의학석사 학위논문

파킨슨병에서 시상하핵 뇌심부자극술의 미세전극기록으로부터 딥러닝을 이용한 임상 결과 예측

Clinical outcome prediction with deep learning from microelectrode recording of subthalamic deep brain stimulation in Parkinson disease

2019 년 4 월

서울대학교 대학원
의학과 중개의학전공
박 광 현

Abstract

Clinical outcome prediction with deep learning from microelectrode recording of subthalamic deep brain stimulation in Parkinson disease

Park Kwang Hyon
Department of Medicine
(Translational Medicine Major)
The Graduate School
Seoul National University

(Objectives)

Deep brain stimulation (DBS) of the subthalamic nucleus (STN) is an effective treatment to improve the motor symptoms of advanced Parkinson disease (PD). Accurate positioning of the stimulation electrodes to STN is mandatory for better clinical outcomes. However, the precise identification of the STN during the microelectrode recording (MER) is not easy. In this study, we analyzed deep learning based MER signals to better predict the clinical outcome of motor function improvement after bilateral STN DBS in patients with advanced PD.

(Methods)

696 left MER segments of 4 seconds length from 34 PD patients with advanced PD who underwent bilateral STN DBS surgery under general anesthesia were included in this study. The datasets of thirty patients were assigned to the training set, and the datasets of four patients were assigned to the test set. The wavelet transformed MER and the ratio of DBS on and off Unified Parkinson's Disease Rating Scale(UPDRS) Part III score of the off-medication state were applied for deep learning. According to the ratio, the patients were divided into two groups, high-responder and moderate-responder group. Visual Geometry Group(VGG)-16 model with multi-task learning algorithm was used to estimate the bilateral effect of DBS. To apply the effect of the contralateral score more than ipsilateral score, the ratio of the loss function was varied. Gradient class activation map was used to marking the lesion of interest of CNN.

(Results)

When we divided MER according to the frequency band and transformed to wavelets, the maximal accuracy was the highest in the 50-500 Hz group, compared with 1-50 Hz and 500-5,000Hz groups. In addition, when the multitask-learning method was applied to 50-500Hz group, the stability of the model was prominently improved. The max accuracy was the highest(80.2%) when the loss ratio of right to left was given as 5:1 or 6:1 in the model. Area under the curve(AUC) was 0.88 in the receiver-operating characteristic(ROC) curve. Gradient class activation map showed that 80-200Hz band was the most commonly referenced area.

(conclusion)

We confirmed that the clinical improvement of PD patients who underwent bilateral STN DBS could be predicted based on multi-task deep learning based MER analysis. The deep learning based MER analysis could be helpful for determining the position of the electrode, by predicting motor function improvement.

………………………………………

keywords **:** Multi-task Learning; Subthalamic Nucleus Deep Brain Stimulation; Microelectrodes Recording; Parkinson Disease; Convolutional Neural Network; Deep Learning

Student Number **: 2013-21727**

Contents

Introduction

PD – DBS – UPDRS

Deep Brain Stimulation(DBS) of the Subthalamic Nucleous(STN) is effective way to improve motor complication in advanced Parkinson disease(PD).(1) Defining and targeting of The STN has been improved due to advance of magnetic resonance imaging(MRI) technique. While brain shifting and intrinsic inaccuracy in frame-based targeting still remain obstacles, intraoperative signal analysis from microelectrode recording(MER) has been usually performed.(2) Although there have been papers that classify and analyze MER and correlate them with clinical outcomes (3),(4),(5), there have been no studies that directly match the MER signal and clinical outcomes of the patients with advanced PD after STN DBS. Among the machine learning methods, CNN can process signal by using imaging processing method.

Signal – CNN – Clinical outcome

In this study, we analyzed outcome-guided deep learning based prediction of motor function improvement after bilateral STN DBS in PD patients. We transformed MER signals to wavelet form which has time axis, compared with Unified Parkinson's Disease Rating Scale(UPDRS) part III scores of the patients with advanced PD at 6 months after STN DBS. We investigated the span of band in MER signals and the deep learning algorithm, by which the clinical outcome of motor function improvement are better predicted in patients with advanced PD after bilateral STN DBS.

Methods

Subjects

44 patients with advanced PD who underwent bilateral STN DBS with MER under general anesthesia between 2014 and 2017 were included in this study. Six patients who had missed clinical data and four patients who had dual electrode stimulation at 6 months after STN DBS were excluded. The clinical information of all patients was retrospectively reviewed. The UPDRS scores of every patient were evaluated preoperatively and at 6 months after surgery by experienced neurologists. The clinical evaluation of UPDRS scores were performed under two conditions (off-medication when the patients had taken no medication for 8 to 12 hours and on-medication when the patients had experienced maximal clinical benefit 1 to 3 hours after the usual morning dose of dopaminergic treatment) before and at 6 months after STN DBS. In this study, only off-medication status was included. The MER was performed at 1 mm intervals from −10 mm to −5 mm and 0.5 mm intervals from −5 mm to + 5 mm. Since the lead was 1.5mm span and MER was performed at intervals of 0.5mm, three input was inevitably matched to the same output value. If stimulated lead depth was changed, the depth of MER which was thought to be stimulated was selected.

Surgical procedure

The detailed surgical procedure was almost same with the previous studies.(6), (7) Anti-Parkinsonian medications were not stopped preoperatively. A stereotactic Leksell(R)-G frame (Elekta Instruments AB, Stockholm, Sweden) was mounted on the head under local anesthesia. A brain MRI with 1.5T system was taken. (General Electric Medical System, Milwaukee, WI, U.S.A.) The FSPGR 3-D sequence was used for anterior commissure (AC) – posterior commissure (PC) calculations. T2 spin-echo images were obtained to define the boundaries of the STN. SurgiPlan® (Eleckta, Stockholm, Sweden) conducted the simulations for targeting the sensorimotor region of the STN and selecting the trajectories. The depth of sedation was monitored by the bispectral index under total intravenous anesthesia with propofol and remifentanil. Propofol concentration was titrated to maintain the bispectral index value of 60-80. MER by LeadPoint (Medtronic, Minneapolis, MN) was performed. The permanent quadripolar electrodes were implanted along the proper trajectory to stimulate more sensorimotor region of the STN, which was localized by a combination of preoperative brain MRI and intraoperative MER. Left DBS was performed first, followed by right DBS. Programmable pulse generator (Medtronic, Minneapolis, MN) were implanted in the subclavicular region and connected to the electrodes.

