**Studies on the usefulness of applying deep learning to interpreting EMG waveforms**

**Introduction**

Electromyography (EMG) is an electrophysiologic test that records electrical activity generated from nerves, muscles, and neuromuscular junctions at rest and during muscle contraction controlled by nervous system through needle inserted into the muscle. (1-6) It is implemented to identify disorders of the nervous system or muscles according to abnormalities in EMG signals that reflect the anatomical and physiological characteristics of the nervous system and muscles. (1-6) Among signal shown on EMG, motor unit action potentials (MUAPs) is used to characterizing whether normal or neuropathy or myopathy. Typical neurogenic MUAPs, commonly observed in neurological diseases, are known to have large amplitudes, long durations, reduced recruitments, and reduced interference patterns, whereas myopathic MUAPs are known to have small amplitudes, short durations, and early recruitments and interference patterns. The diagnostic usefulness of electromyography for identifying peripheral neuropathy and myopathy has been suggested in previous studies.(1, 5-12)

Although electromyography plays an important role in diagnosing normal, neuropathy and myopathy, it has some limitations in that there are discrepancies among examiners, and the accuracy of EMG relies to a lot extent on proficiency of the examiner. Previous studies have reported that sensitivity of EMG in the diagnosis of neuropathy, myopathy, and normal is 47-83%, specificity is 73-81% and inter-rater reliability is 62-81%. (13-15) Additionally, to recognize abnormalities of EMG waveform accurately, considerable time and efforts are needed. As the prevalence of neuropathy and myopathy continues to increase, the frequency of EMG for diagnosing it, the time it takes to interpret it, and the workload of the examiner are bound to increase. (16-19) A new approach may be helpful in clinically diagnosing neuropathy or myopathy through EMG more efficiently and accurately in a shorter time.

Recently, deep learning has been used to analyzing big data in many field, and it is also applied to clinical data including waveform, time series data.(20, 21) Convolutional neural network, one kind of deep learning techniques, has applied to analyzing time series data and waveform data such as electrocardiography, electroencephalography. (20, 22, 23) As a result of the study of reading the results of electrocardiography and electroencephalography using deep learning, the accuracy was similar to or superior to that of medical students or residents, and detect nonobvious abnormalities easily overlooked.(24) Previous studies that analyzed EMG using machine learning were mostly those that analyzed surface EMG or needle EMG signals in resting state. (25-29) To our knowledge, few studies on analyzing EMG data in volitional state have rarely been reported.

In order to overcome the limitations of electromyography and provide a more efficient analyzing method, deep learning was applied to interpreting the EMG waveform. We named our machine learning model nEMGNet with motifs from VGGNet and ResNet, which are known to show good performance in image analysis. (30, 31) To confirm the usefulness of nEMGNet, the accuracy of nEMGNet was compared with the accuracy by physicians who currently interpret EMG.

For this study, we retrospectively reviewed EMG waveforms, which were examined in patients with neuropathy or myopathy or normal, analyzed those by using convolutional neural network built-in Python.

**Methods and materials**

**Data acquisition and preparation**

The data analyzed in this article were from the Seoul National University Hospital database that includes electromyography data of 58 subjects, visited Seoul National University Hospital from Jun, 2015 to Jul, 2020 and divided into 3 datasets of neuropathy, myopathy and normal according to EMG waveforms characteristics which were neurogenic potentials or myopathic potentials or not. The criteria for dividing myopathy, neuropathy, and normal were whether there are one of small amplitude-short duration, high amplitudes-long duration, reduced recruitment, early recruitment, and reduced interference pattern during minimal, moderately and maximally contraction. This study was approved by the Internal Review Board of Seoul National University Hospital (No. 2008-055-1147) and conducted according to the Declaration of Helsinki and its later amendments. Informed consent was not obtained because this study is retrospective analysis. EMG was performed with a monopolar needle electrode from muscles of subjects body (Viking Quest (Natus, USA)). The filter setting was set at 20 Hz (low-cut) and 10 kHz (high-cut). The results of the last 10 seconds of the EMG were recorded and used for analysis. Based on the elbow joint of the upper extremity and the knee joint of the lower extremity, the muscles close to this joint were classified as proximal muscles and distal muscles.

The results of the waveform data of patients stored numerically in the electromyography machine were extracted, and they were made into a waveform through the MATLAB software (version R2020b) program. Among the created waveform data, artifacts occurring in the cases including move of the needle electrode or patients moving among the data at the beginning and at the end were excluded, and some noise in the middle portion was preserved. The raw EMG data, which was originally sampled at 48 kHz, was downsampled to 10 kHz to reduce computational complexity and sliced in fixed window length of 0.4 seconds units and hop size of 0.1 seconds units that were likely to be the most optimal length for post-experimental analysis. After slicing, total segments were composed of 2700 segments from subjects with myopathies, 3664 segments of subjects with neuropathies, and 1706 segments of subjects without neither neuropathy nor myopathy. Consequently, rest of the numerical data was used for analysis.

