

ESTIMATION OF VARIOUS HUMAN EMOTIONS USING LIGHTWEIGHT FNIRS DEVICE

Daisuke Fukui^{1,4}, Takushige Katsura², Masashi Egi¹, Norihisa Komoda³
and Takenao Ohkawa⁴

¹*Hitachi, Ltd., Research & Development Group, Japan*

²*Graduate School of Integrative Science and Engineering Systems, Tokyo City University, Japan*

³*Codesolutions Co. Ltd., Japan*

⁴*Graduate School of System Informatics, Kobe University, Japan*

ABSTRACT

We previously proposed a method for estimating pleasant and unpleasant emotions with high accuracy using only total hemoglobin data measured with a lightweight functional near infrared spectroscopy device. In this study, we used the method to evaluate the accuracy of estimating 20 types of emotions selected as uniformly distributed emotions in Russell's circumplex model. We first divided the 20 types of emotions into four groups, corresponding to the four quadrants of Russell's circumplex model and evaluated the estimation accuracy of each quadrant. The results indicate that the activation quadrant was estimated with high accuracy when the emotion was strongly aroused, with 76.7% recall for the pleasant-activation quadrant and 72.2% recall for the unpleasant-activation quadrant. We then evaluated the estimation accuracy of the 20 emotions individually. The results indicate that "excited" and "lethargic" were estimated with high accuracy, with 73.3% recall for "excited" and 61.5% recall for "lethargic," and recall of "excited" improved to 80% when the emotion was strongly aroused. The results of this study indicate that the more strongly emotions included in activation quadrant in Russell's circumplex model are aroused, the more accurately they can be classified. "Excited" and "lethargic" could be estimated with high accuracy regardless of the degree of emotional arousal.

KEYWORDS

fNIRS, Estimation of Emotions, k-Nearest Neighbor Model, Hemoglobin, Russell's Circumplex Model

1. INTRODUCTION

If the emotions evoked by people in various situations in their daily lives can be estimated at a high classification rate, they can be applied to a variety of purposes. For example, they can be used to improve human computer interaction. This can be achieved by identifying the emotions

people experience when using computers and accordingly optimizing computer behavior. They can also be used to identify people's unconscious evaluations of products and services. If we can identify what emotions people were experiencing while watching a TV commercial, we can improve its content. As more emotions are estimated, the range of applications expands. Russell proposed a model in which various human emotions are arranged in a circle in a two-dimensional space with two axes: pleasant and unpleasant, and activation and deactivation (Russell, 2003).

Functional near infrared spectroscopy (fNIRS), which measures the increase or decrease in hemoglobin (Hb) in blood of the brain based on the absorption rate of NIR light, has been used to estimate pleasant and unpleasant emotions in humans (Fukui, 2021, Yanagisawa, 2015). There is also a method using functional magnetic resonance imaging (fMRI), which measures the increase or decrease in Hb by the change in the magnetic force of deoxygenated Hb. Although these methods can estimate pleasant and unpleasant emotions, they cannot estimate more nuanced emotions as defined using Russell's circumplex model.

There is a method that uses electroencephalography (EEG) sensors to estimate more than two emotions: pleasant and unpleasant. Yuen (2009) computed six statistical features from EEG data and used a neural network to identify five emotions (anger, sad, surprise, happy, and neutral) with a 95% classification rate. Murugappan (2010) used the "db4" wavelet function to derive modified energy-based features from EEG data and identified five emotions (disgust, happy, surprise, fear, and neutral) with an 83.26% classification rate using a k-nearest neighbor (kNN) model. Islam (2021) converted one-dimensional EEG data into Pearson's-correlation-coefficient-featured images of channel correlation of EEG sub-bands and used a convolutional neural network trained on these images to recognize three levels of valence and arousal (low/middle/high) with an accuracy of over 70%.

However, the EEG sensors used in these studies are equipped with 32 to 62 channels, making it difficult to measure EEG in daily life. Also, EEG signals measured at various locations on the scalp are usually contaminated with a large amount of noise. For example, potential changes caused by eye and muscle movements are included in the EEG data as noise. Several methods have been proposed for removing noise in EEG data, but it is difficult to remove it completely.

We previously proposed a method for estimating emotions using a lightweight fNIRS device (Fukui, 2021). Since fNIRS measurement is less affected by noise than EEG measurement, it is more suitable for measuring brain activity of people in their daily lives. On the basis of the results of a previous study (Fukui, 2021), we evaluated the classification rate of the kNN model, which estimates 20 emotions that are selected based on Russell's circumplex model, by cross-validation (Fukui, 2022). To examine the impact of the activation–deactivation axis in Russell's circumplex model on estimation accuracy, we evaluated the classification rate of the kNN model, which estimates emotions in the four quadrants of Russell's circumplex model (pleasant–activation, pleasant–deactivation, unpleasant–activation, and unpleasant–deactivation).

