International Journal of Neural Systems

A Swallowing Decoder based on Deep Transfer Learning: AlexNet Classification of the Intracranial Electrocorticogram

Hiroaki Hashimoto
Department of Neurological Diagnosis and Restoration, Graduate School of Medicine, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan
Department of Neurosurgery, Otemae Hospital
Chuo-ku Otemae 1-5-34, Osaka, Osaka 540-0008, Japan
Endowed Research Department of Clinical Neuroengineering,
Global Center for Medical Engineering and Informatics, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan

Seiji Kameda
Department of Neurological Diagnosis and Restoration, Graduate School of Medicine, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan

Hitoshi Maezawa
Department of Neurological Diagnosis and Restoration, Graduate School of Medicine, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan

Satoru Oshino
Department of Neurosurgery, Graduate School of Medicine, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan

Naoki Tani
Department of Neurosurgery, Graduate School of Medicine, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan

Hui Ming Khoo
Department of Neurosurgery, Graduate School of Medicine, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan

Takufumi Yanagisawa
Department of Neurosurgery, Graduate School of Medicine, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan

Toshiki Yoshimine
Endowed Research Department of Clinical Neuroengineering,
Global Center for Medical Engineering and Informatics, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan

Haruhiko Kishima
Department of Neurosurgery, Graduate School of Medicine, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan

Masayuki Hirata*
Department of Neurological Diagnosis and Restoration, Graduate School of Medicine, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan
Endowed Research Department of Clinical Neuroengineering,

Global Center for Medical Engineering and Informatics, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan
Department of Neurosurgery, Graduate School of Medicine, Osaka university
Yamadaoka 2-2, Suita, Osaka 565-0871, Japan

E-mail: mhirata@ndr.med.osaka-u.ac.jp

To realize a brain-machine interface to assist swallowing, neural signal decoding is indispensable. Eight participants with temporal-lobe intracranial electrode implants for epilepsy were asked to swallow during electrocorticogram (ECoG) recording. Raw ECoG signals or certain frequency bands of the ECoG power were converted into images whose vertical axis was electrode number and whose horizontal axis was time in milliseconds, which were used as training data. These data were classified with four labels (Rest, Mouth open, Water injection, and Swallowing). Deep transfer learning was carried out using AlexNet, and power in the high-γ band (75–150 Hz) was the training set. Accuracy reached 74.01%, sensitivity reached 82.51%, and specificity reached 95.38%. However, using the raw ECoG signals, the accuracy obtained was 76.95%, comparable to that of the high-γ power. We demonstrated that a version of AlexNet pre-trained with visually meaningful images can be used for transfer learning of visually meaningless images made up of ECoG signals. Moreover, we could achieve high decoding accuracy using the raw ECoG signals, allowing us to dispense with the conventional extraction of high-γ power. Thus, the images derived from the raw ECoG signals were equivalent to those derived from the high-γ band for transfer deep learning.

Keywords: swallowing, electrocorticogram, deep transfer learning, AlexNet, brain-machine interface, γ band

1. Introduction

Swallowing is a fundamental function of life and its disturbance (dysphagia) causes malnutrition and aspiration. Moreover, the latter can lead to aspiration pneumonia and high mortality rates.[1] Deaths from pneumonia are increasing rapidly due to population aging, and most cases of pneumonia in the elderly are aspiration pneumonia.[2] Stroke[3, 4] and neural degenerative disease[5, 6] also promote dysphagia. Adopting different postures and modifying the bolus consistency are widely accepted in clinical practice as treatments for dysphagia.[7] Peripheral electrical stimulation[8, 9] and non-invasive brain stimulation[10] are also under study as emerging therapeutic strategies. The final goal of our research is the realization of a swallow-assisting brain-machine interface (BMI) to control a device supporting swallowing, a novel approach to the treatment of dysphagia.

The decoding method is a key technology in the realization of such a swallow-assisting BMI. The prevailing view is that the brainstem is solely responsible for the control of swallowing.[11] However, non-invasive methods such as electroencephalography (EEG),[12, 13] positron emission tomography (PET),[14] near-infrared spectroscopy (NIRS)[15], transcranial magnetic stimulation (TMS),[16] functional magnetic resonance imaging (fMRI),[17-19] and magnetoencephalography (MEG),[20, 21] implicate multiple cortical sites in swallowing, including the sensorimotor cortex, the insula, the premotor cortex, the frontal operculum, the anterior cingulate gyrus,[11, 14, 17-21], etc. The lateral sensorimotor cortex, which corresponds to the orofacial cortex, is particularly involved.[22] Therefore, cortical signals could potentially provide features for the decoding of swallowing intention. Previous studies have detected swallowing motor imagery using scalp EEG signals.[12, 23, 24]

Studies with intracranial electrodes have shown that in machine learning, classification accuracy is better using high-γ-band (> 50 Hz) activities than using lower frequencies.[25] The γ band is a key oscillation that reflects the neural processing of sensory, motor, and cognitive events,[26, 27] and shows better functional localization than does lower-frequency activity.[26] Scalp EEG is not suitable for recording activities in high-γ band and so the high-γ activities must be recorded with intracranial electrodes that yield electrocorticograms (ECoGs). No studies using ECoG signals for the decoding of swallowing intention have been reported.

In previous BMI studies, a support vector machine (SVM)[25] and sparse linear regression (SLiR)[28] have been used to decode hand movements from intracranial

electrode data. Recently, a deep learning method has been used for decoding such signals,[29] but the decoder was trained on individual subjects. Thus, the ECoG datasets available for training were small, and a supervised learning model (i.e., an SVM) was used. However, if we want to use a deep learning method, a large dataset is indispensable, and deep learning from scratch in one individual is difficult. To solve this problem, we chose the approach of transfer learning.

The deep transfer learning method is currently used in medical imaging fields, such as abdominal ultrasound imaging[30], histological imaging[31], and computed tomography (CT) of the brain.[32] Deep learning from scratch is also effective in medical imaging[33]. In these studies, the training dataset is visually meaningful and transfer learning is effective. However, when working with ECoG signals, unaided visual inspection can detect epileptic activities, such as spiking, but cannot determine whether these signals contain motor or swallowing information. The published studies on deep learning of brain signals such as EEG and MEG have all carried out deep learning from scratch.[34-40] Therefore, it remains unclear whether transfer learning is feasible with a visually meaningless ECoG dataset that is only formally an image. In this study, we tested whether transfer learning for the decoding of swallowing intention can be effective with such images.

When a fully deep network structure such as a convolutional neural network (CNN) is trained with small numbers of data, problems of overfitting occur, leading to low performance metrics and low generalization power.[41] In the field of machine learning for the interpretation of medical images, the amount of data available for research is limited—a very common issue. This problem is solved by transfer learning. The concept of transfer learning is that a model already trained with larger datasets is re-used for a new but smaller dataset with which its pretrained learning parameters are shared. During adaptation to the new dataset, only the last fully connected layers are re-trained on the new dataset with initial random weights. Transfer learning can reduce training calculations and memory cost, and provide a new model with powerful feature extraction.[32] Transfer learning on medical images shows efficacies in accuracy, training time, and error rate.[30, 32] In this study, we used AlexNet for transfer learning to subsequently classify swallowing-related ECoG signals.

AlexNet is the first convolutional neural network to achieve the highest classification accuracy, which was recorded at the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012.[42] This deep structure has eight main layers. First five layers are convolutional layers, where the features of images are extracted. Each convolutional layer is followed by a layer of rectified linear units (ReLU), which apply an activation function. After each convolutional layer, a max pooling is used to reduce the network size. The last three layers are fully connected, and the output of the last layer produces 1000 class labels.

In this study, we aimed to decode swallowing intention using ECoG signals. We asked participants to swallow a water bolus at their own pace, and recorded their ECoG signals during this activity. To ensure a clear ECoG without contamination from myoelectric signals, we asked the participants to avoid all movements during the test except swallowing. We used a transfer learning method based on AlexNet for the investigation of swallowing decoding, where the last layer was replaced by four output layers, signaling rest, mouth opening, water injection, and swallowing, respectively. The ECoG signals were converted into 227 × 227 × 3 pixels, which were used as the training dataset. We explored several dataset options, such as using all implanted electrodes versus orofacial electrodes only and using raw ECoG signals versus the power in certain frequency bands. We hypothesized that AlexNet-based transfer learning using ECoG signals would be effective for the classification of swallowing intention (confirmed), and that if we selected the orofacial ECoG signals as features for classification rather than using all electrodes, the decoding accuracy would improve (confirmed). Moreover, we hypothesized that the γ band would be better for decoding than the raw ECoG signal (disconfirmed) or the lower-frequency bands (confirmed).

