

## MIT Open Access Articles

*A system approach for closed-loop assessment of neuro-visual function based on convolutional neural network analysis of EEG signals*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Stock, Simon C., Armengol-Urpi, Alexandre, Kovacs, Balint, Maier, Heiko, Gerdes, Marius et al. 2020. "A system approach for closed-loop assessment of neuro-visual function based on convolutional neural network analysis of EEG signals." Proceedings of SPIE - The International Society for Optical Engineering, 11360.

**As Published:** 10.1117/12.2554417

**Publisher:** SPIE

**Persistent URL:** <https://hdl.handle.net/1721.1/136990>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of Use:** Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



# PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://SPIDigitalLibrary.org/conference-proceedings-of-spie)

## A system approach for closed-loop assessment of neuro-visual function based on convolutional neural network analysis of EEG signals

Stock, Simon, Armengol-Urpi, Alexandre, Kovács, Bálint, Maier, Heiko, Gerdes, Marius, et al.

Simon C. Stock, Alexandre Armengol-Urpi, Bálint Kovács, Heiko Maier, Marius Gerdes, Wilhelm Stork, Sanjay E. Sarma, "A system approach for closed-loop assessment of neuro-visual function based on convolutional neural network analysis of EEG signals," Proc. SPIE 11360, Neurophotonics, 1136008 (1 April 2020); doi: 10.1117/12.2554417

**SPIE.**

Event: SPIE Photonics Europe, 2020, Online Only

# A system approach for closed-loop assessment of neuro-visual function based on convolutional neural network analysis of EEG signals

Simon C. Stock<sup>a</sup>, Alexandre Armengol-Urpi<sup>b</sup>, Bálint Kovács<sup>a</sup>, Heiko Maier<sup>a</sup>, Marius Gerdes<sup>a</sup>, Wilhelm Stork<sup>a</sup>, and Sanjay E. Sarma<sup>b</sup>

<sup>a</sup>Karlsruhe Institute of Technology, Germany

<sup>b</sup>Massachusetts Institute of Technology, USA

## ABSTRACT

We propose a generalized, modular, closed-loop system for objective assessment of human visual parameters. Our system presents periodical visual stimuli to the patient's field of view and analyses the consequent evoked brain potentials elicited in the occipital lobe and recorded through EEG. The analysis of the monitored EEG data is performed in an end-to-end fashion by a convolutional neural network (CNN). We propose a novel CNN architecture for EEG signal analysis that can be trained utilizing the benefits of multi-task learning. The closed-loop attribute of our system allows for a real-time adaptation of the subsequent stimuli to further examine a potentially damaged area or increase the granularity of the exploration. Interchangeability is provided in terms of software modules, stimulus type, visual hardware, EEG acquisition device and EEG electrodes. Initially, the system is designed to monitor visual field loss originating from glaucoma or damage to the optic nerve using a virtual reality (VR) headset for the stimuli presentation. The modular architecture of our system paves the way for the assessment and monitoring of other neuro-visual functions.

**Keywords:** Neuroscience, Brain-Computer-Interface, EEG, Ophthalmology, Deep Learning, Convolutional Neural Networks, System Design

## 1. INTRODUCTION

Visual perception is a particularly interesting research area in neuroscience.<sup>1-3</sup> The correct functioning of vision is crucial for many everyday tasks. We use it to recognize our surroundings and to identify and interact with our environment. However, several conditions can limit vision. Examples for problems with the optical apparatus are myopia, hyperopia, and presbyopia. These can usually be overcome with the help of a personalized vision aid. It is often more severe if the cause of vision impairment is neurological in nature. These cannot be easily corrected and therefore have a high impact on the quality of life.<sup>4-6</sup> In addition to brain tumors in the region of the optic nerve or occipital lobe, common diseases are glaucoma, age-related macular degeneration (AMD), and diabetic retinopathy. The most significant risk to get one of these neurodegenerative diseases is old age. Due to demographic change, the number of neurological eye diseases is, therefore, continually increasing.<sup>7</sup> The restoration of the retina or parts of it is an active field of research. To date, however, there is no successful method for recovering from neurological vision loss.<sup>8</sup> Therefore, the ophthalmological assessment of visual parameters is of utmost importance. Current therapies, therefore, focus on stopping the progression of vision decline. Visual field loss is often slow and gradual. Those affected only recognize it late, which is why it is advisable to examine the neuro-visual field regularly.<sup>9</sup>

Several procedures in Ophthalmology, such as the assessment of visual acuity and visual field, have in common that they tend to take much time. Moreover, they are subjective and require the patient to understand the test and be cooperative.

In this paper, we want to present a system approach for utilizing an EEG-based brain-computer-interface (BCI) to objectively measure neurological responses of the visual cortex to retinal stimulation. The recorded signals are analyzed in real-time using a convolutional neural network (CNN). Our purpose is to outline a combination of state of the art EEGs, displays, and deep learning methods into a closed-loop system. An

exemplary setup of this system is depicted in Figure 1. Furthermore, we want to present a use-case study of the proposed system in the form of visual field testing (perimetry).

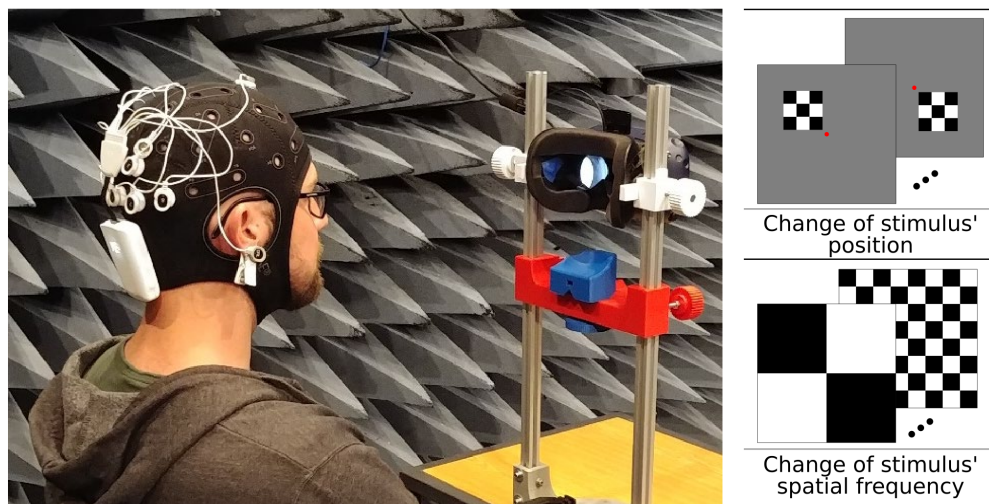


Figure 1: An exemplary measurement setup in the lab. EEG signals are recording using a set of electrodes positioned over the visual cortex while stimuli are displayed on a VR headset. Currently, the system supports checkerboard stimuli for the generation of SSVEPs. Parameters like the temporal and spatial stimulation frequency of the checkerboard and its position and size can be configured.

## 2. RELATED WORK

The subjective nature of most of the current gold standard ophthalmological tests has brought many researchers to explore other methods that involve objective measurements only. The approach of assessing visual defects through the monitoring of brain signals has indeed been one of them. The most common method to implement this concept is by stimulating the visual field of the subject and measuring brain responses elicited by the stimuli, the so-called visual evoked potentials (VEP).<sup>10-12</sup> Exploiting this characteristic of the human neuro-visual pathway, researchers have been able to explore the objective assessment of visual field defects<sup>13,14</sup> and specifically of glaucoma.<sup>15-18</sup> Not only have they been used to evaluate visual defects, but also to study the topography of the VEP properties across the visual field<sup>19</sup> and to investigate the VEP and visual acuity maturation of children.<sup>20,21</sup> When the visual stimuli is presented to the subject's retina in a periodic manner, the evoked brain potentials reach a steady-state where they oscillate at the same rate as the visual stimuli, and they receive the name of steady-state visual evoked potential (SSVEP).<sup>22-24</sup> SSVEPs show high SNR and therefore are widely used to build BCIs<sup>25</sup> and they have also been exploited to evaluate the effective field of view in glaucoma patients<sup>26</sup> and even to monitor the visual pathway function of patients undergoing neurosurgery.<sup>27</sup> A comparison study of the different visual evoked potential-based methods to assess visual acuity has also been conducted.<sup>28</sup> As far as the visual stimulation hardware is concerned, there has been a recent interest in using virtual reality (VR) headsets for the implementation of BCIs.<sup>29,30</sup> As opposed to classic 2D displays, the controlled visual field angles, closed environment and portability of VR systems have also attracted researchers to use this technology for the assessment of glaucomatous visual field defects.<sup>31</sup>

The recent eruption of deep learning techniques to process and extract features from time-series data has also impacted the BCI community. Deep learning has allowed to increase the speed of BCIs<sup>32</sup> and convolutional neural networks (CNN) in particular have also been explored in this field.<sup>33,34</sup> More specifically, CNNs have indeed been used to enhance the detection and classification of SSVEPs,<sup>35-38</sup> to characterize the nature of SSVEPs<sup>39</sup> and even to extract EEG-based biometric information through SSVEP.<sup>40</sup>

### 3. CONCEPT

The aim of the proposed modular system is to carry out SSVEP-based BCI experiments in a closed-loop fashion. Therefore, the system needs to support four main features:

- An easy way to specify stimuli or series of stimuli to be rendered.
- Software components to render the specified stimuli.
- Means to receive recorded data in real-time from an EEG.
- A way to perform real-time inference of machine learning models.

