

# Explainable Deep Learning Model for EMG-Based Finger Angle Estimation Using Attention

Hyunin Lee<sup>id</sup>, Dongwook Kim<sup>id</sup>, and Yong-Lae Park<sup>id</sup>, *Member, IEEE*

**Abstract**—Electromyography (EMG) is one of the most common methods to detect muscle activities and intentions. However, it has been difficult to estimate accurate hand motions represented by the finger joint angles using EMG signals. We propose an encoder-decoder network with an attention mechanism, an explainable deep learning model that estimates 14 finger joint angles from forearm EMG signals. This study demonstrates that the model trained by the single-finger motion data can be generalized to estimate complex motions of random fingers. The color map result of the after-training attention matrix shows that the proposed attention algorithm enables the model to learn the nonlinear relationship between the EMG signals and the finger joint angles, which is explainable. The highly activated entries in the color map of the attention matrix derived from model training are consistent with the experimental observations in which certain EMG sensors are highly activated when a particular finger moves. In summary, this study proposes an explainable deep learning model that estimates finger joint angles based on EMG signals of the forearm using the attention mechanism.

**Index Terms**—Surface electromyography (sEMG), finger angle estimation, deep learning, attention, explainable AI.

## I. INTRODUCTION

ROBOTIC prostheses that enable the disabled or patients with injuries to move their limbs freely are an important goal in biomedical engineering [1]–[3]. Most currently available robotic prostheses are controlled using surface electromyography (sEMG) signals capable of detecting muscle

activities and intentions [4]. Decoding EMG signals into finger or arm motions is one of the essential elements in the development of robotic prostheses, and there have been many studies in this area [5]–[9]. It has been shown that EMG signals can be used to classify hand gestures with an estimation accuracy higher than 95% [10]–[12]. Ramien *et al.* proposed a method of decoding EMG signals to classify individual and combined finger motions using the Bayesian data fusion approach that yielded an accuracy of 90% [13]. Recent advancement of machine learning made possible to extract more complex EMG features, allowing for successful classification of more diverse hand gestures [14]–[16]. However, the current machine learning models have a limitation in predicting hand gestures that were not learned before. To address this issue, there have been studies that predict finger joint angles directly from EMG signals rather than classifying discrete hand gestures [17]–[21]. Afshar and Matusoka showed that the joint angles of the index finger could be estimated from the EMG signals of seven forearm muscles that contributed to flexion of the index finger [17]. Shirao *et al.* also tried to decode the EMG signals of the forearm to predict index finger joint angles [18]. Further studies on predicting angles of multiple fingers have been conducted [19]–[21]. Smith *et al.* succeeded in predicting the metacarpophalangeal (MCP) joint angles of the five fingers using a simple artificial neural network [19]. Hioki *et al.* decoded muscle signals using only four EMG sensors and predicted proximal interphalangeal (PIP) joint angles of all five fingers [20]. Ngeo *et al.* predicted 15 different finger joint angles (three joint angles in each finger) using a Gaussian process (GP) method from eight sEMG sensors [21].

However, these methods still have two major limitations. First, the machine learning algorithms do not provide an explainable artificial intelligence (AI) model, a mechanism to understand and interpret the predictions made by learning, making it difficult to trust the accuracy in general cases. Although predictive modeling is motivated to improve the model performance, a lack of explainability means that the model is not sufficient to make qualitative evaluation for the task for which it was intended. Furthermore, AI models without explainability in robotic systems cannot provide insight into how the results will be utilized to improve the users' mobility [22], [23]. Second, most of the previous machine learning studies for the EMG signal analysis presented the performances based on the discrete hand gestures

Manuscript received 27 December 2021; revised 11 April 2022 and 24 May 2022; accepted 22 June 2022. Date of current version 15 July 2022. This work was supported in part by the National Research Foundation funded by the Korean Government (MSIT) under Grant NRF-2016R1A5A1938472 and Grant NRF-2017H1D8A1030270; and in part by the Korea Medical Device Development Program under Project 9991007365 (Grant KMDF\_PR\_20200901\_0161) funded by the Korean Government. (Corresponding author: Yong-Lae Park.)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institution: Seoul National University, Seoul, South Korea, under Approval No.: 2106/004-019, Dated: June 10, 2021.

Hyunin Lee is with the Department of Mechanical Engineering, Seoul National University, Seoul 08826, Republic of Korea (e-mail: gksdid5@snu.ac.kr).

Dongwook Kim and Yong-Lae Park are with the Department of Mechanical Engineering, Institute of Advanced Machines and Design, Institute of Engineering Research, Seoul National University, Seoul 08826, Republic of Korea (e-mail: ylpark@snu.ac.kr).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TNSRE.2022.3188275>, provided by the authors. Digital Object Identifier 10.1109/TNSRE.2022.3188275

that were already used for training the model, which are not sufficient for relatively dexterous manipulation motions of robotic prostheses. Therefore, it is necessary not only to build a machine learning model that estimates finger joint angles in addition to the hand gestures that the model has not learned before, but also to build an explainable model.

In this study, we propose an explainable AI model that estimates the angles of finger joints by decoding the EMG signals using the attention mechanism [24]. In particular, we show that the inner attention mechanism explains that the proposed AI model learns the complex and nonlinear relationship between the EMG signals and the finger joint angles. This study has two major contributions that have not been addressed before to the best of our knowledge:

- Models are trained with simple data sets of individual finger movements but can predict more complex data sets of random fingers with a relatively high accuracy.
- The after-trained attention matrix demonstrates that the model can learn the relationship between the EMG signals and the finger joint angles, supporting the reason why the proposed model yields a higher estimation accuracy than the models in previous studies.

## II. MATERIALS AND METHODS

### A. Experimental Setup

Eight healthy participants (two females and six males, aged  $24.75 \pm 1.75$  years) who consented in advance to the experimental protocol participated in the experiment. The participants had no history of injuries or surgeries and did not feel any abnormalities when moving their fingers, hands, and wrists freely. The experiment was approved by the Institutional Review Board of Seoul National University (IRB No. 2106/004-019) and all the experiments were conducted in accordance with the approved protocol. The participants were asked to move their fingers without moving their wrists. Their forearm EMG signals were recorded using four wireless multi-channel EMG sensors (Trigno Avanti, Delsys, USA), and the movements of the hands and the fingers were recorded with a camera at 240 frames per second (fps) at the same time as references of the finger joint angles. Each participant was asked to perform two different tasks: moving one finger at a time and moving multiple random fingers simultaneously. In the first task (Task 1), each participant flexed and extended all five fingers individually in the order of thumb, index, middle, ring, and little fingers one by one, which is called “th2pi” in this paper. During the experiments, the participants were asked to maintain a constant speed (one second for flexion and one second for extension). The “th2pi” sequence takes 10 seconds and Task 1 consists of 20 sets of “th2pi” sequences, taking a total of 200 seconds. In the second task (Task 2), each participant was allowed to flex and extend one or more fingers at the same finger flexion-extension speed in Task 1. Each consists of five randomized flexion-extension movements and takes 10 seconds. Task 2 also contains 20 sets of random finger motions, taking 200 seconds. An additional experiment was conducted in which the participants were

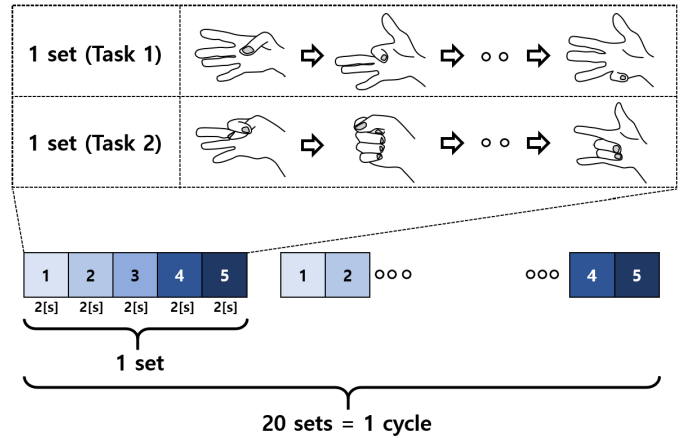
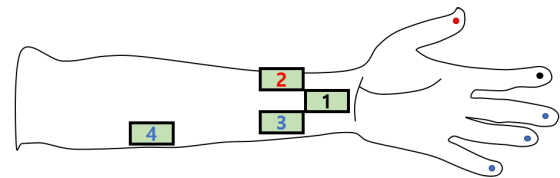


Fig. 1. Experiments composed of two tasks with Task 1 for training the data and Task 2 for testing the data.



