



Universiteit  
Leiden  
The Netherlands

## **Towards high performance and efficient brain computer interface character speller : convolutional neural network based methods**

Shan, H.

### **Citation**

Shan, H. (2020, February 25). *Towards high performance and efficient brain computer interface character speller : convolutional neural network based methods*. Retrieved from <https://hdl.handle.net/1887/85675>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/85675>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/85675> holds various files of this Leiden University dissertation.

**Author:** Shan, H.

**Title:** Towards high performance and efficient brain computer interface character speller : convolutional neural network based methods

**Issue Date:** 2020-02-25

# Chapter 1

## Introduction

A Brain Computer Interface (BCI), also known as mind-machine interface, translates brain signals into computer commands, thereby building communication between the human brain and outside devices. In this way, human-beings can use only the brain to express their thoughts without any real movement. As a result, BCIs become an important communication pathway for the people who lose motor ability, such as patients with Amyotrophic Lateral Sclerosis (ALS) [SD06] or spinal-cord injury. In recent years, BCIs have also been popularly developed for healthy people, in application domains such as entertainments [GP<sup>+</sup>13], mental state monitoring [LTK13], virtual reality [CBJ16] as well as in IoT services [LL<sup>+</sup>14].

A BCI system consists of three components, as shown in Figure 1.1. The first component is the brain signal acquisition. In this component, the brain signals of a subject (person) are recorded by using a brain headset equipped with a number of sensors. The acquired brain signals are sent to the second component for brain signal processing and translation. In this component, a hardware/software platform is used to process and translate brain signals into computer commands. Then, in the last component, the translated commands, i.e., the control signals, are used to control the outside devices, e.g., a prosthesis [MPP08], a computer mouse [Spü15], a mobile phone [CCH<sup>+</sup>10], or a robot [BFL13].



Figure 1.1: Workflow of a typical BCI.

Depending on the placement of the sensors, which are used to acquire brain sig-

nals in the brain headset, BCI systems can be categorized as invasive BCIs, semi-invasive BCIs, and non-invasive BCIs [Wal16]. In invasive BCIs, micro-sensor arrays are placed directly into the cortex [PHP10] to measure action potentials (APs) and local field potentials (LFPs). In semi-invasive BCIs, sensors are placed on the exposed surface of the brain in order to measure electrocorticography (ECoG) signals [SL11]. In non-invasive BCIs, sensors are placed on the scalp in order to acquire electroencephalography (EEG) signals [GS06]. In recent decades, EEG-based BCIs attract most of the BCI research due to their non-invasive, easy, and safe way of acquiring brain signals. EEG-based BCIs can be divided in four main categories [FRAG<sup>+</sup>12], namely P300-based BCIs [FD88], steady state visual evoked potential (SSVEP)-based BCIs [Her01], event related desynchronization (ERD)-based BCIs [PN01], and slow cortical potential-based BCIs [BKG<sup>+</sup>00]. Compared with the other categories of EEG-based BCIs, the P300-based BCIs have the following advantages. The P300-based BCIs are effective for almost every BCI user because the P300 signal, which is the target signal used in the P300-based BCIs, can be evoked in the brain of almost every human being [Els09]. In addition, the P300-based BCIs are relatively fast and straightforward to use. Moreover, the P300 signals work outstandingly well for BCI character spelling applications [GDS<sup>+</sup>09]. Therefore, the P300-based BCIs have attracted a lot of BCI researchers. As the benchmark for a P300-based BCI [FRAG<sup>+</sup>12], the P300 speller [FD88] has been the most-commonly investigated application of the P300-based BCI [FRAG<sup>+</sup>12]. Thus, this dissertation takes the P300 speller as the target BCI application.

## **1.1 Development Trends in P300-based Brain Computer Interface Systems**

P300-based BCIs are still not used in human's daily life and remain in an experimental stage at research labs. In order to bring P300-based BCIs into practical use, currently, there are two development trends for P300-based BCI systems, i.e., to design high performance P300-based BCI systems and to design efficient P300-based BCI systems.

