membership functions, respectively, and  $\mu_{1j}(x_i)$  is membership function of  $i$ th input variable for  $j$ th term.

The third layer is a rule layer. Here number of nodes is equal to the number of rules. Here  $R_1, R_2, \dots, R_r$  represents the rules. The output signals of this layer are calculated using t-norm min (AND) operation:

$$\mu_j(x) = \prod_i \mu_{1j}(x_i) \quad i = 1, \dots, m, \quad j = 1, \dots, r, \quad (3)$$

where  $\prod$  is the min operation.

These  $\mu_j(x)$  signals are input signals for the fifth layer. Fourth layer is a consequent layer. It includes  $n$  linear systems. Here the output values of the rules are determined using linear functions (LF):

$$y_{1j} = \sum_{i=1}^m x_i w_{ij} + b_j. \quad (4)$$

In the fifth layer, the output signals of the third layer are multiplied by the output signals of the fourth layer. The output of  $j$ th node is calculated as  $y_j = \mu_j(x) \cdot y_{1j}$ .

The output signals of FNS are determined as

$$u_k = \frac{\sum_{j=1}^r w_{jk} y_j}{\sum_{j=1}^r \mu_j(x)}. \quad (5)$$

Here  $u_k$  are the output signals of FNS ( $k = 1, \dots, n$ ) and  $w_{jk}$  are weight coefficients of connections used between layers 5 and 6. After calculating the output signal, the training of the network starts.

### 3. Parameter Updates

**3.1. Fuzzy Classification.** The design of FNS (Figure 1) includes determination of the unknown parameters of the antecedent and the consequent parts of the fuzzy if-then rules (1). In fuzzy rules the antecedent part represents the input space by dividing the space into a set of fuzzy regions and the consequent part describes the system behaviour in those regions.

As mentioned above, recently a number of different approaches have been used for designing fuzzy if-then rules. Some of them are based on clustering [20–24, 26], the least squares method (LSM) [20, 22, 30], gradient algorithms [14, 20–23, 26], genetic algorithms [24, 25, 28], and particle swarm optimization (PSO) [31].

In this paper, fuzzy clustering and gradient technique are used for the design of FNS. At first the fuzzy clustering is used to design the antecedent (premise) parts, and then gradient algorithm is used to design the consequent parts of the fuzzy rules. Fuzzy clustering is an efficient technique for constructing the antecedent structures. The aim of clustering methods is to identify a certain group of data from a large data set, such that a concise representation of the behaviour of the system is produced. Each cluster center can be translated into a fuzzy rule for identifying the class. Different clustering algorithms are developed [32–34]. Recently fuzzy c-means [32] and subtractive clustering [33, 34] algorithms have been developed for fuzzy systems. Subtractive is unsupervised clustering [33] which is an extension of the grid based mountain clustering [34]. Here the number of clusters for input data points is determined by the clustering algorithm. Sometimes we need to control the number of clusters in an input space. In these cases, the supervised clustering

algorithms are of primary concern. Fuzzy  $c$ -means clustering is one of them. It can efficiently be used for fuzzy systems [32] with a simple structure and sufficient accuracy. In this paper, the fuzzy  $c$ -means (FCM) clustering technique is used for structuring the premise part of the fuzzy system.

Learning of FNS starts with the update of parameters of antecedent part of if-then rules, that is, the parameters of the second layer of FNS. For this aim FCM classification is applied in order to partition input space and construct antecedent part of fuzzy if-then rules. The following objective function is used in FCM algorithm:

$$J_q = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^q d_{ij}^2, \quad \text{where } d_{ij} = \|x_i - c_j\|, \quad 1 \leq q < \infty, \quad (6)$$

where  $q$  is any real number greater than 1,  $u_{ij}$  is the degree of membership of  $x_i$  in the cluster  $j$ ,  $x_i$  is the  $i$ th of  $d$ -dimensional measured data,  $c_j$  is the  $k$ -dimension center of the cluster, and  $\| * \|$  is any norm expressing the similarity between any measured data and the cluster centers.

