UNIVERSIDAD DEL NORTE

BCI applications based on artificial intelligence oriented to deep learning techniques.

by

Sandra Liliana Cancino Suárez

Thesis Advisor: Norelli Schettini, Ph.D.

Co-Advisor: Jaime Delgado Saa, Ph.D.

Submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Electronics Engineering

Department of Electrical and Electronics Engineering

Agosto 2023

Technical Sheet	
PhD Thesis Title	BCI applications based on artificial intelligence oriented to deep learning techniques
Author	Sandra Liliana Cancino Suárez, MSc.
Thesis Advisor	Norelli Schettini, Ph.D. Associate professor Universidad Del Norte Barranquilla - Colombia
Advisor Approval Signature	
Co-Advisor	Jaime Delgado Saa, Ph.D. Founder SciFork Geneva - Switzerland
Co-Advisor Approval Signature	
Research Group	Biomedical Signal Processing and Artificial Intelligence Laboratory
Research Line	Brain Computer Interfaces BCI
Research Subline	Neuroscience

Contents

List of Figures

List of Tables

Acknowledgements

To my daughter Laura, my motivation to continue when I felt everything was lost. To my lovely parents, my brothers, and my sister-in-law for believing that I would be able to accomplish this huge academic step. To my first advisor Jaime Delgado, an excellent researcher and professor, who show me the world of the brain and BCIs. To my second advisor Norelli Schettini for her patience and wise recommendations for finishing this work. To Juan Manuel López, the best work partner one can have, an excellent professor, and my dear friend who help me to find new ideas for this research. To Marcela Rodríguez, my dear friend, for her words and advice whenever I needed it. To many others in my life. I am so lucky to have you... Thank you so much to all of you!

Abstract

A Brain-Computer Interface, BCI, can decode the brain signals corresponding to the intentions of individuals who have lost neuromuscular connection, to reestablish communication to control external devices. To this aim, BCI acquires brain signals as Electroencephalography (EEG) or Electrocorticography (ECoG), uses signal processing techniques and extracts features to train classifiers for providing proper control instructions. BCI development has increased in the last decades, improving its performance through the use of different signal processing techniques for feature extraction and artificial intelligence approaches for classification, such as deep learning-oriented classifiers. All of these can assure more accurate assistive systems but also can enable an analysis of the learning process of signal characteristics for the classification task. Initially, this work proposes the use of a priori knowledge and a correlation measure to select the most discriminative ECoG signal electrodes. Then, signals are processed using spatial filtering and three different types of temporal filtering, followed by a classifier made of stacked-autoencoders and a softmax layer to discriminate between ECoG signals from two types of visual stimuli. Results show that the average accuracy obtained is 97% (+/- 0.02%), which is similar to state-of-the-art techniques, nevertheless, this method uses minimal prior physiological and an automated statistical technique to select some electrodes to train the classifier. Also, this work presents classifiers analysis, figuring out which are the most relevant signal features useful for visual stimuli classification. The features and physiological information such as the brain areas involved are compared. Finally, this research uses Convolutional Neural Networks (CNN) or Convnets to classify 5 categories of motor tasks EEG signals. Movement-related cortical potentials (MRCPs) are used as a priori information to improve the processing of time-frequency representation of EEG signals. Results show an increase of more than 25% in average accuracy compared to a state-of-the-art method that uses the same database. In addition, an analysis of CNN or ConvNets filters and feature maps is done to find the most relevant signal characteristics that can help classify the five types of motor tasks.