Microelectrode Recordings

We got the signal using a five-array microdrive called ben gun. Each MER signal was filtered at 0.2-5,000Hz. The signal was output through the output board of the Leadpoint(Medtronic, Minneapolis, MN) with a sample rate of 48 KHz gain 1,000. The output signal was acquired using Spike2. These acquired signals were used in the analysis. The MER signal was selected from the trajectory determined to be implanted. Finally, the recorded MER signal was selected from the contact position at 6 months after DBS. We analysed only left MER signals for the unity of input data for deep convolutional neural network(CNN) algorithm. 696 MER segments from 34 PD patients who underwent bilateral STN DBS surgery in general anesthesia were included in this study. The datasets of thirty patients were assigned to the training set, and those of four patients were assigned to the test set.

Wavelet Transformation

There are several ways to visualize the signal, such as Short-Time Fourier Transform(STFT) and wavelet transform(WT). (8) By visualizing the signal, you can see which signal is dominant on frequency domain. The WT is well known approach to the problem of time−frequency analysis of signals such time signal as MER. It is developed for seismic signal analysis(9), and also used in similar application neurophysiologic spike train composition(10). Compared to the Fourier transformation wavelet transformation has ability to reach to the optimal time−frequency resolution due to unfixed window size. (11) In generally Recurrent Neural Network(RNN) has advantages for analyzing sequential or time signal. However, in such case of dealing frequency signal like MER having frequency range of Voice Frequency (VF − 300~3000HZ) and Very Low Frequency (VLF − 3000~30000HZ). Finally, we used wavelet of 50~500Hz, time and frequency resolution factor 8. The length of the wavelet is determined as 4 seconds considering time frequency resolution, learning result and MER segment number. We used Morlet function as a mother wavelet.

Training set and Test set

A total of 696 MER segments were used in 34 patients, with an average of 20.47 segments per person. We distinguished between training set and test set, not data, but patients. Thus, the same patient data can be applied to the MER segment of a new patient, eliminating the case of redundant use in training and testing set. For the training set, 2 patients in the good group and 2 patients in the moderate group were randomly selected.

According to the ratio of DBS on and off score at 6 month after surgery, the patients were divided into two groups, good and moderate. UPDRS score of the right side is better than that of the left side, the classification criteria are determined differently for each side. For left side, 0–0.69 is classified as good response group, 0.7–1.0 as moderate response group. For right side, 0–0.59 is classified as good response group, 0.6–1.0 as moderate response group.

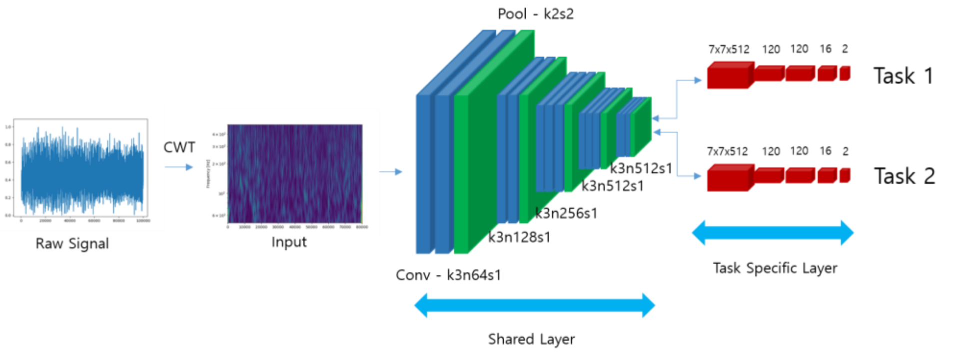
Deep learning

Deep learning is a state of the art algorithm which can extract high, low–level abstractions of data, well known algorithms for Deep Learning includes Recurrent Neural Network(RNN), Convolutional Neural Network(CNN), Generative Adversarial Network(GAN). Among them, CNN has been working best on image processing. One of the CNN structures, Visual Geometry Group(VGG16) (12), has been widely used since the Large Scale Visual Recognition Challenge (ILSVRC) 2014. VGG16 was used as classification network. We modified the network from fully connected layer 4096, 4096, 3(number of class) to 120, 120, 16, 2 with dropout rate(0.2), Learning rate of 0.001 for both multi task learning label with Gradient Descent Optimizer was used. It is difficult to analyze MER by RNN, because the sampling rate is large(20KHz) and the data capacity is large. With wavelet analysis, signal with large capacity can be reduced to small capacity.

Multi−task learning

Multi−task learning has been used successfully in machine learning. And Recently used in computer vision. (13) There are three types of multi−task learning: Joint learning and Learning to learn, Learning with auxiliary task. We adopted joint−learning. With multi−task learning we trained CNN with different two labels. As different tasks, a model that learns two tasks simultaneously is able to learn a more general representation.(14) Learning with label A and label B model generalize better than training with one label. We use convolution layers of VGG16 as a shared layer and train fully connected layer separately with each side labels. Furthermore, we jointly trained single network with two different label with different ratio of loss function. Figure 1 shows whole process schematic diagram. (Figure 1)

Figure 1. Process schematic diagram



Gradient class activation map

Gradient class activation map was known for attention extraction method.(15) We adopted Gradient CAM for finding attention-map for trained CNN layer. We calculate Gradient CAM between last convolution and fully connected layer size of feature map is 14x14x512.

$$\alpha_k^c = \frac{1}{Z}\sum_i \sum_j \frac{\delta y^c}{\delta A_{ij}^k}$$

$$L_{Grad-Cam}^c = ReLu(\sum_k \alpha_k^c A^k)$$

(α_k^c : a partial linearization of the deep network downstream from A, and captures the 'importance' of feature map k for a target class c, the gradient of the score for class c, y^c (before the softmax), with respect to feature maps A_{ij}^k of a convolutional layer, i.e. δy^c , δA^k) calculating $\alpha_k^c A_{ij}^k$ and sum $\alpha_k^c A_{ij}^k$ depthwisely (in thise case k = 512) and passing activation function ReLu to remove the negative value.

