Blending CNNs with Different Signal Lengths for Real-time EEG Classification Sensitive to the Changes

Takashi Ota[a], Kenji Funahashi[a]
[a]Nagoya Institute of Technology, Nagoya 466-8555 Japan

ABSTRACT

Although a lot of BMI research using CNN has been performed, CNN's response to changes in the input EEG is too late to proceed in real-time. We propose a method to improve the real-time performance by blending multiple CNNs with different input signal length. The proposed method generates a classifier which has the advantage of a classifier with short input signal length, i.e., fast response to changes in the input signal, and also the advantage of a classifier with long input signal length, i.e., high classification performance.

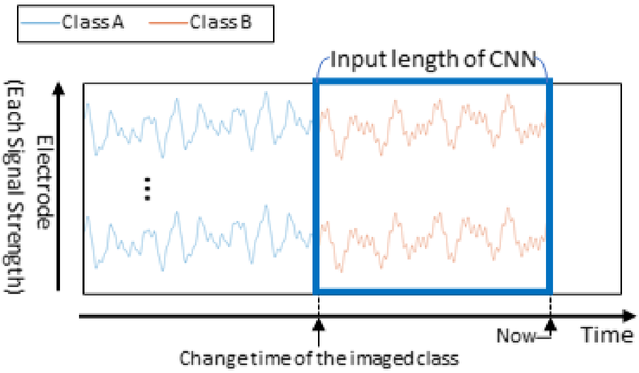
Keywords: Brain Machine Interface, EEG, Neural Network

1. INTRODUCTION

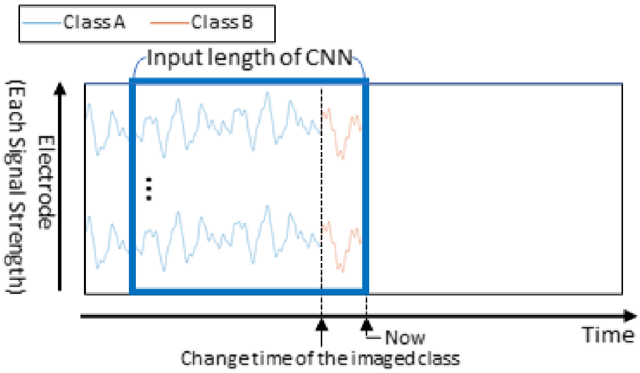
An interface that connects a brain and a machine using brain information such as an electroencephalogram (EEG) is called a brain machine interface (BMI). An EEG is an electrical activity in the brain. It is recorded by electrodes placed on the head, and an EEG can be measured non-invasively. Furthermore, the recent electroencephalograph is small, portable, and high-resolution. Therefore, there is a lot of research about BMI using EEG. For example, the research[1] has been conducted to operate a robot arm with a monkey's EEG. Moreover, there is the research[2] that uses a human's EEG to operate a wheelchair. Schirrmeister et al. showed that[3] CNN, a type of neural network, can classify an EEG with the same high accuracy as FBCSP[4], a conventional EEG decoding technique. This research has attracted attention to the use of neural networks for EEG classification, which are capable of classification of an EEG with high accuracy and automatic feature selection. EEG classification by CNNs involves inputting EEG intensity data on a two-dimensional plane with the horizontal axis as the time axis and the vertical axis as the electrode number into a CNN, and finally obtaining the output of the class with the largest prediction probability. However, there is a disadvantage to classify an EEG in real time that the tracking of the change in the user's imaged class, i.e., the change in an EEG, is delayed if the input signal to a CNN is long in the time axis. Considering the behavior of an CNN classifier for inputs containing changes in imaged class, immediately after the user's imaged class changes (Figure 1 (a)), the input to the classifier contains little EEG of the post-change imaged class, so the prediction probability of the pre-change imaged class is high. When the EEG of the post-change imaged class becomes the majority of the input to the classifier (Figure 1 (b)), the prediction probability of the post-change imaged class becomes high at least. The temporal input of one of the CNN architectures used in the study by Schirrmeister et al.[4] was 522 samples for an EEG of 250 Hz sampling. If this CNN is directly used in a real-time system, the first time all the input become the EEG of the post-changed imaged class like Figure 1 (b) is 522 / 255 ≈ 2.1 seconds after the user's imaged class changes. However, if the input signal length is simply shortened to reduce this delay, the classification performance will be reduced. In our preliminary experiment, we compared the classification performance of a CNN with 128 samples of input signal length and that of a CNN with 256 samples of input signal length. They were trained with 750 training data each for two classes. The classification performance of the former was 64.1% and that of the latter was 71.6%. Thus, there is a trade-off between the speed of response to changes in the imaged class and the classification performance, and it is difficult to classify an EEG in real time by neural networks. In this study, we propose a method to reduce this delay by blending the outputs of CNNs with different input signal lengths in the time axis direction.