The fuzzy classification of input data is carried out through an iterative optimization of the objective function (6), with the update of membership  $u_{ij}$  and the cluster centers  $c_j$ . The algorithm consists of the following steps:

- (1) Initialize  $U = [u_{ij}]$  matrix,  $U^{(0)}$ .
- (2) Calculate the centers vectors  $C^{(t)} = [c_j]$  with  $U^{(t)}$ :

$$c_j = \frac{\left( \sum_{i=1}^N u_{ij}^q \cdot x_i \right)}{\sum_{i=1}^N u_{ij}^q}. \quad (7)$$

- (3) Update  $U^{(t)}$  and  $U^{(t+1)}$ :

$$u_{ij} = \frac{1}{\sum_{k=1}^C (d_{ik}/d_{jk})^{2/(q-1)}}. \quad (8)$$

- (4) If  $\{|U^{(t+1)} - U^{(t)}|\} < \varepsilon$  then stop; otherwise set  $t = t + 1$  and return to Step (2).

In the results of partitioning the cluster centers are determined. These cluster centers will correspond to the centers of the membership functions used in the input layer of FNS. The width of the membership function is determined using the distance between cluster centers.

After the design of the antecedents parts by fuzzy clustering, the parameter update rules are derived for training the parameters of the consequent parts of the fuzzy rules. In the paper, we applied gradient learning with adaptive learning rate. The adaptive learning rate guarantees the convergence and speeds up the learning of the network.

**3.2. Learning Using Gradient Descent.** At the beginning, the parameters of the FNS are generated randomly. To generate a proper FNS model, the training of the parameters has been carried out. For generality, we have given the learning procedure of all parameters of FNS using gradient descent

algorithm. The parameters are the membership function of linguistic values in the second layer of the network and the parameters of the fourth and fifth layers. In the design of FNS cross validation technique is used for separation of the data into training and testing set. Training includes the adjusting of the parameter values. In this paper, a gradient learning with adaptive learning rate is applied for the update of parameters. The adaptive learning rate guarantees the convergence and speeds up the learning of the network. In addition, the momentum is used to speed up the learning processes.

The error on the output of the network is calculated as

$$E = \frac{1}{2} \sum_{k=1}^n (u_k^d - u_k)^2. \quad (9)$$

Here  $n$  is the number of output signals of the network,  $u_k^d$  and  $u_k$  are desired and current output values of the network ( $k = 1, \dots, n$ ), respectively. The parameters  $w_{jk}, a_{ij}, b_j$  ( $i = 1, \dots, m, j = 1, \dots, r, k = 1, \dots, n$ ) in consequent part of network and the parameters of membership functions  $c_{ij}$  and  $\sigma_{ij}$  ( $i = 1, \dots, m, j = 1, \dots, r$ ) in the premise part of FNS are adjusted using the following formulas:

$$\begin{aligned} w_{jk}(t+1) &= w_{jk}(t) - \gamma \frac{\partial E}{\partial w_{jk}} \\ &\quad + \lambda (w_{jk}(t) - w_{jk}(t-1)); \\ a_{ij}(t+1) &= a_{ij}(t) - \gamma \frac{\partial E}{\partial a_{ij}} + \lambda (a_{ij}(t) - a_{ij}(t-1)); \\ b_j(t+1) &= b_j(t) - \gamma \frac{\partial E}{\partial b_j} + \lambda (b_j(t) - b_j(t-1)); \\ c_{ij}(t+1) &= c_{ij}(t) - \gamma \frac{\partial E}{\partial c_{ij}} + \lambda (c_{ij}(t) - c_{ij}(t-1)); \\ \sigma_{ij}(t+1) &= \sigma_{ij}(t) - \gamma \frac{\partial E}{\partial \sigma_{ij}} + \lambda (\sigma_{ij}(t) - \sigma_{ij}(t-1)); \end{aligned} \quad (10)$$

$$i = 1, \dots, m; \quad j = 1, \dots, r; \quad k = 1, \dots, n. \quad (11)$$

Here  $\gamma$  is the learning rate,  $\lambda$  is the momentum,  $m$  is the number of input signals of the network (input neurons) and  $r$  is the number of fuzzy rules (hidden neurons), and  $n$  is the number of output neurons.