2. EMOTION ESTIMATION BY FNIRS

2.1 Definition of Emotions

On the basis of Russell's circumplex model (Russell, 1980, 2003), we selected a total of 20 emotions shown in Figure 1, five in each of the four quadrants of pleasant–activation, pleasant–deactivation, unpleasant–activation, and unpleasant–deactivation. These emotions were selected to have moderate variations in the degree of valence (pleasant–unpleasant) and arousal (activation–deactivation) within each quadrant. The list of selected emotions is shown in Table 1. These 20 emotions were used as research targets in our study.

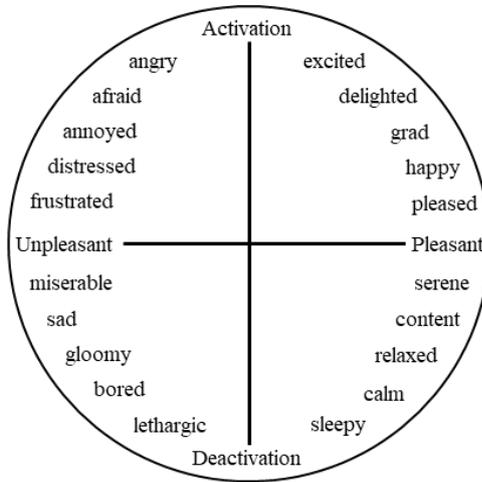


Figure 1. Russell's circumplex model

Table 1. Emotions selected for our study

Quadrants	Emotions
Pleasant-Activation	excited, delighted, grad, happy, pleased
Pleasant-Deactivation	sleepy, calm, relaxed, content, serene
Unpleasant-Activation	angry, afraid, annoyed, distressed, frustrated
Unpleasant-Deactivation	miserable, sad, gloomy, bored, lethargic

2.2 Measurement Method

Figure 2 shows the method we used for measuring brain activity during image viewing using HOT-1000 (Zhang, 2005, Nozawa, 2016), a lightweight fNIRS device manufactured by NeU Co¹. Figure 3 shows the appearance of the HOT-1000 and Table 2 lists its specifications. The participant wears HOT-1000 on his head and views the images on the display. By measuring total-Hb while the participant is viewing the image, we investigated the relationship between the emotion evoked by the participant and brain activity.

¹ <http://neu-brains.site/en/>

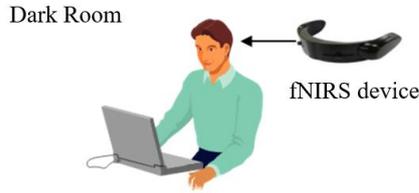


Figure 2. Measurement method



Figure 3. HOT-1000

Table 2. Specifications of HOT-1000

Specification item	Value
Number of channels	2ch (position adjustable)
Light source	LED
Communication protocol	Bluetooth Low Energy
Weight	125 g
Power supply	AAA alkaline batteries x 2 / USB
Operating time	1.5 hours

There are two types of data that can be measured with HOT-1000. One is the total-Hb measured at a depth of about 1 cm from the scalp (S), and the other is the total-Hb measured at a depth of about 3 cm from the scalp (D). During normal brain activity, the increase in total-Hb is not observed near the scalp but only deep in the brain. Therefore, the total-Hb calculated by subtracting S from D can be considered to indicate brain activity (B). However, the amount of change in total-Hb near the scalp and deep in the brain are always affected by heartbeat and other factors as well as brain activity, so it is necessary to eliminate these factors.

First, the baselines of S_i and D_i waveform data are corrected after removing pin and step noise. Next, α , which represents the degree that heartbeat and other factors affect brain activity, is calculated using the following five formulas. The brain-activity waveform data (B_i) are then derived by subtracting α times S_i from D_i . In these formulas, i represents the elapsed time from the start of the brain-activity measurement. The measurement of brain activity begins when HOT-1000 is placed on the participant's head.

$$B_i = D_i - \alpha \cdot S_i$$

- (1) $\alpha = 1$
- (2) $\alpha = \frac{H_i}{h_i}$
- (3) $\alpha = \frac{\sum_{i=T_0}^{T_2} d_i}{\sum_{i=T_0}^{T_2} s_i}$
- (4) $\alpha = \frac{\sum_{i=T_0}^{T_1} d_i}{\sum_{i=T_0}^{T_1} s_i}$
- (5) $\alpha = \text{Corr}(d_i, s_i)$

The H and h in Formula (2) represent the heartbeat components of D and S, respectively. The heartbeat components are calculated using a band-pass filter from 0.7 to 2.0 Hz, which indicates the frequency of the heartbeat. The d and s in Formulas (3) through (5) represent the values of D and S, respectively, in which the heartbeat component has been removed. The T_0 in Formulas (3) and (4) represents the start time of the rest period (see Chapter 3 for details). The T_2 in Formula (3) represents the end time of the image display. The T_1 in Formula (4) represents the end time of the rest period. The Corr function in Formula (5) calculates the correlation coefficient between d_i and s_i in the last 5 s.