2. PROPOSED METHOD

We propose a method to blend the output of CNNs with different input signal lengths in the time axis direction. Blending is a method to improve the overall classification performance by a weighted average of multiple weak classifiers. Figure 2 is the schematic diagram of the proposed method. In this diagram, more than half of the input of CNN_{LONG}, which has a long input signal length, is still EEG data before the change in an imaged class, so the probability of class *A* is shown as 0.8 even though the imaged class has already changed from *A* to *B* at the current time. On the other hand, CNN_{SHORT}, which has a short input signal length, is able to respond to the change in an imaged class because the EEG data after the change in an imaged class already occupies the entire input, and shows a probability of class *B* of 0.95. By blending the output of CNNs with different input signal lengths in the temporal direction in this way, the classifier can respond to changes in imaged class more quickly compared to CNNs with longer input signal lengths. Moreover, the classifier has better classification performance compared to CNNs with shorter input signal length.



(a) Immediately after the imaged class change



(b) When the post-change EEG becomes the majority of the input

Figure 1. Changes in the input to the CNN when the imaged class changes.

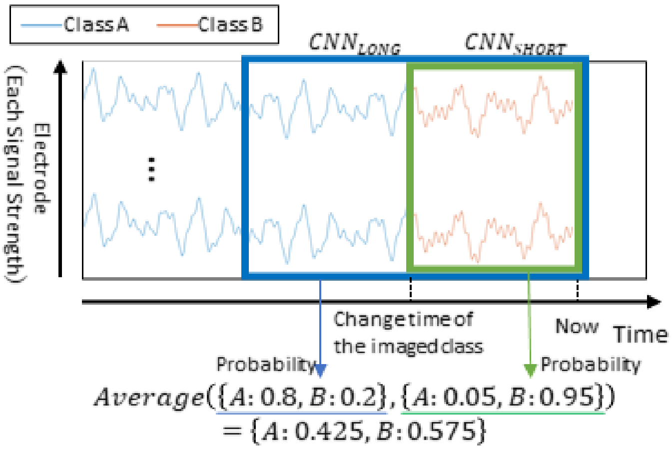


Figure 2. Schematic diagram of the proposed method

3. EXPERIMENT

3.1 Overview

The following procedure was used for the experiment to demonstrate the performance of the proposed method.

1. Training EEG data and test EEG data were recorded. The format of these 2 types of data is explained in section 3.2.
2. CNNs with different input signal lengths were trained with the training EEG data. The conditions for training CNNs are explained in section 3.3.

3. The proposed method was applied to the test EEG data, and the performance of the proposed method was evaluated. The evaluation method of the performance of the proposed method is explained in section 3.4.

