Swallowing Pattern Classification Method Using Multichannel Surface EMG Signals of Suprahyoid and Infrahyoid Muscles

Masahiro Suzuki,* Makoto Sasaki,*, # Katsuhiro Kamata, * Atsushi Nakayama, ** Isamu Shibamoto, *** Yasushi Tamada†

Abstract The ability to fine-tune the movement of swallowing-related organs and change the swallowing pattern to fit the volume of a bolus, texture and the physical properties of the food to be swallowed is referred to as the swallowing reserve. In other words, it is the response capability of food swallowing to avoid choking and aspiration. Herein, we focus on the coordination of the suprahyoid and infrahyoid muscles activities, which are closely related to swallowing movement, as a first step to develop a method to evaluate swallowing reserve, which declines due to neuromuscular disease, muscle weakness caused by aging, to mention a few. First, using two 22-channel electrodes, we measured the surface electromyography (sEMG) signals of suprahyoid and infrahyoid muscles during the following four swallowing conditions: combining two bolus volumes (3 and 15 mL water) and two techniques (normal and effortful swallow). Then, we verified whether the difference in swallowing patterns based on swallowing conditions can be classified from sEMG signals using three machine learning methods; namely, the real-time classification, comprehensive classification, and image recognition method. In the real-time classification method, the mean classification accuracy (MCA) for the four swallowing conditions was as low as 81.5%, indicating that the difference between swallowing conditions performed in a period of approximately 1 s cannot be classified sufficiently by this method. In the comprehensive classification method that applies a majority decision to all the classification results from the start to the end of swallowing, which can be obtained every 16 ms, MCA was 95.1%. Furthermore, in the image recognition method, the change of a series of sEMG signals in the swallowing movement was converted into swallowing pattern image, and the images were classified using a combination of deep convolutional neural networks and support vector machine (SVM). Compared with the comprehensive classification method, the number of training samples for the image recognition method was only 1/26, but the MCA reached 95.7%. This method, which can noninvasively evaluate swallowing patterns that change slightly based on swallowing conditions, could be applied to early detection of reduced swallowing function or a state of frailty (dysphagia potential) in aged individuals.

Keywords: swallowing, pattern classification, suprahyoid and infrahyoid muscles, surface EMG signals.

Adv Biomed Eng. 9: pp. 10–20, 2020.

1. Introduction

Swallowing is a series of movement to transport food from the oral cavity to the stomach, and is divided into three phases; namely, the oral, pharyngeal, and esophageal phases [1]. Swallowing is mediated by a complex physiological mechanism involving “voluntary movement” and “involuntary reflexive movement.” The voluntary movement mainly refers to the intraoral movement to transport the bolus, which is formed by food mastication, to the pharynx via tongue movement (oral phase). Involuntary movement refers to the swallowing reflex-mediated movement to pass the bolus through the pharynx (pharyngeal phase). The swallowing reflex is a highly reproducible and extremely precise movement pattern programmed by central pattern generator (CPG) in the medulla oblongata. Therefore, functional oral movements and intact swallowing reflex are essential for successful swallowing.

Although the swallowing movement pattern induced by the swallowing reflex is highly reproducible, the
movement pattern changes depending on differences in the volume of the bolus (amount of food) [2–6] and the physical properties of the food (such as viscosity and hardness) [7, 8]. Therefore, the ability to fine-tune the movement of swallowing-related organs and change the swallowing pattern to fit objects to be swallowed is considered to be the swallowing reserve, which is the response capability of food swallowing in order to avoid choking and aspiration [9–11]. However, this swallowing reserve declines due to neuromuscular disease, muscle weakness caused by aging, cerebrovascular disease, as well as positional changes of the swallowing-related organs. Furthermore, in patients with swallowing problems secondary to systemic diseases, pharyngeal residue, laryngeal penetration, and bolus aspiration may occur, with an increased risk of developing aspiration pneumonia [12–14]. Generally, decline of the swallowing reserve may cause descent of the hyoid bone and larynx position, associated decrease in the amount of hyoid and larynx elevation in the anterosuperior direction, delay in larynx elevation due to reduced larynx elevation speed, and delay in stimulation of the swallowing reflex [9, 11, 15–17]. However, the specific reference values to detect reduced swallowing function have not yet been clarified. Furthermore, the videofluoroscopic examination of swallowing, which is the gold standard of detailed examination, is essential for the observation and evaluation of these problems. However, this method is not suitable for routine and regular swallowing function evaluation owing to risks of radiation exposure and aspiration of contrast media. To extend the healthy life expectancy of the elderly, it is important to detect elderly individuals who have reduced swallowing function or are in a state of frailty (dysphagia potential) at an early stage [18]. Since these conditions are difficult to perceive by the individuals themselves, it is necessary to develop a safe and convenient evaluation method to monitor daily changes in swallowing reserve.

