Classification of Frequency and Phase Encoded Steady State Visual Evoked Potentials for Brain Computer Interface Speller Applications using Convolutional Neural Networks

JOSHUA JAMES PODMORE

Classification of Frequency and Phase Encoded Steady State Visual Evoked Potentials for Brain Computer Interface Speller Applications using Convolutional Neural Networks

Joshua James Podmore

Abstract

Over the past decade there have been substantial improvements in vision based Brain-Computer Interface (BCI) spellers for quadriplegic patient populations. This thesis contains a review of the numerous bio-signals available to BCI researchers, as well as a brief chronology of foremost decoding methodologies used to date. Recent advances in classification accuracy and information transfer rate can be primarily attributed to time consuming patient specific parameter optimization procedures. The aim of the current study was to develop analysis software with potential 'plug-in-and-play' functionality. To this end, convolutional neural networks, presently established as state of the art analytical techniques for image processing, were utilized. The thesis herein defines deep convolutional neural network architecture for the offline classification of phase and frequency encoded SSVEP bio-signals. Networks were trained using an extensive 35 participant open source Electroencephalographic (EEG) benchmark dataset (Department of Bio-medical Engineering, Tsinghua University, Beijing). Average classification accuracies of 82.24% and information transfer rates of 22.22 bpm were achieved on a BCI naïve participant dataset for a 40 target alphanumeric display, in absence of any patient specific parameter optimization.

Classification of Frequency and Phase Encoded Steady State Visual Evoked Potentials for Brain Computer Interface Speller Applications using Convolutional Neural Networks

Number of Volumes:	1
Name:	Joshua James Podmore
Qualification:	Masters by Research
Department:	Psychology and, School of Engineering and Computing Sciences
Institution:	Durham University
Year:	2017

Table of Contents

List of Abbreviations

BCI:	Brain Computer Interface
AoC:	Accuracy of Classification
AP:	Average Precision
ITR:	Information Transfer Rate
GUI:	Graphic User Interface
SMR:	Sensorimotor Rhythms
ERP:	Event Related Potential
EEG:	Electroencephalography
SSVEP:	Steady State Visual Evoked Potential
PDSA:	Power Density Spectral Analysis
CCA:	Canonical Correlation Analysis
MCCA:	Multiway Canonical Correlation Analysis
PCA:	Principal Component Analysis
KNN:	K-Nearest Neighbours
SVM:	Support Vector Machines
MP:	Max-Pooling
CNN:	Convolutional Neural Networks

Acknowledgements

I would like to thank Dr Jason Connolly (Durham University Psychology Department) and Dr Toby Breckon (Durham University School of Engineering and Computing Sciences) for their support throughout my Masters by Research and in the preparation of the thesis herein.

Aim

The aim of the present thesis is to summarise and then extend upon findings across a vast array of non-invasive Brain-Computer Interface (BCI) research. Here I will scientifically examine the potential of deep learning neural networks to decode Steady State Visual Evoked Potentials (SSVEP) bio-signals, as harnessed via Electroencephalogram (EEG) for BCI communication applications. The majority of BCI spelling systems involve assigning alphanumeric characters to unique stimulus classes which produce reliable changes in bio-signal time-series data. In order to reduce misclassification error researchers often optimize the stimulus and EEG configurations on the single-subject level. The central premise under investigation is whether or not there exists sufficient similarity between 40 classes of SSVEP EEG data **across** trials and critically across **participants** to produce classification accuracy values >=70%. This is a crucial step in the development of a 'plug-in-and-play' BCI speller for rehabilitation and assistive applications. This refers to a BCI which is capable of performing at industry level standards (>=70%) **without**; extensive pre-interfacing data collection, intensive training periods or laborious optimization processes. I will discuss in successive order the alternative bio-signals available for BCI applications, ultimately validating our selection of the SSVEP due to: 1) signal stability, 2) low latency/ refractory period, 3) minimal incidence of BCI illiteracy, and 4) primarily bottom up activation pathway. I will justify the implementation of a matrix style graphical user interface based on: 1) short training period, 2) low user operational fatigue, and 3) potential for maximal information transfer rates. I will finally substantiate the claim that convolutional neural networks (CNNs) possess qualities which surpass the techniques currently employed in terms of: 1) noise resilience, 2) generalisability across participants, 3) user specific data integration, and 4) potential for rapid deployment.

BCI Definition and Origins

BCIs are integrated hardware and software ensembles that harness bio-signals to power assistive devices (Vidal, 1973, Wolpaw, Birbaumer, McFarland, Pfurtscheller & Vaughan, 2002, Blankertz et al, 2006). These differ from Neural Computer Interfaces (NCI) which acquire signals from neurones peripheral to the brain, utilizing such techniques as electromyographic recording from afferent effectors (motor neurones within intact muscle groups) (Mackenzie & Ashtiani, 2011, Vasiljevas, Turčinas, & Damaševičius, 2014). Over the past decade a number of significant developments have fostered a surge in BCI performance. Firstly, bio-signal acquisition devices such as eye trackers, electroencephalographs (EEG) and electromyograms (EMGs) have been introduced to the consumer market. This has improved online trouble-shooting documentation, the quantity of open source datasets and provided manufacturers with greater incentives to optimize their products across platforms, enhancing compatibility. Advances in signal processing algorithms for estimating optimal acquisition device configurations, pre-processing methods for artefacts removal and online device calibration techniques have led to the collection of cleaner, higher fidelity data. Finally, progress in long-life, high capacity, lightweight battery packs have enabled researchers to test BCI during mobile tasks, increasing the practicality of wearable BCI technology.

Synchronous vs Asynchronous/ Reactive vs Active BCI differentiation

There is a multifarious array of BCI formats which researchers must consider in relation to BCI end-users. Synchronous BCIs refer to systems in which the stimuli code is temporally locked. The discrete timing sequences of stimuli events are logged and serve as triggers to begin data acquisition, or operate as markers for the partitioning of data. Subsequently, analysis of the bio-signal time-series data is then performed. Finally, predictions are then made estimating the operation the user intends to perform. The BCI then cycles through these operations, all of which are dependent the initial presentation of a pre-specified time-locked stimulus. (Diez, Mut, Perona & Leber, 2011) These forms of BCI are also often found, however not exclusively, alongside so-called reactive BCI formats (Zander, Kothe, Jatzev & Gaertner, 2010). These systems involve attending external stimuli, such as a visual array, in order to generate reliable bio-signals for

controller systems to interact with afferent effectors. In contrast, asynchronous systems monitor user bio-signals continuously (Krumpe, Walter, Rosentiel & Spüler, 2016). Initiation of a cue is locked to a specific event. Assistive eye-tracking devices, more broadly categorized as Human-Computer Interfaces (HCIs), are often used in these contexts, as user fixation position can be monitored continuously with a high degree of resolution. Events such as the prolonged fixation of a visual target (dwell time), or a systematic pattern of winking and blinking can be programmed to operate different HCI functions. The latter is more problematic in BCI applications due to the inherently noisy characteristics of brain-based bio-signals. In order to maintain data integrity pre-processing must be carried out iteratively via repeated manual calibration, or performed automatically by script-based algorithms. Active BCI formats are frequently paired with these asynchronous systems. These utilize internally generated signals in absence of external stimuli, such as imagined movements to power assistive devices. These methods are therefore theoretically capable of providing any individual, irrespective of medical condition, some level of BCI control. The task requirements of active BCIs to date are however far more cognitively taxing, meaning substantial training is necessary for smooth operation. The current thesis will explore the development of an analysis pipeline for a synchronous reactive style BCI, whereby time-locked SSVEP stimuli provide external stimulation to trigger bottom up brain-based bio-signals.

BCI Configuration

Numerous bio-signals have been identified and tested over the past decades allowing developers to target BCIs to specific patient populations. Signal selection will differ depending on: 1) patient condition, 2) the task in which the individual is being assisted, and 3) the patient's willingness for invasive surgery. Irrespective of the BCI format and target end user population, all BCIs share a network of fundamental components. 1st: a bio-signal sensor. 2nd: signal acquisition hardware. 3rd: signal amplification hardware. 4th: signal pre-processing software. 5th: feature extraction. 6th: device interface. 7th: effector hardware. The core aim of implementing this chain of code-based and hardware components is to 'close the loop' between the human user and their environment. In other words the goal of BCI is to provide real-time feedback between the user and their environment by restoring critical processes such as mobility and communication.

Target Population

The target population for the BCI research conducted herein is quadriplegic patients with anarthria (inability to vocalize) and dexterous vision in at least one eye. There are many aetiological pathways to this condition, the majority of which involve damage to either the brainstem or ventral pons. The most common pathways to injuring these cerebral areas are traumatic cerebral events such as ischemia (principally obstruction of the basilar arteries) (Patterson & Grabois, 1986) or haemorrhage (Kompanje, 2007). Less frequently cited aetiologies include progressive motor neurone diseases, such as amyotrophic lateral sclerosis, multiple sclerosis, polyneuritis with Guillain-Barre Syndrome, central pontine myelinolysis, cerebrospinal injury (thoracic level) and tumour formation (Smith & Delargy, 2005).

Quadriplegic patients are referred to as having "locked-in syndrome" (LIS) when in a state of wakefulness in concert with evidence of awareness inside a non-functioning body (Reigada, Mendes, Paiva, Tavares & Goncalves, 2014). There exist many classification systems for locked in syndrome, ranging from classical (some residual muscle dexterity), incomplete (residual vertical eye movements and blinking retained) and complete locked in syndrome (absence of all voluntary muscular control). The degenerative nature of progressive diseases, such as multiple sclerosis, lead to gradual reductions in the dexterity of muscle groups which initially appear unaffected by the condition. An individual could feasibly receive each of the above mentioned classifications throughout the duration of their life. The BCI research conducted thus far has focused primarily on the restoring functions which individuals with severe forms of paralysis can no longer perform independently. These include the restoration of mobility and communication. This thesis contains references to both these variants of BCI, however as the research conducted here involves the development of a BCI spelling device, the literature discussed mainly explores communication-based BCIs.

During the initial stages of progressive motor diseases, vocalized communication may persist with a lower level of control, meaning communication can still be established with the use of vocal sounds. Emotive context may also be ascertainable from characteristic inflections in the delivery of the vocalized sounds. Eventually, communication may be dependent on the use of a picture or alphabet board containing images or characters. Users can use these types of apparatus to spell out their thoughts or current internal state to an observant conversational partner or clinical assistant. In the later stages of motor degeneration, use of alphabet boards and potentially even simplistic yes/ no answering system based on blinks can become unviable (Smith & Delargy, 2005). Many LIS patients cannot use conventional devices made for patients with severe motor disability, because these require reliable control over at least one muscle group (Khanna, Verma & Richard, 2011). Previous studies into the mortality rate of patients suffering LIS vary depending on the aetiological basis for the condition (Patterson & Grabois, 1986). Associated vascular insults, specifically in relation to damage of the pons and brainstem demonstrate the highest risk of death and lowest probability of functional recovery. The most recent literature review reveals that despite modern therapeutic interventions only between 40-60% of patients survive the first four months after onset (Casonova, Lazzari, Lotta & Mazzuchi, 2003). From these studies it is evident that BCI researchers must work tirelessly to keep open the window of communication between patients and patient families for as long as possible. The target population for which this thesis is aimed at excludes those with severe incomplete or complete locked syndrome and is primarily aimed at classical locked in syndrome patients with retained dexterous eye movements.

Functionality

Surveys conducted on potential and current users of BCIs reveal functionality is the most important aspect of an assistive device. This term constitutes; device usability, usage load (fatigue induction) and system set-up/ calibration time (Lopes, 2001). In the past two decades huge leaps have been made in the development of: 1) signal noise reduction (bio-signal sensor advancements), 2) user interface design (graphic user interface layout) and 3) system portability. In the last decade continual improvements in performance have been demonstrated at lower computational cost which is pushing BCIs to the above mentioned end-goal of closed-loop real-time user control. The steady state visual evoked potential (SSVEP)-based convolutional neural networks (CNNs) defined herein possesses computational qualities which can achieve many of the critical characteristics a functional BCI must present. Primarily in relation to short system deployment duration, minimal training time and reduced mental workload during usage.

Performance Metrics

In order to evaluate BCIs across different use cases, stimuli designs and analyses methods, a number of performance metrics have emerged. Accuracy of Classification (AoC) is determined by calculating the proportion of correct predictions made in relation the total number of predictions. Often, AoC is expressed in an inverted format and referred to as error. In contrast this involves measuring the number of incorrect predictions, in relation to the total number of predictions made. Average Precision (AP) is determined by finding the proportion between the class predicted most frequently (mode) and the total number of predictions made. If an analytical method continuously predicts the same class, the method would be highly precise, this method could however continuously predict the same incorrect class. In this instance the analytical method demonstrates high precision and low accuracy. AoC and AP are primarily used by researchers to optimize the parameters of their BCI, for example these metrics would inform them of an uneven distribution of performance across classes. It is crucial to maintain consistent decode performance across classes in order to prevent interruptions in user workflow during real-time operation. The mean AoC (mAoC) and mean AP (MaP) metrics are used primarily to demonstrate BCI performance to external research groups and describes network performance across all classes tested.

Information transfer rate (ITR) refers to the speed at which users can issue commands, typically expressed in bits per minute. This metric was originally developed to measure the practical output of telecommunication devices. Wolpaw & Ramoser (1998) were the first researchers to introduce

this metric to evaluate the performance of BCIs. The ITR equation includes the following variables: 1) total data acquisition time, 2) number of executable operations (referred to as targets or classes) per trial and 3) the mean error in class prediction. ITR has become a common metric in the field of BCI as it demonstrates the functional capabilities of a system. In other words, ITR demonstrates the ecological validity of a BCI by defining performance over time. In reference to the equation below, N refers to the number of potential targets for classification, P denotes the probability that the target will be classified accurately and T indicates the duration of the data capture period.

$B = \log_2 (N) + P*\log_2 (P) + (1 - P)*\log_2 ((1-P)/ (N-1))$

Bits per minute= B*60/T

Infrequent mentions of analogous terminology such 'letters per minute' are present in BCI speller literature, as well as alternative methods of evaluating communication rate such as the 'effective transfer rate'. However as AoC, mAoC, AP, MaP and ITR at the time of writing have been established as industry standards, these metrics will be used exclusively.

Electroencephalography

The application of Electroencephalography (EEG) is at the forefront of BCI development. These systems are composed of electrodes designed to rest against the scalp, detecting changes in micro-voltage across the lateral surface (Olejniczak, 2006). Neurones undergoing an action potential experience a shift in charged particles through the axon. This generates a primary electrical current and subsequent magnetic field. The harmonious propagation of action potentials leads to the generation of higher power magnetic fields. In contrast, magneto-encephalographs (MEGs) are capable of detecting primary currents due to the use of superconducting quantum interference devices (SQUIDs) (Stufflebeam, Tanaka & Ahlfors, 2011). Non-invasive MEG systems are utilized primarily for pre-surgical mapping (Solomon, Boe & Bardouille, 2015, Ahmed & Rutka, 2016, Alkhalili, Niranjan & Engh, 2016), post-surgical evaluation (Knowlton, 2008a, Knowlton et al, 2008b) and functional connectivity research (Spencer, Niznikiewicz, Shenton & McCarley, 2008). Despite MEG's higher resolution on the temporal axis, as well as increased spatial localization capabilities, these systems are extremely costly. Maintenance of a low magnetic interference environment (Faraday cage), substantial power demands and lack of mobility reduce the applicability of such systems in research.

EEG differs from MEG by detecting so-called 'secondary' currents. Post-excitation a secondary magnetic field is generated which offsets the difference in charge across the neurone. The interaction of these magnetic fields leads to changes in the electrical potentials (micro-voltage) across the scalp. Behavioural, emotional and cognitive phenomena have been linked to changes in EEG signal amplitude and changes in the prevalence of frequency spectra (for review see, Mauss & Robinson, 2009). Traditionally, electrodes are housed in a flexible cap arranged according to the standardized 10-20 international EEG layout (refer to Figure 1). To minimise the impedance level for each sensor a thin medium of conductive gel is injected between the sensor and the skull, these are categorized as wet-systems. Impedance (Ω) defines the degree to which the flow of an alternating current is inhibited as it passes across a conductive surface. This differs from resistance, which defines a similar phenomenon in relation to direct current. Higher impedance levels lead to lower signal amplitudes which have been shown to negatively impact EEG: signal coherence (Tautan et al, 2014) and signal-to-noise-ratios (Kappenmann & Luck, 2010, Chi et al, 2012).

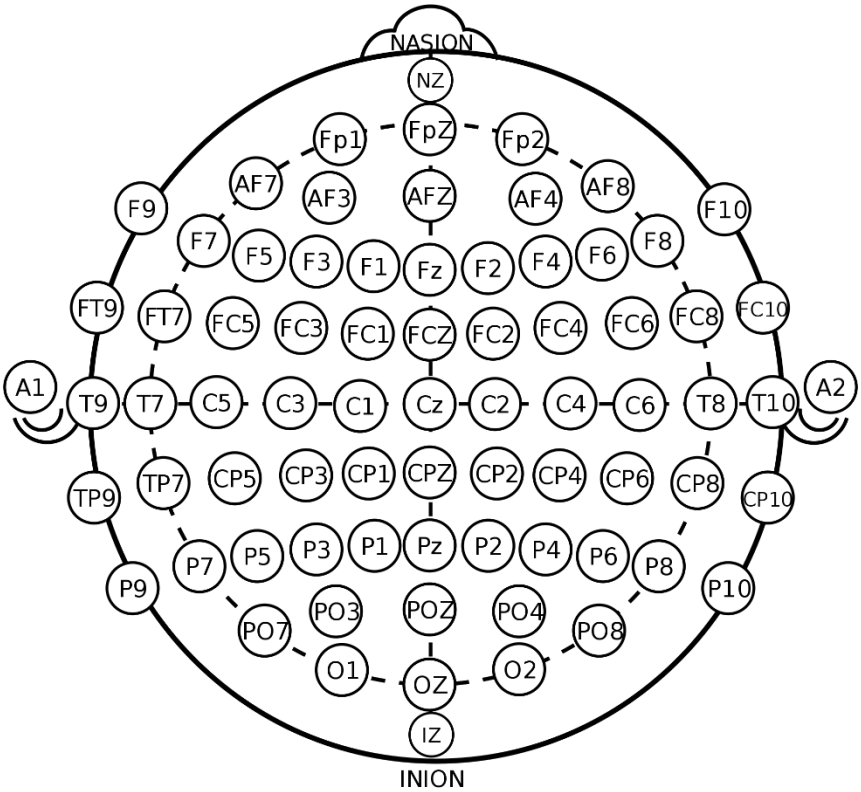


Figure 1. Diagram depicting the standardized layout of EEG electrodes across the lateral surface of the skull for the International 10-20 system.

The process of conductive gel application to each wet-EEG electrode can be time consuming as all electrodes must be adequately coated to minimize impedances. Coupling refers to an instance in which gel from neighbouring electrodes overlaps. The readings from the two overlapping electrodes will blend and invalidate signal localization assumptions, therefore necessitating a complete repetition of the procedure. This costly step-up time can be avoided with recent technological developments in materials science. Dry-EEG setups are now available with sensors coated in silver alloys to maintain sub-5kOhms impedances while keeping electrical interference to a minimum. The sensors for dry-EEG are often constructed to maximise contact against the scalp, with some featuring 'feet' to penetrate through user hair. Other methods utilize foam and spring-loaded housings, as well as intricate tightening mechanisms to accommodate for the unique proportions of the head. Primarily, dry-EEG benefits from rapid deployment, high portability and increased user comfort (Lopez-Gordo, Sanchez-Morillo & Valle, 2014, Mathewson, Harrison & Kizuk, 2017). Conversely, agitation of sensors in these applications result in more exaggerated changes in signal stability.

Development of EEG analysis techniques must consider the non-stationary qualities of the EEG signal. Fundamentally, these signals are non-linear with high dimensionality. Therefore, even when experimental protocols are repeated, signals derived from exposure to the same stimuli will differ over time (Vaid, Singh & Kaur, 2015). Corroboration of these EEG signal characteristics was provided after researchers produced phase portraits of EEG output via the use of reconstructive algorithms (Mayer-Kress, Gotfried & Layne, 1987). Authors identified that EEG time series evolve along a chaotic trajectory where minute changes in the initial conditions of data collection such as; skin conductivity, electrical interference and user concentration can result in divergent outcomes (Layne, Mayer-Kress & Holzfuss, 1986, Blanco, Garcia, Quiroga, Romanelli & Rosso, 1995).

When compared with other techniques such as functional magnetic resonance imaging, EEG kits suffer from poor spatial resolution across each Cartesian axis, with signal acquisition penetrating on average 1-2cm. Moreover, systems are typically restricted to the detection of neural wave patterns between 7 and 70Hz (Nunez & Srinivasan, 2006). Both aforementioned issues are the result of interference from the skull and cerebrospinal fluid during signal acquisition (Ng Logothetis & Kayser, 2013). Electrocorticography (Ecog) circumvents this issue by placing 'depth sensors' typically in a uniform grid format across the cerebral cortex enabling higher frequency sampling of up to 200Hz (Pfurtscheller & Graimann, 2003, Leuthardt et al, 2004). The technique however requires invasive surgical procedures to enable the implantation of sensors (Alcaraz & Manninen, 2017, Panov et al, 2017). Birbaumer (2006) revealed just one complete locked in syndrome patient out of sample group n= 17 indicated their willingness for sub-dural implantation. Additionally, the surgical trauma, clinician costs, hygiene upkeep and patient discomfort reduce the feasibility of implementing such systems on a large scale. Critically, EEG systems lack high spatial resolution; however the temporal sensitivity to changes in neuronal pattern firing is exceptional (sub-millisecond). BCI spellers employing tasks featuring time-locked decision making are therefore highly compatible with this set-up. Moreover, systems with high temporal resolution are arguably more conducive to the development of effective BCI spellers, as such devices are inherently more adept at contributing higher ITR metrics.

EEG-based BCIs

Numerous bio-signals are present within EEG output such as: P300 waves, sensorimotor rhythms and Steady State Visual Evoked Potentials (SSVEPs). Each of the aforementioned signals has been successfully applied in BCI speller device development. In the following sections I will briefly discuss the different approaches research teams across the globe have implemented in the development of assistive devices for individuals with severe forms of paralysis. Primarily this discussion will focus on the different bio-signals utilized and will examine the: practicality, performance and technical obstacles related to each approach. Ultimately, this investigation aims to validate the claim that EEG and more specifically, SSVEP-based EEG, is currently positioned as the optimal method of acquiring bio-signals for BCI applications.

Sensorimotor based-BCIs

Motor Imagery (MI) based BCIs typically harness signals known as sensorimotor rhythms (SMR) using EEG recording hardware. Imagined movements of effector muscles lead to predictable changes in neural firing patterns across corresponding motor regions. Such movements often make use of dominant effector movements to ensure maximum amplification of the signal (Müller-Gerking, Pfurtscheller & Da Silva, 1999, Ang, 2012). Imagined or intended movement of e.g. dominant right hand to perform a power grip has been shown to highly correlate with a decrease in the amplitude of low frequency EEG components such as mu (range: 9-11Hz) and beta (range: 13-18Hz) rhythms in corresponding regions of primary somatosensory cortex (Pfurtscheller, 1999, Yuan et al, 2010). Typically, electrode locations: Cz, C3 and C4 (refer to Figure 1) are analysed due to the high levels of SMR propagation at these locations. This phenomenon is known as event related desynchronization (Pfurtscheller & Aranibar, 1979). Due to the top-down initiation of SMRs, successfully eliciting this signal can provide BCI applications for individuals with severe motor disabilities in combination with conditions impairing visual acuity and dexterity. This signal is therefore extremely flexible as it does not require the dexterous movement of any user muscle groups. Arguably BCI systems utilizing these bio-signals are available to a larger proportion of the most severely paralyzed patient populations, as compared with eye-tracking systems which require dexterous control of the ocular muscles.