3.2 Data Set

In this experiment, we recorded the EEG of two male subjects in their 20s. The electroencephalograph was Emotiv EPOC X, which had 14 electrodes. Its internal sampling rate was 2048Hz, and it was down sampled to 256Hz. Two different types of EEG data: training EEG data and test EEG data were recorded in this study as mentioned above. In the training EEG recording phase, an EEG was recorded while the subject continued to image the instructed action for 16 seconds. There were two types of actions that the subjects were instructed to image: opening and closing of the right hand (right hand class) and opening and closing of the left hand (left hand class). For each of these two classes, an EEG was recorded for 24 times. In the test EEG recording phase the subjects were given instructions on the screen to change the imaged class at the middle of the 8-second EEG recording. There were two types of tasks in the test EEG recording phase: change from right hand class to left hand class and change from left hand class to right hand class. For each of these two tasks, an EEG was recorded for 20 times. The recording tasks were shuffled randomly in order to prevent habituation to the recording task for both the training and test EEG data

3.3 Training CNNs

In order to train the CNNs, the training EEG data was cropped to the input size (the number of samples for data with a sampling rate of 256Hz) of each CNN. Cropping[5] is the process of cutting out a lot of CNN's input size data from data that is longer than the CNN's input size. Cropping is a popular method of data augmentation in the field of image recognition, and Schirrmeister et al. has confirmed that cropping is effective for EEG data as well[4]. Since there was a risk of overfitting if a certain part of the training data was included in many cropped data, we changed the step width of the cropping according to the input size as shown in Table 1. The architecture of the CNNs was determined based on the study of Schirrmeister et al. However, the input sizes of the CNNs used in this study were 180, 256, and 360, and all three types of CNNs had shorter input size than the CNNs used in the study by Schirrmeister et al. Therefore, the size of the convolution window was changed according to the input size, and the dropout ratio and the number of filters were changed according to the amount of training data.

Table 1. Cropping step width and amount of training data for each input size

Input size	Step width	Amount of training data
180	56	2721
256	80	1880
360	120	1201

3.4 Performance evaluation of the proposed method

The input sizes of the CNNs used in this study were set to 180, 256, and 360. For each input size, six classifiers with slightly different CNN architectures were generated and trained. Since the EEG classifier is a weak learner with low classification accuracy, the simple method as blending CNNs with the same input size improves the classification accuracy. Therefore, we compared the blending of classifiers with the same input signal size and the blending of classifiers with different input sizes (our proposed method). First, we obtained a total of 18 classifiers, 6 for each input size, and calculated the transition of the probability of each class over time on the test data using the following procedure.

(1) For each 8-second test data, the corresponding length of data was cut out in 16 sample (approximately 0.06-second) steps and was inputted to each classifier.

(2) The outputs that were obtained in (1) were blended with equal weighting using the following blending conditions (A) through (E). In the following, the abbreviations in parentheses are used.

(A) 6 classifiers with input size 180 (180×6)
(B) 6 classifiers with input size 256 (256×6)
(C) 6 classifiers with input size 360 (360×6)
(D) 2 classifiers with input size 180, 2 classifiers with input size 256, and 2 classifiers with input size 360 (2+2+2)
(E) 3 classifiers with input size 180, 2 classifiers with input size 256, and 1 classifier with input size 360 (3+2+1)

(3) For each of the blending conditions from (A) to (E), the classification performance under normal conditions and the speed of response to changes in an imaged class were examined. Specifically, we divided the 8-second test data into following three phases. Phase 1 before the change in an imaged class at the beginning of the test data, phase 2 during the transition in an imaged class at the middle of the test data, and phase 3 after the change in an imaged class at the end of the test. Then, the percentage of correct responses in each phase was calculated for each subject and each blending condition. Classifiers that are sensitive to changes in an imaged class are expected to have higher accuracy in phase 2. Classifiers with higher classification performance under normal condition are expected to have higher accuracy in phase 1 and phase 3. In the field of brain science, it is said that the reaction time is about 200–300 milliseconds in button pressing experiments that follow visual stimuli. Therefore, the start time of phase 2 was set to about 0.25 seconds (64 samples) after the instruction to change the imaged class, because there was a delay between the time when the instruction to change the imaged class was displayed and the time when the subject started imaging the post-change class. Considering that the delay varies among the data, the duration of phase 2 was set to be long enough, about 2.5 seconds (640 samples).