We chose Qt to develop a graphical user interface (GUI). The GUI allows for an easy specification of the desired stimulation parameters. Additionally, CSV files were chosen as a means to specify a series of stimuli in an easily configurable format. For the display of different stimuli, the need for a graphics engine arises. We chose the Unity game development engine as it supports cross-platform development. Furthermore, it enables the use of VR, conventional monitors and projector-based systems, all integrated into the same software framework. As a means of receiving data from the EEG recording device, the Lab Streaming Layer (LSL) protocol was chosen. EEG equipment generally supports this kind of data streaming either by vendor software or by software included in the LSL distribution. Even if a specific EEG equipment would not support LSL, it can be implemented in an intermediate component translating the data. After the consideration of other systems,<sup>41</sup> we tested our system with a Neuroelectronics Enobio 8 channel EEG<sup>42</sup> with reusable and disposable gel electrodes, and passive dry electrodes. Exploiting the modular architecture of our system, we have tested our concept on high refresh rate consumer LCD monitors, laser projectors, and OLED based VR headsets. Further development would allow for the use of mobile devices as well. However, depending on the use-case and the visual assessment test desired, some displaying hardware will be more appropriate than others.

To reduce motion artifacts, we designed an adjustable chinrest that can be 3D printed. The source files can be found here: <https://github.com/MetisVidere/ChinRest>. The chinrest can be used in combination with a VR headset, which was the visual stimuli display chosen for our proof of concept. The chinrest may also be used to mount a simple phoropter, enabling the use of different lenses during the experiment.

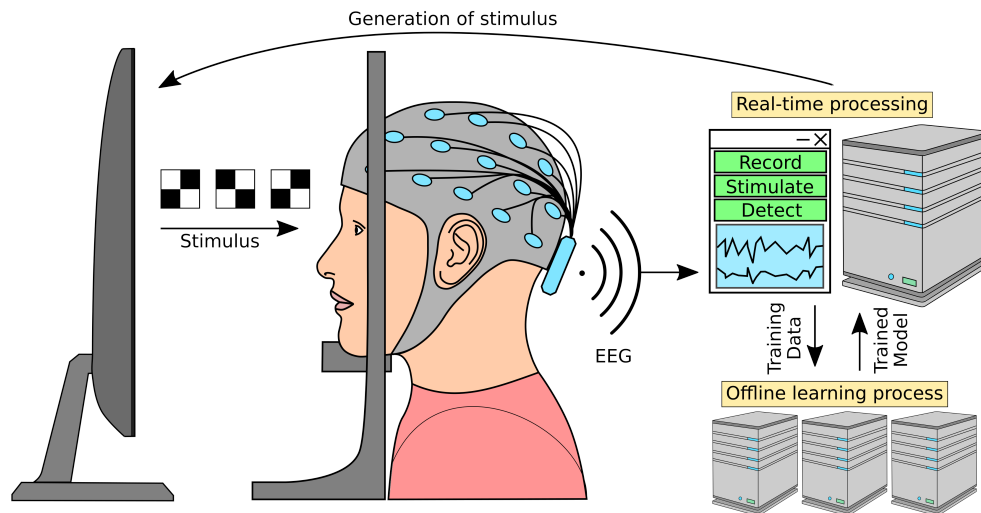


Figure 2: System overview. The subject is placed in front of a display device that is presenting visual stimuli. EEG signals are recorded and transmitted to the local processing unit, capable of running a trained model in real-time. The data is also stored, so it can be used for training new models offline.

## 4. IMPLEMENTATION

The software can be used for two main functionalities. The first is for experimentation and data collection, whereas the second is for closed-loop applications. For the experimentation workflow, the experimenter enters parameters specifying the desired visual stimulus, either by selecting them on the GUI or by defining them in the CSV protocol file. Those parameters are read in by the Backend. They are then used to generate corresponding visual stimuli through a simultaneously running Graphics Engine (in our case: Unity). Meanwhile, the brain signals of the subject are recorded by an EEG device and streamed to the computer. On the computer, this data stream is collected by third-party software from the EEG supplier (in our case: Enobio's NIC2). Whenever a new stimulus is presented, a marker in the form of an integer number is created and sent to the external EEG software via the Lab Streaming Layer (LSL). Here, the marker is synchronized with the EEG data, allowing to reconstruct which periods of the EEG data correspond to a specific visual stimulus. Finally, the EEG data and the synchronized markers are stored on the hard drive for subsequent analysis and/or training of deep learning models. At the same time, the EEG data, along with the markers, are streamed to the Backend using LSL streams. There, they can be evaluated in real-time by a neural network model using an abstract interface and a compiled version of the deep learning framework (in our case: Google TensorFlow). The results of this analysis are then displayed on the GUI and can be used by the Backend to generate further stimuli in a closed-loop fashion. An overview of the system is provided in Figure 2. A more detailed depiction, showing the individual components of the software implementation is given in Figure 3.

### 4.1 Backend

The Backend implements the task of interconnecting the GUI with the Graphics Engine and the external EEG software. Parameters for the next visual stimulus, coming from the GUI or from CSV protocol files, are parsed and passed to the Graphics Engine. Simultaneously, feedback referencing the displayed stimuli is received from the Graphics Engine. This feedback, in turn, is used to generate time synchronized markers, which are sent to the external EEG software through the Lab Streaming Layer. Depending on the application, the Backend also initiates the real-time processing of the received EEG data through the Deep Learning Interface. Subsequently, closed-loop generation of new stimuli based on the results of this processing can be executed.

There are three components in charge of interconnecting the modules: a Controller, connecting GUI and Backend; a segment of shared memory, to communicate between Backend and Graphics Engine; and an interface towards the Deep Learning framework. They are designed to enable communication between the major components while keeping those components exchangeable.

To avoid data loss or a lack of responsiveness, all described tasks need to be fulfilled simultaneously. Thus, the Backend was designed as a multithreaded application. It comprises four threads of execution. Each of the threads fulfills one unique role, and they are synchronized by a set of mutexes and queues. To combine inputs from both the GUI and from a CSV protocol, a queue is used as an input interface. Pushing of stimulation parameters into this queue is provided over a public function call. This call guarantees threadsafe access by using a mutex as an access guard for the queue. When calling the function, the mutex is locked, and the command is pushed into the queue. Then, an internal semaphore is used to signal the availability of new commands to the threads. The four threads then perform their specific tasks as detailed in the following paragraphs:

#### *Main Management Thread*

This thread is responsible for processing all incoming commands from GUI and protocol files and sending them to the Graphics Engine over the Shared Memory module. It waits for the semaphore to indicate the arrival of new commands. Then, it acquires a lock on the queue mutex to accomplish threadsafe access to the queue. Subsequently, it reads all commands from this queue, forwards them into a thread-internal queue, and releases the lock on the mutex. Finally, the new commands from the internal queue are sequentially parsed and their stimulus parameters are sent to the Graphics Engine.

#### *Protocol Reader Thread*

This thread is responsible for reading and parsing commands from a CSV protocol file. Each line is parsed into an encapsulating structure by internal parsing functions. Then, this structure is pushed into the queue.



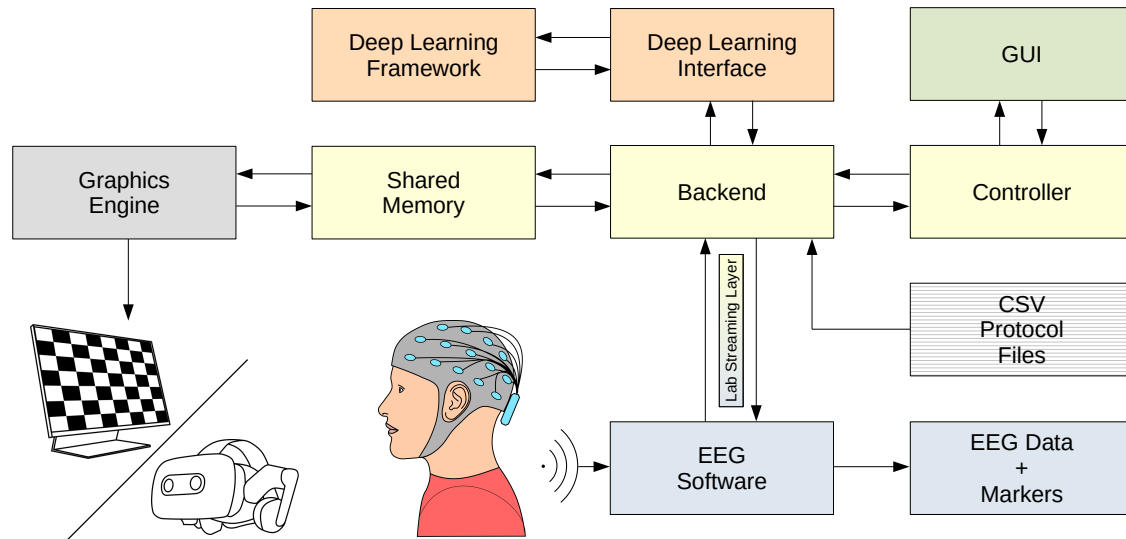


Figure 3: The software system includes several components, of which the Backend forms the core. Stimulation parameters are given over a GUI or parsed from csv protocol files. Real-time inference of neural networks is performed using an abstract interface layer. Communication with the EEG is performed through the Lab Streaming Layer to acquire data and set markers. Stimuli generation is done with a Graphics Engine addressed over a segment of shared memory.

