Expert Systems with Applications 39 (2012) 10764–10771

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Parkinson's Disease tremor classification – A comparison between Support Vector Machines and neural networks

Song Pan [a,*], Serdar Iplikci [b], Kevin Warwick [a], Tipu Z. Aziz [c]

[a] School of Systems Engineering, University of Reading, UK
[b] Department of Electrical and Electronics Engineering, Pamukkale University, Turkey
[c] University Laboratory of Physiology, University of Oxford, UK

ARTICLE INFO

Keywords:
Parkinson's Disease
Deep Brain Stimulation
Intraoperative microelectrode recordings
Radial Basis Neural Network
Multiple Layer Perception
Support Vector Machine

ABSTRACT

Deep Brain Stimulation has been used in the study of and for treating Parkinson's Disease (PD) tremor symptoms since the 1980s. In the research reported here we have carried out a comparative analysis to classify tremor onset based on intraoperative microelectrode recordings of a PD patient's brain Local Field Potential (LFP) signals. In particular, we compared the performance of a Support Vector Machine (SVM) with two well known artificial neural network classifiers, namely a Multiple Layer Perceptron (MLP) and a Radial Basis Function Network (RBN). The results show that in this study, using specifically PD data, the SVM provided an overall better classification rate achieving an accuracy of 81% recognition.

1. Introduction

As life expectancies increase it is a rather unfortunate consequence that diseases associated with aging also increase. Such is the case with Parkinson's Disease (PD). At present the disease is estimated to affect around 180 out of 1,00,000 people, however this proportion is on the increase as people live longer. It is important therefore that we gain more of an insight into the problem and improve our methods of dealing with it. Periodic muscle tremor and rigidity are the most commonly witnessed symptoms however in latter stages, Bradykinesia, Akinesia and Dysathria can also be witnessed.

Several chemical methods are generally used to treat the disease, in its early stages the drug Levodopa (L-dopa) has been widely employed since the early 1970s. However in the vast majority of patients either the effectiveness of L-dopa decreases with time as the disease worsens and/or the severity of side effects due to continually taking L-dopa increase. For those who contract PD at an early age these are particularly pertinent. At such times surgical treatment becomes a considered option as a means to alleviate the symptoms. This can mean lesion however nowadays much more likely is the use of deep brain positioned electrodes to provide regular electronic pulses to a specific area of the brain – a process referred to as Deep Brain Stimulation (DBS).

Deep Brain Stimulation has been used as a medical treatment for Parkinson's Disease (PD) resting tremor, particularly for those with latter stage PD, since the late 1980s (Breit, Schulz, & Benabid, 2004; Cheung & Tagliati, 2010). By placing an array of microelectrodes inside the patient's basal ganglia, specific targets being the subthalamic nucleus (STN) or the globus pallidus internus (GPi) (Benabid, Chabardes, Mitrofanis, & Pollak, 2009; Bour et al., 2010), stimulating pulses can be applied to overcome some or all of the PD symptoms. Typically the array consists of 4 or 6 electrodes positioned lengthways along the shaft of the implant, therefore current can be applied at different depths in the brain. It also means however that the same electrodes can be used to monitor and record electrical activity in this particular part of the brain. Although it is not yet fully understood why DBS prevents tremor, it does give researchers a unique opportunity to understand a little how some of the signals in a specific part of the human brain change before and after the onset of PD tremor.

Currently the collection of LFP signals can only be done during the initial medical DBS operation period (within 1 or 2 weeks of initial implantation before all associated electronics and power are fully implanted), which unfortunately prevents a long-term study of LFP signals. Previous research from Pan (Pan et al., 2007) and Wu (Wu, Warwick, Ma, Burgess, et al., 2010; Wu, Warwick, Ma, Gasson, et al., 2010) has shown that, based on short-term, single patient trials, both Multiple Layer Perceptrons (MLP) and Radial Basis Function Networks (RBN) can be trained to recognise the different LFP patterns which occur at tremor onset.

One important reason for such research using AI techniques is that the present means of stimulation employs continuous current pulses at relatively high frequency, the result being that regular battery replacement is required every 18–24 months, involving further surgery (Wu, Warwick, Ma, Burgess, et al., 2010). Unfortunately inductive battery recharging is not deemed to be a viable

* Corresponding author.
E-mail address: cocoa117@gmail.com (S. Pan).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.02.189

alternative – the passive coil that needs to be implanted is quite large, there is a likelihood of chemical discharges in the body and the patient must lie for long periods of time in an awkward position which patients do not wish to do – even then battery lifetime is only marginally improved. As a result the inductive recharging option is rarely taken up, indeed never at the John Radcliffe Hospital.

On the other hand it appears sensible to attempt to develop an 'intelligent' stimulator, one which only stimulates when it needs to (Gasson, Wang, Aziz, Stein, & Warwick, 2005; Pan et al., 2007). In this sense the stimulator would operate, rather like a heart pacemaker, on a demand only basis. Rough estimations indicate that such a device could extend the battery life to 10 years or even more – possibly further than the life expectancy of a patient diagnosed with PD. The key question however is how does the stimulator 'know' when to stimulate and when not. Hence the requirement for an 'intelligent' method to analyse the LFP signals in real time which is able to clearly indicate when stimulating pulses are required.

The need is therefore for a method of analysis which monitors LFP signals in real time and which indicates when tremors are occurring or, even better, which accurately predicts when tremors are going to occur before they do so. In this case, if the method concludes that tremors are occurring or are about to occur, when this is not the case – a false positive – then this is not so bad. It may result in the stimulator switching on rather more than one would like, but other than that it's not really an issue. Indeed, in the limit it actually approaches the present day operation of the stimulator which effectively assumes that tremors exist all the time, which is not the case at all. However if the method concludes that tremors are not occurring and are not about to, when in fact they already exist or are about to start – a false negative – then this is a major problem which is not acceptable as, based on the analysis, stimulating pulses would not be applied and the patient would actually experience the tremors, which is what we are trying to avoid.