This study focuses on the coordination of suprahyoid and infrahyoid muscles activities as a new approach for evaluating swallowing reserve. These muscles are closely related to the movement of the hyoid and larynx and have the advantage that their activities can be observed from the skin surface of the anterior neck [19–21]. Therefore, if one can detect the differences in swallowing patterns depending on the swallowing conditions from these surface electromyography (sEMG) signals, it may be possible to detect early decline of “the ability to change swallowing pattern (swallowing reserve).”

Herein, sEMG signals of the suprahyoid and infrahyoid muscles are measured under four swallowing conditions, which combines two bolus volumes (3 and 15 mL of water) and two techniques (normal and effortful swallow), as a first step for evaluating the swallowing reserve. Then, we verified whether the difference in swallowing patterns based on swallowing conditions could be classified from sEMG signals using three types of machine learning methods; namely, the real-time classification, comprehensive classification, and image recognition methods.

2. Method
2.1 Participants
Eight healthy young men without any history of dysphagia (22.4 ± 1.2 years of age, 171.4 ± 6.0 cm in height, 62.3 ± 7.4 kg in weight; mean ± SD) participated in the experiment. Approval for the study was obtained in advance by the Ethical Review Board of Iwate University (Number: 201808). Before starting the tests, we explained the study objectives, experimental protocol, and risks to each participant, and received written consent from all the participants.

2.2 Measurement of sEMG signals
The suprahyoid muscles comprise four muscles; namely, digastric, stylohyoid, mylohyoid, and geniohyoid muscles, located above the hyoid bone in the neck (Fig. 1). The infrahyoid muscles are made up of four muscles; sternohyoid, sternothyroid, thyrohyoid, and omohyoid muscles, located below the hyoid bone in the neck. During swallowing, the suprahyoid muscles pull the hyoid bone anterosuperiorly, followed by the thyrohyoid muscle to elevate the larynx. These movements cause epiglottis rotation, thereby helping to prevent aspiration while a bolus passes through the pharynx. Therefore, the activities of suprahyoid and infrahyoid muscles are closely related to the hyoid and laryngeal movements, which play an essential role in the swallowing mechanism.

For measurement of sEMG signals, we designed two 22-channel electrodes based on our previous work [22].

**Fig. 1** Structure of suprahyoid muscles and infrahyoid muscles.
Figure 2 shows the appearance and electrode arrangements of the device. The tips of the electrode, which are made of pure silver rods, are shaped as hemispheres of Φ2 mm to facilitate skin contact. For electrical insulation of the electronic parts, we covered both sides of the flexible polyimide substrate with a layer of silicon. The 22-channel electrode for the suprahyoid muscles was designed as a thin flexible boomerang-shaped patch for attachment to the region above the hyoid bone in the anterior triangle of the neck. The electrode for the infrahyoid muscles was designed as a thin flexible U-shaped patch for attachment to the region under the hyoid bone in the anterior triangle of the neck. The anterior triangle of the neck is the region bounded by the anterior median line of the neck, the inferior border of the mandible, and anterior border of the sternocleidomastoid muscle. Additionally, sEMG signals of the suprahyoid and infrahyoid muscles can be obtained from the skin surface of the anterior triangle of the neck.

For the experiment, we adhered both 22-channel electrodes to the anterior neck of the participant and connected a ground electrode and an indifferent electrode to the left and right earlobes, respectively, using ear clips. The electric potential between each electrode of the 22-channel device and the indifferent electrode was amplified using a separate differential amplifier with reference to the potential of the ground electrode. We set the gain of the differential amplifier to 2,052. Furthermore, we used a band-pass filter with a passband ranging from 16 to 440 Hz bandwidth to remove the noise superimposed on sEMG signals. We digitalized all the 44-channel sEMG signals by synchronizing two 16-bit analog-to-digital converters (USB-6218; National Instruments Corp.). We set the sampling rate to 2,000 Hz.

2.3 Swallowing conditions
The swallowing reflex is induced by "central input from a higher level such as the cerebral cortex," which is involved in voluntary swallowing, and "peripheral input from the pharynx and larynx," which is involved in the reflexive input [1]. In the experiment, to account for these inputs, we performed swallowing under four conditions; combining two bolus volumes and two techniques (Table 1). We used water as a bolus and set the volumes for one swallow to be 3 and 15 mL. In the modified water swallowing test, a screening test to evaluate swallowing function [23], 3 mL (defined volume) of water is used to reduce aspiration risk. Moreover, a healthy adult can swallow at least 20 mL at a time [24]. In a preliminary experiment, all the participants were able to swallow ≥ 20 mL of water at a time. However, 15 mL of water was selected for the experiment considering fatigue of participants and reproducibility of the measurement data. Normal and effortful swallow techniques [1] were used to swallow the boluses of water. An effortful swallow is a common rehabilitation technique to improve swallowing function. For the normal swallow (non-effortful swallow), we instructed the participants to swallow the water naturally. For the effortful swallow, they were instructed to swallow the water as hard as possible ("As you swallow, squeeze hard with all of your muscles" [1]). Furthermore, we delivered water onto the floor of the mouth using a syringe and asked each participant to swallow it in a relaxed state in a sitting position. The experiments were conducted in the following order. First, 3-mL normal swallow (NS3) was performed, followed by 3-mL effortful swallow (ES3). Then, 15-mL normal swallow (NS15), and finally 15-mL effortful swallow (ES15) were performed. Furthermore, we considered these four types of swallow as one set and measured a total of 20 sets of sEMG signals.