### **1.1.1 High Performance P300-based Brain Computer Interface Systems**

The performance of a P300-based BCI system is the communication accuracy and the communication speed between the human brain and a computer. For example, for the P300 speller, the communication accuracy of such BCI system is the character spelling accuracy. The communication speed of such system is the Information Transfer Rate (ITR) [WRMP98]. The current performance of a P300-based BCI system is relatively low because the P300 signals are buried in a lot of noise and thus,

the P300 signals have a very low Signal to Noise Ratio (SNR). This makes it difficult to detect P300 signals evoked in the human's brain, resulting in a low communication accuracy and speed of the P300-based BCI systems. P300-based BCI systems with such low performance are not acceptable for BCI users in their daily life. We take the P300 speller, the most-widely used application of the P300-based BCIs, as an example. Guy [GSB<sup>+</sup>18] explored the usability of the current P300 spellers for disabled people with amyotrophic lateral sclerosis. This report shows that when using a current P300 speller, half of the subjects (persons) cannot spell characters with accuracy that is higher than 90%. To promote P300 spellers to be used in people's daily life, we should try our best to make the subjects spell characters with a P300 speller like the healthy people spell characters with their mouth. This means that we should try to make the subjects who use a P300 speller to achieve the character spelling accuracy that is (or close to) 100%. A P300 speller with accuracy that is much lower than 100% cannot be used in people's daily life. In addition, Guy's report [GSB<sup>+</sup>18] also shows that when using the current P300 spellers, the mean number of characters correctly spelled by the subjects is 3.6 characters per minute. However, a healthy person is able to speak with around 120 characters per minute. Compared with 120 characters per minute, the communication speed of the current P300 spellers, i.e., 3.6 characters per minute, is far from what is needed to be used in people's daily life. Therefore, increasing the performance of the P300-based BCIs is a must in order to promote the P300-based BCIs into people's daily life.

To increase the performance of a P300-based BCI system, efforts are focused on the signal acquisition part and on the signal processing and translation part of a BCI system. In the signal acquisition part, researchers try to improve the recording quality of the EEG signals such that the signals, that contain P300 evoked potentials, have less noise. For example, Koka [KB07] has developed tripolar concentric sensors. These sensors use advanced engineering techniques to enhance the recording capability for brain signals. Unfortunately, the current signal recording techniques cannot provide high enough SNR for P300 signals, thereby not guaranteeing alone very good performance of a P300-based BCI system.

In recent years, massive efforts have been put in the signal processing and translation part of a P300-based BCI system. In order to increase the performance of the P300-based BCI system, a lot of studies have been done in terms of devising preprocessing, feature extraction, and classification methods for P300-based BCI systems. In terms of preprocessing EEG signals, different signal processing techniques are used, such as bandpass filtering [CG11], discrete-wavelet transform (DWT) [SS09], continuous wavelet transform (CWT) [Bos04]. In terms of feature extraction methods for P300-based BCIs, Rivet [RS<sup>+</sup>09] uses an unsupervised algorithm to enhance P300 evoked potentials, Kulasingham [KVDS16] uses Stacked autoencoders (SAEs) to ex-

tract P300-related features. Researchers also try to remove artifacts in order to reduce the noise. For example, Gao [GZW10], Mennes [MWV<sup>+</sup>10], and Gwin [GG<sup>+</sup>10] propose signal processing methods to remove the artifacts caused by the muscle contraction, the eye movement, and the body movement, respectively. In terms of classification methods for P300-based BCIs, researchers have tried different classifiers, such as Support Vector Machine (SVM) [KMG<sup>+</sup>04, RG08], Linear Discriminant Analysis (LDA) [JAB<sup>+</sup>10], Fisher's Linear Discriminants (FLD) [SS09], Stepwise Linear Discriminant Analysis (SWLDA) [JK09], and neural networks (NN) [CG11, MG15, LWG<sup>+</sup>18, SLS18], in order to improve the accuracy of detecting P300 signals. The rapid development of machine learning algorithms for P300-based BCIs boosts the performance improvement for P300-based BCI systems.