EMG Num	Target Muscle	Muscle
EMG 1	Flexor digitorum superficialis	Index, Middle, Ring, Pinky finger flexion
EMG 2	Flexor pollicis longus	Thumb finger flexion
EMG 3,4	Flexor digitorum profundus	Index, Middle, Ring, Pinky finger flexion

Fig. 2. Locations of four EMG sensors and the three muscles on which sensor are placed. All three muscles are involved with finger flexion.

asked to repeat Task 1 for 10 times and Task 2 once. A one-minute break was provided between the tasks. A graphical description of the experiment is shown in Fig. 1.

### B. Data Collection

Four sEMG sensors were attached to different locations of the skin of the forearm, identified as the locations of three extrinsic muscles known to contribute to flexion of at least one of the five fingers: the flexor digitorum superficialis, the flexor pollicis longus, and the flexor digitorum profundus. The flexor pollicis longus muscle is mainly involved with thumb flexion and the other two muscles are with flexion of the remaining four fingers. The EMG signals were obtained at a frequency of 1.26 kHz. The locations of the EMG sensors and the main functions of the three muscles are shown in Fig. 2.

### C. Data Processing

1) *EMG Signal Processing*: The raw EMG signals were offset to make the mean value to zero, and the absolute values of the modified signals were taken. To eliminate the high frequency noise, the EMG signals were filtered with a second-order Butterworth filter (cutoff frequency: 5 Hz) and normalized to fit into the range between 0 and 1 [25]. Finally, the EMG signals sampled at 1.26 kHz were downsampled to 1 kHz. Downsampling and low-pass filtering are common methods of preprocessing of EMG data [26], [27].

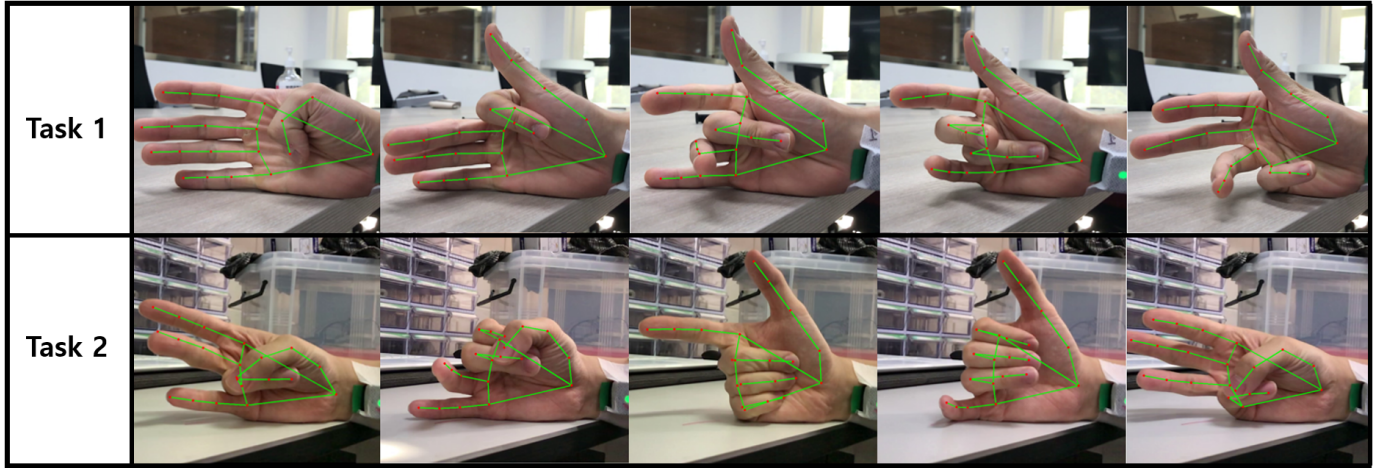


Fig. 3. Examples of finger joint tracking of five different hand motions in Task 1 and Task 2 using a machine learning solution (MediaPipe Hand). The MediaPipe calculates the finger joint position  $(x, y, z)$  from a 2D image.

2) *Video Processing*: The hand movements of the participants were recorded using a 240 fps mobile phone camera (iPhone 7, Apple). The finger joint angles were estimated using an open source framework (MediaPipe Hands, Google Inc., USA) that can calculate the positions and the motions of the fingers [28]. The angles of the three finger joints (metacarpophalangeal (MCP), proximal interphalangeal (PIP), distal interphalangeal (DIP)) were calculated based on the joint positions (angles of thumb MCP (1-MCP), thumb IP (1-IP), index MCP (2-MCP), index PIP (2-PIP), index DIP (2-DIP), middle MCP (3-MCP), middle PIP (3-PIP), middle DIP (3-DIP), ring MCP (4-MCP), ring PIP (4-PIP), ring DIP (4-DIP), little MCP (5-MCP), little PIP (5-PIP), and little DIP (5-DIP)). Then, the angles of each finger were normalized in the range of 0-1. Normalization of each joint angle was performed by subtracting the minimum estimated joint angle from the corresponding joint angle and dividing it by the difference between the maximum estimated joint angle and the minimum estimated joint angle. Fig. 3 shows examples of applying the MediaPipe Hands library to the recorded videos.

### III. PROPOSED MODEL

The main consideration factor for a machine learning model in this study is construction of an algorithm that effectively extracts the characteristics of the EMG signals and the joint angle data. The proposed model considers the following three characteristics of the experiment.

- 1) The EMG signals are time series data (i.e., sequential data).
- 2) The EMG signals from the four sensors are not independent.
- 3) During Task 1, the EMG signals from a specific sensor are strongly activated when the corresponding finger is flexed or extended.

We constructed a machine learning algorithm to reflect the above three features, which is explained in the followings.

#### A. Time Series Data

To take advantage of sequential data, a gate recurrent unit (GRU) was used as a core unit. A GRU is one type

of a recurrent neural network (RNN) which is useful when dealing with sequential data. A GRU is similar to a long short-term memory (LSTM) with a forget gate but has fewer parameters than an LSTM because it does not have an output gate [29], [30].

#### B. Signal Dependency

The encoder-decoder network structure was employed to utilize the dependency of the EMG signals. The encoder consists of four GRUs, with each GRU receiving a single element of the EMG signals ( $x_t \in \mathbb{R}^4$ ) at time  $t$  with the result of the previous GRU as an input value and delivering the output vector and the hidden vector to the next GRU. The hidden vector  $h_t \in \mathbb{R}^{256}$  is calculated as

$$h_t = \tanh(W^{(hh)}h_{t-1} + W^{(hx)}x_t) \quad (1)$$

where  $W^{(hh)} \in \mathbb{R}^{256 \times 256}$  and  $W^{(hx)} \in \mathbb{R}^{4 \times 256}$  are learnable parameters.

The decoder is a stack of several GRUs where each of them estimates an output  $y_t \in \mathbb{R}^{14}$  at time  $t$ . The GRU receives the hidden state of the previous GRU as an input and calculates the output and the next hidden state. The hidden state  $h_t$  and the output  $y_t$  are calculated as

$$h_t = \tanh(W^{(hh)}h_{t-1}) \quad (2)$$

$$y_t = \text{softmax}(W^S h_t) \quad (3)$$

where  $W^{(S)} \in \mathbb{R}^{14 \times 256}$  is a learnable parameter.

#### C. Sensor-Finger Relations

A encoder-decoder network without attention but with a GRU does not fully reflect the above three features. One of the useful experimental observations is that certain EMG sensors respond significantly more than the other sensors when flexing or extending a specific finger. An attention mechanism was used to implement this observation to the model. The main role of the attention mechanism is to induce the model to learn the process of referring to specific EMG sensor ( $EMG_i$ ,  $i \in \{1, 2, 3, 4\}$ ) inputs when predicting the related finger joint angles ( $\theta_i$ ,  $i \in \{1, 2, \dots, 14\}$ ), as described in Fig. 4.

The sub-figures show the activated rectified EMG signals from the four EMG sensors, respectively, when executing three sets of Task 1.

