Investigation of a Deep-Learning Based Brain–Computer Interface With Respect to a Continuous Control Application

MANFRED STRAHNEN AND PHILIPP KESSLER
Department of Computer Science, University of Applied Sciences Ulm, 89075 Ulm, Germany
Corresponding author: Manfred Strahnen (Manfred.Strahnen@thu.de)

ABSTRACT As part of a motor-imagery brain-computer interface (BCI), a deep neural network (DNN) must analyze measured electroencephalogram (EEG) data and identify neural signal patterns characteristic of a particular imagined motor movement. Our studies are intended to investigate the use of such a DNN in asynchronous online applications, where the EEG signals need to be interpreted continuously, as well as to gain insights into the learned neural patterns. We examined EEGNet, a commonly referenced convolutional neural net (CNN). In addition to the impacts of the size and temporal position of the trials used for training and testing the DNN on the classification accuracy, we examined the contributions and temporal behavior of known neural patterns and their effects on the response time of the system and the time period for which the mental state was stably recognized. Because the optimal temporal position of the trials is different for the neural patterns involved, we introduced ‘cropped training’, which is a training method in which the DNN is trained using trials with different temporal positions. This enabled the DNN to learn the neural patterns in the 0-8 Hz frequency range that are important for a short response time and the patterns in the 8-30 Hz frequency range that are important for determining the state duration. We show that cropped training is essential for achieving a good response time as well as for a good state duration.

INDEX TERMS Brain–computer interface (BCI), deep learning, motor-imagery, neural patterns.

I. INTRODUCTION
Brain-computer interfaces (BCI) are designed to establish a connection between the human brain and the environment. One possible application scenario is the control of an arm prosthesis based purely on the power of thought. Ideally, the imagined movement of the nonexistent arm should lead to an equivalent movement of the prosthesis. In most applications, an electroencephalogram (EEG) recording system is used as a non-invasive physical interface for the human brain.

The challenge in setting up a motor-imagery based BCI application is to recognize the signal patterns from the measured EEG that correspond to the subject’s mental intention. Typically, raw EEG data are first pre-processed, for example, by bandpass filtering. Subsequently, the defined signal features, such as the power spectral density present in certain frequency bands, are extracted and fed to a classifier. The result of the classifier can then be used to select the action to be performed [1].

For deep learning based BCIs, the processing stages listed above are ideally replaced by a single deep neural network (DNN). The convolutional neural network (CNN) architecture is the most frequently used architecture in the addressed application area [2]. A CNN commonly used in the field of BCIs is EEGNet [3]. Training such networks requires datasets with many trials. To acquire a trial, a subject must perform a specific mental action, for example, imagining the cyclic opening/closing of the left or right fist. The EEG signals recorded during this time, together with the label that classifies the recording (i.e., describes whether the left or right fist was cyclically opened/closed during the recording), form a trial. Because the acquisition of large datasets (many subjects with many trials each) is time-consuming and costly, commonly available EEG datasets, such as the Physionet...
dataset [4] or the BCI Competition IV 2a dataset [6] are often used for the training and analysis of DNNs.

The BCI system described thus far is a synchronous system, as the user is only allowed to perform mental actions in the time periods specified by a cue signal. An asynchronous BCI system eliminates this restriction. The user is allowed to perform the desired action in a self-paced manner, which brings the system closer to a comfortable practical application. The significantly higher complexity and current state of research on asynchronous BCI systems are described in detail in [7]. A recent review of the use and state of the art of asynchronous BCI systems, especially in the area of BCI-controlled vehicles, is given in [8].

A continuous control application is an application in which a DNN asynchronously recognizes defined self-paced Motor Imagery (MI) actions of the user from a continuously measured EEG data stream and initiates corresponding actions. From the continuously measured EEG data, overlapping trial time slices must be formed, which are passed to the DNN for evaluation. In addition to classification accuracy, the most important performance metrics for such applications are the response time and state duration. Response time is the time from starting the imagination of a certain movement until the DNN recognizes a new mental state and can take an appropriate action. The state duration is the time duration for which the system detects a stable motor imagination.

The results of this work contribute to the design of a deep-learning based asynchronous BCI system in the following respects:

1) We measured the impact of the temporal size of the trials used for training the DNN and the subsequent inference on the achievable classification accuracy and response time of the system. Our study provides guidance in determining the optimal trial size.

2) We investigated which EEG signal patterns, in the following called neural patterns, should be considered with respect to a good classification accuracy, response time, and state duration.

3) It can be shown that the introduction of the so-called ‘cropped training method’ significantly improves both the response time and state duration. In cropped training, the DNN is trained using multiple trials with different time shifts relative to the user action.

4) The influence of the time offset of a trial relative to user action on the overall performance of the system was investigated. This was necessary to choose the time-shifted trials required for cropped training.

The investigations presented here were performed using a Python/PyTorch-based BCI framework that we developed to support the training and analysis of DNNs and EEG datasets. DNNs can be evaluated based on the classification accuracy, i.e. the probability of correct classification, as well as the response time and state duration.

The remainder of this paper is organized as follows. Section II describes the main neural patterns found in the EEG data and their properties, as well as the DNN and EEG datasets used. Section III explains the methods used to obtain the results presented in section IV and discussed in section V. Finally, a conclusion and an outlook on future work are presented in Section VI.