Due to the somatotopic organisation of sensorimotor rhythms it provides the potential for BCI based applications to utilize many of the human body's degrees of freedom. Theoretically, this could enable users to operate BCI devices across multiple dimensions. However, successful fine control of a SMR based BCI under continuous translation with real-world patient users in a 3D context has been documented only a handful of times (Doud, Lucas & Piansky, 2011, LaFleur et al, 2013). Power-based motor imagery of each individual effector: right arm and leg, left arm and leg, are reliably decoded with minimal error. However, inter-effector MI differentiation is problematic due to the restrictions of reduced spatial separation. For example right arm movement is easily distinguished from left arm movement. In contrast, right arm movement leftward vs right arm movement rightward is far more complex. This reduces the navigability of any MI based graphic user interface (GUI) for speller applications to around just 4-8 different command selections. Comparable SSVEP and P300 based systems have recently achieved up-to 40 and 72 possible command selections respectively (Chen et al, 2015, Townsend & Platsko, 2016), making these systems vastly more functional than SMR-based counterparts.

Moreover, there are considerable individual differences in the development of neural structures associated with the propagation of SMR signals. Those lacking the necessary structural integrity and white matter myelination (Halder et al, 2013) are termed 'BCI illiterate' (Kübler & Birbaumer, 2007). Moreover, inter-subject variability in: perception of self-worth (Burde & Blankertz, 2006), mental workload capacity (Hammer et al, 2012) and degree of past dominant hand utilization (Randolph, Jackson & Karmakar2010) have also been shown to correlate significantly with SMR signal power and ultimately user performance metrics. These same factors can also influence the quality of SSVEP and P300 waveform prevalence in EEG time-series; however BCI systems utilizing these bio-signals simply require users to fixate an external stimulus, as opposed to generating a dense mental image of limb movement. Therefore, despite the fact SMR-based BCI systems possess the potential for applications with high degrees of freedom, analytical methods for differentiating between the actuation of these different degrees is restricted. Effectively users would be constrained to sequential GUI command layouts. The distinct disadvantage of such a system is the need to illicit multiple bio-signals successively in order to execute just one target selection. When such systems are paired with SMR-based BCIs, this therefore necessitates multiple user imagined motor movements. Moreover, the impact of misclassification is compounded as the number of operations required for a selection increases.

Additionally, SMR-based BCI spellers require a prolonged training period before reliable classification results are acquired (Ahn & Jun, 2015). Previously researchers gathered substantial amounts of participant specific data in order to fine tune data pre-processing and analysis parameters. Target patients endured many hours of data collection in order to develop low noise reference waveforms with which to compare online EEG data output. For the lowest signal to

noise ratio users must expend vast amounts of energy concentrating on imagined movements which consistently share the same qualities: duration, power, context, handedness and end goal. This process is fatiguing due to the high mental workload and large number of repetitions required. To mitigate patient exhaustion during this data collection period it is possible to collect a small user dataset and then apply generative algorithms to produce artificial data. This is then employed to fine tune classification analysis parameters offline prior to online usage. However, the accuracy of reference signal representations can degrade overtime due to the aforementioned non-stationary properties of the EEG signal. Therefore, resulting data augmentations only provide relevant reference signal templates for short periods of time, requiring consistent re-collection of subject baselines. Researchers have also suggested BCI training prior to the on-set of a complete-locked-in state can also improve online performance in the long-term (Kübler & Birbaumer, 2008). However the mental concentration required during user-device-interfacing cannot be avoided. This therefore makes such a BCI paradigm inaccessible to many individuals with severe quadriplegia as chronic fatigue is often a co-morbid symptom.

In comparable SSVEP and P300-based systems, a short period at the start of an experimental session is used to collect reference data. Users are then tasked with fixating and attending stimuli. In contrast to SMR-based BCI, the provision of simplistic instructions for an intuitive task makes these alternative bio-signals far more accessible for both research and real-world user applications. Crucially, research has also shown that the LIS patient performance on classical motor imagery assessments such as the hand laterality task (Conson et al, 2008) is significantly lower than healthy participant counterparts. Further, research revealed the prevalence of event-related desynchronization in an SMR-based motor imagery assessments was significantly reduced in LIS patient groups (Hill et al, 2006). Importantly, this effect increased in power according to the severity of patient condition. Authors attributed this to a number of potential factors, namely; reduced utilization of relevant motor pathways, cognitive impairments and temporary attentional lapses during the task. Irrespective of the basis for these results, ultimately the findings indicate that the laborious training process required before the implementation of motor imagery based BCI may be at best substantially more problematic in LIS patient populations. Therefore, regardless of the theoretical capabilities of SMR based systems, the feasibility of implementing such BCIs in a real-world setting is at present extremely low. The plethora of justifications provided indicate why the development of a SMR-based multiple step BCI speller was avoided.

Efforts have been made to improve the performance of SMR-based BCI communication devices by introducing novel graphical stimulus designs. The Hex-o-Spell, developed at the Technical University of Berlin, represents arguably the most efficient form of sequential command GUI to date (Blankertz et al, 2006). A display containing six groups of letters, bunched in alphabetical order (Group 1= A-F, Group 2= G-L etc.) is presented on a computer monitor. First users identify which of the 6 groups their target letter resides in. Navigation between targets in the visual array is achieved via imagined movement of either the right (clockwise movement) or left arm (anti-clockwise movement). User intended direction is inferred from bio-signal analysis and the GUI changes accordingly. The 6 locations previously occupied by the 6 alphanumeral groups are then populated with individual characters from the group selected (Muller & Blankertz, 2006). In other words, to select the letter 'A' user must 1st attend Group 1. Then they would have to navigate to the solitary letter A at the 12 o'clock position. After this, a simple two class screen (yes and no box) is presented asking the user if the letter selected is their intended choice. The Hex-o-Spell design is extremely intuitive, due to the alphabetical grouping of targets, minimising user training time. Crucially, this format normalizes the selection time for all targets in the array, however at least 4 seconds are required to derive a reliable prediction from the SMR EEG data. With a minimum of 3 steps per character, maximal performance for such a system peaks at around 7.5 characters per minute.

A truly functional BCI speller should aim at making user commands as efficient as possible; therefore the development of one-step GUIs to increase ITRs is imperative. The most commonly used one-step GUI format is known as a matrix speller, specifically in relation to vision based BCIs. This involves having all targets (alphanumeric characters) presented at once, selecting the intended target after a single data collection period. This is problematic, irrespective of bio-signal

due the issue of target separation. An increase in the number of targets in an array reduces the distance between neighbouring targets. In other words, as the GUI becomes more densely populated the correlation between adjacent targets increases, therefore heightening the risk of misclassification. An increase in misclassification negatively influences the ITR as this necessitates the users to execute more commands to correct the mistakes of the bio-signal decoding algorithms in place.

Eye Tracker based-BCIs

Eye tracking based methods integrating matrix style GUIs have been successful in many applications including BCI wheelchair control, computer cursor control and spelling paradigms, (for review see, Majaranta & Raijha, 2002). Such devices are typically optical (camera based) and calculate the difference in orbital rotation of the eye to estimate user fixation point. This paradigm typically involves users viewing a display with targets spread across a computer monitor. Command execution is conventionally determined by a specific dwell time threshold (period of fixation upon a target). Patients with typical eye movement dexterity would be well suited to such paradigms. State-of-the-art eye tracking systems are capable of reliably distinguishing fixation loci differing by just 0.002^{o} of visual angle. The spatial resolution of this technique therefore compliments a matrix style selection methodology, enabling one step selections from a densely populated target array. Successful attempts at creating ultra-low cost alternatives with high mobility and low computational demand have also been demonstrated, meaning these devices are economically viable (Abbot & Faisal, 2012).

Despite these inherent advantages such devices are impractical for many individuals with progressive motor degeneration conditions and those with severe cerebral motor area damage due to poor ocular motor control. When considering the target population for communicative BCIs eye tracking systems are sub-optimal as the ability to fixate steadily is not often retained by those suffering from congenital or progressively induced quadriplegia (Smith, 2005). It is also common for eye movements to be restricted to the horizontal plane; therefore any target array would have to be arranged in a thin column of characters (Riccio, Mattia, Simione, Olivetti & Cincotti, 2012). Moreover, as fixation density in the human eye reduces as a function of distance from the resting ocular centre point, a feature often exaggerated in quadriplegic patients, it makes such a GUI unfeasible. Eye-Tracker systems are inherently gaze-dependent and therefore not functionally relevant to a large portion of the BCI patient population. It is therefore necessary to explore brain-based bio-signals to power assistive BCI devices. SMR-based systems are not restricted by patient ocular muscle control and are therefore positioned as far more applicable to patients with complete locked-in-syndrome. Further, P300 and SSVEP-based systems are arguably semi-dependent on user gaze, however to a far lower degree than comparable eye-tracking based methods.

Systems incorporating eye-tracking as a peripheral compensation for error however, do seem to have some functional applicability. These operate by locating regional points of fixation, as opposed to the exact spatial location. Lim et al (2015) implemented a hybrid webcam eye-tracker and SSVEP BCI speller paradigm. When compared to a standard SSVEP set-up the hybrid BCI revealed, over the course of a 68 character trial period, an average of 39 typographical errors were avoided. Nevertheless, recent research has directly compared SSVEP based BCI spellers and eye-tracking based spellers on low density visual arrays (4 targets) (Suefusa, 2017). Critically, authors reported SSVEP based methods had significantly higher information transfer rates and mean accuracies for smaller visual targets (40 mm^2 and $20mm^2$). Matrix format arrays currently utilize target stimulus dimensions of around $200mm^2$. However the aforementioned findings illustrate that there is substantively more potential for increasing array density using SSVEP based BCI formats as compared to eye-tracking interfaces. It is for these reasons that the current thesis did not harness eye movement bio-signals for BCI speller applications.

P300

The seminal work of Sutton (1965) revealed that when human users attend to stimuli with a unique time-locked feature, a positive deflection with a latency of around 300ms is exhibited over the

central and parietal regions of the scalp during EEG recording. Farwell and Donchin (1988) exploited this characteristic waveform for BCI speller applications via an oddball paradigm. This involves displaying an array of target stimuli, all of which are programmed to augment at unique time locked points during the data acquisition period. In order for the user to demonstrate their intention to select a specific target they are instructed to attend a target in the array during the data capture period. The EEG output will reveal a distinct positive deflection around 300ms post stimulus augmentation. Analysis involves subtracting the latency of the P300 deflection (300ms) from the EEG time series. This enables researchers to estimate the point in time in which the stimulus augmentation occurred. This unit of time is then compared against the unique time locked values assigned to each target in the array and a prediction for the target intended for selection is deduced. There are numerous forms of stimulus augmentation; colour, size, motion (e.g. rotation) and texture to name a few.

Historically, P300 event-related potentials have been defined as gaze independent bio-phenomena and have therefore been classified as more applicable to locked-in patient populations in relation to eye-tracking or SSVEP-based BCI methods. This is due to the perceived top-down nature of the P300 bio-signal. In other words, the propagation of a P300 is contingent on the user making a conscious decision to attend a specific target in the array. As gaze and attention are not co-dependent, numerous authors assert that covertly attending to the target stimulus is sufficient to produce a viable ERP. Foreseeably, this paradigm therefore possesses a large scope of applicability across many levels of motor disorder severity, specifically in relation to users with poor or absent ocular control. As a result a multitude of studies have utilized these Event Related Potentials (ERPs) for the purposes of BCI spelling (Blankertz, Lemm, Treder, Haufe & Müller, 2011, Yeom, Fazli, Müller & Lee, 2014).

However, the pivotal article from Brunner et al (2010) revealed that P300 propagation is partially dependent on gaze. A typical 36 character array (6x6 matrix) was presented to healthy participants utilizing the popular row/ column paradigm. Each row and column was assigned individual time locked values which control the onset of the visual target augmentation. This results in all targets possessing a unique combination of row and column augmentation onset values. In one condition participants were free to fixate the intended letter for selection and reached and average AoC of 100% using step-wise regression analysis. The second condition required participants to fixate centrally (monitored via eye tracking systems) and were informed to shift their attention to characters intended for selection. Critically, fixation of the central cross was continuously monitored using eye-tracker systems. The latter condition resulted in random classification accuracies for over half the participants involved, also performance decreased as the distance between the fixation cross and the target letter increased. This therefore suggests that the P300 should be re-classified as semi-gaze dependent at best and does not have the potential applicability cited in many BCI speller review articles circulating presently. Critically, these results have been corroborated in additional studies (Treder & Blankertz, 2010).

Additional, yet pertinent issues associated with a P300 based BCI paradigms are related to signal degradation. Repeated exposure to the same stimulus augmentation reduces the amplitude of P300 wave deflection. Therefore, periods of usage must be rationed according to the rate of decay expressed for each user and thus the functionality of such a system drops drastically (Gonsalvez & Polich, 2002). These same obstacles are present in the implementation of SMR-based BCI, which often demand prolonged concentration in order to generate viable motor imagery neural signals. In contrast, SSVEP-based signals do not degrade in amplitude over-time; therefore these bio-signals are arguably better suited to long-term use cases. Some researchers have attempted to remedy this methodological issue in P300-based contexts by assigning alternative stimulus augmentation methods for different targets. The type of augmentation is also randomised over data capture periods to ensure that frequently used letters do not suffer from signal degradation to a greater extent than less frequently used letters. Efforts have also been made to optimise the stimulus presentation format of P300 based BCIs using ‘codebooks’. These are essentially binary representations which indicate distinct points in time at which stimulus augmentation occurs throughout the trial period. Essentially, instead of using one P300 event, multiple P300s are induced across the data acquisition period. The EEG output is then matched to the time-locked

codebook pattern assigned to one of the targets in the display. The implementation of such methods does significantly decrease misclassification error, however this comes at a substantial computational cost.

Impressive ITRs have been achieved utilizing the P300 in a BCI speller context. Townsend (2016), managed to achieve real-online BCI speller performance between 96 and 120 bpm in group of 9 healthy participants. These results were achieved by implementing a stimulus presentation paradigm based on user performance constraints in both the spatial and temporal domain. Optimal oddball temporal parameters were attuned to participant performance online, reducing the total data capture period. Additionally, sophisticated spatial positioning of individual oddball presentations were developed to overcome the issue of neighbouring target misclassification. Traditional methods of stimulus augmentation were replaced with newly corroborated methods which introduce human facial images as the oddball stimulus targets. This amplified the commonly low signal-to-noise ratios inherent to the P300 bio-signal. Authors calculate that the theoretical upper-bounds of this methodology could reach as high as 258 bits per minute and the online results collected herein represent arguably the highest event related potential based BCI ITR to date.

In contrast, comparative SSVEP based bio-signals have achieved significantly higher ITRs in real-time experiments (>180bpm) (Chen et al, 2014, Chen et al, 2015). This is due to the lower latency of the SSVEP bio-signal (140ms) as compared to the P300 waveform (300ms), therefore data capture periods per alphanumeric character can be minimized to a greater extent in the SSVEP-based systems, ultimately increasing the rate at which users can spell. Moreover, as participants were tested over a relatively short time period, therefore the characteristic effects of P300 signal degradation were not fully explored. The study herein aims to define an analytical methodology with maximal generalizability and long term home usage. This rational therefore justifies our selection of the SSVEPs over the P300 and associated event-related potential bio-signals.

Steady State Evoked Potentials

SSVEPs are periodic waveforms that propagate across the lateral surface of the occipital and parietal lobes. These waveforms are a fundamentally bottom up phenomenon measured via EEG. Human users fixating or attending a visual flickering stimulus demonstrate EEG waveforms which mimic the phase and frequency of the stimulus flicker rate (Norcia, Appelbaum, Ales, Cottereau & Rossion, 2015). In other words, attending a stimulus with an 8Hz frequency and a phase angle of 0^o will result in the propagation of an 8Hz EEG waveform with 0^o phase angle across the occipital and parietal areas (Naz & Bawne, 2016) (refer to, Figure 2). Crucially, the SSVEP bio-signal has the lowest latency of any brain-based bio-signal, at just 140ms (Di Russo & Spinelli, 1999, Johansson & Jakobsson, 2000). This low latency in concert with a rapid remission to baseline activity gives SSVEPs a distinct advantage over alternative bio-signals, with SMR paradigms requiring seconds of imagined movement data and P300 systems requiring at least 300ms for waveform elicitation. This positions SSVEPs as the ideal bio-signal for attaining high ITRs in BCI speller applications.

Moreover the absence of signal attenuation over time affords SSVEP based BCI the potential for continual, long term real-world use (Cecotti, 2011). Additionally, the principally bottom up nature of this signal means that no complex, fatigue inducing training period is required prior to device interfacing, as necessitated by SMR-based paradigms. Guger et al (2012) found that SSVEPs could be reliably detected across a sample of 53 participants, achieving a mean accuracy of classification on a 4 target LED array in 95.5% of trials. Similar studies utilizing large sample sizes (n =>80) have been performed using comparable P300-based systems (Guger et al, 2009), showing 89% of subjects achieved >80% AoC and in SMR-based systems (Guger et al, 2003), just 19% of subjects showed an AoC >80%, after a 30 minute training period. This also suggests that the potential risk of BCI illiteracy is low in relation to SSVEP-based BCI systems.

As the SSVEP bio-signal can be reliably decoded in the largest proportion of the general population, it provides BCI utilizing this bio-signal with the greatest scope. However, the stimuli

utilized in SSVEP-based paradigms can introduce obstacles to data collection. During the acquisition of EEG-based SSVEP data for BCI applications, researchers primarily utilize electrodes positioned over the occipital cortex. Data collected at these sites has been shown to contain, on the whole, the highest quality SSVEP signals. However, due to significant individual differences in the responsivity to SSVEP stimuli, parietal regions closely neighbouring the occipital electrodes are also monitored. The literature surrounding the optimal frequency ranges for eliciting SSVEPs with the highest SNR ratio differ. Generally it is accepted this optimum lies between 8 and 16Hz (Herrmann, 2001). It is important to note that the range off stimulation frequencies typically defined as optimal overlap considerably with visual stimulations patterns known to induce epileptic seizure (Fisher, Harding, Erba, Barkley & Wilkins, 2005). Photic sensitivity is estimated to be present in around 15-20% of those diagnosed with epilepsy (Shiraishi, Fujiwara, Inoue & Yagi, 2001). The incidence of epilepsy in the general population is relatively low (1/100), however it is important to note prior to exposure of low frequency stimulatory patterns. This therefore restricts the applicability of such systems to exclusively non-photo-epileptic populations. However, some researchers have recently published findings indicating that higher frequency stimulation patterns outside the ranges specified to induce epileptiform activity may produce waveforms with higher temporal stability (Won, Hwang, Dähne, Müller, Lee, 2016).

Further limitations to the application of SSVEP based-BCI spellers include their historic classification as gaze-dependent bio-signals in BCI literature (Kübler, Kotchoubey, Kaiser, Wolpaw & Birbaumer, 2001, Gao, Cheng & Gao, 2003, Riccio et al, 2012). In other words, the propagation of an SSVEP is reliant on users fixating respective target stimuli, thus necessitating dextrous eye movement in order to form the basis of a reliable BCI speller system. Gaze dependency of the SSVEP signal therefore excludes patients with the most severe forms of quadriplegia from utilizing BCIs based on this waveform. To the opposite effect, a subset of studies directly researching this classification have produced results which suggest SSVEPs are more accurately defined as a semi-gaze-dependent (Muller, Malinowski, Gruber & Hillyard, 2003, Allison et al, 2008). Kelly, Lalor, Reilly, & Foxe, (2005) showed that SSVEPs could be accurately classified when attending one of two flicker visual stimuli in 71% of trials while controlling for fixation deviation via eye tracking systems. The minimal number of studies defining the semi-gaze-dependent nature of the SSVEP signal primarily include low density target displays. The apparent misclassification of SSVEP gaze-dependence has likely stunted the progress of purely attention based SSVEP BCI spellers. Therefore AoC performance of a purely attention based SSVEP-BCI in modern matrix style (>= 36 characters) GUI format is currently unknown. Arguably, the dependency of SSVEP propagation on attentional mechanisms could present obstacles to quadriplegic patients with low working memory capacity or chronic fatigue. However, ultimately the scope of SSVEP based BCI applications is potentially much larger than currently understood. This therefore poses the possibility that quadriplegic patients with low or absent ocular dexterity are in fact eligible for the use of SSVEP based BCI spellers.

In the past GUIs for SSVEP-based BCI used visual displays composed of Light Emitting Diodes (LEDs), as opposed to modern methods which rely on LED-based computer monitors. To reliably produce intended stimulatory patterns, target frequency values selected had to be divisible by the refresh rate of the device. For example, a 60Hz computer monitor is capable of reliably generating an 8.57Hz signal, as this corresponds to 7 frames per second (fps) (60Hz/ 8.57Hz = 7fps). LED devices previously exceeded the refresh rate of readily available computer monitors. Therefore, a larger portion of the optimal frequency range of SSVEPs (8-15Hz) could be utilized. Users were therefore required to attend two devices; an LED stimulator and computer monitor, to track the feedback (predicted target character) from the BCI speller.

In recent years the implementation of LED monitors with higher refresh rates in concert with state-of-the-art presentation algorithms, principally the *approximation method*, have remedied this obstacle (Manyakov, Chumerin & Van Hulle, 2012). Researchers have demonstrated that non-integer fps values, which are therefore non-divisible by the monitor refresh rate can reliably induce SSVEPs. This is achieved by interleaving two stimulation frequencies within a single refresh cycle. In order to generate a 9Hz stimulus frequency, researchers code a target to flicker at either

8.57Hz (7fps) or 10Hz (6fps) for a given portion of the refresh rate cycle. The proportion of 8.57Hz and 10Hz stimulus patterns is calibrated such that when the stimulus signal is averaged over the refresh cycle (e.g. 60Hz) it produces a 9Hz (6.67fps) stimulus pattern.

This same process of averaging is also true for the stimulus patterns embedded in the SSVEP EEG time-series. This affords the ability to increase the array density in matrix style GUI formats providing a larger number of operations to BCIs employing this presentation methodology (Hwang et al, 2012, Chen, Wang, Nakanishi, Jung & Gao, 2015). In conclusion, the capability for high density matrix style arrays with low latency has assisted researchers employing SSVEP-based BCIs in achieving the highest ITRs to date in both speller (Hwang et al, 2012, Chen, Wang, Nakanishi, Jung & Gao, 2014, Chen, Wang, Nakanishi, Jung & Gao, , 2015) and non-speller applications (Kwak, Müller & Lee, 2015). On review, the SSVEP bio-signal possesses the highest potential to form the basis for an out-of-the-box BCI speller for target quadriplegic patient populations.

Machine Learning

Machine Learning involves training a computational algorithm to classify (differentiate) or regress. Regression is used to estimate outcomes for an event in the future. For example, based on the current quality of EEG data (electrode impedance levels), researchers could estimate the likelihood that EEG signals will remain stable over time. Classification in contrast involves generating discrete observations such as estimating the frequency of an SSVEP. A number of the signal processing methods of EEG classification fall under this field of study. All methods incorporate analytical elements that learn features of the data for the purposes of optimization. A thorough discussion of the main analytical methods used for EEG analysis is documented herein.

Decoding Algorithms

Power Spectral Density Analysis

During Power Density Spectral Analysis (PDSA) of SSVEP data the EEG output undergoes a Fast Fourier Transform, which effectively converts EEG data from the time-frequency domain to an exclusively frequency based domain. Targets are determined by finding the frequency with the highest PSD value, then identifying the character corresponding to the predicted frequency value (Wang, Gao, Hong, Jia & Gao, 2008). Successful implementation of PDSA in relation to the decoding of SSVEP data has been demonstrated in numerous studies (Middendorf, McMillan, Calhoun & Jones, 2000). Additionally, the inclusion of harmonic and sub-harmonic component analysis fostered significant increases in classification accuracy (Cheng, Gao, Gao & Wu, 2002, Müller-Putz, Scherer, Brauneis & Pfurtscheller, 2005). The approach however requires substantial time windows in order to produce accurate readings (Lin, 2006). Therefore the ITR of systems implementing PDSA exclusively is typically low. Such analysis may have facilitated the discovery of distinct waveform frequency bands (delta, theta, alpha, beta and gamma), however fast Fourier transforms are highly susceptible to noise. Additionally, the methods of addressing these issues such as the application of autoregressive parametric power spectrum estimation are not optimal. These forms of additional analysis assume that the EEG output is linearly dependent on previous sampled data points. However the non-stationary nature of EEG waveforms over time diminishes the integrity of this assumption. Moreover PDSA is restricted to single channel analysis, whereas the practise of taking a grand average across multiple EEG channels to boost classification accuracy is now standard practise (Friman, Volosyak & Graser, 2007). PSDA, though initially popular in EEG waveform classification, are now less viable than other means. In order to overcome the incidence of misclassification in high density matrix style spellers more powerful methods of classification analysis were developed.