### 1.1.2 Efficient P300-based Brain Computer Interface Systems

P300-based BCI systems have been in an experimental stage at research labs for a long time. Traditional P300-based BCI systems, as shown in Figure 1.2, use a complex EEG headset which utilizes a large number of sensors for brain signal acquisition as well as they use a cumbersome computer for signal processing and translation. Even though such BCI systems may achieve high enough performance in some cases, such complex systems for P300-based BCIs cannot be used in people's daily life. This is because it is impossible for people who wear such complex headset and need such cumbersome computer to move freely everywhere they want. In order to bring P300-based BCIs into practical use, in recent years, researchers have been trying to develop efficient P300-based BCI systems. As shown in Figure 1.3<sup>1</sup>, an efficient P300-based BCI uses a wireless EEG headset for signal acquisition. This headset utilizes a small number of sensors. In addition, such efficient BCI system uses a small mobile platform (e.g., mobile phone) for signal processing and translation. Since nowadays people use mobile phones almost every day and everywhere, it brings BCI users much convenience to use a mobile phone to process brain signals. Therefore, in recent years, a lot of research has been done for efficient P300-based BCI systems that use a wireless EEG headset for signal acquisition and a mobile phone for signal processing.

Concerning the wireless EEG headset, researchers have performed investigations to figure out the type of sensors as well as the number and the position of the sensors placed in the headset in order to build efficient P300-based BCI systems. Regarding the type of the sensors, traditional headsets, used in P300-based BCI systems, utilize wet sensors that operate with specially made conductive gels. The use of gels provides stable and high quality signal recording during a long-term use of a P300-based BCI system. However, the gels are sticky, which makes the BCI users' hair dirty and also

---

<sup>1</sup>This figure is taken from <https://www.emotiv.com/>.

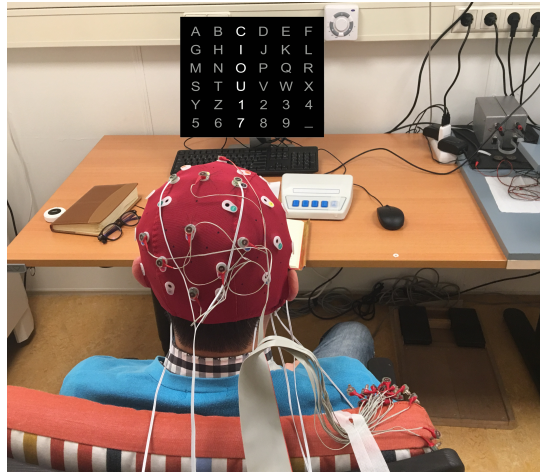


Figure 1.2: An example of a traditional P300-based BCI system.



Figure 1.3: An example of an efficient P300-based BCI system.

makes users not comfortable. In addition, the preparation time of placing wet sensors on the BCI users' scalp is quite long. To make P300-based BCIs more convenient and user-friendly, researchers have developed dry sensors [GVF11, SRC<sup>+</sup>12, CdBV<sup>+</sup>14] for the acquisition of EEG signals. By using dry sensors, users do not need to use the conductive gels any more. To make the BCI users feel more comfortable, researchers also have developed non-contact sensors [HCP02, SDC07, ONB<sup>+</sup>08] for EEG signal acquisition. Non-contact sensors are able to record EEG signals with a certain space between the brain skin and the headset. Unfortunately, dry sensors and non-contact

sensors may impair the performance of the P300-based BCI systems because compared with wet sensors, dry sensors cannot provide the same high quality of recorded EEG signals, and non-contact sensors provide even lower quality of recorded EEG signals because non-contact sensors output a very small signal amplitude and they are very sensitive to artifacts [IS16].