II. BACKGROUND

A. DATASETS AND DATA SEGMENTATION

Our studies are based on the two popular, publicly available EEG datasets, the BCI Competition IV 2a EEG dataset [6], in the following referred to as BCIC, and the Physionet dataset [5], which was recorded with the BC12000 system [4], in the following referred to as PHYS. Both datasets were recorded following a cue-based BCI paradigm where different motor imagery tasks had to be performed. We used a subset of the recorded data, where the subjects had to perform two tasks, namely, the imagination of opening/closing of the left or right fist. The chronological sequence of each trial is shown in Figure 1. After a 2 s rest phase, a visual cue signal instructs the subject to continuously open/close the left or right fist. After an additional 4 s, the cue signal disappears and thereby tells the subject to stop the current task. Again a 2 second rest phase follows. The label indicating which fist had to be opened/closed was synchronously recorded along with the raw EEG data.

Differences between both datasets are given concerning the number of subjects and the number of trials per subject (Table 1). ‘Class’ means the type of task that the subject had to perform. According to Table 1 PHYS dataset is significantly bigger; it roughly provides 3.5 times more trials and 6.6 times more data.

The segmentation of the data, i.e., the partitioning of the data into data for training the DNN and data for testing the DNN, was done as follows: Following the 5-fold cross-validation method, the dataset was divided into five subsets of subjects, with each subset containing the same number of subjects. As shown in Figure 2, for each fold, a different subset was used to test the DNN, while the remaining subsets were used to train the DNN. This type of data segmentation guarantees that the DNN has never seen the data used for testing before.

B. NEURAL PATTERNS

In the following section, some of the most important neural patterns found in the EEG data are briefly described and characterized. Not all patterns are directly related to motor imagery; some are only visible as artifacts in measured EEG data.
TABLE 1. Characteristics of PHYS and BCIC dataset. For PHYS dataset, four of the 109 subjects were discarded due to variability in the number of trials.

| Parameter                  | PHYS dataset | BCIC dataset |
|----------------------------|--------------|--------------|
| Number of subjects         | 109          | 9            |
| Number of classes          | 2            | 2            |
| Trials per subject and class | 21          | 72           |
| Overall number of trials   | 105,221,410  | 9 * 2 * 72 = 1,296 |
| Number of electrodes       | 64           | 22           |
| Electrode positions according to the int. 10-10 system for PHYS, and according to the 10-20 system for BCIC | FCS, FC3, FC1, FCZ, FC2, FC4, FC6, C5, C3, C1, CZ, C2, C4, C6, C5, C3, C1, CZ, CPZ, CP2, CP4, CP6, CP1, CP2, CPZ, CP2, CP4, CP6, CP1, CPZ, CP2, CP4, P1, P2, POZ, FTP, FT8, T7, T8, T9, T10, T7, TP8, P7, P5, P3, P1, P2, P2, P4, P6, P8, PO7, PO3, POZ, PO4, PO8, O1, OZ, O2, IZ | FZ, FC3, FC1, FCZ, FC2, FC4, C5, C3, C1, CZ, CPZ, CP2, CP4, CP6, CP1, CP2, CP4, P1, P2, POZ, FTP, FT8, T7, T8, T9, T10, T7, TP8, P7, P5, P3, P1, P2, P2, P4, P6, P8, PO7, PO3, POZ, PO4, PO8, O1, OZ, O2, IZ |
| Sampling rate              | 160 Hz       | 250 Hz       |
| Overall number of sample points | 4,410,64,4s*160 Hz = 180,633,600 | 1,296,22,4s*250 Hz = 28,512,000 |

FIGURE 2. Data segmentation for 5-fold cross validation.

**P300** is the neural pattern which follows a visual stimulus. There is a large positive deflection of electrical activity \( \sim 300 \) ms after the stimulus. The P300 may occur when the cue signal becomes active.

Slow cortical potential (SCP) is a generic term for slow changes in potentials measured on the scalp, typically in the range of seconds. In the context of this study, movement-related cortical potential (**MRCP**) is of particular interest. This denotes a slow negative deflection of the EEG signal components lying in the frequency range of 0-5 Hz (delta frequency band) before the actual imagination of a movement [9]. The intended motion can be inferred from the measured spatial potential difference. Reference [10] reports a different amount of change in the electrical potential at electrodes C3 and C4 depending on the intended movement (left- or right-hand movement).

Sensor motor rhythm (SMR) is an umbrella term for the changes in the amplitudes of oscillations in different frequency bands observed during the execution or imagination of movements. In particular, the SMRs listed below are well known [11]:

**SMR-MU-DES:** Desynchronization in the alpha frequency band (8 – 12 Hz) over the sensorimotor cortex contralateral to the actual or imagined movement. Desynchronization refers to a decrease in the amplitude of oscillations lying in this frequency range.

**SMR-BETA-DES:** Desynchronization in the beta frequency band (18 – 26 Hz) over the sensorimotor cortex contralateral to the actual or imagined movement.

Table 2 summarizes the neural patterns identified so far and their respective signal characteristics.

Waldert et al. [12], [13] examined the neural patterns found in so-called center-out movement experiments. Their results can be used as a first indication of the correctness of our expectations. It has been reported that the characteristic signatures in the low-frequency band (0–7 Hz) are highly significant for decoding the direction of motion. Significant amplitude changes during motion execution have been reported. The pattern started approximately 250 ms before cue onset and decayed sharply during the first 750 ms after cue onset [13]. Additionally, in the motor cortex, a decrease in amplitude was observed for signals in the frequency range 10-30 Hz.