The final model is a single encoder connected to five different decoders with five different attention matrices, respectively. The five different decoders predict the joint angles of the five fingers (two joints for the thumb and three joints for the index, the middle, the ring, and the little fingers) in parallel. Each of the five attention matrices is applied to the corresponding decoder and does not share the learning weights with the others. The EMG signals at time  $t$ ,  $X(t)$  (i.e.,  $[EMG_1(t), EMG_2(t), EMG_3(t), EMG_4(t)]$ ), are divided into four one-hot vectors,  $X_1(t)$ ,  $X_2(t)$ ,  $X_3(t)$ , and  $X_4(t)$ ,

$$X(t) = \begin{bmatrix} EMG_1(t) \\ EMG_2(t) \\ EMG_3(t) \\ EMG_4(t) \end{bmatrix} \mapsto \begin{bmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \\ X_4(t) \end{bmatrix},$$

and each  $X_i(t)$ ,  $i \in \{1, 2, 3, 4\}$  is one-hot vector that has  $EMG_j(t)$ ,  $j \in \{1, 2, 3, 4\}$  as a non-zero entries where

$$X_1(t) = \begin{bmatrix} EMG_1(t) \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad X_2(t) = \begin{bmatrix} 0 \\ EMG_2(t) \\ 0 \\ 0 \end{bmatrix},$$

$$X_3(t) = \begin{bmatrix} 0 \\ 0 \\ EMG_3(t) \\ 0 \end{bmatrix}, \quad X_4(t) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ EMG_4(t) \end{bmatrix}.$$

$X_1(t)$ ,  $X_2(t)$ ,  $X_3(t)$ , and  $X_4(t)$  are fed into the encoder sequentially in a given order. Since the EMG signals are generated from the neural system (the cortex and the spinal cord), the order of the four one-hot vectors can be determined by the arrival time of the electrical impulses to the four sensors, which is the same as the proximal-to-distal order of the sensor positions. However, this sequence is not necessary and would not hinder the learning process because the depth of the encoder is not too deep to raise the vanishing problem of the input vectors. Fig. 5 overviews the final model.

Fig. 5 shows the inner structure of the shared encoder (Fig. 5-(a)) and the thumb decoder (Fig. 5-(b)). The shared encoder is composed of four encoder cells that are GRUs, as shown in Fig. 5-(a). The  $i^{\text{th}}$  GRU takes  $X_i(t) \in \mathbb{R}^4$  and  $h_{en}^{i-1}(t) \in \mathbb{R}^{256}$  as inputs where  $X_i(t)$  is  $i^{\text{th}}$  the rectified one-hot vector EMG signal and  $h_{en}^{i-1}(t)$  is the hidden vector from the  $(i-1)^{\text{th}}$  GRU. Then it yields  $o_{en}^i(t) \in \mathbb{R}^{256}$  and  $h_{en}^i(t)$  as outputs. The 1<sup>st</sup> GRU takes the hidden vector  $h_{en}^0(t)$  which is a zero vector and  $X_1(t)$  as an input. As the  $i^{\text{th}}$  GRU's input is the same  $(i-1)^{\text{th}}$  GRU's output, the  $h_{en}^i(t)$  contains the previous input information  $X_1(t), \dots, X_i(t)$ .  $O_{en}(1)$  contains only the information of  $EMG_1(t)$ ,  $O_{en}(2)$  contains the information of  $EMG_1(t)$  and  $EMG_2(t)$ ,  $O_{en}(3)$  contains  $EMG_1(t)$ ,  $EMG_2(t)$ , and  $EMG_3(t)$ , and so on. The final outputs of the shared encoder are  $h_{en}^4(t)$  and  $O_{en}(t) \in \mathbb{R}^{256 \times 4}$  where  $h_{en}^4(t)$  contains all the input information in

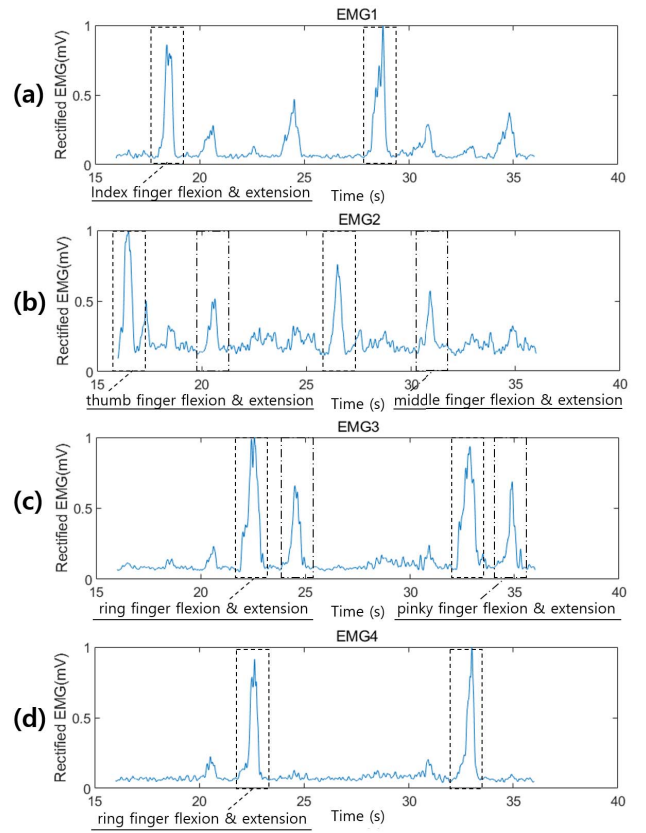


Fig. 4. EMG activation data from Task 1 by showing specific activation of a specific sensor with flexion and extension of a specific finger. (a) Sensor 1, (b) sensor 2, (c) sensor 3, and (d) sensor 4. The dotted box shows the well-marked EMG activation of specific finger movements than compared to the other fingers.

a compressed 256-sized vector and  $O_{en}(t)$  is a matrix that concatenates  $O_{en}^i(t)$ ,  $i \in \{1, 2, 3, 4\}$  in a column direction.

The thumb decoder is composed of two decoder cells, as shown in Fig. 5-(b). The  $i^{\text{th}}$  decoder cell takes three inputs,  $o_{de}^{i-1} \in \mathbb{R}^{14}$ ,  $h_{de}^{i-1} \in \mathbb{R}^{256}$ ,  $O_{en} \in \mathbb{R}^{256 \times 4}$  where the  $o_{de}^{i-1}$  is the estimation vector of the finger joint angle from the  $(i-1)^{\text{th}}$  decoder cell,  $h_{de}^{i-1}$  is the hidden vector from the  $(i-1)^{\text{th}}$  decoder cell, and  $O_{en}$  is the encoder output. In the case of the thumb decoder,  $o_{de}^i$ ,  $i \in \{1, 2\}$  is one-hot vector.  $\theta_i(t)$  is non-zero entries at the  $i^{\text{th}}$  index of the vector  $o_{de}^i$ .

The decoder cells are all identical for the five different decoders.  $o_{de}^{i-1}$  passes a linear layer ( $f_{\text{linear}1}$ ,  $14 \times 256$ ) with a forget dropout ( $p = 0.1$ ). This vector ( $c_1$ ) concatenates with  $h_{de}^{i-1}$ . Then the concatenated vector  $c_2 \in \mathbb{R}^{512}$  passes a linear layer ( $f_{\text{linear}2}$ ,  $512 \times 4$ ) and a softmax layer. This outcome becomes the attention matrix ( $\mathbb{R}^{4 \times 1}$ ).

The key, the value, and the query of the attention mechanism can be defined as the following:

- key:  $O_{en} = [O_{en1}; O_{en2}; O_{en3}; O_{en4}]$
- query:  $c_2 = f_{\text{linear}2}([f_{\text{linear}1}(o_{de}^{i-1}); h_{de}^{i-1}])$
- value = key