The derivatives in (10) are computed using the following formulas:

$$\begin{aligned} \frac{\partial E}{\partial w_{jk}} &= \frac{\partial E}{\partial u_k} \frac{\partial u_k}{\partial w_{jk}} = \frac{(u_k(t) - u_k^d(t)) \cdot y_{1j}}{\sum_{j=1}^n \mu_j}, \\ \frac{\partial E}{\partial a_{ij}} &= \frac{\partial E}{\partial u_k} \frac{\partial u_k}{\partial y_{1j}} \frac{\partial y_{1j}}{\partial a_{ij}} \\ &= \frac{\sum_k (u_k(t) - u_k^d(t)) \cdot w_{kj} \mu_j x_i}{\sum_{j=1}^n \mu_j}, \end{aligned}$$

$$\begin{aligned} \frac{\partial E}{\partial b_j} &= \frac{\partial E}{\partial u_k} \frac{\partial u_k}{\partial y_{1j}} \frac{\partial y_{1j}}{\partial y_j} \frac{\partial y_j}{\partial b_j} \\ &= \frac{\sum_k (u_k(t) - u_k^d(t)) \cdot w_{kj} \mu_j}{\sum_{j=1}^n \mu_j}, \\ &\text{here } i = 1, \dots, m, \quad j = 1, \dots, r, \quad k = 1, \dots, n. \end{aligned} \quad (12)$$

The derivatives in (11) are determined by the following formulas:

$$\begin{aligned} \frac{\partial E}{\partial c_{ij}} &= \sum_k \frac{\partial E}{\partial u_k} \frac{\partial u_k}{\partial \mu_j} \frac{\partial \mu_j}{\partial c_{ij}}, \\ \frac{\partial E}{\partial \sigma_{ij}} &= \sum_k \frac{\partial E}{\partial u_k} \frac{\partial u_k}{\partial \mu_j} \frac{\partial \mu_j}{\partial \sigma_{ij}}. \end{aligned} \quad (13)$$

Here  $i = 1, \dots, m$ ,  $j = 1, \dots, r$ ,  $k = 1, \dots, n$ . Consider

$$\begin{aligned} \frac{\partial E}{\partial u_k} &= u_k(t) - u_k^d(t); \\ \frac{\partial u_k}{\partial \mu_j} &= \frac{y_j - u_k}{\sum_{j=1}^n \mu_j}; \\ \frac{\partial \mu_j(x_i)}{\partial c_{ij}} &= \mu_j(x_i) \frac{2(x_i - c_{ij})}{\sigma_{ij}^2}; \\ \frac{\partial \mu_j(x_i)}{\partial \sigma_{ij}} &= \mu_j(x_i) \frac{2(x_i - c_{ij})^2}{\sigma_{ij}^3}. \end{aligned} \quad (14)$$

Using equations (12)–(14), the derivatives in (10) and (11) are calculated and the correction of the parameters of FNS is carried out.

Convergence is very important problem in learning of FNS model. The convergence of the learning algorithm using gradient descent depends on the selection of the initial values of the learning rate. Usually, the initial value of learning rate is selected in the interval [0-1]. A large value of the learning rate may lead to unstable learning; a small value of the learning rate results in a slow learning speed. In the paper an adaptive approach is applied for updating these parameters. The learning of the FNS parameters is started with a small value of the learning rate  $\gamma$ . During learning,  $\gamma$  is increased if the value of change of error  $\Delta E = E(t) - E(t+1)$  is positive and decreased if negative. This strategy ensures a stable learning for the FNS. In addition a momentum term is used to speed up learning processes. The optimal value of the learning rate for each time instance can be obtained using a Lyapunov function [22, 23]. The derivation of the convergence is given in [22, 23].

#### 4. Simulation Studies

The FNS, described above, is applied for classification of Parkinson's diseases. The people are divided into two classes:

normal and PD. For this aim, the database is taken from University of California at Irvine (UCI) machine learning repository. The data set is donated from hospitals and it has been studied by many researchers. The data set includes biomedical voice measurements of 31 people; 23 were diagnosed with PD. Each row contains the value of the 23 voice parameters. Each column contains 195 items of data for each parameter. The main aim of the data is to discriminate healthy people from those with PD. The parameters that are used for recognition of PD are given in Table 1. These are the parameters of the voice signals recorded directly on the computer using Computerized Speech Laboratory. During modelling the preprocessing have been done on the input data and the input data are normalized in the interval of [0, 1]. The scaling operation helps and makes the training process of the system easy. After normalization, these data are entered as an input signal to the FNS.