4. RESULT AND CONSIDERATION

The average percentages of correct answers for each phase of the blending conditions (A) through (E) introduced in section 3 are shown in Table 2. Comparing (A), (B), and (C), we can see that the classifier with the shorter input size (A) has higher classification accuracy in phase 2, while the classifier with the longer input size (C) has higher classification accuracy in phase 1 and phase 3. As expected, the blended classifier with a short input size responds quickly to changes in the imaged class, while the blended classifier with a long input size has a high classification performance under normal conditions. Next, comparing the blending of classifiers with the same input size ((A), (B), and (C)) with the blending of classifiers with different input sizes ((D) and (E)) for the data of both subjects, we can confirm that the latter shows relatively stable and high probability, and also responds faster to changes in an imaged class. Comparing (B) and (E) for subject 1, it can be confirmed that the proposed method has higher accuracy in all phases. Furthermore, when comparing (B) and (D) of subject 2, and (B) and (E) of subject 2, it can be confirmed that the proposed method is more accurate in all phases as well. This suggests that the proposed method may perform better than the intermediate performance between a classifier with large input size and a classifier with small input size.

Figure 3 shows the change in the output probability of the classifier for an example of test data. By comparing (A), (B), and (C) in this figure, it can be confirmed that the shorter the input signal length, the faster the blended classifier responds to changes in the imaged class. This result agrees with the results of Table 2. The dotted line in Figure 3 indicates the timing of the instruction to change the imaged class. We can see that it takes about 0.3 seconds for (A), about 0.5 seconds for (B), about 0.8 seconds for (C), about 0.4 seconds for (D), and about 0.4 seconds for (E) to actually change the output of the classifier after the instruction to change the imaged class is given. Based on the fact that reaction time is said to be around 200–300 milliseconds, we assume that it took 200 milliseconds for the subject to change the imaged class after the instruction in this experiment. The time between the displayed instruction and the classifier's response to (apparent delay), and the time between the subject's actual change and the classifier's response (actual delay) are shown in Table 3. We can see that the proposed method ((D) and (E)) can reduce the time to response to the change by about 66% compared to the classifier with long input size (C). Finally, the transition of the output probability in Figure 3, that the output probability of the proposed method ((D) and (E)) rarely drops to nearly 50%, and the output is stable compared to the classifiers with short input size (A) and medium input size (B). Even if the estimated probability is not as high as 55%, for example in a two-class classification task, it will be judged as the correct answer, because the estimated probability of 50% is used as the threshold for classification. Therefore, the proposed method with stable output is expected to have higher performance in multiple-class classification tasks.

Table 2. Average percentage of correct answers in each phase

Blending Condition	Subject 1			Subject 2		
	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3
(A) 180×6	0.62	0.61	0.65	0.65	0.64	0.68
(B) 256×6	0.69	0.53	0.72	0.64	0.52	0.75
(C) 360×6	0.75	0.48	0.74	0.75	0.51	0.80
(D) 2+2+2	0.67	0.57	0.74	0.73	0.56	0.76
(E) 3+2+1	0.72	0.60	0.73	0.70	0.61	0.76