3. Swallowing pattern classification using tongue motion classification method

3.1 Classification method
First, to verify the possibility of classifying swallowing patterns using sEMG signals recorded from the suprahy-
swallowing pattern classification using sEMG signals of suprahyoid muscles (Csm); (ii) swallowing pattern classification using sEMG signals of infrahyoid muscles (Cim); and (iii) swallowing pattern classification using sEMG signals of suprahyoid muscles and infrahyoid muscles (Csim).

To classify the swallowing patterns, we set the analysis range of the measured sEMG signals to 6,048 samples (approximately 3 s), which include muscle activity from the start to the end of swallowing. In feature extraction, the overlapped windowing technique of Englishart et al. [25] was performed, which possibly reduces window length to ≤128 ms without a large decrease in classification accuracy, resulting in faster generation of classification result. First, we segmented the sEMG signals in the analysis range into data windows comprising 256 samples (128 ms). Next, we slide the next sampling segment over the current one after a time increment of 16 ms. Thus, feature extraction was performed for a total of 182 windows.

As a feature, we calculated the root mean square (RMS) and cepstrum coefficients (CC) of the sEMG signals for each window and each channel. The RMS features are characteristics of the time domain and provide amplitude information of sEMG signals, whereas CC features are characteristics of the frequency domain. In addition, low-order CC include information about the spectral envelope of the power spectrum of sEMG signals. For the swallowing pattern classification, we selected low-order (1–3) CC as CC features.

The feature extracted from a window is an $M$-dimensional vector, (1-dimensional RMS features + 3-dimensional CC features) × number of channels. Therefore, the feature vector is 88 dimensions for classification conditions Csm and Cim, and 176 dimensions for the classification condition Csim. Training data comprising pairs of feature vectors and correct class labels are required for swallowing pattern classification using the tongue motion classification method. We obtained the correct class labels that correspond to the feature vectors by threshold triggering of sEMG signals.

Furthermore, we utilized the support vector machine (SVM) classifier for swallowing pattern classification. Then, using training data, we created a decision function for classifying each swallowing pattern using an SVM classifier with a radial basis function kernel. We optimized the hyperparameters, i.e., kernel parameter $\gamma$ and penalty parameter $C$, of the SVM using a grid search process within the training data. Besides, using six-fold cross-validation, we defined the optimal combination of the hyperparameters with the highest classification rate. After training the SVM with optimized hyperparameters for $\gamma$ and $C$, we performed swallowing pattern classification using unknown sEMG signals from the test data. Finally, to minimize the effect of possible misclassification, we applied the majority-voting technique to the past $m$ class labels and obtained the final predicted class label with the highest number of wins. We set the majority decision number $m$ to 20 for the real-time classification and 182 for the comprehensive classification. The former ($m = 20$) is the number of decisions for obtaining the final predicted class label in real time during swallowing, and the label is output every 16 ms. When there were fewer than 20 labels at the starting points in the analysis range, the majority of available labels were considered as the final predicted class label. The latter ($m = 182$) is the number of decisions for obtaining only one final predicted class label after the end of swallowing, and the label is determined comprehensively using all the 182 predicted class labels output from the sEMG signals in the analysis range.

The classification accuracy (CA) of the test data was calculated using the following equation:

$$CA = \frac{\text{Number of correctly classified samples}}{\text{Total number of samples}}$$
CA(%) = \frac{\text{number of correctly predicted class label}}{\text{number of predicted class label}} \quad (1)

In addition, to avoid data bias, we divided 20 sets of measured datasets into four groups and utilized two groups (10 sets) as training data and the remaining two groups (10 sets) as test data. We repeated this calculation six times (the number of combinations is \( \binom{4}{2} = 6 \)), and used the average value of all the CAs to evaluate the sensitivity of swallowing pattern classification. Specificity was also obtained as supplementary information to evaluate the reliability of the classification method used [26].