Nomenclature

ANN Artificial Neural Networks

BCI Brain-Computer Interface

CAR Common average reference

CCA Canonical correlation analysis

CNN/ConvNets Convolutional Neural Networks

DBN Deep Belief Networks

DBNs Deep Belief Nets

DNN Deep Neural Networks

ECoG Electrocorticography

EEG Electroencephalography

EMG Electromyography

EOG Electrooculography

ERS Event-related synchronization

fMRI Functional magnetic resonance imaging

Grad-CAM Gradient-weighted Class Activation Mapping

HFC High-Frequency Components

ICA Independent Component Analysis

KNNs K-Nearest Neighbors

LDA Linear Discriminant Analysis

LFC Low-Frequency Components

LRP Layer-wise relevance propagation

LSTM Long-short term memory

MEG Magnetoencephalography

MLP Multi Layer Perceptron

MRCPs Movement Related Cortical Potentials

mVEPs Motion-onset visual evoked potentials

NIRS Near-infrared spectroscopy

PCSA Prior Supervised Convolutional Stacked Auto-encoders

PET Positron emission tomography

PSD Power spectral density

RF Random Forest

RNN Recurrent Neural Networks

SAE Stacked-autoencoders

SCI Spinal cord injury

sLDA Shrinkage Linear Discriminant Analysis

SSVEPs Steady-state visual evoked potentials

STFT Short Time Fourier Transform

SVM Support Vector Machine

Chapter 1

Introduction

1.1 Motivation

The human brain is known as the master controller of our body functions. Through the nervous system, it can send information to perform specific tasks. So when communications are interrupted and are not that easy to restore, people are limited in their daily activities and it can deteriorate their quality of life. A solution may be to acquire brain information and translate it, to control external devices, computers, and robotic prostheses, among others. To achieve this goal, brain information can be obtained as electrical activity in the form of neural signals recorded by electrodes. Signal processing techniques and classifiers altogether may extract relevant signal features to decode brain messages and provide control instructions. Everything described above is performed by a Brain-Computer Interface or BCI. BCI development has increased in the last decades, improving its performance through the use of artificial intelligence approaches, such as deep learning-oriented classifiers, in order to provide more accurate assistive systems, but also to enable an analysis of relevant brain signal characteristics used by the BCI for the classification task.

1.2 BCI or Brain-Computer Interface

BCI refers to the communication and interaction between the human brain and external devices using neural signals and without the use of output pathways such as peripheral nerves and muscles. BCI research and development is important due to its application as assistive technology for disabled people with severe neuromuscular disorders or as a therapeutic tool for people with an impaired neuromuscular function [1]. In the first case, BCI can supply a way of direct communication and control of the environment or alternative locomotion (an electric wheelchair for example), in order to provide more

independence and quality of life for a disabled person and the ones who take care of the patient [2; 3; 4; 5; 6; 7]. In the second case, BCI can be useful for helping people to relearn basic motor functions through neurorehabilitation therapy [7; 8; 9].

1.3 BCI components

BCI systems have basic steps or components: signal acquisition, signal processing, and classification.

1.3.1 Signal acquisition

Brain signals are acquired using electrodes. Then, the signal is amplified and digitized. BCI can use different types of neural signals. Electrical, magnetic, metabolic, and hemodynamic (blood movement) activities of the brain are useful to extract information. Some of the acquisition techniques more popular in BCI applications are magnetoencephalography (MEG) [10], functional magnetic resonance imaging (fMRI) [7], near-infrared spectroscopy (NIRS) [11], positron emission tomography (PET) [12], electroencephalography (EEG), electrocorticography (ECoG) and intracortical methods. Nevertheless, electrophysiological signals acquisition methods of EEG and ECoG have more widespread clinical applications. Also, the development stage of BCI based on these methods is more advanced than the ones based on other types of signals. In the specific case of EEG, a bigger amount of experimental data is available due to the moderate portability and cost of the acquisition equipment [11].

