Towards High-Frequency SSVEP-Based Target Discrimination with an Extended Alphanumeric Keyboard

Sahar Abdelnabi[1], Michael Xuelin Huang[2], Andreas Bulling[3]

***Abstract*— Despite significant advances in using Steady-State Visually Evoked Potentials (SSVEP) for on-screen target discrimination, existing methods either require intrusive, low-frequency visual stimulation or only support a small number of targets. We propose SSVEPNet: a convolutional long short-term memory (LSTM) recurrent neural network for high-frequency stimulation (≥30Hz) using a large number of visual targets. We evaluate our method for discriminating between 43 targets on an extended alphanumeric virtual keyboard and compare three different frequency assignment strategies. Our experimental results show that SSVEPNet significantly outperforms state-of-the-art correlation-based methods and convolutional neural networks. As such, our work opens up an exciting new direction of research towards a new class of unobtrusive and highly expressive SSVEP-based interfaces for text entry and beyond.**

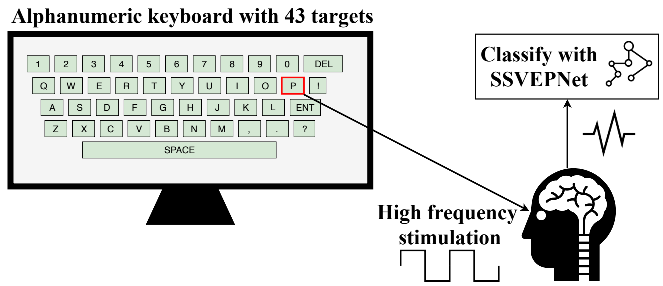

Fig. 1. We propose SSVEPNet: a convolutional LSTM network to discriminate 43 on-screen targets of an extended alphanumeric virtual keyboard from high-frequency Steady-State Visually Evoked Potentials (SSVEP) responses. Details of the architecture are shown in Fig. 4.

I. INTRODUCTION

Steady-State Visually Evoked Potentials (SSVEP) are natural brain responses at a frequency that matches periodic (flickering) visual stimuli [5]. In turn, by comparing the frequency of the SSVEP response to them, different visual targets can be discriminated. SSVEP-based text entry has been explored as a particularly promising use case for SSVEP [11] with recent years having seen increasing research activities in this area. In particular, research recently focused on increasing the number of visual targets (up to 40) using frequency-phase joint encoding and approximations methods [7, 12, 13] as well as on pushing classification accuracy (∼90%) [7, 12].

Despite significant advances, prior works fall short in two aspects that are particularly crucial for future, practical application of SSVEP-based interaction. For one, prior work largely studied low frequency stimulation (8Hz-16Hz). These visual stimuli can induce clear SSVEP responses but are visually intrusive to users and can easily cause fatigue [19]. This is in line with the finding of a prior study [17], which suggested that low frequencies (6–14.9 Hz) caused more discomfort compared with high frequencies (26–34.7 Hz). Second, studies that have explored high frequency stimulation (>20Hz) have either only classified a small number (≤5) of targets [2, 18] or required specialized equipment for visual stimulation, such as external LEDs [6, 17].

At the same time, recent deep learning techniques have been demonstrated to achieve a promising performance for SSVEP classification tasks. Recent breakthroughs include the use of Convolutional Neural Networks (CNNs) [16] and Deep Recurrent Convolutional Networks [3] for SSVEP classification and for top-down SSVEP decoding [1]. All of these studies primarily aimed to identify an optimal feature representation. The most challenging, but also practically most relevant, discrimination of a large number of targets using unobtrusive, high-frequency visual stimulation on an off-the-shelf computer screen remains unexplored.