#### *Feedback From Graphics Engine Thread*

This thread reads feedback placed in the shared memory from the Graphics Engine and generates markers from this feedback, which are then sent to the external EEG application via LSL. Internally, the thread uses a blocking call to go into a waiting state until new feedback is available on the Shared Memory Segment (as signaled by an increment in the respective semaphore). In this case, the call returns and the thread is reawakened. A reference label contained in the feedback is extracted and pushed as a marker value into the LSL stream, transmitting it to the external EEG application. After that, the next blocking call is performed and the thread again enters a waiting state until new feedback is available.

#### *Closed Loop Thread*

This thread continuously pulls samples from the incoming LSL Stream, thereby receiving EEG Data and synchronized labels from the external EEG application. The data is collected and rearranged into an input structure for the currently used neural network. Then, network inference is performed and the results are sent to the GUI over the Controller. Additionally, the thread can generate commands for new stimuli based on the inferred results and push them into the queue.

## 4.2 Graphical User Interface

This module allows the user to control the system. All parameters of the stimuli can be adjusted manually or executed from a selected protocol file. The currently used GUI was designed using Qt. To keep GUI and Backend separable, and exchangeable, we used the Model-View-Controller pattern.<sup>43</sup> The Controller acts as an intermediate component between Backend and GUI. Its purpose is to deliver information sent from the GUI through a signal-based mechanism to the Backend by public method calls and vice versa. This enables the exchangeability of the GUI, for example, if another GUI should be used on phone screens in the future.

## 4.3 Graphics Engine

This component is designed to render the visual stimuli on the current stimulation device, be it a monitor or a VR headset. It is implemented as an external application, written in C#, making use of the Unity game engine. Unity Applications are partitioned into scenes, which form logical sub-units of the application. The current system is supporting two scenes: the *Positional SSVEP Scene* and the *Sweep VEP Scene*. All stimuli

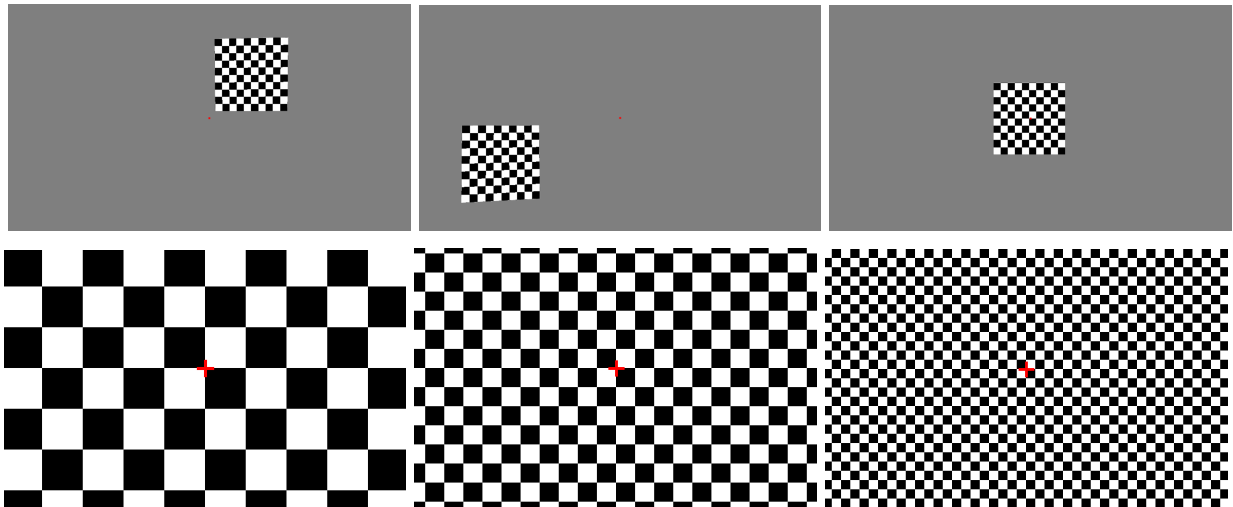


Figure 4: Top row: Screenshots from the *Positional SSVEP Scene*, displaying a stimulus of size  $(10^\circ, 10^\circ)$  at positions  $(6^\circ, 6^\circ)$ ,  $(-16^\circ, -6^\circ)$  and  $(0^\circ, 0^\circ)$  from left to right in a VR screen. Bottom row: Screenshots from the *Sweep VEP Scene*, showing an exemplary progression of spatial frequencies.

were designed as Unity shaders. In addition to these scenes, the class *Parameter Manager*, was designed for communication with the Backend via a Shared Memory handle. With this handle, the Graphics Engine is able to communicate with the Backend almost as if they were a single application.

#### *Positional SSVEP Scene*

Functionally, the scene offers the possibility to display an SSVEP stimulus, whose position (in terms of visual angle relative to the subject), size (in terms of visual angle), flickering frequency and number of squares in the checkerboard can be specified by the experimenter. The stimulus can be configured as a reversal or as a pattern-onset stimulus. To support experiments like perimetry-testing, an optional fixation target, in the form of a red cross or a red dot, can be displayed in the center of the visual field. Additionally, the experimenter can stimulate both eyes or select one eye of interest specifically.

#### *Sweep VEP Scene*

This scene was designed to display sweep VEP stimuli (a sequential variation of the spatial frequency of an SSVEP). It is intended to be used in experiments regarding the objective estimation of visual acuity. The scene displays a screen-wide checkerboard stimulus with adjustable frequency, number of checkers and the possibility to choose between reversal or pattern-onset stimulation. Again, fixation targets in the form of a red dot or cross can be displayed if desired. Exemplary screenshots of both scenes can be seen in Figure 4.

#### *Parameter Manager*

The *Parameter Manager* is a class designed to fetch new stimulus parameters from the Shared Memory module and generate feedback in response to those parameters. The feedback is inherently important as for correct labels; the Backend needs to know what is currently presented through the graphics engine. Additionally, the *Parameter Manager* offers an interface through which objects from all scenes have access to the current stimulation parameters without having to access the Shared Memory segment themselves. The class is designed following a singleton design pattern.<sup>43</sup> Additionally, to adjusting stimuli parameters and providing feedback to Backend, the *Parameter Manager* is starting and stopping individual scenes on demand from the Backend.



#### 4.4 Shared Memory Communication Handle

For the communication between the Graphics Engine (C#) and the Backend (C++), a common interface was designed. The interface is written in C++ and distributed in the form of a dynamic-link library (DLL), whose functions can be imported and called from both C# and C++ code. The widely used C++ Boost library was used for the implementation of this interprocess *Shared Memory Communication Handle* and the respective interprocess semaphores. Within the shared memory segment, two structures reside: one containing information about the currently desired stimulation parameters, and the other containing feedback information from Unity. Feedback hereby includes a label designated to the stimulation parameters chosen, indicating to which stimulus the feedback responds, and a timestamp, for synchronization, indicating when this stimulus was rendered.