In addition to the development of the aforementioned types of sensors, researchers also focus on reducing the number of sensors used in a P300-based BCI system while keeping the performance of this system acceptable [RG08, RS<sup>+</sup>09, RCP<sup>+</sup>10, CRC<sup>+</sup>10, CR<sup>+</sup>11, RCS<sup>+</sup>11, RCMM12, CRT<sup>+</sup>14]. These studies propose sensor selection methods which select an appropriate sensor subset from an initial large set of sensors while keeping an acceptable BCI system performance. Such methods enable substantial reduction of the sensors needed to acquire EEG signals. The reduction of the number of sensors in a P300-based BCI system decreases the price of the EEG headset significantly, reduces the installation time of the P300-based BCI system, and also makes the users feel more comfortable. These advantages of the reduction of the sensors help promoting P300-based BCIs to be used in people's daily life.

After acquiring EEG signals from a wireless EEG headset, an efficient P300-based system uses a small mobile platform, such as a mobile phone, to process these signals. The mobile phone is an example of an embedded resource-constrained computing platform. Thus, the battery and memory of such platform are limited. As a result, the mobile phone cannot support the execution of signal processing algorithms with high complexity because such complex algorithms consume too much energy and memory where the amount of this consumption exceeds the limits of a mobile phone. Therefore, in order to build efficient P300-based BCI systems, signal processing algorithms with low complexity and acceptable performance are in urgent need. Such algorithms should be able to run on a mobile phone and consume a small amount of energy while keeping the system performance acceptable. In addition, in order to build energy-efficient P300-based BCI systems, techniques developed in the embedded system field can be used. Such techniques for energy-efficient task scheduling [LSWS16, CS16, NS17] and energy-efficient application mapping [LSCS15, SLS16] help the mobile phone to work energy-efficiently when used in a P300-based BCI system.

## 1.2 Problem Statement

The important development trends, described in Section 1.1, bring new opportunities to develop P300-based BCI systems. However, they also come with several issues when designing such systems. In this dissertation, we focus on several issues arisen by the aforementioned development trends in the contexts of the performance and the



efficiency of the P300-based BCI systems. The specific problems, we address in this dissertation, are formulated as follows.

### 1.2.1 Problem 1

As discussed in Section 1.1.1, the performance of the P300-based BCI systems is very important to bring these BCIs into people's daily life. Since the P300 speller is the benchmark and the most-commonly investigated application of the P300-based BCIs, we focus on how to improve the performance of the P300 speller. In order to improve the performance of the P300 speller, previous research on P300 spellers uses traditional machine learning methods for the detection of P300 signals and the inference of characters in the P300 speller. The traditional machine learning methods use manually-designed signal processing techniques for feature extraction as well as classifiers like Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA). Unfortunately, manually-designed feature extraction and traditional classification techniques have the following problems: 1) they can only learn the features that researchers are focusing on but lose or remove other underlying features; 2) brain signals have subject-to-subject variability, which makes it possible that methods performing well on certain subjects (with similar age or occupation) may not give a satisfactory performance on others. These problems limit the potential of manually-designed feature extraction and traditional classification techniques for further P300 detection accuracy, character spelling accuracy, and Information Transfer Rate (ITR)<sup>2</sup> improvements for the P300 speller.

Convolutional Neural Networks (CNNs) have the advantage of automatically extracting P300-related features from raw EEG signals. Thus, they can learn not only some features we know but also some features which are important and unknown to us. Automatically learning from raw EEG signals has better ability to achieve good results which are invariant to different subjects (persons). Thus, CNNs are able to boost the full potential of recognizing P300 signals, thereby overcoming the aforementioned shortcomings of traditional machine learning methods.

Therefore, in recent years, researchers have started to design (deep) CNNs for P300-based BCIs [CG11, MG15, LWG<sup>+</sup>18] and achieved better P300 detection, accuracy, character spelling accuracy, and ITR than traditional techniques. However, these CNNs have some limitations in increasing the P300 detection accuracy, the character spelling accuracy, and ITR for the P300 speller. These CNNs first use a spatial convolution layer to learn P300-related spatial features from raw signals. Then, they use several temporal convolution layers to learn P300-related temporal features from the abstract temporal signals generated by the spatial convolution layer (the first layer).