**Building machine learning models**

To find the characteristics of the EMG signal, we used a 1-dimensional convolutional neural network (CNN) based on VGGNet and ResNet, which has been proven effective in image classification, named it nEMGNet. (30, 31) The structure of nEMGNet includes spatial block-1, which reduces the resolution by half, spatial block-2 which reduces the resolution by quarter, and residual block which solves the problem of poor backward propagation as the layer gets deeper by making a residual connection.(30) (Table1)

|  |  |  |
| --- | --- | --- |
| **Spatial reduction block-1**  **(n,k)** | **Spatial reduction block-2**  **(n,k)** | **Residual block(n)** |
| Conv(k)-n,  Stride(1) | Conv(k)-n,  Stride(2) | Conv(5)-n,  Stride(1) |
| BatchNorm | BatchNorm | BatchNorm |
| ReLU | ReLU | ReLU |
| Max-pool(2),  Stride(2) | Max-pool(2),  Stride(2) | Conv(5)-n,  Stride(1) |
|  |  | BatchNorm |

Table 1. convolutional blocks of nEMGNet. n, the number of channel; k, the number of filter; conv, convolutional layer; BatchNorm, batch normalization; Max-pool, max pooling.

The nEMGNet was tested with 4 versions of nEMGNet-A, nEMGNet-B, nEMGNet-C, and nEMGNet-D with different versions according to the number of residual blocks. (Table 2)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **nEMGNet-A** | **nEMGNet-B** | **nEMGNet-C** | **nEMGNet-D** |
| Block 1 | SR block-2  (64, 11) | SR block-2  (64, 11) | SR block-2  (64, 11) | SR block-2  (64, 11) |
| Block 2 | SR block-2  (64, 7) | SR block-2  (64, 7) | SR block-2  (64, 7) | SR block-2  (64, 7) |
| Block 3 | SR block-2  (64, 5) | SR block-2  (64, 5) | SR block-2  (64, 5) | SR block-2  (64, 5) |
| Block 4 |  | Residual block  (64) 2 | Residual block  (64) 4 | Residual block  (64) 6 |
| Block 5 | SR block-1  (128, 5) | SR block-1  (128, 5) | SR block-1  (128, 5) | SR block-1  (128, 5) |
| Block 6 |  | Residual block  (128) 2 | Residual block  (128) 4 | Residual block  (128) 6 |
| Block 7 | SR block-1  (256, 5) | SR block-1  (256, 5) | SR block-1  (256, 5) | SR block-1  (256, 5) |
| Block 8 |  | Residual block  (256) 2 | Residual block  (256) 4 | Residual block  (256) 6 |
| Block 9 | SR block-1  (512, 5) | SR block-1  (512, 5) | SR block-1  (512, 5) | SR block-1  (512, 5) |
| Block 10 |  | Residual block  (512) 2 | Residual block  (512) 4 | Residual block  (512) 6 |
| Block 11 | SR block-1  (1024, 5) | SR block-1  (1024, 5) | SR block-1  (1024, 5) | SR block-1  (1024, 5) |
| Block 12 |  | Residual block  (1024) 2 | Residual block  (1024) 4 | Residual block  (1024) 6 |
|  | FC-512 | FC-512 | FC-512 | FC-512 |
|  | FC-256 | FC-256 | FC-256 | FC-256 |
|  | FC-64 | FC-64 | FC-64 | FC-64 |
|  | FC-16 | FC-16 | FC-16 | FC-16 |
|  | FC-10 | FC-10 | FC-10 | FC-10 |
|  | Softmax | Softmax | Softmax | Softmax |

Table 2. configuration of 4 different versions of nEMGNet. SR block, spatial resolution block; FC, fully connected layer.

A rectified linear unit (ReLU) is applied to the fully connected layer after the convolutional layer. (Figure 1)



Figure 1. nEMGNet structure. Processing composed of convolutional neural network, batch normalization, rectified linear unit (ReLU), max pooling. Light gray box; spatial block-1 and spatial block-2, Dark gray box and bold curved arrow; residual block.

The initial values of nEMGNet hyperparameter were empirically determined based on values that have been widely used. The learning rate, batch size, and epoch were set to 10-3, 32, and 100. Adam optimizer was used for optimizer, and inversely proportional values were used for the class weight to the number of signal segments for preventing erroneous prediction.

The number of muscles tested with EMG is different for each subject, and among the tested muscles, abnormal and normal EMG can coexist. To overcome these limitations, we applied a method called the DiVote (Divide and Vote) technique. DiVote divided each EMG signal into segments of homogeneous length and converted it into 3 signal segment prediction scores through a feature extractor. The muscle signal prediction score was calculated by aggregating the signal segment prediction score and aggregated to derive the subject prediction score through soft voting. When deriving the subject prediction score, two different method were tried. The first method is to calculate and aggregate prediction scores by classifying them by neuropathy, myopathy, and normal without information on the location of the muscle. The second method is to classify whether it is proximal or distal according to the location of the muscle, and calculate and aggregate prediction scores of each neuropathy, myopathy, and normal. After 1500 training, in the learning process for extracting features of neuropathy, myopathy, and normal the learning rate was adjusted to 10-2 and gradient descent was applied.

