A Hybrid Model based on Convolutional Neural Networks and Long Short-term Memory for Rest Tremor Classification

Jihen Fourati\textsuperscript{1,2}, Mohamed Othmani\textsuperscript{3} and Hela Ltifi\textsuperscript{1,4}

\textsuperscript{1}National Engineering School of Sfax, University of Sfax, BP 1173, Sfax, Tunisia
\textsuperscript{2}Research Lab: Technology, Energy, and Innovative Materials Lab, Faculty of Sciences of Gafsa, University of Gafsa, Tunisia
\textsuperscript{3}Faculty of Sciences of Gafsa, University of Gafsa, BP 2100, Gafsa, Tunisia
\textsuperscript{4}Research Groups in Intelligent Machines Lab, BP 3038, Sfax, Tunisia

Keywords: Resting Tremor, Deep Learning, Long-short Term Memory, Convolutional Neural Network, Parkinson’s Disease.

Abstract: Parkinson’s disease is a neurodegenerative disease, in which tremor is the main symptom. Deep brain stimulation can help manage a broad range of neurological ailments such as Parkinson’s disease. It involves electrical impulses delivered to specific targets in the brain, with the purpose of altering or modulating neural functioning. Security is playing a vital role in protecting healthcare gadgets from unauthorized access or modification. Our purpose is to adopt deep learning methodologies to classify resting tremors. To achieve this purpose, a novel approach for resting tremor classification in patients with Parkinson’s disease using a hybrid model based on convolutional neural networks and long short-term memory is proposed. This research exploits the high-level feature extraction of the convolutional neural network model and the potential capacity to capture long-term dependencies of the long short-term memory model. The performed experiments demonstrate that our proposed approach outperforms the best result for other state-of-the-art methods.

1 INTRODUCTION

Parkinson’s disease (PD) is a neurodegenerative disorder, in which patients suffer different symptoms: resting tremor, akinesia and rigidity (LeesAJ, 2009)(Parkinson, 1817). Resting tremor (RT) is a rhythmic and oscillatory involuntary movement that appears in a body part (Abdo et al., 2010). It is the roughest manifested symptom of Parkinson disease tremors, which occurs at a frequency band between 4 and 6 Hz (Lyons and Pahwa, 2005) and disappears when a voluntary movement is performed. Deep brain stimulation (DBS) is an exceedingly used therapy option to lessen the motor signs of advanced PD. However, it has crucial security-related issues. The ability to manipulate the pacemaker-like gadget enables performing numerous randomized trials to assess the efficiency of the device (Rathore et al., 2019). Indeed, an attacker can stop required stimulation or induce some needless shocks in the cerebrum by fake signals production (Choi et al., 2018). Deep brain stimulation affects both tremor amplitude and tremor frequency (Beuter et al., 2001). To quantify the tremor level for PD subjects, (Pedrosa et al., 2018) developed two predictive models to classify Parkinson’s disease’s rest tremor between high or low frequencies. The proposed models have reached a classification accuracy of 92.8%. Furthermore, (Perumal and Sankar, 2016) have studied the impact of using both gait and tremor features for the early detection and monitoring of PD by the use of statistical analysis and machine learning techniques.

To enhance resting tremor detection, a multicharacteristic classification approach depend on the characteristics of the local field potentials (LFP) has been used to recognize tremor-related features in PD patients (Bakstein et al., 2012), and shows that LFPs supplied enough information for detecting resting tremor. The authors of (Shah et al., 2018) proposed a method base on frequency and time domain combined with a logistic regression classifier to detect Parkinsonian rest tremors. Despite that, the delay of tremor detection was not stated and is a substantial parameter for closed-loop DBS implementation. Additionally, (López-Blanco et al., 2019) pro-
posed an android application for PD tremor analysis on a smartwatch. Their application experiments show that this hardware has the prospect to quantify PD subjects’ tremors impartially in the room of consultation. The precise detection of tremor onset in PD is crucial to the success of DBS therapy (Yao et al., 2020). For that, (Yao et al., 2020) proposed a method for detecting tremors during rest by the use of pertinent characteristics combined with machine learning and Kalman. Moreover, (Hssayeni et al., 2019) developed two methods based on deep learning and gradient boosting decision tree model used with wearable sensors to evaluate overall tremor Parkinson. As well, (Oktay and Kocer, 2020) presented a method for the classification of two types of tremors (Essential tremor and Parkinson tremor) based on convolutional long short-term memory. The experiments showed that convolutional long short-term memory provides successful results for differentiation of tremors. The model reached a testing accuracy of 90%. In addition, (Patel et al., 2009) developed a system to measure the gravity of tremor, bradykinesia, and dyskinesia using a wearable sensor platform. Further, (Salarian et al., 2007) presented a method to detect tremors and compare the tremor amplitude measurement to the corresponding unified parkinson’s disease rating scale (UPDRS) score. Moreover, Edwards et al. (Edwards and Beuter, 1999) used tremor features like the amplitude, frequency, and spectral power to differentiate PD tremors. Besides, they combined a set features into one variable to recognize a PD from abnormal tremors effectively (Goldberger et al., 2000).