1.3.1.1 Electroencephalography EEG

EEG is a non-invasive signal acquisition method that uses electrodes placed on the scalp to obtain a spatial and weighted sum of neural electrical activity. This electrical activity comes from synaptic excitations of neuron dendrites in the cortex of the brain. Synaptic currents are produced within dendrites when neurons are activated; this current generates an electric field over the scalp that is recorded through the electrodes and translated to potential differences [13]. Due to the location of EEG electrodes, artifacts from power line interference, signals from cranial muscles activity, and even electrical activity from the heart can affect EEG signal fidelity [14]. Also, frequency range and topographical resolution of signal is not wide (EEg signal is a), raw EEG effective bandwidth is limited to approximately 100 Hz with amplitudes from 10 uV to 100 uV [13]. However, EEG signal acquisition method is safe, simple to use and can help to develop BCI portable systems [7; 15; 16; 17; 18].

1.3.1.2 Electrocorticographic ECoG

Compared to EEG, ECoG is an extremely invasive method that acquires signals from the surface of the neural cortex, beneath the skull and over the cortical surface, through electrode arrays that must be surgically implanted usually in epileptic patients [7]. For that reason, experimental data for ECoG studies is not easy to obtain. On the other hand, ECoG signal has a wider frequency range, from DC to 200 Hz, better topographical resolution and signal fidelity because of a closer location from the electrodes to the brain and better isolation from external electrical disturbances compared to EEG signal [19].

1.3.1.3 Brain waves

Inside brain signals some waves can be recognized due to their different frequency ranges. The Delta wave is defined from 0.5 Hz to 4 Hz and corresponds to human deep sleep. Theta wave can be found in the range between 4 Hz and 7.5 Hz and represents activities related to the human unconscious, deep meditation, and creative inspiration. Alpha wave is defined from 8 Hz to 13 Hz and describes awareness but without concentration. Beta wave lies between 14 Hz and 26 Hz and represents active attention and thinking, focus and process of solving concrete problems. Gamma waves with frequencies between 30 Hz to 45 Hz are useful for detection of certain brain diseases but also can be a good indicator of event-related synchronization (ERS) [13]. An additional wave named broadband or high gamma comes from ECoG signal [20]. As ECoG electrodes are located deeper in the brain and away from the skull which behaves as a low-pass filter, ECoG SNR is better than EEG. Also, gamma waves can reflect event-related activity [21; 22; 23].

1.3.2 Signal processing

Brain signal features that can explain the intent of a person is extracted. Signal information can be represented in the time domain, in the frequency domain, or in the time-frequency domain [24; 25]. Time domain provides signal information about the amplitude variations in specific periods of time. Brain signal features usually used for analysis can be linear such as statistics such as mean, maximum and minimum, variance, skewness, and kurtosis, or non-linear features such as Hjorth parameters, entropy, or fractal dimension [26]. Power spectral density (PSD) estimates the signal power of the different brain waves, in order to extract useful features in the frequency domain. EEG time-frequency analysis can be accomplished with the Short Time Fourier Transform (STFT) or the Continue Wavelet Transform (CWT), extracting brain signal features from spectrograms or scalograms [27].

In addition, ECoG and EEG signals may contain relevant information about the tasks that a person executes. For instance, increases/decreases in power in particular signal

frequency bands are related to motor tasks or imaginary motor tasks [28], alpha (8-12 Hz) and beta (16 - 26 Hz) bands have been extensively used in the design of BCIs based on EEG and ECoG [29; 30]. Additionally, modulation of signal components in the band from 70 - 200Hz, High-Frequency Components (HFC), have been found to be correlated with responses to visual stimulation [20], auditory stimulation [31], and motor activity [32]. These brain signal features are extracted and used to train classification algorithms that are able to determine the intention of the subjects.

1.3.3 Classification Methods

Extracted signal features must be translated to control a desirable output from an assistive device. Classification techniques help to provide a proper output based on an adaptive and robust analysis of brain signal features. Classification methods used in BCI systems could be part of machine learning techniques such as Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM), but also artificial intelligence techniques are used as Artificial Neural Networks (ANN) and Deep Learning [7].

