International Journal of Computer Assisted Radiology and Surgery
https://doi.org/10.1007/s11548-021-02379-0

ORIGINAL ARTICLE

Adapting the listening time for micro-electrode recordings in deep brain stimulation interventions

Thibault Martin[1] · Greydon Gilmore[2] · Claire Haegelen[3] · Pierre Jannin[1] · John S. H. Baxter[1]

Received: 11 January 2021 / Accepted: 12 April 2021

Abstract

Purpose Deep brain stimulation (DBS) is a common treatment for a variety of neurological disorders which involves the precise placement of electrodes at particular subcortical locations such as the subthalamic nucleus. This placement is often guided by auditory analysis of micro-electrode recordings (MERs) which informs the clinical team as to the anatomic region in which the electrode is currently positioned. Recent automation attempts have lacked flexibility in terms of the amount of signal recorded, not allowing them to collect more signal when higher certainty is needed or less when the anatomy is unambiguous.

Methods We have addressed this problem by evaluating a simple algorithm that allows for MER signal collection to terminate once the underlying model has sufficient confidence. We have parameterized this approach and explored its performance using three underlying models composed of one neural network and two Bayesian extensions of said network.

Results We have shown that one particular configuration, a Bayesian model of the underlying network's certainty, outperforms the others and is relatively insensitive to parameterization. Further investigation shows that this model also allows for signals to be classified earlier without increasing the error rate.

Conclusion We have presented a simple algorithm that records the confidence of an underlying neural network, thus allowing for MER data collection to be terminated early when sufficient confidence is reached. This has the potential to improve the efficiency of DBS electrode implantation by reducing the time required to identify anatomical structures using MERs.

Keywords Deep brain stimulation · Micro-electrode recordings · Deep learning · Bayesian models

Introduction

Deep brain stimulation (DBS) is a common treatment for a variety of neurological disorders such as Parkinson's disease in which the abnormal activation of a particular region leads to undesirable characteristic symptomatology. For example, in PD, the degeneration of dopamine producing regions of the *substantia nigra* leads to the abnormal activation of other regions of the basal ganglia which in turn leads to the disorder's characteristic motor symptoms such as the diminished ability to initiative and control voluntary motor actions. DBS is currently the dominant surgical procedure for this pathology and can address the patient's symptoms directly by correcting these neural abnormalities, stimulating particular neural populations in order to suppress their pathological activity [7,13]. This is highly beneficial if other non-surgical techniques, such as pharmaceutical treatment, fails to adequately control the patient's symptoms.

The DBS intervention itself consists of the highly accurate placement of the stimulation electrodes using a stereotaxic frame into a pre-defined region of the patient's subcortical anatomy such as the subthalamic nucleus (STN) or globus pallidus internus (GPi) in the case of Parkinson's disease. Often, the clinical workflow for DBS will consist of multiple

Thibault Martin is supported through a Doctoral Research Grant from Association France Parkinson. John S.H. Baxter is supported by the Institut des Neurosciences Cliniques de Rennes (INCR) and the Natural Sciences and Engineering Research Council of Canada (NSERC) through the Post-Doctoral Fellowship (PDF) program.

✉ John S. H. Baxter
jbaxter@univ-rennes1.fr

[1] Laboratoire Traitement du Signal et de l'Image (LTSI - INSERM UMR 1099), Université de Rennes 1, Rennes, France

[2] Biomedical Engineering Graduate Program, Western University, London, Canada

[3] Department of Neurosurgery, Centre Hospitalier Universitaire de Rennes, Rennes, France

Published online: 11 June 2021

steps, including a preoperative planning phase, an intra-operative electrode positioning phase, and a postoperative stimulation parameter tuning phase.

The preoperative planning phase uses T1- and T2-weighted clinical MR images in which the anatomy of interest is segmented and the potential electrode trajectories can be investigated. Recent approaches in image-guided interventions encourage the use of these preoperative images to assist with the navigation of the DBS electrodes, albeit at a coarser resolution level due to the presence of several contributing sources of error including (1) brain-shift resulting from the craniotomy site, (2) small shifts in the stereotaxic frame's position, and (3) angular error in the insertion angle of the electrode [12]. Exclusive use of MRI would lead to a sub-optimal STN localization in 20% of cases, according to Lozano et al. [13].

Because of these sources of error, an additional data modality must be incorporated to ensure correct positioning during the intra-operative phase. Intra-operative imaging modalities such as intra-operative CT or X-ray have the ability to readily identify the electrodes but, due to a lack of soft-tissue contrast, cannot readily distinguish between the target region and surrounding structures [19]. More advanced intra-operative imaging modalities that have this capability, such as intra-operative MRI, are prohibitively expensive and require specialized operating suites.