2. Methods

2.1. *Participants*

Eight patients with intractable epilepsy were recruited for this study (four males and four females, 15–51 years

(a)

Clinical profiles

Participant	Age/sex	Diagnosis	Number of Orofacial electrodes (Total Electrodes)	Number of Swallowing Instances	Number of Mouth Openings	Number of Water Injections
P1	36y / F	L TLE	18 (68)	31	31	31
P2	30y / F	L TLE	20 (84)	41	41	41
P3	18y / F	R TLE	18 (55)	27	27	27
P4	24y / M	L TLE	15 (69)	34	34	34
P5	51y / M	L TLE	20 (72)	38	38	38
P6	28y / M	R TLE	20 (96)	27	27	27
P7	20y / M	L TLE	20 (94)	33	33	33
P8	15y / F	L TLE	20 (68)	37	37	37

M, Male; F, Female; R, Right; L, Left; TLE, Temporal lobe epilepsy

(b)

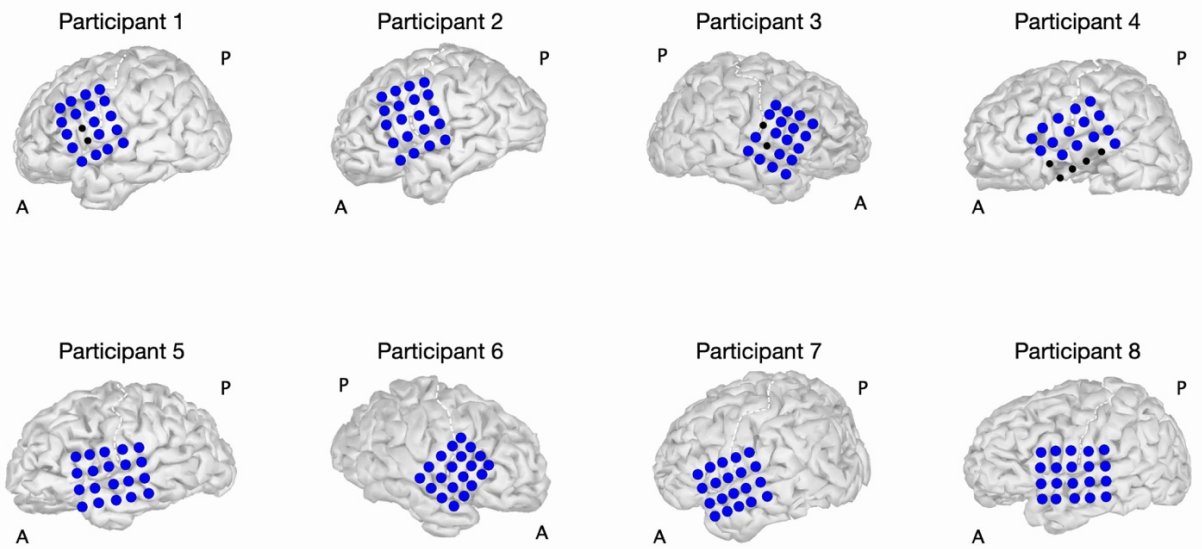

Fig. 1. Clinical profiles. (a) Numbers of orofacial and total electrodes and numbers of three different triggers in each participant are shown in the table. (b) The placements of the implanted electrodes are indicated. Black color indicates electrodes excluded from the analyses because of contamination such as severe noise or epileptic discharges. The central sulcus is indicated by the white dashed line. A, anterior; P, posterior.

of age: 27.8 ± 11.6 y) (Fig. 1a) from April 2014 to August 2019. All participants and their guardians were informed of the purpose and possible consequences of this study, and written, informed consent was obtained. The present study was approved by the Ethics Committee of Osaka University Hospital (Nos. 08061 and 16469).

2.2. *Electrode types and implantation sites*

Several types of electrodes (Unique Medical Co. Ltd., Tokyo, Japan), including grid, stripe, and depth types, were implanted in the subdural space during a conventional craniotomy that was part of clinical surgery for epilepsy. We selected patients with planar-surface platinum grid electrodes (4 × 5 array) placed

over the lateral portion of the central sulcus, which could be used as orofacial electrodes (Fig. 1b). The numbers of total implanted electrodes and orofacial electrodes in each participant are given in Fig. 1a (orofacial electrodes, 18.9 ± 1.8; all electrodes, 75.8 ± 14.3). The diameter of the electrodes was three or five mm, and the inter-electrode center-to-center distance was five, seven, or ten mm.

Preoperative structural magnetic resonance imaging (MRI) was obtained at 1.5 or 3.0 Tesla. Three-dimensional brain renderings were then created by FreeSurfer (https://surfer.nmr.mgh.harvard.edu) using the MRI volume. Images of the implanted electrodes were obtained from postoperative CT scans and were overlaid onto the 3D brain renderings, and the Montreal Neurological Institute (MNI) coordinates of the implanted electrodes were obtained with Brainstorm software (http://neuroimage.usc.edu/brainstorm/). The locations of the implanted electrodes were confirmed by intraoperative photographs.

2.3. *ECoG recording and data preprocessing*

The ECoG signals were measured with a 128-channel digital EEG system (EEG 2000; Nihon Kohden Corporation, Tokyo, Japan) and digitized at a sampling rate of 10,000 Hz. At this point, electrodes providing data contaminated with external noise or epileptic discharges were excluded from further analysis.

The ECoG signals were then down-sampled to 1,000 Hz using BESA Research software (BESA GmbH, Grafelfing, Germany), which was combined with passage through a bandpass filter (0.3–333 Hz) to prevent aliasing and a 60-Hz notch filter to eliminate the AC line noise. In each participant, the ECoG signals were then digitally re-referenced to a common potential averaged across all implanted electrodes.

2.4. *Behavioral procedure*

The experiments were performed approximately one week after surgical electrode placement when all participants had fully recovered from surgery. The participants were asked to sit in a chair and to remain still, and without moving their mouth in particular, for three minutes. We defined this period as the rest state. The participants were then instructed to open their mouths, and the examiner injected 2 mL of water into the mouth with a syringe. We requested that the participants swallow the water bolus at their own pace and without external cueing, to prevent erroneous volitional water swallowing (aspiration). To ensure acquisition of clear ECoGs without contamination from myoelectric signals, we asked the participants to avoid all movements during the test except swallowing. After we had confirmed that the participants had completed one swallowing movement, the next water bolus was administered.

2.5. *Monitoring of swallowing*

To ensure noninvasive monitoring, we used an electroglottograph (EGG), a microphone, and a motion-tracking system. Our EGG was a laryngograph (Laryngograph Ltd, London, UK) configured to record the neck impedance change due to swallowing movements (Fig. 2a).[43] A pair of electrodes was placed on the neck skin below the thyroid cartilage at an interelectrode center-to-center distance of 25 mm and was held in place by an elastic band. Sounds of swallowing due to the bolus passing through the pharynx were detected by a throat microphone (Fig. 2b).[44] We connected the throat microphone (Inkou mike; SH-12iK, NANZU, Shizuoka, Japan) to the laryngograph to record the swallowing sounds on the same trace. The shape of the microphone was arched to fit around the participant's neck. The sampling rate of the laryngograph and throat microphone was 24 kHz.