To design classification model the FNS structure with 23 input and 2 output neurons is generated first. If we use traditional neurofuzzy structure (e.g., [20] or [26]) for 23 inputs and 2 cluster centers,  $\text{pow}(2,23) = 8383608$  rules should be generated. The rules are constructed using all possible combinations of inputs and cluster centers. This is very large number. In this paper the number of rules is selected according to the clustering results, equal to cluster centers.

In the design of FNS, the fuzzy classification is applied in order to partition input space and select the parameters of the premise parts, that is, the parameters of Gaussian membership functions used in the second layer of FNS. FCM clustering is used for the input space with 16 clusters for each input. 16 fuzzy rules are constructed using a different combination of these clusters for 22 inputs. After clustering input space gradient decent algorithm is used for learning of consequent parts of the fuzzy rules, that is, parameters of the 4th layer of FNS. Learning is implemented using cross validation. Cross validation generalizes two independent data sets: training and testing. It is applied to find accurate model of classifier. In the paper 10-fold cross validation is used for separation of the data into training and testing set and for evaluation of classification accuracy. There should be set of experiments in order to achieve required accuracy in the FNS output. The simulation is performed using different number of neurons in hidden layer. The design steps of FNS for the diagnosing PD are given below:

- (1) Read PD data set. Select input and output (target) signals from statistical data. Apply normalization.
- (2) Enter the values of learning rate and momentum. Set the number of clusters. Generate network parameters. Set a maximal number of epochs for learning.
- (3) Apply classification algorithm to the input signals and determine the cluster centers.
- (4) Use cluster centers to determine the centers of membership functions of layer 2.
- (5) Use the centers of membership functions to determine the widths of membership functions.

TABLE I: List of measurement methods applied to acoustic signals recorded from each subject.

Name	ASCII subject name and recording number
MDVP:Fo (Hz)	Average vocal fundamental frequency
MDVP:Fhi (Hz)	Maximum vocal fundamental frequency
MDVP:Flo (Hz)	Minimum vocal fundamental frequency
MDVP:Jitter (%)	
MDVP:Jitter (Abs)	
MDVP:RAP	Five measures of variation in fundamental frequency
MDVP:PPQ	
Jitter:DDP	
MDVP:Shimmer	
MDVP:Shimmer (dB)	
Shimmer:APQ3	Six measures of variation in amplitude
Shimmer:APQ5	
MDVP:APQ	
Shimmer:DDA	
NHR	
HNR	Two measures of ratio of noise to tonal components in the voice
RPDE	
D2	Two nonlinear dynamical complexity measures
DFA	Signal fractal scaling exponent
Spread1	
Spread2	Three nonlinear measures of fundamental frequency variation
PPE	
Status	Health status of the subject: one, Parkinson's; zero, healthy

- (6) Using input statistical data define a random partition for 10-fold cross validation.
- (7) Initialize current number of learning epochs to 1.
- (8) Use PD data set and cross validation and determine training and testing data sets.
- (9) Determine the numbers of rows in training and testing data sets.
- (10) Initialize the number of iterations to 1.
- (11) According to the number of iterations select input data from training data set and send them to the input of FNS.
- (12) Calculate network outputs.
- (13) Determine the values of errors using network output and target output signals. Use these error values to compute the sum of the squared errors (SSE).
- (14) Using error values update the network parameters (learning of network).
- (15) Apply adaptive strategy for updating the learning rate using current and previous values of SSE.
- (16) Compute sum of SSE obtained on each iteration and save as the training error. Repeat Steps (11)–(16) for other remaining training data sets. If the current number of iterations will be less than a number of rows in the training set then go to Step (11), otherwise go to Step (17).
- (17) Select test data set.
- (18) Set number of iterations to 1.
- (19) According to the number of iterations select input data from test data set and send them to the input of FNS.
- (20) Compute the output of FNS.
- (21) Determine the values of errors using network output and target output signals. Compute SSE on the output of the network.
- (22) Compute sum of SSE obtained on each iteration of the loop and save as the testing error. Repeat Steps (19)–(22) for other remaining test data sets.
- (23) Check the value of testing error with the value of testing error obtained in the previous epoch. If the current error value is less than the previous one then go to Step (24), otherwise go to Step (25).
- (24) Save the parameters of the network. Save the values of training and testing errors.
- (25) Use the sum of SSE to find root mean squared error (RMSE). Print the values of testing and training errors; increment the epochs number.
- (26) Check a current number of epochs for the continuation of the learning process. If this number is less than the maximal number of epochs then repeat Steps (8)–(26). Otherwise go to Step (27).
- (27) Print the values of training and testing errors obtained in Step (24).
- (28) Stop the training.