Although several studies have been conducted on the resting tremor classification, not much research has been published about classifying different attack patterns for deep brain implants. Until recently, only two past attempts have been made to classify different attack patterns for brain implants: (Rathore et al., 2019) proposed a deep learning methodology for predicting and forecasting different signal patterns of deep brain stimulation. Typically, the rest tremor velocity is analyzed for evaluating the Parkinson tremor intensity. Various attacks have been introduced in the DBS context to simulate and distinguish between false and authentic stimulations. Moreover, referring to the security of signals from DBS, (Abdaoui et al., 2020) designed a monitoring system for distinguishing false alarms from legitimate ones and classified the attacks using Raspberry Pi3 and deep learning. They achieved an accuracy of 97% for predicting fake signals. In this paper, we propose a novel classification approach, for those who are receiving DBS to relieve tremors, using a hybrid model based on convolutional neural networks and long short-term memory.

This paper studies the pattern of rest tremor velocity (a type of feature observed to evaluate the intensity of neurological disorders) based on the pattern of introduced attack strategies. For this, we studied and examined RTV values to design and train the neural network.

The main contribution of this work is predicting whether the signal is an attack or a genuine signal for deep brain implants.

## 2 MATERIALS AND METHODS

### 2.1 Dataset

For tremor classification, rest tremor in subjects with Parkinson’s disease receiving chronic high frequency electrical deep brain stimulation (DBS) was recorded continuously throughout switching the deep brain stimulator on (at an effective frequency) and off. Data from Physionet online database (Goldberger et al., 2000) were utilized, consisting of readings from the experiments conducted on a group of 16 subjects with PD. Neurophysiological data were acquired by employing a low-intensity laser that was directed to a reflective piece of paper in the subject’s finger for a time period of 60 seconds for data acquisition (Beuter et al., 2001).

### 2.2 Data Acquisition

Data were collected using the MacLab data acquisition system and sampled at 100 Hz. Raw data were exported to S-Plus for analysis and converted from volts to mm/s (Beuter et al., 2001).

The data acquisition are carried out as the following steps:
- The laser was placed at about 30 cm from the subject’s index finger tip.
- The laser beam was directed to a piece of reflective tape placed on the finger tip.
- The velocity-transducer laser captures raw values outputting voltage proportional to the velocity of the finger for 60 seconds.

### 2.3 Data Preprocessing

Before dealing with the data, some signal manipulations were needed. The files of the dataset (Goldberger et al., 2000) vary in tremor velocity units between patients. A few patients presented their data in meters per second, while others presented in millimeters per second. Thus, it is necessary to perform
normalization of the data, such that the posterior processing and obtained results are identical among all patients’ data. A simple formula was used to normalize all data in millimeters per second.

\[ \text{mm/s} = \text{m/s} \times 1,000 \]  

As described by (Rathore et al., 2016), different types of attack strategies can be employed by the attacker. For that, we generated a new dataset by emulating different attack strategies along with modulating learned stimulation patterns. The generated dataset is composed of 4096 genuine and 4096 attacked sequences collected from real measurements. Each sequence contains 300 samples and one label indicating whether the sequence is genuine or attacked.

3 PROPOSED APPROACH

In this section, we present the details of our proposed approach, which contains convolutional and recurrent neural networks. First, we give a general overview of the approach. After that, we discuss each phase in detail.

3.1 Overview

Our main objective is to address the problem of rest tremor classification in patients with Parkinson’s disease using a combination of convolutional neural network (CNN) and long short-term memory (LSTM) based on hybrid deep learning model. The first step of our proposed approach is to improve the data as mentioned in section 2. Then, we extract the time-domain features of the RT data through a one-dimensional convolutional neural network. Next, these features are feeding into LSTM to extract the best representative features. For the last step, the classifier predicts whether the signal is an attack or a genuine signal. The block diagram of our hybrid approach is depicted in Fig.1.