In previous studies in which the possibility of developing an "intelligent" stimulator (Cassidy et al., 2002; Gasson et al., 2005) by analysing the STN LFP signals was first documented, different types of artificial neural network such as Multiple Layer Perceptrons (Pan et al., 2007) and Radial Basis Function Networks (Wu, Warwick, Ma, Burgess, et al., 2010; Wu, Warwick, Ma, Gasson, et al., 2010) were used to "predict" PD tremor based on LFP signals. Initial results showed that both RBN and MLP could be used as classifiers to obtain reasonable results when fed with such tailored LFP signal features as the power spectrum density up to 45 Hz. In such cases both networks were able to separate the features into pre-defined categories with a respectable success rate that clearly could be improved upon.

Much related research has though indicated a clear link in PD between arm EMG signals and STN LFP data specifically between 3–10 Hz (Burgess et al., 2010; Cassidy et al., 2002; Lemstra, Verhagen Metman, Lee, & Lenz, 1999; Levy et al., 2002; Moran, Bergman, Israel, & Bar-Gad, 2008; Weinberger, Hutchison, Lozano, Hodaie, & Dostrovsky, 2009), particularly at tremor onset with nowhere near as strong a relationship in the remainder of the frequency range 0–45 Hz. In this study therefore we focused on using short-term power spectrum at frequencies only in the range 3–10 Hz, by means of band pass filtering, as signal feature input to different types of classifiers. By providing the same training set, the performance of each classifier could then be measured and compared using a Receiver Operating Characteristics (ROC) graph with resultant sensitivity and specificity plots.

In the results described in this paper we focussed not only on MLPs and RBNs, in order to compare directly with previously obtained results, but also, for comparative purposes we also considered the use of Support Vector Machines (SVM) as signal pattern recognition devices, in this application area. SVMs have in fact been widely used with similar types of classification tasks such as drug/nondrug classification (Byvatov, Fechner, Sadowski, & Schneider, 2003) and EEG classification (Zheng, Yang, Li, Zan, & Dong, 2010). However in this paper we examined the performance of the SVM for the classification of PD tremor using LFP signals collected from a real PD patient.

The nature of an MLP as a classifier means that the resultant network and hence its performance is highly dependent on the training data provided even when the same initial set of weights and biases are selected (Warwick, 2011). The RBN meanwhile uses an unsupervised clustering algorithm to determine the centres for the RB neurons, such as *k*-means, fuzzy c-means or a mean tracking algorithm (Sutanto, Mason, & Warwick, 1997). A statistical classifier such as SVM uses a regressive algorithm to arrive at the best curve of specified degree which regresses the data set. Regardless of the number of training epochs related with MLP and problems associated with finding suitable centres for the RB neurons, we focused on each classifier's performance by using an ROC graph to see how well each classifier performed under the same training set. This should in turn give us a better idea as to how well each classifier would perform with real patient LFP data. Essentially both the MLP and RBN were given their best chance.

In the next section we have given an indication of the signal processing carried out. This is followed, in Section 3, by brief descriptions of the two artificial neural networks investigated – rather than describing the basics of what are probably very well known concepts to the reader, the main point with this is to show how we implemented the techniques. In Section 4 we have then indicated the method employed with SVMs. Section 5 then describes in detail the results we obtained such that, in Section 6, some conclusions can be drawn.

2. Tremor frequency analysis

The data used in this study were recorded from a 74 year old, male volunteer who had been diagnosed with Parkinson's Disease and who underwent DBS surgery at the John Radcliffe Hospital, Oxford, UK, with electrodes implanted into his STN. After the surgery a week long recovery period allowed the opportunity to perform recordings of the LFPs when the patient was in a resting (non-tremor) state, through tremor onset and under tremor conditions. The signal data was collected using a bipolar configuration via four contacts of the implanted macrostimulation electrode (Medtronics 3387) implanted in the left subthalamic nucleus.

For experimental purposes surface Electromiograph (EMG) signals were also recorded using disposable adhesive Ag/AgCl electrodes (H27P, Kendall-LTP, MA, USA) positioned over the selected flexor muscle of the patient's right forearm. In this way the physical realisation of both tremor and non-tremor states could be witnessed. Signals were amplified using isolated CED 1902 amplifiers ($\times$10,000 for LFPs and $\times$1000 for EMGs), filtered at 0–500 Hz and digitised using CED 1401mark II at sampling rates of 1000 Hz (Cambridge Electronic Design, Cambridge, UK) (S. Wang, Aziz, Stein, Bain, & Liu, 2006; S.Y Wang, Aziz, Stein, & Liu, 2005). The LFP signal dataset contained digitised signals from all electrodes in a paired/differenced fashion, 0–1, 1–2 and 2–3. By this means the parallel recorded EMG signal could be used to determine the time of tremor onset as can be seen in Fig. 1.

For this research purpose, we manually chose data recording that presented the patient's LFP activity before, during and after tremor onset. This gave a better view on how the power spectrum density (PSD) characteristics changed over time with regard to the event. In a 'perfect' world, control group information would be available to show what normal LFP activity is like in the STN for comparison purposes. However fit, non-PD, volunteers who are

