Automatic detection of Parkinsonism using significance measures and component analysis in DaTSCAN imaging

F.J. Martínez-Murcia, J.M. Górriz, J. Ramírez, I.A. Illán, A. Ortiz, and the Parkinson’s Progression Markers Initiative

Abstract

The study of neurodegenerative diseases has been based for some time on visual and semi-quantitative analysis of medical imaging. This is the case of Parkinsonian Syndrome (PS) or Parkinsonism, which is the second most common neurodegenerative disorder, where $^{123}$I-ioflupane (better known by its tradename, DaTSCAN) images have been of great help. Recently, new developments in machine learning methods and statistics have been applied to the analysis of medical images, yielding to a more operator-independent, objective analysis of them, and thus, setting the Computer Aided Diagnosis (CAD) paradigm. In this work, a new CAD system based on preprocessing, voxel selection, feature extraction and classification of the images is proposed. After preprocessing the images, voxels are ranked by means of their significance in class discrimination, and the first $N$ are selected. Then, these voxels are modelled using Independent Component Analysis (ICA), obtaining a few components that represent each image, which will be used later to train a classifier. The proposed system has been tested on two databases: a 208-DaTSCAN image database from the “Virgen de la Victoria” Hospital in Málaga (VV), Spain and a 289-DaTSCAN image database from the Parkinson Progression Markers Initiative (PPMI). Values of accuracy up to 94.7% and 91.3% for VV and PPMI databases are achieved by the proposed system, which has proved its robustness in PS pattern detection, and significantly improves the baseline Voxel-as-Features (VAF) approach, used as an approximation of the visual analysis.

1. Introduction

Parkinsonian Syndrome (PS) or Parkinsonism is a neurological disorder clinically characterized by the existence of at least two of the following motor symptoms: tremor, hypokinesia and rigidity [1]. From a clinical point of view, the most common cause of PS is Parkinson’s Disease (PD), although a wide-range of other etiologies may lead to a similar set of symptoms. PD is a degenerative disorder of the central nervous system due to the progressive loss of dopamine-generating cells in the substantia nigra, which leads to a corresponding loss of dopamine transporters (DAT) in the nigrostriatal pathway. This disease is the second most common neurodegenerative disorder, with a prevalence of 1–3% in the population over 65 years of age [2].

The introduction of functional brain images techniques has set a milestone in the process of diagnosis of neurological disorders. Particularly in PD, radiotracer-based imaging studies have been applied to assess nigrostriatal presynaptic function. Several approaches such as $T_1$ weighted Magnetic Resonance Imaging ($T_1$-MRI) [3], Magnetization Transfer Imaging (MTI) [4], or functional Magnetic Resonance Imaging (fMRI) [5] have been applied to study neuronal activity, and have been suggested to increase the diagnostic accuracy in the case of Parkinsonian Syndromes [1,6,7].

Dopamine transporters (DAT) are proteins situated at the presynaptic terminal of dopaminergic neurons which are responsible for the re-uptake of dopamine. To measure the integrity of the presynaptic terminals, a measure of the density of DAT using in vivo functional imaging techniques like Single-Photon Emission Computed Tomography (SPECT) or Positron Emission Tomography (PET), along with specific ligands, can be of great help. Various tracers derived from tropane and cocaine analogues have been used for this purpose, among others $^{123}$I-flupane (better known by its tradename, DaTSCAN). After intravenous injection, this drug binds to the dopamine transporters in the striatum. Then, using a SPECT camera, it is
possible to obtain the distribution of the radiopharmaceutical in the brain, and visualize the loss of dopamine transporters [8–10]. The usage of this radiopharmaceutical agent and other tracers has revealed as a suitable method for delimiting patients with and without the involvement of the dopaminergic system in various studies [7,11], and reveal that patients with PS show a reduced uptake of the tracer in the striatum.

Traditionally, the DAT SPECT images obtained are subjected to a visual analysis performed by experienced clinicians. The procedure involves usually a predefined rating [12,13] or the analysis of regions of interest (ROIs) [14,15]. This procedure can be subjective and prone to error, since it relies on gross changes in transporter density throughout the ROI. In contrast, some more automatized methods have been proposed, as [15–17], which establishes semiquantitative parameters in order to index absolute differences between specific/ non-specific uptake in the tomographic examinations.

Lately, some methods based on the machine learning paradigm [18–20] and neural networks [21] have been applied to image analysis procedures, yielding to the construction of computer-aided diagnosis (CAD) systems for several neurodegenerative diseases, such as Alzheimer’s Disease (AD) [22–27] or PD [28,29]. These systems are applied to MRI, SPECT or PET images, in order to extract complex high-dimensional features. Then, these features are used to train an automatic classifier such as SVM [24] to discriminate between normal controls (images from healthy subjects) and pathological images (patients), performing an automatic diagnosis.

The aim of this article is to evaluate the performance of different CAD systems composed of four steps: preprocessing, voxel selection, feature extraction and classification. The CAD systems that we have proposed will use different preprocessing techniques on two different databases: a 208-DaTSCAN image database from the “Virgen de la Victoria” Hospital in Málaga, and a 289-DaTSCAN image database from the Parkinson Progression Markers Initiative (PPMI) [30]. Then, the first N voxels according to their statistical significance will be selected, using three different measures: Student’s t-Test [22,31,32], Mann–Whitney–Wilcoxon U-test [33,34] or Relative Entropy [35]. These methods have been proved very useful for selecting features for classification in several previous works [34,35] (Fig. 1).

In the next step, a feature extraction step will be performed. We will make use of a widely known computational method that has obtained very good results when applied to other neurodegenerative diseases, such as AD diagnosis: Independent Component Analysis (ICA) [23,36]. However, while the method in [23] relies on creating an average image of the individual samples of each class and then compute the mixing matrix, our method uses all the sample images and their class information to compute this matrix. This is possible thanks to the feature selection step, which reduces the number of voxels from over 500,000 to less than 20,000, enabling a faster computation, and hence, allowing us to include all the samples in the calculation of the mixing matrix. This leads to a more consistent estimation of the Independent Components. Finally, we evaluate the performance of the proposed system using SVM classifiers [23,27].