1.3.3.1 Linear Discriminant Analysis LDA

LDA is part of a group of classification algorithms that use linear functions to separate classes. LDA uses hyperplanes to divide data from different classes. When there are only two classes, the expression of the hyperplane is presented in Equation (1.1):

$$g(x) = \theta_x + \theta_0 \,, \tag{1.1}$$

where θ_x is a weight vector, and θ_0 is a bias.

The expression above is useful to allocate data in one class when x makes $g(x > 0)$ and in the other class when x makes $g(x < 0)$.

Parameters of hyperplane are estimated by searching for the projection in which the distance between the inter-class means is maximum and the intra-class variance is minimum. Normal distribution of data and equal covariance matrix for both classes must be assumed. In the cases where there are more than two classes, multiple hyperplanes are used [33; 34]. LDA has been widely used as the classifier for different BCI applications [35; 36; 37; 38; 33].

1.3.3.2 Support Vector Machine SVM

SVM also uses hyperplanes to divide data from different classes, but it works in cases where the classes may not be separable by a linear boundary. SVM is able to find the distance between data groups, using the hyperplane to maximize the gap between them.

A mathematical expression for a hyperplane could follow the equation Equation (1.2):

$$g(x) = W^T x + b, \tag{1.2}$$

where W is a vector of coefficients that represent the slopes of the hyperplane [33; 39].

SVM classifier has been applied in numerous BCI applications [40; 41; 42]. SVM leads to a generalization that is less affected by overtraining. However, a SVM disadvantage is that some of its parameters have to be defined manually [33].

1.3.3.3 Artificial Neural Network ANN

ANN is part of the group of nonlinear classification methods and belongs to artificial intelligence techniques. ANN is an assembly of several units (artificial neurons) that provide a general parameterized non-linear mapping between inputs and outputs. Basic ANN usually has a network conformed by an input layer with its units, an output layer and a middle layer(s) called the hidden layer(s) [34; 43]. The network has connections between each input unit and the output of the previous layer, in such a way that the units of the output layer determine the class of the input feature vector. ANNs are applied to numerous problems, including BCI systems because they are flexible and adaptive classifiers where feature extraction is more automated and less arbitrary than in LDA and SVM [44; 45; 46].

1.3.3.4 Deep Learning

Depending on the number of hidden layers in ANN, the classifier can be a shallow neural network (few hidden layers) or a deep neural network (plenty of hidden layers). Hidden or intermediate layers can be useful to construct multiple levels of feature abstraction, which can be an important advantage in the learning process to solve complex classification problems. However, deep neural networks are harder to train because of the intrinsic instability associated with learning by its gradient descent, which stops the training process in some layers.

Deep Learning models like Stacked-autoencoders (SAE), Convolutional Neural Networks (CNN or ConvNets), Recurrent Neural Networks (RNN), and others have been used as classifiers for BCI systems and other applications [43; 47; 48].

Also, since ANN is a black box concerning the model insides, Deep Learning techniques such as CNN or ConvNets can provide clues about the signal or image characteristics used for decision-making. In other words, CNN or ConvNets can improve performance but also offer information about relevant characteristics used for the predictions [49; 50].

Translated brain information is used to control different external devices such as spelling software [51], move a cursor [52], control an electrical wheelchair [53], or use a neuroprosthesis [54].

1.4 Overview contributions in publications

The thesis contributions in terms of publications are:

1.4.1 Research papers in journals

Title: ConvNets for EEG decoding of attempted arm and hand movements of people with spinal cord injury
Journal: Advanced Intelligent Systems
Impact factor of 7.298, classified as Q1 in ISI and Scopus
Authors: Sandra Cancino, Juan Manuel López, Jaime F. Delgado Saa y Norelli Schettini
Submitted in february 20th 2023 Status: Accepted with minor revisions – waiting for publication