![Block diagram of CNN-LSTM hybrid model-based RT classification](image)

Figure 1: Block diagram of CNN-LSTM hybrid model-based RT classification.

3.2 Convolutional Neural Network Model

CNN is a well-known deep learning architecture inspired by the human neural system (Ko, 2018). It identifies automatically the relevant features without any human supervision (Gu et al., 2018). A typical CNN architecture consists of alternating convolutional layers and pooling layers, followed by one or more fully connected layers. The convolutional layers consist of multiple kernels stacked together that are convolved with their input. It extracts the high-level features from the input signal by a sliding-window technique that outputs feature maps (Pak et al., 2018). The pooling layer provides a typical downsampling operation by applying the pooling operator to aggregate information inside each small region of the input feature channels and then select the most significant feature (Yamashita et al., 2018). These features are fed to the fully connected layer which generates the CNN model’s output data.

3.3 Long Short-term Memory Model

LSTM, a sophisticated version of recurrent neural network which is able to learn long-term dependencies, is designed to solve the long-term dependency problem by means of short-term memory (Pak et al., 2018). LSTM is able to process even the longest sequence data without vanishing the gradient (Pak et al., 2018). Each LSTM unit is composed of a memory cell and three main gates: input, output and forget. The memory cell is designed to selectively add or remove information into/from this cell under the control of these three gates (Bai and Tahmasebi, 2021). The input gate is mathematically represented as following:

\[ i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \]  

The operator \( \cdot \) represents the element-wise multiplication of the vectors. The information to be neglected from the previous memory is controlled by forget gate which is mathematically defined as following:

\[ f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \]  

The cell state is updated by the update gate, expressed mathematically by the following equations:

\[ c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \]  

\[ \tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \]  

The hidden layer of the previous time step is updated by the output gate which is also responsible for the updating the output as it is given by:

\[ o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \]  

\[ h_t = o_t \cdot \tanh(c_t) \]
3.4 Convolutional Neural Network and Long Short-term Memory for Rest Tremor Classification

Artificial intelligence has caught the attention of the scientific community in diverse fields such as intelligent decision support systems (IDSS) (Ellouzi et al., 2017) (Ellouzi et al., 2015), wavelets neural network (Bellil et al., 2008), pattern recognition, function approximation optimization, etc. Deep learning is a subset of machine learning where artificial neural networks, algorithms inspired by the human brain, learn from large amounts of data. The proposed model, i.e., hybrid deep learning strategy, is inspired by biological artificial neural networks.

The proposed model exploits the efficiency of convolutional neural networks for extracting high-level features and learning the internal representation of sequence data as well as the effectiveness of long short-term memory layers for identifying short-term and long-term dependencies. The main purpose of our proposed model is to efficiently combine the advantages of these deep learning techniques. In this study, a hybrid approach of CNN and LSTM has been successfully used as a rest tremor classifier. Our training phase inputs consist of the whole signals (normal and under attack). The inputs go through one convolution layer, one max-pooling layer followed by another convolution layer. The convolution layers (1, 2) are convolved with their respective kernel size and their filter number (2, 32). Between the block of convolution layers, the max-pooling layer is applied to the feature maps. Thus, it was employed to reduce the number of parameters to learn and to minimize the amount of computation performed in the network. After the block of convolution layers and the max-pooling layer, two LSTM layers were applied with recurrent activation function with a dropout and recurrent dropout of 0.2. The total unit of each LSTM layer is 32. Nonlinearity is presented in the model by offering some layers with the hyperbolic tangent (tanh) activation function. The tanh function is the updated version of the sigmoid function on the range, which is a symmetric function centered on zero. Its output is bounded, and it brings nonlinearity to the neural network (Wang et al., 2020). In both LSTM layers, we applied the "tanh" activation function. The mathematical form of tanh function is as follows:

\[ f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \]  

Where \( x \) is the input to a neuron. The details of each layer parameter of the proposed rest tremor classification model are presented in Table 1. For the training parameters of the proposed network, we adopted 80% of samples for training and the remaining 20% for testing. Figure 2 shows the proposed network model. As we are facing a binary classification problem (normal or under attack) and the output of our model is a probability (In the end, we applied the classification layer using a sigmoid function), the best choice is to configure the model with adaptive moment estimation (ADAM) optimizer and the binary cross-entropy loss function. ADAM algorithm ensures that the learning steps, during the training process, are scale-invariant relative to the parameter gradients (Livieris et al., 2020). Finally, the training and testing of the proposed approach is done in 300 epochs and a batch size of 64 samples. A hyperparameter tuning was performed till the optimum performance was reached.