### 3.2 Results

**Tables 2 and 3** show the CA for all the participants. In the real-time classification \((n = 20)\) of swallowing patterns using tongue motion classification method, the mean classification accuracy (MCA) for Csm and Cim was 76.1% and 68.1%, respectively, which was relatively low. Regarding the Csim, a slight improvement in MCA compared to Csm and Cim was observed, but the accuracy was only 81.5%. On the contrary, in the comprehen-

#### Table 2  Classification accuracy (CA) for four swallowing conditions in the real-time classification.

| Sub. | NS3 | ES3 | NS15 | ES15 | Total | NS3 | ES3 | NS15 | ES15 | Total | NS3 | ES3 | NS15 | ES15 | Total |
|------|-----|-----|------|------|-------|-----|-----|------|------|-------|-----|-----|------|------|-------|
| A    | 65.5| 79.9| 70.1 | 84.8 | 75.1  | 46.4| 62.6| 62.1 | 77.2 | 62.1  | 61.9| 81.9| 72.6 | 87.3 | 75.9 |
| B    | 72.9| 69.2| 68.5 | 82.9 | 73.4  | 67.5| 55.6| 67.2 | 73.3 | 65.9  | 81.5| 72.4| 76.2 | 82.7 | 78.2 |
| C    | 76.9| 80.8| 69.4 | 76.5 | 75.9  | 64.5| 77.8| 72.2 | 81.3 | 74.0  | 84.9| 85.6| 94.5 | 89.0 | 88.5 |
| D    | 71.2| 71.9| 74.2 | 80.9 | 74.5  | 67.8| 75.9| 75.3 | 84.3 | 75.8  | 80.3| 83.5| 85.7 | 89.9 | 84.9 |
| E    | 76.0| 78.6| 70.2 | 71.2 | 74.0  | 63.9| 64.1| 64.8 | 65.1 | 64.5  | 80.8| 79.5| 73.6 | 76.4 | 77.6 |
| F    | 72.7| 79.9| 77.5 | 87.4 | 79.4  | 82.7| 69.7| 72.6 | 75.1 | 75.0  | 90.8| 83.2| 85.5 | 86.6 | 86.5 |
| G    | 87.6| 83.4| 77.0 | 82.8 | 82.7  | 72.1| 67.9| 63.8 | 72.7 | 69.1  | 89.3| 81.7| 80.8 | 84.5 | 84.1 |
| H    | 79.9| 71.8| 66.0 | 78.8 | 74.1  | 62.9| 51.5| 57.8 | 60.7 | 58.2  | 82.6| 72.6| 74.2 | 76.7 | 76.5 |
| MCA  | 75.3| 76.9| 71.6 | 80.7 | 76.1  | 66.0| 65.7| 67.0 | 73.7 | 68.1  | 81.5| 80.1| 80.4 | 84.1 | 81.5 |
| MS   | 96.0| 92.8| 93.6 | 91.6 | 93.5  | 95.7| 90.2| 90.7 | 87.9 | 91.1  | 97.7| 94.3| 94.7 | 93.3 | 95.0 |

MCA: Mean classification accuracy, MS: Mean specificity.

#### Table 3  CA for four swallowing conditions in the comprehensive classification.

| Sub. | NS3 | ES3 | NS15 | ES15 | Total | NS3 | ES3 | NS15 | ES15 | Total | NS3 | ES3 | NS15 | ES15 | Total |
|------|-----|-----|------|------|-------|-----|-----|------|------|-------|-----|-----|------|------|-------|
| A    | 78.3| 98.3| 81.7 | 96.7 | 88.8  | 50.0| 78.3| 78.3 | 96.7 | 88.8  | 83.3| 93.3| 86.7 | 98.3 | 90.4 |
| B    | 91.7| 90.0| 93.3 | 100  | 93.8  | 85.0| 75.0| 88.3 | 88.3 | 84.2  | 93.3| 95.0| 100  | 100  | 97.1 |
| C    | 95.0| 95.0| 96.7 | 98.3 | 96.3  | 75.0| 96.7| 96.7 | 95.0 | 90.8  | 95.0| 95.0| 100  | 100  | 97.5 |
| D    | 88.3| 93.3| 91.7 | 93.3 | 91.7  | 83.3| 93.3| 93.3 | 95.0 | 88.7  | 95.0| 95.0| 98.3 | 95.0 | 95.8 |
| E    | 95.0| 98.3| 88.3 | 93.3 | 92.5  | 73.3| 78.3| 81.7 | 88.3 | 80.4  | 95.0| 96.7| 90.0 | 95.0 | 94.2 |
| F    | 90.0| 100 | 96.7 | 98.3 | 96.3  | 95.0| 90.0| 98.3 | 100  | 95.8  | 100| 95.0| 98.3 | 95.0 | 97.9 |
| G    | 100 | 93.3| 90.0 | 96.7 | 95.0  | 85.0| 83.3| 80.0 | 86.7 | 83.8  | 100| 95.0| 91.7 | 98.3 | 96.3 |
| H    | 95.0| 88.3| 76.7 | 98.3 | 89.6  | 86.7| 71.7| 83.3 | 86.7 | 82.1  | 91.7| 93.3| 85.0 | 95.0 | 91.3 |
| MCA  | 91.7| 94.6| 89.4 | 96.2 | 93.0  | 79.2| 82.1| 87.5 | 91.9 | 85.2  | 94.2| 94.8| 93.8 | 97.5 | 95.1 |
| MS   | 98.8| 97.0| 98.1 | 96.7 | 97.7  | 98.0| 95.6| 93.6 | 93.1 | 95.1  | 99.7| 97.8| 98.9 | 97.0 | 98.4 |

MCA: Mean classification accuracy, MS: Mean specificity.
Swallowing is a momentary movement that takes place within one second or less. Nevertheless, the need to classify swallowing conditions in real time is not very high in swallowing function evaluation. Therefore, to classify swallowing pattern with high accuracy, it would be useful to comprehensively determine sEMG signals of both the suprathyroid muscles and the infrahyoid muscles from the start to the end of swallowing and obtain the final predicted class label corresponding to the swallowing condition immediately after swallowing.