1.4.2 Participation in international conferences

Title: Electrocorticographic signals classification for brain computer interfaces using stacked-autoencoders
Conference: SPIE Optics + Photonics 2020
Authors: Sandra Cancino, Jaime Delgado Saa
Date: 24-28 August 2020 (online)
Proceedings Volume 11511, Applications of Machine Learning 2020; 115110J (2020)
https://doi.org/10.1117/12.2568996

Title: Classification of visual stimuli from electrocorticographic recordings using stacked autoencoders
Conference: Neuroscience 2016
Authors: Sandra L. Cancino, Jaime F. Delgado Saa
Date: 12-18 November 2016 (San Diego – USA)

1.5 Thesis Organization

In Chapter 2, there is an overview of Deep learning-based applications in different research areas. Also, some BCI research that uses Deep Learning techniques is described.

The problem statement and the hypothesis are shown in Chapter 3 and Chapter 4.

Chapter 5 presents the first application in this research: Deep learning-based BCI that classifies ECoG signals produced by visual stimuli. This work is focused on the classification performance comparison between the use of the Stacked-autoencoder technique and other state-of-the-art techniques such as SVM.

Chapter 6 presents the second application in this research: Deep learning-based BCI that classifies ECoG signals produced by visual stimuli. This one address the analysis of characteristics learned and used by the Stacked-autoencoder classifier to accomplish predictions.

Chapter 7 presents the third application in this research: Deep learning-based BCI for motor cortex classification of five classes. Here, the proposed method uses some EEG signal processing techniques to enhance time-frequency domain representations, that are used to feed a CNN-based classifier. In addition, the performance of the proposed method is compared to other state-of-the-art research. Finally, an analysis of characteristics learned by the CNN classifier is done.

Finally, Chapter 8 concludes the research work and presents some considerations for possible future work based on the advantages and also limitations of the proposed methods for the described BCI applications.

Chapter 2

Background

2.1 Deep Learning in other applications

Deep Learning techniques have been used in the signal processing area, with applications in speech and audio recognition, computer vision, and physiological signal classification [55; 56]. Applications in speech and audio recognition include speaker and language identification [57], music classification tasks [58], and music recognition [59]. CNN architecture is oriented to computer vision, image classification tasks [60; 61], and scene recognition as a context for object detection [62]. Also, SAE and Deep Belief Networks (DBN) have been used for handwritten single-digit recognition [63; 64]. On the other hand, RNN have been used in character-level language modeling [65], and in designing models that analyze the contents of images and their representation in the domain of natural language in order to provide images descriptions [66].

2.2 Deep Learning in BCI

Since classifiers based on Deep Learning models are used to deal with the numerous variations of brain signals over time and across subjects, its classification performance can be equivalent or superior to state-of-the-art algorithms. Also, Deep Learning classifiers are more automated and do not need the use of handmade features as inputs, as some state-of-the-art classifier algorithms do. But, on the other hand, it is possible to include in the model prior knowledge that is directly related to the problem for reducing model parameters and training time. Finally, Deep Learning techniques can give information from relevant signal characteristics used in the classification BCI stage.

Cecotti et al. [67] propose to use weights of the first hidden layer of a trained CNN to obtain information about the most relevant electrodes, using information from layer features.

Sturm et al. [68] propose a Deep Neural Network (DNN) for classification tasks related to motor-imaginary BCI. Average accuracy from DNN across subjetcs is 74.9% and is comparable to state-of-the-art classifiers however, Layer-wise relevance propagation (LRP) technique [69] is applied to explain DNN classifier decisions through a relevance value for every EEG channel at each time sample named as relevance maps, which are very similar to scalp patterns.

Mirowski et al. [70] proposed CNN for epileptic seizure prediction from invasive EEG signals. Method performance is 100%, better than other techniques as SVM and logistic regression techniques. In this work, the performance is evaluated through false positives per hour and sensitivity estimation. Even though sensitivity results of the three techniques are very similar, CNN has the best performance regarding false positives in the evaluated dataset.