Canonical Correlation Analysis

Multi-Channel optimization significantly improves signal to noise ratios by amplifying relevant data features. Techniques such as canonical correlation analysis (CCA) have demonstrated this definitively (for review see, Vialatte, Maurice, Dauwels & Cichocki, 2010). Previously these methods have been employed on large scale data analytics, assisting researchers in extracting relevant information from vast datasets (von Storch & Zwiers, 2002). CCA, in relation to SSVEP classification, involves generating reference signal patterns with identical properties to the target stimuli patterns. For example if the letter A is assigned a stimulus pattern in which a stimulus square overlaid onto the target letter flashes on and off 8 times per second, a corresponding 8Hz reference signal is generated (refer to Figure 2). Signals are also matched for; length of stimulation period, number of EEG channels, micro-voltage range and sample rate. The first study to utilize CCA for SSVEP detection (Lin et al, 2006) was motivated by the restrictions imposed by PSDA methods, which limit analyses to just one EEG channel. In CCA classification of SSVEP data, each EEG time-series channel is denoted as n^{th} of matrix X. The adjacent dataset, denoted as n^{th} of Y, is made up of a Fourier series for one of the stimulus frequencies used to control the SSVEP targets. Each row of the Y matrix is a harmonic of the stimulus frequency expressed as a vector. Correlational coefficients for Xn^{th} and Yn^{th} are calculated to observe the amount of variance each Xn^{th} and Yn^{th} can count for in the adjacent matrix. CCA then operates to find the linear combination of coefficients which maximises the correlational relationship between datasets X and Y. This is referred to as the correlational variable and is expressed as an ordinary correlation (p).

Typically, during CCA analyses, the number of correlational variables calculated is equal to the depth of the smallest dataset utilized. In the application of CCA for SSVEP detection, only the first correlational variable is calculated as this accounts for the most variance and reduces the computational load of the analysis. Correlational variables are calculated for each stimulus frequency against the EEG input data. The stimulus frequency which produces the highest correlational variable (p) forms the prediction for the SSVEP embedded in the EEG time-series (Lin, Zhang, Wu & Gao, 2007). CCA does not depend on state-of-the-art hardware, in terms of computational load and is therefore relatively well suited for real-time speller applications. Bin, Gao, Yan, Hong & Gao (2009) demonstrated the validity of such analysis, achieving a mean AoC of 95.30% (12 participants) and an ITR of 58 bits per minute during online experiments for a 6 target sequential SSVEP based BCI speller.

The researchers implementing CCA for SSVEP detection must assume that the changes in micro-voltage across the scalp, responsible for SSVEP propagation is the product of a *linear* system in which the SSVEP stimulus signal is the input. This assumption is problematic due to the inherently non-linear and non-stationary characteristics of EEG time-series (Klonowski, 2009). Another primary drawback of implementing standard CCA (as described herein) is the methodology for constructing reference signals. Sine-cosine functions are used to generate template references which are inherently too precise. In other words, as they are not based off real-world samples they lack the presence of latent EEG features making dense matrix style GUI formats unfeasible (Pan, Gao, Duan, Yan & Gao, 2011). Advances in CCA have involved blending artificial random noise into generated reference signals, as well as user specific EEG output. So-called Multiway-CCA (MCCA) builds on this weakness of standard CCA by constructing a reference signals that incorporate relevant data features from a sample of EEG tensor data to produce more flexible reference signals. However such adaptations still incorporate a restricted set of highly precise frequency templates, ignoring natural background noise and other latent variables. This therefore leads to analysis being constrained to representing exclusively within-class relationships, over-looking between-class information (Zhang et al, 2013).

Common Feature Analysis

Common Feature Analysis (CFA) arguably overcomes the inherent issues of CCA to a greater extent than MCCA by generating references signals purely based on participant data. Zhang, Zhou, Jin, Wang & Cichocki (2015) tested the performance of these three classification models using highly simplified SSVEP stimuli; four red squares flickering at 4 distinct frequencies (6, 8, 9 and 10Hz). Authors proceeded under the assumption that there are inherent similarities in the data collected from **one** participant exposed to the same stimulus frequency over multiple trials. Latent features hidden within the participant specific data are computed by finding correlations between within target data. This therefore builds the template arrays with which to correlate raw un-grounded data for prediction. Analysis of performance stats via paired t-tests revealed that CFA and MCCA produce significantly higher classification accuracies than CCA for all time windows (0.5-4 seconds in 0.5 second increments). CFA also significantly outperformed MCCA in terms of accuracy for both the 0.5 (70% vs 43%) and 1 second (78% vs62%) time windows. Critically CFA was also significantly less likely to select non-targets as compared to both MCCA and CCA. In relation to computational efficiency CFA was significantly more accurate than both MCCA and CFA for the 0.5 second time window and matched MCCA for efficiency in the 1sec time window. This decreases dramatically for CFA as the time window increases. Greater data capture requires more computational time for the development of the common feature values. Therefore after this 1 second time window the efficiency advantage is lost. A CFA model would therefore only be applicable for experienced BCI users capable of interacting with the device in decision windows of just 0.5 seconds. Conversely, MCCA would be advisable for first time BCI users. This study is an excellent example of how selection of the classification algorithm is highly dependent on the BCI application and user.

K-Nearest Neighbour

K-Nearest Neighbour (KNN) methods are a simple form of non-parametric analysis with low computational training demands (Cover & Hart, 1967). KNNs firstly plot novel data points into a feature space populated by training data. The area of a regional 'neighbourhood' encompassing the data point is then calculated, the subscript denoting this area being k. The novel input is then classified by determining which target group has the highest number of data points within the regional 'neighbourhood' defined. In the instance of more than one class providing an equal number of neighbouring data points the novel input is assigned to the target group with the most proximal data points in the feature space populated by the training data (Kataria & Singh, 2013). If implemented with a substantial training set and high k value KNN methods can reliably classify non-linear data (Lotte, Congedo, Lécuyer, Lamarche & Amaldi, 2007). These properties initially inspired BCI researchers to employ these methods in the classification of EEG data however attempts were not successful. This can be ascribed to KNNs performing poorly on data with high dimensionality and in the context of data with large ranges. However, such issues can be mitigated with the implementation of principal component analysis and normalization pre-processing stages (Khasnobish, Bhattacharyya, Konar & Tibarewala, 2010, Nasehi & Pourghassem, 2011).

In the past decade KNN methods have not been frequently selected for online BCI spelling devices due to the protracted computational periods involved in predictive search times. However, some success has been shown when KNNs are combined with principal component analysis and multivariate linear discriminant analysis (MLDA) in a processing pipeline. Wang, Chen, Gao & Gao (2016) first gathered data using a simple 4 class paradigm. The participants completed a task which involved fixating one of four flickering red squares (6, 8, 9 and 10Hz) in a randomized sequence. Performance of a MLDA/PCA/KNN, standard CCA, MCCA and CFA was compared across participants for 8 different data capture periods (0.5-4seconds with 0.5second increments). PCA on raw EEG data lead to the identification of multiple variables which in combination accounted for 99% of all data variance. The data was then labelled using the values assigned by the newly defined feature variables. MLDA was then applied which harnesses gradient descent to reduce the sum of squares (error). Finally a k-nearest neighbour's algorithm is applied (k value = 5).

Analysis revealed that MLR/PCA/KNN (83%) was significantly more accurate than CIFA (67%) during the 0.5-second data capture period. MLR also outperformed MCCA on all time windows under 4 seconds and was significantly more accurate than standard CCA in all data capture periods. Critically MLR/PCA/KNN showed the least class imbalance across all data capture periods. This is essential for BCIs which aim to adapt to user performance. If the model changes time window duration based on user performance to enhance ITR it is important that performance is consistent. In other words, similar classification accuracy should be observed for a range of data capture periods (0.5-4seconds). CCA does have a distinct advantage over MLR in that no model training time is necessary, however only 2 or 3 samples of training data are required to ensure the MLR/PCA/KNN significantly outperforms CCA on all time window conditions. Moreover MLR/PCA/KNN can be implemented with an average computational period of just 0.1seconds; making this the system extremely useful in real-time applications.

Support Vector Machines

Support Vector Machines (SVMs) classify data by comparing known-ground truth data values to novel input data. Initially SVMs were designed for the purposes of classifying linearly separable, binary data. Each data point is expressed as a vector based on respective input variables. Adjacent vectors from each group with the shortest distance are used to define the connecting line (2D), plane (3D) or hyperplane (>3D) depending on the number of input variable dimensions (Burke, 1998). Essentially the process of classification is represented as a minimisation problem which aims to bisect the data with a separating line which maximises the margin between target groups. Critically, the calculation is restricted to data points neighbouring the computed hyper-plane (dimensional area in which categorical overlap can occur) (Bennett & Campbell, 2000). The distance in decision space between the un-labelled novel data and the hyperplane is far more

computationally economical as compared to for example linear discriminant analysis based methods as these calculate the distance between all past data and the current input. Past research has revealed the successful implementation of SVMs primarily for classification of SMR-based BCIs (Xu, Guan, Siong, Rangantha & Thulasidas, 2004, Schlögl, Bischof & Pfurtscheller, 2005, Siuly & Li, 2012) and optimization of EEG channel selection (Lal et al, 2004, Schröder, Hinterburger, Bogdan & Birbaumer, 2005). Lotte (2007) also cites the low number of hyper parameters and overtraining resilience as advantageous to the implementation of SVM based analysis.

Kernel Trick

In order to accommodate for non-linear data, SVM feature space can be projected into higher dimensions using transformations such as; polynomial functions, Gaussian radial basis functions and sigmoid functions (Boser, Guyon & Vapnik, 1992). Known as the 'kernel trick', the purpose of projection is to increase the probability of calculating a linear separator which defines a hyperplane with a larger margin. Maximising the width of the margin has the effect of reducing classification error when labelling data which shares the features of more than one variable input (Lotte, 2007). Additionally, margin maximisation between groups affords the SVM higher generalisation capabilities (Jain, Duin & Mao, 2000). Impressive results in relation specifically P300 based-BCI speller have been demonstrated in multiple studies (Kaper, Meinicke, Grossekathoefer, Linger & Ritter, 2004, Li, Guan, Li & Chin, 2008). Importantly however, it must be noted that projection into higher dimensional space greatly increases the burden of computational cost restricting the applicability of such methods in a real-time context. Moreover, repeated increases in dimensionality have been found to introduce a greater degree of generalization error in relation to novel input data classification (Jin & Wang, 2012).

Ensemble SVMs

Ensemble SVMs (ESVMs) where first proposed by Vapnik (1999) in order to provide flexibility to the implementation of SVMs on non-binary (multi-class) problems. It is important to note the increase of prediction calculators, each assigned weighted terms to define the level of contribution to the output answer, does not enable the modelling of non-linear relationships. SVMs in the ensemble are trained separately on randomly selected batches from the training dataset. This results in SVMs becoming selective to a subset of feature space. Localizing the classification capabilities of the SVMs in ensembles has been shown to dramatically increase the performance of such forms of analysis in comparison to individual SVMs. However, ESVMs have been shown to perform poorly on EEG classification problems large numbers of classes (Li, 2014). Therefore, designing a BCI around this form of analysis would result in a device with low operational flexibility and sub-optimal average classification performance.

Filter Bank Analysis

Chen et al (2015) have produced arguably the highest ITR for a complete 40 class target array of alphanumeric characters (5 x 8 stimulation matrix) to date. This was achieved by modulating both the phase and frequency of targets individually. The manipulation of stimulus phase offsets in the context of SSVEP research has been studied previously (Hartmann & Kluge, 2007, Kluge & Hartmann, 2007, Wilson & Palaniappan, 2009, Lee et al, 2010, Jia, 2011, Manyakov et al, 2011). However Chen et al (2015) have maximised the theoretical potential by assigning targets in an SSVEP-based BCI speller unique phase and frequency values to enhance performance of standard CCA classification algorithms. This technique was initially inspired by research in the telecommunications industry and has had a significant effect on the previous restrictions of spectral separation. As the number of targets in the visual array increases, the amount of spectral separation between target frequencies decreases. Therefore the likelihood of misclassifying neighbouring frequency values increases and dramatic reductions in ITR follow. By assigning target stimulation patterns unique frequency and phase values the similarity in resulting SSVEP waveforms is substantially reduced, therefore decreasing misclassification errors.

Initially, authors carried out frequency optimization procedures to determine the frequencies to which individual participants were maximally responsive (Wang, Wang, Gao, Hong & Gao, 2006). Previous research has shown that on healthy human participants are maximally responsive to stimulation patterns within the range of 8-15Hz. However BCI developers have taken advantage of individual differences in the distribution of responsivity within this range. By validating the exact frequencies prospective users are responsive to researchers can optimise the target frequencies presented. Finer incremental increases can be achieved nearer to the locus of optimal signal to noise ratio as the ability to differentiate between neighbouring target frequencies is higher. Further optimization in the form of filter bank analysis was also carried out to establish optimal frequency harmonic ranges for each participant. Filters which proved more effective to the successful classification of EEG data were provided more weight throughout a sub-band training process. The ranges were learned iteratively by feeding EEG data into 5 overlapping frequency. Critically, each arrangement of filter weights was unique based on the input of each participant in the study. Researchers tested the performance of a joint frequency and phase coding method (JFPM) using an array of 40 target alphanumeric characters. JFPM is optimized for the decoding of SSVEP signals generated from a dense array of targets with frequency increments of just 2Hz by integrating evenly spaced corresponding target phases. This was performed in order to maximise the difference between correlational coefficients of neighbouring stimulation patterns.
Each character in the array was tested in a random sequence. Targets for fixation were initially cued by a surrounding red square. Such task manipulations would greatly reduce visual search time and mental workload which would artificially increase the data quality. Therefore, free-spelling tasks were also implemented and demonstrated similar results (4.5 bpm). The JFPM model in online analysis produced an ITR of 170 bits/min and mean accuracy of 88.83%.

Despite the impressive performance stats reported the hours of pre-training data collection for sub-band optimization mean that this system does not possess 'out-of-the-box' functionality. Individual differences in SSVEP responsivity are not stationary, therefore repeated measurements for setting these tailored increments would have to be routinely calculated. The methodology of stimulus presentation has been adopted in this study by utilizing a benchmark dataset produced by the same Beijing Institute (Wang, et al 2016). However a deep neural network was applied in the process of target classification to produce an EEG decoding methodology that functions at industry standards (>=70% accuracy) as soon as a novel user is introduced to a BCI speller. Critically, user-specific parameter optimization was not implemented due to the potential for this step to reduce the deployment time of systems developed at a later date.

Neural Networks- Review

Developments in the field of computational engineering are often fostered by the investigation of neuroscientific phenomena (for review see, James et al, 2017). In the same vain, neural networks are analytical software structures which aim to mimic the classification capabilities inherent to biological systems. Networks are composed of functional units ('nodes') arranged in operational layers which learn to model and classify incoming data, much like neuronal circuits. The characteristic of redundancy (presence of additional components to ensure success in the event of other components failing) which makes the human brain so computationally powerful is harnessed for the classification of many different data inputs. Neural networks are commonly termed 'black boxes', meaning that the data features represented by the nodes which populate the network are not entirely discernible. This characteristic of artificial neural networks (ANNs) also underpins one of the major obstacles of its implementation: architecture building. The process of constructing a neural network can be time consuming as there is no definitive methodology that can guarantee high performance.

The development of such models has many factors; initial weight optimization, number of nodes per layer, variant of normalization performed, degree of down-sampling, network depth, selection of activation function, convolutional filter dimensions and rate of drop-out. All such elements influence the performance of the network. Attempts to investigate structural qualities of a network are problematic as many studies have revealed that models with major differences in terms of the aforementioned factors provide highly similar accuracy results (Sjöberg et al, 1995). Generally, the selection methodology for models in novel applications usually consists of identifying networks used in a similar context as opposed to the bottom up development of a tailored network structure (Prieto et al, 2016). The 'weights' of a network consist of matrices of randomly generated values in typically uniform dimensions (e.g. 6x6). The values which populate this array should contain comparable characteristics to the input data in terms of: variance, range and mean. This process of assigning the base weight values is known as initialization. Improper initialization can lead to exorbitant training periods therefore, this first step is crucial in terms of research time constraints. In order to estimate the quantity of nodes necessary for each layer the data complexity must be taken into account also. Datasets collected with low resolution hardware or sets consisting of data which is inherently noisy may require a greater number of nodes per layer in order to increase the probability of salient data features being represented. A network with too many nodes can lead to over-fitting and 'noise-trailing'. Overfitting occurs when the model becomes highly adept at classifying data solely from the dataset used in training (Piotrowski & Napiorkowski, 2013). Therefore, if the model were to be used in the context of novel data, the classification accuracy would be low or potentially random. Conversely, too few nodes can lead to low classification accuracy when complex data is inputted (inability to analyse non-linear relationships).

Artificial Neural Networks

There exists a multifarious array of ANN architectures and training methods; however the focus of this thesis is to examine the application of convolutional neural networks, trained using supervised learning for the classification of SSVEP data. Each aspect of the network must be rigorously fine-tuned to achieve the maximal error reduction in the prediction of classes. Despite this lack of clarity and laborious development period, higher classification results are consistently found during the application of these networks. Typical ANNs contain an input layer (data) and output (decision) layer, this is a highly simplistic model and classification is only feasible for linear relationship detection. A network typically converges towards global minima (lowest proportion of classification error achievable) via a loss function which aims to maximally reduce the mean squared error (Zhang, 2000). Essentially this involves calculating the distance between a predictor (in this case an ANN nodes) and the input data, then utilizing this information to reduce this gap in future classification operations. Networks can be categorized as either **supervised** or **unsupervised**. The later refers to training a model to classify data in the absence of a ground-truth (data label). This enables the self-organization of the analytical structure, in which the model generates data classes with maximal separation. In contrast, supervised learning functions by feeding the network input data and corresponding class labels with the aim of defining data features which separate explicit target groups.

Applications of neural networks extend too many disparate fields such as; financial solvency estimation (Wilson & Sharda, 1994, Lacher, Coats, Sharma & Fant, 1995), early speech recognition research (Lippmann, 1989, Bourlard & Morgan, 1993) and medical diagnosis (Baxt, 1990, Burke et al, 1997, Amato et al, 2013). Pertinently these analytical methods have shown significant classification capabilities in relation to EEG based data for the detection of epilepsy (Srinivasan, Eswaran & Sriraam, 2007) and mental states (Vuckovic, Radivojevic, Chen & Popovic, 2002). Initially, the efficiency of ANN models to classify EEG based bio-signals for BCI applications was limited in comparison to other linear analytical methods (Dreiseitl & Ohno-Machado, 2002). However, due to the increase in availability of ANN based software performance steadily improved (Robert, Gaudy & Limoge, 2002).

Multi-Layer Perceptrons

Multi-Layer Perceptron Neural Networks (MLPNN) differ from simple ANNs due to the presence of a hidden layer (Schmidthuber, 2015). MLPNNs are highly adaptive to a range of classification problems presenting differing levels of pattern complexity. Light-weight implementations can achieve impressive results analogous to basic parametric analysis for simple pattern recognitions tasks. Likewise, increasing the depth and node quantity allows the network to model of highly complex systems. The application of a non-linear activation functions at the hidden intermediate processing stage affords the analysis of non-linear systems (Sarle, 1994). Conceivably MLPNNs are capable of modelling a limitless number of functions, warranting the description of MLPNNs as 'universal estimators' (Nicolas-Alonso & Gomez-Gil, 2012). Reaching maximal network performance is possible when enough nodes, arranged in the optimal number of layers are formulated. However restrictions arise in terms of processing power. The efficacy of such networks in relation to EEG data has been researched primarily in the context of potential clinical applications; detection of epileptic seizure events (Işik & Sezer, 2012 Artameeyanant, Sultomsanee & Chamnogthai, 2017, Jaiswal & Banka, 2017) and Alzheimer's onset evaluations (Samiee, Kovacs & Gabbouj, 2015, Cecere, Corrado & Polikar, 2014, Er et al, 2017). Moreover, these models have also been employed in BCI relevant fields such as; mental state classification (Anderson, Devulpalli & Stolz, 1995, Nasehi & Pourghassem, 2013) and EEG channel optimization (Ang, Chin, Wang, Guan & Zhang, 2012). Pertinently, numerous researchers have demonstrated that MLPNNs are effective in classifying SMR bio-signals in BCI contexts (Balakrishnan & Puthusserypady, 2005, Chatterjee & Bandyopadhyay, 2016). Importantly, direct performance comparisons between MLPNNs and Logistic Regression techniques revealed MLPNN (L-M) produced higher AoCs (93%) than LR models (89%), higher specificity (92.3% vs 90.3%) and higher sensitivity (92.8% vs 89.2%) ,(Subasi & Ercelebi, 2005). These results suggest that the ability of neural networks to detect latent EEG features with potentially non-linear relationships in the data contribute to higher accuracy of classification. Additional recent research has since corroborated these findings (Shedeed, Issa & El-Sayed, 2013, Nurse, Karoly, Grayden & Freestone, 2015).

Additionally, Manyakov et al (2012) performed an SSVEP classification study modulating both frequency and phase using a shallow feed forward MLPNN. The task involved fixating a centrally located flickering target. Stimuli were presented individually at unique frequency rates. Phase shift to the targets was added post-hoc and the analysis performed offline. 83% classification accuracy was reported on the 16 class implementation with a data acquisition time of 5 seconds. SSVEPs are however also triggered via **covert** attention, meaning that participants are capable of distraction from the intended target for selection if the visual saliency of neighbouring targets reaches a specific threshold. In other words, as the array becomes more densely populated the degree to which **covert** visual attention of non-target neighbours increases. Therefore, the individual presentation of targets means that these results are unlikely to translate into a real-world setting where up to 40 targets are presented simultaneously.

Within the literature there are arguably more frequent accounts of this methodology failing to exceed classification capabilities of less computationally expensive forms of analysis (for review see, Lotte, 2007). Bai et al (2007) compared the performance of numerous analytical methodologies on an offline 2 class task involving the classification of SMR imagined hand movement data. A combination of independent component analysis, power density spectral

estimation and SVMs showed significantly higher accuracies as compared to MLPNNs. Additionally, Ilyas, Saad, Ahmad & Ghani (2016) demonstrated that in the context of a 3 target sensorimotor classification task for MI-based BCI both SVMs and LR outperformed MLPNNs. As previously noted, the capabilities for MLPNN techniques to model highly complex non-linear systems are substantial. The sub-par performance of MLPNN models in the aforementioned instances can be attributed to a number of factors; insufficient data quantities or resolution, poor optimization of network hyper-parameters and unsatisfactory numbers of network node volumes or depth. Recent advances in the utilization of graphic processing units (GPUs) for the calculation of weight updates in neural networks have assisted numerous researchers in tackling these issues. GPUs are traditionally employed to increase the rate at which images are rendered on digital displays. This hardware is designed to function with a high level of parallelism for the purpose of calculating potentially millions of common operations. Oh & Jung (2004) were the first to successfully utilize GPU hardware for the purpose of neural network specific computations. Input data is transformed into a GPU compatible data format, for example vertex values representing colour, position or texture. Matrix multiplications corresponding to derivative calculations for network weight updates are then passed to a pixel shader these compute coded changes in pixel attributes prior to image rendering. Authors were able to achieve convergence at 20 times the rate of a CPU using a benchmark dataset and template neural network code for a text localization task. Similar performance increases have been noted in numerous articles (Luo, Liu & Wu, 2005, Chellapilla Puri & Simard, 2006, Strigl, Kofler, Podlipnig, 2010), and critically it has been shown to scale effectively to more computationally demanding networks (see for review, Raina, Madhavan & Ng, 2009).

With the development of NVIDIA's CUDA software access to high speed network training rates has become an industry standard. The increase in processing power afforded by the aforementioned GPU techniques has enabled researchers to make use of ensembles of MLPNNs. Separate shallow MLPNNs are trained on individually randomised subsets of EEG training data. This process leads to development of networks all of which possess a variety of data feature representations. All networks provide predictions regarding the class of the novel input data, the class with the highest incidence of selection across all models is chosen as the final prediction. These methods have been found as effective means of classification in relation to motor-imagery based BCI (Silva, Barbosa, Viera & Lima, 2016) and SSVEP-EEG data (Chen, Chen & Wu, 2016). Despite these promising results on shallow MLPNN and ensemble networks, only a handful of researchers have endeavoured to develop truly **deep** neural networks for the classification of SSVEP-EEG data.

