Deep Learning Applied to Deep Brain Stimulation in Parkinson's Disease

Pablo Guillén(✉)

Center for Advanced Computing and Data Systems, CACDS,
University of Houston, Houston, TX, USA
pgrondon@uh.edu

Abstract. In order to better model complex real-world data such as biomedical signals, one approach is to develop pattern recognition techniques and robust features that capture the relevant information. In this paper, we use deep learning methods, and in particular multilayer perceptron, to build an algorithm that can predict subcortical structures of patients with Parkinson's disease, based on microelectrode records obtained during deep brain stimulation. We report on experiments using a data set involving 52 microelectrode records for the structures: zona incerta, subthalamic nucleus, thalamus nucleus, and substantia nigra. The results show that the combination of features and deep learning produces 99.2% precision of detection and classification on the average of the subcortical structures under study. In conclusion, based on the high precision obtained in the classification, deep learning could be used to predict subcortical structure, and mainly the subthalamic nucleus for neurostimulation.

1 Introduction

Human information processing mechanisms (e.g. vision and speech) suggest the need of deep architectures for extracting complex structure and building internal representation from rich sensory inputs. For example, human speech production and perception systems are both equipped with clearly layered hierarchical structures in transforming information from the waveform level to the linguistic level and vice versa. It is natural to believe that the state of the art can be advanced in processing these types of media signals if efficient and effective deep learning algorithms are developed. Biomedical signals processing systems with deep architectures are composed of many layers of nonlinear processing stages, where each lower layer's outputs are fed to its immediate higher layer as the input. The concept of deep learning originated from artificial neural network research. Multilayer perceptron with many hidden layers is a good example of the models with deep architectures [1].

Deep learning techniques have been applied to a wide variety of problems in recent years [2–7]. In many of these applications, algorithms based on deep learning have surpassed the previous state-of-art performance. At the heart of all deep learning algorithms is the domain independent idea of using hierarchical layers of learned abstraction to efficiently accomplish high-level task. Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. Deep learning discovers intricate structure in

C.J. Barrios Hernández et al. (Eds.): CARLA 2016, CCIS 697, pp. 269–278, 2017.
DOI: 10.1007/978-3-319-57972-6_20

large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer [1].

Parkinson disease (PD) is thought to affect at least 100 persons in every 100,000. The cardinal symptoms of tremor, bradykinesia, postural instability, and rigor result in substantial disability for patients with PD. During the course of the disease, up to 50% of patients will have symptoms refractory to medication and will experience drug-induced dyskinesias. Over activity of the globus pallidus internus (GPi) and the subthalamic nucleus (STN) is believed to be part of the pathophysiologic mechanism of PD. PD is a chronic progressive neurodegenerative disorder affecting multiple brain circuits leading to motor symptoms such as bradykinesia, rigidity, resting tremor, and loss of postural reflexes [8]. PD also has non-motor manifestations such as neuropsychiatric symptoms, cognitive abnormalities, autonomic disorders, and sleep [9].

PD is primary related to substantia nigra degeneration and, thus, dopamine insufficiency. L-DOPA as a precursor of dopamine is the standard medication in PD. However, disease progression causes L-DOPA therapy efficiency decay (on-off symptom fluctuation), and neurologists often decide to classify patients for DBS (Deep Brain Stimulation) surgery.

DBS [10] is considered a safe and well-tolerated surgical procedure to alleviate PD and other movement disorders symptoms along with some psychiatric conditions. Over the last few decades DBS has been shown to provide remarkable therapeutic effect on carefully selected patients. DBS improves motor functions and therefore quality of life. To date, one main target has emerged in PD patients: the subthalamic nucleus.

DBS involves the surgical implantation of electrodes into deep structures of the brain to modulate brain circuitry in an effort to restore normal physiological function. DBS has been used effectively for the treatment of movement disorders, including PD, Essential tremor (ET), and dystonia, as well as for psychiatric disorders such as obsessive-compulsive disorder (OCD). In addition, DBS may exert its influence via the correction of aberrant neuronal activity. For example, in the setting of DBS for the treatment of PD, the loss of dopamine, which is known to be largely responsible for the pathophysiology of the disease, results in changes in the underlying activity of cells within the basal ganglia [11].