TABLE 2: Fragment from PD data set.

MDVP:Fo (Hz)	119.99200	122.4000	236.20000	237.32300	260.10500	197.56900	151.73700	148.7900
MDVP:Fhi (Hz)	157.30200	148.6500	244.66300	243.70900	264.91900	217.62700	190.20400	158.3590
MDVP:Flo (Hz)	74.99700	113.8190	102.13700	229.25600	237.30300	90.79400	129.85900	138.9900
MDVP:Jitter (%)	0.00784	0.00968	0.00277	0.00303	0.00339	0.00803	0.00314	0.00309
MDVP:Jitter (Abs)	0.00007	0.00008	0.00001	0.00001	0.00001	0.00004	0.00002	0.00002
MDVP:RAP	0.00370	0.00465	0.00154	0.00173	0.00205	0.00490	0.00135	0.00152
MDVP:PPQ	0.00554	0.00696	0.00153	0.00159	0.00186	0.00448	0.00162	0.00186
Jitter:DDP	0.01109	0.01394	0.00462	0.00519	0.00616	0.01470	0.00406	0.00456
MDVP:Shimmer	0.04374	0.06134	0.02448	0.01242	0.02030	0.02177	0.01469	0.01574
MDVP:Shimmer (dB)	0.42600	0.62600	0.21700	0.11600	0.19700	0.18900	0.13200	0.14200
Shimmer:APQ3	0.02182	0.03134	0.01410	0.00696	0.01186	0.01279	0.00728	0.00839
Shimmer:APQ5	0.03130	0.04518	0.01426	0.00747	0.01230	0.01272	0.00886	0.00956
MDVP:APQ	0.02971	0.04368	0.01621	0.00882	0.01367	0.01439	0.01230	0.01309
Shimmer:DDA	0.06545	0.09403	0.04231	0.02089	0.03557	0.03836	0.02184	0.02518
NHR	0.02211	0.01929	0.00620	0.00533	0.00910	0.01337	0.00570	0.00488
HNR	21.03300	19.08500	24.07800	24.67900	21.08300	19.26900	24.15100	24.41200
RPDE	0.414783	0.458359	0.469928	0.384868	0.440988	0.372222	0.396610	0.402591
D2	0.815285	0.819521	0.628232	0.626710	0.628058	0.725216	0.745957	0.762508
DFA	-4.813031	-4.075192	-6.816086	-7.018057	-7.517934	-5.736781	-6.486822	-6.311987
Spread1	0.266482	0.335590	0.172270	0.176316	0.160414	0.164529	0.197919	0.182459
Spread2	2.301442	2.486855	2.235197	1.852402	1.881767	2.882450	2.449763	2.251553
PPE	0.284654	0.368674	0.119652	0.091604	0.075587	0.202879	0.132703	0.160306
Status	1	1	0	0	0	0	1	1