This paper presents the first study that exploits high-frequency stimulation for SSVEP-based desktop applications. Specifically, we investigate a virtual QWERTY-style keyboard input as an example (see Fig. 1). To ensure that our findings can be generalized for other desktop applications, we rely on no dedicated equipment and show visual stimuli on a computer screen. We also study pertinent frequency assignment strategies in such applications, including suitable frequency ranges of all targets, the minimal frequency difference between targets, and different frequency decoding methods. As SSVEP response to high frequency stimulation is much weaker than that of low frequency stimulation, this task is considerably more challenging [19]. To address this, we propose SSVEPNet: a convolutional long short-term memory (LSTM) recurrent neural network for SSVEP-based target discrimination. Our experimental results demonstrate that SSVEPNet achieves an accuracy improvement of 25% over the state-of-the-art methods that used Canonical Correlation Analysis (CCA) [7, 10] and CNNs [4, 16].

The contributions of our work are three-fold. First, we conduct the first study on offline SSVEP-based keyboard input using high frequency stimulation and a large number of visual targets. Second, we record a new dataset for this challenging task (MPII-SSVEP) and we make it publicly available (at: https://www.mpi-inf.mpg.de/MPII-SSVEP) to

[1,2]Sahar Abdelnabi and Michael Xuelin Huang are with the Max Planck Institute for Informatics, Saarland Informatics Campus. E-mail: {selnabi,mhuang}@mpi-inf.mpg.de

[3]Andreas Bulling is with the University of Stuttgart, Institute for Visualisation and Interactive Systems. E-mail: andreas.bulling@vis.uni-stuttgart.de

facilitate future algorithm development and evaluation by the research community. Third, we compare different deep learning architectures for a 43-key classification task based on SSVEP responses and propose SSVEPNet, a new method based on convolutional LSTMs that significantly outperforms the state of the art. While there is still room for improvement in terms of performance at this stage, as the first work on SSVEP-based interaction using high frequency stimulation, this study paves the way for an exciting new research direction and opens up a variety of opportunities for advanced Brain-Computer Interaction (BCI) paradigms that rely on unobtrusive and highly expressive SSVEP-based input.

II. THE MPII-SSVEP DATASET

One of the most important issues for SSVEP-based interaction is the stimulation frequency. Low frequency stimulation, as widely studied in prior research, produces a clear SSVEP response, but it is visually intrusive. In contrast, high frequency stimulation is more visually friendly and comfortable, but the corresponding SSVEP signals are more difficult to identify and recognize. Given that there is no existing dataset for SSVEP with high frequency visual stimulation and a large number of targets, we collected our own high-frequency dataset.

Specifically, we recorded three frequency ranges as shown in Tab. I with three frequency intervals. The first range was from 30Hz to 34.2Hz with a 0.1Hz difference; the second range was from 30Hz to 42.6Hz with a 0.3Hz difference; and the final range was from 30Hz to 69.8Hz with a 0.8Hz difference (with a gap from 46Hz to 54Hz to avoid overlapping with the 50 Hz noise [15]).

These three frequency ranges were designed to study the trade-off between frequency range and difference. For instance, while the first range with the lowest frequencies should induce the strongest SSVEP, it has the smallest frequency difference, which in turn makes it difficult to classify. However, the small difference in the first range is necessary, for it allows for frequency assignment to a large number of targets given a specific frequency range.

In addition to frequency encoding, each target was assigned a phase ranging from 0 to 1.75π with a difference of 0.35π. This setting followed the previous design practice in [12]. The stimulus flickering pattern was implemented using the frequency and phase approximation approach [13], which generates a sequence s of frequency f and phase ϕ as follows:

$$s(f,\phi,i) = square[2\pi f(i/RefreshRate)+\phi] \tag{1}$$

TABLE I

DIFFERENT FREQUENCY ASSIGNMENT STRATEGIES.

Range	Difference	Gap in Assignment
30Hz to 34.2Hz	0.1Hz	No gap
30Hz to 42.6Hz	0.3Hz	No gap
30Hz to 69.8Hz	0.8Hz	46Hz to 54Hz

where *square* creates a 50% duty cycle square wave with levels 0 and 1, i is the frame index and $RefreshRate$ is the refresh rate of the screen. This approach can generate frequencies up to half of the screen refresh rate. Additionally, since trial length represents a compromise between speed and classification accuracy [15], we chose a trial length of 2 seconds as an average of the recently used trial lengths in SSVEP systems (5 seconds in [6], and 0.5 second in [12]).