There two major methods of aggregated probabilities corresponding to neuropathy, myopathy, and normal of the segment from EMG signal data for each muscle, hard voting and soft voting. The former is to select one item with the highest probability for each segment, count the number of selected items and finally select the one that occupies the largest number among them. On the other hand, the latter is to select items with highest probabilities after averaging each probabilities from each segments. In the case of hard voting, where only values with high probability are selected for each segment, data loss might occur, so soft voting was used to reduce data loss. The most probable diagnosis obtained through soft voting in the segment is the muscle prediction score, and the subject prediction score is the soft voting of all muscles by subject after soft voting in the segment.

Classifier는 logistic regression을 사용해서 구한 것임.

**Visualization and evaluation of the result**

The performance of nEMGNet was evaluated as closely distinguishing myopathy, neuropathy, and normal from actual results through 2 steps. First, we evaluated whether the results of classifying EMG data for each muscle match the actual diagnosis, and in the next step, whether the results of classifying each subject match the actual diagnosis. The accuracy, precision, recall, F1, area under receiver operating characteristic curve (AUROC), and Mathew’s correlation coefficient (MCC) were used as indicators for evaluation, and it was calculated using the following formula.

Accuracy=(TP+TN)(TP+TN+FP+FN)

Precision=TP(TP+FP)

Recall=TP(TP+FN)

F1=2×Precision×RecallPrecision+Recall

MCC=klm(CkkClm-CklCmk)k(lCkl)(l'k'≠kCk'l')k(lClk)(l'k'≠kCl'k')

TP, number of true positive; TN, number of true negative; FP, number of false positive; FN, number of false negative, C; confusion matrix from n-class classification result with columns of true labels and rows of predicted labels. Ckk means that actual label is a ‘k’ label and predicted result by nEMGNet is also ‘k’ label.

We used feature visualization to identify the characteristics of each classified waveform learned through nEMGNet. The accuracy of nEMGNet was calculated by cross entropy, and since the number of subjects was small, the actual diagnosis and the predicted diagnosis for each subject were compared with 5-fold cross-validation.

**Comparison the accuracy of nEMGNet with physicians’ interpretations**

In order to evaluate the applicability of nEMGNet to clinical practice, the EMG signal numerical data were transformed to waveform data, which similar to the actual test screen shown on the screen of the EMG machine, and the waveform that provided to the 6 residents of the Department of Neurology and Rehabilitation medicine who currently conduct and interpret EMG. (Suppl. Figure 1) The classified results by 6 residents were compared with results by nEMGNet. The degree of agreement between physicians and nEMGNet and accuracy were obtained.



Supplementary Figure 1. Example of provided EMG waveform data

**Statistical analysis**

Statistical analyses were performed using R statistical software (version 4.1.0; R Foundation for Statistical Computing, Vienna, Austria). The *p*-value less than 0.05 was considered statistically significant. Sensitivity, specificity, and inter-rater reliability were analyzed using the Chi-square test and McNemar test.

**Results**

The data of the subjects used for the analysis were 20 subjects with normal and 19 subjects with neuropathy with whom the diagnosis was radiculopathy, motor axonal polyneuropathy, motor neuron disease, etc., myopathy was 19 subjects with whom the diagnosis was muscular dystrophy and inflammatory myopathy. The number of EMG signal data used for analysis was 125, 161, and 97, respectively, length was 204.31 seconds, 423.12 seconds, and 204.31 seconds. (Table 3)

Table 3. Demographic characteristics of subjects

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Myopathy | Neuropathy | Normal | p-value |
| Number of Subjects | 19 | 19 | 20 |  |
| Female, n(%) | 14 (73.7) | 12 (63.2) | 13 (65) | 0.761 |
| Age (mean±SD) | 52.2±20.1 | 58.4±15.1 | 60.2±16.9 | 0.329 |
| Proportion of EMG according to location of muscle (%) |  | | | <0.001 |
| Distal muscles | 60 (48.4) | 97 (60.2) | 80 (82.5) |  |
| Proximal muscles | 64 (51.6) | 64 (39.8) | 17 (17.5) |  |
| Number of files | 125 | 161 | 97 |  |
| Total signal length (sec) | 313.54 | 423.12 | 204.31 |  |
| EMG signal data number (n (%)) | 2700 (33.5) | 3664 (45.4) | 1706 (21.1) |  |

The results of filtering the entire EMG segment through a convolutional block were plotted through a method called uniform manifold approximation and projection (UMAP). While the data had been passed through convolutional block, the dimension gradually decreased, neuropathy and myopathy were well distinguished, and the normal was located between the two groups. (Figure 2) 

Figure 2. Dimension reduction and clustering of signal segments after passing through convolutional layers of nEMGNet. (a) Initial state, state after passing through (b) 2nd block, (c) 4th block, (d) 6th block, (e) 9th block, (f) 12th block. z1 and z2, reduced dimensions; M, myopathy; N, neuropathy; NL, normal.