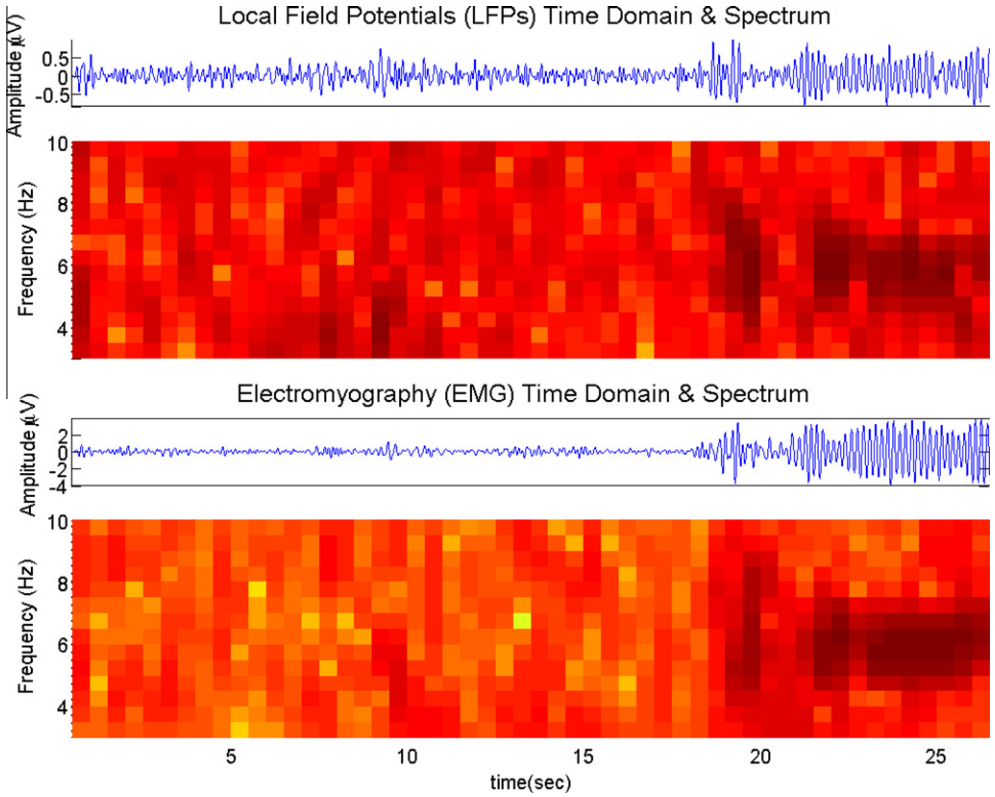

Fig. 1. Patient's LFP and EMG signal.

willing to undergo DBS are difficult to find. It is also unclear if ethical approval for such a venture could possibly be obtained. Hence the data we have is as good as it gets.

Work done by Brown and others (Brown & Williams, 2005; Wang et al., 2006) has shown the LFP tremor frequency gathered from the STN corresponds exactly to the motor unit frequency in the motor neurons – indeed they appear to be covariant. Fig. 1 shows exemplar signals recorded from the patient. The raw LFP and EMG data was first filtered with a 3–10 Hz Chebyshev Type II passband filter. This process removed noise such as 50 Hz mains noise, it also focussed on the details of the power spectrum features in the tremor frequency band. The signal was then normalised using the Matlab built in zscore function. The signal was normalised by using

$$z = \frac{x - u}{q} \tag{1}$$

where, the x is the raw signal data point to be normalised, u is the mean of the dataset and q is the standard deviation of the dataset.

To be clear – the LFP data available is fed in, as described, as input to the classifiers under test. These classifiers do not get to see the EMG data, also shown in Fig. 1. This data is rather for the reader/us to inspect and decide on when the onset of tremor occurs in terms of the increase of energy witnessed at a particular time in a specific part of the power spectrum.

In order to further explore the signal characteristic changes in the temporal domain, the Short-term Fast Fourier Transform (STFT) technique was applied to the gain power spectrum density (PSD). The signal was first equally divided using a sliding moving window with a size of one second period. In particular we used a Hamming window and set the overlap to 1/2 second. The PSD spectrogram function filter generated nearly 14 outputs of power spectrum density values per half second in the frequency range 3–10 Hz.

To further reduce the classifier input dimension and avoid the "curse of dimensionality" problem for MLP and RBN classifiers, we used a Principle Component Analysis (PCA) algorithm. PCA was employed merely to reduce the dimensionality of the input data pattern without losing accuracy in data representation. The number of eigenvectors to be retained was found by trial and error. The PCA was undertaken by employing linear singular value decomposition (SVD). Performing PCA is the equivalent of SVD on the covariance matrix of the data.

The main procedure carried out was as follows:

1. Obtain the input data vector. Subtract the mean from each of the data dimensions. The mean subtracted is the average across each dimension.
2. Calculate the covariance matrix.
3. Calculate the eigenvalues and eigenvectors of the covariance matrix.
4. Choose components and form a feature vector.
5. The top six eigenvalues were then used as classifier inputs.

2.1. EMG evaluation

There are two purposes for the EMG signal gathered from the patient's right forearm flexor muscle. First it has been used to detect the tremor onset time mark, as was described in the previous section. This gives the researcher a better understanding of how the LFP signal activity changes in the pre-tremor period. Secondly, it has also been used to evaluate each classifier's "recognition" and "prediction" performance.

The EMG signal was pre-processed by means of the following procedures.

1. The raw EMG signal was band pass filtered between 3–10 Hz with a Chebyshev Type II bandpass filter.

2. A Fast Fourier Transform was applied to the filtered signal to indicate the precise tremor frequency band.
3. STFT was applied to the filtered signal and the precise tremor frequency band was selected to represent EMG tremor.
4. The STFT EMG tremor band data was normalised using the standard zscore function – see Eq. (1).
5. The normalised EMG Tremor activity signal was divided into two statuses. From experience and visual inspection, PSD activities higher than the threshold value −0.5 were considered to indicate tremor status, whereas below that threshold was considered to be non-tremor status.
6. EMG evaluation data was further processed to determine the tremor onset time mark. From experience it was decided that PSD activities higher than the threshold value for a period of 5 s was marked as tremor onset.

The classifier evaluation dataset was generated based on the EMG PSD activities. A simple filter was applied to the signal, so that tremor status was represent with value 1 and non-tremor status represent with value 0.

3. Artificial Neural Networks

3.1. Multiple Layer Perceptron (MLP)

In this section we give a brief overview of the Multi Layer Perception artificial neural network as we are using it. For those familiar with the network, this section can probably be overlooked.