A. Keyboard Layout

To record the data, we developed an offline BCI speller. We used a simplified virtual QWERTY-based keyboard that consisted of 43 targets, including 26 letter keys, 10 digit keys, three control keys (DEL, ENT, and SPACE), and four symbol keys (",", ".", "?", "!"). A screenshot of the interface is shown in Fig. 2. Each letter and digit key was 110x110 pixels (3.1x3.1 cm^2) and the horizontal and vertical distances between keys were 40 pixels. To fit the interface on our monitor (approximately $54x30cm^2$), our target size was slightly smaller than that of the 40-target interface in [12], since we arranged more targets in the horizontal direction. The widths of the DEL, ENT and SPACE keys (210 , 187, 770 pixels, i.e. 5.9cm, 5.2cm, 21.7cm, respectively) were relatively larger than those of the other keys to resemble the QWERTY keyboard layout.

The interface had a grey background (RGB: 127.5, 127.5, 127.5). The perceived flickering was the result of alternating the targets' color between light green (RGB: 223, 252, 174) and dark blue (RGB: 13, 1, 71). These colors were determined according to [14], which suggested that blue/green stimulus is the least provocative color.

B. Recording Procedure

Our experiment consisted of three conditions, one for each frequency range. There were 10 repetitions in each condition and each repetition contained 43 trials when participants fixated at the 43 targets. Each trial lasted for two seconds. Throughout the trial, all targets were flickering and the participants were instructed to only look at a designated target with a unique frequency. Therefore, there were 10 trials in total for each participant for each frequency.

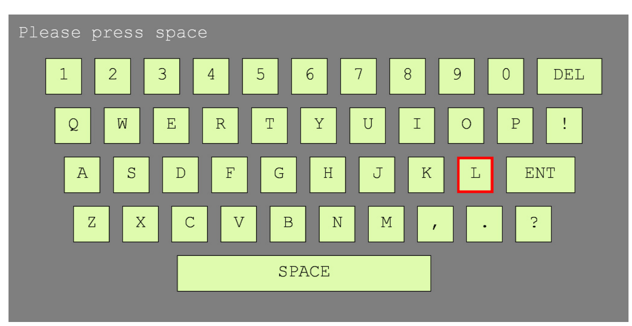

Fig. 2. Screenshot of the virtual keyboard used in our experiment with the red frame indicating the next target. Each target had a height of 3.1cm. The widths of the DEL, ENT and SPACE keys were 5.9cm, 5.2cm 21.7cm, respectively. All other targets had a width of 3.1cm.

Before the actual recording of each trial, we used a red frame to indicate the designated target for the participant to fixate on during the trial (see Fig. 2). To allow for a comfortable pace and sufficient time for the participant to look for the target, we instructed participants to initiate the trial by pressing the SPACE key on a physical keyboard. We encouraged participants to take short breaks at will to avoid eye fatigue, by simply not pressing the SPACE key to start the next trial. Also, we introduced a five-minute break after the completion of half of the total number of trials.

Given that the key sizes were not all the same in the virtual keyboard, key size and location may introduce a bias for SSVEP response. To minimize the biases, we randomly shuffled the 43 frequencies and phases and assigned them to different locations on the virtual keyboard in each repetition. In addition, to avoid learning effects in the experiment, the 43 frequencies were randomly ordered in each repetition. Likewise, we also randomly changed the order of the three frequency ranges for different participants. Randomizing the frequency, phrase, and trial, on the other hand, can allow our network to learn from more diverse data and be more robust to different locations and sizes. A graphical representation of the experiment order is shown in Fig. 3.

The experiment was approved by the ethical review board of the department of computer sciences in Saarland University and participants were asked to sign a consent form with detailed explanation of the experiment and potential risks. We did not proceed with the experiment if a participant was concerned that he/she had or might have a history of seizures, migraines, or light sensitivity. Participants were informed that they could stop the experiment if they felt discomfort.