This works is organized as follows. In Section 2 the databases used to test the system, preprocessing techniques, voxel selection, feature extraction and classification methods used in this paper are presented. In Section 3, we propose some experiments, and different measures of the performance obtained with each combination are detailed. These results are also analysed and discussed here. Finally, the conclusions are drawn in Section 4.

2. Materials and methods

2.1. Databases

The images were obtained after a period of between 3 and 4 h after the intravenous injection of 185 MBq (5 mCi) of DaTSCAN, with prior thyroid blocking with Lugol’s solution. The tomographic study (SPECT) with Ioflupane/FP-CIT-I-123 was performed using a General Electric gamma camera, Millennium model, equipped with a dual head and general purpose collimator. A 360-degree circular orbit was made around the cranium, at 3-degree intervals, 60 images with a duration of 35 s per interval, 128 × 128 matrix. Image reconstruction was carried out using filtered back-projection algorithms without attenuation correction [37,38], application of a Hanning filter (frequency 0.7) and images were obtained with transaxial cuts, following the method proposed in [39].

The images were interpreted by three Nuclear Medicine specialists, with masking of the clinical orientation. Visual assessment was established by exclusively considering the normal/abnormal criterion and after arriving at a consensus report between the three specialists, i.e. whether the FP-CIT SPECT allowed differentiation of a group of conditions with presynaptic involvement from others in which their integrity is assumed, without trying to assign them to different clinical groups within the set of pathological studies. A study was considered to be normal when bilateral, symmetrical uptake appeared in caudate and putamen nuclei, and abnormal when there were areas of qualitatively reduced uptake in any of the striatal structures.

A total of 208 subjects (100 patients and 108 controls), randomly selected from the total studies performed in this center until December 2008 and referred to it because of a movement disorder, were included in the study. Mean age was 70.2 years (41–87) with a standard deviation of 10.2 years (a detailed description of the database can be found in [40]). Clinical diagnosis, a parameter used as ‘gold Standard’ to establish the existence of PS, was made using the diagnostic criteria established previously, with an established minimum follow-up period of 18 months. Those patients who were receiving treatment with drugs that had known or suspected effect on the level of the dopaminergic transporters through direct competitive mechanism were excluded. Although PD is the most representative pathology of PS, there are other medical conditions

![Fig. 1. Detailed schema of the proposed CAD system.](image-url)
which, though they differ clinically from this, are also expressed by this set of symptoms. Some of them are multisystem atrophy (MSA), progressive supra-nuclear palsy (PSP) and corticobasal degeneration (CBD), in which, unlike PD, as well as involvement of the presynaptic terminal, there is involvement at the post-synaptic level of the nigrostriatal pathway.

2.1.2. PPMI database

Data used in the preparation of this article were obtained from the Parkinson’s Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). For up-to-date information on the study visit www.ppmi-info.org.

The images in this database were imaged 4 + 0.5 h after the injection of between 111 and 185 MBq of DaTSCAN. Subjects were also pretreated with saturated iodine solution (10 drops in water) or perchlorate (1000 mg) prior to the injection. All subjects had a supplied $^{57}$Co line marker affixed along the canthomeatal line, which will facilitate subsequent image processing and allow the core lab to accurately distinguish left and right in the face of multiple image file transfers. These markers are only evident in the $^{57}$Co window and hence do not contaminate the $^{123}$I-DaTSCAN brain data [30,41].

Raw projection data are acquired into a $128 \times 128$ matrix stepping each 3 degrees for a total of 120 projection into two 20% symmetric photopeak windows centered on 159 KeV and 122 KeV with a total scan duration of approximately 30–45 min. Other scan parameters (collimation, acquisition mode, etc.) are selected for each site. The images of both the subject’s data and the cobalt striatal phantom are reconstructed and attenuation corrected, implementing either filtered back-projection or an iterative reconstruction algorithm using standard approaches [41]. After the processing, the database contains 289 spatially normalized images, 114 from Normal Control subjects and 175 from PD patients, and of a 91 size.

Fig. 2 shows some selected slices of (a) a PD patient from PPMI database, (b) a PK subject from VV database, (c) a healthy subject from PPMI database, and (d) a healthy control from VV database. It is interesting to note that healthy patients offer a greater contrast of the brain data\[30,41\].

2.2. Image preprocessing

After being registered, the images on both databases are preprocessed. The preprocessing step consist of two steps: spatial and intensity normalization. One additional step can be performed: a binary mask application, which reduces the number of voxels used in the subsequent analysis.

As the PPMI database provides us with spatially normalized images, as described above and in [30], we can only apply an intensity normalization or the addition of a mask to check the impact of these processes in the performance of the system. On the other hand, the “Virgen de la Victoria” hospital provides us with raw images in the VV dataset, which we have already normalized spatially using the SPM8 software [31] with the algorithm that we describe later in Section 2.2.1, but not intensity normalized. This allows us to check the effect of the intensity normalization process on the performance of the whole system, and applying (or not) a mask to the images.

2.2.1. Spatial normalization

All the images in the VV dataset were spatially normalized using the SPM 8 software [31] and yielding, in the case of the VV database, a $73 \times 73 \times 45$ functional activity map for each patient. This method assumes a general affine model with 12 parameters and a Bayesian framework that maximizes the product of the prior function (which is based on the probability of obtaining a particular set of zooms and shears) and the likelihood function, derived from the residual squared difference between the template and the processed image. The template image is computed by registering all control images to a randomly chosen one by affine transformations. This $N_c=108$ controls and its hemisphere mid-plane reflected images are averaged to create the template [42], providing a symmetric image.

$$g = \frac{1}{N_c} \sum_{i \in X_c} t_i(x, y, z) + t_i(-x, y, z)$$  \hspace{1cm} (1)

where $X_c$ denotes the subset of control images, $N_c$ the number of control images, $t_i(x, y, z)$ is the $i$th image and $t_i(-x, y, z)$ is its reflected image in the $x=0$ hemisphere mid-plane.
The images of the PPMI database are already registered and spatially normalized, as commented in the previous section, and following the method proposed in [41].