The dot product was used to define the attention score (also known as the Luong attention [31]). The attention score,  $\text{Att}_{\text{score}} \in \mathbb{R}$  is defined as

$$\text{Att}_{\text{score}}(c_2, O_{eni}) = c_2^T O_{eni}, \quad (4)$$

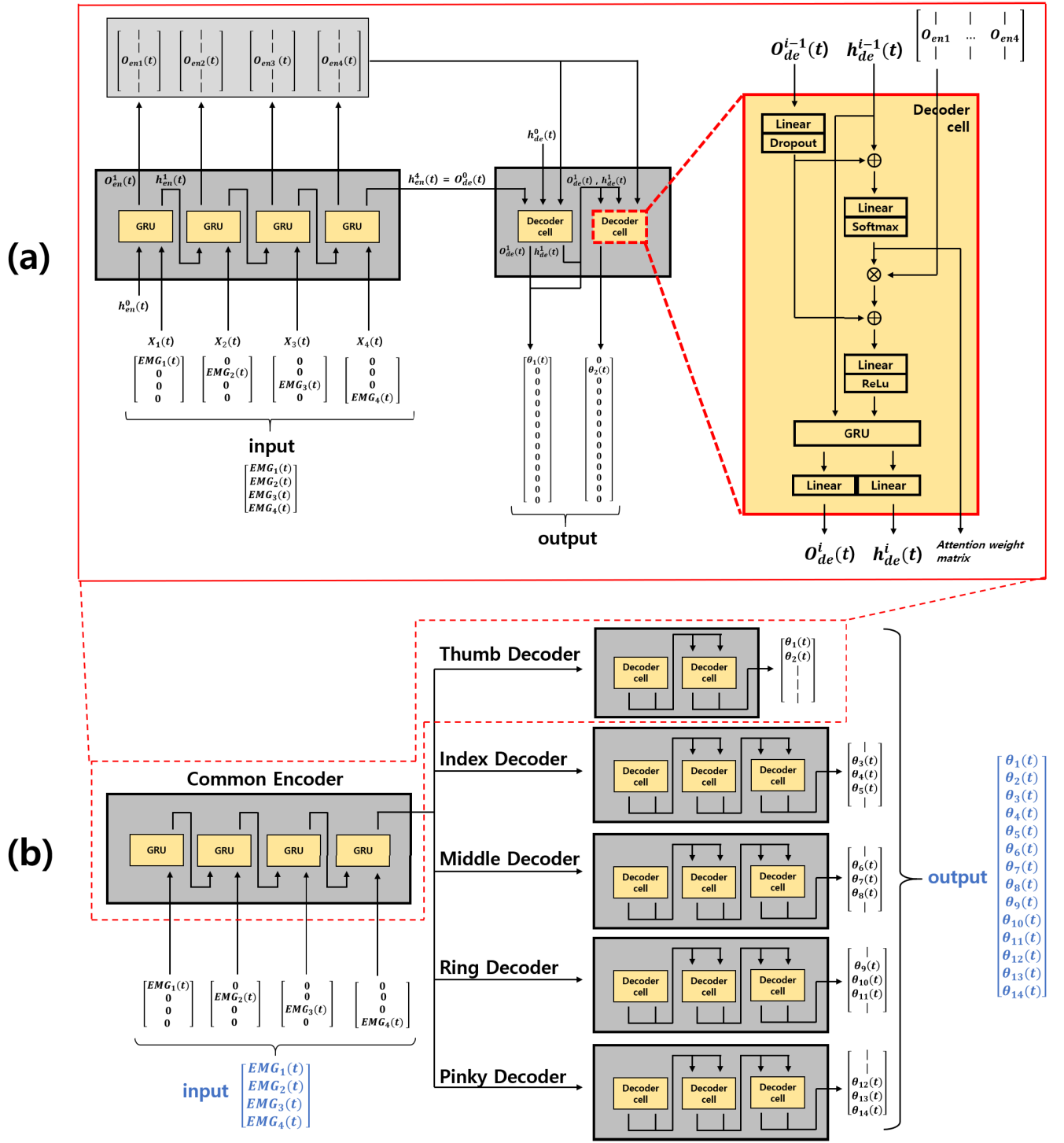


Fig. 5. Overall structure of the proposed network: (a) Encoder-decoder model with an attention matrix. (b) The entire model.

and the attention value  $\text{Att}_{\text{out}} \in \mathbb{R}^{4 \times 1}$  as

$$\text{Att}_{\text{out}} = \text{softmax}([c_2^T O_{en1}, c_2^T O_{en2}, c_2^T O_{en3}, c_2^T O_{en4}]). \quad (5)$$

The attention matrix is calculated by multiplying the key and the query. The matrix multiplication is the crux of the attention mechanism. It means which output  $O_{en}^i$  the decoder cell focuses on when predicting the finger joint angle,  $o_{de}^i$ .

Multiplication of  $\text{Att}_{\text{out}}$  and  $O_{en}$  gives  $c_3$ , i.e.,

$$c_3 = O_{en} \cdot \text{Att}_{\text{out}}. \quad (6)$$

Then the vector  $c_3 \in \mathbb{R}^{256}$  concatenates with the vector  $c_1 \in \mathbb{R}^{256}$  and passes a linear layer ( $f_{\text{linear3}}, 512 \times 256$ ) and a rectified linear activation unit (ReLU) layer [32]. The output vector  $c_4$  can be represented as

$$c_4 = \text{ReLU}(f_{\text{linear3}}([c_3; c_1])). \quad (7)$$

The  $c_4$  and the previous hidden vector  $h_{de}^{i-1}$  are fed into the GRU, which yields an output  $c_5 \in \mathbb{R}^{256}$  and a hidden vector  $h_{de}^i \in \mathbb{R}^{256}$ . Then, the vector  $c_5$  passes the linear layer ( $f_{\text{linear4}}, 256 \times 14$ ) and yields  $o_{de}^i \in \mathbb{R}^{14}$ .

$$c_5, h_{de}^i = \text{GRU}(c_4, h_{de}^{i-1}), \quad (8)$$

$$o_{de}^i = f_{\text{linear4}}(c_5). \quad (9)$$

The training loss is defined as the root-mean-squared error (RMSE) between the estimation angles at time  $t$  ( $o_{de}^i \in \mathbb{R}^{14}$ ) and reference angles at time  $t$  ( $y(t) \in \mathbb{R}^{14}$ ) where  $i \in \{1, 2, \dots, 14\}$ .

$$\text{RMSE} (\%) = \sqrt{\frac{\sum_{i=1}^{14} \{\theta_i(t) - y_i(t)\}^2}{14}} \times 100 \quad (10)$$

The L2 regularization term with a weight decaying  $\lambda_{reg} = 0.00001$ . Since the  $i^{\text{th}}$  entry of  $o_{de}^i$  is  $\theta_i(t)$ , the loss function is expressed as

$$\text{loss} = \sqrt{\frac{\sum_{i=1}^{14} (\theta_i(t) - y_i(t))^2}{14}} + \lambda_{reg} * \sum_{w_i \in W} \|w_i\|^2. \quad (11)$$

where  $W$  is the learnable weight.

## IV. RESULT

### A. Data Preparation

The EMG data from the four sensors ( $X$ ) and their corresponding 14 finger joint angles ( $Y$ ) were preprocessed to make the input data ( $X, Y$ ) to train the model. The train and the test data sets are different. When training the model, the data set ( $X_{\text{train}}, Y_{\text{train}}$ ) is only composed of the EMG signals and the finger joint angles obtained from Task 1. When testing the model, the data set ( $X_{\text{test}}, Y_{\text{test}}$ ) is composed of the EMG signals and the finger joint angles obtained from Task 2. The ranges of the angles in the train data and from the test data are the same.

### B. Model Training

To evaluate the performance of the proposed model, a comparison study was conducted. In this study, we checked the effectiveness of the attention mechanism and the RNN module. Three different neural network models were trained and their best results were compared. The three models are a simple neural network, the proposed model without the attention mechanism, and the proposed model with the attention mechanism. The three models have nearly five million learnable parameters to have the same model complexity, and each model was trained in the same hyperparameter space. The simple neural network is a multilayer perceptron (MLP) with five hidden layers. To check the performance of the RNN that is the deep learning module for time series data, we additionally defined the simple neural network and used it for the comparison study. The size of the hidden layer was fixed to 256 for all three networks. The hyperparameters were teacher-forcing ratios and learning rates. The teacher forcing ratio ( $\gamma_{tfr}$ ) was searched over  $\gamma_{tfr} = \{0.5, 0.7, 0.9\}$ , and the learning ratio ( $\gamma_{lr}$ ) was searched over

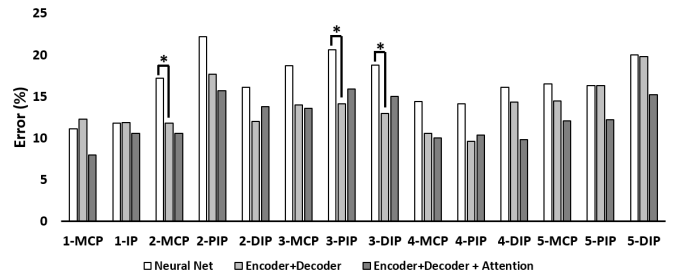


Fig. 6. Error plots in estimating joint angles in Task 2 for three different models. “\*” shows differences of more than 5%.