C. Participants

We recruited 20 participants through university mailing lists (seven female), aged from 23-45 years (mean=28), from different nationalities, most of whom are computer science students. After agreeing to participate and signing the informed consent form, participants were asked to fill in a small questionnaire regarding demographic information. Participants were then assisted to put on the EEG head cap and were instructed to avoid unnecessary movements as much as possible during the recording and to fixate on the red frame-indicated target. Participants on average sat at a distance of approximately 60cm from the screen. The experiment was conducted in a quiet room with dimmed light and no electromagnetic shielding.

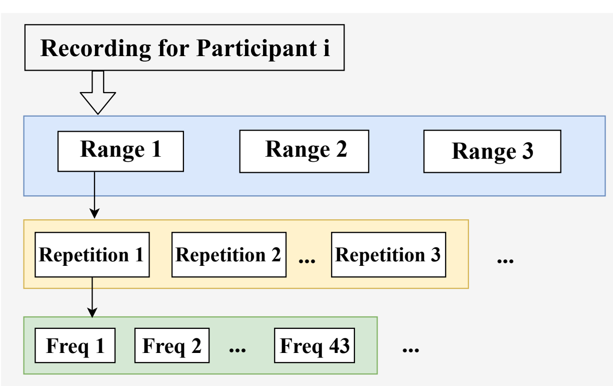

Fig. 3. Illustration of our experimental protocol. Each participant (20 participants) was asked to look at keys on an extended alphanumeric keyboard flickering in three frequency ranges. 10 repetitions were performed for each range (as previously used in [7]), each consisting of 43 individual frequencies. Each trial lasted for two seconds (i.e. 1000 time samples).

D. Apparatus

EEG data were collected using the 32-channel Mobita portable battery operated amplifier (by TMSi, Enschede, The Netherlands), with a sampling rate of 500Hz. The device has no direct way to measure impedance of electrodes; therefore, we visually inspected the quality of signal before experiment recording. The device uses unipolar recording where the reference is the average of all connected unipolar electrodes. The device has a wristband ground, which needs to be in good contact to avoid the 50Hz noise [15]. The TMSi MATLAB interface was used to acquire data and was integrated with the stimulus interface. The recordings were logged with time annotations of trials' start and end. Channels were located according to the international 10–20 electrode system. The stimulus interface was implemented using Matlab 2017b and Psychophysics Toolbox Version 3, and shown on a ViewSonic XG2530 25-Inch LED monitor with adaptable refresh rate up to 144Hz and a pixel resolution of 1920x1080. This toolbox allows for synchronization between refresh rate and color flickering and thus ensures a correct implementation of Eq(1). The experiment was performed on a computer with Intel Xeon E5-1620 V2 3.70 GHz processor, and NVIDIA K2000 1GB graphics card.

III. SSVEPNET ARCHITECTURE

The architecture of our proposed SSVEPNet for target discrimination is shown in Fig. 4. It consists of two layers of 1D convolution and maxpooling for low-level feature extraction, and five LSTM layers to extract temporal characteristics of the SSVEP patterns. Given that dense layers can introduce a large number of weights, we only use one dense layer for classification with softmax to keep our model light and avoid over-fitting. For the same reason, we apply a dropout layer after each pooling layer and each LSTM layer. After the last LSTM layer, we flatten the hidden states of the whole sequence, and feed them to the last dense layer as the last hidden state would not be an enough representation of the whole sequence. We use categorical cross-entropy in the loss function and train the network with Adam optimizer. We use early stopping of training as a regularization method by monitoring the validation accuracy.