Statistical analysis

Statistical analysis was performed with R 3.5.1. For correlation analyses, the correlation coefficient is represented by the letter 'r' with Pearson's correlation analysis.

IRB

This study was conducted adhering to the declaration of Helsinki. The study was performed retrospectively and the records of patients were anonymized and de-identified prior to analysis. Patients gave their written informed consent to have their health information recorded and used for research purposes. The study protocol was approved by the Institutional Review Board of Seoul National University Hospital(IRB No. 1812-086-995).

Results

Patient Data

The MER data of thirty-four patients were included in the study. (table 1) The mean age of patients at diagnosis is 43.1 ± 8.0, the mean age of the patients at surgery is 57.2 ± 8.0. The duration from onset of symptom to surgery is 14.0 ± 5.6. Among the UPDRS part III scores, the subtotals of only the items with left or right segments are denoted by "Left" or "Right", respectively. The mean preoperative UPDRS part III score is 12.4 ± 5.2 for the right side of the body, 12.8 ± 5.4 for the left side of the body. The mean off-DBS UPDRS part III score at 6 months after surgery is 11.2 ± 4.4 for the right side of the body, 11.5 ± 4.4 for the left side of the body. The mean on-DBS UPDRS part III score at 6 months after surgery is 6.8 ± 4.2 for the right side of the body, 7.6 ± 4.1 for the left of the body.

Table 1. Off-Medication Scores in 34 Patients with Bilateral Stimulation of the Subthalamic Nucleus

Score	Score Before STN DBS	Score at 6 Month After STN DBS, Off-Stimulation	Score at 6 Month After STN DBS, On-stimulation
UPDRS Part I	3.9 ± 3.3	NA	3.1 ± 2.1
UPDRS Part II	23.8 ± 10.0	NA	14.1 ± 9.0
UPDRS Part III Right	12.4 ± 5.2	11.2 ± 4.4	6.8 ± 4.2
UPDRS Part III Left	12.8 ± 5.4	11.5 ± 4.4	7.6 ± 4.1
UPDRS Part III Total	40.3 ± 15.0	34.9 ± 12.3	22.3 ± 11.8
UPDRS Total	68.0 ± 23.0	NA	39.5 ± 17.9
Global stage of disease (Hoehn and Yahr)	3.2 ± 0.9	2.8 ± 0.8	2.5 ± 0.7
Global activities of daily living (Schwab and England)	45.6 ± 23.8	NA	64.1 ± 25.8
LEDD	1411.7 ± 547.3	NA	431.5 ± 286.7

The preoperative scores of UPDRS part III of the left and right of the body in 34 patients show linear correlation($r=0.77$, $p<0.001$).(Figure 2) In 6 months after surgery, This correlation also existed in both of DBS on($r=0.79$, $p<0.001$) and off($r=0.74$, $p<0.001$) states.(Figure 3A,3B) Preoperative UPDRS Part III score prior to surgery does not affect DBS On/Off ratio at 6 months after the operation.(Figure 4A, 4B) However, at 6 months, DBS on/off ration of the left body according to that of the right is significantly correlated($r=0.67$, $p<0.001$).(Figure 5)

Figure 2. Left to Right Correlation of Preoperative UPDRS Part III score (Medication Off)

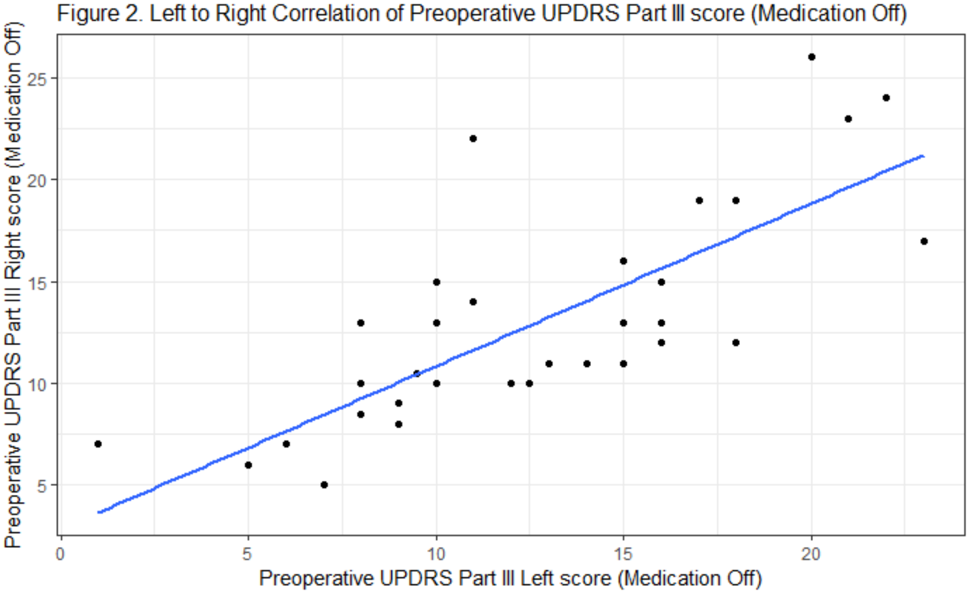


Figure 3A. Right to Left correlation of UPDRS Part III at 6 months (DBS On, Medication Off)