The training of input/output data for the classification system will be a structure whose first component is the twenty-three-dimension input vector and second component is the two-dimension output clusters. Table 2 depicts the fragment from PD data set. The FNS structure is generated with 23 input and two output neurons. After generation fuzzy c-means clustering and gradient descent algorithms are applied for training the parameters of FNS. In the first step, using fuzzy clustering, cluster centers are determined using the input data. These cluster centers are used to organize the membership functions of the inputs of antecedent part of each fuzzy rules. The rule layer is the second layer. The consequent parts of the fuzzy rules are organized using linear functions. Linear functions are determined in fourth layer. After clustering and designing antecedent part the learning of the parameters of consequent part starts. The initial values of the parameters  $w$  and  $b$  of linear functions of consequent part are selected in interval  $[0, 0.2]$ . The initial values of learning rate and momentum are selected as 0.02 and 0.625, correspondingly. During learning the parameters  $w$  and  $b$  of the rule are updated. In the results of learning the fuzzy rules are constructed. The clusters obtained from classification operation will be the centers of Gaussian membership functions used in antecedent parts of fuzzy rules. The consequent parts are constructed on the basis of learning of the parameters of linear functions.

The simulation results of FNS is compared with the simulation results of other models used for classification of PD. For evaluation of the outcomes of the models the Root Mean Square Error (RMSE) is used:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (u_i^d - u_i)^2}. \quad (15)$$

Here  $u_i^d$  are desired values of output and  $u_i$  are actual values of the system output.

To estimate the performance of the FNS clustering systems, the recognition rates and RMSE values of errors between clusters and current output signal are taken. RMSE is computed using formula given above. Recognition rate is computed by the number of items correctly classified divided by the total number of items:

$$\begin{aligned} \text{Recognition\_rate} \\ = \frac{\text{Number of items correctly classified}}{\text{Total number of items}} \cdot 100\%. \end{aligned} \quad (16)$$

During training of FNS, all input data are scaled to interval  $[0, 1]$ . Then fuzzy c-means clustering is applied to input data. The result of clustering is used to set up parameters of the antecedent part of fuzzy rules, that is, parameters of the second layer of FNS structure. The parameters of the consequent part of fuzzy rules are determined by applying gradient learning. The learning has been performed for 2000 epochs. The synthesis of FNS classification system is performed using different number of fuzzy rules. The training has been performed using different number of rules: 2, 5, 8, 12, and 16. Training is performed using 10-fold cross validation. In the results of training the parameters of FNS are determined. Figure 2 depicts the values of RMSE obtained during training. Once the FNS is trained then it has been used for testing. The values of RMSE obtained for train, evaluation, and test stages for FNS having 16 hidden neurons are 0.232154, 0.291636, and 0.283590, correspondingly. The training has been performed with the learning rate 0.01 and momentum rate 0.825. Table 3 describes training and testing results of FNS model obtained using different number of

TABLE 3: Simulation results of FNS.

Number of hidden neurons	RMSE training	RMSE evaluation	RMSE testing	Accuracy (%)
2	0.548520	0.560548	0.551954	81.025641
5	0.397395	0.401047	0.379963	93.333333
8	0.341242	0.435460	0.428456	95.897436
12	0.333357	0.343488	0.335679	97.948718
16	0.232154	0.291636	0.283590	100

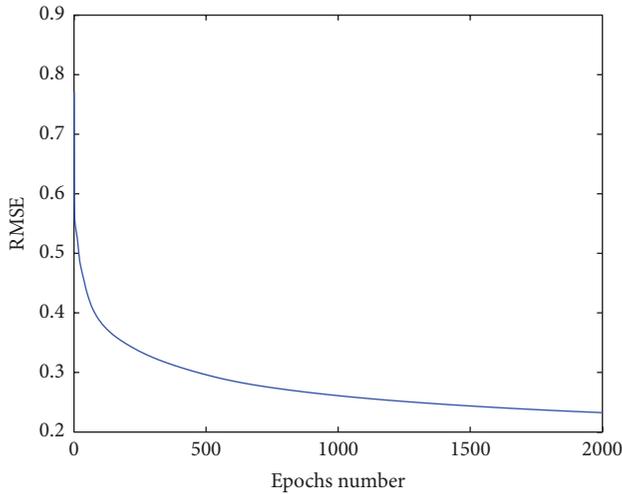


FIGURE 2: RMSE values.

rules: 2, 5, 8, 12, and 16. Simulation results are averaged over ten simulations.