An alternative to these predominantly structural imaging modalities, *functional* electrophysiological modalities are often used instead, demonstrating robust results [20]. The most common of these is the *trial-and-error* method in which the patient remains conscious during the surgery, allowing the clinical team to stimulate regions along the electrode trajectory and determine the optimal electrode positioning from the patient's observable behaviour. This also allows for the clinical team to determine the presence of side-effects, although recent studies have found there to be a disconnect between this intra-operative and later postoperative assessments [3]. In addition, awake surgery is uncomfortable for the patient and potentially infeasible for patients with very advanced PD, whose motor symptoms may prevent them from being adequately still during the procedure. Thus, micro-electrode recording (MER) is often used as the preferred intra-operative data modality for DBS electrode implantation. Instead of stimulating the patient and observing their behaviour, MER electrodes measure the activity of the proximal neural region surrounding them, allowing the clinical team to infer the electrode's position through the characteristic signature of the region of interest [2,17]. In the current standard-of-care, this signature is determined by an expert neurologist by listening to the signal. By integrating different parameters specific to the surgical environment, such as the depth of the electrode, the distance to the target coordinate determined by imaging, and by deciphering the functional neurophysiological characteristics contained in the signal [9], the neurophysiologist establishes the approximate position of the electrode by classifying whether or not the MER arose from inside the STN. This is repeated at regular depth intervals for each trajectory used during the intervention. This requires extensive expertise and is a highly subjective process that could benefit from automation.

Previous work in automatically analysing MER for identifying subcortical DBS target regions have been traditionally feature-based in which the MER signal is represented as a vector of pre-defined, engineered features that represent certain statistical features of the underlying neural population (i.e. firing rate, presence of different frequency bands, etc.) [5,18,21–23]. Although these feature-based approaches are highly automatic, transparent, and can be based on an arbitrary amount of signal, recent data-driven methods operating on the raw signal have proven to have higher accuracy [16]. Among these methods, SepaConvNet, a convolutional neural network, has been able to predict the presence of the STN from one-second MER signals [16]. This fixed time can be made dynamic through the use of recurrent neural networks or Bayesian inference. [14] However, it has yet to be shown how such a dynamic time could be algorithmically integrated and thus validated.

Contributions

In this paper, we propose an approach for optimizing the automatic classification of the STN from intra-operative MER which can accelerate the analysis by reducing the amount of MER signal acquired. The reduction in surgical time is intuitively beneficial for the hospital, as it can reduce the cost of the intervention, the risk of infection and the subsequent patient recovery time. We propose and validate a simple 'listening algorithm' for adapting the amount of MER signal provided to an underlying classification model. This model relies on receiving signal in small chunks which are then provided to the underlying classifier, receiving an updated classification. If the classification results maintain a particular level of confidence for a particular length of time, the only parameters of the proposed method, the listening algorithm produces a final binary classification, thus terminating the listening time early. By varying these two parameters, we have shown how accuracy and efficiency can be effectively traded-off in this framework and show how it can largely improve the efficiency of DBS electrode positioning by reducing the amount of time needed to collect the intra-operative MER signals. Analysis of the dynamics of the listening algorithm over time also give insight into how the underlying machine learning methods could be improved in a way that could directly impact their use.

Methods

Micro-electrode recording database

The data used in this study were collected at the London Health Sciences Centre at Western University Hospital (London, Canada), including 57 Parkinson's Disease patients undergoing a single or bilateral deep brain stimulation intervention. Micro-electrode recording signals were recorded through five channels (anterior, posterior, medial, lateral and central), using a *Leadpoint 5* recording station (Medtronic). A preoperative target was defined by magnetic resonance imaging before surgery, and micro-electrode recordings were captured from 10.0mm to 4.0/5.0mm after target estimation. For each record, sampling was made at 24kHz (8-bits), amplified (gain: 10,000) and digitally filtered (bandpass: 500-5000Hz, notch: 60Hz). The use of these data was led by the collaborative agreement covered by ethical clearance DSA 109045, and was supported by by the Research Ethics Board at the University of Western Ontario (REB # 109045).

The MER database contains 11,162 signals, each containing 9 seconds of recording and have been annotated by an expert neurosurgeon as being either within the STN (2,574 signals) or outside of the STN (8,588 signals). Due to the class imbalance, an oversampling method was designed to increase the amount of the under-represented class. Therefore, a single random second of recording was use per 'OUT' annotated signal, compared to three non-overlapping random seconds of recordings for 'IN' annotated signals, leading to 8,588 'OUT' against 7,722 'IN' samples.

During training, MER samples were grouped by patient (i.e. patients, rather than signals, are divided into training and testing sets) in order to avoid any data leakage and to ensure that evaluation is performed only on unseen patients data. Tenfold cross-validation was performed over the 57 patients, consisting in ten distinct training initializations. For each repetition, onefold is used only for model evaluation with the remaining ninefolds used for model training. The evaluation metrics are then averaged over the repetitions. For the training folds, oversampling of the 'IN' class was used to address class imbalance. For evaluation, metrics were specifically chosen that account for this lack of class balance.