---

<sup>2</sup>For the detailed description of ITR please refer to Section 2.4.3.

In this way, the input to the temporal convolution layers is the abstract temporal signals instead of raw temporal signals. These abstract temporal signals in the feature maps lose raw temporal information. Losing raw temporal information means losing important temporal features because the nature of P300 signals is the positive voltage potential in raw temporal information, see Figure 2.9 explained in Section 2.4.1, as well as many important P300-related features are also embodied in raw temporal information [Pol07]. As a result, these CNNs cannot learn temporal features well. This leads to issues such as: 1) these CNNs prevent further P300 detection accuracy, character spelling accuracy, and ITR improvements for the P300 speller, thereby impairing the performance of the P300 speller; 2) these CNNs have high network complexity to achieve competitive P300 detection accuracy, character spelling accuracy, and ITR for the P300 speller, thereby impairing the efficiency of the P300 speller. Thus, the first problem addressed in this dissertation is:

**Problem 1: How can we design a CNN which achieves high P300 detection accuracy, character spelling accuracy, and ITR for the P300 speller and has low network complexity?**

### 1.2.2 Problem 2

P300 spellers have been in an experimental stage at research labs for a long time. As discussed in Section 1.1.2, P300 spellers are still not used in people's daily life because the efficiency of these P300-based BCI systems is low, even though these systems may achieve high enough performance in some cases. Some reasons for this low efficiency are: 1) Current popular EEG headsets in the BCI systems used for the P300 speller utilize a large number of sensors to achieve high spelling accuracy. The price of the EEG headset is significantly high when the number of sensors is large because a lot of sensors require a complicated electrode cap and a lot of amplifier channels. 2) Utilizing a large number of sensors makes the P300 speller to consume a lot of energy, which is unacceptable for a battery-powered mobile BCI system. Such system utilizes a wireless EEG headset and a resource-constrained hardware platform for data processing. A large number of sensors increases the amount of the data needed to be recorded and processed, thereby increasing the energy consumption of the wireless EEG headset and the hardware platform. This does not allow a mobile P300 speller to work for a long time period on a single battery charge; 3) Utilizing a large number of sensors strengthens the user's discomfort and increases the installation time of the P300 speller.

To address the aforementioned issues caused by the utilization of a large number of sensors, sensor selection methods could be used to select an appropriate sensor subset from an initial large set of sensors while keeping acceptable spelling accuracy. So, a good sensor selection method should enable substantial reduction of

the sensors needed to acquire brain signals. Therefore, good sensor selection methods are in urgent need for designing comfortable, cheap, and energy-efficient P300 spellers and for promoting such P300 spellers into the human’s daily life. Sensor selection methods for the P300 speller have been studied in recent years. For example, [RG08, RSG<sup>+</sup>09, CRC<sup>+</sup>10, CR<sup>+</sup>11] utilize a backward elimination algorithm as a sensor selection strategy. These works propose different ranking functions to evaluate and eliminate sensors such as the P300 signal detection accuracy, the P300 spelling accuracy [CR<sup>+</sup>11], the  $C_{cs}$  score [RG08], the Signal to Signal and Noise Ratio (SSNR) [RSG<sup>+</sup>09, CRC<sup>+</sup>10, CR<sup>+</sup>11], the Area Under the Receiver Operating Characteristic (AUC) [CRT<sup>+</sup>14]. Alternatively, [CG11] and [LWG<sup>+</sup>18] directly select the important sensors for a given user by analyzing the weights of a trained CNN. Unfortunately, the aforementioned sensor selection methods cannot select an appropriate sensor subset such that they can further reduce the number of sensors used to acquire EEG signals while keeping the spelling accuracy the same as the accuracy achieved when the initial large sensor set is used. As a consequence, the cost, energy consumption, and discomfort of a P300 speller are still unacceptably high when using the aforementioned sensor selection methods to design and configure P300 spellers. Therefore, the second problem addressed in this dissertation is:

**Problem 2: How can we design a sensor selection method which is able to further reduce the number of sensors needed to acquire EEG signals while keeping the character spelling accuracy the same as the accuracy achieved when the initial large sensor set is used?**

## 1.3 Research Contributions

In this section, we summarize the research contributions of this dissertation by addressing the research problems outlined in Section 1.2.