The B_i derived using Formula (1) indicates the brain activity calculated simply from the difference between d_i and s_i without removing any effect of heartbeat and other factors. Formula (2) calculates the ratio of the heartbeat components of D and S to remove the effect of heartbeat. Formula (3) calculates the ratio of the integral values of d_i and s_i in the image-display period including the rest period to remove the effects of heartbeat and other factors (such as body movement). Formula (4) calculates the ratio of the integral values in Formula (3) from the rest period only. The reason the correlation coefficients between d_i and s_i for the past 5 s are calculated in Formula (5) is to derive the brain activity using the characteristic that the correlation between total-Hb near the scalp and deep in the brain becomes higher when there is no brain activity and lower when there is brain activity.

Finally, for each of the five types of brain-activity waveform data derived from the five formulas, the seven features shown in Table 3 are calculated. A total of 35 features are used to estimate 20 emotions. We used the kNN model, which is a machine-learning model that can obtain a better classification rate even with a small amount of training data.

Table 3. Features

ID	Feature
F1	Mean
F2	Standard deviation
F3	Median
F4	Maximum value
F5	Time required to reach maximum value
F6	Maximum change in time window (5 s)
F7	Rising slope (1 s from start)

3. EMOTION-ESTIMATION EXPERIMENT

We conducted an experiment using HOT-1000 to measure brain activity when a person viewed 20 images that evoked 20 different emotions. A total of 20 participants, who were asked to cooperate voluntarily, participated in this experiment. Twelve were in their 30s (8 men and 4 women) and 8 were in their 40s (6 men and 2 women). This experiment was approved by the Ethical Review Working Committee of the Research and Development Group, Hitachi, Ltd (Approval No. 20160905-0144).

An overview of the measurement procedure is shown in Figure 4. In the first 5 min, we explained the methods of the experiment, received informed consent, and attached HOT-1000 to the participants. The 20 images and their descriptions were then displayed to the participants in 2 sessions of 10 images each. Each image-display session was followed by a 1-min rest period. Before presenting the ten images to the participants, one neutral image (Figure 5) that was not used for analysis was displayed to the participants to suppress the response of brain activity to the first image displayed. A total of 11 images were displayed in a single “image display” session. The ten images displayed to the participants following the neutral image were arranged randomly. Finally, a questionnaire was administered to the participants to determine whether the images evoked the desired emotions. The reason the images were displayed in two separate sessions was to eliminate the effect of the order in which the images were displayed on the results. The ten images displayed to half the participants in the first image-display session were displayed to the other half in the second image-display session and vice versa.



Figure 4. Measurement procedure



Figure 5. Neutral image

The image-display procedure is shown in Figure 6. First, a 30-s rest period was set up to calm brain activity. During the rest period, only the character "+", which is about 2 cm in length and width, was displayed in the center of the screen, and the participants were told to gaze at it. Next, the description of the image to be displayed was shown for 7 s. By displaying the description of the image immediately before the image display, the target brain activity was stimulated. Finally, an image was displayed to the participants for 10 s to evoke the target emotion. This procedure was repeated 11 times using 1 neutral image and 10 different experimental images.

Estimation of Various Human Emotions using Lightweight fNIRS Device



Figure 6. Image-display procedure



Figure 7. Example of description and image

Figure 7 shows an example of the descriptions displayed to the participants and those displayed following the descriptions. The 20 images used in the experiment were purchased from Getty Images. The method of selecting the 20 images was as follows. First, for each of the 20 emotions (see Table 1), 5 evaluators (all Japanese nationals) selected one candidate image that evoked the emotion through keyword searches on Getty Images². Next, the five evaluators voted to select one image that would strongly evoke that particular emotion. Two evaluators were experts in the field of product and service evaluation (one man and one woman in their 40s). The other three were the researchers (men in their 40s) who have been engaged in fNIRS research for more than two years.

The content of the questionnaire given to the participants at the end of the experiment is shown in Figure 8. The participants were asked whether the emotion described in the description matched the content of the images. The visual analogue scale (VAS) values, which were obtained by normalizing the line lengths of the questionnaire responses to take values from 0 to 10 based on the left edge, were used as the teacher data. The 80 records with a VAS value of less than 5 (data that did not match the target emotion) were excluded from the analysis because the emotion evoked was unknown. The remaining 320 records were used for the analysis. In general, a minimum of 50 records is required for training, and this experiment met that criterion. It is preferable to train 10 times the amount of data as the number of features, but we were able to obtain more than 90% of the recommended data, although it was slightly less than the 350 recommended records for the 35 features in this study.