The exact placement of the stimulator is fundamental for the sensory and motor effects specific to the subthalamus, since small deviations can affect adjacent structures and generate side effects [12]. Technological development has allowed a higher display resolution for the imaging processing of these types of structures [13]. Also, the development of recording systems for the spontaneous or induced electrical activity of such structures, allows defining more accurately their position limits, anatomic relationships with adjacent structures and behavior in relation to the movement or symptoms of Parkinsonism components, which turns in a suitable situation in order to safely establish and characterize the affected area. The risk of errors in the localization of the surgical target for deep brain electrical stimulation requires the use of some form of intraoperative neurophysiological monitoring to confirm the correct destination during surgery. The purpose of the development and implementation of techniques of

classification for the processing microelectrode records (MER) is of great importance today, as it allows the surgical team to determine the optimal location of the lesion or DBS [14, 15].

In this paper we build a set of features each of which measure different signal characteristics, quantify neuronal activity inherent in the subcortical structures, and allows their localization for greater accuracy during DBS. Next, a machine learning algorithm for supervised classification based on deep learning is used with the values of the obtained features, and we show that deep learning can identify and predict with high precision any of the subcortical structures (4 classes): Thalamus (TAL), zona incerta (Zi), subthalamic nucleus (STN) and substantia nigra (SNR).

2 Materials and Methods

2.1 Dataset

Intra-operative microelectrode records were acquired in Parkinsonian patients, awake and unmedicated, subject to deep brain implantation under electrostimulation. Five Parkinsonian patients (4 males and 1 female) aged between 55 ± 6 years old participated voluntarily assuming previously signed consent. Microelectrode records were made using the ISIS MER Inomed system, which is used to obtain an optimal location of the destination (target) through deep brain stimulation. Visualization of neural data started 10 mm on the target data. Each 1 mm a new location was created if the distance between the microelectrode and the destination point was greater than 3 mm. For distances less than 3 mm, the locations were created every 0.5 mm. Specialists in neurosurgery and neurophysiology labeled the obtained signals using the MER system. The acquisition time for each record was 2 s with a sampling frequency of 24 kHz (24,000 samples per second). In total, the database comprises 52 micro recordings, 13 for each of the subcortical structures: Thalamus nucleus (TAL), Zona Incerta (Zi), Subthalamic nucleus (STN) and Substantia Nigra (SNR). These surgical procedures were performed at the Institute of Epilepsy and Parkinson in Pereira, Colombia.

2.2 Features

In the following we present the features to extract the information contained in each of the MER for the different subcortical structures under study.

Curve Length. This feature is useful to know the stability of the values of a signal. If in a given interval, the value of this feature is low, it is an indicative that the signal is stable, otherwise the signal is unstable. Equation (1) defines the calculation of this feature:

$$L = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{1}$$

where each x_i corresponds to a sample in the dataset $X = (x_1, x_2, \ldots, x_N)$.

Threshold. Computation of the threshold is based in the calculation of the deviation of the data in order to know how scattered they are in given window of size *N*. Threshold is calculated as follows:

$$g = \frac{3}{N-1}\sqrt{\sum_{i=1}^{N}(x_i - \overline{X})^2} \tag{2}$$

where $\overline{X}$ is the mean of the dataset.

Peaks. The number of peaks that a given signal has is determined by:

$$\kappa = \frac{1}{2}\sum_{i=1}^{N-2} max\{0, |sgn[x_{i+1} - x_i] - sgn[x_{i+2} - x_{i+1}]|\} \tag{3}$$

where

$$max(a,b) = \begin{cases} a & if \quad a > b \\ b & if \quad a < b \\ a \; o \; b & if \quad a = b \end{cases}$$

$$sgn(x) = \begin{cases} 1 & if \quad x > 0 \\ 0 & if \quad x = 0 \\ -1 & if \quad x < 0 \end{cases}$$