An MLP is based on Rosenblatt's single perceptron (Warwick, 2011). In this study we are using conventional two-layered MLP network as classifier. The network contains an input layer, a hidden layer and an output layer. There is however merely a single neuron in the output layer. The computation of the network's overall output, y, was based on the following Eq. (2):

$$y = g\left(\sum_{j=1}^{m} w_j^2 g\left(\sum_{i=1}^{d} W_{ji}^1 X_i + b_j^1\right) + b^2\right) \tag{2}$$

where, in our case, for the results presented, $d = 6$ and $m = 8$.

All 6 of the input neurons were fully connected to the hidden layer. The single output neuron was, in turn, fully connected with the 8 hidden neurons, each one being defined by (2). The transfer function g for both hidden and output layer neurons was a tan-sigmoid function described by (3).

$$tansig(n) = \frac{2}{(1 + e^{-2n})} - 1 \tag{3}$$

where n is the overall summed input to that neuron respectively.

For the tremor classification problem, the MLP was required to output 1 for tremor status and −1 for non-tremor status. The training data was generated from LFP datasets according to the EMG evaluation, as was discussed in Section 2. For balance, the same amount of pre-tremor and tremor training data was presented to the MLP. The MLP training error output was then compared with the EMG evaluation output, which was used with the Levenberg–Marquardt (LM) backpropagation algorithm to train the network. In the LM method, the change in weights, Δw, was defined by

$$\Delta w = [H + aI]^{-1} g \tag{4}$$

where w is the vector of all adjustable parameters (weights and biases), H is the Hessian matrix and a is a regularising parameter.

The network training was curtailed when the validation output started to increase over a window, the aim being to which will reduce the MLP overfit to the training dataset.

3.2. Radial Basis Function Neural Network

Radial Basis Function Network (RBN) were originally proposed by Lowe and Broomhead (1988). One of the key differences between RBN and MLP is the neuron transfer function inside the hidden and output layer. The RBN generally (and certainly in our case) uses a Gaussian function as shown in (5). The RBN consists of only two layers with a single neuron in the output layer. The output is computed according to (6), essentially it is merely a weighted linear combination of the hidden layer neuron outputs.

$$\varphi_i(x) = exp\left(\frac{-\|X - u_i\|^2}{\sigma_i^2}\right) \tag{5}$$

$$y = \sum_{i=1}^{N} w_i \varphi_i(X) + w_0 \tag{6}$$

where $X = [x_1, x_2, \ldots, x_m]^T \in \mathbb{R}^m$ is the input vector, $u_i = [u_{i1}, u_{i2}, \ldots, u_{im}]^T \in \mathbb{R}^m$ is the centre vector, and σ is the corresponding width of the ith node. N meanwhile, indicates the number of neurons in the hidden layer – in our case, for this comparative study it was found to be best when $N = 300$. A similar type of RBN method has been previously demonstrated to work well with this type of data (Wu, Warwick, Ma, Burgess, et al., 2010). Here however, for the comparison conducted, we have focussed on the specific data considered, in terms of signal frequency. There are two design aspects to an RBF network, (1) Determination of the parameters of the RBF neurons, which can also translate into Gaussian function widths; (2) Calculation of the connection weights between the hidden and output layers. In this study we used a fuzzy c-means clustering algorithm to determine the initial centres.

4. Support Vector Machine classification

In comparison with different artificial neural network classification methods, support vector machine (SVM) is a novel type of learning algorithm which presents a very promising tool for solving pattern recognition problems. In recent years, SVM learning has found a wide range of real-world application due to its good generalisation capabilities, e.g. drug/nondrug classification and the analysis sleep EEG activity during hypopnoea episodes (Byvatov et al., 2003; Ubeyli, Cvetkovic, Holland, & Cosic, 2010).

Support Vector Machine (SVM) approaches are based on Structural Risk Minimisation and Statistical Learning Theory and can handle a classification problem by converting it into either a quadratic programming problem, in the conventional SVM case, or a set of linear equations in the Least-Squares SVM case, respectively. The motivation for using SVM approaches for prediction of tremor onset stems from the fact that SVM models are simple to obtain and that they can possess higher generalisation potential.

In the SVM classification study performed, it was assumed that the data set for obtaining the optimal model was of the form given in (7):

$$\mathcal{D} = \{X_k; y_k\}_{k=1}^{k=N} \tag{7}$$

where $X_k \in \mathbb{R}^n$ is n-dimensional kth input vector, $y_k \in \{-1, +1\}$ is the corresponding binary output, and N is the size/number of available data. The aim was to find an SVM model that captured the association between the input and output data.

The primal form of the SVM classification model was given in (8), a linear model in a higher dimensional feature space $\boldsymbol{F}$.

$$\hat{y}_i = \langle w, \phi(X_i) \rangle \tag{8}$$

where w is a vector in the feature space $\boldsymbol{F}$, $\phi(\cdot)$ meanwhile is a mapping from the input space to the feature space, and $\langle \cdot \rangle$ is the inner

product operation in $\boldsymbol{F}$. The SVM classification algorithm formulated the classification problem as an optimization problem in dual space in which the model was given by (9).

$$\hat{y}_i = \sum_{j=1}^{N_{Tr}} a_j y_j K(X_i, X_j) \tag{9}$$

where N_{Tr} is the size/number of training data, a_j is the Lagrange coefficient corresponding to the training data point X_j, and $K(X_i, X_j)$ is the Gaussian kernel function given by $K(X_i, X_j) = \langle \phi(X_i), \phi(X_j) \rangle = e^{\frac{X_i - X_j^2}{2\sigma^2}} = K_{ij}$. The role of the kernel function is that it handles the inner product in the feature space, thereby eliminating the necessity of the explicit form of $\phi(\cdot)$. In the SVM model given by (9), a training point X_j corresponding to a non-zero a_j value – this being called the *support vector*.

Turning back to the primal form of the classification problem, this was taken to be:

$$\min_{w,b,\xi,\xi^*} P = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N_{Tr}} \xi_i \tag{10}$$

subject to the constraints,

$$y_i \langle w, \phi(X_i) \rangle \leq 1 - \xi_i, \quad i = 1, \ldots, N_{Tr}$$

and

$$\xi_i \geq 0, \quad i = 1, \ldots, N_{Tr} \tag{11}$$

where ξ_i's are slack variables, $\|\cdot\|$ is the Euclidean norm, and C is a regularization parameter that penalizes the errors made by the model.