2.2.2. Intensity normalization

In order to establish comparisons between the uptake value in areas of specific activity (binding to dopaminergic transporters) and areas of non-specific activity (vascular activity) between subjects, some kind of intensity normalization is required. The method used in this work is based on the obtaining of an intrinsic parameter from the image, $I_p$, and the estimation of the binding activity as

$$t' = t/I_p$$

(2)

where $t$ denotes the spatially normalized image, and $t'$ the image normalized spatially and in intensity. The value $I_p$ is determined by a method called “Integral Normalization”. This method approximates the expression $I_p = \int t$ as the sum of all the intensity values of the image, giving an integral value of intensity.

2.2.3. Masking

One operator-dependent step in the system proposed that can be added is the application of a binary mask to select different ROIs that can be significant to the diagnosis of PD. This mask selects all voxels that are higher than a specific intensity threshold $I_{th}$ that must be inserted manually, so that only selected voxels are considered in further processing. The threshold has been established as the central value of intensity in the whole image, and computed as follows:

$$I_{th} = \frac{1}{N}(\text{max}(I_{mean}) - \text{min}(I_{mean})) + \text{min}(I_{mean})$$

(3)

where $I_{mean}$ is the mean of all the images in the database, excluding the tested one. Resulting masks are shown on Fig. 3. These figures illustrate how, using the proposed threshold, the regions of interest for the detection of Parkinsonism using DaTSCAN images (the striatum) are delimited and selected.

The advantages of adding a mask in the process of diagnosis are the considerable reduction of computer load, due to a smaller size of the input vector, as well as introducing a conscious delimitation of areas of great significance in PD diagnosis. On the other hand, the addition of this step also adds a non-automatic task that must be performed manually, which makes this addition not desirable. Anyway, the mask is applied in some of the experiments proposed below to evaluate the impact of a conscious and manual processing of the image in the subsequent results.

2.3. Statistical significance measures

In the first place, a statistical significance measure is computed for each voxel. This measure allow us to rank voxels from the most to the least significant, and then, extract the $N$ first voxels to use them in the subsequent analysis. We make use of three different significance measures: the well-known Student’s $t$-Test [43], the Mann–Whitney–Wilcoxon U-Test [43] and the Relative Entropy (or Kullback–Leibler divergence) [44].

2.3.1. Student’s $t$-Test

Students’ $t$-Test ($t$-Test) is a widely used statistical test which quantifies the differences between two classes $\Omega_1$ and $\Omega_2$. It uses a common estimation of variance for both classes and assumes normal variables. The estimation of statistical $t$ is computed [43] as

$$S_t = \frac{\overline{\Omega}_1 - \overline{\Omega}_2}{\frac{1}{n_1} + \frac{1}{n_2}}$$

(4)

where

$$\sigma^2_{\Omega_1,\Omega_2} = \frac{(n_1-1)\sigma^2_{\Omega_1} + (n_2-1)\sigma^2_{\Omega_2}}{n_1 + n_2 - 2}$$

(5)

$\sigma^2_{\Omega_1,\Omega_2}$ is an estimator of common standard deviation of both samples, $\overline{\Omega}_1$ and $\overline{\Omega}_2$ are the means of each class, $n_1$ is the number of samples in class $\omega_1$ and $n_2$ is the number of samples in class $\omega_2$.

2.3.2. Mann–Whitney–Wilcoxon Test

Mann–Whitney–Wilcoxon U-Test (MWW) uses the absolute value of the statistical $U$ to rank voxels. Calculation of $U$ value is done by the following expression [43]:

$$U_i = R_i - \frac{n_i(n_i + 1)}{2}$$

(6)

where $n_i$ is the sample size for sample $i$, and $R_i$ is the sum of the ranks in sample $i$ (where $i = 1, 2$). Smaller $U_i$ value is taken as the final $U$ value:

$$S_{MWW} = \min(U_1, U_2)$$

(7)

This statistical test measures the dissimilarity between two groups of values, and, although similar to Student’s $t$-Test, is less likely than the latter to spuriously indicate significance because of the presence of outliers and performs no assumption about the images’ statistical distribution. As we make use of real data, this can be an important feature of this method [43].

2.3.3. Relative Entropy

Relative Entropy (also known as Kullback–Leibler divergence) is a non-symmetric measure of the difference between two probability distributions $\Omega_1$ and $\Omega_2$. Due to its lack of symmetry, this measure can also be considered as a measure of significance between two different classes. Relative Entropy can be calculated with Eq. (8) [44].

Let $\Omega_1$ and $\Omega_2$ be two discrete random variables. Relative Entropy is defined as

$$S_{RE} = \sum_i \Omega_1(i) \log \frac{\Omega_1(i)}{\Omega_2(i)}$$

(8)

In words, it is the average of the logarithmic difference between the probability distributions $\Omega_1(i)$ and $\Omega_2(i)$ of class $\omega_1$ and $\omega_2$, respectively, where the average is taken using the probabilities $\Omega_2(i)$. The K–L divergence is only defined if $\Omega_1(i)$ and $\Omega_2(i)$ both sum to 1 and if $\Omega_2(i) > 0$ for any $i$ such that $\Omega_1(i) > 0$. If the quantity $0 \ln 0$ appears in the formula, it is interpreted as zero [44].
2.4. Independent Component Analysis

Independent Component Analysis (ICA) [45], is a statistical technique that represents a multidimensional random vector as a linear combination of non-Gaussian random variables (the so-called “independent components”) to be as independent as possible, and has been used widely on segmentation and clustering of medical images [23,36]. It can be considered as a non-Gaussian version of Factor Analysis [27].

This algorithm is used as a new strategy to avoid the small sample size problem [46]. This phenomenon occurs when the number of input features to the classifier is higher than the number of samples used to train this classifier. As the number of samples is around 200 in both databases, and the number of selected voxels \( N \) ranges between 110 and several thousands (depending on the database), a reduction in the input vector is desirable. Using ICA we obtain a representation of the selected voxels in the IC space, where the number of components \( K \) is lower than 25. This way we ensure that the system is not affected by the small sample size problem.