C. **EEGNet**

For the classification task at hand, there are already several CNN models from other scientific studies [14], [15], [16]. Our motivation for selecting EEGNet [3] is that this DNN is specifically designed for low-power embedded system applications. The core idea behind EEGNet is to combine multiple convolutional layers to extract and subsequently classify temporal and spatial features and patterns specific to the nature of motor-imagery EEG data. A basic overview of the model layers is presented in Figure 3.

The layer **Input** provides EEG data, which contains the time-series data recorded at C electrodes for a certain time (trial time slice), where T is the number of recorded samples per EEG channel. The second layer, **TemConv** is
We used the multitaper method [18] for PSD calculation. For each point, the corresponding mean value was calculated. The average power spectral densities were calculated and displayed from the time-frequency power spectral density. A first overview of the neural patterns present in a dataset can be obtained by calculating the band-specific average relative power spectral density for the frequency bands listed in Table 2.

B. BEST TRIAL SIZE

The duration of a trial is a critical parameter that on the one hand influences the performance of a DNN and on the other hand largely determines the resources that the DNN needs during inference. A trial that is too short would result in low-frequency neural signal patterns not being detected and thus not contributing to classification accuracy. A trial that is too large would result in an unnecessary amount of memory being required to store the trial data and DNN. In addition, excessively large trial time slices would lead to additional computing power requirements. Therefore, choosing the optimal trial size is of great importance for both the performance and resource requirements of the DNN.

The procedure to investigate the best trial size was as follows: EEG data was not pre-processed, except for the application of a notch filter to suppress the interference caused by the power supply (50 Hz or 60 Hz noise). EEGNet was trained and tested using trials starting at 0 s relative to the cue signal. The cross-fold classification accuracy was calculated by taking the mean value of 5 folds. Classification accuracy was measured for trial sizes from 0.2 s to 4 s in 0.1 s steps. The measurement was repeated ten times for each trial size. For each trial size, the mean, minimum, and maximum values were determined.

C. TRIAL TIME SLICE POSITION DEPENDENT ACCURACY

The results of the performed dataset analysis (Chapter IV.A) suggest that the temporal behavior of the identified neural patterns is not constant over the entire time period in which a subject imagines a certain movement. This behavior should also be visible in the time-resolved measured classification accuracy. Time-resolved measurement of the classification accuracy, which means measuring the accuracy for different trial time slice positions relative to the cue signal, is needed for the selection of suitable trials for cropped training (Chapter III.D).

To investigate the temporal behavior of the accuracy, the positions of the trial time slices used for training and testing the net were varied, as illustrated in Figure 4. The network was first trained and tested using the trial time slices with a constant length of 1 s and starting at position -2.0 s relative to the cue signal. Next, the network was trained and tested with the trial time slices starting at \((-2.0 \text{ s} + \text{step}\), where ‘step’ was chosen to be 0.2 s. The then next trial time slice starting position was again shifted by ‘step’ to the right. The last trial time slice position was at +4.2 s, which is directly after switching the cue signal off. For each trial time slice position

\[
t_{\text{off}} = -2.0 \text{ s} + i \cdot \text{step} \quad \text{with} \quad 0 \leq i \leq 31
\]
the 5-fold cross-validation classification accuracy was measured. Because we also wanted to investigate the temporal behavior of the neural patterns described in chapter II.B, EEG data was filtered with the filter configurations specified in Table 3. Filter configuration ‘SCP’ was used to extract the behavior of the MRCP pattern, ‘Alpha’ for the SMR-Mu-Des pattern and ‘Beta’ for the SMR-Beta-Des pattern. Filtering was done using 5th-order Butterworth bandpass filters.

D. POTENTIAL PERFORMANCE OF A CONTINUOUS CONTROL APPLICATION

A continuous control application is an application in which a DNN is supposed to recognize defined motor imagery states from a continuously measured EEG data stream and to initiate corresponding actions. Therefore, overlapping trial time slices are taken from the continuous EEG data stream and passed to the DNN for evaluation. With respect to this application scenario, the mean time-resolved classification accuracy of EEGNet was measured as follows: EEGNet was trained with time slices starting at time 0 s relative to the cue signal, and having a temporal duration of 1 s. The best-fold network, which is the one with the best classification accuracy out of the five folds, was taken and tested with trial time slices with an offset toff relative to the cue signal. toff was varied from -2.0 s to 4.2 s in steps of tstep = 50 ms. For any offset position the mean classification accuracy was calculated. The trial time slices used for testing belong to subjects who, according to the data segmentation plan presented in Chapter II.A, were not considered when creating the time slices used for training. The described procedure allows conclusions to be drawn regarding a continuous control application in that the network, once trained, is tested with test trial time slices that were acquired asynchronously to the trial time slices used for training. For this purpose, each trial (with a time range from -2.0 s to 4.2 s) is treated as a continuous data stream, from which the EEG trial time slices (time duration: 1 s) needed for the continuous inference are cut out. A new inference was made every 50 ms.