Each of our training samples comprises the multi-channel EEG signals in a two-second trial, when a participant is looking at a particular key that flickers at a certain frequency. The training set contains samples from multiple participants and we aim to build a user-independent model for SSVEP pattern classification. In our performance evaluation section, we examine different variants of the proposed architecture

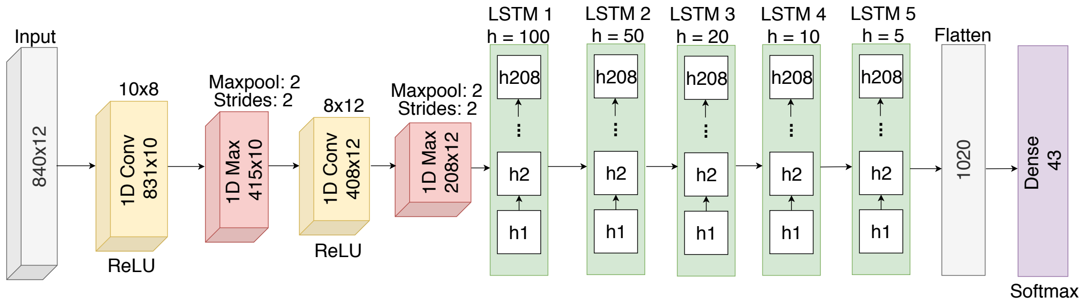

Fig. 4. The architecture of SSVEPNet, which consists of two convolutional layers (with ReLU activations), five LSTMs layers and a dense layer (with Softmax activation). Convolution filters are indicated as Width x Channels and the hidden size of the LSTM is indicated above the LSTM layers. Additionally, the output size of each operation is indicated in the corresponding box.

with different numbers of LSTM and convolution layers and compare the performance of 1D and 2D convolutions.

IV. EXPERIMENTS

We conducted several experiments to evaluate the performance of the proposed method. In the following, we first describe the EEG signal processing procedure, then demonstrate the training setup of SSVEPNet, and finally we describe the examined baseline models.

A. EEG Signal Preprocessing

Trials were extracted and labelled according to the annotations done by the recording and the interface programs with a 150ms delay to account for visual system latency delay [8]. We used a notch filter (between 48Hz and 52Hz) to process each trial to filter out the 50Hz noise [15]. For the deep learning models, each channel in the trial was normalized by removing the mean and dividing by the standard deviation of that channel. Following the previous practise, we used 12 channels (CP1, CP2, CP6, P7, P3, Pz, P4, P8, POz, O1, Oz, O2) that cover the most important channels for SSVEP in the occipital and parietal lobes [16]. All signal preprocessing was done on MATLAB.

B. Training of SSVEPNET

Regarding implementation details of our network, we used Adam optimizer (β_1=0.9 and β_2=0.999), dropout (drop probability=0.4), and $L2$ regularization (0.005) in the last dense layer. We set the learning rate to 0.001 and the patience parameter in early stopping to 5 epochs. Models were implemented in Keras with TensorFlow backend. The dataset was split into training, validation, and test sets with percentages 70%, 10%, and 20% respectively, where the selection of hyperparameters was based on validation performance.

C. Baselines

We compared SSVEPNet to other potential architectures suggested in the related work. The first architecture is the Compact-CNN [16] which is a 2D CNN consisting of a normal convolution, a depth-wise convolution, and a separable convolution. Due to the suitability for the size of our training data, we also compared with another simple CNN architecture [4], which we refer to as Shallow CNN. It contains a 1D convolution with 1D maxpooling and batch norm, followed by a dense layer and a Softmax layer.

Additionally, we compared against CCA and combined-CCA as widely used SSVEP decoding methods. CCA computes the correlation between two multidimensional variables X and Y by finding weights W_x, W_y that maximize the correlation between the projections $X^T W_X$, $Y^T W_y$ [10]. It measures the canonical correlation between multi-channel EEG signals and a set of reference (sines and cosines) signals with different stimuli frequencies and considers the one with maximal correlation as the SSVEP frequency [10]. While CCA requires no learning or supervision, it may not produce optimal results due to the use of artificial references. In contrast, the combined-CCA model [7] leverages the collected signals from a number of trials to form a more realistic reference.