In Jingwei et al. [71], multi-scale deep CNN is used for high-level feature representation in imagined motor EEG. Results show that extracted features are dissimilar between different imagined motor tasks, but they are very similar within the same task across the trials. Classification accuracy from multi-scale deep CNN and Bayesian LDA is the same (100%) but training using the first technique is twice faster.

Drouin-Picaro et al. [72], present CNN and Multi Layer Perceptron (MLP) used for classification of basic saccade directions (up, down, left and right) from frontal EEG data for cursor control without the use of a separate eye tracking device. This work shows an improvement in accuracy results (72.92% and 70.94%) for proposed models compared to two benchmark techniques.

Prior Supervised Convolutional Stacked Auto-encoders (PCSA) to decode finger flexion from ECoG signals is proposed in Wang et al. [73]. Target prior definition complements feature learning of the task of finger flexion. Results show that features learned by PCSA have better performance, in terms of correlation coefficient, than state-of-the-art hand features based on heuristics.

In Martin et al. [74], Deep Belief Nets (DBNs) models use physiological signals of EEG, EOG and EMG for sleeping rate classification. DBNs models use either raw data or handmade features; with a classification accuracy of 72.2%, which is higher than two state-of-the-art algorithms. Authors also show that some learned features in EEG and EMG from raw data DBNs correspond to the handmade selected features.

In Wulsin et al. [75], DBNs with handmade features are used to classify five clinically significant EEG waveforms and measure anomalies. DBNs model has performance comparable in terms of F1 score and classification time to other high-performing classifiers such as SVMs and K-Nearest Neighbors (KNNs).

In Ma et al. [76], a combination between Deep Learning with compressed sensing is applied to obtain motion-onset visual evoked potentials (mVEP) features to perform BCI

task classification. In this work, researchers compare their approach with amplitude-based mVEP feature extraction technique, finding a classification accuracy of 87.5%. A summary of these results can be seen in Table 2.1.

TABLE 2.1: Average accuracy from state-or-the-art research works of BCI using deep learning techniques

Research work	Subjects	Method	Average Accuracy
Cecotti et al. [67]	2	CNN and Multi-CNN	32% and 41%
Sturm et al. [68]	10	DNN	74.9%
Mirowski et al. [70]	21	CNN	100%
Jingwei et al. [71]	1	multi-scale deep CNN	100%
Drouin-Picaro et al. [72]	2	MLP and CNN	72.92% and 70.94%
Martin et al. [74]	25	DBN	72.2%
Ma et al. [76]	11	DL features	87.5%

Chapter 3

Problem Statement

BCI has been a rapidly growing research area, but performance levels are not adequate for BCI applications to reach the reliability of natural communication between human brain and muscle-based function. Therefore, a widely BCI-based clinical application has not been accomplished yet [77]. Specifically, most of the time, research work does not consider prior knowledge as a physiological restriction. Also, in the BCI signal processing stage, there is a wide spectrum of techniques that could be worth exploring to obtain relevant signal information for the next stage. The use of the attributes mentioned above, could provide more relevant signal information as input to the classifier and reduce computational resources and time needed for its training. These could help to accomplish a BCI performance improvement. Finally, there is not much emphasis in the analysis of the extracted features during the classification stage, for most BCI related research.

We propose the design of a deep-neural network based BCI system, that incorporates prior information about brain functionality and proper brain signal processing, before the classification stage; not only to help improve BCI performance, but also to take advantage of the extracted features, used for classification purposes, and further use them to gather more information about signal characteristics, such as frequency range or time-frequency patterns, that can convey meaning about the processes of interest.

Chapter 4

Hypothesis

BCI systems based on Deep Learning techniques that make use of prior physiological information and proper signal processing techniques can provide better performance than state-of-the-art approaches in terms of classification accuracy. Also, the network can provide insight on what the classifier learns in terms of brain signal features and their relations to the processes of interest.