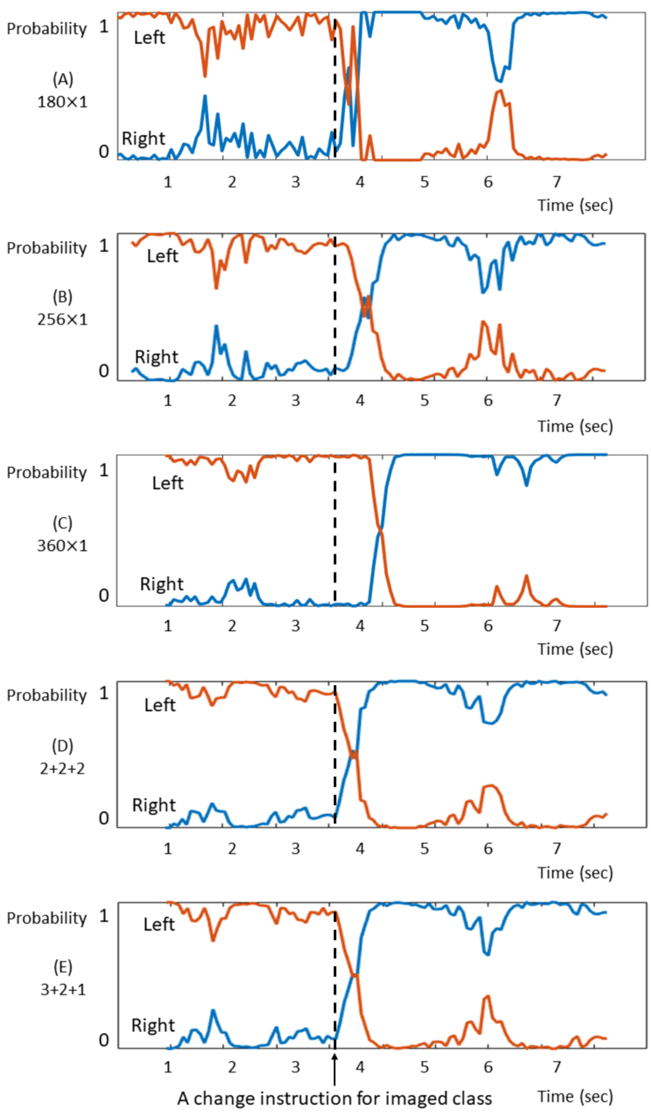


Figure 3. Transition of probability

Table 3. Time taken for the classifier's estimated class to change in the results of Figure 3.

Blending Condition	Apparent Delay	Actual Delay
(A)	0.3	0.1
(B)	0.5	0.3
(C)	0.8	0.6
(D)	0.4	0.2
(E)	0.4	0.2

5. CONCLUSION

In this study, we proposed a method of blending classifiers with different input sizes to achieve a BMI that can respond quickly to changes in an EEG. It is confirmed in experiments that the classification performance changes by changing the blend condition. In the section 4, we presented and discussed test cases that were successfully classified (Figure 3), but in reality, some test cases were not successfully classified. Although CNNs with short input size have the attraction of fast reaction time, their classification performance is still not good enough. Therefore, it is also an important research topic to improve the performance of classifiers with short input size. As a future prospect, we would like to verify whether the operability can be improved by changing the blend condition according to the situation, such as using a blended classifier with a higher ratio of CNNs with smaller input size in situations where there are many changes in an EEG, and using a blended classifier with a higher ratio of CNNs with larger input size in situations where there are few changes in an EEG.

ACKNOWLEDGEMENT

This work was supported in part by JSPS KAKENHI Grant Number JP20K11918.

REFERENCES

[1] Johan Wessberg, et al., “Real-time Prediction of Hand Trajectory by Ensembles of Cortical Neurons in Primates”, Nature, Vol. 408, No. 1, pp. 361-365, (2000).
[2] Vijay Khare, et al., “Controlling Wheelchair Using Electorencephalogram”, IJCSIS, Vol. 8, No. 2, (2010).
[3] Schirrmeister RT, et al., “Deep Learning With Convolutional Neural Networks for EEG Decoding and Visualization”, Human Brain Mapping, Vol. 38, No. 11, pp. 5391-5420, (2017).
[4] Ang KK, et al., “Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface”, IEEE Inter national Joint Conference on Neural Networks (IJCNN) 2008, pp. 2390-2397, (2008).
[5] Andreas Töscher and Michael Jahrer, The BigChaos Solution to the Netflix Grand Prize, https://www.netflixprize.com/assets/ GrandPrize2009_BPC_BigChaos.pdf.