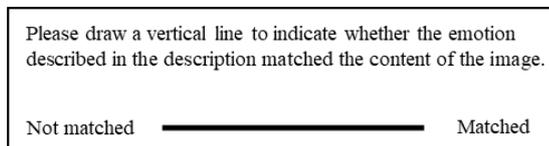


Figure 8. Questionnaire content

² <https://www.gettyimages.com/>

4. EXPERIMENTAL RESULTS

On the basis on the above experimental data, we trained 4 different kNN models for each of quadrants and 20 different kNN models for each of the emotions shown in Table 1. The k in the kNN models was chosen to be 3, which is the most accurate among various k values. The models were evaluated using 10-fold cross-validation using the metrics of Matthews correlation coefficient (MCC), precision, recall, and accuracy.

The 10-fold cross-validation is a validation method with which the training data are randomly divided into ten groups; nine groups are trained and the remaining group is validated, and repeated ten times until all groups are validated. MCC does not depend on data bias. Precision is used to evaluate whether the estimated result from the classification model is correct when the result is positive. Recall is used to evaluate whether the estimated result from the classification model is correct when the correct emotion is positive. Accuracy is used to evaluate whether the estimated result from the classification model is correct for all estimated results.

The formulas for calculating each evaluation metric are described below: TP (true positive) represents the amount of data when the result estimated from the classification model is positive and correct; FP (false positive) represents the amount of data when the result estimated from the classification model is positive and incorrect; TN (true negative) represents the amount of data when the result estimated from the classification model is negative and correct; FN (false negative) represents the amount of data when the result estimated from the classification model is negative and incorrect.

$$\bullet \text{ MCC} = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}}$$

$$\bullet \text{ Precision} = \frac{TP}{TP + FP}$$

$$\bullet \text{ Recall} = \frac{TP}{TP + FN}$$

$$\bullet \text{ Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Table 4 to Table 7 show the evaluation results of the kNN models for each quadrant in Table 1 when the VAS thresholds were changed from 5 to 9. The values in the N column indicate the number of records that exceed the VAS threshold. If there is a large bias in the amount of correct data (corresponding to the target emotion) and the amount of incorrect data (not corresponding to the target emotion), accuracy will be large. Therefore, the baseline value, which is obtained by dividing the greater amount of correct or incorrect data by the total data, is also described in the table. The baseline value represents the accuracy of a meaningless classification model that predicts all data as correct or incorrect and used as a metric to evaluate the validity of accuracy.

Table 4. Classification results of k-NN models for pleasant-activation quadrant

VAS threshold	MCC	Precision	Recall	Accuracy (baseline)	N
5 or more	0.213	0.632	0.145	75.6 (74.1)	83
6 or more	0.224	0.517	0.211	78.1 (77.8)	71
7 or more	0.286	0.56	0.241	82.8 (81.9)	58
8 or more	0.487	0.667	0.444	90.9 (85.9)	45
9 or more	0.689	0.676	0.767	94.4 (90.6)	30

Table 5. Classification results of k-NN models for pleasant-deactivation quadrant

VAS threshold	MCC	Precision	Recall	Accuracy (baseline)	N
5 or more	0.116	0.407	0.237	67.8 (70.9)	93
6 or more	0.32	0.7	0.25	77.5 (73.8)	84
7 or more	0.406	0.697	0.343	84.4 (79.1)	67
8 or more	0.47	0.667	0.431	90.9 (84.1)	51
9 or more	0.498	0.538	0.538	96.3 (91.9)	26

Table 6. Classification results of k-NN models for unpleasant-activation quadrant

VAS threshold	MCC	Precision	Recall	Accuracy (baseline)	N
5 or more	0.165	0.367	0.269	75 (79.1)	67
6 or more	0.424	0.739	0.315	86.6 (83.1)	54
7 or more	0.493	0.833	0.341	90 (86.3)	44
8 or more	0.626	0.8	0.533	95.6 (90.6)	30
9 or more	0.685	0.684	0.722	93.1 (94.4)	18

Table 7. Classification results of k-NN models for unpleasant-deactivation quadrant

VAS threshold	MCC	Precision	Recall	Accuracy (baseline)	N
5 or more	0.218	0.515	0.221	76.3 (75.9)	77
6 or more	0.229	0.457	0.254	79.4 (80.3)	63
7 or more	0.373	0.607	0.327	85.6 (83.8)	52
8 or more	0.477	0.543	0.528	93.4 (88.8)	36
9 or more	0.485	0.5	0.545	96.6 (93.1)	22

For VAS thresholds of 5 or more, the recall of each model ranged from 0.145 to 0.269. This is almost the same as the 0.25 recall of randomly selecting one of the four quadrants. However, as the VAS threshold increased, the recall of each model increased. In particular, the recall of the two quadrants in the activation direction greatly improved, reaching 76.7% for the pleasant-activation quadrant and 72.2% for the unpleasant-activation quadrant when the VAS threshold was set at 9. However, the recall of the two quadrants in the deactivation direction improved only to a small extent: 53.8% for the pleasant-deactivation quadrant and 54.5% for the unpleasant-deactivation quadrant when the VAS threshold was set at 9. This may be due to the fact that emotions in the activation direction have a greater impact on changes in total-Hb than emotions in the deactivation direction.