From Table 3, it was shown that the increase in the number of rules (or the number of hidden neurons) decreases the values of RMSE for training and testing cases and increases recognition rate. The use of clustering and gradient techniques for learning allows quick obtaining of low RMSE value and allows improving performance of FNS for training and testing stages. In the second simulation a comparative analysis of the classification of PD has been performed. The result of the simulation of the FNS classification model is compared with results of simulations of different classification models, such as support vector machine (SVM), neural networks (NN), regression model, decision tree, and FCM based feature weighting. To estimate the performance of the NN, SVM, and FNS clustering systems, the recognition rates and RMSE values of errors between clusters and current output signal are compared. In Table 4, the comparative results of simulations of different models are given. As shown in the table the performance of FNS classification system is better than the performance of the other models.

## 5. Conclusion

The paper presents the diagnosis of Parkinson's diseases using fuzzy neural structures. The structure and learning algorithms of FNS are presented. Fuzzy clustering and gradient descent learning algorithms are applied for the

TABLE 4: Comparative results of different models for classification of PD.

Models	Accuracy (testing)
Decision tree [18]	84.3
Regression [18]	88.6
DMneural [18]	84.3
Neural network [18]	92.9
FCM based feature weighting [17]	97.93
SVM	93.846154
FNS	100

development of the FNS. Learning is performed using 10-fold cross validation data set. The design of the classification system is carried out using different number of fuzzy rules used in FNS. Recognition rate of classification is obtained as 100% with 16 hidden neurons. For comparative analysis, the simulation of PD is performed using different models. The obtained results demonstrate that the performance of FNS is better than the other models used for classification of PD.

## Conflict of Interests

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

## References

- [1] R. Betarbet, T. B. Sherer, and J. T. Greenamyre, "Animal models of Parkinson's disease," *BioEssays*, vol. 24, no. 4, pp. 308–318, 2002.
- [2] A. J. Hughes, S. E. Daniel, L. Kilford, and A. J. Lees, "Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinicopathological study of 100 cases," *Journal of Neurology Neurosurgery and Psychiatry*, vol. 55, no. 3, pp. 181–184, 1992.
- [3] M. A. Little, P. E. McSharry, S. J. Roberts, D. A. E. Costello, and I. M. Moroz, "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection," *BioMedical Engineering OnLine*, vol. 6, article 23, 2007.
- [4] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1015–1022, 2009.
- [5] N. Singh, V. Pillay, and Y. E. Choonara, "Advances in the treatment of Parkinson's disease," *Progress in Neurobiology*, vol. 81, no. 1, pp. 29–44, 2007.
- [6] A. K. Ho, R. Ianse, C. Marigliani, J. L. Bradshaw, and S. Gates, "Speech impairment in a large sample of patients with