In order to attain the dual form of the classification problem, first, we needed the Lagrangian, which was obtained by adding the constraints to the primal form of the problem as

$$L_p = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N_{Tr}} \xi_i - \sum_{i=1}^{N_{Tr}} a_i (y_i \langle w, \phi(X_i) \rangle - 1 + \xi_i) - \sum_{i=1}^{N_{Tr}} \mu_i \xi_i \tag{12}$$

where a_i's and μ_i's are Lagrange multipliers.

Afterwards, the first-order optimality conditions of the primal problem were obtained by taking partial derivatives of L_p with respect to the primal variables and then setting them to zero as follows:

$$\frac{\partial L_p}{\partial w} = 0 \mapsto w = \sum_{i=1}^{N_{Tr}} a_i y_i \phi(X_i) \tag{13}$$

$$\frac{\partial L_p}{\partial w} = 0 \mapsto C - a_i - \mu_i = 0, \quad i = 1, \ldots, N_{Tr} \tag{14}$$

The dual form of the problem becomes a Quadratic Programming (QP) problem as follows:

$$\min_a D = \frac{1}{2}\sum_{i=1}^{N_{Tr}}\sum_{j=1}^{N_{Tr}} a_i a_j y_i y_j K_{ij} - \sum_{i=1}^{N_{Tr}} a_i \tag{15}$$

subject to the constraints,

$$\sum_{j=1}^{N_{Tr}} a_i y_i = 0 \text{ and } 0 \leq a_i \leq C, \quad i = 1, \ldots, N_{Tr} \tag{16}$$

The solution of the QP problem given by (15) and (16) yields the optimum values of the dual variables (a_i's or Lagrange multipliers). Moreover, by considering only the support vectors, the model has a more sparse form as follows:

$$\hat{y}_i = \sum_{\substack{j=1 \\ j \in SV}}^{\#SV} a_j y_j K(X_i, X_j) \tag{17}$$

where #SV stands for the number of support vectors in the model. Thus the whole training data can be represented by only support vectors which constitute a small part of the original training data. The parameters of the SVM classification model are in fact the regularization parameter C and the kernel parameter σ (Alpaydin, 2010; Shawe-Taylor & Cristianini, 2000).

It is clear that different modelling parameters result in different models. Therefore, it is unavoidable to search for the optimal modelling parameters in the parameter space. For this purpose, the data set $\mathcal{D}$ is divided into two subsets: training data ($\mathcal{D}_{Tr}$) and test data ($\mathcal{D}_{Te}$). Then, in order to find the best SVM model, a grid search approach is adopted, where the parameter space is split into grids, and for each node (corresponding to a specific parameter set) on the grid, a model was obtained by using the training data ($\mathcal{D}_{Tr}$). Then, the model that gave the least testing error based on the test data ($\mathcal{D}_{Te}$) was chosen as the optimal model.

5. Experimental results

The data set used in the classification of tremor status, with all three classifiers, was divided into two subsets as shown in Table 1. There were a total of 69 pieces of data exclusively used for training and 53 pieces of data exclusively used for testing. In order to avoid unreliable and over-optimistic performance results, at no time was a piece of testing set data used in any way in the training procedures, and vice versa. In addition both training and testing data were used as continuing recordings without any mixing, dividing or otherwise tampering in order to enhance the results. What we wanted to achieve was a real test, as though working on actual LFP data as collected.

The performance of the three classifiers was evaluated by means of a Receiver Operating Characteristics (ROC) graph. ROCs have been widely used with medical decision and detection algorithms and appeared to be appropriate here. In this study we defined a true positive (TP) as when the clinical EMG signal showed a positive tremor signal and the classifier output also indicated that tremor was present. A true negative (TN) was defined as when the EMG signal showed non-tremor and the classifier output agreed with the clinical state. Meanwhile false positive (FP) indicated when the classifier output said tremor but the clinical EMG signal showed non-tremor to be presented. The last case was when the classifier output indicated a non-tremor status whilst the clinical EMG signals showed that the tremor was actually present in the patients arm.

The ROC confusion matrix is shown in Table 2. There were several performance measures derived from the TP/FP measures, in this analysis we used sensitivity and specificity to present how each classifier performance compared with each other. Sensitivity, also called true positive rate represented how frequently each classifier reported that a tremor existed in the instance where one actually did so exist. The sensitivity can range from as low as 0, indicating that no actual tremor signals were detected, to as high

Table 1
Classifier training and validation dataset.

Dataset	Total length (s)	Non-tremor pieces	Tremor pieces	Sampling frequency (hz)	Contact location	EMG location
Training	34.5	45	24	1000	L STN	Extensor
Validation	26.5	33	20	1000	L STN	Extensor

Table 2
Confusion matrix.

	Classifier decision	
Clinical EMG indication		Tremor presenting	Non-tremor presenting
	Tremor presenting	True positive (TP)	False negative (FN)
	Non-tremor presenting	False positive (FP)	True negative (TN)

as 1, indicating that absolutely all tremor signals were correctly detected with no errors. Clearly, as stated in Section 1, it is important for there to be few false negatives – this is indicated by a high sensitivity value.

The other measurement was the false positive rate, which represented how frequently the tremor status was detected by a classifier when a tremor was not observably apparent and vice versa. The false positive rate also ranged between 0 and 1, where 0 indicates that the classifier correctly identified all incidents and 1 indicated that a classifier misidentified all incidents. In this case a figure close to 0 was preferable, however this statistic was not deemed to be as important as sensitivity.