Deep Learning

Deeping learning refers to a wider field which employs the use of neural networks with multiple layers to enhance the complexity of data features represented by the model. The origins of the deep learning can be traced to Fukushima's (1979, 1980) development of the *NeoCognitron.* This architecture is cited by many as the first true deep neural network (Schmidhuber, 2015). Early research of Mountcastle (1957) and the later investigations of Hubel and Wiesel (1962) both evidenced the presence of a topographical architecture in the mammalian visual brain. Interestingly, deep neural networks, having been initially inspired by human visual brain, are now being implemented to provide a greater understanding of these structures (Dähne et al, 2015, Cichy, Khosla, Pantazis, Torralba & Oliva, 2016). The hyper-interconnectivity between successive layers of the mammalian cortex results in a monotonic increase of receptive field size. This property of the visual brain is harnessed by the *NeoCognitron* to foster invariance to subtle changes in the quality of input data.

Similar levels of node connectivity and depth across layers had previously proven impractical due to computational bottlenecks. The *NeoCognitron* circumvents this obstacle by applying a 'weight sharing' (WS) procedure. Feature maps are data matrices produced by operations which contain some abstraction of the input data that may be useful for the classification task. A weight, otherwise known as a filter, is a matrix or array of random values which is used to produce the feature maps. As the network process the input, data is grouped into windows, which specify a subsection of the input on which the filters operate. Collectively, these filters cover the entire input

data array and can be overlapping. The higher the degree of overlap, the higher the number of operations and therefore the greater the computational cost. The depth of the filter is matched to the depth of the windows covering the data. Feature map values are therefore computed by performing a matrix multiplication, in which the n^{th} value of the filter and the n^{th} value of the data window are multiplied. The same filter values are used in the computation of all feature map values and are not changed throughout the network testing period.

The training of MLPNNs requires the computation of individual weight functions. This is far more computationally expensive and leads to an increase in the amount of training time necessary for the model to converge. Envisage a network which is designed to detect squares within an image. Each filter relates to a different segment of the image. Therefore, every single weight must be individually optimized to detect squares, as foreseeably the squares could appear at any location in the image. It is more computationally efficient to share the weights and optimize the values in the filter array collectively. This saves computational time and allows the network to detect squares independently of square's positioning in the image. These same principles apply in relation to EEG temporal data. It is more efficient to share the weights of filters across an entire input such that data features relevant to the task of classification can be extracted, irrespective of where these are positioned temporally. This constitutes the fundamental difference between simple, shallow multi-layer perceptrons and deeper neural networks. Numerous studies have shown that using shared weights does not drastically decrease network performance, as the huge reduction in computational load allows for a: higher number of epochs and slower learning rate (Schmidhuber, 2015).

Convolutional Neural Networks

CNNs are state-of-the-art deep learning models typically employed primarily in the domain of computer vision. Computer vision refers to the field of developing analytical systems which process visual information to classify or segment salient features or objects (Lee, Grosse, Ranganth & Nh, 2011). Deep Learning CNNs were are amongst the first to demonstrate human detection level performance (LeCun, Bottou, Bengio & Haffner, 1999) and have a proven track record of winning some of the most coveted computer vision competitions in recent history, ImageNet (Krizhevsky, Sutskever & Hinton, 2012), GoogleNet (Lofee & Szegedy, 2015) (LeCun, Bengio & Hinton, 2015).

Unlike simple, shallow MLPNNs, deep CNNs take advantage of the abovementioned weight-sharing principle to minimise computational processing time. Another feature differentiating deep CNNs from MLPNNs is the use of sparse, local connectivity. In traditional MLPNNs, all nodes in adjacent layers are weighted, in other words they are **fully** connected. During the development of deep CNNs, it is assumed that there is a high correlation between neighbouring nodes. In other words the output of convolutions performed on neighbouring windows of data is predicted to be highly correlated, as compared to windows positioned at the start and end of the data matrix. In deep CNNs, weights are combined across these locally positioned nodes and connected to nodes in adjacent layers. Initially, this means nodes in successive layers are only exposed to data from a small region (spatial) or time period (temporal) of the input. However, as the depth of the network increases so too does the receptive field size of respective nodes, as output from distally positioned data is eventually integrated into convolutional operations. In the top layers of the network, node receptive fields can span the entire length of the initial input data. CNNs are therefore not the optimal choice for localization tasks. These nets are highly adept at predicting the presence of patterns within data, as opposed to segmenting or identifying when (timing) a specific pattern occurred during data collection.

Recent applications of CNNs include medical image diagnosis (Bevilacqua, Mastronardi & Marinelli, 2006, Dou et al, 2016, Esteva et al, 2017), facial recognition (Zhang & Zhang, 2014, Farfade, Saberian & Li, 2015, Li, Lin, Shen, Brandt & Hua, 2015) and satellite imagery scanning for interstellar objects (Kim & Brunner, 2016, Kimura et al, 2017). CNNs are however not restricted to exclusively visual information. Impressive performance in; prosthesis kinematic estimation (Allard et al, 2016, Atzori, Cognolato & Müller, 2016), EEG anomaly detection (Wulsin, Blanco, Mani & Lit, 2010), epileptic EEG pattern detection (Antoniades, Spyrou, Took

& Sanei, 2016, Cılasun & Yalçın, 2016, Johansen et al, 2016) have been demonstrated. Also, these same methods have been implemented effectively in the context of BCI bio-signal classification for; mental imagery (Jingwei, Yin & Weidong, 2015, Walker, Disenroth & Faisal, 2015, Tabar, 2016, Schirrmeister et al, 2017) and P300 detection for BCI speller applications (Cecotti, 2011, Lawhern et al, 2016).

The application of CNNs in decoding SSVEP data for BCI applications has however not received significant attention from the academic community. Past research using actual end-point patient populations is extremely scares among the literature, between 16-25 legitimate studies. Clinical applications of assistive devices incorporating neuro-feedback have demonstrated generally positive outcomes (Birbaumer, Murguialday, Weber & Montoya, 2009). However, only a small minority these studies represent long-term home usage of BCI systems, the majority of which employ P300 paradigms for BCI speller applications (Peridikis, 2014). There are many reasons for this dearth of research: patient access, heterogeneity of patient populations and funding grant time constraints (Holz, Botrel, Kaufmann & Kubler, 2015). Even if all cited obstacles can be circumvented it still remains that the real-time BCI hardware and software integration must be seamless in order to ensure the collection of viable data. The difficulty in adapting pre-processing methods for the often irregular neural activity present in patient user EEG outputs makes this process extremely problematic. The justification of CNNs is therefore evident in that irrespective of potential abnormalities network weights can be adjusted in real-time and compensate for the unique variations in EEG signal of the user. The successful development of a generalizable 40 target SSVEP based BCI speller across a large sample of experienced and naïve participants would allow rapid deployment of the system. This would by-pass the need for prolonged data collection periods prior to online use as baseline accuracy of around 70% would allow for an immediate moderate level of control. During online-use training of a parallel network with weighs fine-tuned to the user's unique EEG could be run simultaneously. Therefore fatiguing initial data collection is averted while continuously improving the classification accuracy. A similar process was explored by Williams (2017); with intra-subject transfer learning following an initial inter-specific training period which resulted in higher accuracies of classification.

A handful of notable studies investigating deep CNNs for SSVEP classification are present in the literature and deserve consideration. Cecotti et al (2011) developed a deep CNN for the classification of 4 SSVEP stimulation targets differentiated by frequency. The authors positioned a fast Fourier transform within the network in order to convert EEG data from a time-frequency domain to a purely frequency based format. A mean classification accuracy of 95.61% was achieved; however the sparse target array limits the potential flexibility of BCI spellers built around this model. In a related study, CNNs were also found to outperform alternative methods such as standard CCA and a CCA/ KNN hybrid analysis (Kwak, Müller & Lee, 2017) in the classification of SSVEP in an ambulatory (mobile) environment. These findings corroborate the claims that CNNs possess significant noise resilience in relation to EEG data artefact.

With end-point applications of BCI spellers being their utilization by individuals with quadriplegia it is necessary to consider the decrease in bio-signal quality from such patient populations. Geronimo, Simmons and Schiff (2016) tested a heterogeneous sample of ALS patients on a variety of performance metrics when interfacing with BCIs using both motor imagery and P300 based bio-signals. Authors concluded that behavioural dysfunction as a result of progressive cognitive impairment reduced signal quality significantly. Additional research has also revealed that working memory training prior to use of BCI systems can boost ITR performance (Sprague, McBee & Sellers, 2016). Irrespective, applications should be designed to feature: minimal-no training time, short data capture windows and employ classification techniques which are highly resilient to noise.

During the preliminary literature review in preparation for this thesis, it became apparent that as yet there has not been a successful research article defining a means of accurately classifying SSVEP BCI speller data across a large sub-set of participants for a densely populated visual array. Chen et al (2015) trained participant specific frequency sub-band filters to enhance classification performance on the single subject level. This process necessitates a long data collection period

prior to any real-time classification. Therefore the purpose of this research is to produce a neural network capable of achieving >=70% classification on wet-EEG data for a large sample of patients. Successful development of such a network would accelerate the deployment of SSVEP based BCI spellers and also provide the possibility of integrating user data into later network training sessions to further optimise the system to the individual. In other words the initial obstacle of near random performance during the first few trials of EEG classification would be overcome. Deep CNNs were implemented for the classification of an 8 and a 40 target classification problem. The data used in both model training and validation were derived from a benchmark dataset provided by the Tsinghua Institute (Beijing) (see, Wang et al, 2016 for further details).

Methodology

Data Origins

The data used to train, validate and test all networks outlined in this study was acquired from an online open source repository hosted by the Department of Biomedical Engineering, Tsinghua University (Beijing) (Wang et al, 2016). The following section defines the experimental features of the data acquisition carried out in the above mentioned study. Efforts have been made to expand on the reasoning behind the protocols implemented with the aim of providing additional understanding. For further information, refer to the original article. The accompanying dataset publication gave no suggested analytical methodology. In other words, all data organisation, pre-processing and analysis were performed with custom code unless specified explicitly.

Participants

A total of 35 participants were recruited for the original study (18 males, age range: 17-34 years, mean age= 22 years). All participants tested had normal or corrected to normal vision. A subset (8) of the participants had prior experience with SSVEP-based BCI spellers from partaking in past studies with the respective institute. The remaining 27 participants had no prior experience with such systems and are there classified as naïve. All participants were informed about the study requirements and procedure prior to gaining consent. Ethics approval was provided by the relevant institutional committee (Research Ethics Committee of Tsinghua University).

Stimulus Presentation

The study consists of a 40 target BCI speller experiment. Targets comprise a selection of English alphabetical characters (26), numerical digits (10) and symbols (4) in a 5 x 8 visual target array. The targets were displayed on a 60cm Acer GD245HQ LCD computer monitor with a latency of 2ms and 1920x1080 resolution at a frame rate of 60Hz. Targets (size: 32x32 pixels/ 1.8x1.8cm) were centrally oriented inside stimulus squares (sizes: 140x140 pixels/ 8.1x8.1cm), with a uniform distance between adjacent targets of 50 pixels/ 3cm across both vertical and horizontal planes. The distance between monitor and participant was kept consistent at 70cm throughout the data collection period. Stimulus presentation code was programmed using MATLAB (Mathworks, Inc.) and associated library, Psychophysics Toolbox. Unique stimulus patterns were assigned to individual targets following the joint frequency and phase modulation methodology (refer to Figure 2). Dual frequency and phase combinations were developed as per:

$$(kx,ky)=f0+\Delta f\times[(ky-1)\times 5+(kx-1)]$$

$$\emptyset(kx,ky)=\emptyset 0+\Delta\emptyset\times[(ky-1)\times 5+(kx-1)] \quad (1).$$

In which *kx* refers to the row and *ky* the column index of the visual array. Where $f0$ denotes 8Hz and Δf 0.2Hz (range: 8-15.5). The stimulus presentation followed a sampled sinusoidal stimulation approach. With f and $\emptyset$, a presentation order s(f,$\emptyset$,i) is produced via altering monitor luminance:

$$(f,\emptyset,i)= 12\{1+\sin [2\pi(i \text{ RefreshRate }) + \emptyset]\} \quad (2)$$

sin() creates a sine wave and i represents the frame index within the sequence in which 0 represents dark and 1 represents maximal luminance (refer to Figure 2).

In order to ensure that the coded stimuli patterns match the actual stimuli patterns on screen, researchers calibrated the BCI system using photo-diode measurements. The photo-diode was placed on the monitor used to present target stimulus patterns and generated read-outs expressed in micro-volts. These micro-volt values were then converted into units of luminance to ensure that the changes in stimulus brightness matched the coded values. Further, luminance changes were also measured over time for each coded stimulus pattern to ensure that the frequencies

programmed also matched actual target frequencies presented on screen. These methods were adapted from the research performed by Manyakov et al (2013) and implemented in collected of these data, as well as an additional previous study (Chen, Chen, Gao, Gao, 2014). Authors do not explicitly make note of the difference between expected and actual stimulus patterns during the collection of the benchmark dataset in question.

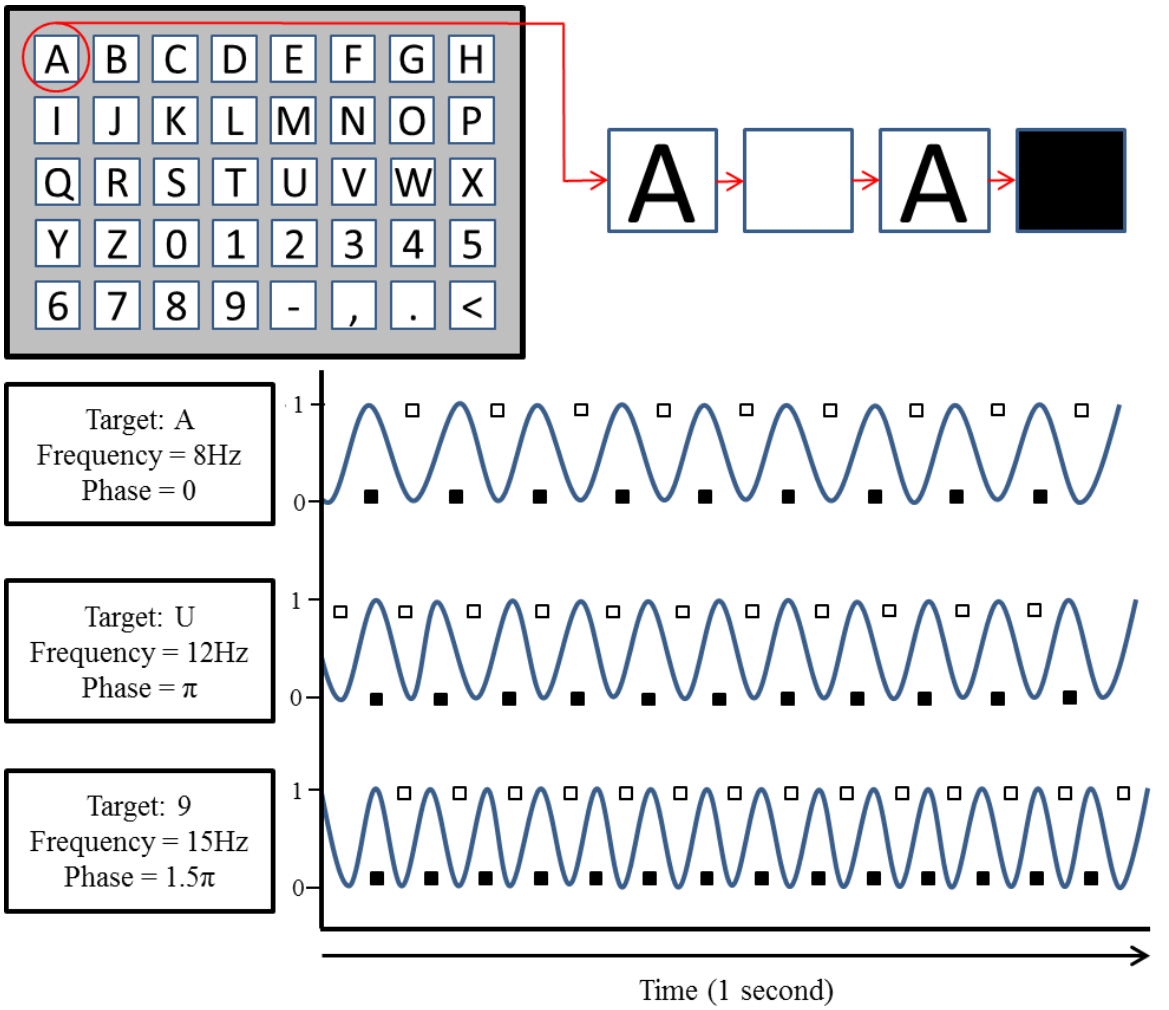


Figure 2. Cartoon of the stimulus patterns used to control the stimulus squares overlaid onto the target letters, numbers and symbols. The top panel shows a representation of the alphanumeric array presented to subjects via a computer monitor. In the top-right section there is a visual representation of how the stimulus frequencies are generated, by alternating an overlaid stimulus square between black and white onto each target character. The bottom panel shows a graphical representation of how the coded stimulus signals alter the behaviour of stimulus squares over time. The frequency of the sinusoid signals define the rate at which the stimulus squares oscillate between black and white. As shown in the figure, stimulus patterns with higher frequencies iterate between black and white brightness more frequently per second. As the signal reaches one, the stimulus square is presented with maximal brightness. As the signal drops to zero, the brightness of the monitor region dedicated to the stimulus square is at minimal brightness. Further, the figure illustrates how the fluctuation in brightness controls the phase offset of the signal.

Experimental Design

Participants were required to perform a cued target selection task. Initially targets are cued via the presentation of a red square at the intended target for selection (0.50 seconds). Participants are instructed to fixate upon the pre-cued target (5 seconds), after which the monitor returns to a resting blank state (0.50seconds). During the fixation period all 40 targets flash concomitantly according to their respective stimulus patterns. The total time period for cueing, acquisition and resting was therefore, 6 seconds. In order to minimise the amount of non-relevant data included into the samples collected participants were instructed to rapidly redirect their gaze to following targets. Moreover, in order to minimise the presence of EEG artefacts participants were asked to

refrain from blinking or jaw clenching during the fixation period. One block consisted of fixating all 40 targets in the visual array once (6seconds*40 targets= 4minutes). A total of 6 blocks were collected with each participant and a mandatory break was implemented after every 2 blocks.

Data Acquisition

The EEG system utilized for SSVEP data acquisition was the 64 channel Synamps2 EEG (Neuroscan) which adheres to the standardised 10-20 electrode arrangement. EEG samples were consistently collected at a rate of 1000Hz. For the purposes of this study analysis was restricted to the following channels: O1, O2, Oz, P3, P4 and Pz. An electrode impedance level threshold of <=10kΩ was maintained throughout the testing period. During data acquisition time stamps were paired to stimulus events in order to correct for temporal instabilities in the collection period with time synchronisation techniques. Subjects were tested in a low-light soundproof room while sat in a padded chair.

Data Pre-Processing

Past research utilizing the aforementioned task and hardware revealed that applying an upper-bound of 90Hz for SSVEP harmonics extraction was optimal; therefore this same procedure was adopted. Post-collection all data were down-sampled by a factor of 4 to minimise memory demands and computational load. Removal of power-line noise was achieved via application of a 50Hz notch filter. Additionally all data was normalized around a mean of zero and scaled between +1 and -1. Finally, data across selected occipital and parietal channels were grounded (via subtraction) using data derived from the Cz (vertex) electrode).

Many implementations of CNNs make use of data augmentation in order to increase the sample size of data on which their neural networks are trained on. This can involve horizontal data flipping, data cropping and the addition of artificial noise/ data blurring. Data augmentation serves the purpose to ensure that the network being trained becomes more resilient/ robust to noisy data input and avoid over-fitting of the model. The reason why such data pre-processing was not performed in this study was due to the nature of the data. Due to the presence of EEG non-stationarities, the EEG is already inherently noisy or 'blurred'. Moreover data flipping or cropping would reduce the spectral separation of our target signals. Flipping the data of an 8.50Hz signal at a 50^{o} degree phase shift would indeed leave the frequency the same, however the phase offset of the data would no longer be consistent across targets of the same class. Effectively this would remove one of the data components that is integral to differentiating between the target signals.

Datasets for Classification

The raw EEG benchmark data was split into two separate sets. A subset of data consisting of EEG output collected during the fixation of 8 targets was collated. These targets were selected to ensure maximal neighbouring frequency separation to populate a dataset with the dimensions; 6 x 1500 x 1 x 1680. Development of a low computational cost implementation attempting to classify targets in the 8 class dataset allowed us to accelerate the initial debugging period as well as providing us with a successful model for the classification of targets in a less densely populated visual array. A secondary dataset consisting of all data present in the benchmark repository was formatted in the following dimensions; 6 (electrodes) x 1500 (samples) x 1 (channel) x 8400 (40 letters * 6 blocks * 35 participants).

Both datasets were spilt according to the following ratio: 70% training, 15% validation, 15% test. Training data refers to the SSVEPs which the network is iteratively learns to model, the validation and test data however were never used in the calculation of weight updates. The performance metrics associated with the validation and test data are analogous to an unsupervised learning procedure whereby the ground truth label is not made evident to the model.

The final AoC, mAoC, AP, MaP and ITR metrics are calculated exclusively in relation to the performance of the network when classifying the test dataset. The test dataset does not include any

data which the neural networks have previously been trained on. Further, these participants included in the test dataset where all BCI-naïve, meaning these individuals had no previous experience with BCI systems. These measure were taken to introduce some of the constraints that are present in the real-world decoding of brain-based bio-signals.

Training Hardware

8 Class network implementations were developed under MATLAB 2017a (MathWorks Inc.) using the MatConvNet convolutional network library (Vedaldi & Lenc, 2015). A template LeNet architecture for classification of the well-established 'mnist' hand-written digits dataset was used as a foundation for the creation of all subsequent models. However, substantial modification of said template code was necessary. The 2GB NVIDIA GTX 750 Ti GPU (1st generation Maxwell architecture) was utilized via CUDA toolkit 7.5 to train the 8 class network defined herein. 40 Class network implementations were trained via remote server on a 12GB NVIDIA Titan X GPU (1st generation Pascal architecture) running CUDA 8.0.

Network Features

The definitions populating this chapter are intended to provide the reader with a general understanding of the processes employed. A detailed mathematical definition of the machine learning operations is beyond the scope of this thesis. However, documentation expanding upon these processes is referenced alongside. Both models utilized the same network architecture, to view the network architecture as plotted using the MatConvNet visualisation function, refer to appendix D, Figure 7. Additionally, the 40 and 8 class network implementations utilized same meta parameters; batch size (5), momentum (0.9) and total number of epochs for training (1500). The models differ in terms of learning rate (LR) with rates of 0.00005 and 0.000001 applied to the 8 and 40 class models respectively.

Back-Propagation

Back-Prop is a highly popular method of optimizing network weights. It is essentially a means of applying updates to weight functions of individual neurones based on the error (delta) or distance between layer output and actual ground truth examples. A process of stochastic gradient descent is applied where labelled ground truth inputs are compared to current network node representations. The functional properties of the node are changed such that the output of the layer and the actual target data become more similar. The degree of change made to the node, based on the error calculation, is known as the Learning Rate (LR). Selection of an appropriate learning rate is critical. Too large and the potential for overshooting past the global minima (lowest error achievable by the model) is high. If set too low, the training period becomes longer and also the ability of the network to move out of local minima (periods in training where error reduces significantly, but not to the point of global minima) is also reduced (Schumacher, Roßner & Vach, 1996). Implementing a fixed learning rate can be problematic. It is theoretically preferable to have large initial jumps, followed by smaller more accurate steps towards the global minima when training neural networks. The development of a momentum value helped overcome this issue. Momentum functions to ensure that the adopted change in weight is proportional to the gradient of the previous change (Rumelhart, Hinton & Williams, 1985). Therefore the process of applying error based weight updating undergoes initial acceleration towards convergence followed by deceleration as the model approaches convergence such that overshooting is avoided.