TABLE I  
ERRORS IN ESTIMATING ANGLES OF 14 FINGER JOINTS  
USING THREE DIFFERENT MODELS (TASK 2)

		simple NN (%)	En-De (%)	En-De-Att (%)
Thumb finger	MCP	11.1 ± 10.4	12.3 ± 5.4	8.0 ± 7.4
	IP	11.8 ± 10.7	11.9 ± 6.4	10.6 ± 6.8
Index finger	MCP	17.2 ± 15.0	11.8 ± 8.9	10.6 ± 10.3
	PIP	22.2 ± 21.1	17.7 ± 14.4	15.7 ± 17.2
	DIP	16.1 ± 14.4	10.1 ± 9.0	13.8 ± 11.8
Middle finger	MCP	18.7 ± 16.4	14.0 ± 10.0	13.6 ± 12.1
	PIP	20.6 ± 17.9	14.1 ± 9.7	15.9 ± 13.3
	DIP	18.8 ± 15.8	13.0 ± 8.7	15.0 ± 11.7
Ring finger	MCP	14.4 ± 12.0	10.6 ± 8.0	10.0 ± 9.2
	PIP	14.1 ± 11.8	9.6 ± 7.7	10.4 ± 10.0
	DIP	16.1 ± 13.6	14.3 ± 10.3	9.8 ± 9.1
Little finger	MCP	16.5 ± 15.4	14.5 ± 9.6	12.1 ± 10.1
	PIP	16.3 ± 15.3	16.3 ± 8.7	12.2 ± 11.4
	DIP	20.0 ± 14.7	19.8 ± 10.6	15.2 ± 8.7

$\gamma_{lr} = \{0.0001, 0.0003, 0.0005, 0.0007\}$ . The code was implemented with an automated deep learning processor, Pytorch 1.8. Minibatch stochastic gradient descent with a size of 1024 was used for optimizing the loss function. The Adam optimizer was used for the optimization process ( $\beta_1 = 0.9, \beta_2 = 0.999$  and  $\epsilon = 10^{-8}$ ) [33].

The test accuracy of the three different models for estimating 14 finger joint angles are shown in Fig. 6 and Table I. The error was defined as an RMSE between the estimation  $y_i(t)$  and the reference  $\theta_i(t)$  values.

The test error of the encoder-decoder network without attention and that with attention are lower than that of the simple neural network except for 1-MCP. Especially, 2-MCP, 3-PIP, and 3-DIP showed relatively large differences in the test errors between the simple network and the network without attention and also between the simple network and the network with attention. Also, the test errors of the network with attention for the thumb angles (1-MCP, 1-IP), the index finger angles (2-MCP, 2-PIP), the ring finger angles (4-MCP, 4-DIP), the little finger angles (5-MCP, 5-PIP, 5-DIP) are lower than that of the network without attention. However, the test errors of the attention model for the middle finger (3-PIP, 3-DIP) are higher than those of the model without attention.

## V. DISCUSSION

### A. Comparison of Prediction Accuracy of Three Different Models

The encoder-decoder networks showed more accurate and stable estimation than the simple neural network

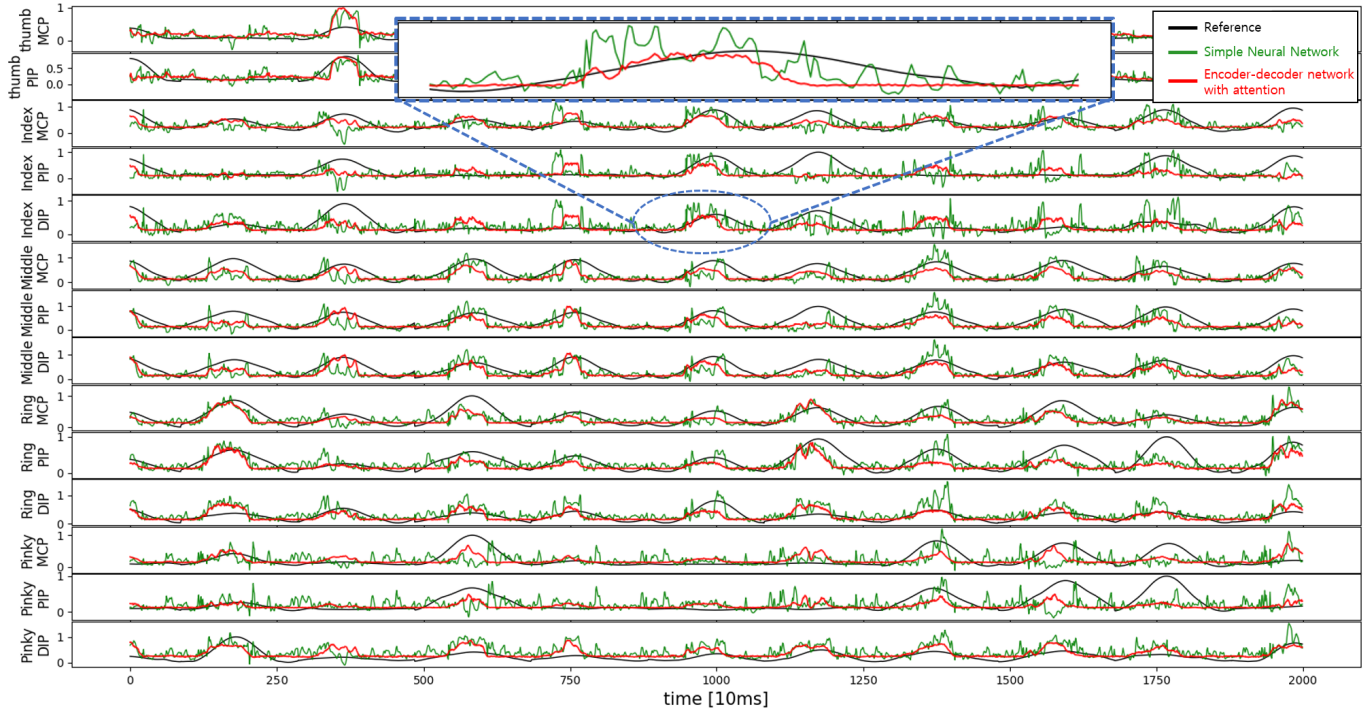


Fig. 7. Estimation results of 14 joint angles (Task 2: random finger flexion and extension). x-axis is time and y-axis is the normalized 14 finger angle between 0 and 1. The black lines are the references obtained from the experiment, the red lines are the estimations using a simple neural network, and the green lines are the estimations of the proposed model.

(Fig. 6 and Table I), implying that it is more appropriate to use a GRU, a general basic unit for sequential data processing, for the analysis of EMG data than a method of simple stacking of hidden layers in simple neural networks (Fig. 7). This result also suggests that the GRU-based encoder-decoder model is suitable for complex classification or prediction tasks, by interpreting complex human biosignals, such as EEG or EMG.

This study also shows the ability of generalization of the proposed model. Few studies have implemented machine learning models to estimate finger joint angles when the train data and the test data are completely different. Previous studies on predicting finger joint angles performed both model training and testing within the same motion data of individual fingers [18]–[21]. However, all these studies that use machine learning to find the optimal relationship ( $f$ ) between the EMG signals ( $X$ ) and the finger joint angles ( $Y$ ) show a strong nonlinearity, indicating that the linear combination of the muscle signals ( $f(x_1 + x_2)$ ) for the joint angles is not the same as the linear combination of  $f(x_1)$  and  $f(x_2)$  [34], [35]. In other words,

$$f(x_1 + x_2) \neq f(x_1) + f(x_2) \quad (12)$$

holds where  $x_1, x_2 \in X$  and  $f(x_1), f(x_2) \in Y$ .

This means that the linear combination of the EMG data and the joint angle data collected during Task 1 cannot fully represent the data from both Task 1 and Task 2. Therefore, the performance of the proposed model validated by the different data sets indicates the higher accuracy in generalization than the previous model.

TABLE II  
PEARSON'S CORRELATION COEFFICIENTS  
OF 14 FINGER JOINT ANGLES

1-MCP	1-IP	2-MCP	2-PIP	2-DIP	3-MCP	3-PIP
<b>0.68</b>	<b>0.66</b>	0.59	0.37	0.50	<b>0.74</b>	<b>0.65</b>
3-DIP	4-MCP	4-PIP	4-DIP	5-MCP	5-PIP	5-DIP
<b>0.73</b>	<b>0.72</b>	<b>0.71</b>	0.56	0.45	0.41	<b>0.63</b>

\*The boldfaced entries are values that satisfy moderate positive correlation.

In addition, we calculated the Pearson correlation coefficients between the predicted angles from the proposed model and the estimated angles to quantify the performance of continuous prediction, and the 14 coefficients are shown in Table II. We also conducted a one-way analysis of variance (ANOVA) test between the three different models to check the differences. The 14 p-values from the ANOVA test are less than 0.05, and we can conclude that there are significant differences in the mean values of the predicted angles from the three different models.