4. Swallowing pattern classification using image recognition method

4.1 Classification method

In the previous section, we showed that swallowing pattern classification should focus on the coordinated movement of the suprathyroid muscles and infrahyoid muscles from the start of swallowing to the end. Therefore, in the next study, we converted the change of a series of sEMG signals during swallowing movement into an image and verified whether it is possible to classify the swallowing pattern by image recognition.

In this analysis, the information of “88-dimensional feature vector $\times 182$ windows” extracted from the analysis range consisting of 6,048 samples were imaged as one image. This means that the sEMG signals of the suprathyroid and infrahyoid muscles are imaged individually. This information used for imaging was the same as the features used in the previous section. First, the feature vector consisting of RMS and CC features extracted in each window of 128 ms was converted into a color map to obtain the array shown in Fig. 4. Since RMS and CC features have different scales, they were normalized respectively in advance using the maximum and the minimum values in the training dataset. This normalization method is widely known as the min–max normalization technique. However, this method is easily affected by outliers. Hence, each value was arranged in descending order, and the average of the top 100 values was considered as the maximum value for normalization and the average of the lower 100 values as the minimum value for normalization. Next, a swallowing pattern image was

![Swallowing pattern images of (a) suprathyroid muscles and (b) infrahyoid muscles.](Image)
created by arranging all the color maps created from the 182 feature vectors. Finally, this image was resized to an image of 227 × 227 pixels for image recognition.

For classification of swallowing pattern images, we used AlexNet [27], which is one of the convolutional neural networks (CNNs) with outstanding classification performance in the field of image recognition (Fig. 5). AlexNet has been trained using over 1.2 million images from the ImageNet dataset and can classify the images into 1,000 categories of objects. In the deep learning applications, the transfer learning that fine-tunes a pre-trained AlexNet with new categories of images is commonly used to classify categories of images [28]. However, in this study, there were only ten datasets of sEMG signals that can be used for transfer learning, and the number of swallowing pattern images was not sufficient to fine-tune a pretrained AlexNet.

Therefore, we used a combination of AlexNet and SVM classifier [29–31] for classification of swallowing pattern images, since AlexNet is robust for feature learning and feature expression. Hence, AlexNet was used for feature extraction, and SVM classifier was used for classification. Based on previous research, we applied SVM after the fully connected layer fc6 in AlexNet (Fig. 5). The dimensions of the feature vector extracted from the fully connected layer fc6 is 4,096. When we performed the swallowing pattern classification using the swallowing pattern images of both the suprahyoid muscles and the infrahyoid muscles (i.e., Csm), an 8,192-dimensional vector—in which two feature vectors extracted respectively were serially connected—was used as the SVM input signal (Fig. 6).

The training procedure of SVM, method of selecting training data and test data, and method of calculating the CA were the same as in Subsection 3.1. Additionally, while the number of learning samples in comprehensive classification was 7,280 (182 windows × 4 swallowing conditions × 10 sets), there were only 40 (1 image × 4 swallowing conditions × 10 sets) in the image recognition method. Therefore, there is a possibility that the SVM training cannot be performed sufficiently. Hence, for the training data, we created the swallowing pattern image while shifting the analysis range (6,048 samples) by 32 samples, up to 96 samples in forward and backward directions, and increased the number of learning samples sevenfold. Thus, the number of learning samples was increased to 1/26 of that in the comprehensive classification.

### 4.2 Results

Table 4 shows the CA for all the participants. The MCA of the swallowing pattern using image recognition was 92.9%, 92.8% and 95.7% for Csm, Cim and Csim, respectively. Additionally, the MCA exceeded 90% under all the classification conditions. When examining individual swallowing conditions, the MCA exceeded 90% in all swallowing conditions for Csm, Cim and Csim. The mean specificity was 97.6%, 97.6% and 98.6% for Csm, Cim and Csim, respectively. Table 5 shows the confusion matrix in image recognition for Csim, which has the highest MCA among all the classification conditions. Consequently, it was found that the misclassification rate between NS3 and ES15 (differed by both bolus volume and swallowing technique) was 0%, showing no misclassification between the two conditions. Additionally, the misclassification rate was ≤1.0% between ES3 and NS15, showing rare misclassification between the two conditions. On the contrary, when either the bolus volume or the swallowing techniques was the same, the misclassification rate increased. Nevertheless, the highest misclassification rate was only 3.1%.