Adaptive listening time

The MER listening algorithm, shown in Algorithm 1, is used to determine the final prediction for a given signal. This algorithm listens to the output of the neural network (P), counting the number of consecutive times that it is under a particular threshold α (which indicates a confident conclusion that the signal comes from outside the STN) or above the threshold $1-\alpha$ (which indicates a confident conclusion that the signal comes from inside the STN). After a pre-determined number of this conclusions is made, the algorithm is permitted to terminate early with a prediction. Else, the algorithm returns '*None*' which means a longer listening time would still be required. Given the retrospective nature of this study, '*None*' is a valid result of this algorithm, as we do not have the capability to acquire more signal if Algorithm 1 does not return a positive or negative prediction in the 9 seconds. However, in clinical use, the algorithm could always be given more signal, which is often the case in regions along the boundary of the STN where it is difficult, even for trained neurophysiologists, to confidently determine the electrode location.

Algorithm 1: MER Listening Algorithm.

Data: $P[t] = P_{(t)}(X = \text{'IN'})$, $\alpha \in (0, 1)$, patience > 0
Result: Final prediction: prediction $\in$ {'None','OUT','IN'},
Termination time: t where patience $\leq t \leq length(P)$

```
begin
    prediction ← 'None';
    C'OUT' ← 0;
    C'IN' ← 0;
    for t = 1 → length(P) do
        /* If we are 'certain', add to the
           number of seconds above/below
           threshold                          */
        if P[t] < α then
            C'OUT' ← C'OUT' + 1;
            C'IN' ← 0;
        else if P[t] > 1 − α then
            C'IN' ← C'IN' + 1;
            C'OUT' ← 0;
        else
            C'OUT' ← 0; C'IN' ← 0;

        /* Determine if we can terminate early
           */
        if C'OUT' = patience then
            prediction ← 'OUT';
            break;
        else if C'IN' = patience then
            prediction ← 'IN';
            break;

    end
    return prediction, t
end
```

Underlying Bayesian and neural network architectures

The first network used is a reproduction of SepaConvNet (SCN), the convolutional neural network proposed by Peralta et al. [16]. The signals undergo some preprocessing in order to remove any artefacts related to the acquisition, clipping its amplitude to be within [-249:250]. A short-term Fourier transform algorithm was used with a Hann window of 512 samples and a hop length of ten samples resulting in a spectrogram with 21,600 time points by 257 frequency bands.

Each frequency band is then normalized to further remove artefacts using approximate min–max scaling with the 95th and 5th percentiles as the maximum and minimum values. The signals are then fed into a convolutional neural network as shown in Fig. 1 which classifies them into either 'IN' (the network output being close to 1) or 'OUT' (the network output being close to 0). The benefit of this network is that the amount of signal used is relatively short and can be processed almost in real-time, meaning that the predictions given by SepaConvNet for each time point can be used.

In order to extend SepaConvNet which process independently one second spectrograms, the Bayesian framework presented by Martin et al. [14] was also reproduced. This method is based on Bayes' Theorem, and update the probability of the signal arising from the STN given consecutive predictions from SepaConvNet. The general formula is presented by the authors in the following form:

$$P_{(t)}(X = p) \propto P_{(t-1)}(X = p) \times P(f_{(t)}|X = p) \tag{1}$$

X	the electrodes location as either 'in' or 'out' of the STN.
p	the classes to be distinguished, specifically {in, out}.
t	the length of MER used, i.e. an integer in the range [1..9].
$P_{(t)}(X = p)$	is the probability distribution of X, given past predictions up to and including time t.
$P(f_{(t)}\|X = p)$	is the probability of SepaConvNet generating some feature at time t (represented by $f_{(t)}$) conditioned on the anatomy the MER arose from.

Equation 1 thus allows to obtain for each successive listening second a prediction dependent on the behaviour of the underlying: $P(f_{(t)}|X = p)$, associated with the prior probability calculated at time $t-1$, $P_{(t-1)}(X = p)$. At the initialization, the prior term is calculated according to the ratio of the classes within the validation data. Two definitions of the term $f_{(t)}$ proposed by Martin et al. have been re-implemented in the framework of this study:

1. *A simple Bayesian extension of SCN*, which uses the non-thresholded SepaConvNet output as an indication of the certainty of the network. In this case, $f_{(t)} \in [0, 1]$, and $P(f_{(t)}|X = p)$ density is computed over a Gaussian Mixture Model which models SepaConvNet output according to each class. With two Gaussian components per class, this approach adds 12 parameters to SepaConvNet.
2. *An advanced Bayesian extension of SCN*, which also computes $P(f_{(t)}|X = p)$ using a Gaussian mixture model to represent the probability distribution of the nonnegative activation vector from the second to last layer SepaConvNet, with four components per class. This architecture increases the number of parameters by 8,456 beyond that of SepaConvNet.

Evaluation system

For each network, α and patience terms were set by using a GridSearch algorithm over validation data. This learning method is used to refine the parameters of a predictive model, by trying all combinations of its parameters passed in arguments, and by associating a prediction score to each combination. GridSearch was used for incremental values of 0.025 for $\alpha \in (0, 0.5)$, and patience domain was described as follows: $0 < \text{patience} \leq 9$.