We captured the motion of the participants at 30 frames per second with a motion-tracking system newly developed by us using Kinect v2 (Microsoft, Redmond, Washington, USA), which we termed the simple swallow-tracking system (SSTS).[45] The participants were seated facing Kinect v2, which was placed on a tripod at a distance of one meter, and their mouth and throat movements were captured. An electric stimulator (NS-101; Unique Medical, Tokyo, Japan) supplied digital synchronizing signals to the laryngograph and a 128-channel digital EEG system. The signals were also converted into light flashes by an LED, which were

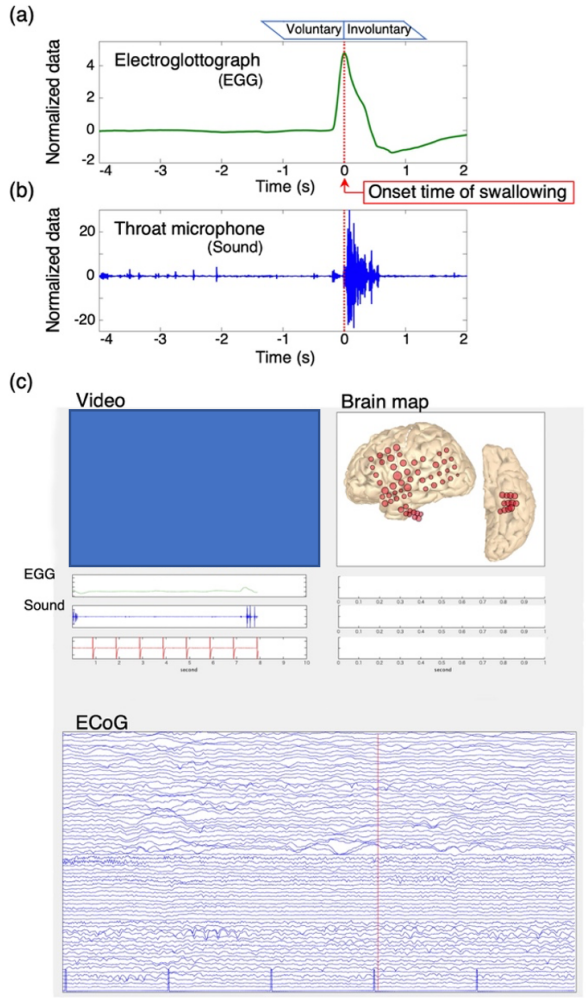

Fig. 2. Across-trials averaged impedance waveforms of an electroglottograph (a) and a throat microphone (b), all from one participant (P1). For analysis, the onset of swallowing was defined as the peak time of an impedance waveform that had been marked as a swallow. (c) The graphical user interface we used for synchronizing multimodal data during display.

captured by the RGB camera of the Kinect v2 system. The digital triggers and LED flashes enabled us to place in time registry the multimodal data streams produced by the EGG, microphone, motion-tracking system, and ECoG. To facilitate this operation, we programmed an original graphical user interface (GUI) in MATLAB (MathWorks, Natick, MA, USA) that enabled us synchronize the multimodal data streams and display them at the same time (Fig. 2c).

2.6. *Signal segmentation based on swallowing-related events*

2.6.1. Mouth-open triggers

All swallowing-related events were captured by the Kinect v2 camera.[45] Because of the precise synchronization of the video and ECoG signals, we could use the video to detect the time at which the participant opened their mouth, and we inserted mouth-open triggers into the ECoG data at that time.

2.6.2. Water-injection triggers

Using the video, we could also detect the time when the examiner injected water into the participants' mouth, and we inserted water-injection triggers into the ECoG data at that time.

2.6.3. Swallow triggers

Swallowing activities caused swallowing-related impedance changes in the neck, and this impedance waveform was clearly associated with the swallowing activities.[44] The laryngograph was used to identify the signal most reliably related to swallowing.[43] The swallowing onset time was determined visually at the time when the impedance waveform reached the peak (Fig. 2a). Swallowing sounds occurred frequently as the bolus of water passed through the pharynx,[44] and their evaluation in conjunction with the EGG helped us to judge whether the impedance change was caused by swallowing. Additionally, we confirmed that the changes in impedance and the sounds corresponded to water swallowing using the video stream captured by the Kinect v2. Based on these convergent data, we inserted swallow triggers, which were taken as the swallowing onset times, into the ECoG data.

The numbers of each type of trigger were the same for a given participant, as shown in Fig. 1a (average: 33.5 ± 5.1 times per participant for all triggers).

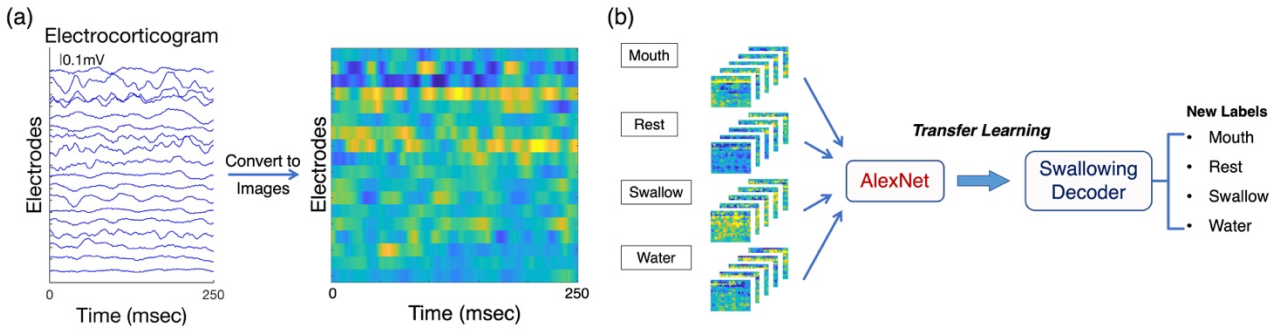

Fig. 3. Transfer deep learning with AlexNet. (a) Raw ECoG signals were converted into color-scaled images. Vertical axis, electrode number; horizontal axis, time (ms). (b) Image datasets labeled as mouth opening, rest, swallow, and water injection were created and used for transfer deep learning with AlexNet. The swallowing decoder classified the unlabeled data to four labels. The performance of transfer learning was evaluated by ten-fold cross-validation.

2.6.4. Training images

AlexNet assumes that the input data are images. Therefore, we converted the waveform data into color-scale images using the imagesc.m function in MATLAB R2019b. The vertical axis of an image is electrode number and the horizontal axis is time (Fig. 3a). The data were z-normalized. Data epochs 250 ms long were extracted starting 250 ms before each trigger and after each 100-ms interval thereafter. Five data of 250 ms were acquired at each trigger. One datum consisted of *n* electrodes × 250 data points. This was converted by MATLAB R2019b into a color-scaled image using the imagesc.m function and saved as a .jpg file of 875 × 1167 × 3 pixels (Fig. 3a). We prepared three image datasets using the mouth-open, swallow, and water-injection triggers, respectively. Moreover, the same number of images was also created from the rest-state data. Therefore, four types of labels (Mouth-open, Rest, Swallow, and Water-injection) were used. The number of images created were: 620 in participant one (P1); 820, P2; 540, P3; 680, P4; 760, P5; 540, P6; 660, P7; and 740, P8 (mean ± standard deviation (SD), 670 ± 101.4).

Two types of features were used for preparing the training images, namely the raw ECoG signals and the power in certain frequency bands. A bandpass filter using a two-way least-squares finite impulse response (FIR) filter (pop_eegfiltnew.m in the EEGLAB toolbox) was applied to the total ECoG signal before data extraction (δ band, 1–4 Hz; θ band, 4–8 Hz; α band, 8–13 Hz; β band, 13–30 Hz; low-γ band, 30–50 Hz; high-γ band, 75–150 Hz). To power signals, we used the Hilbert transform.[46] We created several alternative decoders trained with all implanted electrodes, with only the orofacial electrodes, or with the orofacial electrodes less one.

2.7. *Spectral analysis*

The ECoG signals of each participant were time-locked to each trigger, defined as 0 s, and extracted. A time–frequency analysis of the time-locked ECoG signals was performed in EEGLAB version 14.1.2b (http://sccn.ucsd.edu/eeglab/) with the frequency range set at 1–200 Hz and the spectral power in dB calculated in 1-Hz bins with 200 data points (*i.e.*, from -1.5 to 2.0 s in every 17-ms window). The baseline for the time–frequency analysis of each electrode was the initial 0.5 s. The power spectrum was computed using the short-time Fourier transform[47] over a sliding latency window and averaged across data trials.

2.8. *Deep transfer learning with AlexNet*

2.8.1. Transfer learning

We used the open source deep learning framework AlexNet[42] with MATLAB for training and testing the neural networks. All training and testing were

performed on a Windows 64-bit workstation with a 2.3 GHz six-core Inter Xeon central processing unit, 48 GB of memory, and an NVIDIA GeForce GTX 1080 graphical processing unit.

We used publicly available weights for AlexNet. The final fully connected layer was replaced with four outputs layers corresponding to the four image categories (*i.e.*, mouth open, rest, swallow, and water injection) (Fig. 3b), and initialized with random weights. The final layer was replaced with a classification output layer. For training, the weights for the five convolutional layers were frozen to extract features, and the learning rates for the fully connected layers were fixed at 0.001.

The training images of 875 × 1167 × 3 pixels were re-sized to 227 × 227 × 3 pixels for data input. For transfer learning, we re-trained the network with MATLAB. We used the stochastic gradient descent with momentum optimizer.[48] Swallowing decoders were trained either with the image data sets of individual participants or with the total image data set of all participants.