**Contribution 1: Proposing a CNN architecture which has low complexity and achieves high P300 detection accuracy, character spelling accuracy, and ITR for the P300 speller.**

To address Problem 1 in Section 1.2.1, we propose a simple, yet effective CNN architecture, called One Convolution Layer Neural Network (OCLNN), for the P300 speller. This CNN has only one convolution layer which is the first layer of the network. This layer performs both the spatial convolution and the temporal convolution at the same time, thereby learning very useful P300-related features from both raw temporal information and raw spatial information. Our OCLNN exhibits very low network complexity because it uses only one convolution layer and does not use fully-connected layers before the output layer. We perform experiments on three benchmark

datasets and compare our results with those in previous research works that report the best results. The comparison shows that our proposed CNN can increase the P300 signal detection accuracy with up to 14.23% and the character spelling accuracy with up to 35.49%. The comparison also shows that our proposed CNN achieves comparable ITR with the related BN3 method [LWG<sup>+</sup>18]. Moreover, our CNN achieves higher ITR compared to other state-of-the-art related methods [CG11, MG15, RG08, Bos04]. However, our OCLNN still has certain limitations to extract some important features related to P300 signals. OCLNN extracts P300-related spatial and temporal features at the same time in its single convolution layer, thereby extracting only P300-related joint spatial-temporal features through the spatial-temporal convolution. OCLNN does not extract P300-related separate temporal features and separate spatial features. These separate temporal features and separate spatial features have proven to be very important for the P300 speller [FTM<sup>+</sup>88, Pol07, PNCB11, HVE06].

**Contribution 2: Proposing an ensemble of different CNNs, we have devised, for the P300 speller.**

Our OCLNN proposed in **Contribution 1** has the limitation that it cannot extract separate temporal features and separate spatial features related to P300 signals. Adding some temporal or spatial convolution layers following the first spatial-temporal convolution layer of OCLNN is a potential method to enable OCLNN to learn P300-related separate spatial or separate temporal features. Nevertheless, such potential method cannot learn P300-related separate temporal or spatial features well due to the loss of raw information. The raw information loss happens because the input to these added temporal or spatial convolution layers for OCLNN is the abstract signals generated by the first spatial-temporal convolution layer instead of raw signals. To address properly the aforementioned limitation of OCLNN (proposed in **Contribution 1**), we propose an ensemble of two novel CNNs, we have devised, together with OCLNN in order to learn well the aforementioned P300-related separate spatial and separate temporal features, which are not extracted by OCLNN, together with the spatial-temporal features extracted by OCLNN. Our proposed ensemble of CNNs is called Ensemble of Convolutional Neural Networks (EoCNN). Our proposed two novel CNNs used in EoCNN are called One Spatial Layer Network (OSLN) and One Temporal Layer Network (OTLN), respectively. OSLN and OTLN has only one convolution layer. OTLN performs the temporal convolution in the first layer to learn P300-related separate temporal features. OSLN performs the spatial convolution in the first layer to learn P300-related separate spatial features. In this way, the input to OSLN and OTLN is raw signals, thus these two novel CNNs are able to learn features from raw signals. As a consequence, OTLN and OSLN can learn well P300-related separate temporal features and separate spatial features, respectively. Our EoCNN