To confirm that these results were not biased toward any particular participant or image, a test of uniform distribution using the chi-square method was conducted. The data were randomly grouped so that the expected frequency was greater than 5, since it is known that an approximation to the chi-square distribution becomes poorer when the expected frequency is less than 5.

Figure 9 shows the number of records per participant for pleasant-activation emotions with a VAS threshold of 9 or more. The test of uniform distribution of this graph revealed a chi-square value of 3.668 and a p-value of 0.453, rejecting the null hypothesis that the amount of data is biased toward any particular participant.

Figure 10 shows the number of records per image for pleasant–activation emotions with a VAS threshold of 9 or more. The test of uniform distribution of this graph revealed a chi-square value of 2.335 and a p-value of 0.674, rejecting the null hypothesis that the number of data is biased toward any particular image.

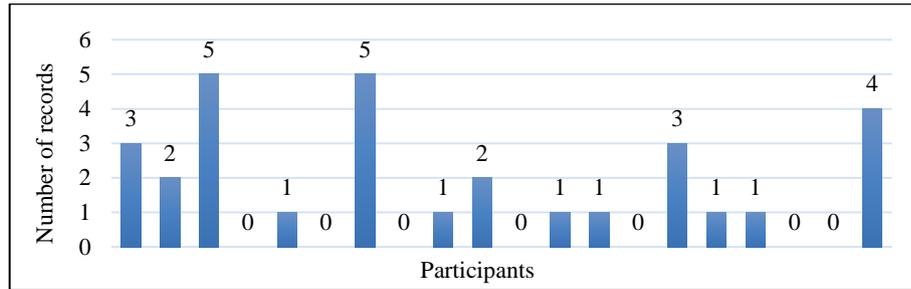


Figure 9. Number of pleasant-activation records per participant

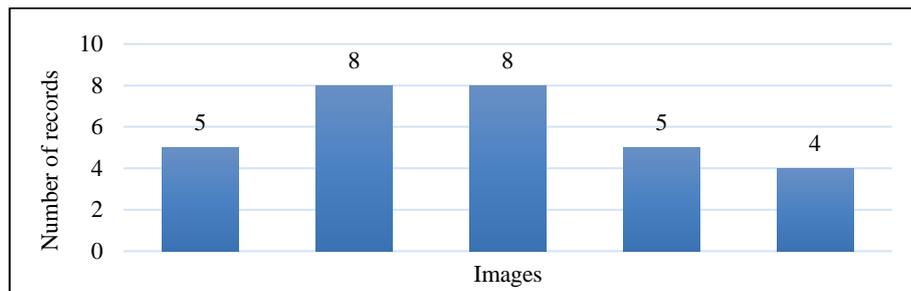


Figure 10. Number of pleasant-activation records per image

Figure 11 shows the number of records per participant for pleasant–deactivation emotions with a VAS threshold of 9 or more. The test of uniform distribution of this graph revealed a chi-square value of 3.616 and a p-value of 0.460, rejecting the null hypothesis that the amount of data is biased toward any particular participant.

Figure 12 shows the number of records per image for pleasant–activation emotions with a VAS threshold of 9 or more. The test of uniform distribution of this graph revealed a chi-square value of 2.462 and a p-value of 0.651, rejecting the null hypothesis that the amount of data is biased toward any particular image.

