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

**Introduction**

Needle electromyography (EMG) is obtained electrical signals from nerves, muscles and neuromuscular junctions, and is used to identifying any abnormalities in neuromuscular system, helping to estimate the location and extent of lesions. In case of patients with disorders of nerves, muscles, EMG shows waveforms made by damaged muscles or nerves including abnormal spontaneous potentials or action potentials generated by damaged muscles associated with nerve or muscle disorder. The diagnostic usefulness of electromyography for identifying peripheral neuropathy and myopathy has been suggested in previous studies. (1-7)

Previous studies have been reported that the accuracy of electromyography in patients with radiculopathy, the sensitivity was 36-64%, specificity 54-58%, inter-rater reliability was about 60-90% which is different according to type of abnormality.(1, 4, 6, 8-14) To recognize abnormalities of EMG waveform accurately, considerable time and efforts are needed. For diagnosing neuropathy and myopathy, higher accuracy of electromyography is required, and objective interpretation is required by decreasing the discrepancy between raters. As the importance and frequency of EMG, the workload of physicians and time taken for EMG get increased. This hardship might be alleviated by automated algorithms such as deep learning algorithms which assist physicians’ interpretations.

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. Convolutional neural network, one kind of deep learning techniques, has applied to analyzing time series data and waveform data such as electrocardiography, electroencephalography. 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.(15) To our knowledge, few studies on analyzing volitional EMG data have rarely been reported.

In order to increase the accuracy of electromyography, we applied deep learning to interpreting the EMG waveforms. 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 59 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 monopolar needle electrode from muscles of upper extremity or lower extremity (Viking Quest (Natus, USA). The filter setting was set at 20 Hz (low-cut) and 10 kHz (high-cut). During the test, 10 s of signals were averagely recorded.

**Training and validation**

Raw nEMG signals were used as input without any transformations. To reduce computational complexity the signal was downsampled to 10kHz. Downsampled frequency was chosen after inspecting the downsampled signals under different frequencies. Since CNN followed by DNN must receive an input of fixed length, each waveform is sliced into multiple frames by slicing the waveform with fixed window length $T$ and hop size $d$. $T=0.4s$, $d=0.1s$ was chosen empirically.

The objective of the task is to classify the input signals into 3 different classes.

, or distinguish the input signals

Raw nEMG signals.

Each patient contains multiple waveforms which contains

For example, a Neuropathy patient may have recordings from normal muscles as well as pathogenic muscles.

Thus distinct features for classifying the waveforms are mixed within class labels, and the model may learn

This does not mean there are no dinstinct features among the subjects, as the majority of the features each subject contains would be the designsted diagnosis label just like how medical doctors classify each subject.

Thus we expected the model to learn

(soft voting)

After each waveform frames are converted into diagnosis predictions of size 3, we must aggregate the prediction results for each frame into a single diagnosis prediction for the subject. The prediction aggregation was conducted in 2 stages. First, the prediction results were averaged per wavefile, which results in a vector of size (3,) per wavefile, equivalent to having single prediction for each muscle recording available. Each frames sliced from the wavefile were assumed to have equal contribution to the prediction result.

Next, after the prediction vectors for each muscle were acquired, the prediction vectors were aggregated by averaging the prediction results for the muscle. The muscle aggregation was performed in 2 methods, which we compared the results. The first method was to average predictioin results across all available muscle types, resulting in subject prediction feature vector of size (3,). The second method was to average prediction results across each musle subtype, namely Proximal (P), and Distal (D) muscles, resulting in subject prediction feature vector of size (6,).

As each subject contains mixed features with majority being the featues from their class, the subject feature vectors had mixed festures in after feature extraction. However since majority of the features each subject contains determines the class each subject is assigned to, we apply a classifier after feature extraction in order to move the decision boundary in feature space.