### 4.3 Discussion

The suprahyoid muscles are involved in voluntary movements such as tongue movement to transport a bolus to the pharynx, whereas the thyrohyoid muscle (infrahyoid muscles) is involved in involuntary reflex movements.
and follows the activity of the suprahyoid muscles to elevate the larynx. Since participants were instructed to swallow the water as hard as possible for the effortful swallow, which closely relates to voluntary movement, we speculated that classification of swallowing patterns using only sEMG signals of the infrahyoid muscles would be difficult. As suspected, the MCA for Cim was 68.1% in the real-time classification, and 85.2% in the comprehensive classification; both exhibited low accuracy. However, when using the image recognition method, the MCA for Cim and Csm were both approximately 93%, which indicates that sufficient information for classification can be extracted from the sEMG signals of the infrahyoid muscles. This finding shows that pretrained AlexNet is robust and has excellent feature extraction even for untrained swallowing pattern images. Also, for both Csm and Cim in comprehensive classification, CA was less than 80% under several swallowing conditions in several participants. However, with the image recognition method, the CA was 80% or more under all conditions in all the participants, indicating the possibility of stable classification regardless of whether the sEMG signals of the suprahyoid muscles or those of the infrahyoid muscles were used.

As mentioned above, although the number of training samples was only 1/26 of that in comprehensive classification, the image recognition method can classify the four swallowing conditions with high and stable accuracies. Herein, AlexNet pretrained for classifying 1,000 categories of objects, was adopted as a feature extractor. Moreover, SVM was used to classify four swallowing conditions from the features extracted by AlexNet. However, fine-tuning of a pretrained AlexNet for classifying new category of objects, which is generally performed in transfer learning, cannot be performed because the number of swallowing pattern images that can be used as training data was small. Nevertheless, the fact that the MCA for Csim reached 95.7% indicates the robustness of AlexNet as a feature extractor and also suggests the possibility of further development of swallowing condition classification using swallowing pattern images. Therefore, CA may be further improved by increasing the swallowing pattern images as training data and performing fine-tuning of AlexNet or adopting another CNNs such as ResNet [32]. Additionally, there is room for optimizing the method of creating swallowing pattern images, and further improvement in CA is expected for swallowing pattern classification using image recognition.

5. Conclusion

The objective of this study was to confirm whether the

| Actual class | Predicted class |
|--------------|----------------|
|              | NS3 | ES3 | NS15 | ES15 |
| NS3          | 95.2| 2.5 | 2.3  | 0.0  |
| ES3          | 1.5 | 95.4| 0.0  | 3.1  |
| NS15         | 0.4 | 1.0 | 96.0 | 2.5  |
| ES15         | 0.0 | 1.9 | 1.9  | 96.3 |

MCA: Mean classification accuracy, MS: Mean specificity.

Table 4 CA for the four swallowing conditions in the image recognition method.

| Classification using sEMG signals of the suprahyoid muscles (Csm) | Classification using sEMG signals of the infrahyoid muscles (Cim) | Classification using sEMG signals of the suprahyoid and infrahyoid muscles (Csim) |
|-------------------------------------------------------------------|-------------------------------------------------------------------|--------------------------------------------------------------------------------|
| Sub.                  NS3 | ES3 | NS15 | ES15 | Total    NS3 | ES3 | NS15 | ES15 | Total    NS3 | ES3 | NS15 | ES15 | Total |
| A                     86.7 | 91.7| 91.7 | 95.0 | 91.3 | 85.0 | 93.3| 93.3 | 90.0 | 90.4 | 91.7 | 96.7| 93.3 | 95.0 | 94.2 |
| B                     95.0 | 96.7| 91.7 | 90.0 | 93.3 | 90.0 | 91.7| 98.3 | 96.7 | 94.2 | 95.0 | 95.0| 100 | 100 | 97.5 |
| C                     93.3 | 95.0 | 100 | 100 | 97.1 | 95.0 | 93.3| 100 | 100 | 97.1 | 95.0 | 95.0| 100 | 100 | 97.5 |
| D                     88.3 | 96.7| 93.3 | 90.0 | 92.1 | 93.3 | 91.7| 91.7 | 91.7 | 92.1 | 95.0 | 100 |96.7| 91.7 | 95.8 |
| E                     91.7 | 80.0| 91.7 | 85.0 | 87.1 | 86.7 | 81.7| 85.0 | 95.0 | 87.1 | 91.7 | 90.0 | 91.7 | 93.3 | 91.7 |
| F                     100 | 95.0 | 98.3| 100 | 98.3 | 98.3 | 91.7| 100 | 93.3 | 95.8 | 100 | 96.7 | 100.0 | 95.0 | 97.9 |
| G                     96.7 | 100 | 91.7 | 96.7 | 96.3 | 100 | 93.3 | 95.0 | 95.0 | 95.8 | 100 | 100 | 96.7 | 100 | 99.2 |
| H                     95.0 | 80.0| 85.0 | 90.0 | 87.5 | 88.3 | 85.0| 86.7 | 98.3 | 89.6 | 93.3 | 90.0 | 90.0 | 95.0 | 92.1 |
| MCA                   93.3 | 91.9| 92.9 | 93.3 | 92.9 | 92.1 | 90.2 | 93.8 | 95.0 | 92.8 | 95.2 | 95.4 | 96.0 | 96.3 | 95.7 |
| MS                    98.8 | 97.4| 97.6 | 96.7 | 97.6 | 98.9 | 97.2| 97.5 | 96.8 | 97.6 | 99.4 | 98.2 | 98.6 | 98.1 | 98.6 |