Estimation of Various Human Emotions using Lightweight fNIRS Device

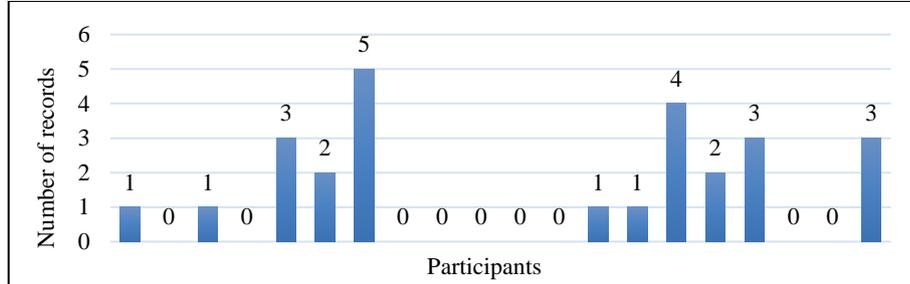


Figure 11. Number of pleasant-deactivation records per participant

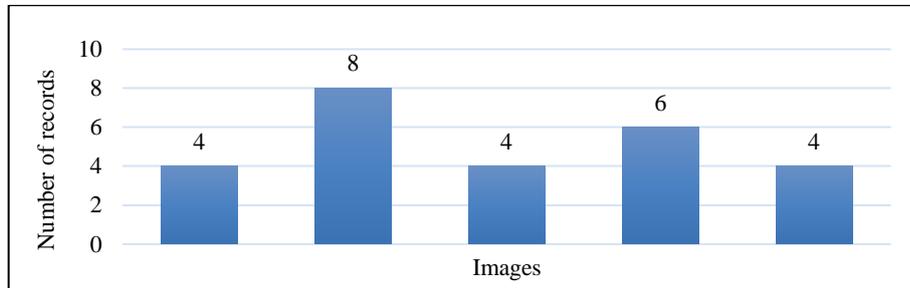


Figure 12. Number of pleasant-deactivation records per image

Figure 13 shows the number of records per participant for unpleasant-activation emotions with a VAS threshold of 9 or more. The test of uniform distribution of this graph revealed a chi-square value of 2.334 and a p-value of 0.311, rejecting the null hypothesis that the amount of data is biased toward any particular participant.

Figure 14 shows the number of records per image for unpleasant-activation emotions with a VAS threshold of 9 or more. The test of uniform distribution of this graph revealed a chi-square value of 1.001 and a p-value of 0.606, rejecting the null hypothesis that the amount of data is biased toward any particular image.

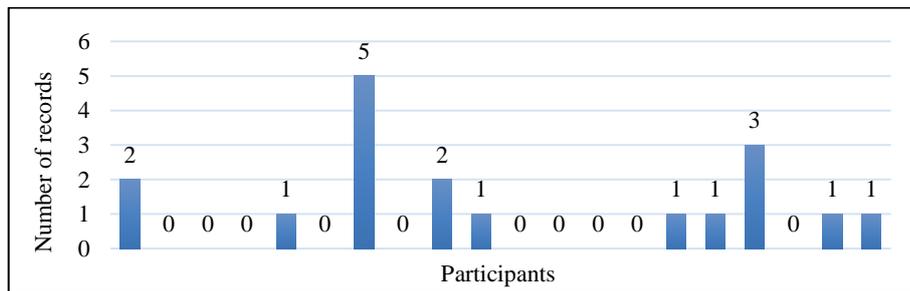


Figure 13. Number of unpleasant-activation records per participant

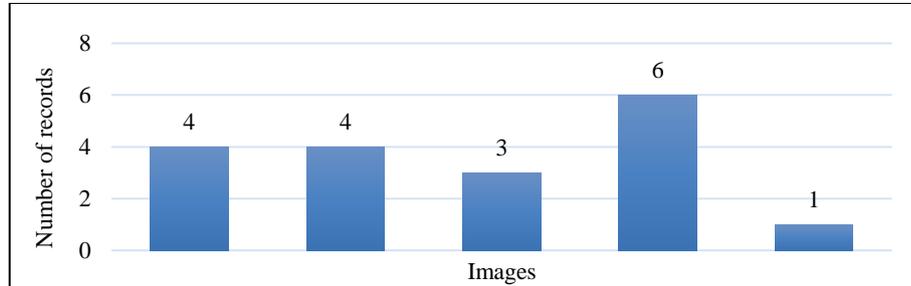


Figure 14. Number of unpleasant-activation records per image

Figure 15 shows the number of records per participant for unpleasant–deactivation emotions with a VAS threshold of 9 or more. The test of uniform distribution of this graph revealed a chi-square value of 1.999 and a p-value of 0.573, rejecting the null hypothesis that the amount of data is biased toward any particular participant.

Figure 16 shows the number of records per image for unpleasant–deactivation emotions with a VAS threshold of 9 or more. The test of uniform distribution of this graph revealed a chi-square value of 0.908 and a p-value of 0.823, rejecting the null hypothesis that the amount of data is biased toward any particular image.

The results of these tests confirm that in all four quadrants, the data with a VAS threshold of 9 or more were not biased toward any particular participant or image.