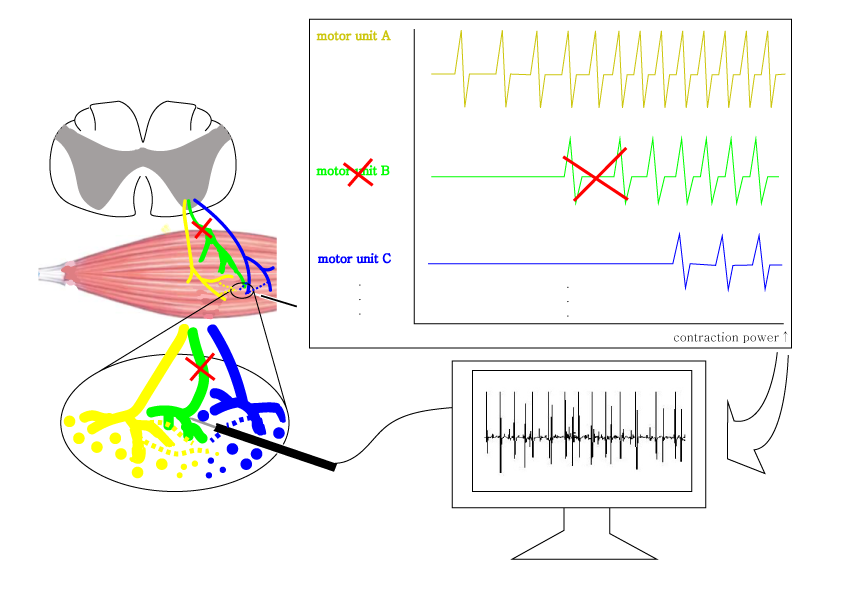
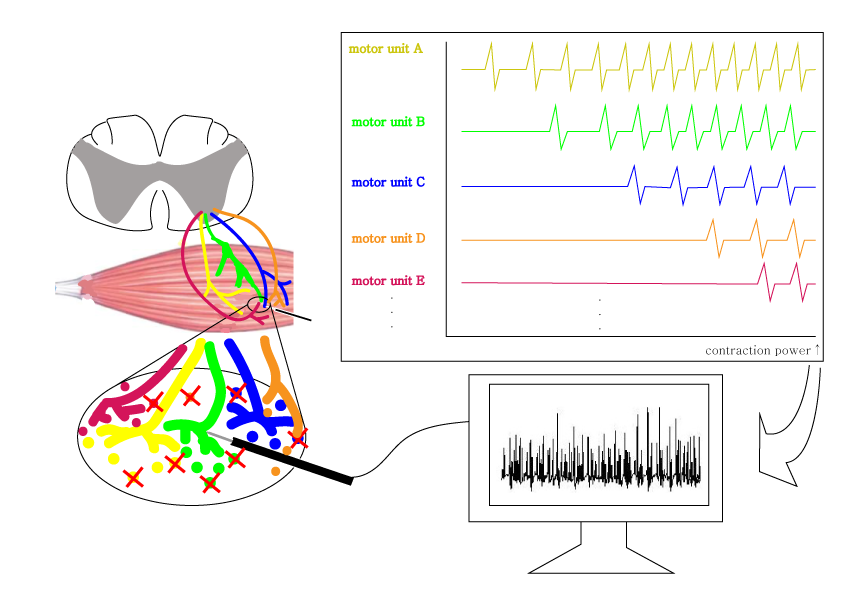
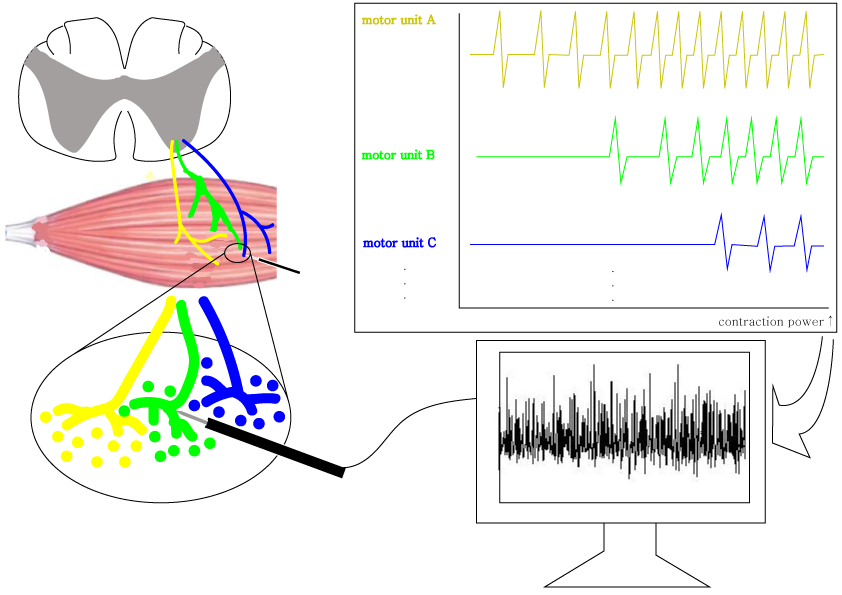
Normal subjects showed ambiguous festures whereas Myopathic and Neuropathic subjects were easily distinguishable. We believe this is because Neuropathic and Myopathic subjects contain their own dinstinct features as well as normal waveforms whereas normal subjects do not have a dinstinct waveform. Thus the model would predict normal waveforms to be ambiguous since normal waveforms can be found in all 3 types of patients, and there’s no other dinstinct feature to distinguish normal subjects from the others.