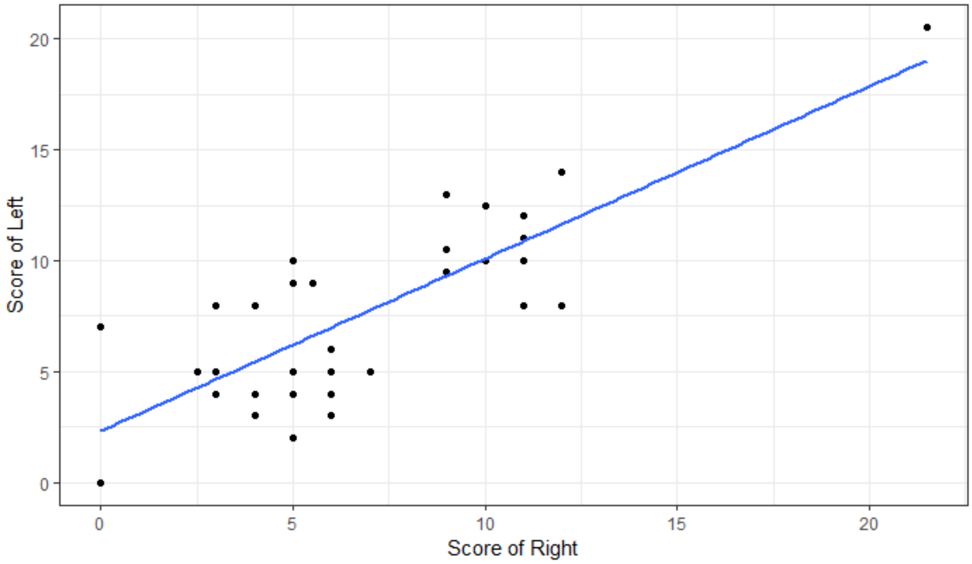


Figure 3B. Right to Left correlation of UPDRS Part III at 6 months (DBS Off, Medication Off)

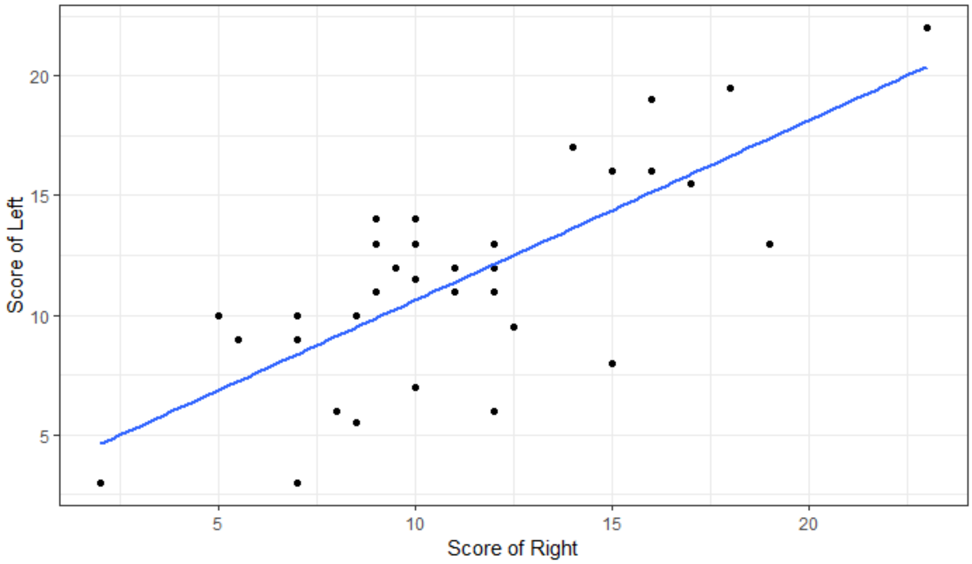


Figure 4A. Left DBS On/Off ratio at 6 months according to Preoperative Part III score (Medication Off)

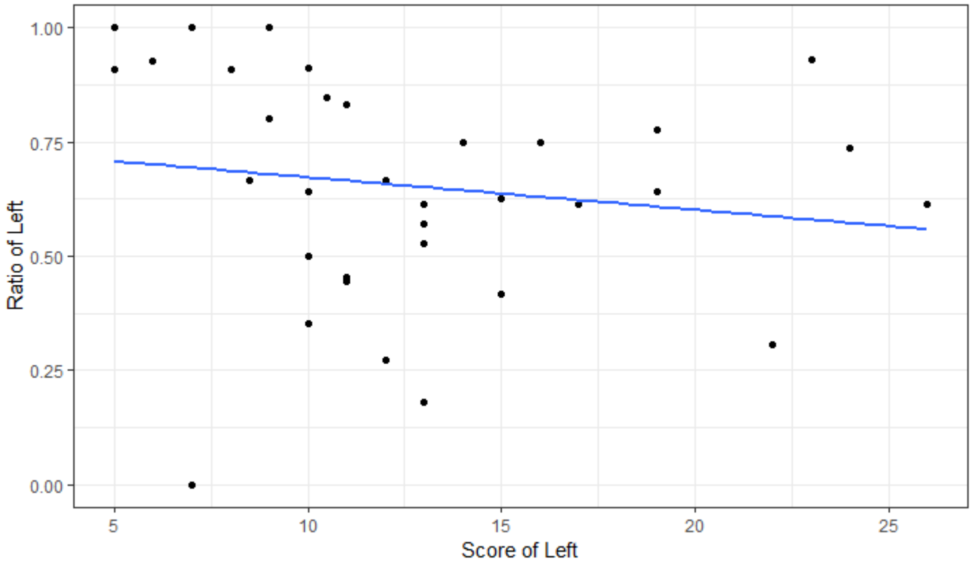


Figure 4B. Right DBS On/Off ratio at 6 months according to Preoperative part III score (Medication Off)