Root Mean Square. It is defined as the square root of the mean of the squares of the values of the signal. The root mean square (also known as quadratic mean) is determined by:

$$d = \sqrt{\frac{\sum_{i=1}^{N} x_i^2}{N}} \tag{4}$$

Average Nonlinear Energy. The average nonlinear energy is computed as:

$$ANE = \frac{1}{N-2}\sum_{i=2}^{N-1}(x_i^2 - x_{i-1}x_{i+1}) \tag{5}$$

Zero Crossings. The amount of zero crossings k for a given signal is determined through the formula:

$$k = \frac{1}{2}\sum_{i=1}^{N-1} |sgn(x_{i+1}) - sgn(x_i)| \tag{6}$$

2.3 Deep Learning

The concept of deep learning originated from artificial neural network research. Multilayer perceptron with many hidden layers is a good example of the models with deep architectures. Unlike the neural networks of the past, modern deep learning has cracked the code for training stability and generalization and scale on big data. It is often the algorithm of choice for highest predictive accuracy, as deep learning algorithms performs quite well in a number of diverse problems.

There are several theoretical frameworks for deep learning, and here we summarize the feedforward architecture used by H20 [16]. Multilayer perceptron (MLP) are feed-forward neural networks with architecture composed of the input layer, the hidden layer and the output layer. Each layer is formed from small units known as neurons. Neurons in the input layer receive the input signals X and distribute them forward to the rest of the network. In the next layers, each neuron receives a signal, which is a weighted sum of the outputs of the nodes in the previous layer. Inside each neuron, an activation function is used to control the input, (Fig. 1 shows an example). Such a network determines a non-linear mapping from an input vector to the output vector, parameterized by a set of network weights, which are referred to as the vector of weights W. The first step in approximating the weight parameters of the model is finding the appropriate architecture of the MLP, where the architecture is characterized by the number of hidden units, the type of activation function, as well as the number of input and output variables. The second step estimates the weight parameters using the training set. Training estimates the weight vector W to ensure that the output is as close to the target vector as possible. The structure of a MLP network is shown in Fig. 2. This basic framework of MLP neural networks can be used to accomplish deep learning task. Deep learning architectures are models of hierarchical feature extraction, typically involving multiple levels of nonlinearity.

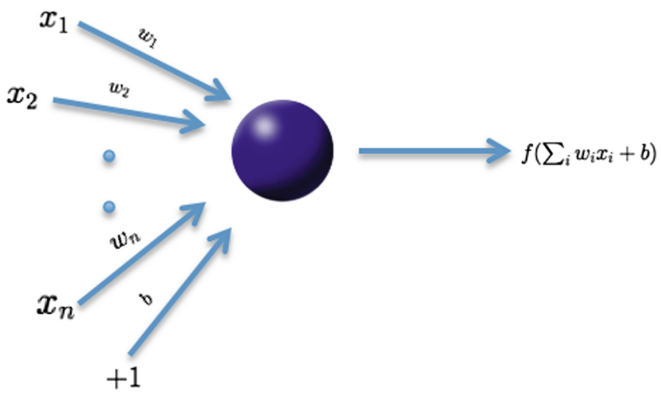

Fig. 1. The function *f* represents the nonlinear activation function used throughout the network and the bias *b* represents the neuron's activation threshold.

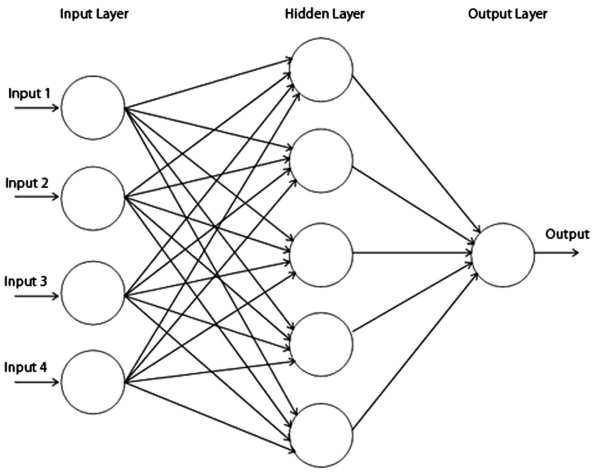

Fig. 2. Structure of an architecture multilayer perceptron.

2.4 Tools

The features described before were implemented in Matlab R2015, and the Library H20 [16] was used in order to perform the classification through deep learning.