상지는 elbow joint를 기준으로 하지는 knee joint를 기준으로 이보다 가까운 쪽 근육은 proximal, 먼쪽 근육은 distal로 구분하여 분석을 진행함. 머신러닝은 convolutional neural network를 이용하여 근전도 파형을 분석하였음. 48Hz의 frequency로 분석하였고 0.2초 정도의 시간간격으로 데이터를 분할하여 학습시키고 validation및 적용을 하였음. 이후 환자의 최종임상 진단 중 myopathy, neuropathy, normal어디에 해당하는지의 결과와 머신러닝을 이용해 분석한 결과를 비교하여 정확도를 구했음 (Figure 1)

**Deep learning architecture**

****

이런식으로 그림 만들어서.



**Statistical analysis**

**Results**

신경병은 radiculopathy, axonal neuropathy, motor neuron disease 의 환자를 대상으로 하였고 근육병은 muscular dystrophy, inflammatory myopathy 등의 환자를 대상으로 하였음. (Table 1)

Table 1. Demographic characteristics

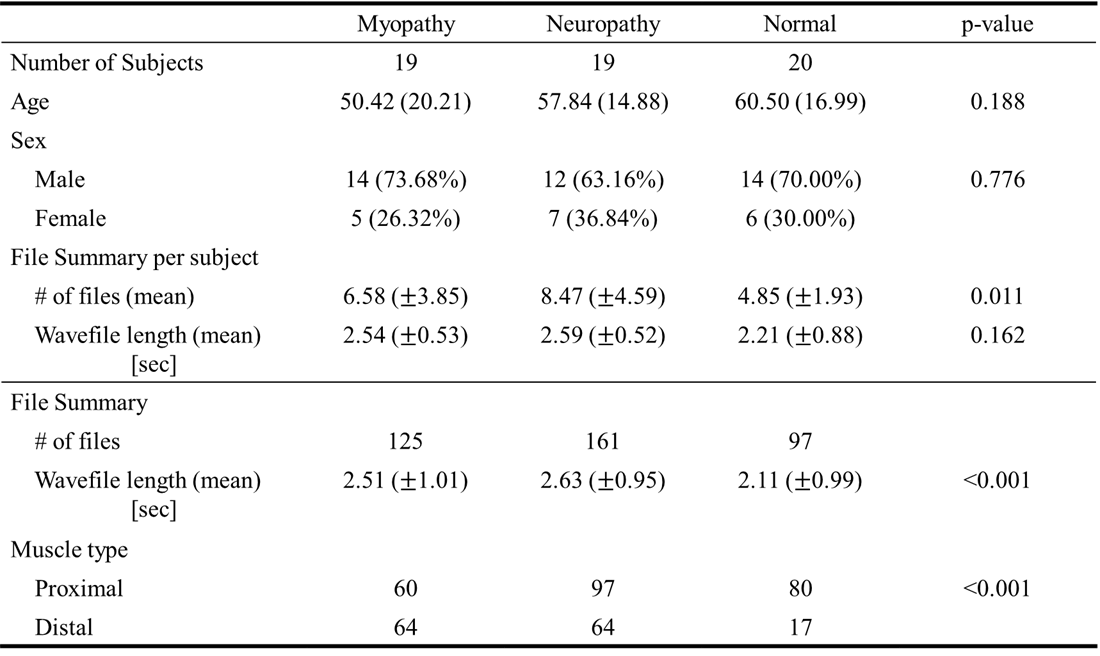


Figure 3. Performance of Feature Extractor + Classifier

Table 2. Accuracy of deep learning

Table 3. Confusion matrix

Figure 4. Extracted features of waveforms by using the convolutional neural network

Figure 5. Learned features of waveforms on feature extractor by using the convolutional neural network

Figure 6. EMG waveform classifier overview

근전도 데이터를 머신러닝을 이용해서 분석한 결과 60%후반에서 80%초반의 정확도를 보였음. 정상과 근육병을 감별하는 결과는 ~%, 근육병과 신경병을 구분하는 결과는 ~%, 정상과 신경병을 구분하는 결과는 ~%의 결과를 보였음.

**Conclusions**

Recently, deep learning has been successfully applied to assisting diagnosis of medical diseases in so many ways.

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.

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

In most cases, the accuracy and reliability of EMG interpretation by physician has been reported about 36-64% of the sensitivity, 54-58% of the specificity, 60-90% of inter-rater reliability. Interpreting the EMG might be difficult task for experienced physicians. From 6 different residents, the sensitivity was ~%, the specificity was ~%, and inter-rater concordance was ~%. With applying the deep learning algorithms, those were ~%, ~%, and ~%.