4 EXPERIMENTAL RESULTS

Our model was evaluated on a laptop computer with a 4 GHz CPU, 64 GB of memory. All methods were executed using Jupyter Notebooks (Python 3.7.10).

The proposed approach was tested on the Physionet dataset. Physionet "tremor-DB" is made up of the original rest tremor velocity signals, the ground truths for benchmark the RT classification. All data recordings are considered genuine measurements. In order to study the effect of deep brain stimulation on amplitude and frequency characteristics of rest tremor in Parkinson’s disease (Beuter et al., 2001), different attack strategy was carried out. We have emulated different attack strategies introduced in the DBS framework by changing the learnt stimulation pattern. To improve the computational efficiency of the proposed model, we generate a new dataset by creating faulty signals that correspond to the attack patterns. The dataset is composed of 4096 genuine and 4096 attacked sequences collected from real measurements. Each sequence contains 300 samples and one label indicating if the sequence is genuine (label 0) or attacked (label 1).

Cross-validation (CV) is a standardized test commonly used to test the ability of the classification system using various combinations of the testing and training datasets. The proposed approach was trained with the k-fold cross-validation procedure (k=10) by using train data. This procedure divides arbitrarily the set of observations into 10 approximately equal folds. The first fold is used as a validation set, and the method fits on the remaining k-1 folds.

We opted for using 10-fold cross-validation technique since this is one of the most using method for estimating accuracy due to its relatively low bias and
Table 1: The architecture of our proposed model.

<table>
<thead>
<tr>
<th>Layers No.</th>
<th>Type</th>
<th>Kernel size</th>
<th>No. kernels/Units</th>
<th>Output shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer 1</td>
<td>Conv1D</td>
<td>2</td>
<td>32</td>
<td>2 x 32</td>
</tr>
<tr>
<td>Layer 2</td>
<td>MaxPooling1D</td>
<td>2</td>
<td>32</td>
<td>2 x 32</td>
</tr>
<tr>
<td>Layer 3</td>
<td>Conv1D</td>
<td>2</td>
<td>32</td>
<td>2 x 32</td>
</tr>
<tr>
<td>Layer 4</td>
<td>LSTM</td>
<td>32</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Layer 5</td>
<td>LSTM</td>
<td>32</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Layer 6</td>
<td>Dense</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: The proposed CNN-LSTM model.

Figure 3: Loss and accuracy curves (training and validation).

The cross-validation process was repeated until every tremor dataset were included in the testing dataset with 10-fold. The models obtained from the training phase are tested section-by-section and then averaged of recall, precision, accuracy, and F1-score (given in Table 5). The 10-fold cross-validation process results are shown in Table 2. The mean and standard values of the training accuracy and validation accuracy after 10-fold cross-validation are evaluated as 97.01% (±0.0119898) and 94.90% (±0.0193959), respectively. In Table 3, we summarize the average and standard deviation for the calculated k-fold cross-validation results process. The best model results are obtained in epoch number 287. It takes 109 ms/step and leads to training loss = 0.1760, training binary accuracy = 0.9701, validation loss = 0.1760, and validation binary accuracy = 0.9593, as shown in Fig. 3.

To compare our experiments with previous works, we defined four metrics: accuracy, recall, precision, and F1-score. The most frequent classification evaluation metric is accuracy. However, it can be misleading when the target classes are imbalanced. In such circumstances, recall and precision are more appropriate metrics.
Table 2: 10-fold cross-validation results.