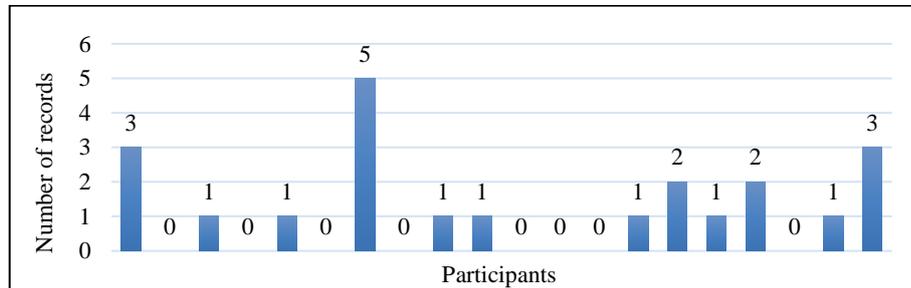


Figure 15. Number of unpleasant-deactivation records per participant

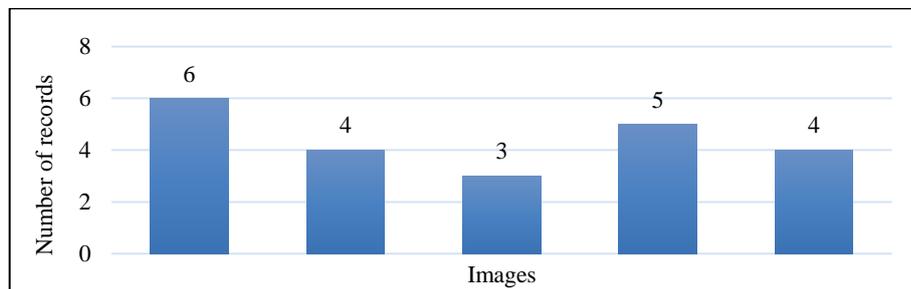


Figure 16. Number of unpleasant-deactivation records per image

Table 8 shows the evaluation results of the kNN models for 20 emotions in Table 1 when the VAS threshold is greater than 5. The values in the N column indicate the number of records corresponding to each emotion. The results indicate that the "excited" kNN model's MCC was 0.695, precision was 68.8%, recall was 73.3%, and accuracy was 97.2% (baseline was 95.3%). The "lethargic" kNN model's MCC was 0.656, precision was 72.7%, recall was 61.5%, and accuracy was 97.5% (baseline was 95.9%). Both models had high MCC with recall exceeding 60% and accuracy exceeding the baseline. All the other kNN models showed lower accuracy than the baseline and their recall was less than 50%.

"Excited" is placed as a highly active emotion in the "pleasant" direction in Russel's circumplex model, and "lethargic" is placed as a less active emotion in the "unpleasant" direction in that model. It is possible that the degree of activation in Russel's circumplex model is reflected in the amount of total-Hb.

Table 8. Classification results of k-NN models for individual emotions

Emotion	MCC	Precision	Recall	Accuracy (baseline)	N
excited	0.695	0.688	0.733	97.2 (95.3)	15
lethargic	0.656	0.727	0.615	97.5 (95.9)	13
calm	0.319	0.333	0.400	91.3 (93.8)	20
frustrated	0.255	0.278	0.313	92.5 (95.0)	16
afraid	0.234	0.250	0.273	94.7 (96.6)	11
gloomy	0.211	0.273	0.214	94.1 (95.6)	14
serene	0.154	0.211	0.200	90.3 (93.8)	19
sad	0.156	0.182	0.235	90.3 (94.7)	17
distressed	0.137	0.200	0.143	93.8 (95.6)	14
content	0.132	0.200	0.158	91.3 (94.1)	20
annoyed	0.118	0.130	0.214	90.3 (95.6)	14
grad	0.103	0.182	0.111	92.2 (94.4)	18
relaxed	0.099	0.150	0.158	89.7 (94.1)	19
sleepy	0.097	0.143	0.133	92.2 (95.3)	15
miserable	0.080	0.111	0.143	91.3 (95.6)	14
pleased	0.065	0.100	0.133	90.3 (95.3)	15
happy	0.063	0.111	0.118	90.3 (94.7)	17
delighted	0.053	0.105	0.111	89.7 (94.4)	18
angry	0.000	0.000	0.000	91.6 (96.3)	12
bored	-0.013	0.048	0.053	88.1 (94.1)	19

Tables 9 and 10 show the classification results when the VAS thresholds were changed from 6 to 9 for "excited" and "lethargic", respectively. For "excited," when the VAS threshold exceeded 8, recall improved to 0.8. There was no improvement in the recall for "lethargic" by changing the VAS threshold. Similar to the results of the four-quadrant evaluation, the recall of deactivation-emotion estimation does not improve with changes in the VAS threshold, and it is thought that the recall conversely decreases as the amount of data decreases.

These results indicate that our method can estimate brain activity when "excited" or "lethargic" is evoked with a classification rate of more than 60%. The results also indicate that "excited" with a VAS of more than 8 can be classified with a classification rate of 80%. It may be possible to classify the brain activity of people who feel strongly "excited" with a high classification rate.