V. PERFORMANCE EVALUATION

In this section, we report the performance of SSVEPNet against the examined baselines, and answer key questions regarding high frequency stimulation, including the suitable frequency assignment strategy for a large number of targets and the appropriate deep learning architecture and parameters for the classification task.

A. Performance of Frequency Assignment Strategies

Fig. 5 shows the performance comparison of SSVEPNet against CCA and combined-CCA in the classification of the three studied frequency ranges. Most importantly, we see that the lowest frequency range with the smallest frequency difference (0.1Hz) is the most suitable frequency assignment strategy. Comparison within the same classification method indicates that the use of "Range 1" yielded the highest accuracy, followed by that of "Range 2" and "Range 3". Specifically, SSVEPNet achieved a classification accuracy of 31% for "Range 1" (30Hz-34.2Hz), 4% for "Range 2" (30Hz-42.6Hz), and 3% for "Range 3" (30Hz-69.8Hz). It is worth noting that the difference in accuracy between "Range 2" and "Range 3" is relatively modest and significantly lower than that between "Range 1" and the others, meaning that using frequencies ranging from 30Hz-42.6Hz and 30Hz-69.8Hz for SSVEP-based interaction can be similarly difficult. This

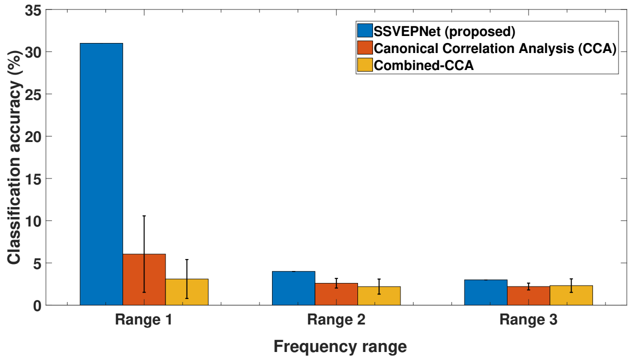

Fig. 5. A comparison between the classification accuracy of SSVEPNet (blue), CCA (dark orange) and Combined-CCA methods (light orange) for the three frequency ranges. Error bars for CCA and Combined-CCA indicate the standard deviation of accuracy across all participants, while training of SSVEPNet is performed using cross-participant concatenation of data.

result implies that the use of a relatively lower frequency that can produce stronger SSVEP responses is still more preferable, even though it results in a smaller frequency difference among different targets. This can be a useful guideline to identify the trade-off between frequency range and difference.

Our comparison demonstrates that SSVEPNet yields the best performance across different frequency assignment strategies. In particular, it gives a significantly higher accuracy than CCA and combined-CCA in "Range 1". Although there is still much room for improvement, our method is promising and outperforms the chance level (only 2.3%) and the state of the art methods.

B. *Ablation Study*

Given the above results, this section focuses on "Range 1" to further investigate the model architecture and parameters, such as different depths and layer types. Tab. II shows the performances. The first four rows give the results of LSTMs with different depths, whose hidden states of each step in the whole sequence are flattened and fed to one dense layer. The accuracy increased with the increase of depth, ranging from 18.6% (1 Layer) to 26.7% (5 Layers), and it saturated after that. Interestingly, applying one and two additional convolution layers at the feature extraction stage before the five LSTM layers further increased accuracy to 29.4% and 31%, respectively (fifth and sixth (SSVEPNet) rows in Tab. II).

Moreover, the last two rows in Tab. II show the accuracy of the studied Compact-CNN and Shallow-CNN. However, neither of these models could achieve comparable results as LSTMs. The Compact-CNN network yielded a low accuracy (3.6%). We also noticed that the training accuracy of Compact-CNN converged at a relatively low value compared with those of the LSTMs. Although the training accuracy of the Shadow CNN reached a higher value than that of the Compact-CNN, it still had a low accuracy (4.2%). This result implies that LSTMs can be a better fit to process SSVEP signals compared with CNN.

TABLE II

PERFORMANCE OF DIFFERENT VARIANTS OF DEEP LEARNING MODELS.