### B. Explainable Model by Attention

One of the significant implications of this study is interpretation of after-trained attention matrices ( $\text{Att}_{\text{out}}$ ). The result of the attention matrices and our interpretation provide a reason on the high estimation accuracy of the encoder-decoder network with attention than that without attention.

Fig. 8 is the color-map of after-trained attention matrix ( $\text{Att}_{\text{out}}$ ) composed of a basic attention matrix ( $\text{Att}_{\text{out}}^p$  where  $p = \{1, 2, \dots, 14\}$ ) stacked in the column direction.  $\text{Att}_{\text{out}}^p$  is

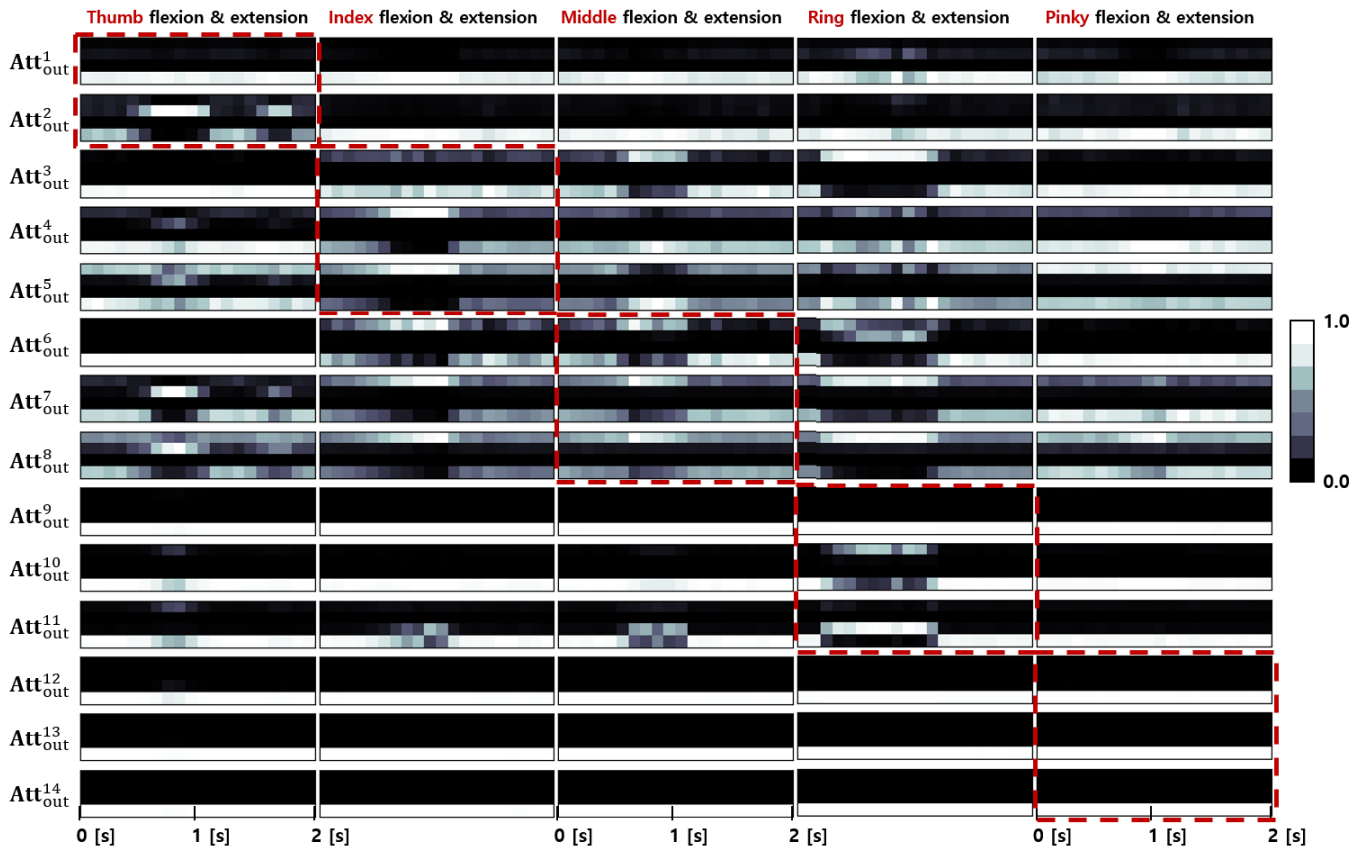


Fig. 8. Attention matrix result when moving individual finger (Task 1). This explains why the proposed network yields higher accuracy than previously proposed all machine learning method.

an attention matrix, one of the outcomes of a decoder cell that estimates  $\theta_p$ . The width of  $\text{Att}_{\text{out}}^p$  depends on the experiment time. For example, if the subject performed only thumb flexion and extension (the total experiment time is two seconds, and the data are resampled at 10 Hz for visualization), we obtain two attention matrices,  $\text{Att}_{\text{out}}^1 \in \mathbb{R}^{4 \times 20}$  and  $\text{Att}_{\text{out}}^2 \in \mathbb{R}^{4 \times 20}$ . The  $x$ -axis of  $\text{Att}_{\text{out}}^p$  is time ( $10^{-1} \text{ s} = 1/f_{\text{resample}} \text{ s}$ ) and  $y$ -axis is the encoder output values ( $O_{eni}(t)$ ,  $i \in \{1, 2, 3, 4\}$ ) referenced by the decoder cells to predict the joint angle  $\theta_p$ . For fixed time  $t_0$  ( $j = t_0$ ), each  $\text{Att}_{\text{out}}^p(i, j = t_0)$  for  $i \in \{1, 2, 3, 4\}$  indicates the contribution of  $O_{eni}(t)$  when predicting the angle  $\theta_p$ . These values are normalized, i.e.,  $\sum_{i=1}^4 \text{Att}_{\text{out}}^p(i, j = t_0) = 1$ . Fig. 8 is the after-trained attention matrix ( $\text{Att}_{\text{out}}$ ) for evaluation data (one set of Task 1).  $\text{Att}_{\text{out}}$  consists of a total of 14  $\text{Att}_{\text{out}}^p \in \mathbb{R}^{4 \times 100}$  where  $p \in \{1, 2, \dots, 14\}$ .

The following explains the white areas in the attention matrix and how they are related to the experimental observation.

1)  $\text{Att}_{\text{out}}^1$  and  $\text{Att}_{\text{out}}^2$ : They usually show attentions (bright color) on  $O_{en4}(t)$ , but for thumb flexion and extension,  $\text{Att}_{\text{out}}^2$  shows active referencing to  $O_{en2}(t)$  during the time at 0.4 s  $\sim$  1.05 s and at 1.55 s  $\sim$  1.75 s, which is the compressed information on  $EMG_1(t)$  and  $EMG_2(t)$ . This matches the experimental observation that  $EMG_2(t)$  highly activates when flexing and extending the thumb (Fig. 4).

2)  $\text{Att}_{\text{out}}^3$ ,  $\text{Att}_{\text{out}}^4$ , and  $\text{Att}_{\text{out}}^5$ : For flexion and extension of index finger,  $\text{Att}_{\text{out}}^3$ ,  $\text{Att}_{\text{out}}^4$ , and  $\text{Att}_{\text{out}}^5$  show attentions on  $O_{en4}(t)$  and specially on  $O_{en1}(t)$  that is the compressed information of  $EMG_1(t)$ . This matches the experimental observation that  $EMG_1(t)$  highly activates when flexing and extending the index finger. The relatively bright color of  $\text{Att}_{\text{out}}^4(1, t)$  (2.3 s  $\sim$  2.55 s), and  $\text{Att}_{\text{out}}^5(1, t)$  (2.3 s  $\sim$  2.6 s) also matches the experimental result that  $EMG_1(t)$  is more activated with the index finger than with the middle or the ring finger.

3)  $\text{Att}_{\text{out}}^9$ ,  $\text{Att}_{\text{out}}^{10}$ , and  $\text{Att}_{\text{out}}^{11}$ : For flexion and extension of the ring finger, the attention matrix ( $\text{Att}_{\text{out}}^{11}$ ) shows a different attention pattern of a bright color on  $O_{en3}(t)$  that is the compressed information of  $EMG_1(t)$ ,  $EMG_2(t)$ , and  $EMG_3(t)$  for flexion and extension of the middle, the ring, and the little fingers.  $\text{Att}_{\text{out}}^{11}(4, t)$  activates during flexion and extension of the index (2.3 s  $\sim$  2.55 s), the middle (2.3 s  $\sim$  2.55 s), the ring (2.1 s  $\sim$  2.6 s) fingers. A relatively bright color and a longer time for  $\text{Att}_{\text{out}}^{11}(4, t)$  of the ring finger also matches the experimental result of the highly activated  $EMG_3(t)$  during flexion and extension of the ring finger.