The ROC graph represents the two values (sensitivity and specificity) in a plane where the true positive rate (sensitivity) is indicated on the x-axis and the false positive rate is indicated on the y-axis. The ideal classifier should have pair of values close to top left hand corner of the graph, to show that it correctly classified tremor status, and exhibited a low false alarm rate.

where,

$$\text{True Positive Rate} = \frac{TP}{TP + FN}$$

and

$$\text{False Positive Rate} = \frac{FP}{FP + TN} \tag{18}$$

All three classifiers operated on a classification output range set between -1 and 1. We therefore used a threshold value of 0 to classify the output as shown below

$$tremor = classifier \geq 0$$

$$nontremor = classifieroutput < 0$$

Fig. 2 presents a comparison of the output selection from all three classifiers. The EMG evaluation – taken from visual inspection – is indicated with a blue line, value 0 indicates non-tremor status and 1 indicates tremor status. The red line is the classifier output base on the power spectrum density feature input at the corresponding time scale. The output of all three classifiers is presented between -1 and 1, where the threshold value for tremor detection was set to 0. The data input is from the validation shown in Table 1.

It is apparent from the EMG data that the PD tremor starts at around 17 s. In general all the classifiers managed to indicate the difference between tremor and none tremor status. In addition however all three classifiers gave some indication of tremor status between 5–12.5 s even though the actual tremor only began at 17 s. This indicates the PSD of the STN LFP starts to appear close to the tremor PSD of the LFP around 5 s before the actual tremor starts, this result matches with the medical experience of the neurosurgeons involved.

The results shown in Fig. 3 and Table. 3 indicate that the MLP had the lowest overall accuracy rate when compared with the other two classifiers – accuracy being given by TP + TN divided by all classifications. However it did have a similar false positive rate when compared with the SVM and importantly its sensitivity was significantly better that that of the RBN.

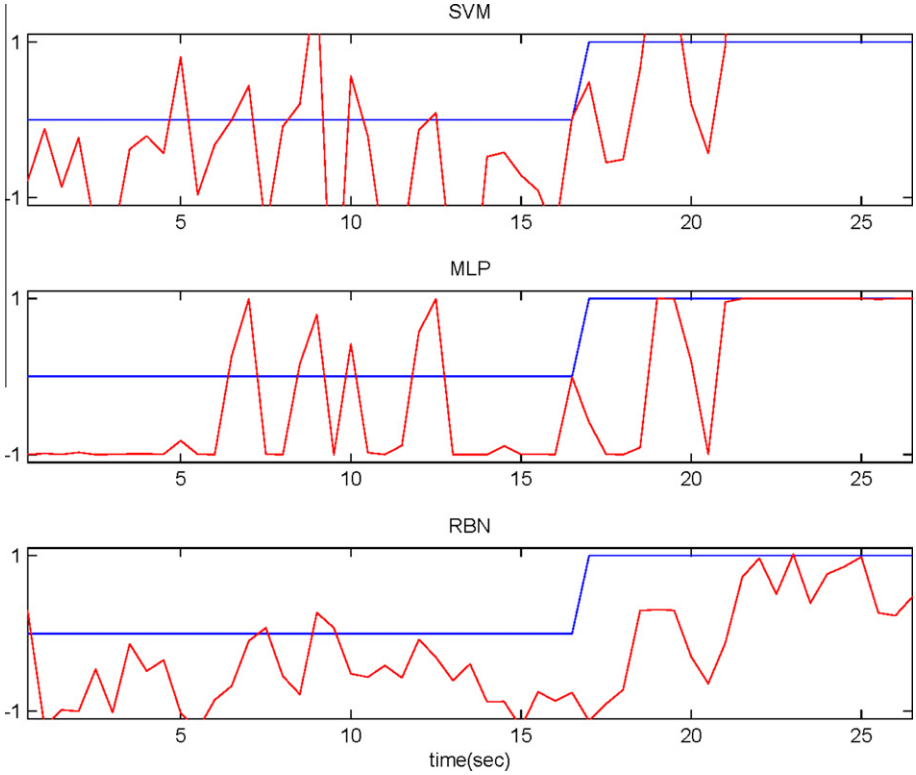

Fig. 2. Three classifiers validation output.

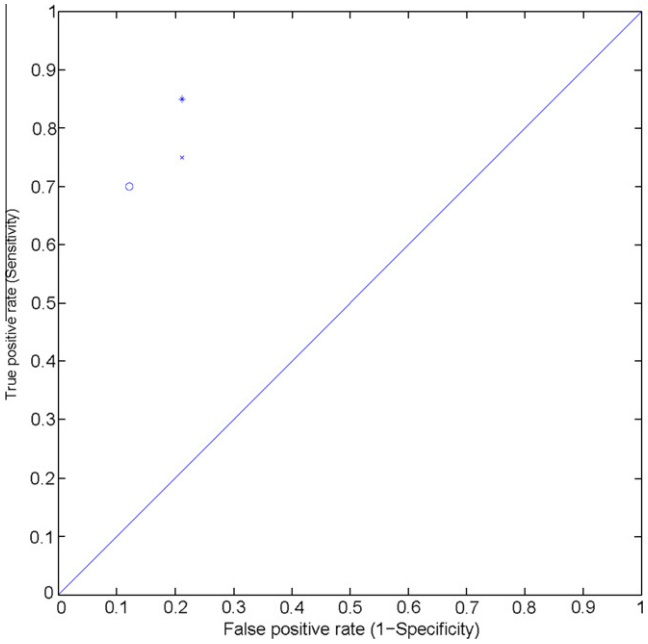

Fig. 3. SVM (∗), RBN (o) and MLP (*x*).

Table 3
Classifier comparison.

Classifier	Input dimensions	Hidden neurons	Threshold value	True positive rate (%)	False positive rate (%)	Accuracy (%)
MLP	6	8	0	75.01	21.21	79.25
RBN	6	300	0	70.04	12.17	80.13
SVM	6	xxx	0	85.06	20.95	81.14

In combining both visual inspection and numerical results it does appear that the MLP seems to be able to detect the PSD change in STN-LFP before the tremor actually occurs – something that has previously been suggested as a predictive quality. However it was not so good at detecting the PSD changes when the true event actually occurred. This is perhaps why the MLP produced the lowest accuracy of all three classifiers.