The classified results were depicted as a heatmap and a 3-dimensional plot. (Figure 3) The predicted result with the largest value among the muscle signal prediction scores through the DiVote pipeline is expressed in 3 different color, and the highter the probability, the darker the color. In the muscle signal prediction score, the predicted result with the largest value among the subject prediction scores that passed through the DiVote pipeline again was denoted by N for neuropathy, M for myopathy, and NL for normal. (Figure 3A) The subject prediction score was depicted as a 3-dimensional plot with the probability of being classified as myopathy, neuropthy, and normal as each axis. (Figure 3B) The classifier measured using logistic regression and argmax function was added as a decision boundary that distinguishes myopathy, neuropathy, and normal in figure 3B. Neuropathy and myopathy were distinguished relatively well, however, normal was directed toward the center and not well differentiated. (Figure 3C) The classifier measured by adding the information about location of muscles, which was divided into proximal or distal muscle, was added to figure 3B as shifted decision boundary, as a result, the normal was better distinguished compared to figure 3C. (Figure 3D) 

Figure 3. The results of subject classification through DiVote pipeline and decision boundary (A) Heatmap of the most probable diagnosis among muscle signal prediction scores. Each square box represents the most probable diagnosis value in color after aggregating the signal segment prediction scores predicted by nEMGNet. (B) The subject prediction scores as dots in a 3-dimensional plot. (C) Decision boundary calculated through simple argmax function. (D) Shifted decision boundary after adding the information on location of muscles. M, myopathy; N, neuropathy; NL, normal; P, proximal muscle; D, distal muscle.

The accuracy of total prediction over all, myopathy, neuropathy, and normal segments without processing of DiVote pipeline was 62.35%, 71.58%, 63.2%, and 52.26%. 각 subject를 classification하는데 추가정보 없이 분류했을 때에는 약 67-75%였는데 subject feature를 모두 넣었을 때는 약 76-81%로 조금 더 향상됐고 근육의 위치를 나타내는 proximal/distal에 대해서 추가로 정정보 넣었을 때 정확도가 더 향상돼서 약 76-83%의 정확도를 보였습니다. Subject feature type에 무관하게 추가 정보를 주었을 때 performance가 더 향상되는 결과를 보였고 nEMGNet별로 비교하면 SR block-2내에서 2개의 residual block을 포함하는 nEMGNet-B가 가장 좋은 performance를 보였고 이와 대조적으로 residual block을 포함하지 않는 nEMGNet-A가 가장 떨어지는 performance를 보였습니다. (Table 4)

|  |  |  |  |
| --- | --- | --- | --- |
| nEMGNet | No classifier  (argmax) | Subject features  (all) | Subject features  (proximal/distal) |
| A | 67.1710.75% | 76.064.90% | 76.5710.23% |
| B | 73.647.27% | 81.924.83% | 83.695.28% |
| C | 69.957.77% | 81.266.35% | 81.876.80% |
| D | 75.356.93% | 81.266.35% | 80.815.31% |

Table 4. Subject diagnosis accuracy of nEMGNet and classifier pipeline

분석한 모든 nEMGNet의 performance를 evaluation metrics를 기준으로 기존에 발표된 다른 연구결과와 비교한 것입니다. (Table 5)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Reference | Feature extractor | Evaluation Metrics (%) | | | | | |
| Accuracy | F1 | Precision | Recall | AUROC | MCC |
| Proposed | nEMGNet -A | 76.57 | 74.67 | 79.17 | 76.57 | 89.43 | 68.33 |
|  | nEMGNet -B | **83.69** | **83.59** | **87.96** | **83.69** | **91.45** | **77.70** |
|  | nEMGNet -C | 81.87 | 81.61 | 85.74 | 81.87 | 91.21 | 74.66 |
|  | nEMGNet -D | 80.81 | 80.56 | 86.65 | 80.81 | 90.53 | 74.16 |
| Nam et al. [28] | Inception-v4 | 57.47 | 51.38 | 56.53 | 57.47 | 78.39 | 42.03 |
| Nodera et al. [27] | ResNet50 | 73.84 | 72.59 | 81.94 | 73.84 | 81.27 | 64.76 |
|  | ResNet152 | 75.61 | 74.90 | 81.11 | 75.61 | 85.30 | 65.50 |
|  | VGG16 | 68.23 | 65.47 | 66.86 | 68.23 | 79.85 | 55.42 |
|  | VGG19 | 72.32 | 69.66 | 74.71 | 72.32 | 81.10 | 62.91 |
|  | Inception-v3 | 71.92 | 70.64 | 79.29 | 71.92 | 83.48 | 61.60 |

Table 5. Evaluation results of 4 different nEMGNets and previous reported results.