The following two metrics were used to evaluate the results: the response time treact70 is the time from the activation of the cue signal, which is supposed to be the point in time when the user starts to imagine the requested task, until the DNN detects the corresponding mental state with a classification accuracy of at least 70%. According to Figure 6, toff70 marks the beginning of the test trial time slices of duration tsize, which lead to a mean classification accuracy of at least 70%. These test trial time slices are submitted to the DNN for evaluation at time toff70 + tsize. After the processing time tproc required for the evaluation, the result is available, and the first reaction can be performed. The following applies to treact70:

\[ t_{react70} = toff70 + t_{size} + t_{proc} \]

The second parameter used as a metric is the state duration time tstate70, which is also marked in Figure 6. tstate70 is the duration for which the DNN continuously detects a classification accuracy of at least 70%. This value should be as close as possible to the time for which the cue signal is active, which would mean that the DNN...
detects a given movement (with an accuracy of 70%) for almost the entire time that the subject imagines the movement. The results presented in chapter IV.D show that the DNN trained as described above delivers a very poor accuracy of almost the entire time that the subject imagines the movement. For this reason, the so-called cropped training was introduced, which was already used by Schirrmeister et al. [14] to optimize the classification accuracy of a DNN. In cropped training, the DNN is not trained with trials at a fixed temporal position, but, as shown in Figure 5, with several overlapping trial time slices starting at different times. The underlying idea is that the DNN learns the characteristics of not just one, but multiple temporal trial time slice positions. Similar approaches, where a classifier must learn the features valid at different points in time, have already been used by others [19], [20].

IV. RESULTS
A. DATASETS ANALYSIS
The methodology presented in chapter III.A) for examining the EEG dataset was applied to the BCIC and PHYS dataset. Specifically, the existence of the expected neural patterns was investigated. The upper left part of Figure 7 shows the time-frequency PSD map for the BCIC dataset, the lower left part the one for the PHYS dataset. The state at time \( t = -1.2 \text{ s} \) was defined as the resting state. The upper and lower right parts of Figure 7 show the average relative power spectral densities for the frequency bands where the presumed characteristic neural patterns were expected to occur. Here, the average power spectral density in the 0-7 Hz frequency band at time \( t = -1.2 \text{ s} \) was used to calculate the average relative power spectral density.

The P300, MRCP, SMR-Mu-Des, and SMR-Beta-Des neural patterns are clearly visible in Figure 7. Since the P300 pattern is not specific to the type of movement performed, only MRCP, SMR-Mu-Des, and SMR-Beta-Des remain potential candidates for building a BCI system. MRCP can be clearly identified as the ‘strongest’ neural pattern, about twice as strong as SMR-Mu-Des, which in turn is significantly stronger than SMR-Beta-Des.

B. BEST TRIAL SIZE
To determine the best trial size, 5-fold classification accuracy was measured for different trial sizes, as described in chapter III.B. Figure 9 shows the mean, minimum, and maximum classification accuracies for the BCIC and PHYS datasets. For the BCIC dataset the maximum mean accuracy of 80.7% was achieved with a trial size of 2.5 s. It is worth noting that relatively high-accuracy values were measured even for small trial sizes. An accuracy of 77.9% was achieved with a trial size of only 400 ms, which is only 2.8% below the maximum. However, the necessary amount of input data is only 16% of the amount of data required to reach the maximum value. For the PHYS dataset, an accuracy of 77.4% was measured at 400 ms trial size. Small trial sizes significantly reduce the memory and computational requirements of the computer system that must perform the inferencing, but at the cost of lower classification accuracy. As a compromise, we have chosen a trial time slice size of 1 s for our investigations, unless otherwise specified.

C. TRIAL TIME SLICE POSITION DEPENDENT ACCURACY
The temporal behavior of the 5-fold classification accuracy was measured by training and testing EEGNet with trial time slices with a temporal offset relative to the cue signal, as described in chapter III.C. To extract the individual temporal behavior of each neural pattern, measurements were performed for each of the filter configurations given in Table 3. Accordingly, in Figure 8, the results labeled ‘SCP’ are due to the MRCP patterns, the results labeled ‘Alpha’ due to the SMR-Mu-Des pattern, and those labeled ‘Beta’ due to the SMR-Beta-Des pattern.

The following statements can be made for the BCIC dataset (Figure 8):
- The overall accuracy (filter configuration ‘All’) strongly depends on the position of the trial time slices used for training and testing. It is best when the trial time slice starting point is in the range from \( -0.7 \text{ s} \) to \( +0.4 \text{ s} \). (Notice that a trial time slice starting at \(-0.7 \text{ s}\) means that 70% of the included data has been recorded before, and 30% has been recorded after cue onset!)
- For trial time slice offsets in the range \( -0.7 \text{ s} \) to \( 0.4 \text{ s} \) the accuracy is dominated by the information given in the SCP frequency band, whereas for trial time slice offsets above \( 0.4 \text{ s} \), ‘Alpha’ and ‘Beta’ band contributions are more significant.
- Neural patterns learned from the ‘Alpha’ and ‘Beta’ frequency bands are delayed by approximately \( 0.5 \text{ s} \) compared to those learned from the ‘SCP’ band.

These results provide a first indication that the neural pattern in the SCP band seems to be important for a short response time \( t_{\text{react70}} \), whereas the patterns lying in the alpha and beta bands contribute more to a good state duration \( t_{\text{state70}} \).