2.8.2. Model testing

The performances of the swallowing decoders were evaluated by 10-fold cross-validation, that is, by randomly splitting the image datasets into training data and testing data. In each validation, we calculated accuracy, and sensitivity.

2.9. *Comparison with an SVM decoding model*

To compare the accuracy of transfer learning with a reference method, we used a multi-class SVM to classify the high-γ power images to the four labels. From the high-γ power signals calculated by the Hilbert transform, 250-ms-long data epochs were extracted starting 250 ms before each trigger and after each 100-ms interval thereafter. The epochs for each orofacial electrode were summed and the result treated as a decoding feature. The number of decoding features were 744 in P1, 984 in P2, 648 in P3, 816 in P4, 912 in P5, 648 in P6, 792 in P7, and 888 in P8 (mean ± SD, 804 ± 121.7). The performances of the SVM decoders were evaluated by 10-fold cross-validation.

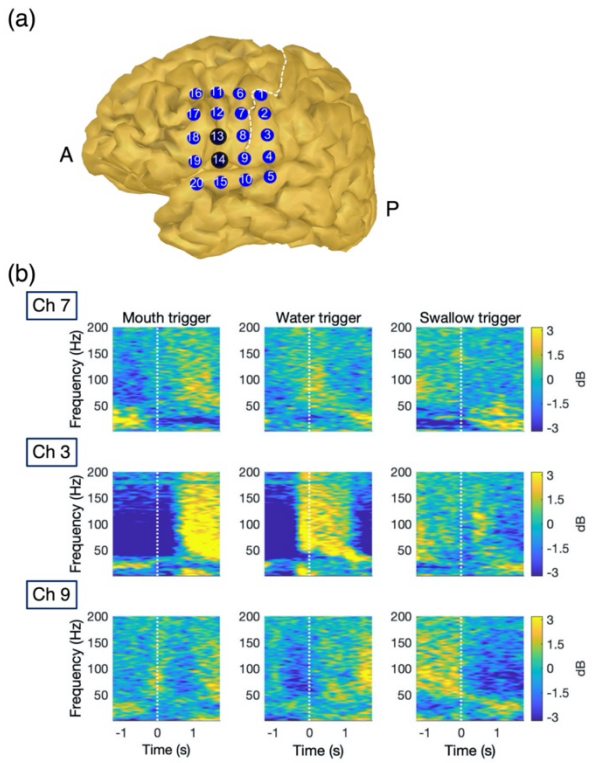

Fig. 4. Oscillatory activities evoked by mouth opening, water injection, and swallowing in the cerebral cortex in a representative participant (P1). (a) Shown is the reconstructed MRI with orofacial electrodes. The numbers correspond to the electrode numbers. The central sulcus is indicated by the white dashed line. (b) Time–frequency maps are shown from -1.5 to 2.0 s around the mouth-open, water-injection, and swallow triggers. The mouth-open-triggered time–frequency plot from Ch 7, which was attached to the precentral gyrus, shows decreases in β-band power at 0 s. Notable power increases in the high-γ band are also observed at 0 s in the water-injection-triggered time–frequency plot from Ch 3, which was attached to the postcentral gyrus. High-γ power increases specific to swallowing appear from -1.5 s to 0 s in Ch 9, which was attached to the subcentral area. A, anterior; P, posterior; Ch, channel.

2.10. *Statistics*

For the statistical evaluation of the swallowing decoders, we used the single-sided Wilcoxon rank-sum test in individual analyses, and the single-sided Wilcoxon

signed-rank test in group analyses. We applied a conservative Bonferroni correction to correct for multiple comparisons and used a corrected significance threshold of $p < 0.05$.

3. Results

3.1. *Time–frequency analyses*

The time–frequency maps (TFM) were calculated from the orofacial electrodes using mouth-open, water-injection, and swallow triggers in a representative participant (P1). The third electrode (Channel 3; Ch 3), Ch 7, and Ch 9 were placed on the postcentral gyrus, the precentral gyrus, and the subcentral area, respectively. The last is a narrow gyrus between the caudolateral extreme of the central sulcus and the lateral sulcus (Fig. 4a). In the P1 mouth-open data, the low frequency bands of less than 30 Hz were depressed in the precentral gyrus at 0 s (Fig. 4b: TFM of Ch 7 with the mouth-open trigger). When the water bolus was injected into the mouth, obvious high-γ activities were observed in the postcentral gyrus at 0 s (Fig. 4b: TFM of Ch 3 with the water-injection trigger). High-γ activities were localized in the subcentral area immediately before the swallow trigger (Fig. 4b: TFM of Ch 9 with the swallow trigger).

From mouth opening to swallowing, several oscillatory activities were observed with different timings in different locations. Almost all participants exhibited generally the same spatiotemporal profiles. This result implies that the orofacial electrodes were informative for decoding the timing of mouth opening, water injection, and swallowing.

3.2. *Decoding using the raw ECoG signal from orofacial electrodes*

3.2.1. Orofacial electrodes vs. all electrodes

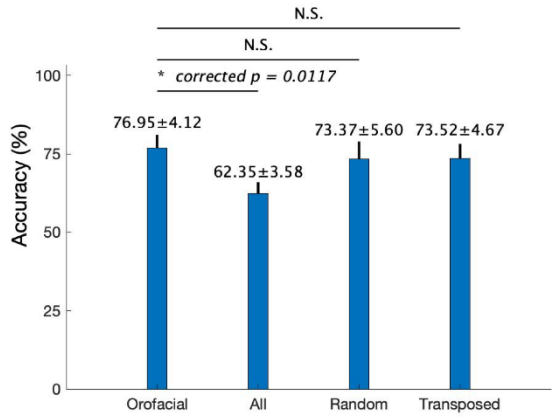

Fig. 5. Classification accuracy for raw ECoG signals. The inclusion of only orofacial electrodes in creating the training image dataset allows significantly higher accuracy than using all electrodes (single-sided Wilcoxon signed-rank test with Bonferroni correction, corrected $p < 0.05$). Changing the order in which the electrodes are presented on the vertical axis ("Random") and transposing the vertical and horizontal axes ("Transposed") has no influence on decoding accuracy (single-sided Wilcoxon signed-rank test with Bonferroni correction). N.S., not significant.

.

Training images generated from the raw ECoG signal were used to prepare four types of dataset. Two of them were derived from orofacial electrodes only and from all electrodes. Transfer learning was carried out with each participant, and the individual classification accuracies for labelling the four different types of images were calculated. The average accuracy of the orofacial-electrodes group was significantly higher than that of the all-electrodes group (76.95 ± 4.12 vs. 62.35 ± 3.58, mean ± standard error (SE), corrected $p = 0.0117$, single-sided Wilcoxon signed-rank test) (Fig. 5). Therefore, the exclusion of the non-orofacial electrodes from the dataset results in more effective classification than using all electrodes.

3.2.2. Changing the order of the electrodes

The vertical axis of the training images indexed the electrodes used (Fig. 3a). By default, the electrodes were arranged in ascending order from the top of the

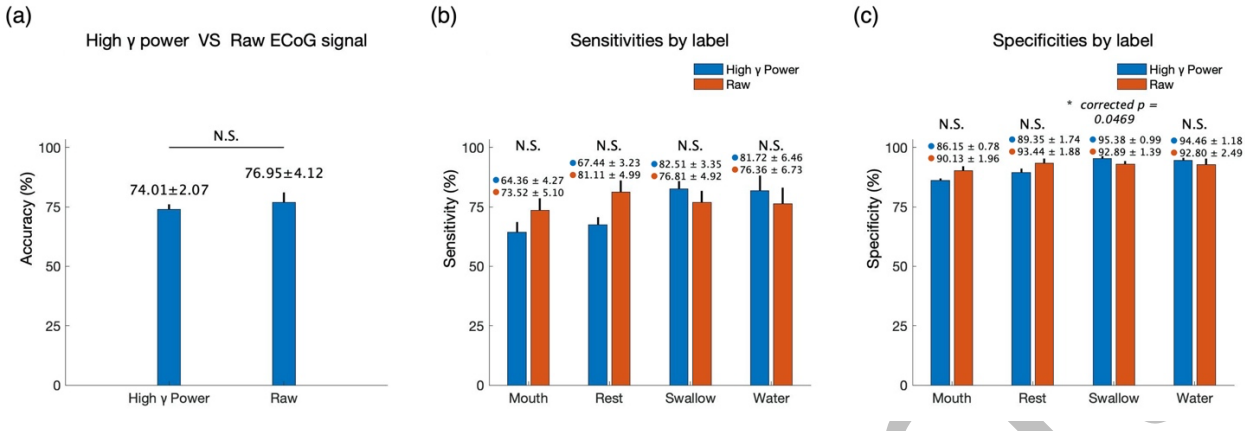

Fig. 6. A performance comparison between use of high-γ power and raw ECoG signals for training. (a) No significant accuracy difference is observed between the high-γ power and raw ECoG groups (single-sided Wilcoxon signed-rank test). (b) No significant sensitivity difference is observed in any labeling group between high-γ power and raw ECoG signals (single-sided Wilcoxon signed-rank test with Bonferroni correction). (c) A significant specificity difference is observed in the swallowing group between high-γ power and raw ECoG (single-sided Wilcoxon signed-rank test with Bonferroni correction, corrected $p < 0.05$).