### 5. EEG DATA PROCESSING

#### 5.1 State of the Art: The Canonical Correlation Analysis (CCA)

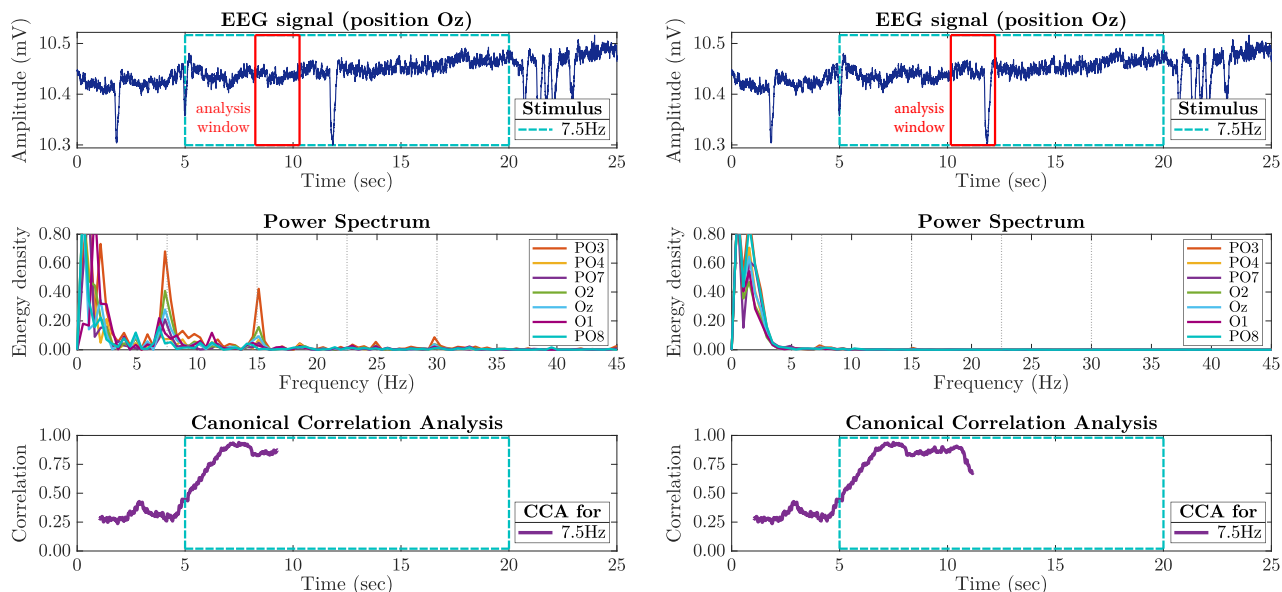
For the classification of SSVEPs, the stimulus frequency  $f_{st}$  and its harmonics need to be identified from the EEG signals recorded from the visual cortex. A robust approach for SSVEP detection is based on canonical correlation analysis (CCA). Initially proposed in 1936,<sup>44</sup> it was used the first time for SSVEP analysis in 2006.<sup>45</sup> The principal idea is to use the EEG signal channels  $X \in \mathbb{R}^{C \times N}$  as a first set of variables. Furthermore, one creates a reference signal  $Y \in \mathbb{R}^{2N_h \times N}$  containing sine and cosine waves at the stimulus frequency and its harmonics:

$$Y = \begin{bmatrix} \sin(2\pi f_{st}) \\ \cos(2\pi f_{st}) \\ \sin(2 \cdot 2\pi f_{st}) \\ \cos(2 \cdot 2\pi f_{st}) \\ \vdots \\ \sin(N_h \cdot 2\pi f_{st}) \\ \cos(N_h \cdot 2\pi f_{st}) \end{bmatrix}, \quad n = \frac{1}{T_s}, \frac{2}{T_s}, \dots, \frac{N}{T_s}, \quad (1)$$

where  $C$  is the number of EEG channels,  $N$  is the number of samples,  $N_h$  is the number of harmonics and  $T_s$  is the EEG sampling frequency. In this way, CCA generates a vector of canonical correlations with vector length of  $\min\{C, 2N_h\}$ :  $\rho = (\rho_1, \rho_2, \dots, \rho_{\min\{C, 2N_h\}})$ . The Euclidean norm of this correlation vector

$$\|\rho\| = \sqrt{\sum_{i=1}^{\min\{C, 2N_h\}} \rho_i^2} \quad (2)$$

is suitable to use as a feature for the classification of SSVEP signals.<sup>46</sup> A comparison between the CCA and the Fourier spectrum is illustrated in figure 5 for an EEG signal with 7.5 Hz stimulus. From the results in figure 5a, we can conclude that the CCA is a powerful and appropriate technique to analyze and classify the presence of known SSVEP features, which is, however, influenced by signal artifacts like blinking.



(a) SSVEP peaks are visible at 7.5 Hz and its harmonics in the Fourier power spectrum. For a good signal in the analysis window, the CCA value is respectively high. (b) The power of SSVEP peaks is negligible compared to other signal components in the Fourier power spectrum. The CCA value is considerably decreased due to the blinking artifact.

Figure 5: The figures show the Fourier power spectrum and canonical correlation value of the EEG signal at the time of the red analysis window with a length of 2 s. During the blue window a 7.5 Hz stimulus was presented.

## 5.2 Deep Learning Approach

Over recent years, the interest in applying deep neural networks to process EEG signals steadily increased.<sup>36,47</sup> This is due to their ability to handle signal abnormalities and artifacts as they use multiple levels of abstraction in comparison to traditional techniques. Moreover, it is suggested that deep neural networks are able to do correct classifications even in noisy conditions.<sup>48</sup>

Deep convolutional neural networks (CNNs) have been reliably used in many applications, especially when involving images. However, EEG signals have quite different characteristics in comparison to an image. In contrast to a static 2D image with large spatial dimensions, EEG signals have a large temporal dimension and they are obtained from the 3D scalp surface with rather small spatial resolution. However, features in a multi-channel EEG signal have a significant dependence on adjacent values, which is analogous to images or videos. Therefore, CNNs can also be applied for the classification of EEG signals. As non-task-relevant sources affect the EEG signal significantly, they usually have a low signal-to-noise ratio. Thus, we are exploring the use of convolutional neural network deep learning techniques for EEG signal analysis.

Since EEG signals are highly stochastic in nature,<sup>49</sup> the convergence of the learning process is difficult due to many local minima. Therefore, we increased the batch size during training to avoid local minima at the beginning and to smooth the feature surface in the feature space for regularization.<sup>50</sup> Spatial filtering over EEG channels was employed to eliminate stochastic noise. Furthermore, it enables the weighting of electrodes and, therefore, can enhance channels containing more information than others.<sup>51</sup> Due to these reasons, we proposed the input of the neural network to resemble the spatial arrangement of the electrode positions on the scalp. A schematic of this is presented in figure 6.

When working with deep learning algorithms, the size of the dataset has a huge influence on the training results. To artificially increase the size of our dataset, we augmented it by cropping the signal records into overlapping time slices. This generates a larger amount of input samples. It also forces the convolutional neural network to learn the SSVEP frequency components, present in all crops independently from artifacts or signal abnormalities it could otherwise use to overfit the training data.<sup>52,53</sup>

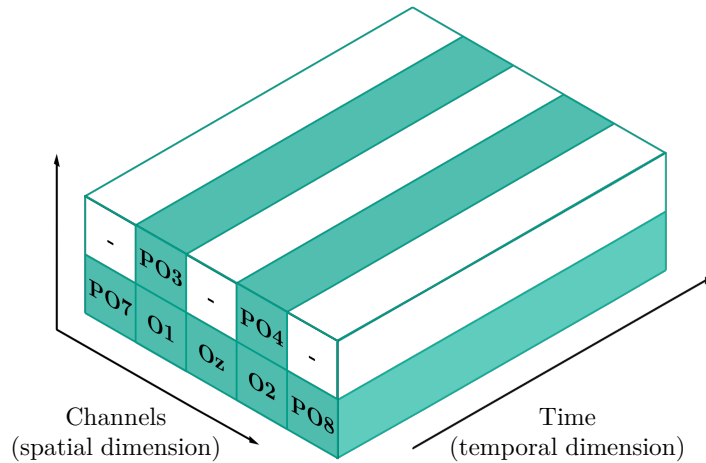


Figure 6: The channel dimensions of the EEG are projected to a 2D plane resembling an image. The temporal dimension of the signal is presented as a 3rd dimension. Using 3D convolution filters, one is able to extract common signal morphology for further processing.

In contrast to images, EEG signals are challenging to find online in large, annotated databases. Therefore we had to use the available data efficiently. Especially for small training sets, neural network approaches tend to overfit. This can be prevented using different regularizers. For this purpose, we used L2, dropout, and batch normalization. Additionally, we used Multi-Task Learning.<sup>54</sup> This is a form of regularization by which the network learns to work abstractly in the encoder part to solve several related tasks using joint training. In our case, we classify whether the signal is pathologically relevant and simultaneously perform a CCA regression (compare figure 7). We can describe the combined loss function as

$$\mathcal{L}_{\text{Multi}} = \alpha_1 \mathcal{L}_{\text{Cla}} + \alpha_2 \mathcal{L}_{\text{CCA}}, \quad (3)$$

with  $\mathcal{L}_{\text{Cla}}$  being the classification and  $\mathcal{L}_{\text{CCA}}$  being the CCA regression loss. The  $\alpha_n$  multiplier describes the weighting of the individual losses towards the combined Multi-Task loss  $\mathcal{L}_{\text{Multi}}$ .

To maximize the performance of our neural network structures, we used the Bayesian optimization process.<sup>55</sup> In each optimization cycle, a neuronal network with new hyperparameters is initialized and trained using the training set. The hyperparameters of the consecutive run are selected based on the performance on the validation set. The general structure of the hyperparameter optimization workflow is presented in Figure 8.

