**Study on the usefulness of deep learning in diagnosing with electromyography**

**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%. As a result of applying the DiVote pipeline, the accuracy improved to 76-81%, and the accuracy was further improved to 76-83% when the location of muscle was also considered. As a result of comparing the accuracy of 4 versions with different number of residual blocks among nEMGNets, nEMGNet-B, which included 2 residual blocks between spatial blocks, showed the best accuracy, while nEMGNet-A, which included no residual block, showed the poorest accuracy. (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. The accuracy of subject classification according to nEMGNet version with and without DiVote (Divide and Vote) pipeline processing and additional information of muscle location

The accuracy of 4 versions of nEMGNet showed better than accuracy of results with reported algorithms from previous studies. (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. [27] | Inception-v4 | 57.47 | 51.38 | 56.53 | 57.47 | 78.39 | 42.03 |
| Nodera et al. [28] | 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.

The accuracy of nEMGNet-B is depicted as confusion matrices by muscle prediction results and subject prediction results for each diagnosis. (Figure 4)



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

Based on the results of training through nEMGNet, the characteristics of the waveform of myopathy, neuropathy, and normal were similar to characteristics of the actual EMG waveform. Waveform of myopathy showed small amplitude and short duration and neuropathy showed high amplitude and long duration. (Figure 5)



Figure 5. Trained waveforms of myopathy, neuropathy, and normal based on nEMGNet. Left columns, waveform based on learned features by nEMGNet; Right colomn, actual waveform; First row, myopathy; second row, neuropathy; third row, normal. Note that (A), (B), (E), (F) were plotted with 5mV y-axis limit and (C), (D) was plotted with 20mV y-axis limit to show the overall shape of the neuropathy EMG signal of high amplitude.

The prediction results of nEMGNet and the results of classification by 6 physicians were compared. The sensitivity and specificity of the former were ~, ~ and those of the latter were ~, ~ by median and interquartile range. The inter-rater reliability between physicians and the inter-rater reliability between physicians and nEMGNet were ~, ~ (Fleiss kappa). The time taken for classification was ~ second, ~ second for 6 physicians and nEMGNet per muscle and ~ second, ~second for 1 subject. (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. The results of classification by physicians (mean result of 6 physicians) and nEMGNet. Result was shown with sensitivity, specificity, inter-rater reliability, elapsed time.

**Conclusions**

The aim of this study was to evaluate the accuracy of detecting the presence of peripheral neuropathy or myopathy by analyzing EMG waveform data using machine learning, and to confirm its applicability in clinical practice. For that purpose, we applied deep learning named nEMGNet to interpreting the EMG waveforms, and assess the performance and compare the accuracy with classification by 6 physicians. As a result of analysis with nEMGNet, the time required was shorter, and the accuracy was superior to accuracy of the physician’s analysis.

When EMG data is classified as machine learning, there are some limitations in that the number of muscles tested for each subject and time length of EMG data for each muscle is different, and the some EMG data of all muscles showed abnormalities. DiVote pipeline was used to overcome these limitations and contributed to the improved accuracy. Additionally, considering that peripheral neuropathy mainly shows abnormalities in the distal part muscle, whereas, myopathy mainly shows abnormalities in the proximal part muscles, addtional information on muscle location, which means whether the muscles are located close to the trunk or not, was added to the nEMGNet and this contributed to improved accuracy.

Previously, there have been reports that machine learning showed good performance when applied to image analysis, surface EMG, and needle EMG. [25-31] Previous studies that analyzed EMG data as 2 dimensional data using machine learning were studies to analyze gestures using surface EMG or signals during resting state using needle EMG. [25-29] For the purpose of diagnosing neuromuscular disorders, needle EMG is useful rather than surface EMG, and not only the signal during resting state but also the signal of during volitional state should be considered among needle EMG signal. [1-6, 8, 32, 33]. After minimal noise were removed at the beginning and the end of the EMG data, the rest EMG data as the 1-dimensional numerical data during the volitional state for minimizing the data loss that may occur while using the 2-dimensional data as in the previous studies. To confirm the clinical applicability of nEMGNet, the diagnostic accuracy of physicians was measured and compared with that of nEMGNet. Finally, we found that the accuracy and time-taken of diagnosing neuropathy, myopathy, and normal were 83.69% and 40 seconds in using only EMG data by nEMGNet, which is better than that of the machine learning model found in previous studies or physicians.

Interestingly, the diagnostic accuracy of physicians was ~% lower than expected, and the reason for this is thought to be mainly due to 2 reasons. First, in the data used in this study, the proportion of peripheral neuropathy and myopathy is out of distribution, which is much higher than the prevalence in population. Second, It is thought that the pre-test probability of diagnosing only with the EMG data without clinical information such as the patient’s age and symptoms, as in clinical practice, may have worked. 머신러닝 분석과정에서 neuropathy, myopathy를 고려할 때 검사된 muscle이 proximal muscle인지 distal muscle인지를 같이 고려할 때 머신러닝도 neuropathy의 경우에는 전자의 weight는 ~, 후자의 weight는 ~로 후자를 더 많이 고려하고 myopathy의 경우에는 전자의 weight는 ~, 후자의 weight는 ~로 전자를 더 많이 고려했는데 이 점은 실제 neuropathy에서는 distal muscle에 주로 이상이 발생하고 myopathy에서는 proximal muscle에 주로 이상이 발생한다는 점과 어느정도 맥락을 같이 하는 것으로 생각됩니다.