vertical axis (Normal order). We evaluated whether electrode order was meaningful for decoding. The accuracy calculated from the normal-order dataset ("Orofacial" in Fig. 5) was compared with that calculated from a random-order dataset ("Random" in Fig. 5). Raw ECoG signals were used, and transfer learning was done in each participant individually. We found no significant differences. Therefore, the order of arrangement of the electrode data in the training images does not influence decoding accuracy.

3.2.3. Effect of transposing the axes of the training images

By default, the vertical axis indexed the electrodes used and the horizontal axis was time in milliseconds (Fig. 3a). We evaluated whether image orientation was meaningful for decoding. We transposed the vertical and horizontal axes ("Transposed" in Fig. 5) and compared the resulting accuracy with that of the default dataset ("Orofacial" in Fig. 5). Raw ECoG signals were used, and transfer learning was performed in each participant individually. We found no significant differences. Therefore, whether the vertical or horizontal axis is an electrode index or time in milliseconds does not influence decoding accuracy.

3.3. *High-γ power vs. raw ECoG signal*

Accuracy calculated after training with high-γ power datasets was compared with that calculated after training with datasets derived from the raw ECoG signals. Transfer learning was done in each participant. We found no significant differences (Fig. 6a). We then evaluated the two groups for sensitivity for each label and found no significant differences using a Bonferroni correction (Fig. 6b). Next, we evaluated the two groups for specificity and found a significant difference in the Swallowing label (corrected $p = 0.0469$, single-sided Wilcoxon signed-rank test with Bonferroni correction) (Fig. 6c). The sensitivity and specificity for swallowing calculated from high-γ power were higher than those calculated from the raw ECoG.

3.4. *Frequency band analyses*

The results calculated with datasets derived from the power feature in six different frequency bands were compared for accuracy. Transfer learning was done in each participant individually. The maximum values of accuracy were obtained from the high-γ band (74.01% ± 2.07% [mean ± SE]) and the differences from the other bands were significant ($p = 0.0195$, single-sided

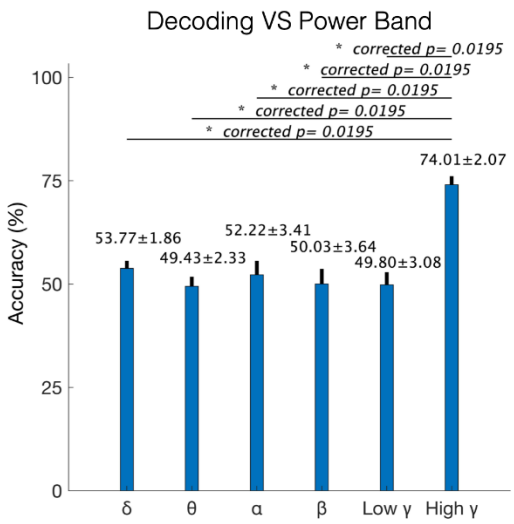

Fig. 7. Shown are the results of transfer learning using images derived from the high-γ band. The accuracy of classification resulting from transfer learning using images derived from the power feature of the high-γ band is significantly higher than that of the power feature in other bands (single-sided Wilcoxon signed-rank test with a Bonferroni correction).

Wilcoxon signed-rank test with a Bonferroni correction) (Fig. 7). Therefore, high-γ power was informative for decoding.

3.5. *Electrodes informative for classification, by label type*

For mouth-open, water-injection, and swallow labels in each participant, we compared sensitivities between two image datasets calculated from high-γ power. One dataset comprised images derived from the orofacial electrode data, and the other comprised images derived from all orofacial electrodes but one. We repeated this analysis for each possible choice of the deleted electrode. In P1, for example, the participant had 18 implanted electrodes, leading to 18 comparisons. In each label, we evaluated how much the sensitivities were affected by the deletion, using the single-sided Wilcoxon rank-sum test with a Bonferroni correction to decide significance.

3.5.1. Results for a representative participant (P1)

In P1, electrode deletions resulting in decreased mouth-open label sensitivity were located mainly on the superior temporal gyrus (STG) and anterior to the central sulcus. For the water-injection label, the important electrodes were located on the lateral postcentral gyrus, and for the swallow label, along or close to the Sylvian fissure (Fig. 8a). There were no significant changes in all label, however. The data indicated the different distribution patterns of electrodes

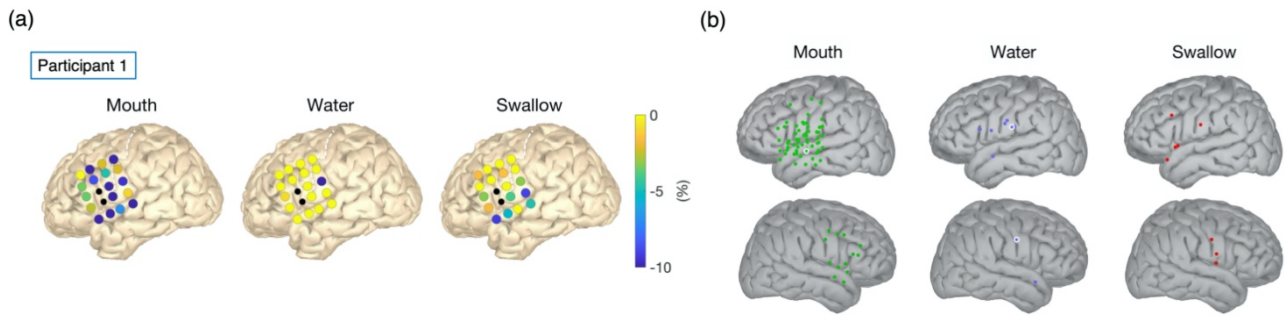

Fig. 8. Shown are electrodes informative for decoding. (a) In representative participant 1, the percent decrease in sensitivity resulting from exclusion of a given electrode's data from the analysis is shown on that electrode's location. In mouth-open decoding, a greater than ten percent decrease is observed in the superior temporal gyrus in regions anterior to the central sulcus. In water-injection decoding, greater than ten percent decreases are confined to the postcentral gyrus. In swallow decoding, greater than ten percent decreases are observed in regions along the Sylvian fissure. (b) Electrodes showing greater than ten percent decreases or the greatest decrease of all participants in each label are shown in overlay on the MNI standard brain. Top row, left hemisphere; bottom row, right. The distribution of electrodes informative for mouth opening is diffuse. Four out of nine electrodes informative for water injection are located on the postcentral gyrus. Electrodes informative for swallowing lie along the Sylvian fissure. Significant electrodes are indicated with white circles (single-sided Wilcoxon rank-sum test with Bonferroni correction).

informative for classification among the three labels.

3.5.2. Electrodes informative for classification in terms of the MNI standard brain

In each participant, the electrodes informative for classification, defined as those resulting in a greater than ten percent decrease or the greatest decrease in sensitivity when excluded from the analysis, were overlain on the Montreal Neurological Institute (MNI) standard brain (Fig. 8b). The MNI coordinates of the electrodes were obtained for all participants. For mouth opening, the distribution of informative electrodes was diffuse. One electrode located on the STG showed significance (p = 0.0313, single-sided Wilcoxon rank-sum test with a Bonferroni correction). Four electrodes out of nine informative for water injection were located on the postcentral gyrus, and two electrodes showing significance were located on the precentral and postcentral gyri (p = 0.0026 and 0.0313, respectively, single-sided Wilcoxon rank-sum test with a Bonferroni correction). The electrodes informative for swallowing distributed along the Sylvian fissure, including the subcentral area and the frontal operculum. However, no significant changes in sensitivity were observed for this label type.