가장 좋은 performance를 보인 nEMGNet-B이 각 EMG signal segment별로 classification하는 정확도를 class로 나눠서 실제 class와 예측한 결과 class로 confusion matrix로 나타냈습니다. (Figure 4)



Figure 4 Confusion matrix of prediction by nEMGNet-B. Left; Accuracy of predicted results for each EMG waveform, Right; Accuracy of predicted results by considering all EMG for each muscle of each patient together

흥미롭게도 nEMGNet을 통해 훈련된 근전도 데이터를 토대로 neuropathy, myopathy, normal의 파형 데이터의 특징을 시각화해서 본 결과 실제 대상자의 근전도 파형과 유사했는데 myopathy는 small amplitude, short duration, early recruitment의 특징을 보였고 neuropathy는 high amplitude의 long duration, reduced recruitment의 특징을 보였습니다. (Figure 5)

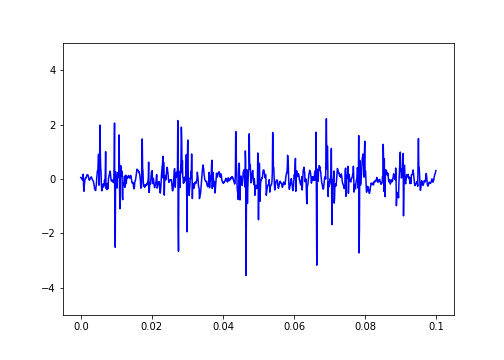
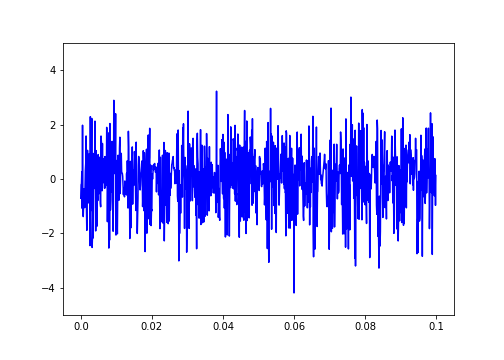
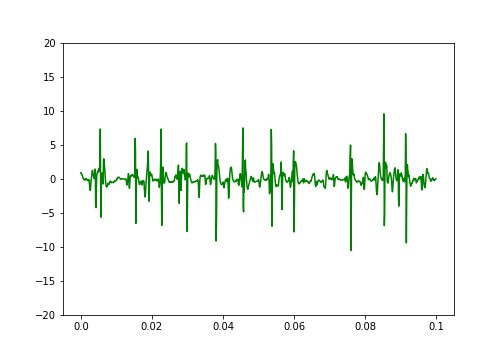
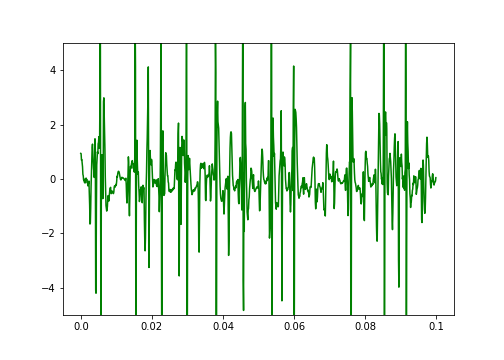
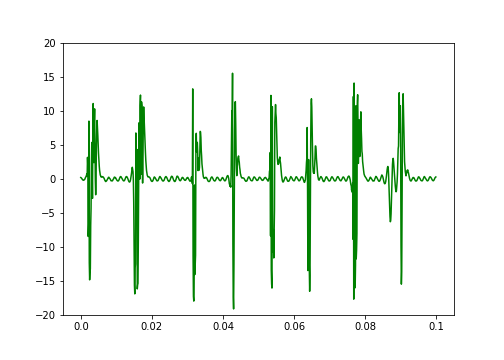
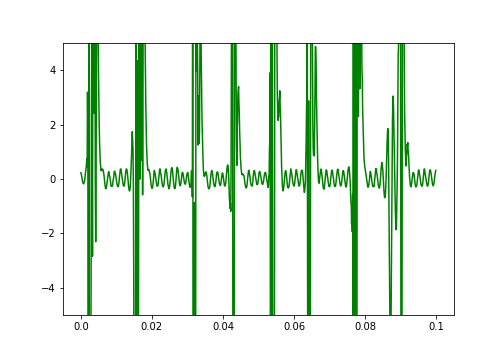
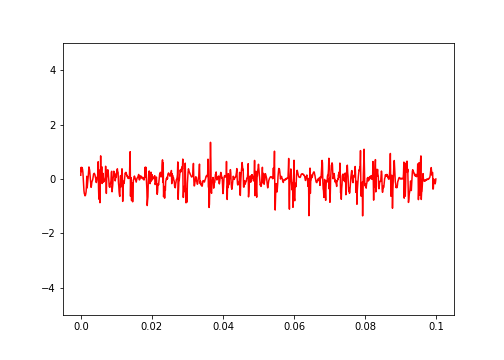
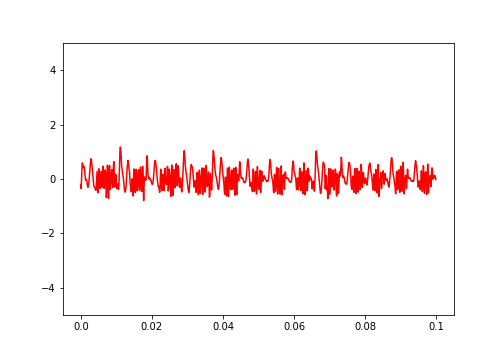