Evaluation metrics

At any given time-point, the current prediction of Algorithm 1 is one of 'None' (indicating no certain result), 'Out' (indicating that the signal arose from outside the STN) and 'In' (indicating that the signal arose from inside the STN). Because of the ternary, rather than binary, nature of the classification problem, there are four metrics that should be evaluated:

$$\begin{aligned} SENS &= \frac{True\ Positive}{Positive} & SPEC &= \frac{True\ Negative}{Negative} \\ MISS &= \frac{False\ Negative}{Positive} & FALLOUT &= \frac{False\ Positive}{Negative} \end{aligned} \tag{2}$$

Sensitivity (SENS) and specificity (SPEC) are measures of positive quality—how well the network correctly classifies signals, whereas the miss rate (MISS) and fall-out rate (FALLOUT) are measures of negative quality—how frequently the network mis-classifies signals. It should be noted that classifying a signal as 'None' lowers all of these rates, neither correctly classifying nor incorrectly classifying a signal. Thus, to take into account both positive and negative quality, we use a balanced ternary quality (BTQ) with the formula:

$$BTQ = \frac{1}{2}(SENS + SPEC - MISS - FALLOUT) \tag{3}$$

This formula has several advantages including being inherently class balanced with a random classifier, or a classifier not producing any predictions, having an expected BTQ of 0. Additionally, if there are no 'None' predictions, there is a relationship between the BTQ and the balanced accuracy (BACC), specifically $BTQ = 2 \times BACC - 1$.

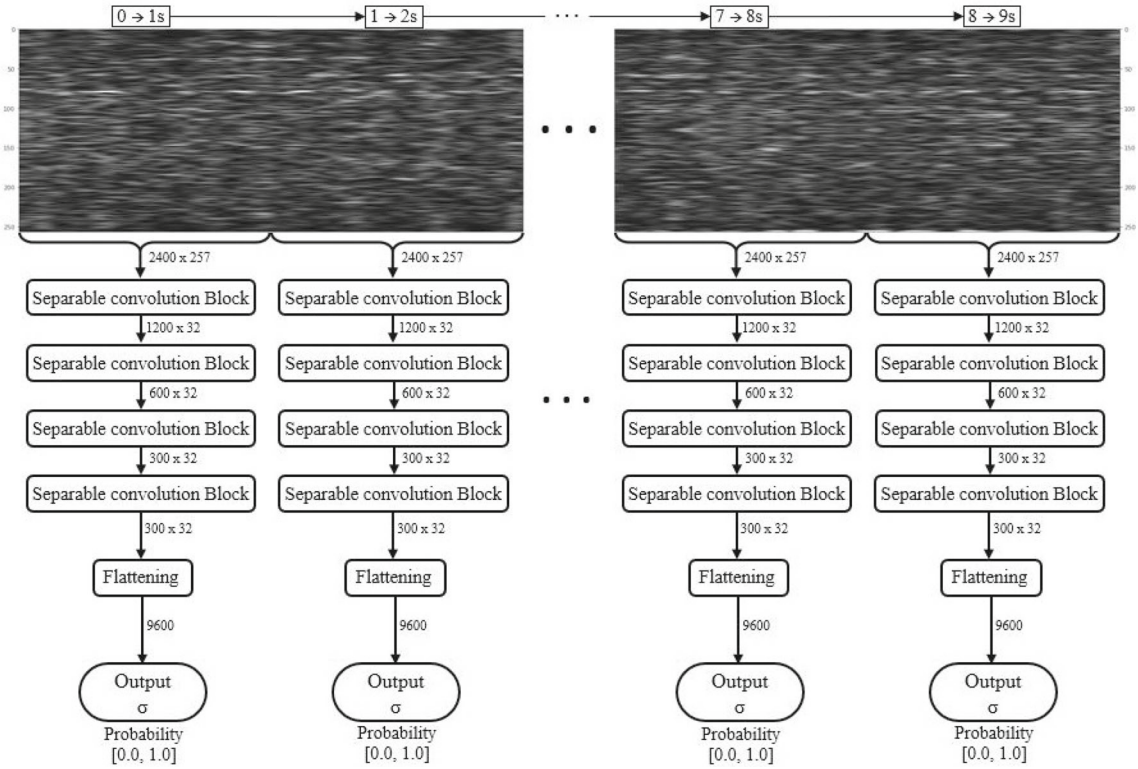

Fig. 1 SepaConvNet architecture

Results

The parameter space for each of the three comparative methods has been explored using a grid search and the resultant BTQ metric is shown in Fig. 2. Note that the colour scale is the same for each subfigure and has been scaled to reflect the maximal performance of the methods investigated. The performance of the advanced Bayesian SepaConvNet is consistently higher than the alternatives.

In order to investigate the behaviour of these methods over the course of the signal, the optimal $patience$ and α parameters were defined independently for each method, and the behaviour of each approach is visualized in Fig. 3.