Assume that we observe \( n \) linear mixtures \( x_1, x_2, \ldots, x_n \) of length \( N \) that can be modelled as an expression of \( K \) independent components (IC). These independent components are defined as \( S = (s_1, s_2, \ldots, s_K) \), where each \( s_k \) vector has a length of \( N \). So, each random vector \( x_k \) can be described as a linear combination of \( K \) independent components:

\[
x_k = a_{k1}s_1 + a_{k2}s_2 + \ldots + a_{kN}s_N
\]

Without loss of generality we can assume that both the observed vectors and the independent components are zero mean. If the previous conditions are not met, the \( x \) variables can be centered by subtracting the sample mean. To use a vector-matrix notation, more convenient in this case, we denote as matrix \( X \) the random vector whose elements are \( x_1, \ldots, x_n \). We also denote as \( A \) the matrix that contains all \( a_{kn} \) elements, the “mixing matrix” that projects each image into the space defined by the IC. Using this notation, the mixing model above remains as follows:

\[
X = AS
\]

The starting point of ICA is the assumption that all components \( s_k \) are statistically independent. To measure independence, we assume that all independent components have a non-gaussian statistical distribution. It is assumed that a sum of independent signal trends to gaussianity, so if non-gaussianity is maximized with any independence criteria \( F \), for instance, the kurtosis or negentropy, we obtain signals that are more independent than the previous ones [45,47]. After estimating the matrix \( A \), we can compute its inverse, \( W \) and obtain the projection \( S \) of the images in the dataset into the IC space with

\[
S = WX
\]

2.4.1. FastICA

Adaptive algorithms based on gradient descend can be problematic when they are used on an environment in which adaptation is not necessary, like this case. The convergence is often slow, and depends on the choice of convergence parameters. As a solution to this problem, block algorithms based on kurtosis is introduced. In [49], this algorithm, known as FastICA, is generalized to general contrast functions. The single unit FastICA algorithm has the following form:

\[
w(k) = E(xg(w(k-1)^T x) - E(g'(w(k-1)^T x))w(k-1)
\]

where the loadings vector \( w \) is normalized to unit norm in each iteration, and the function \( g(x) \) is a derivative of the contrast function \( G \) defined in [47]. The expected values are estimated in practice by using the mean of a significantly high number of samples of the input data. The speed of convergence of the fixed-point algorithms is clearly superior to more neural algorithms. Improvements between 10 and 100 times the speed are observed frequently [50]. The FastICA MATLAB® package, provided by A. Hyvärinen in his website at the Aalto University (http://research. ics.aalto.fi/ica/fastica/) has been used to perform the Independent Component Analysis in this article.

2.5. Support Vector Machines

Support Vector Machines (SVM), introduced in the late 70s [51], are a set of related supervised learning methods widely used in pattern recognition [52], voice activity detection (VAD) [53] and classification [27]. SVM with linear discriminant functions define decision hypersurfaces or hyperplanes in a multidimensional feature space, that is:

\[
g(x) = w^T x + w_0 = 0,
\]

where \( w \) is known as the weight vector and \( w_0 \) as the threshold. The weight vector \( w \) is orthogonal to the decision hyperplane and the optimization task consists of finding the unknown parameters \( w_0 = 1, \ldots, n \) defining the decision hyperplane.

Let \( x_i, i = 1, 2, \ldots, n \) be the feature vectors of the training set, \( X \). These belong to either \( w_1 \) or \( w_2 \), the two classes. If the classes were linearly separable, the objective would be to design a hyperplane that classifies correctly all the training vectors. Among the different design criteria, the maximum-margin hyperplane is usually selected since it leaves the maximum margin of separation between the two classes. Since the distance from a point \( x \) to the hyperplane is given by \( z = \|g(x)/\|w\|\| \), scaling \( w \) and \( w_0 \) so that the value of \( g(x) \) is +1 for the nearest point in \( w_1 \) and -1 for the nearest points in \( w_2 \) the optimization problem is reduced to minimizing a cost function \( f(\omega) = 1/2\|\omega\|^2 \) subject to

\[
f_{svm}(x) = \sum_{i=1}^{N} \alpha_i y_i \Phi(s_i) \cdot \Phi(x) + w_0
\]

where

\[
0 \leq \alpha_i \leq C \quad \forall i \quad \text{and} \quad \sum_{i=1}^{N} \alpha_i y_i = 0
\]

In these equations, \( \alpha_i \) is the solution of a quadratic optimization problem that is usually determined by quadratic programming or the well-known sequential minimal optimization algorithm, and \( \Phi(s) \) or \( \Phi(x) \) denotes the transformation of the feature vectors into the effective feature space. This basic SVM classifier produces a linear separation hyperplane. A more general expression for SVM can be defined with the addition of kernels, that substitute the dot product of the transformed feature vectors \( \Phi(s) \cdot \Phi(x) = K(s,x) \). Thus, the expression of the hyperplane in Eq. (13) can be replaced by

\[
g(x) = \sum_{i=1}^{N} \alpha_i y_i K(s_i, x) + b
\]

The kernel functions that we use in this work are the linear kernel \( K(s, x) = s \cdot x \) (on Eq. (13)) and a Gaussian radial basis function (RBF) kernel \( K(s, x) = \exp(-\|s-x\|^2/(2\sigma^2)) \), as implemented in MATLAB® in the functions svmtrain and svmclassify of the Bioinformatics Toolbox™.

3. Experiments and results

Three different experiments are proposed, considering different preprocessing strategies in the databases:

- **Raw experiment.** This experiment applies the predefined spatial normalization for each database (commented in Section 2.2), but uses neither intensity normalization, nor any mask.
• **Experiment 1.** In this experiment, apart from using the spatial normalization of the Raw experiment, a binary mask (based on a threshold and proposed in Section 2.2.3) is applied to the whole database to select the regions of interest.

• **Experiment 2.** Finally, in this experiment, the whole database is intensity normalized, using the normalization method proposed in Section 2.2.2. No mask is used in this case.

These experiments are applied to both databases, obtaining the six combinations that appear in the following sections (PPMI-Raw and VV-Raw for the Raw experiment applied to PPMI and Virgen de la Victoria—VV-databases, PPMI-1 and VV-1 for the Experiment 1 applied to PPMI and VV databases and PPMI-2 and VV-2 for the Experiment 2 applied to either PPMI or VV database, respectively).

### 3.1. Validation methods and parameters

In this paper two relevant system parameters are defined: the number of voxels selected \( N \) by means of the significance measures obtained and the number of features extracted from this selection \( K \) using Independent Component Analysis (ICA). The performance of the classifier has been tested on values in the range of \( 200 < N < 1400 \) for VV database and \( 200 < N < 2200 \) for PPMI database, and on values in \( 1 < K < 25 \) in both databases.