The differences between the BCIC and PHYS datasets can be described as follows:
- For the PHYS dataset the accuracies we get when ‘All’ frequency bands are used and when exclusively the information given in the ‘SCP’ band is used are nearly the same. ‘Alpha’ and ‘Beta’ band information seem to have no significant impact on accuracy.
- For BCIC dataset there is a delay of \( \sim 0.5 \text{ s} \) between the neural pattern learned from ‘Alpha’ or ‘Beta’ bands and the ones learned from ‘SCP’ band. For PHYS dataset there is only a delay between the patterns learned from Alpha and SCP bands, but no delay between those learned from Beta and SCP.
- For BCIC, a sharp drop in accuracy is observed at a trial time slice offset of more than \( 1 \text{ s} \). For PHYS dataset, it remains at a high level until the end of the trial.

D. POTENTIAL PERFORMANCE OF A CONTINUOUS CONTROL APPLICATION
In the scenario investigated here, an EEGNet instance trained with trial time slices with a fixed offset of \( t_{\text{df}} = 0 \text{ s} \) was tested with trial time slices with an offset in the range
of: $-2.0 \text{s} < t_{\text{off}} < 4.2 \text{s}$. $t_{\text{off}}$ was varied in steps of $t_{\text{step}} = 50 \text{ms}$. The choice of $t_{\text{step}} = 50 \text{ms}$ is reasonable because even low-power microcontrollers can perform an inference in this time period [21]. Figure 10 shows the measured classification accuracy as a function of $t_{\text{off}}$ for the BCIC dataset. A significant classification accuracy is only obtained when the temporal positions of the training and test trials are almost identical. More precisely, the response time introduced in chapter III.D is $t_{\text{react}70} = -0.05 \text{s} + 1.0 \text{s} + 50 \text{ms} = 1.0 \text{s}$ and the state duration is $t_{\text{state}70} = 150 \text{ms}$.

These very poor values mean that the DNN detects the imagined motion after a delay of one second for a very short period of only 150 ms. Theoretically, the subject should have imagined the motion for 4 s. The results documented in chapter IV.A confirm this assumption. As shown in Figure 7, the neural patterns resulting from an imagined movement are active for significantly longer than 150 ms. However, the results from section IV.A also show that the intensities of the neural patterns vary with time, which ultimately led us to the method of cropped training, which should allow the DNN to learn the time course of the neuronal patterns present in the EEG signals. The cropped training configurations given in Table 4 were investigated. The contribution of each neural pattern was examined by applying the bandpass filter indicated in Table 3 for each. The DNN was trained using different partially overlapping crops. For example, in the All-Crop1 configuration, the DNN was trained using only unfiltered trial time slices with an offset of 0 s. In the Beta-Crop13 configuration, EEG data was first filtered with a 12-30 Hz bandpass filter and then EEGNet was trained with 13 crops with the offset values given in Table 4.

The temporal locations of the crops used for training were selected according to the following rules:

- When only one crop is used, its temporal position corresponds to the position of the associated maximum, as shown in Figure 8.
M. Strahnen, P. Kessler: Investigation of a Deep-Learning Based Brain–Computer Interface

FIGURE 8. Temporal behavior of the frequency band specific cross-fold classification accuracies for a time slice size of 1 s.

FIGURE 9. Classification accuracy depending on the trial time slice size.

• When using multiple crops, the temporal positions of these crops are such that for each of these crops,

a classification accuracy of at least 60% is achieved, as shown in Figure 8.

Figure 12 shows the time course of the measured continuous application classification accuracy for the BCIC dataset, for each of the configurations given in Table 4. Based on this, the metrics listed in Table 5 were calculated. The findings that can be derived from this study are:

• According to Figure 8, the neural patterns lying in the low-frequency SCP band are dominant in the time range $-0.7 \text{s} < t_{\text{off}} < 0.4 \text{s}$. When trained with only one crop lying within this time range, the DNN learns only the characteristics of this crop. With only a small time shift of the test trial time slices compared to the trial time slices used for training, the accuracy decreased significantly (see Figure 12, configurations All-Crop1 and SCP-Crop1). The use of multiple closely spaced crops leads to a significant improvement in

FIGURE 10. BCIC dataset: Mean classification accuracy in a continuous control application, where EEGNet is trained only with trial time slices starting at position $t_{\text{off}} = 0 \text{s}$.

TABLE 4. Configurations used for cropped training.

| Configuration | Bandpass region | Number of crops | Crop $t_{\text{off}}$ values (in s) |
|---------------|-----------------|-----------------|-----------------------------------|
| All-Crop1     | -               | 1               | 0.0                               |
| All-Crop20    | -               | 20              | -0.75, -0.625, -0.5, -0.375, -0.25, -0.125, 0.0, 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, 0.875, 1.0, 1.125, 1.25, 1.375, 1.5, 1.675 |
| SCP-Crop1     | 0-8 Hz          | 1               | 0.0                               |
| SCP-Crop13    | 0-8 Hz          | 13              | -0.5, -0.375, -0.25, -0.125, 0.0, 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, 0.875, 1.0                               |
| Alpha-Crop1   | 8-12 Hz         | 1               | 0.75                              |
| Alpha-Crop13  | 8-12 Hz         | 13              | 0.0, 0.125, 0.25, 0.375, 0.5, 0.75, 0.875, 1.0, 1.125, 1.25, 1.375, 1.5, 1.675                           |
| Beta-Crop1    | 12-30 Hz        | 1               | 0.5                               |
| Beta-Crop13   | 12-30 Hz        | 13              | 0.0, 0.125, 0.25, 0.375, 0.5, 0.75, 0.875, 1.0, 1.125, 1.25, 1.375, 1.5, 1.675                     |
TABLE 5. Continuous application response and state duration times for different ‘cropped training’ configurations, with $t_{size} = 1.0$ s and $t_{proc} = 50$ ms for BCIC dataset.