3.6. ***Individual decoder vs. total decoder***

Using high-γ power, we compared the classification accuracy of transfer learning between that in an individual and that in the total participant sample. The average accuracy obtained from the individualized decoders was greater than that obtained from training one decoder with the all-images dataset. However, the observed difference did not reach significance (single-sided Wilcoxon rank-sum test) (Fig. 9).
The decoder in this study took average 0.02 s to decode one image file.

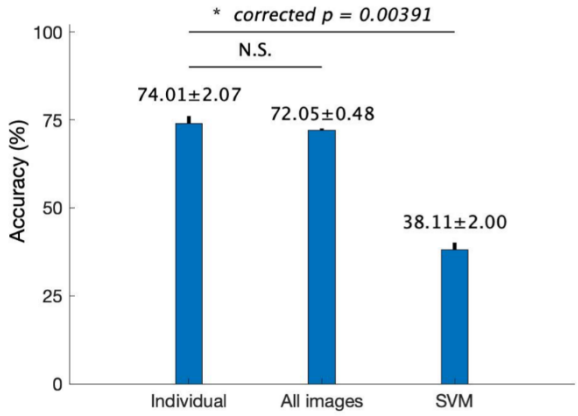

Fig. 9. Comparison of classification accuracy between individualized transfer learning and transfer learning using all images. No significant difference is observed (single-sided Wilcoxon rank-sum test with Bonferroni correction). N.S., not significant. The accuracy obtained from the SVM model was significantly less than that obtained from individual transfer learning (single-sided Wilcoxon signed-rank test with Bonferroni correction, corrected $p < 0.05$)

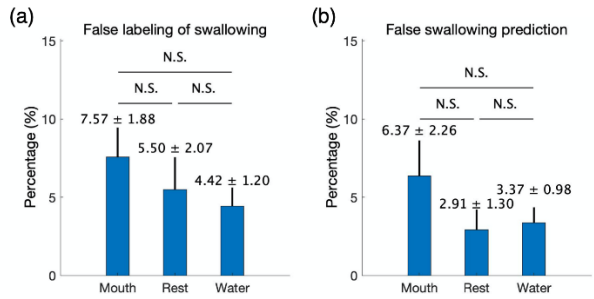

Fig. 10. (a) Type II error rate, Swallowing label. (b) Type I error rate, Swallowing label. (single-sided Wilcoxon signed-rank test with Bonferroni correction) N.S., not significant.

3.7. ***The deep learning decoder versus an SVM decoder***

Using high-γ power calculated from orofacial electrodes, we compared classification accuracy between transfer learning ("Individual" in Fig. 9) and the SVM model. The accuracy obtained from transfer learning was significantly greater than that of the SVM (corrected p = 0.004, single-sided Wilcoxon signed-rank test with Bonferroni correction) (Fig. 9).

3.8. *False labeling and false prediction for swallowing*

Using a decoder trained on high-γ power recorded from orofacial electrodes, we evaluated the type II error rate for Swallowing (the decoder misses the swallowing event). Figure 10a shows the percentage of swallowing data that were incorrectly labeled as Mouth, Rest, or Water. Figure 10b shows the percentage incorrect among swallowing predictions, where Mouth, Rest, or Water labels are misclassified as Swallowing (type I errors). The results show a tendency for the decoder to incorrectly label Swallowing as Mouth open and Mouth open as Swallowing. However, these tendencies did not reach significance (single-sided Wilcoxon signed-rank test with Bonferroni correction).

4. Discussion

In this study, deep transfer learning by AlexNet was used for classification of ECoG signals by four data labels, namely Rest, Mouth open, Water injection, and Swallowing. First, we selected raw ECoG signals as training features, and using only the orofacial electrodes, achieved higher accuracy than we did using all electrodes. We confirmed that in the training images, the order of the electrodes indexed by the vertical axis, and which variable, time or electrode, was plotted on the vertical axis, had no influence on decoding accuracy (Fig. 5). Next, we compared between raw ECoG signals and power in the high-γ band. Accuracy was equivalent in the two groups. However, using high-γ power yielded higher sensitivity and specificity for the Swallowing and Water labels (Fig. 6). High-γ power also yielded significantly greater accuracy than did lower-frequency power bands (Fig. 7). The above results were obtained with individual-specific transfer learning. Finally, with high-γ power, transfer learning with the all-participants dataset showed accuracy equivalent to that of individual-specific transfer learning. Our transfer-learning model also demonstrated significantly higher accuracy than did an SVM (Fig. 9).

AlexNet was intended to be trained by meaningful images. However, the study showed that nonvisual images derived from ECoG signals were suitable as training images for transfer learning. Moreover, including only data from electrodes on the orofacial cortex in the images yielded significantly greater decoding accuracy than using the data from all electrodes. The interesting finding of this study is that the accuracy achieved using raw ECoG signals for deriving the images was comparable to that achieved when using only the power feature of the high-γ band (75–150 Hz).

Previous studies have established the feasibility of transfer learning for classification of medical images, such as those obtained with ultrasound,[30] MRI,[49] and CT.[32] These training images are meaningful and can be classified visually. However, the training images used in this study were derived from ECoG signals had no visually meaningful content. Our study showed that AlexNet-based transfer learning works well with visually meaningless images.

If an individual-specific decoder based on a CNN is trained up from scratch for use as a BMI, the small dataset will be a problem. Even if that problem were overcome, the relatively long training time would still be an issue. Transfer learning has the potential to solve both these problems.

The orofacial cortex, a lateral region of the central sulcus, is activated by mouth movements,[50, 51] tongue movements,[20, 22, 52] and swallowing.[17-19, 53] In the present study, participants first opened the mouth, then the experimenter injected a water bolus into the mouth, then the participants swallowed. Our time–frequency maps showed dynamic oscillatory changes in the orofacial cortex from the time of mouth opening to swallowing. Suppression of lower-frequency bands, including the β band, were observed in the precentral gyrus at mouth opening. Obvious high-γ activities appeared in the postcentral gyrus when the water bolus was injected, and swallowing was accompanied by high-γ activities in the subcentral area. Since the different activities involved in swallowing were all evoked in the orofacial area, we inferred that exclusion of non-orofacial electrodes from the training data for transfer learning could potentially achieve higher accuracy than using all electrodes.

Previous studies have shown that decoding using high-γ ECoG activity is more accurate than using ECoG bands of lower frequency,[25, 54] which agrees with our results. Transfer learning using high-γ power showed significantly greater accuracy than using lower-frequency power bands, and we consider that this result

reflects a relatively more focal spatiotemporal distribution and functional localization of the high-γ activity.[26, 55]

Contrary to our expectations, transfer learning using high-γ features showed no greater decoding accuracy than that using the raw ECoG signals. In this study, signals to be used as training data were converted into images, and we infer that the images derived from the raw ECoG signals contained all the information of the high-γ band. Therefore, the achievement of high decoding accuracy without preprocessing for extraction of high-γ features is feasible.

Moreover, we investigated for each label the percent decrease in classification sensitivity occasioned by the exclusion of the data from a particular electrode. We inferred that the higher the percent decrease, the more important the removed electrode was for classification of a certain label, and we term the most important electrodes the informative electrodes. The distribution of informative electrodes was different for each labeling.

Electrodes informative for water injection and swallowing were distributed mainly on the postcentral gyrus and in regions along the Sylvian fissure, respectively, and these distributions agreed with the somatotopy of water injection and swallowing.[20] These results were obtained after training on the power feature of the high-γ band, indicating that specific high-γ activities were evoked in these regions by water injection and swallowing, respectively, and that these activities are effective for classification. Moreover, the sensitivities of swallowing and water injection after transfer learning using high-γ power were higher than those obtained using the raw ECoG signal. This result might also indicate that swallowing and water injection evoked specific high-γ activities in the cortex. We demonstrated that the order of presentation of the electrode data along the vertical axis of the training images had no influence on decoding accuracy. When we transposed the vertical and horizontal axes, we likewise observed no change in accuracy. However, for achieving high accuracy, it is important that the informative electrodes be included in the training images.

On the other hand, the distribution of electrodes informative for mouth opening was diffuse (*i.e.*, not localized). This result indicates that the high-γ activities related to mouth opening are not better localized than those related to water injection or swallowing. The number of electrodes informative for classification of mouth opening events was greater than for water injection or swallowing, and we infer from this that the decoding of mouth opening using high-γ power is not robust and is relatively difficult compared with analogous classifications to the other labels.