Table 9. Classification results of k-NN model for “excited”

VAS threshold	MCC	Precision	Recall	Accuracy (baseline)	N
5 or more	0.695	0.688	0.733	97.2 (95.3)	15
6 or more	0.673	0.667	0.714	96.7 (94.9)	14
7 or more	0.648	0.667	0.667	96.4 (94.6)	12
8 or more	0.787	0.8	0.8	97.5 (93.8)	10
9 or more	0.714	0.667	0.8	96.9 (94.8)	5

Table 10. Classification results of k-NN model for “lethargic”

VAS threshold	MCC	Precision	Recall	Accuracy (baseline)	N
5 or more	0.656	0.727	0.615	97.5 (95.9)	13
6 or more	0.264	0.4	0.2	96.0 (96.3)	10
7 or more	0.431	0.6	0.333	96.4 (95.9)	9
8 or more	0.28	0.286	0.333	94.4 (96.3)	6
9 or more	0.478	0.5	0.5	95.8 (95.8)	4

5. CONCLUSION

To estimate various emotions in daily life, we conducted an experiment to estimate 20 emotions that are selected based on Russell’s circumplex model using HOT-1000, which is a lightweight fNIRS device. The 35 features representing brain activity were first calculated from the total-Hb data measured using HOT-1000. Using those 35 features, we then trained kNN models to predict various emotions and evaluated the classification rate of those emotions by 10-fold cross validation.

The 20 emotions were first divided into 5 groups of 5 each, corresponding to the 4 quadrants of Russell’s circumplex model, and their estimation accuracy was evaluated. The results indicate that the recall of all kNN models for each quadrant improved as the VAS threshold increased. The recall of kNN models for the two quadrants in the activation direction greatly improved, with recall of the kNN model for the pleasant–activation quadrant improving from 14.5 to 76.7% and recall of the kNN model for the unpleasant–activation quadrant improving from 26.9 to 72.2% when the VAS threshold was set at 9. The emotions in the activation quadrant may have a stronger effect on brain activity with higher degrees of arousal.

We then evaluated the estimation accuracy of each of the 20 emotions. The recall of “excited” was 73.3% and the recall of “lethargic” was 61.5% when the VAS threshold was greater than 5. The recall of all other emotions were less than 50%. As shown in Figure 1, “excited” was the most active emotion in the pleasant–activation quadrant, while “lethargic” was the most inactive emotion in the unpleasant–deactivation quadrant of the Russell’s model. The recall of “excited” tended to improve as the VAS threshold increased, with a recall of 80% when the VAS threshold was more than 8. However, the recall of the kNN model for “lethargic” did not improve when the VAS threshold increased. These results also indicate that the degree of activation or deactivation in Russell’s circumplex model affects brain activity.

Previous studies have shown that EEG is effective in predicting various human emotions. However, the EEG sensors used in these studies are equipped with 32 to 62 channels, making it difficult to measure EEG in daily life. Also, EEG signals measured at various locations on the scalp are usually contaminated with a large amount of noise. Therefore, we attempted to estimate emotions using a lightweight device that can measure brain activity in daily life and

measure fNIRS data, which is robust against noise. The results of this study indicate that the more strongly emotions included in activation quadrant in Russell’s circumplex model are aroused, the more accurately they can be classified. “Excited” and “lethargic” could be estimated with high accuracy regardless of the degree of emotional arousal in daily lives. However, the number of participants and amount of data collected were not sufficient, so continued data collection is necessary for future work.

REFERENCES

- Cover, T. and Hart, P., 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, Vol. 13, Issue 1, pp 21–27.
- Fukui, D. et al., 2021. Estimation of human emotion using total-hemoglobin data measured by a lightweight fNIRS device. *IEEJ Transactions on Electronics, Information and Systems*, Vol. 141, Issue 6, pp. 735–742. (in Japanese)
- Fukui, D. et al., 2022. Estimation of human emotion defined with Russell’s circumplex model using lightweight fNIRS device. *Proceedings of the International Conferences on Interfaces and Human Computer Interaction 2022*, pp. 41–48.
- Islam, M. R. et al., 2021. EEG channel correlation based model for emotion recognition. *Computers in Biology and Medicine*, 136, 104757.
- Murugappan, M. et al., 2010. Classification of human emotion from EEG using discrete wavelet transform. *Journal of biomedical science and engineering*, 3(04), pp. 390–396.
- Nozawa, T. et al., 2016. Interpersonal frontopolar neural synchronization in group communication: An exploration toward fNIRS hyperscanning of natural interactions. *NeuroImage*, Vol. 133, June 2016, pp. 484–497.
- Russell, J. A., 1980. A circumplex model of affect. *Journal of Personality and Social Psychology*, Vol. 39, No. 6, pp. 1161–1178.
- Russell, J. A., 2003. Core affect and the psychological construction of emotion. *Psychological Review*, Vol. 110, pp. 145–172.
- Yanagisawa, K. and Tsunashima, H., 2015. Evaluation of pleasant and unpleasant emotion with visual stimulus using NIRS. *Transactions of the Japanese Society for Medical and Biological Engineering*, Vol. 53, Supplement, pp. S379–S382.
- Yuen, C. T. et al., 2009. Classification of human emotions from EEG signals using statistical features and neural network. *International Journal of Integrated Engineering*, 1(3), pp. 71–79.
- Zhang, Y. et al., 2005. Eigenvector-based spatial filtering for reduction of physiological interference in diffuse optical imaging. *J. Biomed Opt*, Vol. 10, No. 1, pp. 11–14.