4)  $\text{Att}_{\text{out}}^6$ ,  $\text{Att}_{\text{out}}^7$ , and  $\text{Att}_{\text{out}}^8$ : On the other hand, for flexion and extension of the middle finger,  $\text{Att}_{\text{out}}^6$ ,  $\text{Att}_{\text{out}}^7$ , and  $\text{Att}_{\text{out}}^8$  show different patterns especially on  $O_{en1}(t)$ . In Fig. 4,  $EMG_1$  slightly activates with the middle finger.



5)  $Att_{out}$ : The above interpretations represent that the attention matrix after training learns the relationship between the EMG signals and the finger joint angles. This means that the attention mechanism induces the encoder-decoder network to refer to experimental observations for estimating the finger joint angles with a higher accuracy. Therefore, the attention mechanism contributes to the high estimation accuracy by providing the model with the opportunity to learn the relationship between the finger joint angles and the corresponding muscle signals. This is consistent with a previous study that has reported on the definition of explainable AI (XAI) in the medical field [36]. Although prior research has shown a high estimation accuracy in decoding the biosignals (EEG and EMG) using machine learning, they were not able to effectively explain why their learning algorithms yielded the results [37]–[40].

### C. Challenges of Real Time Control of Prosthesis

The availability of real time control of an upper-limb robotic prosthesis in a clinical setting is a challenge due to the complex activation of multiple muscles and multiple DOFs of the upper arm movements. The performance of robotic prosthesis with machine learning is related to the time complexity of the device, which is composed of algorithmic complexity, computational time, and hardware costs [41]. The algorithmic (time) complexity of our model is composed of the complexity of the attention mechanism,  $O(nd)$ , and the complexity of GRU,  $O(nd^2)$ , where  $n$  is the number of the EMG sensors and  $d$  is the hidden size of the attention matrix and the GRU [24], [42]. According to prior research, the time complexities of machine learning models that control a robot arm using EMG signals, such as a convolutional neural network (CNN) or a multilayer perceptron (MLP) are  $O(knd^2)$  and  $O(n)$ , respectively, where  $n$  is the number of the sensors,  $d$  is the hidden size, and  $k$  is the kernel size [43], [44]. We think the time complexity of the proposed model will not be a problem if we select proper  $n$ ,  $d$  and robot hardware. However, the real-time control strategy is still limited due to the tradeoff between the complexity of the mechanical configuration and the complexity of the control systems [45]–[48].

### D. Future Work

One of the immediate areas of future work will be implementation of transfer learning [49] or incremental learning [50] for training models using the joint angle data and the forearm EMG data collected at specific arm positions and for predicting the finger joint angles using the EMG signals from different arm postures. Another important area of future work will be experimental validation of the proposed approach using physical systems, including robotic hands [51]–[53] and wearable robots [54]–[57] that require seamless interfaces between the user intentions and actuation. In addition, it would be beneficial to check the practicality of the proposed AI model in real-world applications, such as clinical rehabilitation, where the ability of real-time control is critical. In this way, we can also verify the long-term effect of the proposed method.

## VI. CONCLUSION

This study proposed a new machine learning model for estimating finger joint angles based on using raw EMG signals. The proposed model has a structure of a shared encoder - five parallel decoders with a basic unit of GRU, and an attention mechanism is applied to construct an explainable machine learning model. The encoder consists of four GRU cells, and the decoder consists of two or three decoder cells with the same number of cells as that of estimation angles, and each of the five decoders has an attention matrix. We first predicted complex data (Task 2 - random finger movements) using the model learned with only simple data (Task 1 - individual finger movements), which demonstrated the ability of generalization of the proposed model. In addition, the attention mechanism applied in this study provides the reason why the proposed model allows for accurate and stable estimation compared to two other models: a simple neural network and an encoder-decoder network without attention. The color map results of the attention matrix after training proved that the learning was proceeded by reflecting the physiological relationship between the forearm EMG signals and the finger angles during model training.

## ACKNOWLEDGMENT

The authors thank the research participants for their dedication that made this study possible and Yuna Yoo for discussions about the experimental protocol and the results.

## REFERENCES

- [1] G. S. Dhillon and K. W. Horch, "Direct neural sensory feedback and control of a prosthetic arm," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, no. 4, pp. 468–472, Dec. 2005.
- [2] G. N. Saridis and T. P. Gootee, "EMG pattern analysis and classification for a prosthetic arm," *IEEE Trans. Biomed. Eng.*, vol. BME-29, no. 6, pp. 403–412, Jun. 1982.
- [3] I. Vujaklija, D. Farina, and O. Aszmann, "New developments in prosthetic arm systems," *Orthopedic Res. Rev.*, vol. 8, p. 31, Mar. 2016.
- [4] N. A. Alibeji, V. Molazadeh, F. Moore-Clingenpeel, and N. Sharma, "A muscle synergy-inspired control design to coordinate functional electrical stimulation and a powered exoskeleton: Artificial generation of synergies to reduce input dimensionality," *IEEE Control Syst.*, vol. 38, no. 6, pp. 35–60, Dec. 2018.
- [5] D. R. Merrill, J. Lockhart, P. R. Troyk, R. F. Weir, and D. L. Hankin, "Development of an implantable myoelectric sensor for advanced prosthesis control," *Artif. Organs*, vol. 35, no. 3, pp. 249–252, Mar. 2011.
- [6] T. Tamei and T. Shibata, "Fast reinforcement learning for three-dimensional kinetic human-robot cooperation with an EMG-to-activation model," *Adv. Robot.*, vol. 25, no. 5, pp. 563–580, 2011.
- [7] N. A. Alibeji, N. A. Kirsch, and N. Sharma, "A muscle synergy-inspired adaptive control scheme for a hybrid walking neuroprosthesis," *Frontiers Bioeng. Biotechnol.*, vol. 3, p. 203, Dec. 2015.
- [8] H. Rehbaum, N. Jiang, L. Paredes, S. Amsuess, B. Graimann, and D. Farina, "Real time simultaneous and proportional control of multiple degrees of freedom from surface EMG: Preliminary results on subjects with limb deficiency," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 1346–1349.
- [9] S. Muceli and D. Farina, "Simultaneous and proportional estimation of hand kinematics from EMG during mirrored movements at multiple degrees-of-freedom," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 3, pp. 371–378, May 2012.
- [10] J. Kim, S. Mastnik, and E. André, "EMG-based hand gesture recognition for realtime biosignal interfacing," in *Proc. 13th Int. Conf. Intell. User Interface (IUI)*, 2008, pp. 30–39.
- [11] X. Zhang, X. Chen, Y. Li, V. Lantz, K. Wang, and J. Yang, "A framework for hand gesture recognition based on accelerometer and EMG sensors," *IEEE Trans. Syst., Man, Cybern., A, Syst. Humans*, vol. 41, no. 6, pp. 1064–1076, Nov. 2011.