Model Description	Accuracy
3 Layers LSTM (size: 100,50,10)	18.6%
4 Layers LSTM (size: 100,50,10,5)	23.3%
5 Layers LSTM (size: 100,50,20,10,5)	26.7%
6 Layers LSTM (size: 100,50,20,10,5,3)	26.3%
1D conv + 5 Layers LSTM	29.4%
SSVEPNet	**31**%
Compact-CNN (2D convolution) [16]	3.6%
Shallow CNN (1D convolution) [4]	4.2%

VI. DISCUSSION

Our work represents the first study to exploit high-frequency stimulation for SSVEP-based desktop interaction with a large number of targets. We evaluated three frequency assignment strategies and our experiment results suggest that the relative lower frequency range with a smaller frequency difference is the most suitable strategy in our setting. Therefore, we suggest that future related studies should apply frequencies from 30Hz-34.2Hz for a task with a similar number of targets (43 in our case). To facilitate future research on this challenging task, we make our dataset publicly available.

In addition, we proposed SSVEPNet, a novel Conv-LSTM architecture, to identify gaze target based on SSVEP responses. Our method significantly outperforms the state-of-the-art CCA as well as CNN methods.

VII. CONCLUSION

This paper reports the first study on SSVEP-based discrimination of a large number of on-screen targets and high-frequency visual stimulation. We proposed SSVEPNet, a Conv-LSTM network, to discriminate the SSVEP patterns of 43 keys on a on-screen virtual keyboard. Experimental results on a newly collected dataset show that discrimination difficulty increases significantly with increasing stimulation frequency. That is, given the same starting frequency, a relatively narrower frequency range can be more suitable for such applications than a wider range that covers higher frequencies. Our method shows promising performance improvements on this challenging task of about 25% accuracy higher than the best performing CCA methods as well as a CNN-only method. As such, our work makes a significant step towards more unobtrusive and expressive, and thus more practically useful, SSVEP-based brain-computer interfaces. The implication is that convolution can produce suitable features to capture useful information in SSVEP signals, however, LSTMs play a critical role to identify the temporal features to distinguish different SSVEP frequencies. As such, we also recommend that interested readers can explore the use of Conv-LSTM for their SSVEP-based applications.

Despite the encouraging results, we see that there is still room for improvement. Due to the large number of targets, the neighbouring frequency difference is merely

0.1Hz. Furthermore, the variances of individual differences and the noise in signals also make our task very challenging. However, we foresee that there are a few promising solutions. In future, we plan to explore Generative Adversarial Networks (GANs) [9] on our current dataset to see if this can better handle the small training data limitation. Besides, we plan to continue to enlarge our dataset, so that we can experiment with models with higher capacity or more complex architectures. We plan to study advanced frequency assignment strategies for different targets, e.g. to maximize the frequency difference between physically adjacent targets, as well as to study the effect of different design parameters, such as the number of targets and the length of trials.

References

[1] Min-Hee Ahn and Byoung-Kyong Min. "Applying deep-learning to a top-down SSVEP BMI". In: *Brain-Computer Interface (BCI), 2018 6th International Conference on*. IEEE. 2018, pp. 1–3.

[2] Saba Ajami, Amin Mahnam, and Vahid Abootalebi. "Development of a practical high frequency brain–computer interface based on steady-state visual evoked potentials using a single channel of EEG". In: *Biocybernetics and Biomedical Engineering* 38.1 (2018), pp. 106–114.

[3] Mohamed Attia, Imali Hettiarachchi, Mohammed Hossny, and Saeid Nahavandi. "A time domain classification of steady-state visual evoked potentials using deep recurrent-convolutional neural networks". In: *Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on*. IEEE. 2018, pp. 766–769.

[4] Nik Khadijah Nik Aznan, Stephen Bonner, Jason D Connolly, Noura Al Moubayed, and Toby P Breckon. "On the Classification of SSVEP-Based Dry-EEG Signals via Convolutional Neural Networks". In: *arXiv preprint arXiv:1805.04157* (2018).