The RBN meanwhile managed to score highest/best on the false positive rate by some margin but achieved a much lower score on the much more important aspect of sensitivity at 70% – significantly lower than the other two techniques. In essence it appears that the RBN was rather more (perhaps too much) "conservative" with PSD changes when compared with the other two classifiers – only giving a response when it was "sure".

Overall the SVM achieved the highest accuracy rate of 81.14% with also the highest sensitivity rate at 85%, as shown in Fig. 3. Remember the sensitivity was the important factor in this study. The SVM was able to detect both pre-tremor PSD changes and also appeared to correctly detect the tremor onset at 17 s.

It is not clear why all three classifiers failed to successfully/accurately detect the continuation of EMG indicated tremor after onset. Indeed all three classifiers output non-tremor values at between 18–19 and 21 s which indicated perhaps that the PSD showed pattern matching more closely aligned to non-tremor status than tremor status. This point is discussed further in the next section.

6. Conclusions

The aim of this study was to compare the performance of SVM, MLP and RBN on a specific medical application, namely the classification of a Parkinson's Disease patient's STN-LFP signal as being associated with tremor or non-tremor periods. In comparison with previous research, in this study we focused on classifying two different tremor states in the STN-LFP signal using signal frequency in the band 3–10 Hz as the only input feature as this region had been indicated as being of major significance in related research. The results obtained showed that the SVM generally out performed MLP and RBN in terms of overall accuracy and, most importantly, with a much lower false negative rate.

In previous studies of this kind the focus has been on using artificial neural networks such as MLP and RBN to find a pattern that can be used to predict tremor onset. In this study we were primarily focusing on using classifiers to distinguish between two states of the STN-LFP signal. In the process of this study however, we found that all three classifiers identified a tremor pattern of some kind approximately 8–12 s prior to actual onset. It appears that all three classifiers spike at a similar point in time, which has led us to believe this might not be merely coincidence and may indeed be used as warning pattern.

It is worth pointing out that in this particular study all three classifiers incorrectly classified the first few seconds after tremor onset as still being in a non-tremor state. Strictly speaking this type of false negative classification should be avoided at all times, since as a result no DBS stimulation might be applied resulting in the patient suffering from tremors. The actual underlying reason for this is unclear however as all three classifiers had issues at this point it may well indicate the underlying nature of the PD signal.

The point here is that the study carried out may well not only assess the performance of three classifiers on a particular dataset but also open up a better understanding of the PD problem and the nature of tremor onset itself. As can be seen from the signals in Fig. 1 it is not a simple case of there being no tremor at all at one point in time and then suddenly there are definitely tremors. The patient's brain is adjusting and locking in to different cycles over a relatively short period of time and hence there are indeed probably indications that tremors are about to start before they 'officially' do but also once tremors have 'officially' started there are periods when they weaken and may even almost stop.

What this means is that the function which we want the classifiers to adhere to is a very poor one because it is not simply a case of yes or no with the human brain in the situation of PD onset. So the blue line[1] in Fig. 2 is merely arrived at due to our best guess at what the situation in the STN is and how that relates to PD tremors. Hence the three classifiers are doing their best under difficult circumstances. In light of this a rate of 85% success must be seen as an absolutely excellent result. Part of the remaining 15% being just as much (maybe more) due to inaccuracies in our blue line as in the performance of the classifier in question.

It is acknowledged that we did not directly compare here the computational power required for training and operating each classifier and that this would, in the end result practical stimulator, affect battery life. Rather we were more concerned here with differences in performance. Indeed battery life and other practical issues will be considered further in the future studies. We also have presented here a comparative analysis on one particularly difficult patient data study. There is in fact a lot more to be learnt before such classifiers can be used in a practical feedback stimulator where an AI classifier actually switches on and off the stimulator.

The nature of signals measured during the onset of PD tremor vary considerably between patients – partly due to the individuals concerned and partly due to electrode placement in the brain, but results indicate that it may also be the case that the disease is

[1] For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.

exhibited differently in different patients. This is understandable as "Parkinson's Disease" is a name given to a wide class of neurological disorders that affect different people in different ways. So studies on broader classes of patients, e.g. a 10 patient study, may well be of interest in due course but may well also be extremely confusing because one cannot be at all clear as to whether differences in the results are down to differences in classifier performance or due to differences in the nature of the disease in different patients. It is essentially very easy in multi-patient studies to jump to very wrong conclusions. What we have tried to show here is merely how well (or badly) some different classifiers can perform in identifying tremor onset in a typical patient.

Acknowledgements

The research results presented here are based on data obtained by Professor Tipu Aziz and his team at the University Laboratory of Physiology, University of Oxford and the Functional Neurosurgery Group, Department of Surgery, Radcliffe Infirmary, Oxford, UK.

References

Alpaydin, E. (2010). *Introduction to machine learning*. The MIT Press.

Benabid, A. L., Chabardes, S., Mitrofanis, J., & Pollak, P. (2009). Deep brain stimulation of the subthalamic nucleus for the treatment of Parkinson's disease. *The Lancet Neurology, 8*(1), 67–81. http://dx.doi.org/10.1016/S1474-4422(08)70291-6.

Bour, L. J., Contarino, M. F., Foncke, E. M. J., Bie, R. M. A., Munckhof, P., Speelman, J. D., et al. (2010). Long-term experience with intraoperative microrecording during DBS neurosurgery in STN and GPi. *Acta Neurochirurgica, 152*, 2069–2077. http://dx.doi.org/10.1007/s00701-010-0835-y.

Breit, S., Schulz, J. B., & Benabid, A.-L. (2004). Deep brain stimulation. *Cell and Tissue Research, 318*(1), 275–288.

Brown, P., & Williams, D. (2005). Basal ganglia local field potential activity: Character and functional significance in the human. *Clinical Neurophysiology, 116*(11), 2510–2519. http://dx.doi.org/10.1016/j.clinph.2005.05.009.