- [12] E. Rahimian, S. Zabihi, S. F. Atashzar, A. Asif, and A. Mohammadi, "Surface EMG-based hand gesture recognition via hybrid and dilated deep neural network architectures for neuroprosthetic prostheses," *J. Med. Robot. Res.*, vol. 5, no. 2, 2020, Art. no. 2041001.
- [13] R. N. Khushaba, S. Kodagoda, M. Takruri, and G. Dissanayake, "Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals," *Expert Syst. Appl.*, vol. 39, pp. 10731–10738, Sep. 2012.
- [14] Y. Huang, K. B. Englehart, B. Hudgins, and A. D. C. Chan, "A Gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 11, pp. 1801–1811, Nov. 2005.
- [15] J.-U. Chu, I. Moon, Y.-J. Lee, S.-K. Kim, and M.-S. Mun, "A supervised feature-projection-based real-time EMG pattern recognition for multi-function myoelectric hand control," *IEEE/ASME Trans. Mechatronics*, vol. 12, no. 3, pp. 282–290, Jun. 2007.
- [16] E. Rahimian, S. Zabihi, A. Asif, D. Farina, S. F. Atashzar, and A. Mohammadi, "FS-HGR: Few-shot learning for hand gesture recognition via electromyography," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1004–1015, 2021.
- [17] P. Afshar and Y. Matsuoka, "Neural-based control of a robotic hand: Evidence for distinct muscle strategies," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, vol. 5, Apr. 2004, pp. 4633–4638.
- [18] N. A. Shirao, N. P. Reddy, and D. R. Kosuri, "Neural network committees for finger joint angle estimation from surface EMG signals," *Biomed. Eng. OnLine*, vol. 8, no. 1, pp. 1–11, Dec. 2009.
- [19] R. J. Smith, F. Tenore, D. Huberdeau, R. Etienne-Cummings, and N. V. Thakor, "Continuous decoding of finger position from surface EMG signals for the control of powered prostheses," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2008, pp. 197–200.
- [20] M. Hioki and H. Kawasaki, "Estimation of finger joint angles from sEMG using a recurrent neural network with time-delayed input vectors," in *Proc. IEEE Int. Conf. Rehabil. Robot.*, Jun. 2009, pp. 289–294.
- [21] J. G. Ngeo, T. Tamei, and T. Shibata, "Continuous and simultaneous estimation of finger kinematics using inputs from an EMG-to-muscle activation model," *J. Neuroeng. Rehabil.*, vol. 11, no. 1, pp. 1–14, Dec. 2014.
- [22] J. Wiens *et al.*, "Do no harm: A roadmap for responsible machine learning for health care," *Nature Med.*, vol. 25, no. 9, pp. 1337–1340, Sep. 2019.
- [23] E. Halilaj, A. Rajagopal, M. Fiterau, J. L. Hicks, T. J. Hastie, and S. L. Delp, "Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities," *J. Biomech.*, vol. 81, pp. 1–11, Nov. 2018.
- [24] A. Vaswani *et al.*, "Attention is all you need," in *Proc. NeurIPS*, 2017, pp. 5998–6008.
- [25] P. Konrad, "The ABC of EMG: A practical introduction to kinesiological electromyography," Ver. 1.4, Noraxon Scottsdale, AZ, USA, 2006, pp. 5–30.
- [26] J. B. Stephansen *et al.*, "Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy," *Nature Commun.*, vol. 9, no. 1, pp. 1–15, Dec. 2018.
- [27] T. Lajnef *et al.*, "Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines," *J. Neurosci. Methods*, vol. 250, pp. 94–105, Jul. 2015.
- [28] C. Lugaresi *et al.*, "MediaPipe: A framework for building perception pipelines," 2019, [arXiv:1906.08172](https://arxiv.org/abs/1906.08172).
- [29] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, [arXiv:1406.1078](https://arxiv.org/abs/1406.1078).
- [30] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Comput.*, vol. 12, no. 10, pp. 2451–2471, Oct. 2000.
- [31] M.-T. Luong, H. Pham, and C. D. Manning, "Effective approaches to attention-based neural machine translation," 2015, [arXiv:1508.04025](https://arxiv.org/abs/1508.04025).
- [32] A. Fred Agarap, "Deep learning using rectified linear units (ReLU)," 2018, [arXiv:1803.08375](https://arxiv.org/abs/1803.08375).
- [33] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- [34] S. K. Dwivedi, J. Ngeo, and T. Shibata, "Extraction of nonlinear synergies for proportional and simultaneous estimation of finger kinematics," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 9, pp. 2646–2658, Sep. 2020.
- [35] J. Ngeo, T. Tamei, K. Ikeda, and T. Shibata, "Modeling dynamic high-DOF finger postures from surface EMG using nonlinear synergies in latent space representation," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 2095–2098.
- [36] E. Tjoa and C. Guan, "A survey on explainable artificial intelligence (XAI): Toward medical XAI," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 11, pp. 4793–4813, Nov. 2021.
- [37] U. Cote-Allard *et al.*, "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 760–771, Apr. 2019.
- [38] S. Chambon, M. N. Galtier, P. J. Arnal, G. Wainrib, and A. Gramfort, "A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 4, pp. 758–769, Apr. 2018.
- [39] P. Wang, A. Jiang, X. Liu, J. Shang, and L. Zhang, "LSTM-based EEG classification in motor imagery tasks," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 11, pp. 2086–2095, Nov. 2018.
- [40] D. Zhang, L. Yao, K. Chen, S. Wang, P. D. Haghghi, and C. Sullivan, "A graph-based hierarchical attention model for movement intention detection from EEG signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 11, pp. 2247–2253, Nov. 2019.
- [41] P. Maciejasz, J. Eschweiler, K. Gerlach-Hahn, A. Jansen-Troy, and S. Leonhardt, "A survey on robotic devices for upper limb rehabilitation," *J. Neuroeng. Rehabil.*, vol. 11, no. 1, pp. 1–29, Dec. 2014.
- [42] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, [arXiv:1409.0473](https://arxiv.org/abs/1409.0473).
- [43] B. Guo, Y. Ma, J. Yang, Z. Wang, and X. Zhang, "Lw-CNN-based myoelectric signal recognition and real-time control of robotic arm for upper-limb rehabilitation," *Comput. Intell. Neurosci.*, vol. 2020, pp. 1–12, Dec. 2020.
- [44] N. Parajuli *et al.*, "Real-time EMG based pattern recognition control for hand prostheses: A review on existing methods, challenges and future implementation," *Sensors*, vol. 19, no. 20, p. 4596, Oct. 2019.
- [45] Y. Jiang *et al.*, "Shoulder muscle activation pattern recognition based on sEMG and machine learning algorithms," *Comput. Methods Programs Biomed.*, vol. 197, Dec. 2020, Art. no. 105721.
- [46] M. Cracchiolo *et al.*, "Decoding of grasping tasks from intraneural recordings in trans-radial amputee," *J. Neural Eng.*, vol. 17, no. 2, Apr. 2020, Art. no. 026034.
- [47] S. Wendelken *et al.*, "Restoration of motor control and proprioceptive and cutaneous sensation in humans with prior upper-limb amputation via multiple Utah slanted electrode arrays (USEAs) implanted in residual peripheral arm nerves," *J. Neuroeng. Rehabil.*, vol. 14, no. 1, pp. 1–17, Dec. 2017.
- [48] V. Mendez, F. Iberite, S. Shokur, and S. Micera, "Current solutions and future trends for robotic prosthetic hands," *Annu. Rev. Control, Robot., Auto. Syst.*, vol. 4, no. 1, pp. 595–627, May 2021.
- [49] D. W. Kim, J. Kwon, B. Jeon, and Y.-L. Park, "Adaptive calibration of soft sensors using optimal transportation transfer learning for mass production and long-term usage," *Adv. Intell. Syst.*, vol. 2, no. 6, 2020, Art. no. 1900178.
- [50] F. Schiel, A. Hagengruber, J. Vogel, and R. Triebel, "Incremental learning of EMG-based control commands using Gaussian processes," in *Proc. Conf. Rob. Learn.*, 2020, pp. 1137–1146.
- [51] R. Mutlu, G. Alici, M. het Panhuis, and G. M. Spinks, "3D printed flexure hinges for soft monolithic prosthetic fingers," *Soft Robot.*, vol. 3, no. 3, pp. 120–133, Sep. 2016.
- [52] S. J. Yoon, M. Choi, B. Jeong, and Y.-L. Park, "Elongatable gripper fingers with integrated stretchable tactile sensors for underactuated grasping and dexterous manipulation," *IEEE Trans. Robot.*, early access, Feb. 7, 2022, doi: [10.1109/TRO.2022.3144949](https://doi.org/10.1109/TRO.2022.3144949).
- [53] L. Jiang, K. Low, J. Costa, R. J. Black, and Y.-L. Park, "Fiber optically sensorized multi-fingered robotic hand," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2015, pp. 1763–1768.
- [54] T. Proietti *et al.*, "Sensing and control of a multi-joint soft wearable robot for upper-limb assistance and rehabilitation," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 2381–2388, Apr. 2021.
- [55] Y.-L. Park *et al.*, "Design and control of a bio-inspired soft wearable robotic device for ankle-foot rehabilitation," *Bioinspiration Biomimetics*, vol. 9, no. 1, 2014, Art. no. 016007.
- [56] J. Kwon, J.-H. Park, S. Ku, Y. Jeong, N.-J. Paik, and Y.-L. Park, "A soft wearable robotic ankle-foot-orthosis for post-stroke patients," *IEEE Robot. Autom. Lett.*, vol. 4, no. 3, pp. 2547–2552, Jul. 2019.
- [57] T. H. Hong, S. H. Park, J. H. Park, N. J. Paik, and Y.-L. Park, "Design of pneumatic origami muscle actuators (POMAs) for a soft robotic hand orthosis for grasping assistance," in *Proc. 3rd IEEE Int. Conf. Soft Robot. (RoboSoft)*, May/Jul. 2020, pp. 627–632.