Chapter 5

Application 1, Deep learning-based BCI with visual stimuli: Stacked-autoencoders with ECoG

This chapter is based on the poster titled "Classification of visual stimuli from Electrocorticographic recordings using stacked auto-encoders" [78], and the research paper titled "Electrocorticographic signals classification for brain computer interfaces using stacked-autoencoders" [79].

5.1 Introduction

The use of deep learning classification techniques has been incorporated in the BCI community [68; 80], aiming to improve its performance. In [81] a method based on stacked autoencoders or SAE was used to improve the detection of Event-Related Potentials (P300 ERPs) showing a significant improvement (accuracy of 69.2%) compared to traditional methods like LDA and multilayer perceptrons (MLP). In [82], authors describe a classifier with stacked autoencoders for EEG-based emotion recognition. The results show higher accuracies of valence and arousal of 49.52% and 46.03%, respectively, compared to SVM and naive Bayes techniques.

ConvNets or CNN have been used for BCI applications. Kwak et al. [83] proposed to used them for steady-state visual evoked potentials (SSVEPs) classification paradigm. Results showed an accuracy of 99.28% and 94.03% in the static and ambulatory conditions, higher than other techniques like Canonical correlation analysis (CCA). Xie et al. [84], used CNN and RNN for individual finger flexion decoding, and obtained an

accuracy of 33% in the test set. Tabar et al. [85] used CNN and SAE in combination to classify EEG Motor Imagery signals. A significant performance (accuracy of 77.6%) compared to other decoding techniques such as SVM and separately CNN and SAE was obtained.

In the training process, stacked-autoencoders can learn a more compact representation of input signal, with an increasing abstraction as the network depth also increases [47]. This way, an analysis of this compact representation can give information on useful signal characteristics employed for the classification process. A summary of these results can be seen in Table 5.1.

ConvNets or CNN are also useful for this analysis, but their computational cost is significantly bigger because they require millions of parameters and operations in the training process [86; 87]. Nevertheless, there are some techniques like transfer learning that allow the use of pre-trained models in order to reduce training time and the computational resources needed for the process [88]. This approximation will be further explored in Chapter 7.

TABLE 5.1: Average accuracy from state-or-the-art research works of BCI using deep learning techniques as SAE and ConvNets or CNN

Research work	Subjects	Method	Average Accuracy
Vareka et al. [81]	15	SAE	69.2%
Jirayucharoensak et al. [82]	32	SAE (valence/arousal)	49.5% and 46%
Kwak et al. [83]	7	CNN (SSVEPs)	99.3% and 94%
Xie et al. [84]	3	CNN and RNN	33%
Tabar et al. [85]	9	CNN and SAE	77.6%

Despite the fact that traditionally deep learning techniques require a considerable amount of data to perform well given the number of parameters that need to be learned; this application and the others described above show that with a limited amount of data, it is possible to obtain similar or sometimes a significant improvement compared to state of the art approaches.

5.2 Method

5.2.1 Dataset Description

Ethics statement: All patients participated in a purely voluntary manner, after providing informed written consent, under experimental protocols approved by the Institutional Review Board of the University of Washington (#12193). All patient data was anonymized according to IRB protocol, in accordance with HIPAA mandate. These

data originally appeared in the manuscript "Spontaneous Decoding of the Timing and Content of Human Object Perception from Cortical Surface Recordings Reveals Complementary Information in the Event-Related Potential and Broadband Spectral Change" published in PLoS Computational Biology in 2015 [20].

The selected dataset consists of brain electrical activity acquired from ECoG arrays located on the sub-temporal cortical surface of epileptic patients. Brain signals were recorded when visual stimuli were presented to patients. An amount between 31 and 60 electrodes were used for every subject. A random sequence of gray-scale images from faces and houses was presented in intervals of 400ms, with 400ms blank screen inter-stimulus interval. There were 3 experimental trials for each subject, with 100 visual stimuli in each trial (50 different house images and 50 different face images). Patients performed a basic visual stimuli discrimination task and ECoG potentials were recorded at a sample rate of 1000 Hz. A bandpass filter with cut frequencies between 0.15 to 200 Hz was used for signal pre-processing As information from blank screen inter-stimulus interval is not useful for our system, we only used the 400 ms interval where visual stimuli are shown to patients. See Figure 5.1 for the time diagram description in the experiment.