Figure 5. Learned features of nEMGNet. Top row are real signals and bottom row are generated signals through feature visualization. nEMGNet from the 1st fold was used to plot the figure. (a) Real myopathic signal. (b) Generated myopathic signal. (c) Real neuropathic signal. (d) Generated neuropathic signal. (e) Real neuropathic signal with 20mV y-axis limit. (f) Generated neuropathic signal with 20mV y-axis limit. (g) Real normal signal, (h) Generated normal signal. Note that (a), (b), (c), (d), (g), (h), is plotted with 5mV y-axis limit for better comparison between signals of different labels. (e), (f) is plotted with 20mV y-axis limit to show the overall shape of the neuropathic signal. (c)&(e) and (d)&(f) are identical signals.

nEMGNet과 phsycian의 classification 결과를 비교하기 위해 physician에게 EMG signal data를 가지고 근전도 기계와 거의 유사하게 파형을 재생하는 방식으로 보여주고 physician별로 각각 근육별로, subject별로 classification하게한 뒤 결과를 저장하고 소요된 시간을 측정했습니다. nEMGNet의 예측결과와 6명의 physician이 classification한 결과를 비교한 결과 physician들의 sensitivity는 median, IQR%, specificity는 median,IQR%였고 nEMGNet의 예측결과의 sensitivity는 ~%, specificity는 ~%였습니다. 그리고 physician간의 inter-rater reliability (Fleiss κ)는 ~였고 nEMGNet과 각 physician간의 inter-rater reliability는 ~였습니다. 그리고 classification에 소요된 시간(mean±SD)은 근육 1개당 소요된 시간은 physician이 ~초±~초였고 nEMGNet은 ~초±~초, subject 1명당 소요된 시간은 physician이 ~초±~초, nEMGNet이 ~초±~초로 nEMGNet이 훨씬 적은 시간에 classification을 완료하는 것으로 나타났습니다. (Table 6)

|  |  |  |
| --- | --- | --- |
|  | Physician | nEMGNet |
| Sensitivity (median, IQR) |  |  |
| Specificity (median, IQR) |  |  |
| Inter-rater reliability (Fleiss kappa)  Between physician  Between Physician and nEMGNet |  |  |
|  | |
| Elapsed time (sec) |  | 21.62 (GPU)  43.28 (CPU) |

Table 6. result of classification by physicians (mean result of 6 physicians) and nEMGNet. Result was shown with sensitivity, specificity, inter-rater reliability, elapsed time.

**Conclusions**

We applied deep learning to interpreting the EMG waveforms, and assess the accuracy of machine learning based-EMG interpretation and compare the results done by physicians. EMG signal의 분석만 했을 때보다 subject feature를 모두 넣었을 때 performance가 더 향상되는 결과를 보였고 subject feature를 모두 넣은 것에 비해 muscle의 위치에 따라서 proximal인지 distal인지의 추가 정보를 넣었을 때 정확도가 더욱 향상되는 결과를 보였습니다.

기존에 이미지 분석과 surface EMG, needle EMG에 적용하여 머신러닝을 적용하여 좋은 효과를 확인한 논문들이 있었습니다. (25-31) 이전 논문은 needle EMG의 volitional state의 signal에 적용한 논문이 없었는데 근전도를 이용해서 neuropathy, myopathy, normal을 감별하는데 volitional state의 근전도가 유용하다는 것은 널리 알려진 바 있습니다. (5, 6, 32) 우리는 convolutional neural network를 이용해서 EMG signal data를 분석했고 이를 nEMGNet이라는 이름을 붙였습니다. 근전도 검사 결과를 2차원 이미지 데이터로 분석한 기존의 연구와 다르게 근전도 검사 결과를 수치화 시킨 1차원 데이터를 분석해서 데이터 손실을 줄였고 각 환자별로 검사의 길이가 다른 heterogeneity를 DiVote pipeline을 통해서 subject별로 검사된 EMG signal을 조금씩 뽑아서 합친뒤 분석하는 방식으로 극복했습니다. 또한 DiVote pipeline을 통해서 subject별로 일부의 EMG는 정상의 결과를 보이고 일부의 EMG에서는 이상이 나오므로 전반적인 EMG를 모두 고려해서 진단을 하는 과정을 반영하려고 했습니다. 또한 neuropathy는 주로 distal part muscle에서 이상이 발견되고 반면에 myopathy는 주로 proximal part muscle에서 이상이 발견된다는 점을 고려해서 EMG signal data를 얻은 근육의 위치에 따라 proximal, distal로 나눈 추가 information을 같이 고려해서 머신러닝의 performance를 더욱 개선했고 DiVote pipeline에 classifier를 사용함으로써 neuropathy, myopathy, normal을 구분하는 decision boundary를 조정함으로써 performance를 더욱 개선해서 83.69%의 accuracy를 얻었습니다.