3 Results

As mentioned before when we presented the database, each record was acquired for 2 s at a sampling frequency of 24 kHz, which leads to each record having 48,000 samples. If we consider a trajectory of 13 records for each of the subcortical structures: Thalamus nucleus (TAL), Zona Incerta (Zi), Subthalamic nucleus (STN) and Substantia Nigra (SNR), the final trajectory is made up of 52 records and has a total of 2,496,000 samples. Next, the final trajectory is divided into windows 4992 consecutive samples and for each of these windows the six features are determined, yielding a total of 500 instances (patterns) by feature. The decomposition described above is presented in matrix form as follows (Fig. 3):

$$X = \begin{array}{c} \\ X_1 \\ X_2 \\ \vdots \\ \vdots \\ X_n \end{array} \begin{array}{c} \begin{array}{ccccc} V_1 & V_2 & \cdots & \cdots & V_p \end{array} \\ \begin{bmatrix} x_{11} & x_{12} & \cdots & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & \cdots & x_{np} \end{bmatrix} \end{array}$$

Fig. 3. Feature matrix.

Once the feature matrix is assembled, we proceed to label the first 125 instances as Class 1, which correspond to the Thalamus; the next 125 instances as Class 2, corresponding to Zona Incerta; the following 125 instances as Class 3, representing the Subthalamic nucleus; and finally, the last 125 instances as Class 4, corresponding to the Substantia Nigra. In our approach, we train a MLP on a set of randomly selected features, approximately 60%, extracted from the feature matrix, then approximately 20% are used as the validation set, and approximately 20% are used as the testing set. The purpose of this training is to learn the multilayer architectures by simple stochastic gradient descent. The backpropagation procedure to compute the gradient of an objective function with respect to the weights of a multilayer stack of modules is nothing more than a practical application of the chain rule for derivatives. The key insight is that the gradient of the objective with respect to the input of a module can be computed by working backward from the gradient with respect to the output of that module. The backpropagation equation can be applied repeatedly to propagate gradients through all modules, starting from the output at the top (where the network produces its prediction) all the way to the bottom (where the external input is fed) [1]. Multilayer feedforward neural networks consist of many layers of interconnected neuron units, starting with an input layer to match the feature space, followed by multiple layers of nonlinearity, and ending with a linear regression or classification layer to match the output space. The weights linking neurons and biases with other neurons fully determine the output of the entire network, and finally learning occurs when these weights are adapted to minimize the error on labeled training data. To go from one layer to the next, the weighted sum of their inputs from the previous layer pass the result through a non-linear function. At the present, the most popular non-linear function is the rectified linear unit (ReLU). Variable epochs correspond to the numbers of passes over the training data set. Table 1 shows the size of the architecture multilayer perceptron and parameters used on the experiments to evaluate the classification.

Table 1. Parameters of the architecture multilayer perceptron.

Variables	Parameters
Input	500
Hidden	(32,32,32)
Output	4
Activation function	ReLU
Loss function	Mean squared error
Epochs	1000000

Tables 2 and 4 show the output report by H20 on train data and validation data, respectively, and Tables 3 and 5 show their confusion matrices.

We can see that the results in the training set and validation set are consistent with the architecture and variables used for classification.

Table 2. Precision reported on train data.

Deep learning ** Reported on train data. **
MSE: 2.19526e-11 R^2: 0.9999

Table 3. Confusion matrix on train data.

Class_2	Class_1	Class_3	Class_4	Error	Rate
77.0	0.0	0.0	0.0	0.0	0/77
0.0	77.0	0.0	0.0	0.0	0/77
0.0	0.0	79.0	0.0	0.0	0/79
0.0	0.0	0.0	73.0	0.0	0/73
77.0	77.0	79.0	73.0	0.0	0/306

Table 4. Precision reported on validation data.

Deep learning ** Reported on validation data. **
MSE: 0.0098 R^2: 0.9925

Table 5. Confusion matrix on validation data.

Class_2	Class_1	Class_3	Class_4	Error	Rate
23.0	0.0	0.0	0.0	0.0	0/23
0.0	22.0	0.0	0.0	0.0	0/22
1.0	0.0	23.0	0.0	0.0416	1/24
0.0	0.0	0.0	33.0	0.0	0/33
24.0	22.0	23.0	33.0	0.0098	1/102

The testing set is used to predict the variable CLASS, which contains labels for each class (Class 1, Class 2, Class 3, and Class 4), and a predictive accuracy of 99.2% for the different classes is obtained.