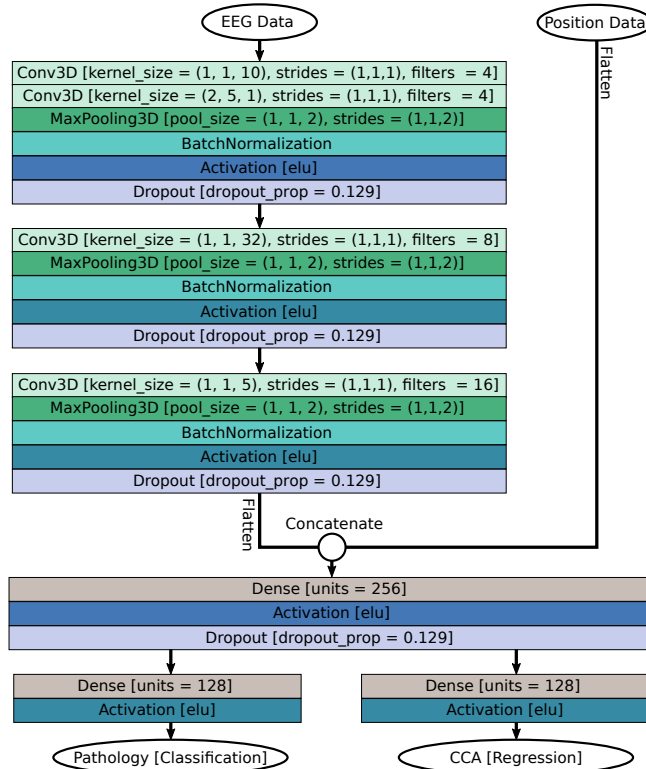


Figure 7: Structure of the CNN used in the multi-task learning process. EEG Data is processed through a series of convolutional layers, followed by MaxPooling, BatchNorm, Activation and Dropout layers. The resulting features are concatenated with the input indicating the stimuli position. After a series of dense layers, both a classification (healthy/pathological) and a regression of the CCA are computed.

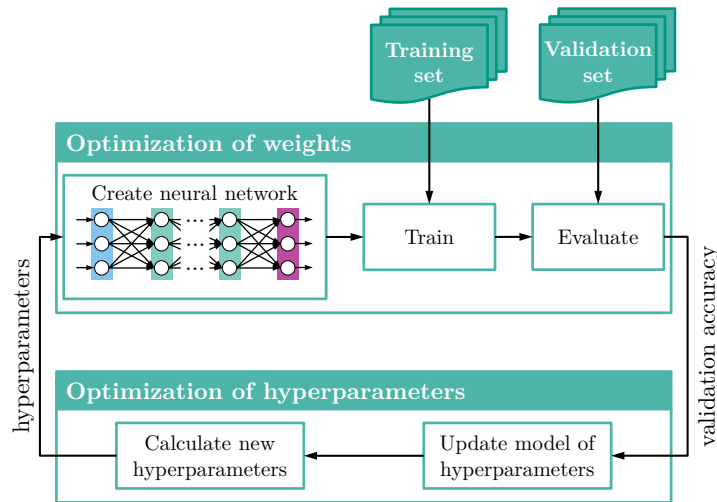


Figure 8: The general flowchart for hyperparameter optimization. In each optimization cycle, new hyperparameters are calculated by the optimizer. A neural network with the iterated hyperparameters is built and trained on the training set and evaluated on the validation set. Based on the performance on the validation set, the hyperparameters are updated for the next training cycle.

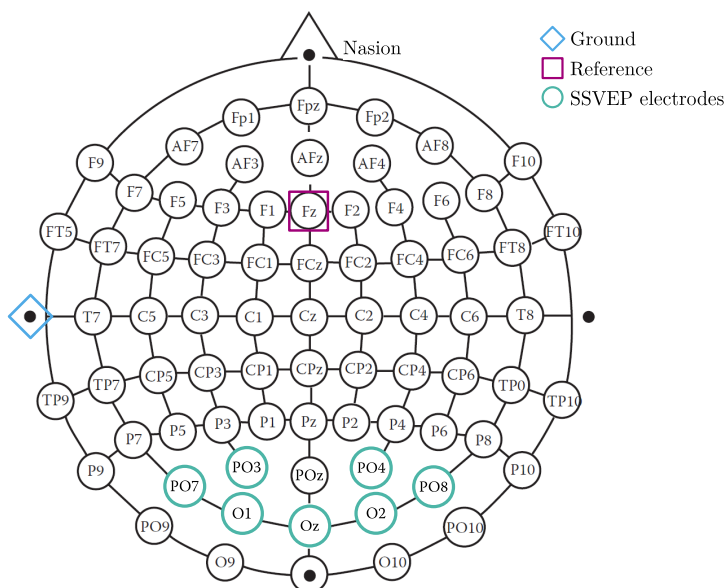


Figure 9: The figure shows the chosen electrode placement on the international 10/10 system. We used seven electrodes for the detection of signals over the visual cortex. In addition there are the ground and reference electrodes. We placed the ground electrode on the ear, to reduce biosignal activity. The optional reference electrode is usually placed onto the location of Fz to remove irrelevant frontal brain activities from the signal.

## 6. EXPERIMENT SETUP

In this chapter, we want to present a use case for the proposed system. We collected EEG data from 15 individuals using our proposed system. The data was collected at the Massachusetts Institute of Technology, USA, and at the Karlsruhe Institute of Technology, Germany.

### 6.1 EEG Setup

We considered dry electrodes during trials, but although the system setup time was drastically improved, we struggled with a bad signal to noise ratio. Thus, for the recordings described here, we are using standard wet electrodes, which provided a better signal to noise ratio. When positioning the electrodes on the head, it is important to consider the later application. We aim to quantify the functionality of vision. Therefore, we placed the electrodes over the occipital lobe. Figure 9 shows the electrode setup used in our experiments. We chose  $F_z$  as the reference electrode in compliance with the current ISCEV standard for clinical visual evoked potentials.<sup>56</sup> Furthermore, we utilized the chinrest presented in Chapter 3 to position the patient in a comfortable way in order to minimize artifacts.

As 0.3–3% of the human population carry the risk of photosensitive epilepsy,<sup>57</sup> there are official guidelines for its prevention.<sup>58</sup> As stimulation frequency, we chose 7.5 Hz as it lies outside the highly epilepsy provocative range of frequencies between 15–25 Hz.

### 6.2 Study Protocol

Initially, we asked the subjects if they had neurological conditions like Epilepsy, as well as whether they have had any eye or visual pathway infections in the past. We only proceeded to record when there was no risk involved. Furthermore, the participants were told that they could quit at any time without stating any reasons. One measurement routine has a duration of about 2 minutes. We performed five (one without mask and four with the masks presented in section 6.4) of those measurements for the monocular case and the binocular case for each subject respectively. The electrode placement was as described in Figure 9. Furthermore, we used a chin rest like the one shown in Figure 1.

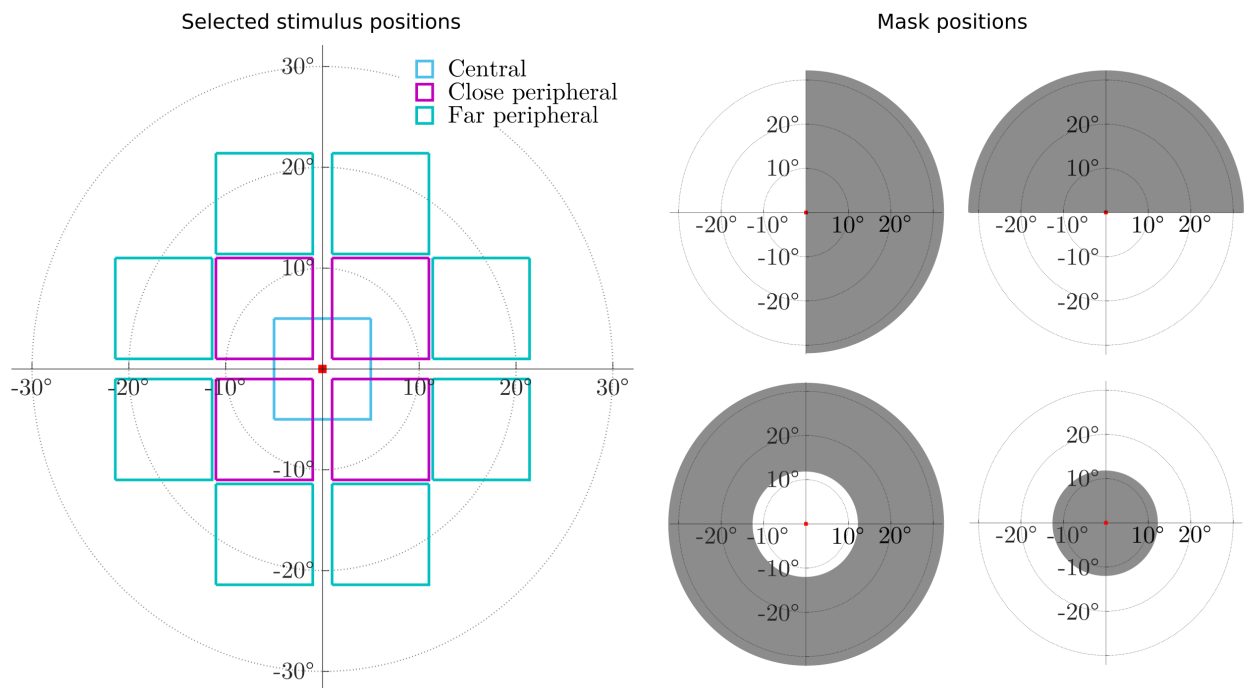


Figure 10: On the left side, all used stimulus positions for data collection are presented. In the experiment, we tested one time without any overlay mask, followed by tests with each of the four vision impairment masks displayed on the right.