Feature Extraction: Convolutional Layer

Fukushima (1979, 1980) pioneered a revolutionary data operation now known as a 'convolutional layer', which invariably constitutes the first layer of any CNN. This involves initially specifying a rectangular array (for example 3x3) constituting a 'window' with which to isolate and operate on subsections of input data. A weight vector (also known as a filter) is then applied to the data occupied by this subsection of input. The distance between windows is specified by the stride value, higher stride values are more computationally efficient as this reduces the total number of

operations carried out. Conversely, decreasing the stride effectively increases the degree of overlap between neighbouring windows (for example shifting the array just one data point at a time, as opposed to three at a time) increases the richness of representational information included. A stride value of 1 was maintained throughout all network implementations defined herein. Filters consist of an array with the same dimensions as the sliding window. The values within this filter are usually initialized with random numbers between -1 and 1, the same protocol was followed herein. The dot-product between the subsection of input data and the values in the filter array is computed and finally the weighted sum of this element-wise multiplication produces a single data point. The newly computed values populate an adjacent output array. Filters are conditioned iteratively over a series of training epochs to represent some feature of the input data. Convolutional layers positioned deeper within the network build up higher level representations allowing for the classification of more complex features.

The primary task of the networks developed in this research is to extract frequency information from SSVEPs embedded in EEG time series data. During the initial phase of neural network development, uniform 6x6 filter dimensions where selected. This filter size would move across all 6 channels sampled, as well as 6 data points from the initial raw input EEG data. As data the EEG time-series were sampled at 250Hz, this means a micro voltage value was recorded every 4ms. Therefore, a 6x6 filter would capture 24ms (6x4ms) worth of temporal EEG data. It was theorized that, given sufficient network depth and node density, that the models would develop representations of the gradient of the frequency stimulus patterns, as opposed to representations of the entire phase cycle of a stimulus pattern. These dimensions however proved unsuccessful, despite rigorous attempts to optimize meta parameters and modify pre-processing procedures.

The filters were therefore redesigned in order to capture the entire phase cycle of the respective stimulus patterns. The lowest target frequency assigned to a stimulus pattern was 8Hz. This means that peak-to-peak, the waveform could only be expressed in its entirety over the course of 32 samples. This is due to the fact that cycle of an 8Hz signal has a latency of 125ms. Therefore, the minimal data depth for representing a full 8Hz cycle, at a 250Hz sampling rate, is 32 samples. In order to accommodate for this, the first convolutional filters which received the raw EEG input data were set at the dimensions 6x30. This ensured that each convolutional filter contained roughly enough data for the expression of the lowest target frequency stimulation patterns. The height of the first convolutional filter was also maintained at 6. This was based off of the well-established process of data averaging in EEG analysis. Therefore, instead of actively pre-processing the data in this manner, information across channels was collapsed into a singleton dimension after the first convolutional operation, as evident from the network architecture diagram (see appendix D, Figure 7).

Normalization and Regularization: Batch Normalization

Batch Normalization aims to remedy what Loffe and Szegedy (2015) termed internal covariate shift. This refers to the phenomenon where deep neural network inputs transmitted between subsequent layers can change dramatically based on minor shifts in the distribution of training data. This problem is analogous to improperly randomising training data prior to learning. Initially, network weights shift to more accurately model input data and thus the parameters of the model rapidly become accustomed to the characteristics of the input. Minor changes in the input data can lead to significant reductions in convergence rate as parameters must go through an adjustment period to become accustomed to the new qualities of the input data. This is usually addressed by setting extremely low learning rates and thus results in long training periods. Internal covariate shifting becomes more problematic with increased network depth as more time is required to alter layer parameters further up the hierarchy. Batch normalization acts as a regularization layer in that filter values are re-distributed according to a population-wise view of the input data as opposed to a batch-wise view. This has the effect of allowing for higher learning rates and mitigates the need for the excessive application of drop-out layers. This technique has been successfully implemented in many networks, namely the GoogleNet and inception networks (Szegedy et al, 2015).

Drop-Out

These operational layers address the ever present issue of network over-fitting. During training, networks learn iteratively to classify input data. However, during this process it is necessary to avoid allowing the network to become adept at exclusively classifying input data. If the network over-fits to input data the generalizability of the network is diminished. Therefore, networks will only be able to accurately classify data from the set on which it is trained. Researchers ultimately intend for models to be applied to novel datasets and overfitting can mitigate or entirely impede this possibility. Drop-out layers function as a form of data augmentation, restricting the number of nodes on which weight updates are performed (Srivastava, Hinton, Krizhevsky, Sutskever & Salakhutdinov, 2014). The rate of 'drop-out' is traditionally set at 0.5, meaning only 50% of nodes in subsequent layers are included in the weight update process. This drastically increases the amount of time required to train the model as nodes are effectively trained at half the pace. By restricting the number of nodes included in the weight updates the diversity of data features to which nodes become responsive towards increases. By increasing the number of data features represented by network nodes it defends against the possibility of nodes becoming exclusively selective to features primarily found in the training set.

Down-sampling: Max Pooling

Customarily in CNN architectures contain an aggressive down-sampling layers known as max pooling (MP) layers. This can be conceptualised in a similar manner to the convolutional layer were a **non**-overlapping rectangular window isolates a subsection of the input array. The highest value in this window then populates a corresponding point in an adjacent output array. The size of this window will influence the resolution of the output data, with larger windows resulting in a smaller output array. The first applications of this operation can be found in the Cresceptron (Weng, Ahuja & Huang, 1992) and HMAX (Riesenhuber & Poggio, 1999) architectures. MP CNNs trained via backpropagation are arguably the most successful computer vision competition winners, (Schmidhuber, 2015). Other methods such as average pooling are available. These layers effectively create a mean of all values in a sliding filter window instead of simply identifying the highest value as per max pooling. This technique has been used successfully in recent CNN architectures (Szegedy, 2015), however the computational cost of such pooling is far higher, and therefore this technique was not selected in this study.

Activation Function: Rectified Linear Activation

In simplistic terms the purpose of an activation function is to classify incoming data as stimulatory or non-stimulatory. Duch and Jankowski (1999) identified over 500 such functions. In CNN architectural arrangements the activation step is primarily performed following a down-sampling operation. Output values from neurones occupying a preceding max pooling layer are 'squashed' into a specified range via a transformation function. This assists in differentiating between active and non-active nodes and ultimately helps to predict whether the original input data does or does not contain features that are relevant for classification. The activation functions are therefore crucial in the training of weights during back-propagation as they can be used to determine the error between targets and current output values.

Envisage the output of each node being squashed through a sigmoid function between the range of 0 and 1. The advantage of applying a non-linear activation function is their characteristic differentiation of values. In other words each node's activity is represented across a sigmoid output curve occupying a unique position and therefore has an individual corresponding error value. A major drawback in the application of sigmoid based activation functions is their characteristic 'saturation' of gradient when approaching 0 or 1. Weight updates in back-propagation are calculated using the gradient between total error demonstrated (x axis) and the activation of the node (y-axis). If training is successful the node will become highly selective, therefore resulting in a high probability of positioning on the 0 or 1 plateaus. In the later stages of training when applying this non-linear activation function to node outputs it leads to the clustering of many nodes along these plateaus. If the gradient calculated is minimal then the corresponding derivatives calculated to update the weights based on delta (error) scores will also be small. This

leads to the highly problematic phenomenon of convergence failure or 'vanishing gradient'. Other activation functions such as tahn and even modern function adaptations such as the 'leaky sigmoid' do not suffer the same vanishing gradient pitfalls, however the computational cost of numerous calculations featuring differentiable non-linear functions makes their application limited.

A viable alternative to the aforementioned activation functions are Rectified Linear Units (ReLUs), now the most popular for application in recent competition winning MPCNNs. The method was initially introduced by Hahnloser (2000), effectively applied and eventually endorsed by the neural network community (Glorot, Bordes & Bengio, 2011, LeCeun, Bengio & Hinton, 2015). ReLUs apply a **near** linear transformation on incoming node stimulation data and position them across an unbounded range. This means that lines plotted in this mapping space do not plateau, therefore eliminating the issue of the vanishing gradient problem. Moreover ReLUs do not feature calculations of exponential functions therefore they are significantly more computationally efficient. Krizhevsky et al (2012) demonstrated that deep nets trained using the ImageNet database (1.2 million images) reached 25% error performance 6 times faster when using ReLUs as compared to a tanh activation function counterpart. Therefore, these very same methods were used to deduce node activations in the networks defined herein.

Decision Function: Softmax

The final layer of a CNN is fully connected and functions in a similar manner to logistic regression by assigning a class value to each of the groups specified for classification. This class value, as per the softmax operation is calculated in a similar manner to the previously discussed convolutional layers (via dot-product). However the dimensions of the filter and subsequent output is restricted to the number of groups the model is trained to classify. For example in a binary classification task the final output would be represented by a 1x1x2 array. The SoftMax function increases the contrast in class values transforming them over the range 0:1. This assists in the differentiation between targets for tasks with a high number of classes. In its simplest form these final values indicate the degree to which the input has activated the learned representations of each group hidden within the network.

Summary

The networks are comprised of operational layers arranged in clusters consisting of a convolutional layer, a batch normalization layer, a drop out layer, a rectified linear unit and a pooling layer. The network consists of five such clusters with a final convolution and soft max layer. In the training configuration the network contains a total of 22 operational layers (refer to appendix D, Figure 7). In the testing configuration batch normalization and drop out layers are by passed as these are only necessary during the network training phase. Therefore in the testing configuration the network is comprised of just 16 operational layers.

Hypothesis

It is expected that the networks trained on the 8 class problem datasets will produce higher AoCs and APs in comparison to the 40 class problem networks. This is due to the lower correlational coefficients between target stimulation patterns in the 8 class set. The potential differences in bit-rate exhibited by prospective network implementations are problematic to deduce. 8 class networks may demonstrate higher AoCs, allowing for a heightened bit-rate in comparison to 40 class networks. However, the higher number of targets in the 40 class implementation could amplify AoCs beyond those of the 8 class networks.

Results

Alongside ITR, AoC and MaP the training graphs for each dataset are shown, as well as a confusion matrix detailing test results. The training graphs plot network error over training epochs. The error refers to a metric which is essentially the inverse of AoC, as it represents the probability of *incorrectly* predicting the letter users intended to spell. If a network demonstrated an error of zero, the network would be classed as 100% accurate, as it would have never provided an incorrect prediction. The term 'error', is used primarily in relation to the performance of the networks during training. The term AoC, is used primarily when commenting on the network performance in the test phase.

When using the soft max operational layer (final layer in the network architecture), all classes are assigned a probability value from zero to one indicating which target character the user is most likely to be fixating. Classes are then ranked using these values in ascending order from least, to most likely. The top 1 error refers to the probability that the correct letter is positioned at the top of this ranked list. In contrast, the top 5 error refers to the probability that the correct letter is positioned within the 5 highest ranked positions on this list. There is a higher likelihood that the correct prediction will lie within the 5 highest ranked classes, as opposed to occupying just the first class. Therefore, the top 5 error values are typically far lower for the majority of the training period.

The confusion matrix is a graphical representation of all predictions made by the networks for the test dataset and is used widely in the evaluation of network performance. The predictions made by the network for each class are cross-tabulated against the actual ground truth labels. In this study the ground truths relate to the character subjects were cued to fixate and the predictions refer to the classifications the network made based on the EEG time-series data. Confusion matrices are used in the evaluation of networks to see if certain classes are being consistently misclassified, or which classes are 'confused', as well as revealing the direction of this misclassification. For instance, classes with stimulus patterns which are similar in terms of frequency, such as a 8.0Hz and 8.2Hz signal, would likely be confused as they are more highly correlated, as compared to more distally positioned stimulus patterns, such as 8.0Hz and 15.0Hz.

Confusion matrices are often coloured in order to communicate these relationships more clearly, these coloured figures are often referred to as heat maps. The MatLab heat map function used to generate the figures in this work uses a blue-red colour scale, with red indicating a higher number of class selections and blue indicating a low number of class selections. Each value represents a percentage of the total number of predictions made. Therefore, when all values in the heat map are summed it totals 100. The intersection of each class shows the percentage of correct predictions made for that class. For example, the first row and first column for both heat maps herein shows the percentage of letter A predictions, when the user was actually fixating the letter A.

The optimum percentages each of these class intersections for the 8 class problem is 12.5, as 100/8 = 12.5. The optimum percentages for the 40 class problem is 2.5, as 100/40 = 2.5. Any percentages shown to lie outside the diagonal points of target and ground truth intersection represent misclassifications, or points of class 'confusion'. The heat maps herein can be used to illustrate network accuracy by examining the 'redness' of cells positioned along the class **intersections**. A solid red colour at these intersecting points indicates that the classes were selected only when the corresponding targets were fixated. Network precision can be inferred by examining the degree of confusion for each class. The number of cells which are *not* presented as solid blue and are *not* positioned on the points of class intersection indicate the network was not operating with maximal precision.

8 Class Results

After 1500 epochs of training, the model shows a top 1 **error** of 0.01%, with a top 5 error of 0%. Additionally, the model produced a top one **validation** error of 22.22% (refer to Figure 3) and a top five error of 0.06%. ITR calculations using the final test mAoC (82.94%) revealed a bit rate of 19.74 bits per minute.

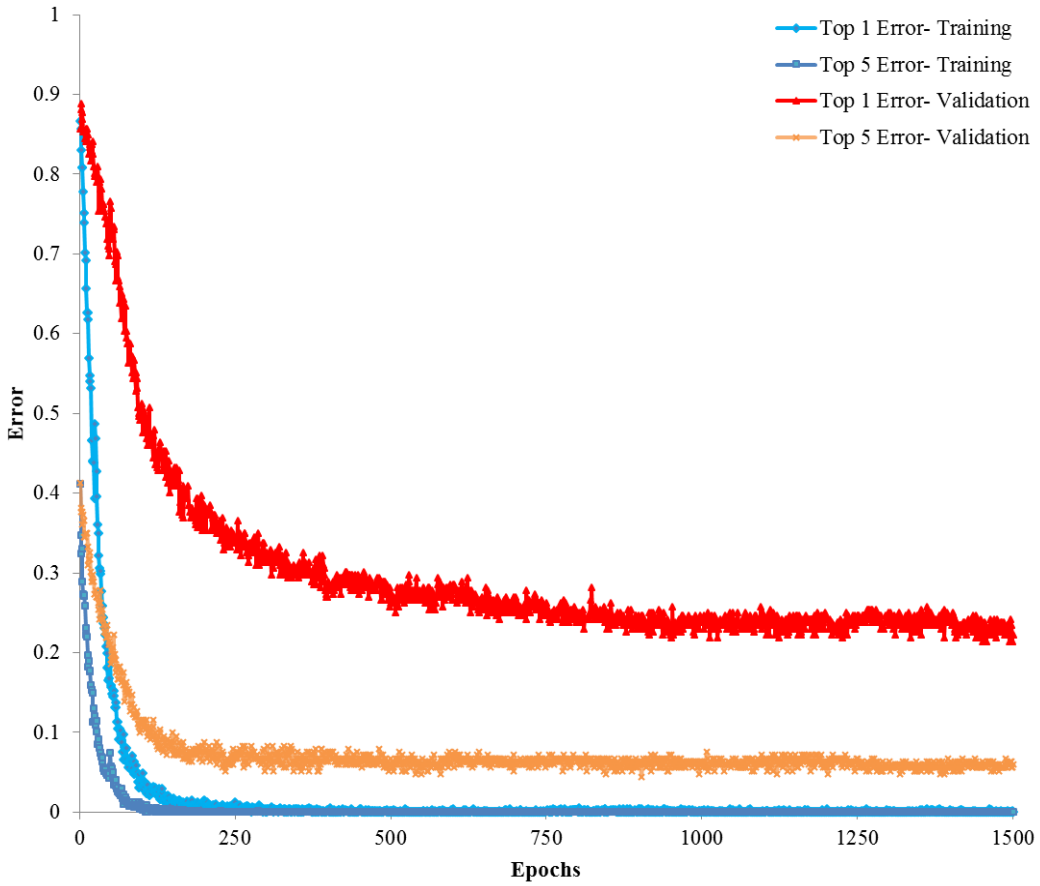


Figure 3. This graph illustrates the decrease in classification error throughout the training period for the 8 Class network in terms of both the training and validation datasets. Classification error is an inverse measure of AoC, indicting the probability of incorrectly selecting the actual target a user was cued to fixate. The top 1 error refers to the likelihood of the target class being ranked as the most probable prediction estimated by the network. The top 5 error refers to the likelihood of the target class being ranked in the top 5 predictions out of a possible 40 classes.

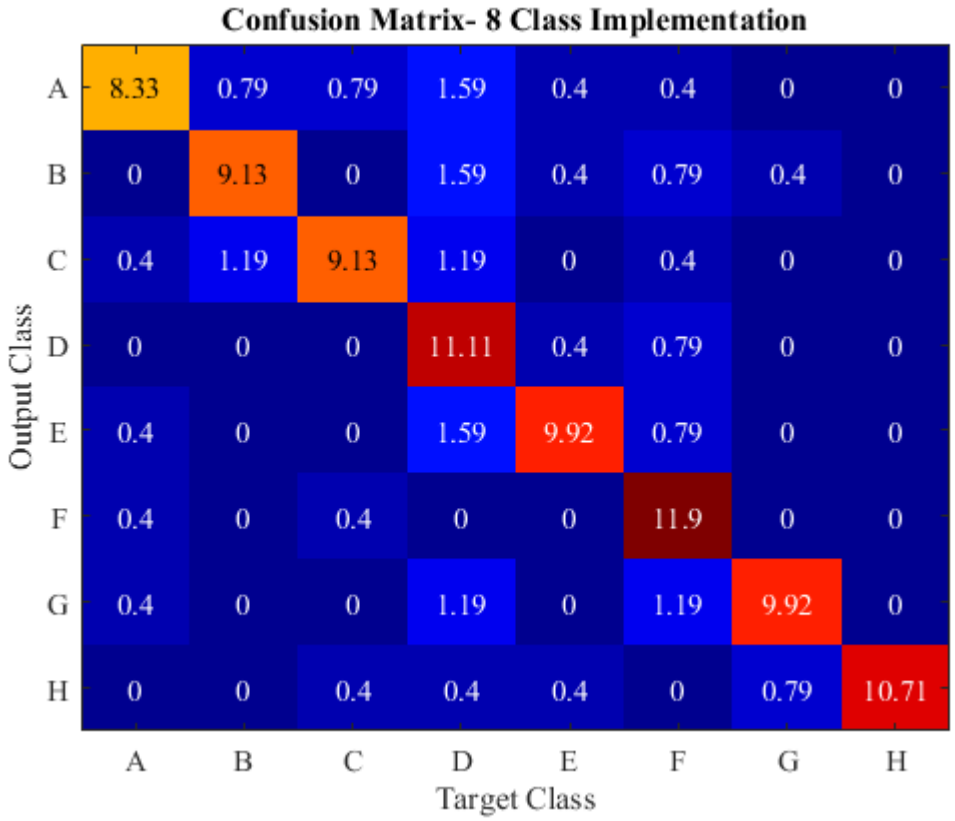


Figure 4. This is a heat map populated with percentages which relate to the cross-tabulation of network predictions (y axis) and actual ground truth labels (x axis). A shift of colour from blue to red indicates a general decrease in classification error. For more detailed information refer to appendix A, Table 1.

As shown in the heat map, there is an even distribution of misclassification errors across the target stimulation frequencies (refer to Figure 4). However target class D demonstrates a particularly low average precision (0.59). Class D was selected as the target prediction in substantially more instances than other target classes which has evidently inflated the corresponding target AoC percentage (90.32%). In contrast target class H was classified with 100% precision. The mean average precision of the network across all classes was calculated at 0.82. The lowest AoC observed is associated with target A at just 67.74%, whereas the highest AoC is attained by target F (93.75%).

40 Class Results

After 1500 epochs of training, the model's top 1 training error of 0.01%, with a top 5error of 0%. Additionally, the network produced a top one validation error of 16.82% (refer to Figure 5) and a top five error of 1.08%. ITR calculations utilized metrics from the performance of the model in relation to the test dataset (mAoC= 82.94%) and revealed a bit rate of 21.22 bits per minute.

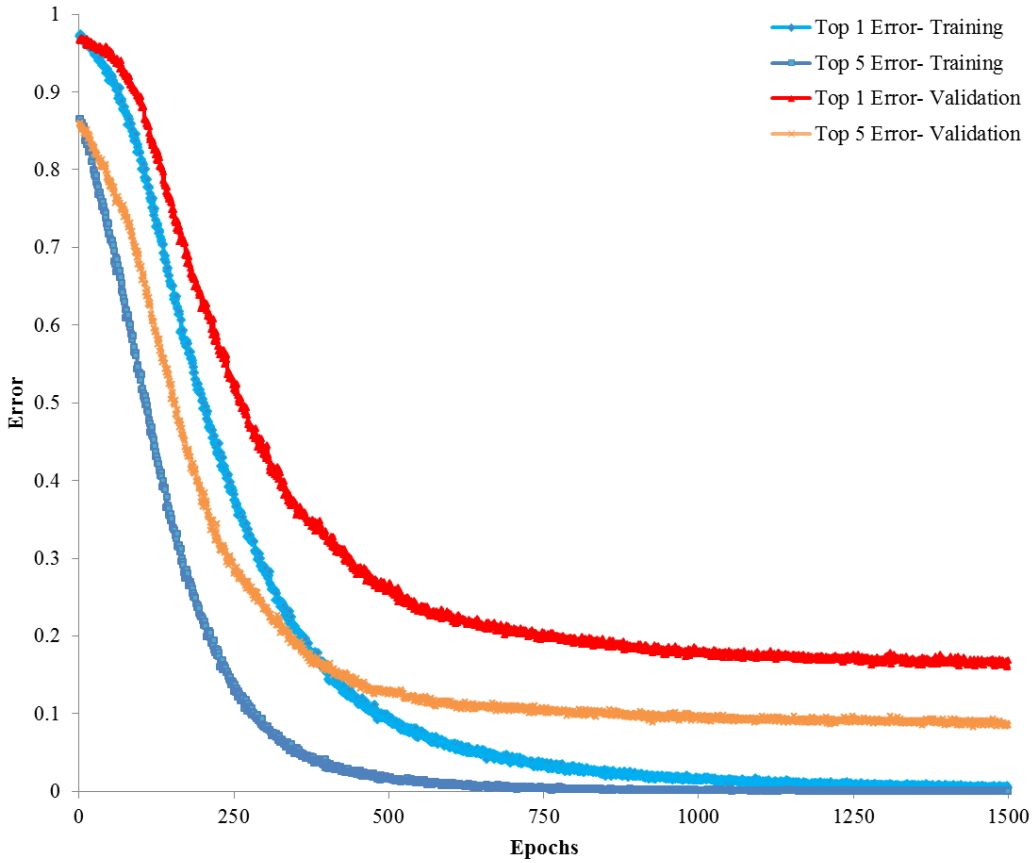


Figure 5. This graph illustrates the decrease in classification error throughout the training period for the 40 class network in terms of both the training and validation datasets. Classification error is an inverse measure of AoC, indicting the probability of incorrectly selecting the actual target a user was cued to fixate. The top 1 error refers to the likelihood of the target class being ranked as the most probable prediction estimated by the network. The top 5 error refers to the likelihood of the target class being ranked in the top 5 predictions out of a possible 40 classes.