<table>
<thead>
<tr>
<th>Experimental trials</th>
<th>Training accuracy</th>
<th>Training loss</th>
<th>Validation accuracy</th>
<th>Validation loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9630</td>
<td>0.1494</td>
<td>0.9277</td>
<td>0.2369</td>
</tr>
<tr>
<td>2</td>
<td>0.9690</td>
<td>0.1247</td>
<td>0.9444</td>
<td>0.1619</td>
</tr>
<tr>
<td>3</td>
<td>0.9815</td>
<td>0.1144</td>
<td>0.9628</td>
<td>0.1566</td>
</tr>
<tr>
<td>4</td>
<td>0.9690</td>
<td>0.1317</td>
<td>0.9630</td>
<td>0.1513</td>
</tr>
<tr>
<td>5</td>
<td>0.9444</td>
<td>0.2487</td>
<td>0.9442</td>
<td>0.2041</td>
</tr>
<tr>
<td>6</td>
<td>0.9815</td>
<td>0.1020</td>
<td>0.9649</td>
<td>0.1483</td>
</tr>
<tr>
<td>7</td>
<td>0.9815</td>
<td>0.0848</td>
<td>0.9669</td>
<td>0.1531</td>
</tr>
<tr>
<td>8</td>
<td>0.9815</td>
<td>0.1138</td>
<td>0.9628</td>
<td>0.1805</td>
</tr>
<tr>
<td>9</td>
<td>0.9669</td>
<td>0.1379</td>
<td>0.9074</td>
<td>0.3349</td>
</tr>
<tr>
<td>10</td>
<td>0.9630</td>
<td>0.1574</td>
<td>0.9463</td>
<td>0.1856</td>
</tr>
</tbody>
</table>

Table 3: Statistical analysis of 10-fold cross-validation process.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training accuracy</td>
<td>0.97013</td>
<td>0.01198</td>
</tr>
<tr>
<td>Training loss</td>
<td>0.13648</td>
<td>0.04499</td>
</tr>
<tr>
<td>Validation accuracy</td>
<td>0.94904</td>
<td>0.01936</td>
</tr>
<tr>
<td>Validation loss</td>
<td>0.19132</td>
<td>0.05772</td>
</tr>
</tbody>
</table>

Table 4: Confusion matrix of binary rest tremor classification problem.

<table>
<thead>
<tr>
<th>Predicted class</th>
<th>Normal</th>
<th>Under Attack</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual class</td>
<td>TP</td>
<td>FN</td>
</tr>
<tr>
<td>Normal</td>
<td>FP</td>
<td>TN</td>
</tr>
</tbody>
</table>

cases, other evaluation metrics should be considered in addition to the accuracy. Recall, precision and F1-score are excellent quantification measures for binary classification.

The model classification performance is best presented in a two-class confusion matrix consisting of a 2x2 matrix, with True Positives (TP), False Negatives (FN), False Positives (FP), and True Negatives (TN) described in Table 4.

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \quad (10)
\]

\[
\text{Precision} = \frac{TP}{TP + FP} \quad (11)
\]

\[
F1 - \text{score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (12)
\]

Table 5 illustrates the comparison between our proposed approach and other previous studies. The proposed approach achieved an accuracy, a recall, a precision, f1-score, and area under the receiver operating characteristic curve (AUC) of 94.9%, 96.42%, 90.0%, 93.0%, and 93.75% respectively. It can be noted that the best accuracy, recall, and F1-score were achieved by the proposed method, i.e., 94.9%, 96.42%, and 93.0%, which demonstrates its excellent capacity for addressing the challenge. The area under the curve reaches a high value.

Figure 4 shows the receiver operating characteristic curve (ROC curve) of the classification with resting tremor. The ROC curve plots between the values of the true positive rate (sensitivity) to the false positive rate (1-specificity) at various classification thresholds.

The area under the curve was achieved as 93.75%, which quantifies the overall ability of the algorithm to distinguish between a subject with normal rest tremor velocity and a subject under attack.
Table 5: Comparison results on Physionet dataset.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDA (Perumal and Sankar, 2016)</td>
<td>0.869</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>KNN+SVM (Pedrosa et al., 2018)</td>
<td>0.928</td>
<td>1</td>
<td>0.833</td>
<td>0.909</td>
</tr>
<tr>
<td>Ours</td>
<td><strong>0.949</strong></td>
<td>0.90</td>
<td><strong>0.964</strong></td>
<td><strong>0.93</strong></td>
</tr>
</tbody>
</table>

5 CONCLUSIONS

In this paper, a novel approach based on convolutional neural networks and a long short-term memory hybrid model for resting tremor classification is presented. The aim of this study was to exploit the high-level feature extraction of the convolutional neural network model and the potential capacity to capture long-term dependencies of the long short-term memory. A comparison study is reported to demonstrate the performance and the effectiveness of the novel proposed approach among the methods in previous literature. As exhibited in experiments, the proposed approach outperforms state-of-the-art methods in terms of recall, accuracy, and F1-score.

REFERENCES