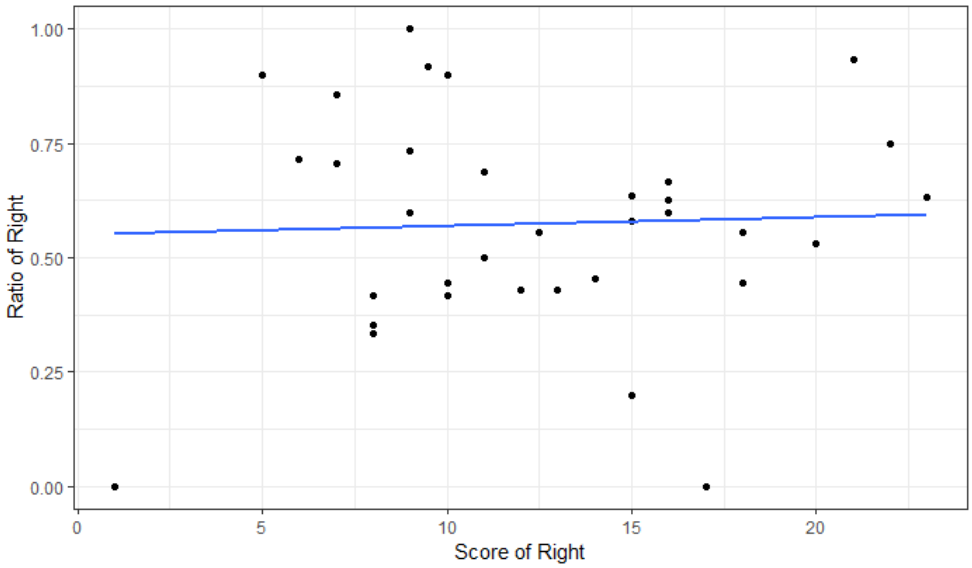
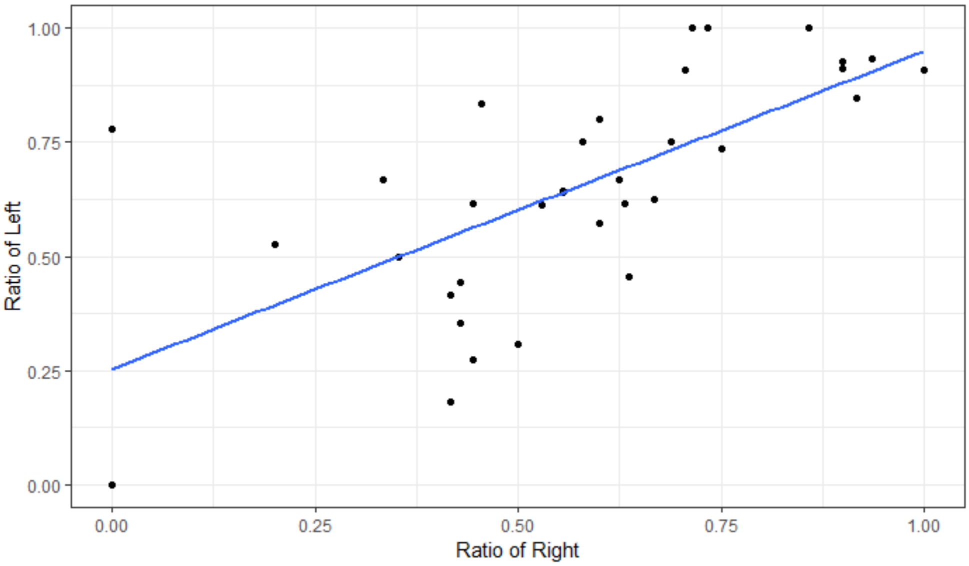


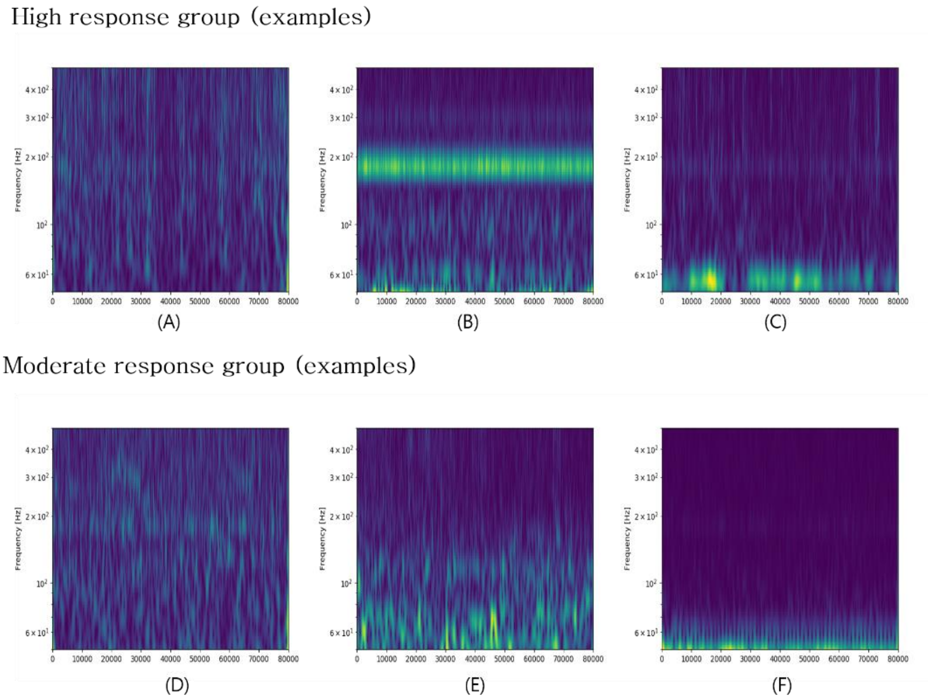
Figure 5. Right to Left correlation of UPDRS Part III DBS On/Off ratio at 6 months (Medication Off)



MER & clinical outcome relation

The MER data is passed through a 50−500 Hz bandpass filter, cut into 4−second lengths, normalized, and transformed into wavelets.(Figure 6A−F) It is classified according to clinical outcome.

Figure 6. Examples of Wavelet for High and Moderate response group



Single−task learning and Multi−task learning

When we divided MER according to the frequency band and transformed to wavelets, the max accuracy was the highest in the 50−500 Hz group(75.0%), compared with 1−50 Hz group(33.0%) and 500–5,000 Hz group(62.5%). However, as the model repeats epochs, the results become unstable.(Figure 7A) We used multi–task learning to take into account the effect of DBS on both bodies and found the ratio of maximal accuracy to the ratio of left and right loss functions. The max accuracy was the highest(80.2%) when the loss ratio of right to left was given as 5:1 and 6:1 in the model(77.1% at 4:1, 71.9% at 3:1, 65.6% at 2:1).(Table 2) The stability of the model was prominently improved, compared with single–task learning.(Figure 7B) The receiver–operating characteristic(ROC) curve was obtained from prediction for the clinical outcome.(Figure 8) Area under the curve(AUC) was 0.88.