Our results suggest that machine learning has the possibilities to be embedded in EMG machines, reducing errors in EMG interpretation and the workload of physicians, and potentially preventing personal medial information leakage that can arise when EMG data is uploaded online for EMG analysis. There have been some limitations.

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. Epub 2009/01/16. doi: 10.1002/mus.21180. PubMed PMID: 19145648.

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 &amp; Psychiatry. 2005;76(suppl 2):ii32-ii5. doi: 10.1136/jnnp.2005.069211.

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. Epub 2019/07/07. doi: 10.1016/b978-0-444-64032-1.00016-3. PubMed PMID: 31277852.

6. Whittaker RG. The fundamentals of electromyography. Pract Neurol. 2012;12(3):187-94. Epub 2012/06/05. doi: 10.1136/practneurol-2011-000198. PubMed PMID: 22661353.

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. Epub 1985/10/01. doi: 10.1212/wnl.35.10.1514. PubMed PMID: 2993952.

8. Bromberg MB. The motor unit and quantitative electromyography. Muscle Nerve. 2020;61(2):131-42. Epub 2019/10/04. doi: 10.1002/mus.26718. PubMed PMID: 31579956.

9. Gerardo Gutiérrez Gutiérrez CBLFNAMM. Use of Electromyography in the Diagnosis of Inflammatory Myopathies. Reumatología Clínica (English Edition). 2012;8(4):195-200. doi: 10.1016/j.reumae.2011.10.004.

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-. doi: 10.1212/wnl.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. Epub 2019/11/09. doi: 10.1002/pmrj.12281. PubMed PMID: 31702870.

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. Epub 1981/03/01. doi: 10.1002/ana.410090317. PubMed PMID: 6261675.

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. Epub 2005/12/02. doi: 10.1097/01.brs.0000188400.11490.5f. PubMed PMID: 16319753.

14. Kendall R, Werner RA. Interrater reliability of the needle examination in lumbosacral radiculopathy. Muscle Nerve. 2006;34(2):238-41. Epub 2006/04/13. doi: 10.1002/mus.20554. PubMed PMID: 16609977.

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. Epub 1995/02/01. doi: 10.1016/0924-980x(94)00248-6. PubMed PMID: 7533715.

16. Arthur KC, Calvo A, Price TR, Geiger JT, Chiò A, Traynor BJ. Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. Nat Commun. 2016;7(1):12408. doi: 10.1038/ncomms12408.

17. Longinetti E, Fang F. Epidemiology of amyotrophic lateral sclerosis: an update of recent literature. Curr Opin Neurol. 2019;32(5):771-6. doi: 10.1097/WCO.0000000000000730. PubMed PMID: 31361627.

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). doi: 10.1093/rap/rky035.

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. doi: 10.1371/journal.pone.0210574.

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). doi: 10.3389/fphy.2019.00103.

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. doi: <https://doi.org/10.1016/j.neuroimage.2020.117021>.

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. doi: 10.1088/1741-2552/ab260c.

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. Nat Commun. 2020;11(1):1760. doi: 10.1038/s41467-020-15432-4.

25. Akef Khowailed I, Abotabl A. Neural muscle activation detection: A deep learning approach using surface electromyography. J Biomech. 2019;95:109322. doi: <https://doi.org/10.1016/j.jbiomech.2019.109322>.

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-. doi: 10.3389/fnbot.2016.00009. PubMed PMID: 27656140.

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. Epub 2019/04/30. doi: 10.4258/hir.2019.25.2.131. PubMed PMID: 31131148.

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. Epub 20190223. doi: 10.1016/j.clinph.2019.01.024. PubMed PMID: 30870796.

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. doi: 10.1109/TBME.2019.2899222.

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. Haig AJ, Gelblum JB, Rechtien JJ, Gitter AJ. Technology assessment: the use of surface EMG in the diagnosis and treatment of nerve and muscle disorders. Muscle Nerve. 1996;19(3):392-5. doi: 10.1002/(sici)1097-4598(199603)19:3<392::Aid-mus21>3.0.Co;2-t. PubMed PMID: 8606710.

33. Meekins GD, So Y, Quan D. American Association of Neuromuscular & Electrodiagnostic Medicine evidenced-based review: use of surface electromyography in the diagnosis and study of neuromuscular disorders. Muscle Nerve. 2008;38(4):1219-24. doi: 10.1002/mus.21055. PubMed PMID: 18816611.