To further illustrate the performance of Algorithm 1 using different underlying methods, Fig. 3 presents the evolution of validation metrics across the database of 5,584 signals. It should be noted that the bars in Fig. 3 are calculated over the signals predicted at the given time-point, and thus $SENS + MISS = SPEC + FALLOUT = 1$. Similarly, the number of signals classified at each time step is shown using a logarithmic scale as approximately 90% of signals are classified as soon as possible for all of the proposed methods and a linear scale would render subsequent time points difficult to distinguish.

As shown in Fig. 2, the advanced Bayesian method shows the best performance among the three methodologies studied. In order to determine if early termination has a negative effect on classification accuracy, the advanced Bayesian extension was evaluated both with and without early termination. The hypothesis was that the same method but with access to the full signal would perform better than the same method without said full access. However, this effect was very slight that early termination decreased the BACC score from 83.5% without early termination to 83.0% with early termination. This is despite having unclassified signals, which BACC, a binary classification metric, considers to all be incorrectly classified. Although this improvement is not statistically significant (under a paired Student's t test across all 57 patients in the cross-evaluation), it still indicates that early termination very likely does not have a negative effect on classification performance.

Discussion

The evaluation of the SepaConvNet neural network architecture has already highlighted the relevance of this data-driven methodology for the recognition of patterns within a MER signal [16]. In order to further optimize prediction performance, the SepaConvNet extension then focused on the iterative integration of an arbitrary number of seconds of MER to increase the certainty of the underlying neural

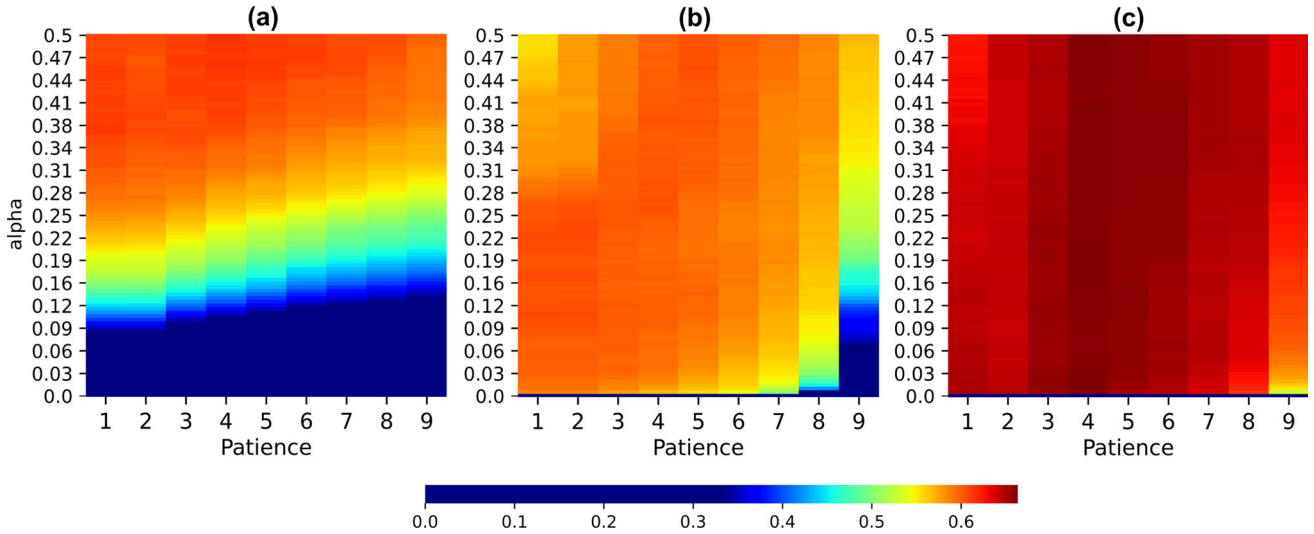

Fig. 2 Final BTQ of the three comparative methods at the end of the 9 second MER signal averaged across all k folds of cross-validation. Adaptive Listening time with SepaConvNet (**a**), the Simple Bayesian Extension (**b**), and the Advanced Bayesian Extension (**c**)

network [14]. The algorithm presented in this study aims at dynamically establishing the listening time required to optimize the trade-off between listening shortening and prediction quality. This work shows that the use of the presented early stopping algorithm does not alter the overall quality of prediction, indicating that it is possible to stop listening prematurely for some signals, without penalty, and thus accelerate the task of MER identification. Furthermore, the early stopping listening algorithm could also be extended to any classification approach based on the use of intra-operatives features. [6,10]

In terms of interpreting the algorithm parameterization, particularly *via* Fig. 2, there are two dynamics to take into account: the misclassification rate and the non-prediction rate. For example, as the patience increases, it should have strongly increase the non-prediction rate (i.e. more signals are missed due to insufficient MER length) with a weaker negative effect on the misclassification rate (i.e. fewer signals near the boundaries are misclassified as they need to stay on the same side of the boundary for longer). Similarly, as α increases, Algorithm 1 becomes more forgiving, thus increasing the misclassification rate and lowering the non-prediction rate.