Moreover, our results showed a tendency for the decoder to confuse swallowing with mouth opening. This result may reflect the overlap of neural function between mouth opening and swallowing because the same oral muscles are used for both. In daily situations, we usually swallow after mastication, which underscores the need for a precise classification of swallowing and mouth movement in a swallow-assisting BMI. Moreover, chewing causes noise contamination of brain signals because of masticatory muscular activity. However, we infer that this problem may be solved if the decoder could distinguish myoelectric from brain signals, using a training set of electromyographic recordings. The oral muscles are also involved whenever we speak or make a facial expression. The orofacial cortex is also activated by speaking[56, 57]. The problem of distinguishing between speech and swallowing will be our future theme and we consider that the transfer learning approach has potential in that area.

Transfer learning using all 5360 training images showed a decoding accuracy equivalent to that using individual images (670 ± 101.4 per participant). This indicates that obtaining a generalized swallowing decoder is feasible. However, our results were obtained with data from the ECoG, which is an invasive technology. For realizing a swallow-assisting BMI, a non-invasive method of recording brain signals is indispensable. Moreover, our results were obtained by offline analysis, and the next step will be to move to online analysis, which is needed for evaluating the feasibility of this BMI project. Moreover, deep learning models are data-driven systems; that is, accuracy improves with the number of training images. In this study, the training images were few. Therefore, further studies with greater numbers of training images are needed. We confirmed that transfer learning with ECoG training datasets and based on AlexNet is effective. However, use of a higher-performance CNN such as VGG[58] might enable the achievement of still higher decoding accuracies.

5. Conclusions

In this study, we first demonstrated that deep transfer learning trained by ECoG signals enables classification of signals indicating either rest, mouth opening, water injection, or swallowing. We also demonstrated that AlexNet, which has been trained with a large dataset of visually meaningful images, can be used for transfer learning trained by visually meaningless images derived from ECoG signals. When high-γ power time series and raw ECoG signals were converted into training images, the best accuracy achieved was 74–76%, and our study showed that the raw ECoG signal is comparable to the high-γ activities for this purpose. In particular, the best sensitivity and specificity of swallowing classification that we obtained (82% and 95% respectively) was achieved using high-γ power training images. In conclusion, transfer learning trained by signals obtained from the ECoG is efficient for classification of swallowing.

6. Acknowledgments

This work was supported by JSPS KAKENHI [Grant nos. JP26282165, JP18H04166, and JP18K18366], by the Ministry of Internal Affairs and Communications, and by a grant from the National Institute of Information and Communications Technology (NICT).

7. Author Contributions

MH designed the study. HH performed the experiments, assisted with the medical treatments for epilepsy, created the MATLAB program, analyzed the data, created all figures, and was responsible for writing the manuscript. SK, HM, and MH assisted in the acquisition of the data. SO, NT, HMK, TaYa, HK, and MH clinically cared for and evaluated the patients. ToYo, HK, and MH supervised the study.

8. References

1. C. Cabib, O. Ortega, H. Kumru, E. Palomeras, N. Vilardell, D. Alvarez-Berdugo, D. Muriana, L. Rofes, R. Terre, F. Mearin and P. Clave, Neurorehabilitation strategies for poststroke oropharyngeal dysphagia: from compensation to the recovery of swallowing function, *Ann N Y Acad Sci* **1380**(1) (2016) 121-138.
2. S. Ebihara, H. Sekiya, M. Miyagi, T. Ebihara and T. Okazaki, Dysphagia, dystussia, and aspiration pneumonia in elderly people, *J Thorac Dis* **8**(3) (2016) 632-9.
3. D. G. Smithard, N. C. Smeeton and C. D. Wolfe, Long-term outcome after stroke: does dysphagia matter?, *Age Ageing* **36**(1) (2007) 90-4.
4. P. E. Marik, Aspiration pneumonitis and aspiration pneumonia, *New England Journal of Medicine* **344**(9) (2001) 665-671.
5. V. Boccardi, C. Ruggiero, A. Patriti and L. Marano, Diagnostic Assessment and Management of Dysphagia in Patients with Alzheimer's Disease, *J Alzheimers Dis* **50**(4) (2016) 947-55.
6. M. P. Jani and G. B. Gore, Swallowing characteristics in Amyotrophic Lateral Sclerosis, *NeuroRehabilitation* **39**(2) (2016) 273-6.
7. R. Speyer, L. Baijens, M. Heijnen and I. Zwijnenberg, Effects of therapy in oropharyngeal dysphagia by speech and language therapists: a systematic review, *Dysphagia* **25**(1) (2010) 40-65.
8. S. Suntrup, I. Teismann, A. Wollbrink, M. Winkels, T. Warnecke, C. Pantev and R. Dziewas, Pharyngeal electrical stimulation can modulate swallowing in cortical processing and behavior - magnetoencephalographic evidence, *Neuroimage* **104**(2015) 117-24.
9. C. L. Ludlow, I. Humbert, K. Saxon, C. Poletto, B. Sonies and L. Crujido, Effects of surface electrical stimulation both at rest and during swallowing in chronic pharyngeal Dysphagia, *Dysphagia* **22**(1) (2007) 1-10.
10. E. Verin and A. M. Leroi, Poststroke dysphagia rehabilitation by repetitive transcranial magnetic stimulation: a noncontrolled pilot study, *Dysphagia* **24**(2) (2009) 204-10.
11. C. Ertekin and I. Aydogdu, Neurophysiology of swallowing, *Clinical Neurophysiology* **114**(12) (2003) 2226-2244.
12. H. Yang, C. Guan, K. S. Chua, S. S. Chok, C. C. Wang, P. K. Soon, C. K. Tang and K. K. Ang, Detection of motor imagery of swallow EEG signals based on the dual-tree complex wavelet transform and adaptive model selection, *J Neural Eng* **11**(3) (2014) 035016.
13. I. Jestrovic, J. L. Coyle and E. Sejdic, Differences in brain networks during consecutive swallows detected using an optimized vertex-frequency algorithm, *Neuroscience* **344**(2017) 113-123.