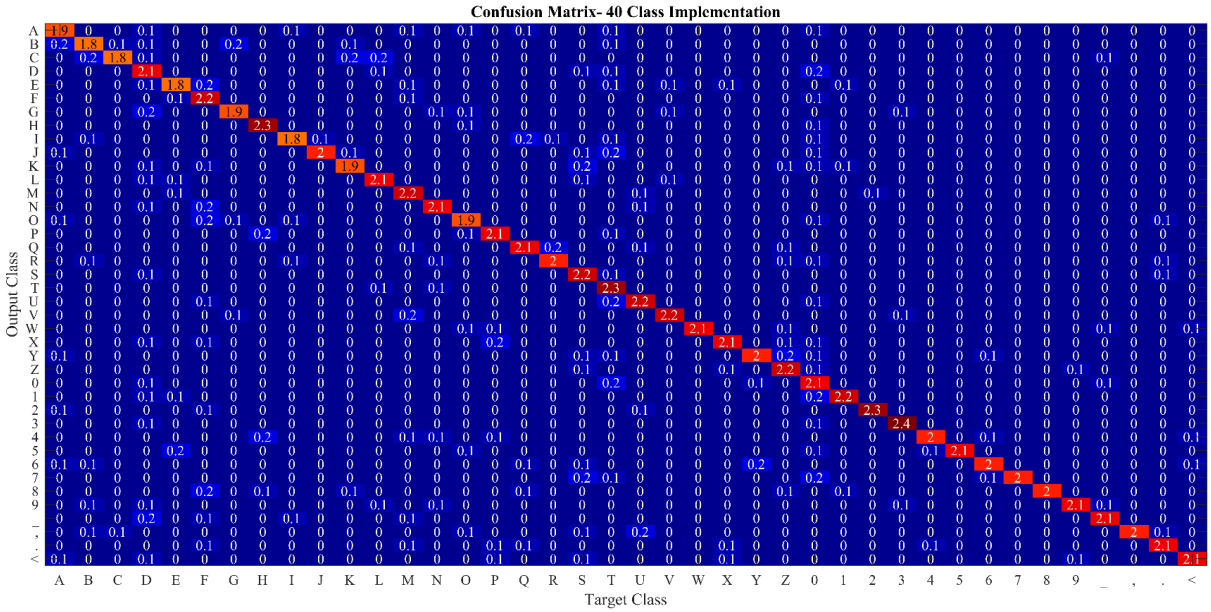


Figure 6. Heat map showing the distribution of predictions (represented by percentage) made by the 40 class network implementation for each target class during the test period. For the purposes of presentation the table has been reduced, a larger reference can be found in appendix C. More detailed information on the relevant data values presented in appendix B, Tables 2a-2d.

The confusion matrix indicates that the instances of misclassification are evenly distributed across the alternative stimulation patterns utilized (see, Figure 6 & appendix C). The lowest AoCs reported herein corresponded to targets B, C, E and I with all respective stimulatory patterns yielding 74.19% AoCs. Targets D (11.00Hz, 4.71 radians, AP= 0.58) and 0 (10.60Hz, 3.14 radians, AP= 0.57) represent the lowest average precision (AP) metrics calculated across the dataset (refer to, appendix B: Tables 2a-2d). Both values are substantially lower than the mean average precision of 0.85 (k=1). The highest precision values achieved are related to targets: 5 (15.60Hz, 4.71 radians), 7 (9.80Hz, 2.57 radians), 8 (10.80Hz, 3.14 radians) and ‘,’ (13.80Hz, 1.57 radians) all of which reached an AP of 1.00.

Discussion

Results Summary

Critically, all metrics defined herein represent the results of test data classified via the aforementioned networks. Participant data populating the test set were used exclusively in this regard. Meaning at no point was data from these participants included in the training process. Moreover, the participant data populating both the validation and test sets were classified by the authors of the Tsinghua benchmark dataset as 'unexperienced'. Therefore the results herein represent the closest approximation of a real-world implementation as feasibility possible with the data available. From the results it is evident that the network implementations were mostly successful. The 40 class implementation produced higher test mAoC (82.94%), MaP (0.85) and ITR (22 bpm) values when compared to the 8 class implementation (mAoC= 80.15%, MaP= 0.82, ITR= 19.74 bpm). This may initially seem counter-intuitive as the 8 class dataset, contained waveforms with distinctly lower between class correlational coefficients than the 40 class dataset. However, in light of previous studies which employ correlation based analysis, coefficients do not give a true indication of data similarity. It is therefore clear that the two datasets contain similar levels of latent variables such as; harmonic frequency components which are weighted more heavily in the accurate classification of corresponding character waveforms.

Accuracy of Classification

Critically, AoC did not drop below 70% for any of the 40 class targets. Targets yielding the highest AoC values were H (15 Hz, 4.71 radians) and T (11.4 Hz, 4.71 radians) suggesting larger phase offsets could boost network capabilities for target differentiation (refer to table 2a-2d, in appendix B). Moreover, there appears to be a trend for higher misclassification of target frequencies with integer value, however this is likely circumstantial. Interestingly, all targets possess either 0 or 1.57 degree phase offsets. It could be argued that at these lower phase offsets increase the probability of neighbouring target stimulation patterns sharing harmonic components, therefore increase the chance of misclassification. However this did not significantly influence the AoCs or precision metrics of numerous other targets sharing similar properties.

Differences in the AoCs between the 8 and 40 class implementations can be attributed to unique learning rates. The 40 class implementation was assigned a learning rate (LR) of just 0.000001, as compared to the 8 class LR of just 0.00005. This has substantially increased the generalisability of the model, as evident from the comparatively lower difference between training, validation and test AoCs of the 40 class network. Furthermore, the higher performance metrics of the 40 class model could also be attributed to the patently larger dataset used in training. Providing more ground truth data inevitably decreases classification error as the model is able fine tune meta-parameters to a broader palate of real-world examples by enhancing the number of EEG waveform representations embedded in the model. Despite the test mAoC for the 40 class implementation being greater than the research standard of 70%, around 1 in every 5 words intended for selection will represent a misclassification error. Frequent character misclassification would undoubtedly reduce user workflow during device interfacing. The standard procedure when designing SSVEP based BCI stimuli is to restrict the frequency range of potential target patterns to between 8-15.5Hz. This is due to a large body of past research indicating that this sub-band in frequency space provides optimal signal to noise ratio across a normally distributed population. Despite these findings application of such target frequencies can either isolate or endanger some potential end-users due to the risk of inducing epileptic seizures. Current research however suggests that utilization of higher frequency values (>15 and <70Hz) can nullify the inherent risks of epileptiform activity. Additionally, these stimulation patterns produce EEG output with higher signal stability and also dramatically reducing participant perceptions of stimulus induced discomfort (Won et al, 2016). Moreover, this increases the range of frequency values available when selecting target stimulus patterns, allowing for large incremental differences between neighbouring frequency values. The inevitable dramatic decrease in resulting SSVEP waveform similarity between targets would undoubtedly increase classification accuracies. Also it is important to note, the networks defined herein represent the **bottom-end** of the theoretical deep

CNN classification capabilities. User data from continual online use could be used to train the network, effectively tailoring it to the specific EEG characteristics of the patient. Real-time performance using similar EEG device configurations implementing these networks online is necessary to fully validate these results.

Information Transfer Rate

The ITR reported herein is comparatively low when considering recent BCI speller research. This can be attributed to the 6 second data capture period. Previous studies have explored the possibility of modulating data capture periods in real-time in response to user performance (Yin, Zhou, Jiang, Yu & Hu, 2015). This technique has the benefit of retaining high classification accuracies independent of user processing speed. The optimization can be performed prior to substantial online analysis by collecting a small sub-set of user SSVEP EEG signals. The prevalence of necessary data features is then calculated and an optimal data capture period is estimated. This initial assessment functions to speed up the system-user calibration process, allowing more time for task training prior to online-usage. This process of optimizing the data acquisition period duration has also been performed online to ensure changes in performance throughout the interfacing period do not impede classification accuracy (Kha, Nguyen, Kha & Dutkiewicz, 2017). Optimization of algorithms which dictate these parameters could potentially be achieved in real-time through the implementation of a fairly shallow CNN.

The performance of the 8 and 40 class networks were not tested using shorter time windows as the CNN architecture requires input data of consistent dimensions. This is a potentially limiting factor to the maximal ITR attainable by such networks. A potential means of circumventing this obstacle would be to train 6 separate neural networks using iteratively larger input dimensions (1:6 seconds). During the training phase a model configured for large input dimensions corresponding to 6 seconds of data capture could be utilised. If consistently high classification accuracies are achieved, the use of a model accepting smaller input arrays could then be substituted as the SSVEP decoder. Decreases in classification accuracy due to user fatigue or low SNR would trigger the utilization of networks tuned for larger data capture windows in order to maintain functionality. A more sophisticated solution however would be to implement a hybrid recurrent-convolutional neural network. Traditionally Recurrent Neural Networks have been used in the domain of semantic language processing and speech recognition. RNNs are arguably more flexible architectures as they are capable of accepting data input sequences of varying length. Development of a hybrid CNN-RNN was beyond the scope of this thesis; however this is clearly a potentially fruitful investigation for future research. Additional methods of increasing BCI speller ITRs include the integration of language models to improve user performance and experience. Techniques such as: auto-correct and auto-completion have been studied extensively (for review see, Speier, Arnold & Pouratian, 2016). However, there have as yet been no efforts made to integrate emoticons into BCI speller visual target arrays. The potential for increasing ITR as well as providing BCI speller users more tools to add directional valence to their printed text is substantial. Others means of embellishing the printed text from BCI spellers with emotional content include digital vocalisation of the decoded characters. Arguably, the ability to introduce more emotional context to the text printed represents the next challenge in the development of BCI spellers.

Confusion Matrix

When inspecting the confusion matrices of both 8 and 40 class implementations, both indicate a preponderance for selection of target class D leading to distinctly lower average precision values. This may seem abnormal in consideration of the fact that the stimulation frequency (11Hz) is positioned centrally in the optimal range for low noise SSVEP signals. However, as previously discussed it is not uncommon for there to be significant individual differences in relation to the responsivity of participants to certain stimulation frequencies. The test participant population may contain a number of individuals sharing a distinctly lower responsivity to the 11Hz frequency. This therefore suggests that the EEG data collected during fixation of the class D target may contain significantly lower quality SSVEP signals. Increased noise can result in the random expression of data features typically associated with other target waveforms, therefore increasing

misclassification. Alternatively, this phenomenon may be product of the target's positioning on the computer monitor which displays both the flickering targets and character predictions after each 6 second stimulation period. The D character occupies a central position neighbouring the region of the screen which prints visual feedback (characters predicted). During the task, participants may use region of the screen occupied by letter D as a default resting position. This would explain the over-classification of letter D due to the inclusion of a small but potentially relevant amount of SSVEP signals generated in response to the flickering of D during the initial moments of the flicker period. This could be remedied by instructing participants to avoid this behaviour and placing a fixation cross centrally while monitoring the deviance of gaze during the inter-trial period using an eye-tracker. Another possibility is to retrain the networks defined herein on the same datasets after excluding the first 100ms of data. This would ensure that unnecessary data associated with feedback monitoring and target visual search would not be feed into the model.

Ultimately, irrespective of the methods it is imperative that each operation potentially executable by the user possesses the same likelihood for selection. This is critical in order for user work flow to be consistent during device interfacing. Such functionality would foster seamless application of the BCI in real-world applications. Interestingly, the results reported herein suggest the integration of more targets in the visual array is feasible as the jump from 8 classes to 40 classes did not result in a net decrease of performance statistics. It is clear however that in order to accurately classify SSVEP data from densely populated targets arrays a substantial training set is required. Inclusive to the current Tsinghua dataset the application of data augmentation techniques was overlooked. Authors outline a procedure whereby the phase off-sets applied to the data can be reversed, therefore feasibly a larger range of phase offset values could be utilized to produce significantly more frequency and phase pairings for additional targets. This would necessitate the repeated usage of the stimulus frequencies employed, however with sufficient phase offsets the correlational coefficients could be reduced to similar levels. In addition, to increase network resilience subsets of the data could be augmented by applying noise prior to training. This would ensure that the models maintain high rates of classification during prolonged periods of use. Specifically this added robustness would be useful when networks are paired with wet-EEG systems as SNRs have been shown to decrease when conductive gel begins to evaporate.

Tsinghua Benchmark Data

The hardware and meticulous procedure employed by the developers of the Tsinghua benchmark dataset (Wang et al, 2016) were key to the acquisition of the performance metrics reported herein. The use of a high fidelity wet-EEG system in low electrical interference conditions lead to the production of EEG data with minimal noise and movement artefacts. This undoubtedly assisted in the definition of network parameters sufficient for the differentiation of target waveform patterns. The use of wet-systems in real-world user settings may be impractical as often target patients are housed in clinical environments which are densely populated with devices producing electrical interference. Often such devices are necessary to assist patients in critical processes which they are no-longer able to sustain independently. Consequently, the removal of these machines from the environment is not a realistic solution. Therefore, effective implementation of BCI spellers in a real-world context requires substantial advances in the development of EEG sensors with more protective shielding. Despite the execution of high quality data collection by the Tsinghua research group, the open source data available was down-sampled prior to release. Authors justify this pre-processing by asserting that this step removes a significant amount of high frequency noise from the samples (Wang et al, 2016). However, it would have been interesting to explore the classification capabilities of a network implementation trained on this higher resolution dataset. In future, in-house data collection using dry mobile kits (Cognionics Quick 20) at higher sampling rates (1kHz) will be performed. The inevitable increase in EEG noise may be offset by the higher resolution of the EEG data input. Outcomes of such experimentation are difficult to predict due to difference in data collection hardware, pre-processing software and participant experience level. However, the potential maintenance or increase in classification capabilities in concert with increased kit mobility would produce a BCI speller with a higher degree of flexibility.

Meta Parameters

Through a process of experimentation it became evident that lower batch sizes significantly reduced the tendency for models to over-fit as well as increasing the final top one validation accuracy. Traditionally, batch sizes (the number of training examples included in the weight update calculations) are set at around 20, constituting a typical memory capacity of around 1MB. However, this is of course entirely arbitrary in relation to the format of data used as the initial network input. A reduction in batch size contingently increases in the amount of processing time required per epoch. Training at lower batch sizes has also been found to significantly improve final validation error results, however as previously alluded this comes with a higher price in terms of research time. All networks including were trained with batch sizes of 5. As reported by Breuel (2015), the combination of low batch sizes with; ReLUs, convolutional layers and softmax output operations in the context of deep network architectures produced the highest performance metrics across a number of classic image processing benchmark datasets. It is however clear that further manipulation of the meta-parameters could lead to the acquisition of significantly higher mAoCs. However, the aim of the project was to achieve >=70% mAoC across all 35 participants for a 40 class problem, therefore this took priority in terms of research time allocation.

With a reduction in batch size, typically a reduction in the learning rate is also necessitated. The inordinately low learning rates of 0.00005 and 0.000001 were used in the 8 and 40 class implementations respectively. This step was critical in order to avoid over-fitting across the learning period to the training data set. Additional measures to mitigate overfitting include the deployment of a drop out layer after the first convolutional layer. Interestingly, despite recommendations from the developers of the batch normalization technique, the inclusion of a solitary drop out layer improved mAoC dramatically in both the 40 class and 8 class implementations. However, the addition of further drop-out layers to the model had a negative impact on the global minima attained. This is likely due to the drop-out layers reducing the ability of ReLUs to convey the incidence of node stimulation in relation to the input data. Such a phenomenon is common when combining both drop-out and ReLUs and forms the basis of the argument set out in Loffe & Szegedy (2015) as to why inclusion of drop-out layers in tandem with batch normalization is unadvisable.

Network Improvements

Network Initialization: Unsupervised Pre-Training

In order to speed up the rate of convergence researchers have shown the utility of deploying networks on a simplified unsupervised learning task for a low number of epochs prior to the primary classification task. The process involves pre-training each layer to represent a non-linear transformation of the outputs from the previous layer, characterising the primary differences of the input. This essentially acts as a form of regularization, by reducing the variance in node values across the network while encouraging expedient arrangements of parameter space (Erhan et al, 2010). Unsupervised pre-training has been employed in a number of the most successful image classification and segmentation CNNs to date (Krähenbühl, Doersch, Donahue & Darrell, 2015). Authors of the aforementioned models attest to the capacity for this form of pre-training to ultimately reduce global minima and the instance of model over-fitting (Jarrett, Kavukcuoglu, LeCun, 2009). Future optimization of the networks should consider employing this means of parameter initialization.

Local Response Normalization

The regularization method implemented in the networks defined herein was a by-product of the batch-normalization layer operations. Regularization protocols have also been implemented more directly in the form of discrete normalization layers such as Local Response Normalization (LRN). LRN attempts to engender the characteristic lateral inhibition (LI) expressed in biological neurones in artificial networks. LI defines the process whereby the net-neurone activity reduces in response to increased activity of neighbouring neurones. Moreover, excitation of individual neurones in the absence of neighbouring activity leads to significantly higher levels of activity in the said neurone. This effect has been demonstrated countless times in the mammalian visual

system and has been shown to drastically improve the perceptible differences in fundamental features of visual space such as texture and edge. LRN layers in the context of artificial neural networks therefore function to amplify the detectable differences in node activity as the data moves through the network. Previous studies have asserted the efficacy of implementing LRN layers specifically when also utilizing ReLUs due their un-bounded activations. Replacing batch normalization with LRNs would undoubtedly increase the total training time required to meet convergence. However this may come with the benefit of a lower global minima. Experimentation into finding an arrangement whereby both forms of normalization are utilized by the model may permit the retention of positive characteristics from both.

CNNs for Pre-Processing

The use of CNNs as a data-pre-processing tool has been previously established (Wulsin, 2010). During the collection of EEG data, despite some models being marketed as 'mobile' EEG signal quality deteriorates rapidly when electrodes experience friction and un-alignment with the scalp. Movement artefacts may be less of a contributing factor when in use by real-world target patient populations. However, during the intermediate research stage where data collection from healthy participants is necessary to assess network efficacy the implementation of CNNs as movement artefact tools could increase the yield of usable data per testing session and thus increase the rate of research. Moreover, these same forms of pre-processing could eventually be utilized as eye blink removal software during online use by quadriplegic patients.

GPU Deployment

Critical to attaining >70% mAoC on both implementations of the network was the utilization of GPUs during the training period. Specifically in relation to the 40 class implementation, without the use of the aforementioned NVidia Titan X, training a model of the network described herein would have been unachievable. The 40 class implementation is arguably a success, however the memory demands to run a single test character are significant. This therefore places added cost to the BCI hardware used to decode the SSVEP bio-signals. However, with significant advances in the field of computational processing and efforts from technology retailers to market tailored CNN training hardware into convenient bundles means this limitation should not represent a major obstacle.

Real-World Application

Despite impressive results the scope of application for SSVEP based BCI spellers are restricted due to a number of factors. Primarily, individuals suffering visual defects from traumatic, congenital or progressive disorders would not be able to utilize such systems. Attempts at developing BCIs which harness related bio-signals such as Steady State Auditory Evoked Potentials, or Steady State Somatosensory Potentials (Hori, 2017) have been explored. However, information transfer rates of such systems are significantly lower than current vision based BCI spellers. Critical to the assessment of SSVEP based BCI speller applicability is an investigation of SSVEP gaze independence in the context of a dense target array. As previously mentioned SSVEP gaze dependency is a divisive topic in the field of BCI. Establishing the level of real-time performance attainable with the appropriate gaze fixation controls is crucial to understanding the upper limits of the SSVEP bio-signal as a means for powering BCI spellers. The powerful computational methods defined are highly noise resilient as evident from the results and may represent the best methods of developing a high speed purely attention based BCI speller.

Moreover, research has demonstrated that there are age related differences in SSVEP based BCI performance (Volosyak, 2017). This has been attributed a number of factors; reduced learning rates, maximal performance speed and bio-physiological differences. Expanding upon the later, EEG requires high conductivity between the electrode and the surface of a user's scalp. As human users age, the reduction of collagen in the skin attenuates the signal quality received during EEG data collection. This obstacle is not as exaggerated during collection of wet-EEG systems data, as compared to dry-EEG systems. Recently so-called semi-dry electrodes have been developed in an attempt to capture the benefits of the aforementioned systems while minimising the disadvantages. Yang et al (2016) have created sensors constructed from porous ceramic. The absorbent properties of the materials utilized allow researchers to load sensors with small quantities of conductive gel

which slowly release across the areas where electrodes contact the user's scalp. This ensures that the signal instability inherent to dry-EEG systems is overcome. However it is questionable to what extent to the impracticalities of repeatedly applying conductive gel have been superseded. User discomfort would indeed be improved, however not entirely ameliorated. As mentioned previously the dry-mobile-EEG systems currently present the most practical format for long-term BCI spellers. It may be possible through the application of more powerful classification techniques as described herein to overcome these inherent age-related performance decreases. Nevertheless, long-term ecological studies are required to validate such potentialities.

Conclusion

In conclusion the results herein define a deep CNN for SSVEP based BCI with the potential for plug-and-play functionality. To reiterate, the aim of this study was to produce a CNN capable of =>70% classification accuracy on all targets occupying a dense visual array. Critically, this was achieved in the absence of **any user specific** parameter optimization. However, it is important to note that the application of deep CNNs for the classification of SSVEP based bio-signals is still in a period of infancy. In order for researchers to continue advancing these methods of analysis larger benchmark datasets captured at higher resolutions are required in order to provide the resulting networks with increased generalisability. Additionally, researchers should begin using stimulation patterns harnessing a larger range of the frequency spectrum to reduce user fatigue and target waveform data similarity. Moreover; network parameter optimization, the integration of the unsupervised pre-training weight initialization procedures and local response normalization layers are critical to enhancing network accuracy and precision. Finally, an evaluation of SSVEPs gaze-dependency in the context of a dense target array is crucial to establishing the applicability of such systems.

Appendix

Appendix A

Table 1

8 Class Network Performance Metrics

Character Index	A	B	C	D	E	F	G	H
Frequency (Hz)	8.00	9.00	10.00	11.00	12.00	13.00	14.00	15.00
Phase (radians)	0.00	1.57	3.14	4.71	0.00	1.57	3.14	4.71
Accuracy of Classification	67.74	74.19	74.19	90.32	78.12	93.75	78.12	84.37
Average Precision	0.84	0.82	0.85	0.59	0.86	0.73	0.89	1.00

Appendix B

Table 2a

40 Class Network Performance Metrics

Character Index	A	B	C	D	E	F	G	H	I	J
Frequency (Hz)	8.00	9.00	10.00	11.00	12.00	13.00	14.00	15.00	8.20	9.20
Phase (radians)	0.00	1.57	3.14	4.71	0.00	1.57	3.14	4.71	0.00	1.57
Accuracy of Classification	77.42	74.19	74.19	83.87	74.19	90.32	77.42	93.55	74.19	80.65
Average Precision	0.75	0.77	0.92	0.58	0.79	0.67	0.86	0.83	0.85	0.96

Table 2b

40 Class Network Performance Metrics

Character Index	K	L	M	N	O	P	Q	R	S	T
Frequency	10.20	11.20	12.20	13.20	14.20	15.20	8.40	9.40	10.40	11.40
Phase	3.14	4.71	0.00	1.57	3.14	4.71	0.00	1.57	3.14	4.71
Accuracy of Classification	77.42	87.1	90.32	87.1	77.42	83.87	83.87	80.65	90.32	93.55
Average Precision	0.83	0.84	0.76	0.84	0.77	0.81	0.79	0.89	0.68	0.66