| Configuration   | Response time $t_{react70}$ (s) | State duration time $t_{state70}$ (s) |
|-----------------|---------------------------------|--------------------------------------|
| All-Crop1       | 1.0                             | 0.15                                 |
| All-Crop20      | 0.8                             | 1.75                                 |
| SCP-Crop1       | 1.0                             | 0.15                                 |
| SCP-Crop13      | 0.55                            | 1.3                                  |
| Alpha-Crop1     | 1.45                            | 1.45                                 |
| Alpha-Crop13    | 1.25                            | 1.55                                 |
| Beta-Crop1      | 1.3                             | 1.0                                  |
| Beta-Crop13     | 1.0                             | 2.1                                  |

The state duration. If only one crop is used, the state duration is 150 ms (All-Crop1), and if 20 crops are used (All-Crop20), the value of $t_{state70} = 1.75$ s is achieved.

- According to Figure 8, the neural patterns located in the alpha and beta bands are dominant in the time range $t_{off} > 0.4$ s. In contrast to the scenario described above, in the case of the alpha and beta bands, no significant drop in accuracy is observed with a small time shift of the test trial time slices compared to the trial time slices used for training (see Figure 12, configurations Alpha-Crop1 and Beta-Crop1). The use of multiple crops leads to a less significant improvement in the state duration $t_{state70}$.

The SCP-Crop13 configuration provided the best result for the response time and the Beta-Crop13 configuration for the state duration. The All-Crop20 configuration can be regarded as a good compromise. With this configuration, the DNN achieved a response time of 0.8 s. The state was detected with a continuous classification accuracy of at least 70% for a duration of 1.75 s. In addition, the effort required for bandpass filtering is omitted.

Similar results were obtained with the PHYS dataset. Based on the findings obtained with the BCIC dataset, the studies were performed only for not bandpass filtered EEG data. The left part of Figure 11 shows the mean classification accuracy when the DNN was trained only with trial time slices that are at $t_{off} = 0$ s. The response time is $t_{react70} = 850$ ms and the state duration time is $t_{state70} = 350$ ms. When using 20 crops evenly distributed over the period $-0.75 < t_{off} < 3.5$ s, the response time improves to $t_{react70} = 550$ ms and the state duration time to $t_{state70} = 1.15$ s (Right part of Figure 11).

Trial slice sizes of 1.0 s were used to achieve all the results presented so far. The duration of a trial time slice has a large impact on the computational power and memory requirements of the computer used and, as described in chapter IV.B, on the classification accuracy. To investigate the influence of the trial time slice size on the response time and the state duration time, we additionally measured these values with trial time slice sizes of 400 ms and 1800 ms. The investigations were performed for the scenario identified so far as optimal (not bandpass filtered EEG signals, use of the cropped training method). The results listed in Table 6 lead to the following conclusions:

TABLE 6. Response time and state duration time for different trial slice sizes. (EEGNet always has been trained with 20 crops of not bandpass filtered data).

| Trial slice size | BCIC | PHYS |
|------------------|------|------|
| (in ms)          | $t_{react70}$ (in ms) | $t_{state70}$ (in ms) | $t_{react70}$ (in ms) | $t_{state70}$ (in ms) |
| 400              | 700  | 700  | 550  | 500  |
| 1000             | 800  | 1750 | 550  | 1150 |
| 1800             | 900  | 2500 | 700  | 1350 |
The trial slice size has a small effect on the response time. The response time achievable with an EEGNet trained according to the cropped training method is 700-900 ms for the BCIC dataset, and 550-700 ms for the PHYS dataset.

In contrast to the response time, the trial slice size has a large impact on the state duration time. For the BCIC dataset, the state duration time improves from 700 ms (for \( t_{\text{size}} = 400 \) ms) to 2500 ms (for \( t_{\text{size}} = 1800 \) ms), for the PHYS dataset from 500 ms to 1350 ms.

V. DISCUSSION AND SUMMARY

Using the method presented in chapter III.A to qualitatively examine EEG data sets, important neural patterns that may occur during the imagination of movements were identified and qualified for both data sets in terms of changes in their signal strength and temporal location. The strongest neural pattern in terms of change in signal strength was identified in the low-frequency range of 0–8 Hz. For patterns located in the adjacent alpha range (8–12 Hz), a 50% smaller change in signal strength was observed. In the beta range (12–30 Hz), it was significantly lower.

The detection of dominant neural patterns located in the low-frequency range requires a certain minimum temporal size of the trials. According to our investigations, this is approximately 400 ms. In this case, the accuracy achieved with the DNN was only 3-5% below the maximum achievable value. At the same time, however, only 16% of the data required to achieve the maximum accuracy is needed, which in turn has a direct impact on the memory and compute power requirements of the executing computer.

The question of which neural patterns contribute most to a high classification accuracy is important from a neurophysiological as well as from a technical point of view. From a technical point of view, the effort involved in filtering out certain neural patterns can only be justified if a higher classification accuracy is achieved. Our investigations show that for a simple left/right classification, the low-frequency SCP band (0-8 Hz) is dominant and seems to be sufficient.