Figure 7. Accuracy of Single-task Learning and Multi-task Learning According to Learning Epoch

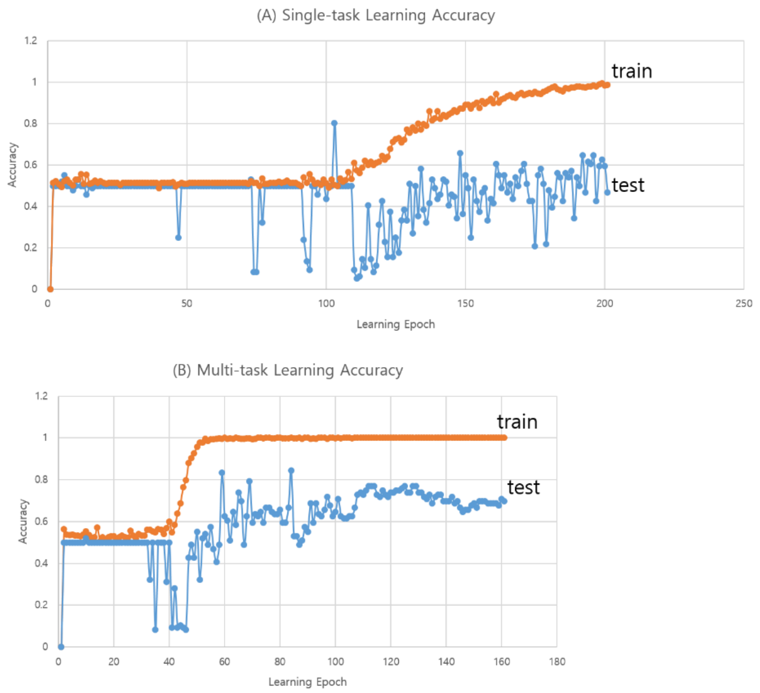
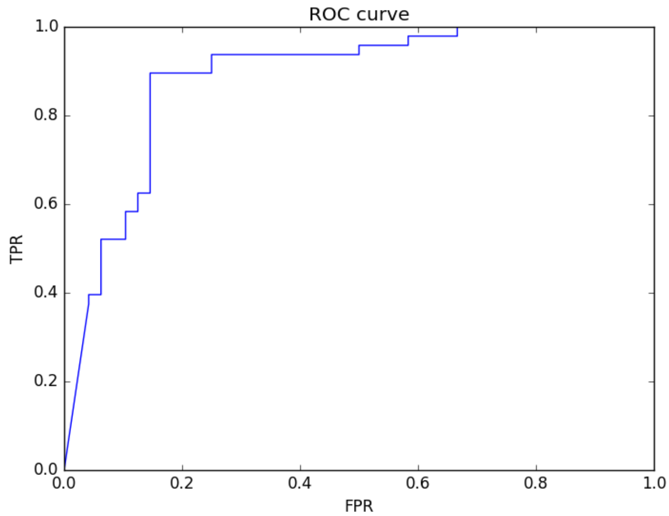


Table 2. Maximal Accuracy according to Loss Ratio of Right to Left

Loss Ratio (Rt.:Lt.)	1:5	1:4	1:3	1:2	2:3	1:1	3:2	2:1
Max. Accuracy (%)	65.6	59.4	67.7	68.8	64.6	64.6	66.7	65.6

Loss Ratio (Rt.:Lt.)	5:2	3:1	4:1	5:1	6:1	7:1	8:1	9:1	10:1
Max. Accuracy (%)	66.7	71.9	77.1	80.2	80.2	77.1	67.7	67.7	67.7

Figure 8. ROC curve

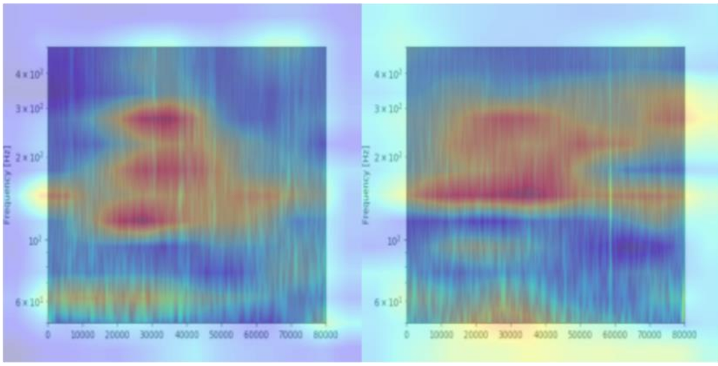


Gradient Class Activation Map

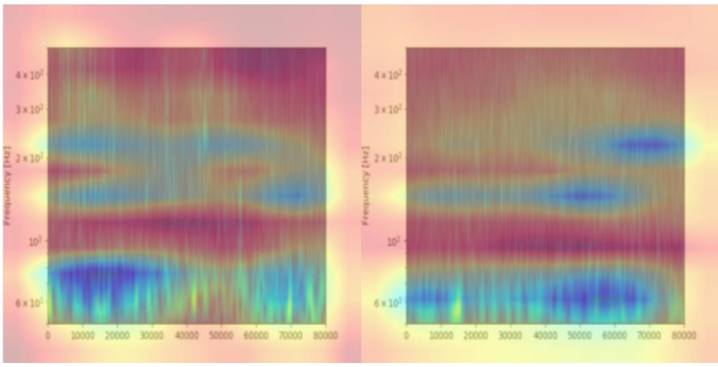
When we tested the wavelets in the frequency range of 50–500Hz in the test set, the class activation map result as shown in the figure was obtained.(Figure 9) Although varying according to the test data, the most commonly observed attention area is in 80−200Hz.