For comparison among the different procedures and experiments, we use the evaluation parameters known as accuracy, sensitivity, specificity, positive likelihood (PL) and negative likelihood (NL) ratios. Apart from the well-known accuracy rate of a classification, which computes the proportion between correctly classified samples and total samples, sensitivity and specificity are the most widely used parameters to describe a diagnosis test. These parameters are defined as follows:

\[
\text{Sens} = \frac{TP}{TP + FN}, \quad \text{Spec} = \frac{TN}{TN + FP}
\]  

where \( TP, TN, FP \) and \( FN \) are the number of true positives, true negatives, false positives and false negatives, respectively. Sensitivity and specificity are used to measure the proportion of actual positives or negatives which are identified correctly (e.g. the percentage of PD patients, or normal controls who are identified as such).

These measures reveal the ability of a system to detect PS/NOR patterns. In this case, the best CAD system will be the one that leads to similar striatum values that do not allow our system to correctly interpret the results. As can be observed in Fig. 2a, similar values of intensity are localized in the striatum, while the values of the rest of the brain areas are more discriminative than those. So, selecting the obvious ROI – the striatum – could be here detrimental for the performance of the system.

The case of the VV database is different, due to its lack of intensity normalization. Although the subjects depicted on Fig. 2b and d have a similar scale – in fact, they have been chosen to better interpret the results – most of them do not show this attribute. Therefore, selecting the striatum in these images is also negative for the performance of the system, because sometimes a low-intensity striatum corresponds to PD subject and sometimes to a low-intensity brain of a non-affected subject due to, for example, a lower drug uptake or attenuation problems with the hardware. Here, the most relevant feature of these unnormalized images is the contrast between the striatum and the rest of the brain, and not the individual intensity values in these zones.

The behavior of the VAF system with these strategies of preprocessing highlights the benefits of using a normalization, which allow us to compare the images voxel to voxel, assuming that a

### 3.2. Baseline

The first method applied to the databases is the Voxels-as-Features (VAF) approximation [22]. VAF is considered as a baseline in many works, as many studies suggest that this method is, at least, comparable with the visual exam performed by experts [22]. This approximation uses all voxels in each image as a feature vector, which is used as an input to the classifiers.

This baseline has been applied to both databases, PPMI database and the database from the “Virgen de la Victoria” Hospital, using the Raw images (PPMI-Raw and VV-Raw), a mask to select only the regions of interest (PPMI-1 and VV-1), or the intensity normalization discussed in Section 2.2.2 (PPMI-2 and VV-2). We have used a 30-Fold validation method on the whole system and the evaluation parameters are presented in Table 1. Only SVM linear has been used to compute the results, due to the large number of input features to the classifiers, to obtain more generalizable results and avoid the Small Sample Size problem (see Section 2.4).

In this Table, the use of intensity normalization on the images shows a significant improvement of the performance results over the same VAF approach, but using the raw images. The addition of a mask is, on the contrary, inadvisable. This is probably due to the attenuation correction in the PPMI database and to the lack of intensity normalization in the VV database. In the first case, the attenuation correction leads to similar striatum values that do not allow our system to correctly interpret the results. As can be observed in Fig. 2a, similar values of intensity are localized in the striatum, while the values of the rest of the brain areas are more discriminative than those. So, selecting the obvious ROI – the striatum – could be here detrimental for the performance of the system.

<table>
<thead>
<tr>
<th>Database</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PL</th>
<th>NL</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPMI-Raw</td>
<td>0.837</td>
<td>0.833</td>
<td>0.840</td>
<td>5.21</td>
<td>0.198</td>
</tr>
<tr>
<td>PPMI-1</td>
<td>0.757</td>
<td>0.631</td>
<td>0.840</td>
<td>3.95</td>
<td>0.438</td>
</tr>
<tr>
<td>PPMI-2</td>
<td>0.840</td>
<td>0.807</td>
<td>0.862</td>
<td>5.88</td>
<td>0.224</td>
</tr>
<tr>
<td>VV-Raw</td>
<td>0.831</td>
<td>0.796</td>
<td>0.870</td>
<td>6.13</td>
<td>0.234</td>
</tr>
<tr>
<td>VV-1</td>
<td>0.808</td>
<td>0.778</td>
<td>0.840</td>
<td>4.86</td>
<td>0.265</td>
</tr>
<tr>
<td>VV-2</td>
<td>0.884</td>
<td>0.870</td>
<td>0.900</td>
<td>8.70</td>
<td>0.144</td>
</tr>
</tbody>
</table>

Values of accuracy, sensitivity, specificity, PL and NL are calculated using \( k \)-Fold cross-validation method [55,56]. \( k \)-Fold considers \( k \) randomly selected sets of AD patients and NOR. It iteratively holds out a set for testing purposes, while training the classifier with the remaining sets, so that each set is left out once. This is repeated for each of the \( k \) sets, and an average value of the evaluation parameters is computed. In this work, a number of folds \( k=30 \) is used.

Another method to measure the trade-off between sensitivity and specificity is the Receiver Operating Characteristic (ROC) curve [57]. This curve is computed by varying any of the input parameters, and plotting the values obtained for sensitivity versus 1 – specificity.
similar value of intensity in two different subjects corresponds to a similar value of the drug uptake.

3.3. Independent Component Analysis

The second system tested is the one which combines selection of $N$ voxels, ranked by means of t-Test, Wilcoxon or Entropy criteria, the modeling of these feature vector using Independent Component Analysis (ICA) and a posterior classification step. Most representative results are shown in Table 2.

Table 2 shows the resulting evaluation parameters obtained with each experiment, database and selection criterion used. The classifier that performs better with each combination is also cited, which gives us an idea of the linearity of the distribution of the features in the ICA space.