For the patience parameter, one consistent result can be seen across all methods: at extremely high patience values, such as 7, 8 or 9, the BTQ swiftly decreases at low values of alpha. This is likely because the signal is simply not long enough for such confidence levels to have been maintained for such an amount of time.

For SepaConvNet with early stopping, the effect of the alpha parameter seems to dominate the algorithm's behaviour, being sensitive to it at all values of patience. This is likely because the output of SepaConvNet, as noted by Martin et al. [14] is not a true likelihood reflecting the model's underlying certainty and thus the range of values is reduced rather than full ranging between 0 and 1.

For the Bayesian methods, the dynamics of early termination are somewhat different due to the convergence behaviour of said methods, i.e. that they converge towards either 0 or 1 relatively quickly for the majority of signals [14]. At low values of alpha, lower values of patience tend to be optimal, as both methods take time to reach said confidence levels. As the value of alpha increases and this minimal convergence time therefore lessens, we see that middling patience values become preferable, indicating that there is, in fact, a population of signals that originally start to converge in one direction and then 'change their mind' and ultimately converge to the opposite. This interesting behaviour raises a number of hypothesis regarding how frequently characteristic signal patterns occur which could inform surgical practice in terms of minimum listening times for each of the 'IN' and 'OUT' classes.

Particular attention has been given to the advanced Bayesian extension, as it was the best performing in both this study and a previous study [14]. As shown in Fig. 3c, it is possible to reduce the average listening time for MER signals by more than half, which is a strong argument for improving the quality of clinical support when using a real-time prediction tool. In addition, the prediction quality of MER signals within the first stop timestep demonstrates that the early stopping algorithm is capable of highlighting signals with "simple" discriminative features, for which it provides a higher prediction score than that of the underlying Advanced Bayesian Extension given the entire signal, increasing the balanced accuracy from 83.5% to 85.45% for those signals. Therefore, this approach would sepa-