[5] Fabrizio Beverina, Giorgio Palmas, Stefano Silvoni, Francesco Piccione, Silvio Giove, et al. "User adaptive BCIs: SSVEP and P300 based interfaces." In: *PsychNology Journal* 1.4 (2003), pp. 331–354.

[6] Anna Chabuda, Piotr Durka, and Jarosław Żygierewicz. "High frequency SSVEP-BCI with hardware stimuli control and phase-synchronized comb filter". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26.2 (2018), pp. 344–352.

[7] Xiaogang Chen, Yijun Wang, Masaki Nakanishi, Tzyy-Ping Jung, and Xiaorong Gao. "Hybrid frequency and phase coding for a high-speed SSVEP-based BCI speller". In: *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE. 2014, pp. 3993–3996.

[8] Chengcheng Han, Guanghua Xu, Jun Xie, Chaoyang Chen, and Sicong Zhang. "Highly Interactive Brain–Computer Interface Based on Flicker-Free Steady-State Motion Visual Evoked Potential". In: *Scientific reports* 8.1 (2018), p. 5835.

[9] Kay Gregor Hartmann, Robin Tibor Schirrmeister, and Tonio Ball. "EEG-GAN: Generative adversarial networks for electroencephalograhic (EEG) brain signals". In: *arXiv preprint arXiv:1806.01875* (2018).

[10] Zhonglin Lin, Changshui Zhang, Wei Wu, and Xiaorong Gao. "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs". In: *IEEE transactions on biomedical engineering* 53.12 (2006), pp. 2610–2614.

[11] ST Morgan, JC Hansen, and SA Hillyard. "Selective attention to stimulus location modulates the steady-state visual evoked potential". In: *Proceedings of the National Academy of Sciences* 93.10 (1996), pp. 4770–4774.

[12] Masaki Nakanishi, Yijun Wang, Xiaogang Chen, Yu-Te Wang, Xiaorong Gao, and Tzyy-Ping Jung. "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis". In: *IEEE Transactions on Biomedical Engineering* 65.1 (2018), pp. 104–112.

[13] Masaki Nakanishi, Yijun Wang, Yu-Te Wang, Yasue Mitsukura, and Tzyy-Ping Jung. "A high-speed brain speller using steady-state visual evoked potentials". In: *International journal of neural systems* 24.06 (2014), p. 1450019.

[14] Jaime Parra, Fernando H Lopes da Silva, Hans Stroink, and Stiliyan Kalitzin. "Is colour modulation an independent factor in human visual photosensitivity?" In: *Brain* 130.6 (2007), pp. 1679–1689.

[15] Ivan Volosyak, Hubert Cecotti, and Axel Gräser. "Steady-state visual evoked potential response-impact of the time segment length". In: *Proc. on the 7th international Conference on Biomedical Engineering BioMed2010, Innsbruck, Austria, February*. 2010, pp. 17–19.

[16] Nicholas R Waytowich et al. "Compact Convolutional Neural Networks for Classification of Asynchronous Steady-state Visual Evoked Potentials". In: *arXiv preprint arXiv:1803.04566* (2018).

[17] Dong-Ok Won, Han-Jeong Hwang, Sven Dähne, Klaus-Robert Müller, and Seong-Whan Lee. "Effect of higher frequency on the classification of steady-state visual evoked potentials". In: *Journal of neural engineering* 13.1 (2015), p. 016014.

[18] Wang Yijun, Wang Ruiping, Gao Xiaorong, and Gao Shangkai. "Brain-computer interface based on the high-frequency steady-state visual evoked potential". In: *Proceedings of the First International Conference on Neural Interface and Control*. IEEE. 2005, pp. 37–39.

[19] Danhua Zhu, Jordi Bieger, Gary Garcia Molina, and Ronald M Aarts. "A survey of stimulation methods used in SSVEP-based BCIs". In: *Computational intelligence and neuroscience* 2010 (2010), p. 1.