Table 5 Confusion matrix for the Csim in image recognition.
difference in swallowing patterns based on swallowing conditions can be classified from multichannel sEMG signals of the suprathyroid muscles and infrahyoid muscles, as a first step to develop a method to evaluate swallowing reserve. A total of four swallowing conditions combining two bolus volumes (3 and 15 mL water) and two techniques (normal and effortful swallow) were classified using three machine learning methods; namely, the real-time classification, comprehensive classification, and image recognition methods. In the real-time classification method, the SVM-based tongue motion classification method developed for mobility support of severely disabled people was applied to swallowing pattern classification. The MCA of swallowing conditions was as low as 81.5%, indicating that the difference between the swallowing conditions performed in a short time of about 1 s cannot be classified sufficiently by this method. In the comprehensive classification method, a majority decision is applied to all the classification results, i.e., predicted class labels, from the start to the end of swallowing obtained every 16 ms using tongue motion classification method. The MCA obtained by this method was 95.1%. In the image recognition method, the change of a series of sEMG signals in the swallowing movement was converted into a swallowing pattern image, and the images were classified using a combination of CNNs and SVM. Compared with the comprehensive classification, the number of training samples in the image recognition method was 1/26, but the MCA reached 95.7%. These results confirm that the differences in swallowing patterns can be classified with high accuracy from multichannel sEMG signals of the suprathyroid muscles and the infrahyoid muscles. Additionally, in the image recognition method, the MCA for Csm and Cim were both approximately 93%, indicating that the four swallowing conditions can be classified stably, even if sEMG signals of either the suprathyroid muscles or infrahyoid muscles are used.

Differences in bolus volume and physical properties of food cause changes in swallowing movement patterns induced by the swallowing reflex. Swallowing reserve is the ability to adjust the movement of organs for swallowing and to change the swallowing pattern for objects to be swallowed without choking and aspiration. Using the image recognition method, swallowing patterns could be classified with high accuracy of 95.7%, suggesting that the sEMG signals of the suprathyroid and infrahyoid muscles contain sufficient information regarding the changes in movement pattern of swallowing-related organs depending on swallowing conditions. From the results of the swallowing pattern classifications analyzed in this study, CA increases as the change in swallowing movement pattern becomes greater according to swallowing conditions. However, if swallowing reserve decreases, the changes in swallowing movement pattern are less likely to appear on the swallowing pattern image, and thus CA is expected to decrease. If this hypothesis is correct, decline in swallowing reserve may be detected early by focusing on the CA of swallowing patterns. To prove this hypothesis and establish an evaluation method for swallowing reserve, how the CA of swallowing patterns (obtained using the method proposed in this study) changes based on aspects such as age, gender, presence or absence of dysphagia, and type of dysphagia needs to be evaluated. This study classified swallowing patterns only in healthy young men with no history of dysphagia. Hence, measuring and analyzing swallowing patterns in elderly persons and dysphagic patients are an essential next step. For an in-depth analysis of individual swallowing reserve, swallowing pattern classification focusing on the changes in muscle activity of the entire anterior neck and detailed studies of physiological changes in each muscle are needed.

Acknowledgement

This study was supported by the Grant-in-Aid for Scientific Research (B), Japan Society of Promotion of Science (Principal Investigator: Makoto Sasaki, Grant Number: 18H03557).

References

1. Logemann JA: Evaluation and treatment of swallowing disorders, 2nd ed. Pro-Ed, Austin, 1998.
2. Dodds WJ, Man KM, Cook IJ, Kahrilas PJ, Stewart ET, Kern MK: Influence of bolus volume on swallow-induced hyoid movement in normal subjects. AJR Am J Roentgenol. 150(6), pp. 1307–1309, 1988.
3. Cook IJ, Dodds WJ, Dantas RO, Kern MK, Massey BT, Shaker R, Hogan WJ: Timing of videofluoroscopic, manometric events, and bolus transit during the oral and pharyngeal phases of swallowing. Dysphagia. 4(1), pp. 8–15, 1989.
4. Jacob P, Kahrilas PJ, Logemann JA, Shah V, Ha T: Upper esophageal sphincter opening and modulation during swallowing. Gastroenterology. 97(6), pp. 1469–1478, 1989.
5. Leonard RJ, Kendall KA, McKenzie S, Gonçalves MI, Walker A: Structural displacements in normal swallowing: a videofluoroscopic study. Dysphagia. 15(3), pp. 146–152, 2000.
6. Ueda N, Nohara K, Kotani Y, Tanaka N, Okuno K, Sakai T: Effects of the bolus volume on hyoid movements in normal individuals. J Oral Rehabil. 40(7), pp. 491–499, 2013.
7. Dantas RO, Dodds WJ, Massey BT, Kern MK: The effect of high- vs low-density barium preparations on the quantitative features of swallowing. AJR Am J Roentgenol. 153(6), pp. 1191–1195, 1989.
8. Ishida R, Palmer JB, Hiemae KM: Hyoid motion during swallowing: factors affecting forward and upward displacement. Dysphagia. 17(4), pp. 262–272, 2002.
9. Logemann JA, Pauloski BR, Rademaker AW, Colangelo LA,
Kahrilas PJ, Smith CH: Temporal and biomechanical characteristics of oropharyngeal swallow in younger and older men. J Speech Lang Hear Res. 43(5), pp. 1264–1274, 2000.