4 Conclusions

In this paper we investigated the use of deep learning for the classification and prediction of subcortical structures of patients with Parkinson's disease, based on microelectrode records (MER), obtained during deep brain stimulation (DBS).

We proposed six types of input features and a corresponding architecture to precisely predict subcortical structures. First, we showed that the network can learn surprisingly well. Second, we showed that the network can classify the different subcortical structures with high efficiency, and finally, a high precision is achieved in the task of predicting the different classes.

Our experiments indicate that a deep learning approach in combination with input features, has the potential to capture subcortical structures patterns, which may boost the classification performance. These investigations could be further improved in future studies by carrying out more exhaustive searches for the parameters in the architectures. Moreover, the overall performance of these systems could be further improved.

We conclude that deep learning could be used to monitor in real time the location of subcortical structures, reducing the uncertainty that exists during surgery, and representing a valuable support tool for neurosurgeons and electrophysiologists during electrical stimulation and deep brain electrode implantation for the treatment of parkinsonian patients.

Acknowledgments. This work was funded by Center for Advanced Computing and Data Systems, CACDS, at the University of Houston, Houston, TX, USA.

References

1. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**, 436–444 (2015)
2. Langkvist, M., Karlsson, L., Loutfi, A.: A review of unsupervised feature learning and deep learning for time series modeling. Pattern Recogn. Lett. **42**, 11–24 (2014)
3. Yu, D., Deng, L.: Deep learning and its applications to signal and information processing. IEEE Signal Process. Mag. 45–54 (2011)
4. Langkvist, M., Karlsson, L., Loutfi, A.: Sleep stage classification using unsupervised feature learning. Adv. Artif. Neural Syst. 1–9 (2012)
5. Yuste, R.: From the neuron doctrine to neural networks. Nat. Rev. Neurosci. **16**(8), 487–497 (2015)
6. Rubinov, M.: Neural networks in the future of neuroscience research. Nat. Rev. Neurosci. **16** (12), 767 (2015)
7. Yamins, D., DiCarlo, J.J.: Using goal-driven deep learning models to understand sensory cortex. Nat. Neurosci. **19**(3), 356–365 (2016)
8. Jankovic, J.: Parkinson's disease: clinical features and diagnosis. J. Neurol. Neurosurg. Psychiatry **79**(4), 368–376 (2008)
9. Chaudhuri, K., Schapira, A.: Non-motor symptoms of Parkinson's disease: dopaminergic pathophysiology and treatment. Lancet Neurol. **8**(5), 464–474 (2009)
10. Okun, M.: Deep-brain stimulation for Parkinson's disease. New Engl. J. Med. **367**(16), 1529–1538 (2012)
11. Galvan, A., Devergnas, A., Wichmann, T.: Alterations in neuronal activity in basal ganglia-thalamocortical circuits in the parkinsonian state. Front. Neuroanat. **9**(5), 1–21 (2015)
12. Shin-Yuan, C., Sheng-Huang, L., Shinn-Zong, L.: Subthalamic nucleus deep brain stimulation for Parkinson's disease – an update review. Tzu Chi Med. J. **17**(4), 205–212 (2005)

13. Pollo, C., Vingerhoets, F., Pralong, E., Ghika, J., Maeder, P., Meuli, R., Thiran, J., Villemure, J.: Localization of electrodes in the subthalamic nucleus on magnetic resonance imaging. J. Neurosurg. **106**, 36–44 (2007)
14. Guillén, P., Martínez-de-Pisón, F., Sánchez, R., Argaez, M., Velásquez, L.: Characterization of subcortical structures during deep brain stimulation utilizing support vector machines. In: 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2011), Conference Proceedings, Boston, MA, USA (2011)
15. Guillén, P., Barrera, J., Martínez-de-Pisón, F., Argaez, M., Velásquez, L.: Data mining in the process of localization and classification of subcorticals structures. In: EATIS, Conference Proceedings, Valencia, España (2012)
16. Aiello, S., Click, C., Roark, H., Rehak, L.: Machine Learning with Python and H20. H20, Gravesend (2016). Edited by Lanford, J.