In this study, we focused only on the differentiation between neuropathy or myopathy in general. However, myopathy includes different types of diseases such as inflammatory myositis, myotonic dystrophy, muscular dystrophies, congenital myopathies, or others. Although they share common myopathic MUAP features, they can exhibit characteristic MUAP traits which can differentiate a specific diagnosis. Likewise, neuropathy includes entrapment neuropathy, Guillain-Barre syndrome, radiculopathies, or Lou Gehrig's disease whose electromyographic findings are different. If a larger dataset including various specific conditions is available in the future, ML algorithms may be able to determine a detailed diagnosis based on EMG patterns.

**References**

1. 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.

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

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

4. 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-.

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

6. 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.

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

8. Haldeman S, Shouka M, Robboy S. Computed tomography, electrodiagnostic and clinical findings in chronic workers' compensation patients with back and leg pain. Spine (Phila Pa 1976). 1988;13(3):345-50.

9. Khatri BO, Baruah J, McQuillen MP. Correlation of electromyography with computed tomography in evaluation of lower back pain. Arch Neurol. 1984;41(6):594-7.

10. Kuruoglu R, Oh SJ, Thompson B. Clinical and electromyographic correlations of lumbosacral radiculopathy. Muscle Nerve. 1994;17(2):250-1.

11. Nardin RA, Patel MR, Gudas TF, Rutkove SB, Raynor EM. Electromyography and magnetic resonance imaging in the evaluation of radiculopathy. Muscle Nerve. 1999;22(2):151-5.