그동안의 EMG분석에 machine learning을 적용한 결과와 비교한 결과 nEMGNet의 장점은 첫째, 1-dimension의 data를 image로 변환시켜서 2차원으로 분석하지 않고 1차원으로 분석함으로써 변환과정에서 잃을 수 있는 정보의 양을 줄였다는 점입니다. 둘째, EMG를 분석할 때 각 환자마다 검사된 총 시간과 검사된 근육의 수가 다르므로 분석에 한계가 있을 수 있지만 이러한 점을 DiVote를 통해서 극복해서 서로 다른 개수의 EMG data를 분석할 수 있다는 점입니다. 셋째, 환자마다 EMG data중 어떤 근육에서는 이상이 나오고 어떤 근육에서는 이상이 나오지 않을 수 있는데 이러한 부분을 DiVote를 통해서 조금씩 데이터를 뽑아서 합친뒤 분석하는 방법으로 subject의 진단의 정확도를 높였다는 점입니다. 비록 하나하나의 근육의 EMG data가 이상이 있는 것도 정상인 것도 있어 종합적으로 생각해서 진단할 때 혼동될 수 있지만 전체적으로 고려해서 진단한다면 이러한 점을 극복해서 진단을 더 용이하게 할 수 있다는 점이 장점입니다.

The diagnostic usefulness of electromyography for identifying peripheral neuropathy and myopathy has been suggested in previous studies.

Recently, deep learning has been successfully applied to assisting diagnosis of medical diseases in so many ways. The diagnostic usefulness of electromyography for identifying peripheral neuropathy and myopathy has been suggested in previous studies.

We applied convolutional neural network to classification of volitional EMG waveforms via making numerical EMG data into waveform, editing and re-transforming into waveform data. Result of analysis showed that EMG waveform was well classified by deep learning algorithm. Our deep learning model might reduce error rate of EMG interpretation and physicians workload. This study suggest that the models built on deep learning-based EMG waveform interpretation might be complementary for physicians’ interpretation and promising.

Deep learning has shown good performance in many medical data including waveform and time series data such as electrocardiography and electroencephalography. In the analysis of waveforms of EMG, deep learning algorithm showed favorable performance compared with analyzed by physician and residents.

이번 연구는 임상정보없이 근전도만으로 분석했고 포함된 데이터 셋이 현실에서의 분포와는 다소 다른 out of distribution의 문제와 clinical information없이 판단한 pre-test probability라는 점 때문에 physician의 정확도가 다소 낮게 나온 것으로 보인다. 그럼에도 불구하고 nEMGNet은 이전에 연구된 근전도만으로 진단하는 연구보다 더 높은 정확도를 보였다는 장점이 있고 컴퓨터에서 프로그램을 로딩하고 데이터를 모두 분석해서 총 58명의 진단을 내리는데 40초밖에 소요되지 않았음. 소요시간이 길지 않아 실제 환자에게 근전도를 검사하는 과정에서 실시간으로 프로그램을 이용해서 진단에 도움을 줄 수 있을 것으로 생각되며 분석을 위해 환자의 정보가 포함된 검사결과를 온라인 데이터베이스에 업로드를 하지 않아도 되어 환자의 정보도 보호할 수 있다는 장점도 있겠다.

Until now, a few studies on analyzing EMG waveforms by deep learning have been documented.

Our study shed lights on diagnosis patient of suspected neuropathy or myopathy by nEMGNet which might help with EMG signal classification. However, our study has several limitations. First, study number is not enough to demonstrate perfect usefulness of deep learning on EMG classification. Second, we focused only on dividing EMG signal into neuropathy, myopathy, and normal. However, more specialized diagnosis could be identified with more concise machine learning algorithms. Third, we analyzed data from only one center data. Future study with much more data and multicenter data will show potential of applying machine learning to EMG interpretation.

**References**

1. Daube JR, Rubin DI. Needle electromyography. Muscle Nerve. 2009;39(2):244-70.

2. Kimura J. Electrodiagnosis in Diseases of Nerve and Muscle: Principles and Practice: Oxford University Press; 2013.

3. Mills KR. The basics of electromyography. Journal of Neurology, Neurosurgery & Psychiatry. 2005;76(suppl 2):ii32-ii5.

4. Oh SJ. Clinical Electromyography: Nerve Conduction Studies: Lippincott Williams & Wilkins; 2003.

5. Rubin DI. Needle electromyography: Basic concepts. Handb Clin Neurol. 2019;160:243-56.

6. Whittaker RG. The fundamentals of electromyography. Pract Neurol. 2012;12(3):187-94.

7. Aminoff MJ, Goodin DS, Parry GJ, Barbaro NM, Weinstein PR, Rosenblum ML. Electrophysiologic evaluation of lumbosacral radiculopathies: electromyography, late responses, and somatosensory evoked potentials. Neurology. 1985;35(10):1514-8.