uses the ensemble of OSLN and OTLN together with OCLNN, thereby extracting more useful P300-related features than OCLNN alone. As a result, our EoCNN can achieve higher P300 signal detection accuracy, character spelling accuracy, and ITR for P300 speller than OCLNN. Experimental results on three benchmark datasets show that our proposed EoCNN is able to increase the P300 signal detection accuracy, the character spelling accuracy, and the ITR achieved by OCLNN with up to 4.32%, 5%, and 6.05 bits/min, respectively. Also, our proposed EoCNN outperforms other related methods with a significant P300 signal detection accuracy improvement up to 18.55%, a significant character spelling accuracy improvement up to 38.72%, and a significant ITR improvement up to 21.75 bits/min. In terms of network complexity, the complexity of our EoCNN is lower than the complexity of the CNN in [MG15], and higher than the complexity of OCLNN and the CNNs in [CG11, LWG<sup>+</sup>18].

**Contribution 3: Proposing a CNN-based method for sensor reduction in the P300 speller.**

To address Problem 2 in Section 1.2.2, we propose a novel CNN-based sensor selection method, called Spatial Learning based Elimination Selection (SLES). Compared with the state-of-the-art sensor selection methods [RG08, RS<sup>+</sup>09, RCP<sup>+</sup>10, CRC<sup>+</sup>10, CR<sup>+</sup>11, RCS<sup>+</sup>11, RCMM12, CRT<sup>+</sup>14], our SLES is able to further reduce the number of sensors needed to acquire EEG signals in the P300 speller while keeping the character spelling accuracy the same as the accuracy achieved when an initial large set of sensors is used. Our SLES uses a novel parameterized CNN, we have devised, to evaluate and rank the sensors during the sensor selection process. This method features an iterative, parameterized, backward elimination algorithm to eliminate and select sensors. The parameter configured in this algorithm controls the training frequency of the CNN and the number of sensors to eliminate in every iteration. We perform experiments on three benchmark datasets and compare the minimal number of sensors selected by our SLES method and other selection methods needed to acquire brain signals while keeping the spelling accuracy the same as the accuracy achieved when the initial large set of sensors is used. The results show that, compared with the minimal number of sensors selected by other methods, our method can reduce this number with up to 44 sensors.

**Contribution 4: Proposing an improved ensemble of CNNs for the P300 speller with a small number of sensors.**

As a result of **Contribution 2**, our EoCNN is able to achieve higher spelling accuracy and ITR compared to other state-of-the-art methods for the P300 speller. As a result of **Contribution 3**, our SLES method can reduce the number of sensors needed

to acquire EEG signals in a EoCNN-based P300 speller while keeping the character spelling accuracy and the ITR the same as the character spelling accuracy and the ITR achieved by EoCNN when an initial large set of sensors is used in the P300 speller. We call the character spelling accuracy and the ITR, achieved by EoCNN for the P300 speller with a large number of sensors (e.g., 64 sensors), the state-of-the-art character spelling accuracy and ITR of the P300 speller. The experimental results mentioned in **Contribution 3** also show that in most cases, in order to preserve the state-of-the-art character spelling accuracy and ITR, we need to use more than 16 sensors to acquire EEG signals in the EoCNN-based P300 speller. Unfortunately, popular low-complexity and relatively cheap (affordable) BCI systems utilize a small number of sensors for the acquisition of EEG signals. Typically, such small number of sensors is less than or equal to 16 sensors. For example, BCI systems such as MUSE [MUS], EMOTIV Insight [Ins], Quick-8 [Qui], B-Alert X10 [B-A], EMOTIV EPOC+ [EMO], and OPEN BCI Mark IV [Mar] utilize only 4, 5, 8, 10, 14, and 16 sensors, respectively. Therefore, it is a challenge to achieve the state-of-the-art character spelling accuracy and ITR of the P300 speller with popular low-complexity and relatively cheap BCI systems that use a small number of sensors, i.e., less than or equal to 16 sensors, to acquire EEG signals.