The first interesting result that we infer from the table is that values obtained for the Experiment 1 and Raw are very similar in both databases, with similar values of accuracy, PL and NL, being the results obtained with Experiment 1 (with a mask) slightly higher than those obtained without the mask. This is very interesting, because it involves a relevant computational load reduction, which increases the performance of the live system, and also performs slightly better in some cases. On the other hand, the similarities between the results obtained with Experiment 1 and Raw show that the hypothesis testing methods used as selection criterion perform in a similar way to the mask. This suggests that both using a supervised method (a mask) or an unsupervised method (computed significance over the whole image) the results are almost the same, and then, our choice must be based on a trade-off between the computational load of the system and the addition (or not) of a supervised step. The difference of computational load varies between computing the significance maps for the whole brain, which contains 902,629 voxels in the PPMI database and 239,805 for the VV database, or the computation of those maps for the 1125–2606 selected voxels of the striatum (depending on the database), which, as commented above, can be of great help in live systems. Both the usage of full images and the mask-based selection of voxels are equally valid, and both obtain similar results.

Regarding the Experiment 2, the obtained results behave in a different way depending on the database used. While the PPMI database obtains better results when there is no intensity normalization (in experiments PPMI-Raw and PPMI-1), the system obtains its highest performance when applied to a normalized VV database.

Before this behavior is analyzed, some details about the PPMI database must be pointed out. Due to the spatial normalization performed in the Raw PPMI images, some PD affected images show a deformation of the shape of the brain. This occurs because of the lower intensity of the striatum, due to lower amounts of dopamine in this area. When a template is used to register these images, the shape of the striatum differs substantially from that of the template (this is smaller, and of irregular shape), so that the resulting transformation could lead to a deformation of the striatum, and thus, of the whole brain, in order to match the template. Thereby some brain coordinates do not match the same anatomical position, which can lead to a misclassification.

Also, in PPMI database, an attenuation correction using a cobalt line marker (as commented in Section 2.1.2) has been performed. This line marker facilitates the image processing and distinguishes left and right sides for manual processing. However, this adds a non-relevant area in the images with values similar to those of the striatum, which experiments Raw and 2 can take into account, unlike experiment 1 that uses a mask to select the ROI. This combination of high-intensity non-relevant areas, brain deformations due to low intensity striatum and attenuation correction of the images could cause the lower performance values obtained in the Experiment 2 when applied to PPMI database.

On the other hand, best performance results are obtained when using an intensity normalization step in the VV database. As only spatially normalized images are used, the intensity normalization procedure features an increment in the accuracy, and a significant decrement in the NL values. Compared to Experiment 1, where only a mask is applied to the images, values obtained using either t-Test or MWW perform significantly better. Due to the aforementioned intensity normalization, values of the striatum in the Experiment 2 are comparable, while those of the Experiment 1 are not. However, the Raw images should present differences, since high accuracy values are also obtained with Experiments VV-Raw and VV-1.

Table 2
Best results of the complete system applied to the experiments proposed. Evaluation parameters of accuracy, sensitivity, specificity, PL and NL are shown. The classifier used to compute these results appears along with them and its corresponding experiment and selection criterion.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Method</th>
<th>Classifier</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PL</th>
<th>NL</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPMI-Raw</td>
<td>Entropy</td>
<td>SVM-Lin</td>
<td>0.903</td>
<td>0.982</td>
<td>0.851</td>
<td>6.61</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>t-Test</td>
<td>SVM-Lin</td>
<td>0.903</td>
<td>0.974</td>
<td>0.857</td>
<td>6.82</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon</td>
<td>SVM-Lin</td>
<td>0.900</td>
<td>0.956</td>
<td>0.863</td>
<td>6.97</td>
<td>0.051</td>
</tr>
<tr>
<td>PPMI-1</td>
<td>Entropy</td>
<td>SVM-Lin</td>
<td>0.907</td>
<td>0.974</td>
<td>0.863</td>
<td>7.10</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>t-Test</td>
<td>SVM-Lin</td>
<td>0.907</td>
<td>0.982</td>
<td>0.857</td>
<td>6.88</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon</td>
<td>SVM-Lin</td>
<td>0.913</td>
<td>0.991</td>
<td>0.863</td>
<td>7.23</td>
<td>0.010</td>
</tr>
<tr>
<td>PPMI-2</td>
<td>Entropy</td>
<td>SVM-Lin</td>
<td>0.900</td>
<td>0.965</td>
<td>0.857</td>
<td>6.75</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>t-Test</td>
<td>SVM-RBF</td>
<td>0.896</td>
<td>0.956</td>
<td>0.857</td>
<td>6.69</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon</td>
<td>SVM-Lin</td>
<td>0.900</td>
<td>0.956</td>
<td>0.863</td>
<td>6.97</td>
<td>0.051</td>
</tr>
<tr>
<td>VV-Raw</td>
<td>Entropy</td>
<td>SVM-Lin</td>
<td>0.928</td>
<td>0.935</td>
<td>0.920</td>
<td>11.69</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td>t-Test</td>
<td>SVM-Lin</td>
<td>0.923</td>
<td>0.917</td>
<td>0.930</td>
<td>13.10</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon</td>
<td>SVM-Lin</td>
<td>0.923</td>
<td>0.926</td>
<td>0.920</td>
<td>11.57</td>
<td>0.081</td>
</tr>
<tr>
<td>VV-1</td>
<td>Entropy</td>
<td>SVM-Lin</td>
<td>0.928</td>
<td>0.935</td>
<td>0.920</td>
<td>11.69</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td>t-Test</td>
<td>SVM-Lin</td>
<td>0.923</td>
<td>0.917</td>
<td>0.930</td>
<td>13.10</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon</td>
<td>SVM-Lin</td>
<td>0.923</td>
<td>0.926</td>
<td>0.920</td>
<td>11.57</td>
<td>0.081</td>
</tr>
<tr>
<td>VV-2</td>
<td>Entropy</td>
<td>SVM-RBF</td>
<td>0.938</td>
<td>0.954</td>
<td>0.920</td>
<td>11.92</td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td>t-Test</td>
<td>SVM-RBF</td>
<td>0.947</td>
<td>0.981</td>
<td>0.910</td>
<td>10.91</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon</td>
<td>SVM-RBF</td>
<td>0.942</td>
<td>0.981</td>
<td>0.900</td>
<td>9.81</td>
<td>0.021</td>
</tr>
</tbody>
</table>
The last comment of this table concerns the classifiers that better match each experiment and voxel selection criterion. The best classifier is very representative of the spatial distribution that the projected features present on the ICA space. Linear classifiers such as SVM usually fit linearly separable classes, which can result in a good projection of the features into the ICA space. But SVM Linear performs as the best classifier when using random distributed classes too. On the contrary, non-linear classifiers such as the SVM-RBF, reveal classes that are non-linearly separable, and therefore, achieve best performance because they conform to best shape the spatial distribution of ICA projections in space. While using SVM-Lin or RBF in experiments Raw and 1 does not vary the performance of the system (e.g, using SVM-RBF in PPMI-Raw leads to accuracy results around 0.90 and using it with VV-1 achieves an accuracy around 0.93), the intensity normalization procedure seems to vary the spatial distribution of the independent components, specially with the VV database, making a non-linear kernel (RBF) to better fit those and obtain better results.