8. Bromberg MB. The motor unit and quantitative electromyography. Muscle Nerve. 2020;61(2):131-42.

9. Gerardo Gutiérrez Gutiérrez CBLFNAMM. Use of Electromyography in the Diagnosis of Inflammatory Myopathies. Reumatología Clínica (English Edition). 2012;8(4):195-200.

10. Leblhuber F, Reisecker F, Boehm-Jurkovic H, Witzmann A, Deisenhammer E. Diagnostic value of different electrophysiologic tests in cervical disk prolapse. Neurology. 1988;38(12):1879-.

11. Sawada K, Horii M, Imoto D, Ozaki K, Toyama S, Saitoh E, et al. Usefulness of Electromyography to Predict Future Muscle Weakness in Clinically Unaffected Muscles of Polio Survivors. PM R. 2020;12(7):692-8.

12. Tonzola RF, Ackil AA, Shahani BT, Young RR. Usefulness of electrophysiological studies in the diagnosis of lumbosacral root disease. Ann Neurol. 1981;9(3):305-8.

13. Haig AJ, Tong HC, Yamakawa KS, Quint DJ, Hoff JT, Chiodo A, et al. The sensitivity and specificity of electrodiagnostic testing for the clinical syndrome of lumbar spinal stenosis. Spine (Phila Pa 1976). 2005;30(23):2667-76.

14. Kendall R, Werner RA. Interrater reliability of the needle examination in lumbosacral radiculopathy. Muscle Nerve. 2006;34(2):238-41.

15. Nirkko AC, Rösler KM, Hess CW. Sensitivity and specificity of needle electromyography: a prospective study comparing automated interference pattern analysis with single motor unit potential analysis. Electroencephalogr Clin Neurophysiol. 1995;97(1):1-10.

16. Arthur KC, Calvo A, Price TR, Geiger JT, Chiò A, Traynor BJ. Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. Nature Communications. 2016;7(1):12408.

17. Longinetti E, Fang F. Epidemiology of amyotrophic lateral sclerosis: an update of recent literature. Curr Opin Neurol. 2019;32(5):771-6.

18. Parker MJS, Oldroyd A, Roberts ME, Ollier WE, New RP, Cooper RG, et al. Increasing incidence of adult idiopathic inflammatory myopathies in the City of Salford, UK: a 10-year epidemiological study. Rheumatol Adv Pract. 2018;2(2).

19. Rose L, McKim D, Leasa D, Nonoyama M, Tandon A, Bai YQ, et al. Trends in incidence, prevalence, and mortality of neuromuscular disease in Ontario, Canada: A population-based retrospective cohort study (2003-2014). PLoS One. 2019;14(3):e0210574.

20. Alfaras M, Soriano MC, Ortín S. A Fast Machine Learning Model for ECG-Based Heartbeat Classification and Arrhythmia Detection. Frontiers in Physics. 2019;7(103).

21. Lu X, Wu Y, Yan R, Cao S, Wang K, Mou S, et al., editors. Pulse waveform analysis for pregnancy diagnosis based on machine learning. 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC); 2018 12-14 Oct. 2018.

22. Gemein LAW, Schirrmeister RT, Chrabąszcz P, Wilson D, Boedecker J, Schulze-Bonhage A, et al. Machine-learning-based diagnostics of EEG pathology. Neuroimage. 2020;220:117021.

23. Roy Y, Banville H, Albuquerque I, Gramfort A, Falk TH, Faubert J. Deep learning-based electroencephalography analysis: a systematic review. J Neural Eng. 2019;16(5):051001.

24. Ribeiro AH, Ribeiro MH, Paixão GMM, Oliveira DM, Gomes PR, Canazart JA, et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. Nature Communications. 2020;11(1):1760.

25. Akef Khowailed I, Abotabl A. Neural muscle activation detection: A deep learning approach using surface electromyography. J Biomech. 2019;95:109322.

26. Atzori M, Cognolato M, Müller H. Deep Learning with Convolutional Neural Networks Applied to Electromyography Data: A Resource for the Classification of Movements for Prosthetic Hands. Front Neurorobot. 2016;10:9-.

27. Nam S, Sohn MK, Kim HA, Kong H-J, Jung I-Y. Development of Artificial Intelligence to Support Needle Electromyography Diagnostic Analysis. Healthc Inform Res. 2019;25(2):131-8.

28. Nodera H, Osaki Y, Yamazaki H, Mori A, Izumi Y, Kaji R. Deep learning for waveform identification of resting needle electromyography signals. Clin Neurophysiol. 2019;130(5):617-23.

29. Wei W, Dai Q, Wong Y, Hu Y, Kankanhalli M, Geng W. Surface-Electromyography-Based Gesture Recognition by Multi-View Deep Learning. IEEE Trans Biomed Eng. 2019;66(10):2964-73.

30. He K, Zhang X, Ren S, Sun J, editors. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 27-30 June 2016.

31. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 14091556. 2014.

32. Preston DC, Shapiro BE. Electromyography and Neuromuscular Disorders: Clinical-Electrophysiologic-Ultrasound Correlations: Elsevier; 2020.