Figure 9. Gradient Class Activation Map

(A) High response group



(B) Moderate response group



Discussion

Single-task Learning

When the wavelet transform of the entire MER is performed, it is divided into three frequency zones considering the resolutions available as input of CNN. It was the same as frequency band pass because wavelet was divided according to frequency. As is well known, artifacts were most common at 1-50 Hz wavelet. At 500-5,000 Hz wavelet, the spike seemed the most noticeable, but it was also seen at 50-500 Hz wavelet. Typical firing patterns, particularly asymmetrical spikes at high frequency with bursting patterns, and proprioceptive responses to passive has been used to determine the boundary of STN.(1) On the other hand, because β-oscillatory activity(13-30Hz) was limited to dorsolateral oscillatory region of the STN, the length of the dorsolateral oscillatory region recorded in the macroelectrode-implanted trajectory predicted a favorable response to STN DBS.(4)

There is also a study that high-frequency(500-2,000Hz) background reflects the location of the STN.(3) Unexpectedly, the results of single-task learning in 50-500Hz wavelet were best. It seems that 50-500Hz wavelet shows most consistent feature that is best to CNN.

Multi−task Learning

Preoperative UPDRS Part III score prior to surgery does not affect DBS On/Off ratio at 6 months after the operation.(Figure 4A, 4B) In other words, there is no significant correlation between the preoperative condition and postoperative improvement. We believe the reason is that the position of the electrode affected the clinical outcome mostly. In this study, the ratios of left and right are correlated with each other at 6 months.(Figure 5) This is probably because if either electrode is well inserted, it also has a positive effect on the ipsilateral motor symptoms. Unilateral STN DBS is able to improve the bilateral UPDRS motor scores of PD patients. The ratio of ipsilateral side symptom improvement to contralateral side is approximately 1.63−2.09:1 according to the paper.(16),(17),(18) Although the study did not show DBS Off scores at 6 months, similar results can be deduced from the fact that there is not a large difference between the preoperative score and DBS off scores of postoperative 6 month. To overcome the instability and the inaccuracy of the single−task learning model, multi−task learning using the ipsilateral symptom data would be a good option.

The maximal accuracy over 70% was right to left loss ratio of 3−6:1. STN DBS has a bilateral effect, but it has more effect on the contralateral side. It seems that the same result as the papers reporting the clinical outcome in unilateral DBS. Multi-task learning is more stable than single-task learning. This is because the bilateral effect is considered, but on the other hand, multi-task learning itself has the effect of preventing overfitting. Considering such a point, it is meaningful that the maximal accuracy is higher in one model to refer to the contralateral score more.

Gradient Class Activation Map

We used Gradient class activation map, a technique for marking the lesion which affects the result value more in CNN image recognition. It was to try to deduce what lesion was used as the basis of judgment in the model. There was no perfect consistency, but it was the most frequent reference to the 80-200Hz band. However, due to the nature of CNN, it is dangerous to link these frequency bands has the basis of clinical judgment.

Expected Clinical Usefulness

In our center, STN is considered as the place where the typical firing pattern is visually observed in the MER, and the final lead position is determined. In the case of awake DBS, we could check the results by macrostimulation the selected target, but it is not possible at present because of asleep DBS. It would be helpful if we could predict the clinical outcome of DBS from the single MER. In particular, when a typical firing pattern is invisible, this will be helpful. The method of determining lead position by a surgeon is different, but this can be used as common reference material. Recently, there have been published article that predict the clinical outcome from electrode coordinate using deep learning,(19) There were no articles that used deep learning to measure clinical outcome while leaving time information of MER data.

Limitation

The left DBS was performed first and the right DBS was performed on the same day. The right side MER was excluded because it was judged that bias due to brain shifting was more severe on the right DBS. Furthermore, only asleep DBS with the same medication was included to reduce the anesthesia and awake bias, such as patient movement or snoring. Future studies will include all the excluded data. It will be good to find a way to input the left and right MER signal simultaneously to the learning algorithm. Because we used the data on 6 months after surgery, we cannot reflect the surgical and medical problems such as infection, disease progression. CSF leakage or head position can make brain shifting during surgery, but we ignored the brain shifting. Brain shifting can change the exact location of lead at 6 months after surgery. We expect to obtain better results by finding the way to reflect each brain shifting. Since the lead was 1.5mm span and MER was performed at intervals of 0.5mm, three input was inevitably given as one output value. In the further study, we will get better results by labeling each one at 1:1, or adding information about the spatial relationship.

Conclusion

Clinical Improvement of PD patients who underwent bilateral STN DBS can be predicted from MER using multitask deep learning. In the meantime, we have analyzed the MER signal with the characteristic such as spikes and frequency. The position of the appropriate electrode has been determined by comparing each clinical outcome. We then have adjusted the stimulus position and intensity to give the best results in follow up. In this study, the MER signal itself and the clinical outcome were matched end-to-end. Especially, we used multi-task learning to consider the bilateral effect of unilateral DBS. The results of the experiment with different ratio of loss function shows that 5:1 and 6:1 has best maximal accuracy(80.2%). We visualized the basis for judgement using gradient class activation map. We expected to be able to use machine learning to find out the differences that people could not find in the conventional way. It is not possible to accept the result as it is due to various limitations. After correcting the limitations one by one and satisfying the results, we will be able to figure out the hidden factor in the MER data. We believe that this may be helpful for determine the location of electrode in new patients.

References

1. Benabid AL, Chabardes S, Mitrofanis J, Pollak P. Deep brain stimulation of the subthalamic nucleus for the treatment of Parkinson's disease. Lancet Neurol. 2009;8(1):67-81.

2. Li Z, Zhang JG, Ye Y, Li X. Review on Factors Affecting Targeting Accuracy of Deep Brain Stimulation Electrode Implantation between 2001 and 2015. Stereotact Funct Neurosurg. 2016;94(6):351-62.

3. Novak P, Daniluk S, Ellias SA, Nazzaro JM. Detection of the subthalamic nucleus in microelectrographic recordings in Parkinson disease using the high-frequency (> 500 hz) neuronal background. Technical note. J Neurosurg. 2007;106(1):175-9.

4. Zaidel A, Spivak A, Grieb B, Bergman H, Israel Z. Subthalamic span of beta oscillations predicts deep brain stimulation efficacy for patients with Parkinson's disease. Brain : a journal of neurology. 2010;133(Pt 7):2007-21.