Because the differences between Experiments Raw and 1 have not been significant beyond the computational load, we will concentrate on the difference of the performance in using intensity-normalized images or not. As the PPMI database was attenuation corrected and, regarding the results obtained for Experiments Raw and 2, this can be considered as an intensity normalization procedure, we will concentrate on the results obtained for PPMI-Raw. In the case of VV database, we will analyze both the results obtained for experiment Raw (unnormalized) and 2 (normalized), and compare the performance obtained in both experiments.

To better illustrate the variations of the performance values along the different variables $N$ (number of selected voxels) and $K$ (number of independent components used), the accuracy obtained for each of the experiments proposed above (PPMI-Raw, VV-Raw and VV-2) when varying the input parameters of the system is shown in Fig. 4.

Fig. 4a and b shows the accuracy values obtained for experiment PPMI-Raw and each of the selection criteria. Those images

![Accuracy vs Number of Selected Voxels](image1)

![Accuracy vs Number of Independent Components](image2)

![Accuracy vs Number of Selected Voxels](image3)

![Accuracy vs Number of Independent Components](image4)

![Accuracy vs Number of Selected Voxels](image5)

![Accuracy vs Number of Independent Components](image6)

Fig. 4. Accuracy values over $K$ and $N$ for (a and b) experiment PPMI-Raw; (c and d) experiment VV-Raw; and (e and f) experiment VV-2.
feature, as commented, an attenuation correction that performs similar to an intensity normalization procedure, leading to similar results in the three experiments proposed (Raw, 1 and 2). The number of selected voxels that we have evaluated is higher than that used for VV database, due to the higher size of images in the PPMI database. The accuracy plots shows that performance values remain almost constant in values around 0.9 for every experiment and selection criterion (in selected operation ranges). Particularly, when regarding the N values in Fig. 4a, there is an interesting behavior in the graphs showing how different selection methods lead to different operation ranges. A sort of voxel threshold divides the accuracy in two values: around 0.85 and 0.9, and this threshold varies between 400 voxels for MWW selection criterion and 1400 voxels for Relative Entropy. Therefore, the method that better measures the significance of voxels is the one that achieves better values of accuracy with a smaller number of voxels.

However, the behavior in Fig. 4b, that depicts how the performance values vary with a different number of independent components (K), offers different results. In fact, best values are obtained with lower K, especially when using a Relative Entropy selection criterion. This means that the independent components are better modeled when it takes into account the voxels selected using Relative Entropy significance measures. On the other hand, either t-Test and MWW offer similar results, as expected. But the point here is the high accuracy values around 0.9 and with low variation along the number of components extracted.

In Fig. 4c and d, accuracy results for the VV-Raw experiment are shown. The images used on VV-Raw were only spatially normalized, so no comparisons can be made between intensity levels a priori. However, the relationships between striatum values and the remaining areas of the brain are very similar among Parkinson-affected images, and so it happens in normal control images. As significance values have been obtained using three different hypothesis testing methods among all images, and these significance measures are computed according to intensity values – which are not comparable a priori – it is possible to think that there will be a number of noisy voxels which will reduce the performance of the classifier. That explains the variations of the accuracy on Fig. 4c and d. In Fig. 4c low accuracy values are obtained with a lower number of voxels, but the performance increases as a higher number of voxels are considered. Due to the aforementioned reasons, it is also reasonable to assume that the estimation of the independent components will be better as the number of selected voxels increases, and so will the performance of the classifier.

Regarding the accuracy variations in function of K, it is also interesting to note that we need a high number of components (more than 6) to achieve good accuracy results. This is probably due to the intensity differences between different images of the same class, and between images of different classes. As has been commented before, we needed a higher number of voxels to obtain good performance results, and probably some of those were non-relevant noisy voxels that obtained higher significance values due to the lack of intensity normalization. Therefore, a higher number of independent components are needed to better model these vectors that contain both significant and noisy voxels. Finally, to focus on the operating ranges, for experiment VV-Raw we obtain good results of accuracy in a range of N > 600 and K > 6, although it might slightly vary depending on the selection criterion used.

Fig. 4e and f depicts accuracy values for experiment VV-2. Here, it is important to highlight the high values and robustness of the accuracy values over N, which are obtained using either of the significance computation methods proposed. Due to the intensity normalization applied to the images of the database, the system outperforms the VV-Raw, obtaining a robust performance and almost totally avoiding dependence on the number of voxels selected N. Accuracy values over K perform slightly different to the previous two experiments, mainly due to the different classifier used here. Best values are obtained in the first 5 independent components, and from this point on, the performance of the system degrades, which is a typical behavior of using non-linear kernels. A number of components of K=4,5 is optimum for this application. This is probably due to the peaking phenomenon [58]. The peaking phenomenon describes how the use of complex classifiers leads to a better fitting of the training data, but degrading its generalization properties. Then, an increment in the size of the training data (a new IC added) leads to an increment in the complexity of the classifier, at the expense of the generalization ability of the classifier. Therefore, as the number of IC increases, the generalization capabilities of the classifier decrease, and so does the performance of the system. Since this problem is exacerbated by the complexity of the system, this will occur much faster with an RBF kernel than with a Linear one.

As seen in these figures, the method used to estimate voxel significance has no relevant impact on later outcomes. To illustrate better this result, Table 3 summarizes the average parameters for the PPMI-Raw and VV-2 experiments. As we have tested the benefits of using the intensity normalization algorithm in the database, experiment VV-Raw will be obviated from this moment. These average values are computed over N and over all K for PPMI-Raw and the first 10 independent components for VV-2 for each significance estimation method, using either Linear or RBF kernel in both databases, to better illustrate the behavior of the system.