### 6.3 The Stimulus

In order to assess the overall measurement quality, we use a 10-second recording of the individual with eyes closed to record the corresponding alpha waves. Subsequently, we performed the measurements with stimuli and a fixation target present. The tested positions are displayed in Figure 10 on the left. Close peripheral stimuli were centered on visual angles of  $(6^\circ, 6^\circ)$  while far peripherals were centered on  $(6^\circ, 16^\circ)$  or  $(16^\circ, 6^\circ)$ , respectively. All stimuli had a width and height of  $10^\circ$  of visual angle. We chose pattern-onset stimuli at 7.5 Hz with a size of 10 degrees. Each stimulus was presented for 7 seconds on and 3 seconds off. The pattern had a checkerboard size of 1 degree. We chose this relatively large grating size due to the decreasing visual acuity in the periphery and to compensate for the spherical aberration effects of the lenses used in VR-devices in the far periphery.

### 6.4 Simulated Vision Defects

As this is a pre-clinical study, we were limited to healthy individuals for our tests. We used a post-processing filter framework to simulate vision defects. For this task, we used a software framework we introduced in 2018<sup>59</sup> and clinically evaluated in 2019.<sup>60</sup> We applied some modifications to obtain four post-processing masks for this study. Pictograms of the used masks are presented in Figure 10 on the right. Field defects like those can appear because of severe problems with perception due to glaucoma or damage to the optic nerve.<sup>61-63</sup>

## 7. RESULTS

In the following chapter, we are showcasing the use-case of perimetry using our proposed system. We want to perform the signal processing using convolutional neural networks and compare the results to a CCA-based threshold approach. For the deep learning approach, we utilized the QUA<sup>3</sup>CK machine learning development process.<sup>64</sup>



## 7.1 Question

With this study, we wanted to implement a convolutional neural network for objective monitoring of visual perception in central, near-peripheral, and far-peripheral vision. The aim was to objectively distinguish between 'healthy vision' and 'pathologically relevant vision loss'. We defined the latter to be present if a filter mask occluded more than 50% of the stimulus. The dataset for this task was recorded as outlined in Chapter 6. To increase both the amount of available data and the challenge rating (CR) of this task, slices consisting of only 2.5 s of the recorded 7 s periods were used for each classification.

## 7.2 Understanding the Data

For this experiment, we collected EEG data from a total of 15 subjects. Of these, ten were measured at the Karlsruhe Institute of Technology and five at the Massachusetts Institute of Technology. Subject age ranged from 22 to 33 years (mean age: 26.5 years,  $\sigma = 3.2$  years). 20% of the subjects were female. Whilst analyzing the data, it was noticeable that female subjects had stronger SSVEP responses in comparison with most males, as has been found by other research as well.<sup>65</sup> For the binocular case, we collected  $15 \times 5 \times 13 = 975$  seven-second stimulus-snippets at different locations, either with or without a visual impairment overlay mask present. Monocular data was collected for almost all participants and filter masks as well, further increasing the size of the data library. With a step size of 536 ms, we obtain a total of 9 sample-snippets per seven-second interval. This rendered a total of 8775 samples in the data set for the binocular case. After data acquisition, the initial step was to perform a train, test, validation split. Four subjects were randomly selected for the test set. For validation, we picked three subjects from the remaining data. Due to the used impairment masks shown in figure 10, we collected more data for the healthy case than for the pathological case, leading to a slight imbalance in the dataset. Therefore we had to weigh the data accordingly during training. For testing, we used a balanced sub-sample of the recorded test data.

## 7.3 Algorithm and Optimization process

We chose to use the Multi-Net proposed in Figure 7. EEG signals and the position of the stimulus were presented to the network as input data. Position was encoded as a tuple consisting of the distance to the center and the visual quadrant in which the stimulus was positioned. Despite the common standardization of the input data, no further preprocessing was applied and the network was trained in an end-to-end fashion. The outputs of the Multi-Net are the classification in healthy or impaired vision and a regression of the euclidean norm of the calculated CCA vector of the input data. We chose to use the CCA value as an auxiliary output for the network as a means of regularization (see Chapter 5). This regularization by multitask-learning was intended to support the network in finding features that generalize well to unseen data. Finally, we used a Bayesian hyperparameter optimization process based on the python library *scikit-optimize* to find a well-suited set of hyperparameters for the training, again in order to generalize well to unseen data. We optimized for dropout probability, weighting of  $\alpha_1$  to  $\alpha_2$  (compare equation 3), initial learning rate, L2 regularization in the convolutional layers and the amount of neurons in the fully connected layers.

## 7.4 Conclude and Compare

In this section, we will compare the CNN-based results with a CCA-threshold-based state of the art algorithm. For the estimation of the threshold, we used a combination of the training and validation set. Thresholds were defined for nine stimulation positions individually. As the test dataset was imbalanced towards healthy cases, we used a random subset of those for testing and performed the test five times. The means and standard deviations of the resulting accuracies on the test set are presented in Table 1, whereas Figure 11 shows a graphical comparison.

Both algorithms performed best on central stimuli (position 1) while having a decreased performance on peripheral stimuli (positions 2-5). For far peripheral stimuli (6-9), accuracies were slightly above random choice. Notably, there is a difference in accuracy between the monocular case and the binocular case: SSVEPs are harder to detect in the monocular case for both the CCA and the CNN approach.

In general, the differences between the state of the art CCA algorithm and the CNN approach are pretty narrow. Difference between mean CNN accuracy over all stimulus positions and mean CCA accuracy is -1.75% in the binocular case and +0.07% in the monocular case.

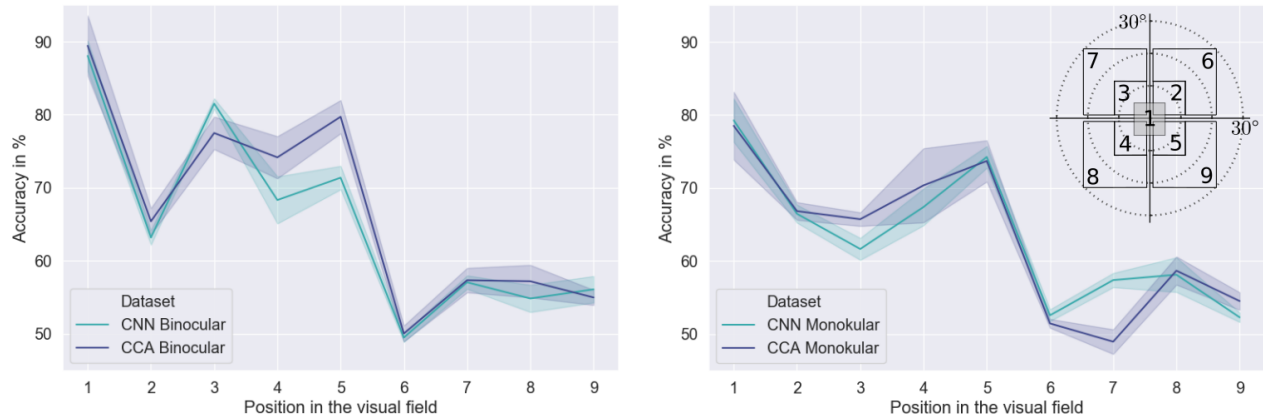


Figure 11: Left plot: Classification accuracies on the test set for the binocular case. Right plot: Accuracies on the monocular test case. Results are partitioned based on the stimulus' position. The shaded area shows the standard deviation over five test runs.

Position	CNN (Bi)	CCA (Bi)	CNN (Mono)	CCA (Mono)
Center (1)	88.1 ( $\sigma = 2.1$ )	89.4 ( $\sigma = 4.1$ )	79.3 ( $\sigma = 3.0$ )	78.5 ( $\sigma = 4.8$ )
Top right (2)	63.19 ( $\sigma = 1.0$ )	65.42 ( $\sigma = 1.8$ )	66.5 ( $\sigma = 1.8$ )	66.9 ( $\sigma = 1.2$ )
Top left (3)	81.5 ( $\sigma = 0.7$ )	77.5 ( $\sigma = 2.2$ )	61.7 ( $\sigma = 1.5$ )	65.7 ( $\sigma = 0.9$ )
Bottom left (4)	68.3 ( $\sigma = 3.2$ )	74.2 ( $\sigma = 2.9$ )	67.4 ( $\sigma = 2.5$ )	70.4 ( $\sigma = 5.1$ )
Bottom right (5)	71.4 ( $\sigma = 1.6$ )	79.7 ( $\sigma = 2.3$ )	74.3 ( $\sigma = 1.5$ )	73.7 ( $\sigma = 2.8$ )
Far top right (6)	49.5 ( $\sigma = 0.6$ )	50.1 ( $\sigma = 1.1$ )	52.6 ( $\sigma = 0.8$ )	51.5 ( $\sigma = 0.6$ )
Far top left (7)	57.1 ( $\sigma = 0.9$ )	57.4 ( $\sigma = 1.7$ )	57.4 ( $\sigma = 1.0$ )	49.0 ( $\sigma = 1.7$ )
Far bottom left (8)	54.9 ( $\sigma = 1.9$ )	57.2 ( $\sigma = 2.2$ )	58.1 ( $\sigma = 2.4$ )	58.7 ( $\sigma = 1.9$ )
Far bottom right (9)	56.1 ( $\sigma = 1.8$ )	55.0 ( $\sigma = 1.0$ )	52.3 ( $\sigma = 0.7$ )	54.5 ( $\sigma = 1.2$ )

Table 1: Comparison of mean accuracies reached with a CNN and CCA approach.