Burgess, J. G., Warwick, K., Ruiz, V., Gasson, M. N., Aziz, T. Z., Brittain, J.-S., et al. (2010). Identifying tremor-related characteristics of basal ganglia nuclei during movement in the Parkinsonian patient. *Parkinsonism & Related Disorders, 16*(10), 671–675. http://dx.doi.org/10.1016/j.parkreldis.2010.08.025.

Byvatov, E., Fechner, U., Sadowski, J., & Schneider, G. (2003). Comparison of support vector machine and artificial neural network systems for drug/nondrug classification. *Journal of Chemical Information and Computer Sciences, 43*(6), 1882–1889.

Cassidy, M., Mazzone, P., Oliviero, A., Insola, A., Tonali, P., Lazzaro, V. D., et al. (2002). Movement-related changes in synchronization in the human basal ganglia. *Brain, 125*(6), 1235–1246. http://dx.doi.org/10.1093/brain/awf135.

Cheung, T., & Tagliati, M. (2010). Deep brain stimulation: Can we do it better? *Clinical Neurophysiology, 121*(12), 1979–1980. http://dx.doi.org/10.1016/j.clinph.2010.05.024.

Gasson, M. N., Wang, S. Y., Aziz, T. Z., Stein, J. F., & Warwick, K. (2005). Towards a demand driven deep-brain stimulator for the treatment of movement disorders. In: *3rd IEE International Seminar on Medical Applications of Signal Processing* (pp. 83–86). London.

Lemstra, A. W., Verhagen Metman, L., Lee, J. I., Dougherty, P. M., & Lenz, F. A. (1999). Tremor-frequency (3–6 Hz) activity in the sensorimotor arm representation of the internal segment of the globus pallidus in patients with Parkinson's disease. *Neuroscience Letters, 267*(2), 129–132. 16/S0304-3940(99)00343-2.

Levy, R., Ashby, P., Hutchison, W. D., Lang, A. E., Lozano, A. M., & Dostrovsky, J. O. (2002). Dependence of subthalamic nucleus oscillations on movement and dopamine in Parkinson's disease. *Brain, 125*(6), 1196–1209. http://dx.doi.org/10.1093/brain/awf128.

Lowe, D., & Broomhead, D. (1988). Multivariable functional interpolation and adaptive networks. *Complex systems, 2*, 321–355.

Moran, A., Bergman, H., Israel, Z., & Bar-Gad, I. (2008). Subthalamic nucleus functional organization revealed by parkinsonian neuronal oscillations and synchrony. *Brain, 131*(12), 3395–3409. http://dx.doi.org/10.1093/brain/awn270.

Pan, S., Warwick, K., Stein, J., Gasson, M. N., Wang, S. Y., Aziz, T. Z., & Burgess, J. (2007). Prediction of Parkinson's disease tremor onset using artificial neural networks. In *Proceedings of the fifth IASTED International Conference: biomedical engineering* (pp. 341–345). Presented at the 5th IASTED international conference on biomedical engineering, Innsbruck, Austria: ACTA Press. Retrieved from http://portal.acm.org/citation.cfm?id=1295494.1295556.

Shawe-Taylor, J., & Cristianini, N. (2000). *An introduction to support vector machines and other kernel-based learning methods*. UK: Cambridge University Press.

Sutanto, E., Mason, J., & Warwick, K. (1997). Mean-tracking clustering algorithm for radial basis function centre selection. *International Journal of Control, 67*(6), 961–977. http://dx.doi.org/10.1080/002071797223884.

Ubeyli, E. D., Cvetkovic, D., Holland, G., & Cosic, I. (2010). Analysis of sleep EEG activity during hypopnoea episodes by least squares support vector machine employing AR coefficients. *Expert Systems with Applications, 37*(6), 4463–4467. http://dx.doi.org/10.1016/j.eswa.2009.12.065.

Wang, S., Aziz, T. Z., Stein, J. F., Bain, P. G., & Liu, X. (2006). Physiological and harmonic components in neural and muscular coherence in Parkinsonian tremor. *Clinical Neurophysiology, 117*(7), 1487–1498.

Wang, S.-Y., Aziz, T. Z., Stein, J. F., & Liu, X. (2005). Time-frequency analysis of transient neuromuscular events: dynamic changes in activity of the subthalamic nucleus and forearm muscles related to the intermittent resting tremor. *Journal of Neuroscience Methods, 145*(1–2), 151–158. http://dx.doi.org/10.1016/j.jneumeth.2004.12.009.

K., Warwick (2011). *Artificial intelligence: The basics*. Taylor & Francis.

Weinberger, M., Hutchison, W. D., Lozano, A. M., Hodaie, M., & Dostrovsky, J. O. (2009). Increased gamma oscillatory activity in the subthalamic nucleus during tremor in Parkinson's disease patients. *Journal of Neurophysiology, 101*(2), 789–802. http://dx.doi.org/10.1152/jn.90837.2008.

Wu, D., Warwick, K., Ma, Z., Burgess, J. G., Pan, S., & Aziz, T. Z. (2010). Prediction of Parkinson's disease tremor onset using radial basis function neural networks. *Expert Systems with Applications, 37*(4), 2923–2928. http://dx.doi.org/10.1016/j.eswa.2009.09.045.

Wu, D., Warwick, K., Ma, Z., Gasson, M. N., Burgess, J. G., Pan, S., et al. (2010). Prediction of Parkinson's disease tremor onset using a radial basis function neural network based on particle swarm optimization. *International Journal of Neural Systems, 20*(02), 109. http://dx.doi.org/10.1142/S0129065710002292.

Zheng, X., Yang, B., Li, X., Zan, P., & Dong, Z. (2010). Classifying EEG Using Incremental Support Vector Machine in BCIs. In K. Li, L. Jia, X. Sun, M. Fei, & G. W. Irwin (Eds.), *Life System Modeling and Intelligent Computing* (Vol. 6330, pp. 604–610). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://www.springerlink.com/content/1488160k6w556181/.