Table 2c

40 Class Network Performance Metrics

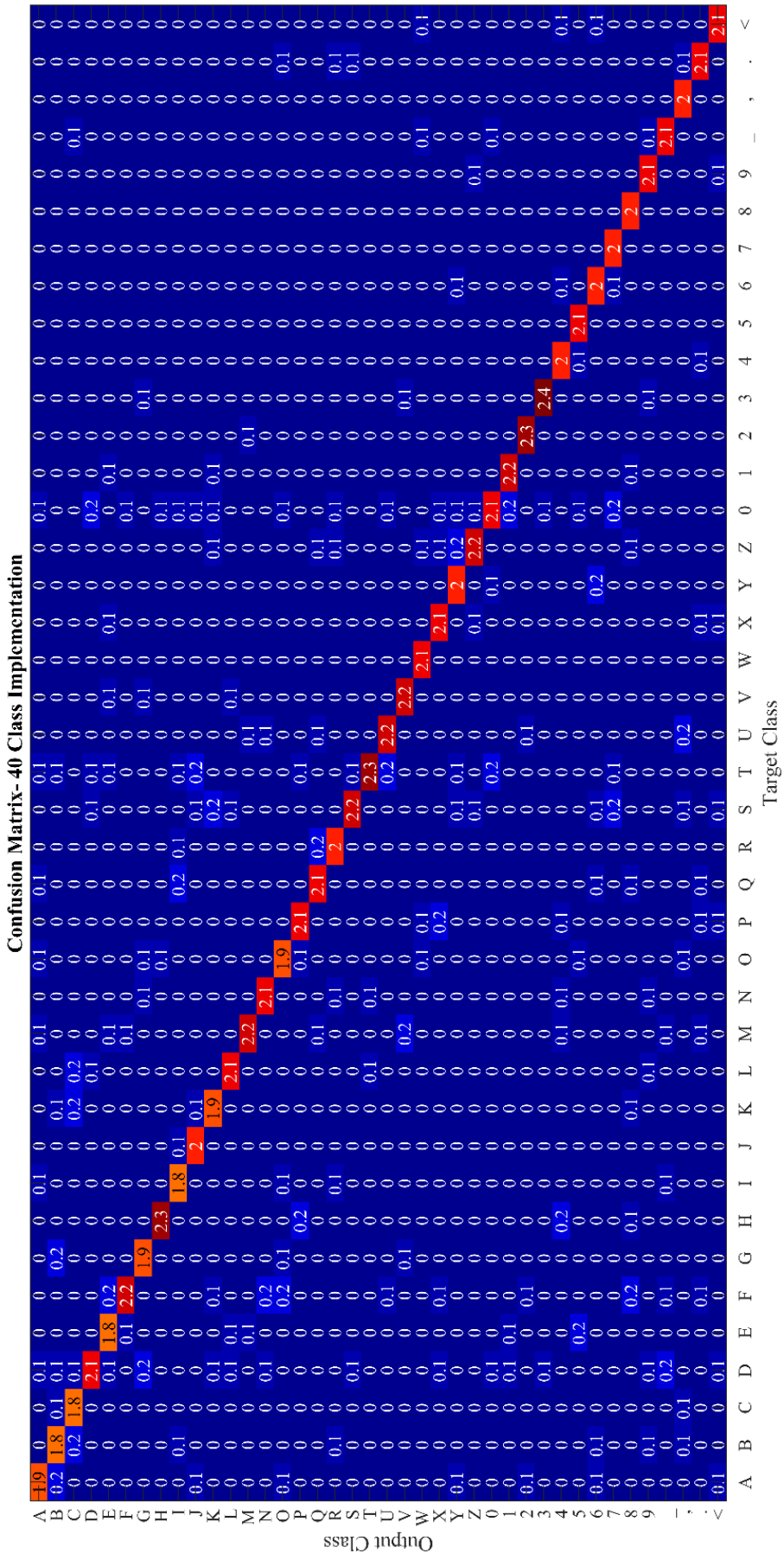
Character Index	U	V	W	X	Y	Z	0	1	2	3
Frequency	12.40	13.40	14.40	15.40	8.60	9.60	10.60	11.60	12.60	13.60
Phase	0.00	1.57	3.14	4.71	0.00	1.57	3.14	4.71	0.00	1.57
Accuracy of Classification	87.50	87.50	84.38	81.25	78.13	87.5	84.38	87.50	90.63	93.75
Average Precision	0.82	0.90	1.00	0.87	0.89	0.78	0.57	0.90	0.97	0.91

Table 2d

40 Class Network Performance Metrics

Character Index	4	5	6	7	8	9	_	,	.	<
Frequency	14.60	15.60	8.80	9.80	10.80	11.80	12.80	13.80	14.80	15.80
Phase	3.14	4.71	0.00	1.57	3.14	4.71	0.00	1.57	3.14	4.71
Accuracy of Classification	78.13	84.38	78.13	78.13	78.13	81.25	84.38	78.13	81.25	81.25
Average Precision	0.93	1.00	0.89	1.00	1.00	0.93	0.87	1.00	0.87	0.90

Appendix C


Confusion Matrix- 40 Class Implementation
Output Class
Target Class

Appendix D

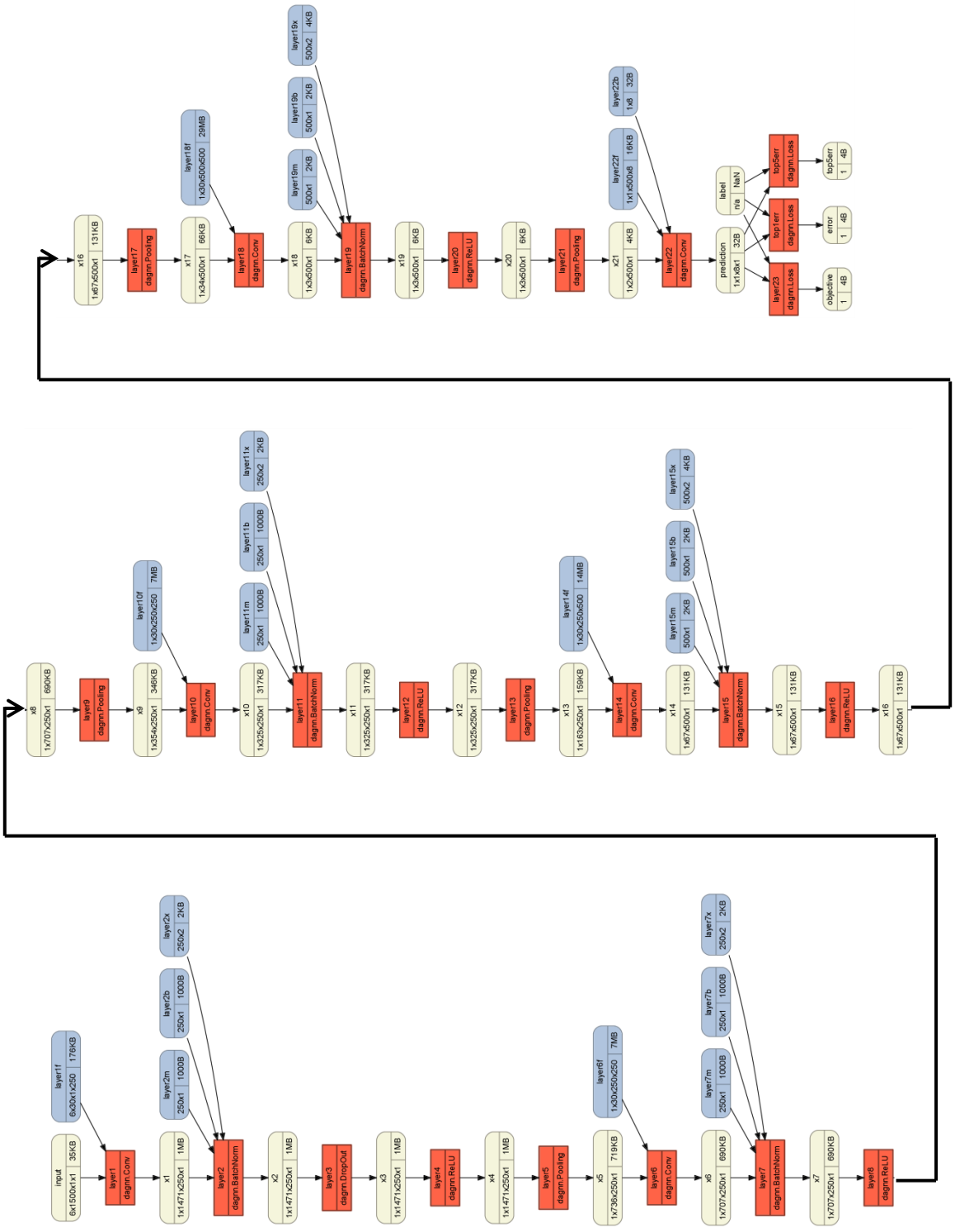


Figure 7. Neural Network architecture diagram displayed using relevant MatConvNet print functions.

References

Abbott, W. W., & Faisal, A. A. (2012). Ultra-low-cost 3D gaze estimation: an intuitive high information throughput compliment to direct brain–machine interfaces. *Journal of neural engineering*, *9*(4), 046016.

Ahmed, R., & Rutka, J. T. (2016). The role of MEG in pre-surgical evaluation of epilepsy: current use and future directions. *Expert review of neurotherapeutics*, *16*(7), 795-801.

Ahn, M., & Jun, S. C. (2015). Performance variation in motor imagery brain–computer interface: A brief review. *Journal of neuroscience methods*, *243*, 103-110.

Alcaraz, G., & Manninen, P. (2017). Intraoperative electrocorticography. *Journal of Neuroanaesthesiology and Critical Care*, *4*(4), 9.

Alkhalili, K., Niranjan, A., & Engh, J. (2016). Preoperative Magnetoencephalography Improves Tumor Resection Safety in Awake Craniotomy: Our Initial Experience. *Journal of Neurological Surgery Part B: Skull Base*, *77*(S 01), P057.

Allard, U. C., Nougarou, F., Fall, C. L., Giguère, P., Gosselin, C., Laviolette, F., & Gosselin, B. (2016, October). A convolutional neural network for robotic arm guidance using sEMG based frequency-features. In *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on* (pp. 2464-2470). IEEE.

Allison, B. Z., McFarland, D. J., Schalk, G., Zheng, S. D., Jackson, M. M., & Wolpaw, J. R. (2008). Towards an independent brain–computer interface using steady state visual evoked potentials. *Clinical neurophysiology*, *119*(2), 399-408.

Amato, F., López, A., Peña-Méndez, E. M., Vaňhara, P., Hampl, A., & Havel, J. (2013). Artificial neural networks in medical diagnosis.

Anderson, C. W., Devulapalli, S. V., & Stolz, E. A. (1995). Determining mental state from EEG signals using parallel implementations of neural networks. *Scientific programming*, *4*(3), 171-183.

Ang, K. K., Chin, Z. Y., Wang, C., Guan, C., & Zhang, H. (2012). Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. *Frontiers in neuroscience*, *6*.

Antoniades, A., Spyrou, L., Took, C. C., & Sanei, S. (2016, September). Deep learning for epileptic intracranial EEG data. In *Machine Learning for Signal Processing (MLSP), 2016 IEEE 26th International Workshop on* (pp. 1-6). IEEE.

Artameeyanant, P., Sultornsanee, S., & Chamnongthai, K. (2017). Electroencephalography-based feature extraction using complex network for automated epileptic seizure detection. *Expert Systems*.

Atzori, M., Cognolato, M., & Müller, H. (2016). Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands. *Frontiers in neurorobotics*, *10*.

Bai, O., Lin, P., Vorbach, S., Li, J., Furlani, S., & Hallett, M. (2007). Exploration of computational methods for classification of movement intention during human voluntary movement from single trial EEG. *Clinical Neurophysiology*, *118*(12), 2637-2655.

Balakrishnan, D., & Puthusserypady, S. (2005, April). Multilayer perceptrons for the classification of brain computer interface data. In *Bioengineering Conference, 2005. Proceedings of the IEEE 31st Annual Northeast* (pp. 118-119). IEEE.

Baxt, W. G. (1990). Use of an artificial neural network for data analysis in clinical decision-making: the diagnosis of acute coronary occlusion. *Neural computation*, *2*(4), 480-489.

Bennett, K. P., & Campbell, C. (2000). Support vector machines: hype or hallelujah?. *Acm Sigkdd Explorations Newsletter*, *2*(2), 1-13.

Bevilacqua, V., Mastronardi, G., & Marinelli, M. (2006). A neural network approach to medical image segmentation and three-dimensional reconstruction. *Lecture notes in computer science*, *4113*, 22.

Bin, G., Gao, X., Yan, Z., Hong, B., & Gao, S. (2009). An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method. *Journal of neural engineering*, *6*(4), 046002.

Birbaumer, N. (2006). Breaking the silence: brain–computer interfaces (BCI) for communication and motor control. *Psychophysiology*, *43*(6), 517-532.

Birbaumer, N., Murguialday, A. R., Weber, C., & Montoya, P. (2009). Neurofeedback and brain–computer interface: clinical applications. *International review of neurobiology*, *86*, 107-117.

Blanco, S., Garcia, H., Quiroga, R. Q., Romanelli, L., & Rosso, O. A. (1995). Stationarity of the EEG series. *IEEE Engineering in medicine and biology Magazine*, *14*(4), 395-399.

Blankertz, B., Dornhege, G., Krauledat, M., Schröder, M., Williamson, J., Murray-Smith, R., & Müller, K. R. (2006). The Berlin Brain-Computer Interface presents the novel mental typewriter Hex-o-Spell.

Blankertz, B., Lemm, S., Treder, M., Haufe, S., & Müller, K. R. (2011). Single-trial analysis and classification of ERP components—a tutorial. *NeuroImage*, *56*(2), 814-825.

Boers, E. J., & Kuiper, H. (1992). Biological metaphors and the design of modular artificial neural networks.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992, July). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory* (pp. 144-152). ACM.

Bourlard, H., & Morgan, N. (1993). Continuous speech recognition by connectionist statistical methods. *IEEE Transactions on Neural Networks*, *4*(6), 893-909.

Breuel, T. M. (2015). The effects of hyperparameters on SGD training of neural networks. *arXiv preprint arXiv:1508.02788*.

Brunner, P., Joshi, S., Briskin, S., Wolpaw, J. R., Bischof, H., & Schalk, G. (2010). Does the 'P300'speller depend on eye gaze?. *Journal of neural engineering*, *7*(5), 056013.

Burde, W., & Blankertz, B. (2006). Is the locus of control of reinforcement a predictor of brain-computer interface performance?. In *Proceedings of the 3rd International Brain-Computer Interface Workshop and Training Course* (Vol. 2006, pp. 108-109).

Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, *2*(2), 121-167.

Burke, H. B., Goodman, P. H., Rosen, D. B., Henson, D. E., Weinstein, J. N., Harrell, F. E. & Bostwick, D. G. (1997). Artificial neural networks improve the accuracy of cancer survival prediction. *Cancer*, *79*(4), 857-862.

Casanova, E., Lazzari, R. E., Lotta, S., & Mazzucchi, A. (2003). Locked-in syndrome: improvement in the prognosis after an early intensive multidisciplinary rehabilitation. *Archives of physical medicine and rehabilitation*, *84*(6), 862-867.

Cecotti, H. (2011). Spelling with non-invasive Brain–Computer Interfaces–Current and future trends. *Journal of Physiology-Paris*, *105*(1), 106-114.

Cecere, C., Corrado, C., & Polikar, R. (2014, April). Diagnostic utility of EEG based biomarkers for Alzheimer's disease. In *Bioengineering Conference (NEBEC), 2014 40th Annual Northeast* (pp. 1-2). IEEE.

Chatterjee, R., & Bandyopadhyay, T. (2016, January). EEG based Motor Imagery Classification using SVM and MLP. In *Computational Intelligence and Networks (CINE), 2016 2nd International Conference on* (pp. 84-89). IEEE.

Chellapilla, K., Puri, S., & Simard, P. (2006, October). High performance convolutional neural networks for document processing. In *Tenth International Workshop on Frontiers in Handwriting Recognition*. Suvisoft.

Chen, X., Chen, Z., Gao, S., & Gao, X. (2014). A high-itr ssvep-based bci speller. *Brain-Computer Interfaces*, *1*(3-4), 181-191.

Chen, X., Wang, Y., Nakanishi, M., Jung, T. P., & Gao, X. (2014, August). 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* (pp. 3993-3996). IEEE.

Chen, X., Wang, Y., Nakanishi, M., Gao, X., Jung, T. P., & Gao, S. (2015). High-speed spelling with a noninvasive brain–computer interface. *Proceedings of the national academy of sciences*, *112*(44), E6058-E6067.

Chen, Y. J., Chen, S. C., & Wu, C. M. (2016, May). Using modular neural network to SSVEP-based BCI. In *Applied System Innovation (ICASI), 2016 International Conference on* (pp. 1-3). IEEE.

Cheng, M., Gao, X., Gao, S., & Xu, D. (2002). Design and implementation of a brain-computer interface with high transfer rates. *IEEE transactions on biomedical engineering*, *49*(10), 1181-1186.

Chi, Y. M., Wang, Y. T., Wang, Y., Maier, C., Jung, T. P., & Cauwenberghs, G. (2012). Dry and noncontact EEG sensors for mobile brain–computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *20*(2), 228-235.

Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A., & Oliva, A. (2016). Deep neural networks predict hierarchical spatio-temporal cortical dynamics of human visual object recognition. *arXiv preprint arXiv:1601.02970*.

Cılasun, M. H., & Yalçın, H. (2016, May). A deep learning approach to EEG based epilepsy seizure determination. In *Signal Processing and Communication Application Conference (SIU), 2016 24th* (pp. 1573-1576). IEEE.

Conson, M., Sacco, S., Sarà, M., Pistoia, F., Grossi, D., & Trojano, L. (2008). Selective motor imagery defect in patients with locked-in syndrome. *Neuropsychologia*, *46*(11), 2622-2628.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, *13*(1), 21-27.

Dähne, S., Bießmann, F., Samek, W., Haufe, S., Goltz, D., Gundlach, C. & Müller, K. R. (2015). Multivariate machine learning methods for fusing multimodal functional neuroimaging data. *Proceedings of the IEEE*, *103*(9), 1507-1530.

Diez, P. F., Mut, V. A., Perona, E. M. A., & Leber, E. L. (2011). Asynchronous BCI control using high-frequency SSVEP. *Journal of neuroengineering and rehabilitation*, *8*(1), 39.

Di Russo, F., & Spinelli, D. (1999). Electrophysiological evidence for an early attentional mechanism in visual processing in humans. *Vision research*, *39*(18), 2975-2985.

Dou, Q., Chen, H., Yu, L., Zhao, L., Qin, J., Wang, D. & Heng, P. A. (2016). Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. *IEEE transactions on medical imaging*, *35*(5), 1182-1195.

Doud, A. J., Lucas, J. P., Pisansky, M. T., & He, B. (2011). Continuous three-dimensional control of a virtual helicopter using a motor imagery based brain-computer interface. *PloS one*, *6*(10), e26322.

Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics*, *35*(5), 352-359.

Duda, R. 0., Hart, PE, and Stork, DG 2001. Pattern Classification.

Duch, W., & Jankowski, N. (1999). Survey of neural transfer functions. *Neural Computing Surveys*, *2*(1), 163-212.

Er, F., Iscen, P., Sahin, S., Çinar, N., Karsidag, S., & Goularas, D. (2017). Distinguishing age-related cognitive decline from dementias: A study based on machine learning algorithms. *Journal of Clinical Neuroscience*.

Erhan, D., Bengio, Y., Courville, A., Manzagol, P. A., Vincent, P., & Bengio, S. (2010). Why does unsupervised pre-training help deep learning?. *Journal of Machine Learning Research*, *11*(Feb), 625-660.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115-118.

Farfade, S. S., Saberian, M. J., & Li, L. J. (2015, June). Multi-view face detection using deep convolutional neural networks. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval* (pp. 643-650). ACM.

Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, *70*(6), 510-523.

Fisher, R. S., Harding, G., Erba, G., Barkley, G. L., & Wilkins, A. (2005). Photic-and pattern-induced seizures: a review for the Epilepsy Foundation of America Working Group. *Epilepsia*, *46*(9), 1426-1441.

Friman, O., Volosyak, I., & Graser, A. (2007). Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces. *IEEE transactions on biomedical engineering*, *54*(4), 742-750.

Fukushima, K. (1979). Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position- Neocognitron. *ELECTRON. & COMMUN. JAPAN*, *62*(10), 11-18.

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, *36*(4), 193-202.

Gao, X., Xu, D., Cheng, M., & Gao, S. (2003). A BCI-based environmental controller for the motion-disabled. *IEEE Transactions on neural systems and rehabilitation engineering*, *11*(2), 137-140.

Geronimo, A., Simmons, Z., & Schiff, S. J. (2016). Performance predictors of brain–computer interfaces in patients with amyotrophic lateral sclerosis. *Journal of neural engineering*, *13*(2), 026002.

Glorot, X., Bordes, A., & Bengio, Y. (2011, April). Deep Sparse Rectifier Neural Networks. In *Aistats* (Vol. 15, No. 106, p. 275).

Gonsalvez, C. J., & Polich, J. (2002). P300 amplitude is determined by target-to-target interval. *Psychophysiology*, *39*(3), 388-396.

Guger, C., Edlinger, G., Harkam, W., Niedermayer, I., & Pfurtscheller, G. (2003). How many people are able to operate an EEG-based brain-computer interface (BCI)?. *IEEE transactions on neural systems and rehabilitation engineering*, *11*(2), 145-147.

Guger, C., Daban, S., Sellers, E., Holzner, C., Krausz, G., Carabalona, R., ... & Edlinger, G. (2009). How many people are able to control a P300-based brain–computer interface (BCI)?. *Neuroscience letters*, *462*(1), 94-98.

Guger, C., Allison, B. Z., Großwindhager, B., Prückl, R., Hintermüller, C., Kapeller, C. & Edlinger, G. (2012). How many people could use an SSVEP BCI?. *Frontiers in neuroscience*, *6*.

Halder, S., Varkuti, B., Bogdan, M., Kübler, A., Rosenstiel, W., Sitaram, R., & Birbaumer, N. (2013). Prediction of brain-computer interface aptitude from individual brain structure. *Frontiers in human neuroscience*, *7*.

Hammer, E. M., Halder, S., Blankertz, B., Sannelli, C., Dickhaus, T., Kleih, S. & Kübler, A. (2012). Psychological predictors of SMR-BCI performance. *Biological psychology*, *89*(1), 80-86.

Hahnloser, R. H., Sarpeshkar, R., Mahowald, M. A., Douglas, R. J., & Seung, H. S. (2000). Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature*, *405*(6789), 947-951.

Hartmann, M., & Kluge, T. (2007, May). Phase coherent detection of steady-state evoked potentials: theory and performance analysis. In *Neural Engineering, 2007. CNE'07. 3rd International IEEE/EMBS Conference on*(pp. 179-183). IEEE.

Herrmann, C. S. (2001). 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.

Hill, N. J., Lal, T. N., Schroder, M., Hinterberger, T., Wilhelm, B., Nijboer, F. & Kubler, A. (2006). Classifying EEG and ECoG signals without subject training for fast BCI implementation: comparison of nonparalyzed and completely paralyzed subjects. *IEEE transactions on neural systems and rehabilitation engineering*, *14*(2), 183-186.

Holz, E. M., Botrel, L., Kaufmann, T., & Kübler, A. (2015). Long-term independent brain-computer interface home use improves quality of life of a patient in the locked-in state: a case study. *Archives of physical medicine and rehabilitation*, *96*(3), S16-S26.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of physiology, 160(1), 106-154.

Hwang, H. J., Lim, J. H., Jung, Y. J., Choi, H., Lee, S. W., & Im, C. H. (2012). Development of an SSVEP-based BCI spelling system adopting a QWERTY-style LED keyboard. *Journal of neuroscience methods*, *208*(1), 59-65.

Ilyas, M. Z., Saad, P., Ahmad, M. I., & Ghani, A. R. I. (2016, November). Classification of EEG signals for brain-computer interface applications: Performance comparison. In *Robotics, Automation and Sciences (ICORAS), International Conference on* (pp. 1-4). IEEE.

Işik, H., & Sezer, E. (2012). Diagnosis of epilepsy from electroencephalography signals using multilayer perceptron and Elman artificial neural networks and wavelet transform. *Journal of medical systems*, *36*(1), 1-13.

Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(1), 4-37.

Jaiswal, A. K., & Banka, H. (2017). Local pattern transformation based feature extraction techniques for classification of epileptic EEG signals. *Biomedical Signal Processing and Control*, *34*, 81-92.

James, C. D., Aimone, J. B., Miner, N. E., Vineyard, C. M., Rothganger, F. H., Carlson, K. D. & Naegle, J. H. (2017). A historical survey of algorithms and hardware architectures for neural-inspired and neuromorphic computing applications. *Biologically Inspired Cognitive Architectures*.

Jarrett, K., Kavukcuoglu, K., & LeCun, Y. (2009, September). What is the best multi-stage architecture for object recognition?. In *Computer Vision, 2009 IEEE 12th International Conference on*(pp. 2146-2153). IEEE.

Jin, C., & Wang, L. (2012). Dimensionality dependent PAC-Bayes margin bound. In *Advances in neural information processing systems* (pp. 1034-1042).

Jingwei, L., Yin, C., & Weidong, Z. (2015, July). Deep learning EEG response representation for brain computer interface. In *Control Conference (CCC), 2015 34th Chinese* (pp. 3518-3523). IEEE.

Johansen, A. R., Jin, J., Maszczyk, T., Dauwels, J., Cash, S. S., & Westover, M. B. (2016, March). Epileptiform spike detection via convolutional neural networks. In *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on* (pp. 754-758). IEEE.

Johansson, B., & Jakobsson, P. (2000). Fourier analysis of steady-state visual evoked potentials in subjects with normal and defective stereo vision. *Documenta Ophthalmologica*, *101*(3), 233-246.

Kappenman, E. S., & Luck, S. J. (2010). The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology*, *47*(5), 888-904.

Kaper, M., Meinicke, P., Grossekathoefer, U., Lingner, T., & Ritter, H. (2004). BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm. *IEEE Transactions on Biomedical Engineering*, *51*(6), 1073-1076.