5. Vallabhajosula S, Haq IU, Hwynn N, Oyama G, Okun M, Tillman MD, et al. Low-frequency versus high-frequency subthalamic nucleus deep brain stimulation on postural control and gait in Parkinson's disease: a quantitative study. Brain Stimul. 2015;8(1):64-75.

6. Kim W, Song IH, Lim YH, Kim MR, Kim YE, Hwang JH, et al. Influence of propofol and fentanyl on deep brain stimulation of the subthalamic nucleus. J Korean Med Sci. 2014;29(9):1278-86.

7. Lee WW, Ehm G, Yang HJ, Song IH, Lim YH, Kim MR, et al. Bilateral Deep Brain Stimulation of the Subthalamic Nucleus under Sedation with Propofol and Fentanyl. PLoS One. 2016;11(3):e0152619.

8. Daubechies I. The wavelet transform, time−frequency localization and signal analysis. IEEE transactions on information theory. 1990;36(5):961−1005.

9. Goupillaud P, Grossmann A, Morlet J. Cycle−octave and related transforms in seismic signal analysis. Geoexploration. 1984;23(1):85−102.

10. Przybyszewski A. An analysis of the oscillatory patterns in the central nervous system with the wavelet method. Journal of neuroscience methods. 1991;38(2−3):247−57.

11. Ren Z, Qian K, Wang Y, Zhang Z, Pandit V, Baird A, et al. Deep scalogram representations for acoustic scene classification. IEEE/CAA Journal of Automatica Sinica. 2018;5(3):662−9.

12. Simonyan K, Zisserman AJapa. Very deep convolutional networks for large−scale image recognition. 2014.

13. Girshick R, editor Fast r−cnn. Proceedings of the IEEE international conference on computer vision; 2015.

14. Ruder S. An overview of multi−task learning in deep neural networks. arXiv preprint arXiv:170605098. 2017.

15. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D, editors. Grad−cam: Visual explanations from deep networks via gradient−based localization. Proceedings of the IEEE International Conference on Computer Vision; 2017.

16. Walker HC, Watts RL, Guthrie S, Wang D, Guthrie BL. Bilateral effects of unilateral subthalamic deep brain stimulation on Parkinson's disease at 1 year. Neurosurgery. 2009;65(2):302−9; discussion 9−10.

17. Chung SJ, Jeon SR, Kim SR, Sung YH, Lee MC. Bilateral effects of unilateral subthalamic nucleus deep brain stimulation in advanced Parkinson's disease. Eur Neurol. 2006;56(2):127-32.
18. Sobstyl M, Zabek M, Zaczynski A, Gorecki W, Mossakowski Z, Brzuszkiewicz-Kuzmicka G. Unilateral Subthalamic Nucleus Stimulation in the Treatment of Asymmetric Parkinson"s Disease with Early Motor Complications. Turk Neurosurg. 2017;27(2):294-300.
19. Baumgarten C, Haegelen C, Zhao Y, Sauleau P, Jannin P. Data-Driven Prediction of the Therapeutic Window during Subthalamic Deep Brain Stimulation Surgery. Stereot Funct Neuros. 2018;96(3):142-50.

요약(국문초록)

파킨슨병에서 시상하핵 뇌심부자극술의 미세전극기록으로부터 딥러닝을 이용한 임상 결과 예측

박광현

의학과 중개의학전공

서울대학교 대학원

연구 배경

시상하핵의 뇌심부자극술은 진행된 파킨슨병에서 운동 증상을 호전시키는 효과적인 치료이다. 좋은 임상적인 결과를 위해 자극 전극을 정확하게 위치시키는 것이 필요하다. 하지만 미세전극측정을 통해서도 시상하핵을 정확하게 식별하는 것이 쉽지 않다. 이 연구에서는 진행된 파킨슨병 환자에서 딥러닝을 기반으로 미세전극측정을 분석하여 양측 시상하핵 뇌심부자극술 후의 운동기능 호전 정도를 예측하였다.

연구 방법

이 연구에는 전신마취 하에서 양측 시상하핵 뇌심부자극술을 시행받

은 34명의 환자로부터 측정된 4초 길이의 좌측 미세전극측정 분절이 포함되었다. 30명의 환자는 훈련군으로 4명의 환자는 실험군으로 구분하였다. 웨이브릿(wavelet) 변환된 미세전극측정 자료와 UPDRS(Unified Parkinson's Disease Rating Scale) 파트 III 중 오프-약물(Off-medication) 시기의 뇌심부자극/비자극 점수가 딥러닝에 사용되었다. 그 비율에 따라 고반응군과 중반응군으로 분류하였다. 다중작업학습 알고리즘을 이용한 VGG-16 모델이 DBS의 양측성 효과를 추정하기 위해 사용되었다. 동측의 점수보다 반대측의 점수를 크게 반영하도록 하기 위해 손실함수(loss function)의 비율을 다양하게 적용 하였다. CNN이 참조한 영역을 표시하기 위해 Grad-CAM을 사용하였다.

연구 결과

미세전극측정신호를 주파수 대역 별로 나누어 웨이브릿 변환하였을 때, 최대정확도는 1-50Hz와 500-5,000Hz와 비교하여 50-500Hz에서 가장 높았다. 게다가 다중작업학습을 적용하였을 때 모델의 안정도가 더 개선되었다. 최대 정확도는 좌우 손실함수의 비율이 5:1과 6:1 때 80.2%로 가장 높았다. 수신자 조작 특성 곡선(ROC curve)에서 곡선하 면적(AUC) 값은 0.88이었다. Grad-CAM에서는 80-200Hz 대역을 가장 흔히 참조한 것을 보여주었다.

연구 결론

미세전극측정의 다중작업학습을 통한 분석으로 파킨슨병 환자에서

양측 시상하핵 뇌심부자극술 시행 후 임상적 호전에 관한 예측이 가능할 것으로 판단하였다. 딥러닝으로 미세전극측정신호를 분석하여 수술 후 운동기능향상을 예측함으로써, 전극의 위치를 결정하는 데에 도움이 될 것으로 기대한다.

……………………………………

주요어 : 다중작업학습; 시상하핵 뇌심부자극술; 미세전극측정; 파킨슨병; 콘볼루션 신경망; 딥러닝

학 번 : 2013-21727