Ohmoe Y, Sugiuira M, Mogitite M: Swallowing function in the very old (85+ years): changes during swallowing physiology with normal aging. J Jpn Bronchoesophagol Soc. 54(1), pp. 1–7, 2003. (in Japanese)

Namasivayam-MacDonald AM, Barbón CEA, Steele CM: A review of swallow timing in the elderly. Physiol Behav. 184, pp. 12–26, 2018.

Perlman AL, Booth BM, Grayhack JP: Videofluoroscopic predictors of aspiration in patients with oropharyngeal dysphagia. Dysphagia. 9(2), pp. 90–95, 1994.

Power ML, Hamdy S, Goulermas JY, Tyrrell PJ, Turnbull I, AA, Zoratto DC: The relationship between hyoid and laryngeal closure. Dysphagia. 24(3), pp. 257–264, 2009.

Steele CM, Bailey GL, Chau T, Molfenter SM, Oshalla M, Waito K, Yüceyar N: Piecemeal deglutition and dysphagia. J Voice. 67(6), pp. 755–764, 2016.

Yamazaki Y, Tohara H, Hara K, Nakamura Y, Hayashi T: Development of a biometric system for the measurement of swallowing function. Japan Med Assoc J. 54(1), pp. 31–34, 2011.

Masahiro Suzuki, et al: Swallowing Pattern Classification Method

Masahiro Suzuki received B.E. degree from Iwate University in 2018. He is currently a student at Master’s Program in Graduate School of Science and Engineering, Iwate University. His research interests include biosignal processing, artificial intelligence, pattern recognition. He is a member of Japanese Society of Stomatognathic Function.

Makoto Sasaki

Makoto Sasaki received Ph.D. degree from Akita University in 2005. He is currently an associate professor at Graduate School of Science and Engineering, Iwate University. His research interests include dysphagia rehabilitation, biosignal processing, artificial intelligence, pattern recognition, and human computer interface. He is a member of Japanese Society for Medical and Biological Engineering, Japanese Society of Dysphagia Rehabilitation, Japanese Society of Stomatognathic Function, Society of Biomechanisms Japan, Society of Life Support Engineering, Japan Society of Mechanical Engineering, and IEEE. He received the Best Research Award at JSMBE symposium 2011 and 2019, respectively.
Katsuhiro Kamata
Katsuhiro Kamata received B.E. degree from Nippon Institute of Technology in 2009. He is currently a student at Doctoral Program in Graduate School of Science and Engineering, Iwate University. His research interests include dysphagia rehabilitation, biosignal processing, artificial intelligence, and embedded system. He is a member of Japan Society of Mechanical Engineering.

Atsushi Nakayama
Atsushi Nakayama received Ph.D. degree from Nagoya University in 2006. He is currently a professor at Department of Engineering for Future Innovation, National Institute of Technology, Ichinoseki College. His research interests include human machine interface, biosignal processing, and dysphagia rehabilitation. He is a member of Japanese Society of Mechanical Engineering, Society of Biomechanisms Japan, Japan Society for Clinical Biomechanics, and IEEE.

Isamu Shibamoto
Isamu Shibamoto received his Ph.D. degree from Tokyo Medical and Dental University in 2007. He currently works as a professor in Speech Language and Hearing Sciences at Graduate School of Rehabilitation Sciences, Seirei Christopher University. He is focusing on dysphagia rehabilitation, cortical representation and surface EMG analysis during swallow in his research. He is an editor and a member of Asia Pacific Society of Speech Language and Hearing, Japanese Society of Dysphagia Rehabilitation, and Society of Swallowing and Dysphagia of Japan.

Yasushi Tamada
Yasushi Tamada received Ph.D. degree from Iwate Medical University in 2011. He is currently an Assistant Professor at School of Dentistry, Iwate Medical University. His research interests include dysphagia rehabilitation, gerodontology and prostodontics. He is a member of Japanese Society of Stomatognathic Function, Japanese Society of Dysphagia Rehabilitation, Japanese society of gerodontology, Japan Prostodontic Society, Japanese Society for Parenteral and Enteral Nutrition.