Kataria, A., & Singh, M. D. (2013). A review of data classification using k-nearest neighbour algorithm. *International Journal of Emerging Technology and Advanced Engineering*, *3*(6), 354-360

Kelly, S. P., Lalor, E. C., Reilly, R. B., & Foxe, J. J. (2005). Visual spatial attention tracking using high-density SSVEP data for independent brain-computer communication. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *13*(2), 172-178.

Kha, V. A., Nguyen, D. N., Kha, H. H., & Dutkiewicz, E. (2017, February). Dynamic stopping using eSVM scores analysis for event-related potential brain-computer interfaces. In *Medical Information and Communication Technology (ISMICT), 2017 11th International Symposium on* (pp. 82-85). IEEE.

Khanna, K., Verma, A., & Richard, B. (2011). "The locked-in syndrome": Can it be unlocked?. *Journal of Clinical Gerontology and Geriatrics*, *2*(4), 96-99.

Khasnobish, A., Bhattacharyya, S., Konar, A., & Tibarewala, D. N. (2010). K-Nearest neighbor classification of left-right limb movement using EEG data. In *International conference on Biomedical Engineering and assistive technologies, NIT Jalandhar*.

Kim, E. J., & Brunner, R. J. (2016). Star-galaxy classification using deep convolutional neural networks. *Monthly Notices of the Royal Astronomical Society*, stw2672.

Kimura, A., Takahashi, I., Tanaka, M., Yasuda, N., Ueda, N., & Yoshida, N. (2017, June). Single-epoch supernova classification with deep convolutional neural networks. In *Distributed Computing Systems Workshops (ICDCSW), 2017 IEEE 37th International Conference on* (pp. 354-359). IEEE.

Klonowski, W. (2009). Everything you wanted to ask about EEG but were afraid to get the right answer. *Nonlinear Biomedical Physics*, *3*(1), 2.

Kluge, T., & Hartmann, M. (2007, May). Phase coherent detection of steady-state evoked potentials: experimental results and application to brain-computer interfaces. In *Neural Engineering, 2007. CNE'07. 3rd International IEEE/EMBS Conference on* (pp. 425-429). IEEE.

Knowlton, R. C. (2008). Can magnetoencephalography aid epilepsy surgery?. *Epilepsy currents*, *8*(1), 1-5.

Knowlton, R. C., Elgavish, R. A., Bartolucci, A., Ojha, B., Limdi, N., Blount, J. & Kankirawatana, P. (2008). Functional imaging: II. Prediction of epilepsy surgery outcome. *Annals of neurology*, *64*(1), 35-41.

Krähenbühl, P., Doersch, C., Donahue, J., & Darrell, T. (2015). Data-dependent initializations of convolutional neural networks. *arXiv preprint arXiv:1511.06856*.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

Krumpe, T., Walter, C., Rosenstiel, W., & Spüler, M. (2016). Asynchronous P300 classification in a reactive brain-computer interface during an outlier detection task. *Journal of neural engineering*, *13*(4), 46015-46024.

Kübler, A., Kotchoubey, B., Kaiser, J., Wolpaw, J. R., & Birbaumer, N. (2001). Brain–computer communication: Unlocking the locked in. *Psychological bulletin*, *127*(3), 358.

Kübler, A., & Müller, K. R. (2007). An introduction to brain-computer interfacing. *Toward brain-computer interfacing*, 1-25.

Kübler, A., & Birbaumer, N. (2008). Brain–computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients?. *Clinical neurophysiology*, *119*(11), 2658-2666.

Kwak, N. S., Müller, K. R., & Lee, S. W. (2015). A lower limb exoskeleton control system based on steady state visual evoked potentials. *Journal of neural engineering*, *12*(5), 056009.

Kwak, N. S., Müller, K. R., & Lee, S. W. (2017). A convolutional neural network for steady state visual evoked potential classification under ambulatory environment. *PloS one*, *12*(2), e0172578.

Lacher, R. C., Coats, P. K., Sharma, S. C., & Fant, L. F. (1995). A neural network for classifying the financial health of a firm. *European Journal of Operational Research*, *85*(1), 53-65.

LaFleur, K., Cassady, K., Doud, A., Shades, K., Rogin, E., & He, B. (2013). Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface. *Journal of neural engineering*, *10*(4), 046003.

Lal, T. N., Schroder, M., Hinterberger, T., Weston, J., Bogdan, M., Birbaumer, N., & Scholkopf, B. (2004). Support vector channel selection in BCI. *IEEE transactions on biomedical engineering*, *51*(6), 1003-1010.

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., & Lance, B. J. (2016). EEGNet: A Compact Convolutional Network for EEG-based Brain-Computer Interfaces. *arXiv preprint arXiv:1611.08024*.

Layne, S. P., Mayer-Kress, G., & Holzfuss, J. (1986). Problems associated with dimensional analysis of electroencephalogram data. In *Dimensions and entropies in chaotic systems* (pp. 246-256). Springer, Berlin, Heidelberg.

Lee, P. L., Sie, J. J., Liu, Y. J., Wu, C. H., Lee, M. H., Shu, C. H. & Shyu, K. K. (2010). An SSVEP-actuated brain computer interface using phase-tagged flickering sequences: a cursor system. *Annals of biomedical engineering*, *38*(7), 2383-2397.

Lee, H., Grosse, R., Ranganath, R., & Ng, A. Y. (2011). Unsupervised learning of hierarchical representations with convolutional deep belief networks. *Communications of the ACM*, *54*(10), 95-103.

LeCun, Y., Haffner, P., Bottou, L., & Bengio, Y. (1999). Object recognition with gradient-based learning. *Shape, contour and grouping in computer vision*, 823-823.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444.

Leuthardt, E. C., Schalk, G., Wolpaw, J. R., Ojemann, J. G., & Moran, D. W. (2004). A brain–computer interface using electrocorticographic signals in humans. *Journal of neural engineering*, *1*(2), 63.

Li, Y., Guan, C., Li, H., & Chin, Z. (2008). A self-training semi-supervised SVM algorithm and its application in an EEG-based brain computer interface speller system. *Pattern Recognition Letters*, *29*(9), 1285-1294.

Li, H., Lin, Z., Shen, X., Brandt, J., & Hua, G. (2015). A convolutional neural network cascade for face detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 5325-5334).

Lin, Z., Zhang, C., Wu, W., & Gao, X. (2007). Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. *IEEE Transactions on Biomedical Engineering*, *54*(6), 1172-1176.

Lippmann, R. P. (1989). Review of neural networks for speech recognition. *Neural computation*, *1*(1), 1-38.

Loffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning* (pp. 448-456).

Lopes, J. B. (2001, May). Designing user interfaces for severely handicapped persons. In *Proceedings of the 2001 EC/NSF workshop on Universal accessibility of ubiquitous computing: providing for the elderly* (pp. 100-106). ACM.

Lopez-Gordo, M. A., Sanchez-Morillo, D., & Valle, F. P. (2014). Dry EEG electrodes. *Sensors*, *14*(7), 12847-12870.

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain–computer interfaces. *Journal of neural engineering*, *4*(2), R1.

Luo, Z., Liu, H., & Wu, X. (2005, July). Artificial neural network computation on graphic process unit. In *Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on* (Vol. 1, pp. 622-626). IEEE.

Mayer-Kress, Gottfried, & Layne, S. P. (1987). Dimensionality of the human electroencephalogram. *Annals of the New York Academy of Sciences*, *504*(1), 62-87.

Manyakov, N. V., Chumerin, N., Combaz, A., Robben, A., van Vliet, M., & Van Hulle, M. M. (2011, September). Decoding Phase-Based Information from Steady-State Visual Evoked Potentials with Use of Complex-Valued Neural Network. In *IDEAL* (pp. 135-143).

Manyakov, N. V., Chumerin, N., & Van Hulle, M. M. (2012). Multichannel decoding for phase-coded SSVEP brain–computer interface. *International journal of neural systems*, *22*(05), 1250022.

Manyakov, N. V., Chumerin, N., Robben, A., Combaz, A., van Vliet, M., & Van Hulle, M. M. (2013). Sampled sinusoidal stimulation profile and multichannel fuzzy logic classification for monitor-based phase-coded SSVEP brain–computer interfacing. *Journal of neural engineering*, *10*(3), 036011.

Mathewson, K. E., Harrison, T. J., & Kizuk, S. A. (2017). High and dry? Comparing active dry EEG electrodes to active and passive wet electrodes. *Psychophysiology*, *54*(1), 74-82.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and emotion*, *23*(2), 209-237.

Middendorf, M., McMillan, G., Calhoun, G., & Jones, K. S. (2000). Brain-computer interfaces based on the steady-state visual-evoked response. *IEEE transactions on rehabilitation engineering*, *8*(2), 211-214.

Mountcastle, V. B. (1957). Modality and Topographic Properties of Single Neurones of Cats Somatosensory Cortex. *Journal of neurophysiology*, *20*(4), 408-434.

Muller, M. M., Malinowski, P., Gruber, T., & Hillyard, S. A. (2003). Sustained division of the attentional spotlight. *Nature*, *424*(6946), 309.

Muller, K. R., & Blankertz, B. (2006). Toward noninvasive brain-computer interfaces. *IEEE Signal Processing Magazine*, *23*(5), 128-126.

Müller-Gerking, J., Pfurtscheller, G., & Flyvbjerg, H. (1999). Designing optimal spatial filters for single-trial EEG classification in a movement task. *Clinical neurophysiology*, *110*(5), 787-798.

Müller-Putz, G. R., Scherer, R., Brauneis, C., & Pfurtscheller, G. (2005). Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components. *Journal of neural engineering*, *2*(4), 123.

Nasehi, S., & Pourghassem, H. (2011, May). A novel effective feature selection algorithm based on S-PCA and wavelet transform features in EEG signal classification. In *Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on* (pp. 114-117). IEEE.

Nasehi, S., & Pourghassem, H. (2013, April). Mental task classification based on HMM and BPNN. In *Communication Systems and Network Technologies (CSNT), 2013 International Conference on* (pp. 210-214). IEEE.

Naz, S., & Bawane, N. G. (2016). Recent Trends in BCI based Speller System: A Survey Report. *International Journal of Engineering Science*, *8596*.

Ng B, Logothetis N, Kayser C. EEG phase patterns reflect the selectivity of neural firing. Cerebral Cortex. 2013;23(2):389–398. pmid:22345353

Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors*, *12*(2), 1211-1279.

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: a review. *Journal of vision*, *15*(6), 4-4.

Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA.

Nurse, E. S., Karoly, P. J., Grayden, D. B., & Freestone, D. R. (2015). A generalizable brain-computer interface (bci) using machine learning for feature discovery. *PloS one*, *10*(6), e0131328.

Obermaier, B., Neuper, C., Guger, C., & Pfurtscheller, G. (2001). Information transfer rate in a five-classes brain-computer interface. *IEEE Transactions on neural systems and rehabilitation engineering*, *9*(3), 283-288.

Oh, K. S., & Jung, K. (2004). GPU implementation of neural networks. *Pattern Recognition*, *37*(6), 1311-1314.

Olejniczak, P. (2006). Neurophysiologic basis of EEG. *Journal of clinical neurophysiology*, *23*(3), 186-189.

Pan, J., Gao, X., Duan, F., Yan, Z., & Gao, S. (2011). Enhancing the classification accuracy of steady-state visual evoked potential-based brain–computer interfaces using phase constrained canonical correlation analysis. *Journal of neural engineering*, *8*(3), 036027.

Panov, F., Levin, E., de Hemptinne, C., Swann, N. C., Qasim, S., Miocinovic, S. & Starr, P. A. (2017). Intraoperative electrocorticography for physiological research in movement disorders: principles and experience in 200 cases. *Journal of neurosurgery*, *126*(1), 122-131.

Patterson, J. R., & Grabois, M. (1986). Locked-in syndrome: a review of 139 cases. *Stroke*, *17*(4), 758-764.

Pfurtscheller, G., & Aranibar, A. (1979). Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement. *Electroencephalography and clinical neurophysiology*, *46*(2), 138-146.

Pfurtscheller, G., & Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology*, *110*(11), 1842-1857.

Pfurtscheller, G., Graimann, B., Huggins, J. E., Levine, S. P., & Schuh, L. A. (2003). Spatiotemporal patterns of beta desynchronization and gamma synchronization in corticographic data during self-paced movement. *Clinical neurophysiology*, *114*(7), 1226-1236.

Piotrowski, A. P., & Napiorkowski, J. J. (2013). A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling. *Journal of Hydrology*, *476*, 97-111.

Prieto, A., Prieto, B., Ortigosa, E. M., Ros, E., Pelayo, F., Ortega, J., & Rojas, I. (2016). Neural networks: An overview of early research, current frameworks and new challenges. *Neurocomputing*, *214*, 242-268.

Raina, R., Madhavan, A., & Ng, A. Y. (2009, June). Large-scale deep unsupervised learning using graphics processors. In *Proceedings of the 26th annual international conference on machine learning* (pp. 873-880). ACM.

Randolph, A. B., Jackson, M. M., & Karmakar, S. (2010). Individual characteristics and their effect on predicting Mu rhythm modulation. *Intl. Journal of Human–Computer Interaction*, *27*(1), 24-37.

Reigada, C., Mendes, O., Paiva, C., Tavares, M., & Gonçalves, E. (2014). ALS patients in locked-in syndrome: a systematic review of the literature. *Palliative Medicine*, *28*(6), 688-689.

Riccio, A., Mattia, D., Simione, L., Olivetti, M., & Cincotti, F. (2012). Eye-gaze independent EEG-based brain–computer interfaces for communication. *Journal of neural engineering*, *9*(4), 045001.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature neuroscience*, *2*(11), 1019-1025.

Robert, C., Gaudy, J. F., & Limoge, A. (2002). Electroencephalogram processing using neural networks. *Clinical Neurophysiology*, *113*(5), 694-701.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). *Learning internal representations by error propagation* (No. ICS-8506).

Samiee, K., Kovacs, P., & Gabbouj, M. (2015). Epileptic seizure classification of EEG time-series using rational discrete short-time Fourier transform. *IEEE transactions on Biomedical Engineering*, *62*(2), 541-552.

Sarle, W. S. (1994). Neural networks and statistical models.

Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M. & Ball, T. (2017). Deep learning with convolutional neural networks for EEG decoding and visualization. *Human Brain Mapping*.

Schlögl, A., Lee, F., Bischof, H., & Pfurtscheller, G. (2005). Characterization of four-class motor imagery EEG data for the BCI-competition 2005. *Journal of neural engineering*, *2*(4), L14.

Schröder, M., Lal, T. N., Hinterberger, T., Bogdan, M., Hill, N. J., Birbaumer, N. & Schölkopf, B. (2005). Robust EEG channel selection across subjects for brain-computer interfaces. *EURASIP Journal on Applied Signal Processing*, *2005*, 3103-3112.

Schumacher, M., Roßner, R., & Vach, W. (1996). Neural networks and logistic regression: Part I. *Computational Statistics & Data Analysis*, *21*(6), 661-682.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85-117.

Shedeed, H. A., Issa, M. F., & El-Sayed, S. M. (2013, November). Brain EEG signal processing for controlling a robotic arm. In *Computer Engineering & Systems (ICCES), 2013 8th International Conference on* (pp. 152-157). IEEE.

Shiraishi, H., Fujiwara, T., Inoue, Y., & Yagi, K. (2001). Photosensitivity in relation to epileptic syndromes: a survey from an epilepsy center in Japan. *Epilepsia*, *42*(3), 393-397.

Silva, V. F., Barbosa, R. M., Vieira, P. M., & Lima, C. S. (2017, February). Ensemble learning based classification for BCI applications. In *Bioengineering (ENBENG), 2017 IEEE 5th Portuguese Meeting on* (pp. 1-4). IEEE.

Siuly, S., & Li, Y. (2012). Improving the separability of motor imagery EEG signals using a cross correlation-based least square support vector machine for brain–computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *20*(4), 526-538.

Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Delyon, B., Glorennec, P. Y. & Juditsky, A. (1995). Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, *31*(12), 1691-1724.

Smith, E., & Delargy, M. (2005). Locked-in syndrome. *Bmj*, *330*(7488), 406-409.

Solomon, J., Boe, S., & Bardouille, T. (2015). Reliability for non-invasive somatosensory cortex localization: Implications for pre-surgical mapping. *Clinical neurology and neurosurgery*, *139*, 224-229.

Speier, W., Arnold, C., & Pouratian, N. (2016). Integrating language models into classifiers for BCI communication: a review. *Journal of neural engineering*, *13*(3), 031002.

Spencer, K. M., Niznikiewicz, M. A., Shenton, M. E., & McCarley, R. W. (2008). Sensory-evoked gamma oscillations in chronic schizophrenia. *Biological psychiatry*, *63*(8), 744-747.

Sprague, S. A., McBee, M. T., & Sellers, E. W. (2016). The effects of working memory on brain–computer interface performance. *Clinical Neurophysiology*, *127*(2), 1331-1341.

Srinivasan, V., Eswaran, C., & Sriraam, N. (2007). Approximate entropy-based epileptic EEG detection using artificial neural networks. *IEEE Transactions on information Technology in Biomedicine*, *11*(3), 288-295.

Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *Journal of machine learning research*, *15*(1), 1929-1958.

Strigl, D., Kofler, K., & Podlipnig, S. (2010, February). Performance and scalability of GPU-based convolutional neural networks. In *Parallel, Distributed and Network-Based Processing (PDP), 2010 18th Euromicro International Conference on* (pp. 317-324). IEEE.

Stufflebeam, S. M., Tanaka, N., & Ahlfors, S. P. (2009). Clinical applications of magnetoencephalography. *Human brain mapping*, *30*(6), 1813-1823.

Subasi, A., & Ercelebi, E. (2005). Classification of EEG signals using neural network and logistic regression. *Computer methods and programs in biomedicine*, *78*(2), 87-99.

Sutter, E. E. (1992). The brain response interface: communication through visually-induced electrical brain responses. *Journal of Microcomputer Applications*, *15*(1), 31-45.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187-1188.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D. & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-9).

Tautan, A. M., Mihajlovic, V., Chen, Y. H., Grundlehner, B., Penders, J., & Serdijn, W. A. (2014, March). Signal Quality in Dry Electrode EEG and the Relation to Skin-electrode Contact Impedance Magnitude. In *BIODEVICES*(pp. 12-22).

Tabar, Y. R., & Halici, U. (2016). A novel deep learning approach for classification of EEG motor imagery signals. *Journal of neural engineering*, *14*(1), 016003.

Townsend, G., & Platsko, V. (2016). Pushing the P300-based brain–computer interface beyond 100 bpm: extending performance guided constraints into the temporal domain. *Journal of neural engineering*, *13*(2), 026024.

Treder, M. S., & Blankertz, B. (2010). (C) overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and brain functions*, *6*(1), 28.

Vaid, S., Singh, P., & Kaur, C. (2015, February). EEG signal analysis for BCI interface: A review. In *Advanced Computing & Communication Technologies (ACCT), 2015 Fifth International Conference on* (pp. 143-147). IEEE.

Vedaldi, A., & Lenc, K. (2015, October). Matconvnet: Convolutional neural networks for matlab. In *Proceedings of the 23rd ACM international conference on Multimedia* (pp. 689-692). ACM.

Vialatte, F. B., Maurice, M., Dauwels, J., & Cichocki, A. (2010). Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Progress in neurobiology*, *90*(4), 418-438.

Vidal, J. J. (1973). Toward direct brain-computer communication. *Annual review of Biophysics and Bioengineering*, *2*(1), 157-180.

von Storch, H., & Zwiers, F. W. (2002). Statistical analysis in climate research.

Vuckovic, A., Radivojevic, V., Chen, A. C., & Popovic, D. (2002). Automatic recognition of alertness and drowsiness from EEG by an artificial neural network. *Medical engineering & physics*, *24*(5), 349-360.

Walker, I., Deisenroth, M., & Faisal, A. (2015). Deep convolutional neural networks for brain computer interface using motor imagery. *IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE DEPARTMENT OF COMPUTING*.

Wang, Y., Wang, R., Gao, X., Hong, B., & Gao, S. (2006). A practical VEP-based brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *14*(2), 234-240.

Wang, Y., Gao, X., Hong, B., Jia, C., & Gao, S. (2008). Brain-computer interfaces based on visual evoked potentials. *IEEE Engineering in medicine and biology magazine*, *27*(5).

Wang, Y., Chen, X., Gao, X., & Gao, S. (2016). A Benchmark Dataset for SSVEP-Based Brain-Computer Interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*.

Weng, J., Ahuja, N., & Huang, T. S. (1992, June). Cresceptron: a self-organizing neural network which grows adaptively. In *Neural Networks, 1992. IJCNN., International Joint Conference on* (Vol. 1, pp. 576-581). IEEE.

Williams, J. M. (2017). Deep Learning and Transfer Learning in the Classification of EEG Signals.

Wilson, R. L., & Sharda, R. (1994). Bankruptcy prediction using neural networks. *Decision support systems*, *11*(5), 545-557.

Wilson, J. J., & Palaniappan, R. (2009, April). Augmenting a SSVEP BCI through single cycle analysis and phase weighting. In *Neural Engineering, 2009. NER'09. 4th International IEEE/EMBS Conference on* (pp. 371-374). IEEE.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical neurophysiology*, *113*(6), 767-791.

Wolpaw, J. R., Ramoser, H., McFarland, D. J., & Pfurtscheller, G. (1998). EEG-based communication: improved accuracy by response verification. *IEEE transactions on Rehabilitation Engineering*, *6*(3), 326-333.

Won, D. O., Hwang, H. J., Dähne, S., Müller, K. R., & Lee, S. W. (2015). Effect of higher frequency on the classification of steady-state visual evoked potentials. *Journal of neural engineering*, *13*(1), 016014.

Wulsin, D., Blanco, J., Mani, R., & Litt, B. (2010, December). Semi-supervised anomaly detection for EEG waveforms using deep belief nets. In *Machine Learning and Applications (ICMLA), 2010 Ninth International Conference on* (pp. 436-441). IEEE.

Xu, W., Guan, C., Siong, C. E., Ranganatha, S., Thulasidas, M., & Wu, J. (2004, August). High accuracy classification of EEG signal. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on* (Vol. 2, pp. 391-394). IEEE.

Yang, J., Singh, H., Hines, E. L., Schlaghecken, F., Iliescu, D. D., Leeson, M. S., & Stocks, N. G. (2012). Channel selection and classification of electroencephalogram signals: an artificial neural network and genetic algorithm-based approach. *Artificial intelligence in medicine*, *55*(2), 117-126.

Yeom, S. K., Fazli, S., Müller, K. R., & Lee, S. W. (2014). An efficient ERP-based brain-computer interface using random set presentation and face familiarity. *PloS one*, *9*(11), e111157.

Yin, E., Zhou, Z., Jiang, J., Yu, Y., & Hu, D. (2015). A dynamically optimized SSVEP brain–computer interface (BCI) speller. *IEEE Transactions on Biomedical Engineering*, *62*(6), 1447-1456.

Yuan, H., Liu, T., Szarkowski, R., Rios, C., Ashe, J., & He, B. (2010). Negative covariation between task-related responses in alpha/beta-band activity and BOLD in human sensorimotor cortex: an EEG and fMRI study of motor imagery and movements. *Neuroimage*, *49*(3), 2596-2606.

Yuan, P., Gao, X., Allison, B., Wang, Y., Bin, G., & Gao, S. (2013). A study of the existing problems of estimating the information transfer rate in online brain–computer interfaces. *Journal of neural engineering*, *10*(2), 026014.

Zander, T. O., Kothe, C., Jatzev, S., & Gaertner, M. (2010). Enhancing human-computer interaction with input from active and passive brain-computer interfaces. In *Brain-computer interfaces* (pp. 181-199). Springer London.

Zhang, G. P. (2000). Neural networks for classification: a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *30*(4), 451-462.

Zhang, Y., Zhou, G., Jin, J., Wang, M., Wang, X., & Cichocki, A. (2013). L1-regularized multiway canonical correlation analysis for SSVEP-based BCI. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *21*(6), 887-896.

Zhang, C., & Zhang, Z. (2014, March). Improving multiview face detection with multi-task deep convolutional neural networks. In *Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on* (pp. 1036-1041). IEEE.

Zhang, Y., Zhou, G., Jin, J., Wang, X., & Cichocki, A. (2015). SSVEP recognition using common feature analysis in brain–computer interface. *Journal of neuroscience methods*, *244*, 8-15.