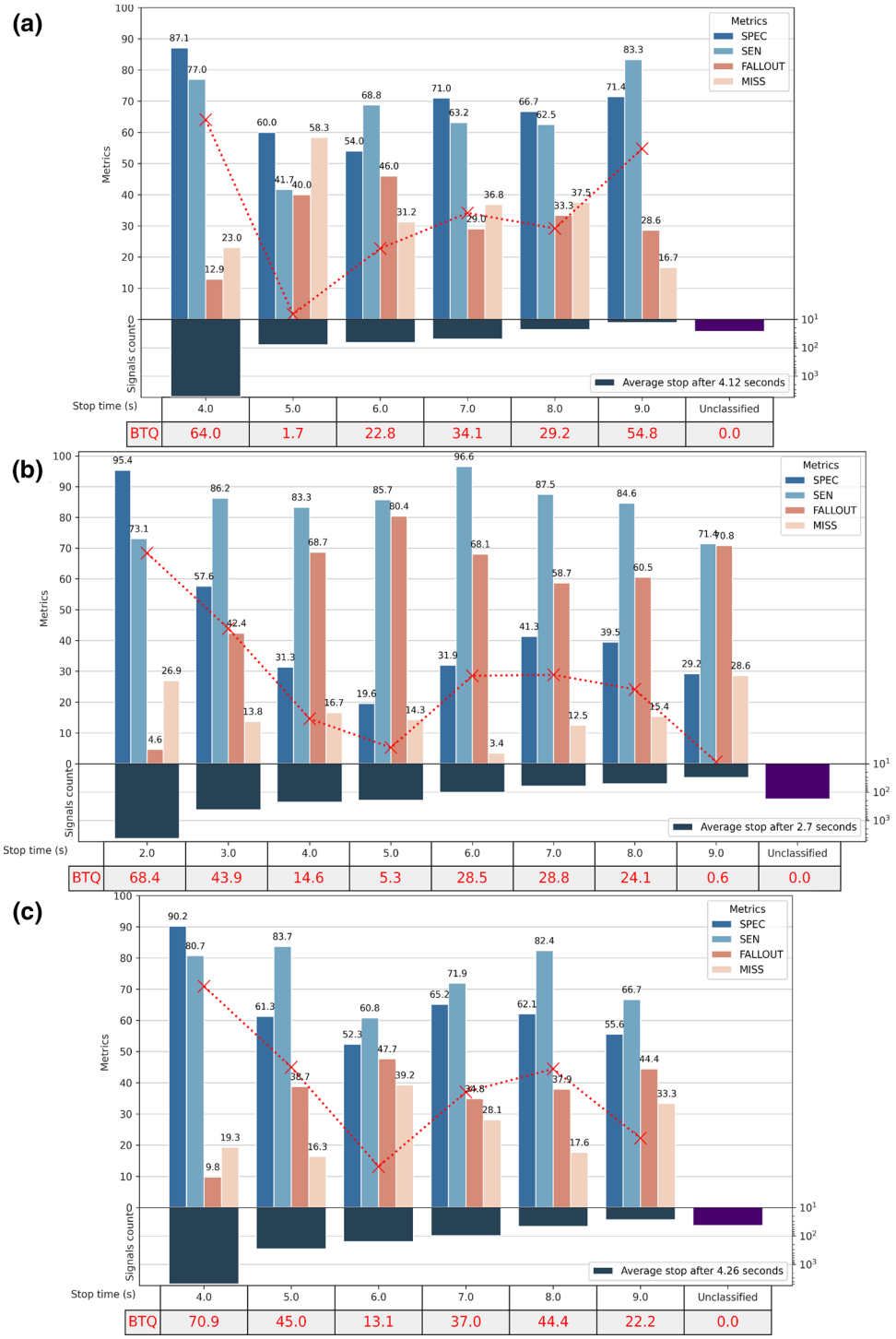

Fig. 3 Distribution of signals and classification metrics with respect to the listening time determined by the Early Stopping algorithm. Each column pools the set of MERs across the database associated with the same listening duration by the early stopping algorithm. The red line represents the evolution of the average BTQ score for each possible stopping time observed. The three comparative methods were evaluated using k-fold cross-validation to determine the optimal alpha and patience parameters. **a** Adaptive listening time with SepaConvNet ($patience = 4; \alpha = 0.5$), **b** Adaptive listening time with simple Bayesian extension ($patience = 2; \alpha = 0.225$)), **c** Adaptive listening time with advanced Bayesian extension ($patience = 4; \alpha = 0.18$)

rate the simple cases for identification from the complex ones. The complexity of identification may be explained by the nature of the physiological boundaries which are not clear-cut, and therefore potentially difficult to interpret. Lastly, this approach is designed to work with signals that are not denoised. Numerous sources of mechanical and physiological noise have been characterized for electrophysiological signals [1,8,11], which can render the signal more complex and difficult to classify. In order to integrate this work in a surgical context, this algorithm could be used to identify simple signals in a much shorter time, and propose longer signal acquisitions and flag more complex signals.

Limitations and future work

Despite the improvement in predictions for all signals that follow a rapid convergence towards certainty (i.e. an output of 0 or 1), which can be seen in the first column of each subfigure in Fig. 3, signals that are classified after the minimum amount of time show a noticeable decrease in accuracy. This is likely because all signals that are completely unambiguous are classified at the very first opportunity, with only the more difficult signals remaining.

Therefore, improving the predictive model's accuracy specifically for this family of signals remains a priority for future work in terms of the underlying neural network, its Bayesian extensions, and the early termination algorithm. One potential avenue for this is to integrate more information related from the DBS planning stage (i.e. the rough location of the STN and the desired electrode trajectories determined from pre-operative MRI) which would likely allow for more unambiguous signals occurring further away from the boundaries of the STN to be quickly classified based primarily on depth information, leaving the neural network to learn distinctive patterns for the more difficult boundary cases rather than the more frequent easier cases. Depth information has been used in many automatic MER signal analysis approaches [4,15,24], and the significant improvement of predictions resulting from its inclusion has recently been validated [4].

Conclusion

Through this study, we proposed a listening algorithm for MER signal classification, capable of adapting to the level of certainty of the predictions made by a predictive model. The purpose of this algorithm is to determine if listening to additional MER signal would improve the classification of a particular MER signal based on an underlying model. Due to the simplicity of this algorithm (with only two parameters, both of which are bounded) we could evaluate and explore a variety of configurations both in terms of the algorithm's parameter space and the underlying machine learning model performing the predictions. The evaluation of this system has highlighted the performance of the advanced Bayesian extension of SepaConvNet, which seems to be well-adapted to the logic of the listening algorithm, outperforming comparative methods, while being fairly insensitive to parameterization. In terms of performance, the combination of the advanced Bayesian method with the early stopping algorithm does not reduce the prediction quality compared to using the entire acquired signal, indicating that such efficiencies can be found without worrying as to incurring speed-vs.-performance tradeoff. Thus, the simple early termination algorithm proposed in this paper would be, as an adjunct to deep learning, a relevant implementable and low-risk addition, allowing for an improvement in surgical workflow efficiency.

Funding This work was funded by Association France Parkinson.

Availability of data and materials Data is not available for this study.

Declarations

Conflict of interest Authors do not have any conflicts of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

Code availability Code is currently not made available.