- Parkinson's disease," *Behavioural Neurology*, vol. 11, no. 3, pp. 131–137, 1998.
- [7] J. H. L. Hansen, L. Gavidia-Ceballos, and J. F. Kaiser, "A nonlinear operator-based speech feature analysis method with application to vocal fold pathology assessment," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 3, pp. 300–313, 1998.
- [8] B. Boyanov and S. Hadjitodorov, "Acoustic analysis of pathological voices," *IEEE Engineering in Medicine and Biology Magazine*, vol. 16, pp. 4–82, 1997.
- [9] S. Hadjitodorov, B. Boyanov, and B. Teston, "Laryngeal pathology detection by means of class-specific neural maps," *IEEE Transactions on Information Technology in Biomedicine*, vol. 4, no. 1, pp. 68–73, 2000.
- [10] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge, UK, 1999.
- [11] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction: With 200 Full-Color Illustrations*, Springer, New York, NY, USA, 2001.
- [12] T. Khan, J. Westin, and M. Dougherty, "Classification of speech intelligibility in Parkinson's disease," *Biocybernetics and Biomedical Engineering*, vol. 34, no. 1, pp. 35–45, 2014.
- [13] R. Prashanth, S. Dutta Roy, P. K. Mandal, and S. Ghosh, "Automatic classification and prediction models for early Parkinson's disease diagnosis from SPECT imaging," *Expert Systems with Applications*, vol. 41, no. 7, pp. 3333–3342, 2014.
- [14] S. Pan, S. Iplikci, K. Warwick, and T. Z. Aziz, "Parkinson's Disease tremor classification—a comparison between Support Vector Machines and neural networks," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10764–10771, 2012.
- [15] G. Singh and L. Samavedham, "Unsupervised learning based feature extraction for differential diagnosis of neurodegenerative diseases: a case study on early-stage diagnosis of Parkinson disease," *Journal of Neuroscience Methods*, vol. 256, pp. 30–40, 2015.
- [16] J. Zhang, X. Luo, and M. Small, "Detecting chaos in pseudoperiodic time series without embedding," *Physical Review E*, vol. 73, no. 1, Article ID 016216, 2006.
- [17] K. Polat, "Classification of Parkinson's disease using feature weighting method on the basis of fuzzy C-means clustering," *International Journal of Systems Science*, vol. 43, no. 4, pp. 597–609, 2012.
- [18] R. Das, "A comparison of multiple classification methods for diagnosis of Parkinson disease," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1568–1572, 2010.
- [19] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning—I," *Information Sciences*, vol. 8, no. 3, pp. 199–249, 1975.
- [20] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, *Neuro-Fuzzy And Soft Computing*, chapter 7, Prentice-Hall, New Jersey, NJ, USA, 1997.
- [21] R. H. Abiyev and O. Kaynak, "Fuzzy wavelet neural networks for identification and control of dynamic plants—a novel structure and a comparative study," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 8, pp. 3133–3140, 2008.
- [22] R. H. Abiyev, "Fuzzy wavelet neural network based on fuzzy clustering and gradient techniques for time series prediction," *Neural Computing & Applications*, vol. 20, no. 2, pp. 249–259, 2011.
- [23] R. H. Abiyev, O. Kaynak, T. Alshamleh, and F. Mamedov, "A type-2 neuro-fuzzy system based on clustering and gradient techniques applied to system identification and channel equalization," *Applied Soft Computing*, vol. 11, no. 1, pp. 1396–1406, 2011.
- [24] R. H. Abiyev, "Credit rating using type-2 fuzzy neural networks," *Mathematical Problems in Engineering*, vol. 2014, Article ID 460916, 8 pages, 2014.
- [25] R. H. Abiyev, R. Aliev, O. Kaynak, I. B. Turksen, and K. W. Bonfig, "Fusion of computational intelligence techniques and their practical applications," *Computational Intelligence and Neuroscience*, vol. 2015, Article ID 463147, 3 pages, 2015.
- [26] Q. H. Do and J.-F. Chen, "A neuro-fuzzy approach in the classification of students' academic performance," *Computational Intelligence and Neuroscience*, vol. 2013, Article ID 179097, 7 pages, 2013.
- [27] M. Subhi Al-Batah, N. A. Mat Isa, M. F. Klaiib, and M. A. Al-Betar, "Multiple adaptive neuro-fuzzy inference system with automatic features extraction algorithm for cervical cancer recognition," *Computational and Mathematical Methods in Medicine*, vol. 2014, Article ID 181245, 12 pages, 2014.
- [28] J.-T. Lu, Y.-C. Chang, and C.-Y. Ho, "The optimization of chiller loading by adaptive neuro-fuzzy inference system and genetic algorithms," *Mathematical Problems in Engineering*, vol. 2015, Article ID 306401, 10 pages, 2015.
- [29] Z. Yang, Y. Wang, and G. Ouyang, "Adaptive neuro-fuzzy inference system for classification of background EEG signals from ESES patients and controls," *The Scientific World Journal*, vol. 2014, Article ID 140863, 8 pages, 2014.
- [30] N. K. Kasabov and Q. Song, "DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction," *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 2, pp. 144–154, 2002.
- [31] H. Melo and J. Watada, "Gaussian-PSO with fuzzy reasoning based on structural learning for training a Neural Network," *Neurocomputing*, vol. 172, pp. 405–412, 2016.
- [32] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, NY, USA, 1981.
- [33] S. L. Chiu, "Fuzzy model identification based on cluster estimation," *Journal on Intelligent Fuzzy Systems*, vol. 2, no. 3, pp. 267–278, 1994.
- [34] R. R. Yager and D. P. Filev, "Generation of fuzzy rules by mountain clustering," *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology*, vol. 2, no. 3, pp. 267–278, 1994.