Using only low-frequency signal components would have significant technical advantages, but unfortunately, this approach is too simplistic. The results presented in chapter IV.D and summarized below show that the neural patterns lying in the alpha and beta bands are important for practical use in the context of a BCI application in that they can be used to optimize the state duration time.

The question about the best temporal positions of the time slices used for training the DNN can be answered as follows: The classification accuracy of the DNN used here is best if the trial time slices used for training have a temporal offset to the cue signal that is in the range \(-0.4\) s to 0 s (Figure 8). In addition, the highly variable temporal behavior of the contributions of the identified neuronal patterns is important. For the BCIC dataset, the contributions coming from the alpha and beta bands are shifted by about 0.5 s, which means that when using trial time slices starting at time \( t = 0.5 \) s, the contributions to classification accuracy from these bands are higher than those from the SCP band. This finding was taken into account in the choice of appropriate trial time slice locations for the cropped training, as described in chapter IV.D.

For the practical application of a DNN in the context of a BCI application, the measurement of the continuous classification accuracy as presented in chapter III.D is particularly important. The method presented there additionally supports the measurement of the response time and state duration time. The results presented in chapter IV.D clearly show that without the use of cropped training, the performance of the DNN is rather modest. With the help of cropped training, the response time could be improved by \( \sim 200 \) ms for the BCIC and by 300 ms for the PHYS dataset. The state duration time could be improved by 1.6 s for the BCIC and by 0.8 s for the PHYS dataset. In addition, the impact of the trial slice size on the response time and the state duration time was investigated. In contrast to the state duration time, the trial slice size has only a minor impact on the response time. Thus, if the focus is on determining the start time of an imagined motion, a trial slice size as small as possible (400 ms) should be chosen. If the duration of the imagined movement is also important, a larger trial slice size (>1 s) should be taken.

The measured response time values roughly agree with those published in the literature. Lopez-Larraz et al. [22] measured a response time of 400-800 ms for spinal cord injury patients. Salavanis et al. [23] report a response time in the range of 500-700 ms. Furthermore, both sources report a dependency of response time on trial slice size, which is comparable to the dependency we found. However, neither team used deep-learning methods, so our research can be considered the first indication that deep-learning-based BCIs provide comparable results in terms of response time to those using conventional feature extraction and classification methods as described in [1].

Furthermore, the effects of different neural patterns on the response time and duration of the detected state were investigated. The results obtained in this respect confirm the investigations conducted by Waldert et al. [12] and Hehenberger et al. [24], according to which the direction of an intended movement is essentially encoded by signal components lying in the low-frequency signal range. One explanation for the poor state duration value when using the low-frequency signal components (SCP-Crop1, \( t_{\text{state70}} = 150 \) ms) could be that the phase of the signal is essential for information encoding in this frequency range. With only one crop, the DNN only learns the phase that is valid in this period. For example, if the DNN is confronted with a test-trial time slice shifted by 200 ms, this would be equivalent to a signal with a significantly different phase and could therefore no longer be detected. The duration of the imagined movement can be estimated much better with the neural patterns located in the alpha and beta bands. In addition, training with only one crop provides a much better value for the state duration than the scenario using the neural patterns in the low-frequency range. A possible explanation could be that the duration of the movement is
essentially encoded by the amplitudes of the brain waves in the alpha and beta bands. A slight temporal shift of the test trials compared to the training trials has almost no effect on the resulting classification accuracy.

VI. CONCLUSION

Since the neural patterns induced by motor imagery are not constant in time, cropped training was introduced. This allows the DNN to learn both, the neural patterns in the low-frequency band (SCP band), which are important for a short response time, and the patterns in the alpha and beta bands, which are needed for mental state duration detection. With cropped training, both parameters improved significantly.

We are currently planning to build a system to measure EEG signals induced by imagined movements and to use this system as part of a continuous control application that incorporates the findings published in this paper.

REFERENCES

[1] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yage, “A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update,” J. Neural Eng., vol. 15, no. 3, Jun. 2018, Art. no. 031005, doi: 10.1088/1741-2552/AAB2FA.

[2] X. Zhang, L. Yan, X. Wang, J. Monaghan, D. Mcclpine, and Y. Zhang, “A survey on deep learning-based non-invasive brain signals: Recent advances and new frontiers,” 2019, arXiv:1905.04149.

[3] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lanece, “EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces,” J. Neural Eng., vol. 15, no. 5, Oct. 2018, Art. no. 056013, doi: 10.1088/1741-2552/aacec8.

[4] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, “BICE2000: A general-purpose brain-computer interface (BCI) system,” IEEE Trans. Biomed. Eng., vol. 51, no. 6, pp. 1034–1043, Jun. 2004.

[5] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, “PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals,” Circulation, vol. 101, no. 23, pp. e215–e220, 2000.

[6] C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller, “BCI competition 2008—Graz data set A,” Inst. Knowl. Discovery, Lab. Brain-Comput. Interfaces, Graz Univ. Technol., Graz, Austria, 2008. Accessed: Dec. 12, 2022. [Online]. Available: http://bnci-horizon-2020.eu/database/data-sets.