## 7.5 Knowledge Transfer

The CNN was able to classify the unseen test data with comparable accuracy to the CCA. Thus, we assume that the CNN was not only able to learn a transformation from the time into the frequency domain with a focus on the stimulated frequencies over multiple channels, but furthermore was able to learn which thresholds are considered healthy and which are pathologic for individual stimulus positions.

Investigating why the CNN was not able to outperform the CCA based classification, two possible reasons come to mind. Firstly, it might be the case that the EEG data during SSVEP stimulation contains no relevant features other than the elicited harmonics of the stimulation frequency. Thus, both CNN and CCA would rely on the same set of basic features and their performance should be equal. Another hypothesis might be that the EEG signal does contain features despite the harmonics, but that these features are not present in all subjects or vary across subjects so that a larger dataset would be needed in order to identify the relationships between these features and the desired classification outcome. From the current results, we can only conclude that our CNN based approach is not finding features serving classification better than a state of the art CCA approach.

One approach to further increase the performance of the CNN might thus be found in increasing the size of the dataset. Furthermore, more complex stimulation patterns, which, for example, make use of nonlinear electrophysiological processes in the human visual system,<sup>66</sup> may favor a CNN approach in the future.

## 8. CONCLUSION AND OUTLOOK

In this paper, we presented a generalized, modular, closed-loop system approach towards the objective assessment of neuro-visual function. We showcased that such a system has to fulfill a multitude of requirements in order to work well within a machine learning research environment. There is the need to define and display stimuli and sequences of such in various displaying technologies. It needs to connect and synchronize EEG and labels in order to collect data for the learning process. Moreover, it needs the means to perform inferences with trained machine learning models to run in a closed-loop fashion. We proposed an implementation serving those requirements and carried out a use-case study aiming towards objective visual field testing. A Multi-Task CNN for the classification of pathological and healthy cases was trained and we compared the results against a CCA-threshold algorithm baseline. Despite the comparably small dataset used for training, the CNN-based approach could indeed perform on the same level like the state of the art threshold-based CCA approach.

As this study used only the data of 15 individuals, we assume that a machine learning model trained on a larger dataset would likely outperform our presented solution. For future implementations one major goal is to implement high frequency stimulus embedding analog to 'Sublime'<sup>29</sup> in an attempt to further improve patient comfort. We hope that our work can contribute to EEG classification and objective measurement of the neuro-visual function using machine learning. We see great potential for brain-computer-interface-based automatic assessments in ophthalmology in the future.

## ACKNOWLEDGMENTS

We want to thank everyone who has contributed to making this project a success. Special thanks go to the participants of our studies, who were willing to get their retinas and visual cortices stimulated. We would also like to thank the Massachusetts and Karlsruhe Institutes of Technology for enabling the research on this collaborative project.

## REFERENCES

- [1] Marr, D., [*Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*], MIT Press Ltd (2010).
- [2] Zhaoping, L. and Li, Z., [*Understanding vision: theory, models, and data*], Oxford University Press, USA (2014).
- [3] Chalupa, L. M. and Werner, J. S., [*The visual neurosciences, Vols. 1 & 2*], MIT press (2004).
- [4] Cahill, M. T., Banks, A. D., Stinnett, S. S., and Toth, C. A., "Vision-related quality of life in patients with bilateral severe age-related macular degeneration," *Ophthalmology* **112**(1), 152 – 158 (2005).
- [5] Gutierrez, P., Wilson, M. R., Johnson, C., Gordon, M., Cioffi, G. A., Ritch, R., Sherwood, M., Meng, K., and Mangione, C. M., "Influence of glaucomatous visual field loss on health-related quality of life," *Archives of ophthalmology* **115**(6), 777–784 (1997).
- [6] Nelson, P., Aspinall, P., and O'Brien, C., "Patients' perception of visual impairment in glaucoma: a pilot study," *British Journal of Ophthalmology* **83**(5), 546–552 (1999).
- [7] Wolfram, C., "Demografie und altersabhängigkeit bei augenerkrankungen," *Der Ophthalmologe* **112**(1), 73–81 (2015).
- [8] Holmes, D., "Reconstructing the retina," *Nature* **561**(7721), S2–S2 (2018).
- [9] Wong, E. Y., Keeffe, J. E., Rait, J. L., Vu, H. T., Le, A., McCarty, C., and Taylor, H. R., "Detection of undiagnosed glaucoma by eye health professionals," *Ophthalmology* **111**(8), 1508–1514 (2004).
- [10] Courchesne, E., Hillyard, S. A., and Galambos, R., "Stimulus novelty, task relevance and the visual evoked potential in man," *Electroencephalography and clinical neurophysiology* **39**(2), 131–143 (1975).
- [11] Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S., and Hillyard, S. A., "Cortical sources of the early components of the visual evoked potential," *Human brain mapping* **15**(2), 95–111 (2002).
- [12] Clark, V. P., Fan, S., and Hillyard, S. A., "Identification of early visual evoked potential generators by retinotopic and topographic analyses," *Human brain mapping* **2**(3), 170–187 (1994).

- [13] Klistorner, A. I., Graham, S. L., Grigg, J. R., and Billson, F. A., "Multifocal topographic visual evoked potential: improving objective detection of local visual field defects.," *Investigative ophthalmology & visual science* **39**(6), 937–950 (1998).
- [14] Hood, D. C. and Zhang, X., "Multifocal erg and vep responses and visual fields: comparing disease-related changes," *Documenta Ophthalmologica* **100**(2-3), 115–137 (2000).
- [15] Hood, D. C., Thienprasiddhi, P., Greenstein, V. C., Winn, B. J., Ohri, N., Liebmann, J. M., and Ritch, R., "Detecting early to mild glaucomatous damage: a comparison of the multifocal vep and automated perimetry," *Investigative ophthalmology & visual science* **45**(2), 492–498 (2004).
- [16] Hood, D. C. and Greenstein, V. C., "Multifocal vep and ganglion cell damage: applications and limitations for the study of glaucoma," *Progress in retinal and eye research* **22**(2), 201–251 (2003).
- [17] Graham, S. L., Klistorner, A. I., Grigg, J. R., and Billson, F. A., "Objective vep perimetry in glaucoma: asymmetry analysis to identify early deficits.," *Journal of glaucoma* **9**(1), 10–19 (2000).
- [18] Graham, S. L., Klistorner, A. I., and Goldberg, I., "Clinical application of objective perimetry using multifocal visual evoked potentials in glaucoma practice," *Archives of ophthalmology* **123**(6), 729–739 (2005).
- [19] Baseler, H., Sutter, E., Klein, S., and Carney, T., "The topography of visual evoked response properties across the visual field," *Electroencephalography and clinical neurophysiology* **90**(1), 65–81 (1994).
- [20] Lenassi, E., Likar, K., Stirn-Kranjc, B., and Breclj, J., "Vep maturation and visual acuity in infants and preschool children," *Documenta Ophthalmologica* **117**(2), 111–120 (2008).
- [21] Norcia, A. M. and Tyler, C. W., "Spatial frequency sweep vep: visual acuity during the first year of life," *Vision research* **25**(10), 1399–1408 (1985).
- [22] Herrmann, C. S., "Human eeg responses to 1–100 hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena," *Experimental brain research* **137**(3-4), 346–353 (2001).
- [23] Hillyard, S. A., Hinrichs, H., Tempelmann, C., Morgan, S. T., Hansen, J. C., Scheich, H., and Heinze, H.-J., "Combining steady-state visual evoked potentials and f mri to localize brain activity during selective attention," *Human brain mapping* **5**(4), 287–292 (1997).
- [24] Morgan, S., Hansen, J., and Hillyard, S., "Selective attention to stimulus location modulates the steady-state visual evoked potential," *Proceedings of the National Academy of Sciences* **93**(10), 4770–4774 (1996).
- [25] Zhu, D., Bieger, J., Garcia Molina, G., and Aarts, R. M., "A survey of stimulation methods used in ssvep-based bcis," *Computational intelligence and neuroscience* **2010** (2010).
- [26] Wai, A. A. P., Goh, Z., De Foo, S., and Guan, C., "A study of ssvep responses in case of overt and covert visual attention with different view angles," in [2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)], 3847–3852, IEEE (2018).
- [27] Bergholz, R., Lehmann, T., Fritz, G., and Rütther, K., "Fourier transformed steady-state flash evoked potentials for continuous monitoring of visual pathway function," *Documenta ophthalmologica* **116**(3), 217–229 (2008).
- [28] Kurtenbach, A., Langrová, H., Messias, A., Zrenner, E., and Jägle, H., "A comparison of the performance of three visual evoked potential-based methods to estimate visual acuity," *Documenta Ophthalmologica* **126**(1), 45–56 (2013).
- [29] Armengol-Urpi, A. and Sarma, S. E., "Sublime: a hands-free virtual reality menu navigation system using a high-frequency ssvep-based brain-computer interface," in [Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology], 1–8 (2018).
- [30] Zao, J. K., Jung, T.-P., Chang, H.-M., Gan, T.-T., Wang, Y.-T., Lin, Y.-P., Liu, W.-H., Zheng, G.-Y., Lin, C.-K., Lin, C.-H., et al., "Augmenting vr/ar applications with eeg/eog monitoring and oculo-vestibular recoupling," in [International Conference on Augmented Cognition], 121–131, Springer (2016).
- [31] Nakanishi, M., Wang, Y.-T., Jung, T.-P., Zao, J. K., Chien, Y.-Y., Diniz-Filho, A., Daga, F. B., Lin, Y.-P., Wang, Y., and Medeiros, F. A., "Detecting glaucoma with a portable brain-computer interface for objective assessment of visual function loss," *JAMA ophthalmology* **135**(6), 550–557 (2017).
- [32] Nagel, S. and Spüler, M., "World's fastest brain-computer interface: Combining eeg2code with deep learning," *PloS one* **14**(9) (2019).
- [33] Cecotti, H. and Graser, A., "Convolutional neural networks for p300 detection with application to brain-computer interfaces," *IEEE transactions on pattern analysis and machine intelligence* **33**(3), 433–445 (2010).