12. 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.

13. Wu ZA, Tsai CP, Yang DA, Chu FL, Chang T. Electrophysiologic study and computerized tomography in diagnosis of lumbosacral radiculopathy. Zhonghua Yi Xue Za Zhi. 1987;39(2):119-25.

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

15. 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. Nat Commun. 2020;11(1):1760.

Figure 1. overview of training and inference process

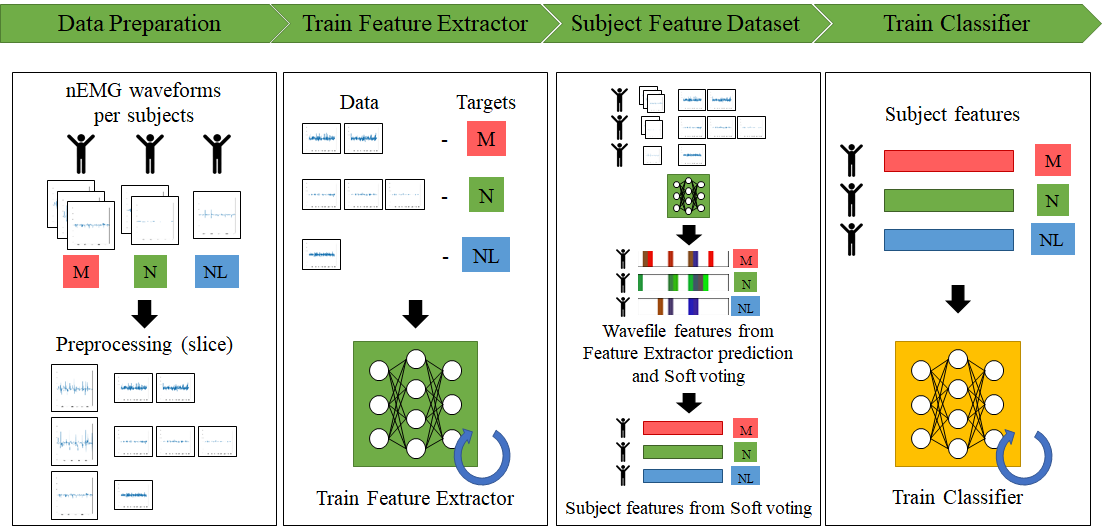


Figure 2. Convolutional neural network model architecture

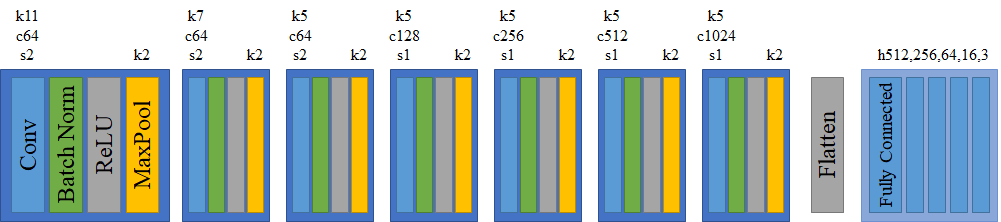


Figure 3. Performance of Feature Extractor + Classifier

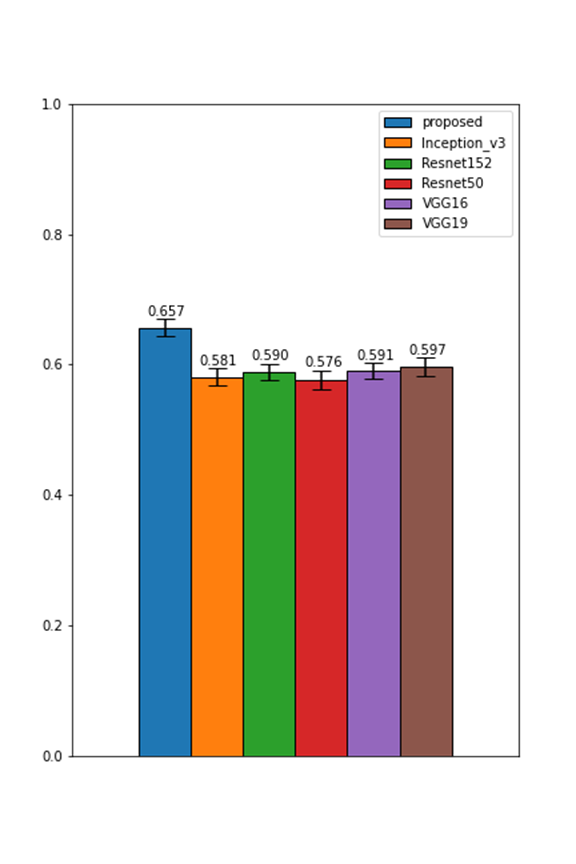


Figure 4. Extracted features of waveforms by using the convolutional neural network

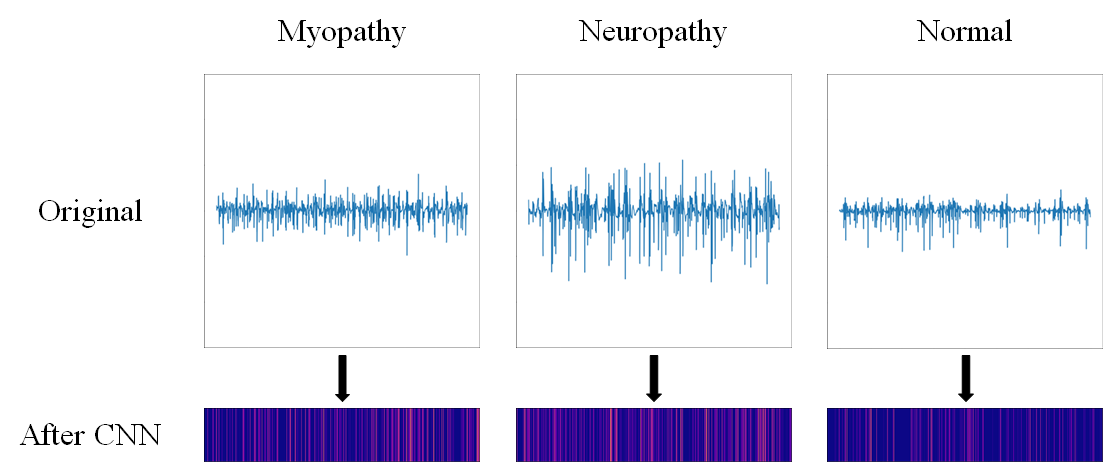


Figure 5. Learned features of waveforms on feature extractor by using the convolutional neural network