References

1. Bakštein E, Sieger T, Wild J, Novák D, Schneider J, Vostatek P, Urgošík D, Jech R (2017) Methods for automatic detection of artifacts in microelectrode recordings. J Neurosci Methods 290:39–51
2. Benazzouz A, Breit S, Koudsie A, Pollak P, Krack P, Benabid AL (2002) Intraoperative microrecordings of the subthalamic nucleus in Parkinson's disease. Mov Disord 17(S3):S145–S149
3. Blume J, Schlaier J, Rothenfusser E, Anthofer J, Zeman F, Brawanski A, Bogdahn U, Lange M (2017) Intraoperative clinical testing overestimates the therapeutic window of the permanent dbs electrode in the subthalamic nucleus. Acta Neurochirurgica 159(9):1721–1726
4. Cao L, Li J, Zhou Y, Liu Y, Zhao Y, Liu H (2019) Online identification of functional regions in deep brain stimulation based on an unsupervised random forest with feature selection. J Neural Eng 16(6):066015
5. Coelli S, Levi V, Vecchio JDVD, Mailland E, Rinaldo S, Eleopra R, Bianchi AM (2020) Characterization of Microelectrode Recordings for the Subthalamic Nucleus identification in Parkinson's disease. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC). pp. 3485–3488 (Jul 2020)
6. Coelli S, Levi V, Vecchio JDVD, Mailland E, Rinaldo S, Eleopra R, Bianchi AM (2020) An intra-operative feature-based classification of microelectrode recordings to support the subthalamic nucleus functional identification during deep brain stimulation surgery. J Neural Eng
7. Dostrovsky JO, Lozano AM (2002) Mechanisms of deep brain stimulation. Mov Disord 17(S3):S63–S68
8. Hosny M, Zhu M, Gao W, Fu Y (2020) A novel deep LSTM network for artifacts detection in microelectrode recordings. Biocybern Biomed Eng 40(3):1052–1063
9. Hutchison WD, Lozano AM (2000) Microelectrode recordings in movement disorder surgery. Mov Disord Surg 15:103–117
10. Khosravi M, Atashzar SF, Gilmore G, Jog MS, Patel RV (2020) Intraoperative localization of stn during dbs surgery using a data-driven model. IEEE Journal of Translational Engineering in Health

and Medicine 8:1–9, conference Name: IEEE Journal of Translational Engineering in Health and Medicine
11. Klempíř O, Krupička R, Bakštein E, Jech R (2019) Identification of microrecording artifacts with wavelet analysis and convolutional neural network: an image recognition approach. Measur Sci Rev 19(5):222–231
12. Li Z, Zhang JG, Ye Y, Li X (2016) Review on factors affecting targeting accuracy of deep brain stimulation electrode implantation between 2001 and 2015. Stereotactic Funct Neurosurg 94(6):351–362
13. Lozano CS, Tam J, Lozano AM (2018) The changing landscape of surgery for Parkinson's Disease. Mov Disord 33(1):36–47
14. Martin T, Peralta M, Gilmore G, Sauleau P, Haegelen C, Jannin P, Baxter JSH (2021) Extending convolutional neural networks for localizing the subthalamic nucleus from micro-electrode recordings in Parkinson's disease. Biomed Signal Process Control 67:102529
15. Moran A, Bar-Gad I, Bergman H, Israel Z (2006) Real-time refinement of subthalamic nucleus targeting using Bayesian decision-making on the root mean square measure. Mov Disord 21(9):1425–1431
16. Peralta M, Quoc A, Ackaouy A, Martin T, Gilmore G, Haegelen C, Sauleau P, Baxter J, Jannin P (2020) SepaConvNet for Localizing the Subthalamic Nucleus using One Second Micro-Electrode Recordings. Montreal, Canada (Jul
17. Rodriguez-Oroz MC, Rodriguez M, Guridi J, Mewes K, Chockkman V, Vitek J, DeLong MR, Obeso JA (2001) The subthalamic nucleus in Parkinson's disease: somatotopic organization and physiological characteristics. Brain J Neurol 124(Pt 9):1777–1790
18. Schiaffino L, Muñoz AR, Martínez JG, Villora JF, Gutiérrez A, Torres IM, Kohan, y.D.R. (2016) STN area detection using K-NN classifiers for MER recordings in Parkinson patients during neurostimulator implant surgery. J Phys Conf Ser 705:012050
19. Servello D, Zekaj E, Saleh C, Pacchetti C, Porta M (2016) The pros and cons of intraoperative ct scan in evaluation of deep brain stimulation lead implantation: a retrospective study. Surg Neurol Int 7(Suppl 19):S551
20. Sterio D, Zonenshayn M, Mogilner AY, Rezai AR, Kiprovski K, Kelly PJ, Beric A (2002) Neurophysiological refinement of subthalamic nucleus targeting. Neurosurgery 50(1):58–69
21. Valsky D, Marmor-Levin O, Deffains M, Eitan R, Blackwell K, Bergman H, Israel Z (2017) Stop! border ahead: automatic detection of subthalamic exit during deep brain stimulation surgery. Mov Disord 32(1):70–79
22. Wan KR, Maszczyk T, See AAQ, Dauwels J, King NKK (2018) A review on microelectrode recording selection of features for machine learning in deep brain stimulation surgery for Parkinson's disease. Clin Neurophysiol 130:145–154
23. Wong S, Baltuch GH, Jaggi JL, Danish SF (2009) Functional localization and visualization of the subthalamic nucleus from microelectrode recordings acquired during DBS surgery with unsupervised machine learning. J Neural Eng 6(2):026006
24. Zaidel A, Spivak A, Shpigelman L, Bergman H, Israel Z (2009) Delimiting subterritories of the human subthalamic nucleus by means of microelectrode recordings and a Hidden Markov Model. Mov Disor 24(12):1785–1793