- [34] Sakhavi, S., Guan, C., and Yan, S., “Learning temporal information for brain-computer interface using convolutional neural networks,” *IEEE Transactions on Neural Networks and Learning Systems* **29**(11), 5619–5629 (2018).
- [35] Kwak, N.-S., Müller, K.-R., and Lee, S.-W., “A convolutional neural network for steady state visual evoked potential classification under ambulatory environment,” *PloS one* **12**(2), e0172578 (2017).
- [36] Aznan, N. K. N., Bonner, S., Connolly, J., Al Moubayed, N., and Breckon, T., “On the classification of ssvep-based dry-eeeg signals via convolutional neural networks,” in [2018 *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*], 3726–3731, IEEE (2018).
- [37] Bevilacqua, V., Tattoli, G., Buongiorno, D., Loconsole, C., Leonardis, D., Barsotti, M., Frisoli, A., and Bergamasco, M., “A novel bci-ssvep based approach for control of walking in virtual environment using a convolutional neural network,” in [2014 *International Joint Conference on Neural Networks (IJCNN)*], 4121–4128, IEEE (2014).
- [38] Ravi, A., Manuel, J., Heydari, N., and Jiang, N., “A convolutional neural network for enhancing the detection of ssvep in the presence of competing stimuli,” in [2019 *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*], 6323–6326, IEEE (2019).
- [39] Ibáñez-Soria, D., Soria-Frisch, A., Garcia-Ojalvo, J., and Ruffini, G., “Characterization of the non-stationary nature of steady-state visual evoked potentials using echo state networks,” *PloS one* **14**(7) (2019).
- [40] El-Fiqi, H., Wang, M., Salimi, N., Kasmarik, K., Barlow, M., and Abbass, H., “Convolution neural networks for person identification and verification using steady state visual evoked potential,” in [2018 *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*], 1062–1069, IEEE (2018).
- [41] Ratti, E., Waninger, S., Berka, C., Ruffini, G., and Verma, A., “Comparison of medical and consumer wireless eeg systems for use in clinical trials,” *Frontiers in Human Neuroscience* **11**, 398 (2017).
- [42] <https://www.neuroelectronics.com/solutions/enobio>. Accessed: 2020-03-04.
- [43] Gamma, E., Helm, R., Johnson, R., and Vlissides, J., [*Design Patterns: Elements of Reusable Object-Oriented Software*], Addison-Wesley Longman Publishing Co., Inc., USA (1995).
- [44] Hotelling, H., “Relations between two sets of variates,” in [*Breakthroughs in statistics*], 162–190, Springer (1992).
- [45] Lin, Z., Zhang, C., Wu, W., and Gao, X., “Frequency recognition based on canonical correlation analysis for ssvep-based bcis,” *IEEE transactions on biomedical engineering* **53**(12), 2610–2614 (2006).
- [46] Mondini, V., Mangia, A. L., Talevi, L., and Cappello, A., “Sinc-windowing and multiple correlation coefficients improve ssvep recognition based on canonical correlation analysis,” *Computational intelligence and neuroscience* **2018** (2018).
- [47] Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., and Lance, B. J., “Eeg-net: a compact convolutional neural network for eeg-based brain-computer interfaces,” *Journal of neural engineering* **15**(5), 056013 (2018).
- [48] Raghunath, S. M., Ulloa Cerna, A., Jing, L., vanMaanen, D., Stough, J. V., Hartzel, D., Leader, J., Good, C., Fornwalt, B. K., and Haggerty, C. M., “A deep neural network for predicting incident atrial fibrillation directly from 12-lead electrocardiogram traces,” *Circulation* **140**(Suppl.1), A14407–A14407 (2019).
- [49] Wright, J. and Liley, D., “Dynamics of the brain at global and microscopic scales: Neural networks and the eeg,” *Behavioral and Brain Sciences* **19**(2), 285–295 (1996).
- [50] Smith, S. L., Kindermans, P.-J., Ying, C., and Le, Q. V., “Don’t decay the learning rate, increase the batch size,” *arXiv preprint arXiv:1711.00489* (2017).
- [51] Cecotti, H., “A time–frequency convolutional neural network for the offline classification of steady-state visual evoked potential responses,” *Pattern Recognition Letters* **32**(8), 1145–1153 (2011).
- [52] Schirrneister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggenesperger, K., Tangermann, M., Hutter, F., Burgard, W., and Ball, T., “Deep learning with convolutional neural networks for eeg decoding and visualization,” *Human brain mapping* **38**(11), 5391–5420 (2017).
- [53] Tibor Schirrneister, R., Gemein, L., Eggenesperger, K., Hutter, F., and Ball, T., “Deep learning with convolutional neural networks for decoding and visualization of eeg pathology,” *arXiv preprint arXiv:1708.08012* (2017).

- [54] Caruana, R. A., “Multitask connectionist learning,” in [*In Proceedings of the 1993 Connectionist Models Summer School*], Citeseer (1993).
- [55] Snoek, J., Larochelle, H., and Adams, R. P., “Practical bayesian optimization of machine learning algorithms,” in [*Advances in neural information processing systems*], 2951–2959 (2012).
- [56] Odom, J. V., Bach, M., Brigell, M., Holder, G. E., McCulloch, D. L., Mizota, A., Tormene, A. P., for Clinical Electrophysiology of Vision, I. S., et al., “Iscv standard for clinical visual evoked potentials:(2016 update),” *Documenta Ophthalmologica* **133**(1), 1–9 (2016).
- [57] Erba, G., “Preventing seizures from “pocket monsters”: a way to control reflex epilepsy,” (2001).
- [58] Fisher, R. S., Harding, G., Erba, G., Barkley, G. L., and Wilkins, A., “Photic-and pattern-induced seizures: a review for the epilepsy foundation of america working group,” *Epilepsia* **46**(9), 1426–1441 (2005).
- [59] Stock, S., Erler, C., and Stork, W., “Realistic simulation of progressive vision diseases in virtual reality,” in [*Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*], *VRST '18*, Association for Computing Machinery, New York, NY, USA (2018).
- [60] Stock, S. C., Erler, C., Stork, W., Labuz, G., Son, H. S., Khoramnia, R., and Auffarth, G. U., “Suitability of virtual reality for vision simulation—a case study using glaucomatous visual fields,” *Investigative Ophthalmology & Visual Science* **60**(9), 2441–2441 (2019).
- [61] Kim, J. M., Kyung, H., Shim, S. H., Azarbod, P., and Caprioli, J., “Location of initial visual field defects in glaucoma and their modes of deterioration,” *Investigative ophthalmology & visual science* **56**(13), 7956–7962 (2015).
- [62] Oyster, C. W., “The human eye,” *Sunderland, MA: Sinauer* (1999).
- [63] Lang, G. K., [*Ophthalmology: A Short Textbook.*], Thieme (2000).
- [64] Becker, J., Grimm, D., Hotfilter, T., Meier, C., Molinar, G., Stang, M., Stock, S., and Stork, W., “Qua<sup>3</sup>ck a machine learning development process,” in [*Symposium Artificial Intelligence for Science, Industry and Society (AISIS 2019)*], (2019).
- [65] Skosnik, P. D., Krishnan, G. P., Vohs, J. L., and O’Donnell, B. F., “The effect of cannabis use and gender on the visual steady state evoked potential,” *Clinical neurophysiology* **117**(1), 144–156 (2006).
- [66] Nguyen, K. T., Liang, W.-K., Lee, V., Chang, W.-S., Muggleton, N. G., Yeh, J.-R., Huang, N. E., and Juan, C.-H., “Unraveling nonlinear electrophysiologic processes in the human visual system with full dimension spectral analysis,” *Scientific reports* **9**(1), 1–13 (2019).