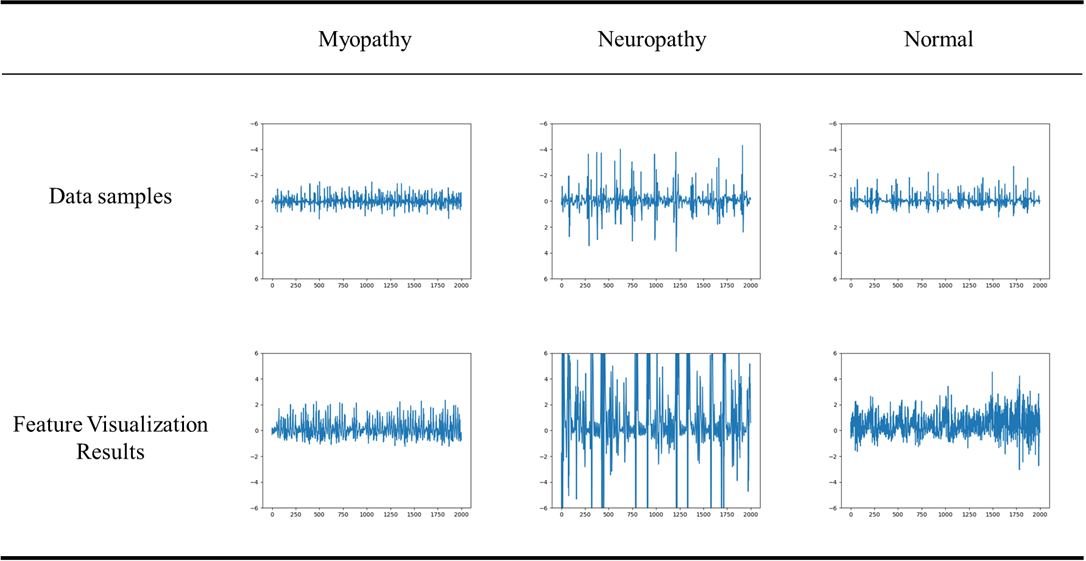


Figure 6. EMG waveform classifier overview

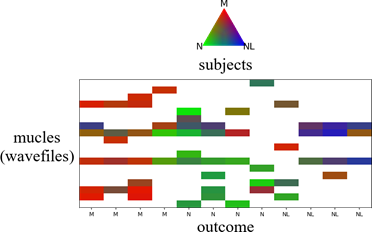


Table 1. Demographic characteristics

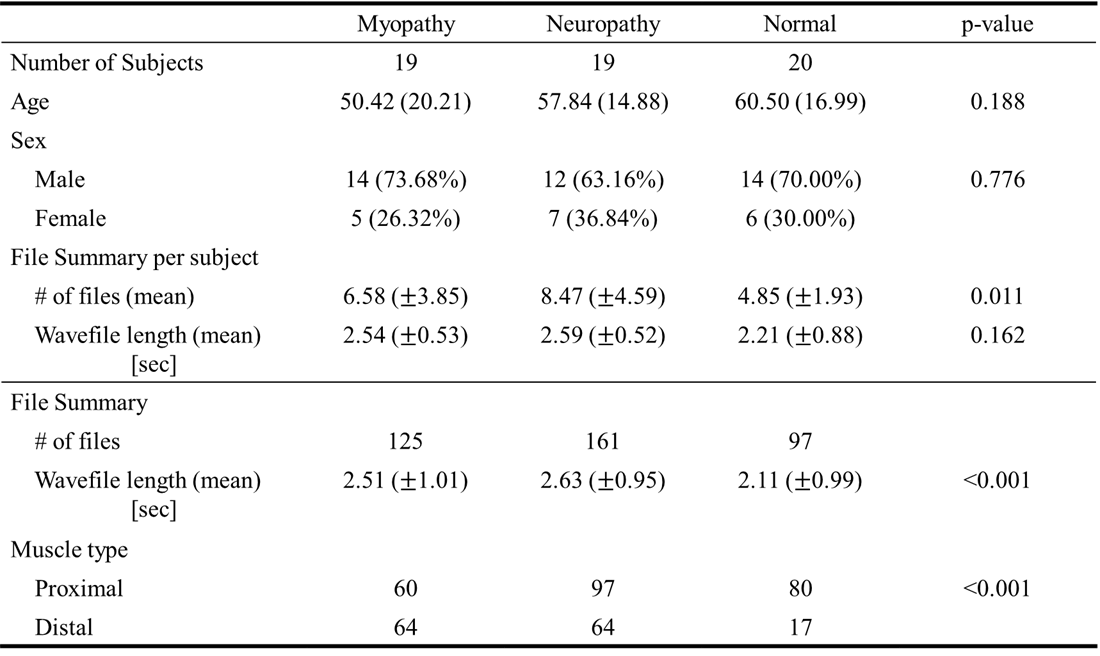


Table 2. Accuracy of deep learning

Table 3. Confusion matrix