These average accuracy values range between 0.88 and 0.89 for PPMI-Raw and over 0.92 when computed over N for experiment VV-2 (SVM-RBF). Average values computed over K for experiment VV-2 are not very relevant because, as commented before, the performance of the classifier decreases when the number of components increases due to the peaking phenomenon. A slight decrease in the performance of the system when we use the Linear kernel is noticeable in this case, although maintaining good performance values. Both systems hold small dependence of the number of voxels N and, when a linear classifier is used, of the number of components K. Positive and Negative Likelihood values obtained are lower in PPMI database than in VV database. As commented in Section 3 values of PL over 8 are very likely for the patient to suffer from a specific diagnosis, so values obtained for VV-2 experiment are very representative (Fig. 5).

Regarding the PPMI database, although accuracy, sensitivity specificity, PL and NL values are good, and the classifier shows its robustness against the variation in the input parameters K and N, the performance showed by this system does not equal to that obtained by the VV-2 experiment. This can be due to both the aforementioned attenuation correction and the spatial normalization difficulties. Using an RBF kernel here provides no improvements, and also decreases the performance values in most cases. The best average values for PPMI database are obtained when considering a system composed by Relative Entropy significance measure, ICA for feature extraction when using SVM-Lin and the Mann–Whitney–Wilcoxon U-test when using SVM-RBF. For VV database, similar average values are obtained by any of the selection criteria and using SVM-RBF. We finally choose Relative Entropy criterion by analogy with the experiment PPMI-Raw. Average values of accuracy, sensitivity and specificity for each of these systems (Relative Entropy + ICA + SVM Classifier), computed along N and K are shown in Fig. 6.

Finally, we look at the ROC curves obtained by the proposed system in Fig. 6a PPMI database and 6b VV database. The performance of the system by means of its sensitivity–specificity trade-off is displayed using the three alternative significance estimation methods proposed previously, along with the performance obtained by the VAF approach, using the provided images in the case of PPMI and the intensity-normalized images in the
It is interesting to note that, as we have commented above, the performance obtained by the three significance estimation methods is very similar, with similar operation points, although the differences among them are bigger in the PPMI database. The systems proposed perform in both cases better than the VAF-based systems, which is widely considered as an

case of VV.
estimation of the performance of a visual analysis performed by expert clinicians [22].

4. Conclusion

In this paper, a new approach to Computed Aided Diagnosis (CAD) for the Parkinson’s Disease (PD) is proposed. The system makes use of a preprocessing step, in which the images are spatially and intensity normalized, and using an optional mask. Then, the voxels of the images are ranked by means of their significance, and the first $N$ are selected. These voxels are modeled in terms of Independent Component Analysis (ICA) to extract a low number of Independent Components which work as feature vectors for each image. Finally, a SVM classifier is trained to detect these PD patterns, and its performance is evaluated.

The resultant system performs significantly well in both databases, although slightly better with VV database, probably due to the flexibility of the intensity normalization process. Further analysis reveals that, when normalization is applied to the database, the significance measures perform in a more similar way. The system demonstrates its ability and robustness in PS pattern detection as it provides high accuracy, sensitivity and positive likelihood values with a wide range of operation, remaining almost $N$-independent and with a low $K$ dependence – depending on the classifier used – within a defined range. Furthermore, the system clearly outperforms the VAF-based systems, which is considered to reproduce the visual procedure performed by clinicians, in both databases.

Acknowledgments

This work was partly supported by the MICINN under the TEC2008-02113 and TEC2012-34306 projects and the Consejería de Innovación, Ciencia y Empresa (Junta de Andalucía, Spain) under the Excellence Projects P07-TIC-02566, P09-TIC-4530 and P11-TIC-7103. DaTSCAN image database is courtesy of the Department of Nuclear Medicine of the “Virgen de la Victoria” Hospital in Málaga (Spain).

PPMI – a public–private partnership – is funded by The Michael J. Fox Foundation for Parkinson’s Research and funding partners, including Abbott, Biogen Idec, F. Hoffman-La Roche Ltd., GE Healthcare, Genentech and Pfizer Inc.

References


F.J. Martínez-Murcia received the B.Sc. degree in Telecommunications Engineering and the Master degree in Computer Architecture and Networking from the University of Granada, Spain in 2010 and 2011 respectively. Then, he earned a Ph.D. grant with the SIPBA research group, in the Dept. Signal Theory, Networking and Communications at the University of Granada (2012). His research interests at present lie in the field of Signal Processing and Biomedical Applications.

J.M. Górriz received a B.Sc. degree in Physics (2000) and a B.Sc. in Electronic Engineering (2001) from the University of Granada, Spain. He then earned a PhD from the University of Cádiz, Spain (2003) and a PhD from the University of Granada, Spain (2006). He is currently a full-time Professor with the Department of Signal Theory, Networking and Communications at the University of Granada. He has coauthored more than 200 technical journals and conference papers in these areas and has served as an editor for several journals and books. His present interests lie in the field of statistical signal processing and its application to speech and medical image processing.

J. Ramírez received the M.A.Sc. degree in Electronic Engineering in 1998, and the Ph.D degree in Electronic Engineering in 2001, both from the University of Granada. Since 2001, he is a full-time professor at the Dept. of Signal Theory Networking and Communications of the University of Granada (Spain). His research interest includes signal processing and biomedical applications including brain image processing, robust speech recognition, speech enhancement, voice activity detection, seismic signal processing and implementation of high performance digital signal processing systems. He has coauthored more than 150 technical journal and conference papers in these areas. He has served as a reviewer for several international journals and conferences.

I.A. Illán received the B.Sc. Degree in Physics from the Complutense of Madrid University, Spain and the Ph.D from the University of Granada, Spain in 2004 and 2009. Actually, he is working under a local fellowship in the Department of Signal Theory, Networking and Communications at the University of Granada. His present research interests include supervised learning, signal processing, independent component analysis and biomedical applications.

Andres Ortiz received the M.Sc. degree in electronics in 2000, and the Ph.D. degree in 2008, both from the University of Granada. From 2000 to 2005 he was working as Systems Engineer with Telefonica, Madrid, Spain, where his work areas were high performance computing and network performance analysis. Currently he is an Associate Professor at the Department of Communication Engineering of the University of Malaga. His research interests include high performance networks, bioinspired systems and biomedical signal processing.