[7] G. R. Müller-Putz, “Feel your reach: An EEG-based framework to continuously detect goal-directed movements and error processing to gate kinesiethetic feedback informed artificial arm control,” Frontiers Hum. Neurosci., vol. 16, p. 110, Jan. 2022. Accessed: May 15, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnhum.2022.841312.

[8] A. Hekmatmanesh, P. H. J. Nardelli, and H. Handroos, “Review of the state-of-the-art of brain-controlled vehicles,” IEEE Access, vol. 9, pp. 110173–110193, 2021, doi: 10.1109/ACCESS.2021.3100700.

[9] R. Cunnington, R. Iansik, J. Bradshaw, and J. Phillips, “Movement-related potentials associated with movement preparation and motor imagery,” Exp. Brain Res., vol. 111, no. 3, pp. 429–436, Oct. 1996, doi: 10.1007/BF00228732.

[10] G. Dornhege, B. Blankertz, and G. Curio, “Speeding up classification of multi-channel brain-computer interfaces: Common spatial patterns for slow cortical potentials,” in Proc. 1St Int. IEEE EMBS Conf. Neural Eng., Capri, Italy, 2003, pp. 595–598, doi: 10.1109/CNE.2003.1196989.

[11] H. Yuan and B. He, “Brain–computer interfaces using sensorimotor rhythms: Current state and future perspectives,” IEEE Trans. Biomed. Eng., vol. 61, no. 5, pp. 1425–1435, May 2014, doi: 10.1109/TBME.2014.2312397.

[12] S. Waldert, H. Preis, E. Demandt, C. Braun, N. Birbaumer, A. Aertsen, and C. Mehring, “Hand movement direction decoded from MEG and EEG,” J. Neurosci., vol. 28, no. 4, pp. 1000–1008, Jan. 2008, doi: 10.1523/JNEUROSCIENCE.1571-07.2008.

[13] S. Waldert, T. Pistohl, C. Braun, T. Ball, A. Aertsen, and C. Mehring, “A review on directional information in neural signals for brain-machine interfaces,” J. Physiol., vol. 103, nos. 3–5, pp. 244–254, May 2009, doi: 10.1016/J.JPHYSIOPARIS.2009.08.007.

[14] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, “Deep learning with convolutional neural networks for EEG decoding and visualization: Convolutional neural networks in EEG analysis,” Hum. Brain Mapping, vol. 38, no. 11, pp. 5391–5420, Nov. 2017, doi: 10.1002/HBM.23730.

[15] B. Abibullaev, I. Dolzhikova, and A. Zollanvari, “A brute-force CNN model selection for accurate classification of sensorimotor rhythms in BCIs,” IEEE Access, vol. 8, pp. 101014–101023, 2020, doi: 10.1109/ACCESS.2020.2997681.

[16] H. Dose, J. S. Møller, H. K. Iversen, and S. Puthusserypady, “An end-to-end deep learning approach to MI-EEG signal classification for BCIs,” Expert Syst. Appl., vol. 114, pp. 532–542, Dec. 2018, doi: 10.1016/J.ESWA.2018.08.031.

[17] Xiayang Wang, (Nov. 20, 2020), Q-EEGNET_Torch, Accessed: Dec. 12, 2021. [Online]. Available: https://github.com/xiaywang/q-eeegnet_torch.

[18] B. Babadi and E. N. Brown, “A review of multitaper spectral analysis,” IEEE Trans. Biomed. Eng., vol. 61, no. 5, pp. 1555–1564, May 2014, doi: 10.1109/TBME.2014.2311996.

[19] E. López-Larraz, L. Montesano, A. Gil-Agudo, and J. Minguez, “Continuous decoding of movement intention of upper limb self-initiated analytic movements from pre-movement EEG correlates,” J. Neuroeng. Rehabil., vol. 11, no. 1, p. 153, 2014, doi: 10.1186/1743-0003-11-153.

[20] D. Liu, W. Chen, K. Lee, R. Chavarriaga, F. Iwane, M. Bouri, Z. Pei, and J. D. R. Millan, “EEG-based lower-limb movement onset decoding: Continuous classification and asynchronous detection,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 26, no. 8, pp. 1626–1635, Aug. 2018, doi: 10.1109/TNSRE.2018.2855053.

[21] X. Wang, M. Hersche, B. Tönnke, K. Baya, M. Magnno, and L. Benini, “An accurate EEGNet-based motor-imagery brain–computer interface for low-power edge computing,” 2020, arXiv:2004.00077.

[22] E. López-Larraz, J. M. Antelis, L. Montesano, A. Gil-Agudo, and J. Minguez, “Continuous decoding of motor attempt and motor imagery from EEG activity in spinal cord injury patients,” in Proc. Ann. Int. Conf. IEEE Eng. Med. Biol. Soc., Aug. 2012, pp. 1798–1801, doi: 10.1109/EMBC.2012.6464299.

[23] M. Salvaris and P. Haggard, “Decoding intention at sensorimotor timescales,” PLoS ONE, vol. 9, no. 2, Feb. 2014, Art. no. e85100, doi: 10.1371/JOURNAL.PONE.0085100.

[24] L. Hehenberger, L. Bastitic, I. A. Strublea, and G. R. Müller-Putz, “Directional decoding from EEG in a center-out motor imagery task with visual and vibrotactile guidance,” Frontiers Hum. Neurosci., vol. 15, Sep. 2021, Art. no. 687252, doi: 10.3389/FNHUM.2021.687252.