To address the aforementioned challenge, we perform a study on our EoCNN (**Contribution 2**) as well as the three CNNs used in EoCNN, i.e., OTLN, OSLN, and OCLNN, for the P300 speller with different number of sensors in order to find the reason why EoCNN cannot achieve the state-of-the-art character spelling accuracy and ITR for a P300 speller with a small number of sensors. This study reveals that the reason for this is that EoCNN has the problem of putting equal importance on OSLN, OTLN, and OCLNN, when combining the outputs from these three CNNs for the P300 speller, irrespective of the number of sensors used to acquire EEG signals. In order to solve this problem of EoCNN, we propose an improved EoCNN for the P300 speller called PEOCNN. In PEOCNN, first, we parameterize the process of combining the outputs from OSLN, OTLN, and OCLNN. Then, we use the Sequential Model-based Algorithm Configuration (SMAC) [HHLB11] to automatically find and set values for the parameters depending on the number of sensors utilized in the P300 speller. In this way, PEOCNN is able to adapt/configure the importance of using the outputs from OSLN, OTLN, and OCLNN for the P300 speller depending on the number of sensors that are utilized. Experiments on three benchmark datasets show that when using our PEOCNN for the P300 speller, the state-of-the-art character spelling accuracy and ITR can be achieved in a BCI system with less than or equal to 16 sensors to acquire EEG signals.

## 1.4 Dissertation Outline

In this section, we give an outline of this dissertation:

**Chapter 2** introduces some background information on Convolutional Neural Networks (CNNs), the P300 signal, the P300 speller, the Information Transfer Rate (ITR), and the datasets used in this dissertation.

**Chapter 3 - 6** describe in details the contributions introduced in Section 1.3. Each chapter is organized in a self-contained way. That is, each chapter has its specific introduction, related work, proposed method, experimental evaluation, and conclusions.

**Chapter 3** presents our proposed simple, yet effective, CNN architecture for the P300 signal detection and P300-based character spelling. This chapter is based on the following publication:

- **Hongchang Shan**, Yu Liu, and Todor Stefanov,  
 "A Simple Convolutional Neural Network for Accurate P300 Detection and Character Spelling in Brain Computer Interface",  
*In Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI'18)*, pp. 1604-1610, Stockholm, Sweeden, July 13-19, 2018.

**Chapter 4** presents our proposed ensemble of CNNs for the P300 signal detection and P300-based character spelling. This chapter is based on the following publication:

- **Hongchang Shan**, Yu Liu, and Todor Stefanov,  
 "Ensemble of Convolutional Neural Networks for P300 Speller in Brain Computer Interface",  
*In Proceedings of the 28th International Conference on Artificial Neural Networks (ICANN'19)*, pp. 376-394, Munich, Germany, September 17-19, 2019.

**Chapter 5** presents our proposed sensor reduction method for the P300 speller. This chapter is based on the following publications:

- **Hongchang Shan**, and Todor Stefanov,  
 "SLES: A Novel CNN-based Method for Sensor Reduction in P300 Speller,"  
*In Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'19)*, pp. 3026-3031, Berlin, Germany, July 23-27, 2019.
- **Hongchang Shan**, and Todor Stefanov,  
 "A Novel Sensor Selection Method based on Convolutional Neural Network for P300 Speller in Brain Computer Interface",  
*The 56th ACM/IEEE Design Automation Conference (DAC'19) WIP session*, Las Vegas, NV, USA, June 2-6, 2019.

**Chapter 6** presents our proposed improved ensemble of CNNs for a P300 Speller with a small number of sensors. This chapter is based on the following publication:

- **Hongchang Shan**, Yu Liu, and Todor Stefanov,  
"An Empirical Study on Sensor-aware Design of Convolutional Neural Networks for P300 Speller in Brain Computer Interface,"  
In *Proceedings of "12th IEEE International Conference on Human System Interaction (IEEE HSI'19)"*, pp. 5-11, Richmond, Virginia, USA, June 25-27, 2019

**Chapter 7** ends this dissertation by providing summary and conclusions regarding the work presented in this dissertation.