14. S. Hamdy, J. C. Rothwell, D. J. Brooks, D. Bailey, Q. Aziz and D. G. Thompson, Identification of the cerebral loci processing human swallowing with H2(15)O PET activation, *J Neurophysiol* **81**(4) (1999) 1917-26.
15. S. E. Kober and G. Wood, Changes in hemodynamic signals accompanying motor imagery and motor execution of swallowing: a near-infrared spectroscopy study, *Neuroimage* **93 Pt 1**(2014) 1-10.
16. S. Hamdy, Q. Aziz, J. C. Rothwell, K. D. Singh, J. Barlow, D. G. Hughes, R. C. Tallis and D. G. Thompson, The cortical topography of human swallowing musculature in health and disease, *Nat Med* **2**(11) (1996) 1217-24.
17. S. Hamdy, D. J. Mikulis, A. Crawley, S. Xue, H. Lau, S. Henry and N. E. Diamant, Cortical activation during human volitional swallowing: an event-related fMRI study, *Am J Physiol* **277**(1 Pt 1) (1999) G219-25.
18. R. E. Martin, B. G. Goodyear, J. S. Gati and R. S. Menon, Cerebral cortical representation of automatic and volitional swallowing in humans, *J Neurophysiol* **85**(2) (2001) 938-50.
19. J. A. Toogood, R. C. Smith, T. K. Stevens, J. S. Gati, R. S. Menon, J. Theurer, S. Weisz, R. H. Affoo and R. E. Martin, Swallowing Preparation and Execution: Insights from a Delayed-Response Functional Magnetic Resonance Imaging (fMRI) Study, *Dysphagia* (2017).
20. P. L. Furlong, A. R. Hobson, Q. Aziz, G. R. Barnes, K. D. Singh, A. Hillebrand, D. G. Thompson and S. Hamdy, Dissociating the spatio-temporal characteristics of cortical neuronal activity associated with human volitional swallowing in the healthy adult brain, *Neuroimage* **22**(4) (2004) 1447-55.
21. R. Dziewas, P. Soros, R. Ishii, W. Chau, H. Henningsen, E. B. Ringelstein, S. Knecht and C. Pantev, Neuroimaging evidence for cortical involvement in the preparation and in the act of swallowing, *Neuroimage* **20**(1) (2003) 135-44.
22. S. Brown, E. Ngan and M. Liotti, A larynx area in the human motor cortex, *Cerebral Cortex* **18**(4) (2007) 837-845.
23. I. Jestrovic, J. L. Coyle and E. Sejdic, Decoding human swallowing via electroencephalography: a state-of-the-art review, *J Neural Eng* **12**(5) (2015) 051001.
24. H. Yang, K. K. Ang, C. Wang, K. S. Phua and C. Guan, Neural and cortical analysis of swallowing and detection of motor imagery of swallow for dysphagia rehabilitation-A review, *Prog Brain Res* **228**(2016) 185-219.
25. T. Yanagisawa, M. Hirata, Y. Saitoh, T. Goto, H. Kishima, R. Fukuma, H. Yokoi, Y. Kamitani and T. Yoshimine, Real-time control of a prosthetic hand using human electrocorticography signals, *Journal of Neurosurgery* **114**(6) (2011) 1715-1722.
26. N. E. Crone, D. L. Miglioretti, B. Gordon and R. P. Lesser, Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. II. Event-related synchronization in the gamma band, *Brain* **121 (Pt 12)**(1998) 2301-15.
27. H. Hashimoto, Y. Hasegawa, T. Araki, H. Sugata, T. Yanagisawa, S. Yorifuji and M. Hirata, Non-invasive detection of language-related prefrontal high gamma band activity with beamforming MEG, *Sci Rep* **7**(1) (2017) 14262.
28. Y. Nakanishi, T. Yanagisawa, D. Shin, H. Kambara, N. Yoshimura, M. Tanaka, R. Fukuma, H. Kishima, M. Hirata and Y. Koike, Mapping ECoG channel contributions to trajectory and muscle activity prediction in human sensorimotor cortex, *Sci Rep* **7**(2017) 45486.
29. X. R. N. Wang, A. Farhadi, R. Rao and B. J. a. p. a. Brunton, AJILE Movement Prediction: Multimodal Deep Learning for Natural Human Neural Recordings and Video, (2017).
30. P. M. Cheng and H. S. Malhi, Transfer learning with convolutional neural networks for classification of abdominal ultrasound images, *Journal of digital imaging* **30**(2) (2017) 234-243.
31. Y. Yu, J. Wang, C. W. Ng, Y. Ma, S. Mo, E. L. S. Fong, J. Xing, Z. Song, Y. Xie, K. Si, A. Wee, R. E. Welsch, P. T. C. So and H. Yu, Deep learning enables automated scoring of liver fibrosis stages, *Sci Rep* **8**(1) (2018) 16016.
32. A. M. Dawud, K. Yurtkan and H. Oztoprak, Application of Deep Learning in Neuroradiology: Brain Haemorrhage Classification Using Transfer Learning, *Computational Intelligence and Neuroscience* **2019**(2019).
33. O. M. Manzanera, S. K. Meles, K. L. Leenders, R. J. Renken, M. Pagani, D. Arnaldi, F. Nobili, J. Obeso, M. R. Oroz and S. Morbelli, Scaled Subprofile Modeling and Convolutional Neural Networks for the Identification of Parkinson's Disease in 3D Nuclear Imaging Data, *International journal of neural systems* **29**(9) (2019) 1950010-1950010.
34. J. Aoe, R. Fukuma, T. Yanagisawa, T. Harada, M. Tanaka, M. Kobayashi, Y. Inoue, S. Yamamoto, Y. Ohnishi and H. Kishima, Automatic diagnosis of neurological diseases using MEG signals with a deep neural network, *Scientific reports* **9**(1) (2019) 5057.
35. A. Emami, N. Kunii, T. Mtasuo, T. Shinozaki, K. Kawai and H. J. N. C. Takahashi, Seizure detection by convolutional neural network-based analysis of scalp electroencephalography plot images, (2019) 101684.
36. H. Yang, S. Sakhavi, K. K. Ang and C. Guan 2015, "On the use of convolutional neural networks and augmented CSP features for multi-class motor imagery of EEG signals classification," in 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), ed.^eds. Editior (IEEE, pp. IEEE.
37. A. Antoniades, L. Spyrou, D. Martin-Lopez, A. Valentin, G. Alarcon, S. Sanei and C. C. Took, Deep neural architectures for mapping scalp to intracranial EEG, *International journal of neural systems* **28**(08) (2018) 1850009.

38. C. Hua, H. Wang, H. Wang, S. Lu, C. Liu and S. M. Khalid, A Novel Method of Building Functional Brain Network Using Deep Learning Algorithm with Application in Proficiency Detection, *International journal of neural systems* **29**(01) (2019) 1850015.
39. A. H. Ansari, P. J. Cherian, A. Caicedo, G. Naulaers, M. De Vos and S. Van Huffel, Neonatal seizure detection using deep convolutional neural networks, *International journal of neural systems* **29**(04) (2019) 1850011.
40. U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, H. J. C. i. b. Adeli and medicine, Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals, **100**(2018) 270-278.
41. M. Long, Y. Cao, J. Wang and M. I. Jordan, Learning transferable features with deep adaptation networks, *arXiv preprint arXiv:1502.02791* (2015).
42. A. Krizhevsky, I. Sutskever and G. E. Hinton 2012, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, ed.^eds. Editior.
43. H. Firmin, S. Reilly and A. Fourcin, Non-invasive monitoring of reflexive swallowing, *Speech Hearing and Language* **10**(1997) 171-184.
44. T. Kusuhara, T. Nakamura, Y. Shirakawa, K. Mori, Y. Naomoto and Y. Yamamoto, Impedance pharyngography to assess swallowing function, *J Int Med Res* **32**(6) (2004) 608-16.
45. H. Hashimoto, M. Hirata, K. Takahashi, S. Kameda, Y. Katsuta, F. Yoshida, N. Hattori, T. Yanagisawa, J. Palmer, S. Oshino, T. Yoshimine and H. Kishima, Non-invasive quantification of human swallowing using a simple motion tracking system, *Scientific Reports* **8**(1) (2018) 5095.
46. M. X. Cohen, Assessing transient cross-frequency coupling in EEG data, *J Neurosci Methods* **168**(2) (2008) 494-9.
47. A. Delorme and S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, *Journal of neuroscience methods* **134**(1) (2004) 9-21.
48. D. J. Im, M. Tao and K. Branson, An empirical analysis of the optimization of deep network loss surfaces, *arXiv preprint arXiv:1612.04010* (2016).
49. S. Lu, Z. Lu and Y.-D. Zhang, Pathological brain detection based on AlexNet and transfer learning, *Journal of computational science* **30**(2019) 41-47.
50. R. Salmelin, M. Hamalainen, M. Kajola and R. Hari, Functional segregation of movement-related rhythmic activity in the human brain, *Neuroimage* **2**(4) (1995) 237-43.
51. P. G. Mihai, O. von Bohlen Und Halbach and M. Lotze, Differentiation of cerebral representation of occlusion and swallowing with fMRI, *Am J Physiol Gastrointest Liver Physiol* **304**(10) (2013) G847-54.
52. K. J. Miller, E. C. Leuthardt, G. Schalk, R. P. Rao, N. R. Anderson, D. W. Moran, J. W. Miller and J. G. Ojemann, Spectral changes in cortical surface potentials during motor movement, *J Neurosci* **27**(9) (2007) 2424-32.
53. P. G. Mihai, M. Otto, T. Platz, S. B. Eickhoff and M. Lotze, Sequential evolution of cortical activity and effective connectivity of swallowing using fMRI, *Hum Brain Mapp* **35**(12) (2014) 5962-73.
54. T. Yanagisawa, M. Hirata, Y. Saitoh, H. Kishima, K. Matsushita, T. Goto, R. Fukuma, H. Yokoi, Y. Kamitani and T. Yoshimine, Electrocorticographic control of a prosthetic arm in paralyzed patients, *Ann Neurol* **71**(3) (2012) 353-61.
55. S. S. Dalal, A. G. Guggisberg, E. Edwards, K. Sekihara, A. M. Findlay, R. T. Canolty, M. S. Berger, R. T. Knight, N. M. Barbaro, H. E. Kirsch and S. S. Nagarajan, Five-dimensional neuroimaging: localization of the time-frequency dynamics of cortical activity, *Neuroimage* **40**(4) (2008) 1686-700.
56. D. A. Moses, N. Mesgarani, M. K. Leonard and E. F. Chang, Neural speech recognition: continuous phoneme decoding using spatiotemporal representations of human cortical activity, *J Neural Eng* **13**(5) (2016) 056004.
57. D. F. Conant, K. E. Bouchard, M. K. Leonard and E. F. Chang, Human sensorimotor cortex control of directly-measured vocal tract movements during vowel production, *Journal of Neuroscience* (2018) 2382-17.
58. K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556* (2014).