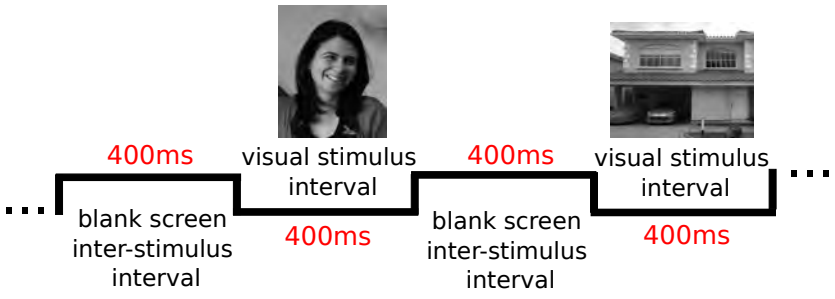

FIGURE 5.1: Time diagram in the experiment.

5.2.2 Signal Processing

5.2.2.1 Spatial Filtering

Spatial filtering was applied to ECoG signals through common average reference (CAR). CAR calculates the signal average across the whole set of electrodes and subtracts it from each of them (see Equation (5.1)) :

$$x(t) = s(t) - \frac{\sum_{i=1}^{N} s_i(t)}{N}, \tag{5.1}$$

where $s(t) = s_1(t), s_2(t), \ldots, s_N(t)$ are the different electrode signals, and N is the number of channels or electrodes.

5.2.2.2 Temporal Filtering

Temporal filtering was done to extract High Frequency components (HFC) and Low Frequency Components (LFC) from signals x(t). A band-pass butterworth filter of 4th order with cut frequencies between 70 Hz and 170 Hz was selected to obtain HFC. Other types of IIR filters like chebyshev and elliptic with different parameters were tested, but the 4th order butterworth filter showed a magnitude and a phase behavior appropriate for our application. Then, signal envelope from HFC was estimated through the absolute values of the analytic signal $x_{analytic}(t)$ (see Equation (5.2)):

$$x_{analytic}(t) = x_{bandpass}(t) + jH\{x_{bandpass}(t)\}, \tag{5.2}$$

where H represents the Hilbert transform.

A low-pass butterworth filter of 4th order with a cut frequency of 20Hz was used to smooth HFC signal envelope, downsampled by a factor of 10 corresponding to a sufficient sample rate according to Nyquist-Shannon sampling theorem. The low-pass filtering is done in order to have a smoothed version of HFC signal envelope, to make its comparison with the other signals easier, thus aiding the classifier to discriminate these signals among classes. Since filtering restricts the signal's highest frequency, it is possible to downsample it.

5.2.2.3 Electrode Selection

Electrodes were selected in accordance with physiological priors and a discriminability measure based on the coefficient of determination. First, brain areas related to visual processing were selected based on the literature [89; 90; 91], which led to the selection of electrodes in the fusiform gyrus, lingual gyrus, and inferior occipital gyrus for each subject. In Fig. 5.2, there is a graphical and not quantitative example corresponding to subjects 1 and 6. The trial average for each of the two classes is presented on the top-left (subject 1) and on the bottom-left (subject 6). It can be seen that the trial average shows different characteristics from one class to the other.

Following this procedure, the degree of discriminability between classes based on the signals of each of the selected electrodes was determined using the coefficient of determination r^2. The r^2 was calculated across trials for each time point on each electrode and it is an indication of the linear relationship between the variations in the signal magnitude and the classes of stimulus presented. Only electrodes with significant differences in r^2 between the baseline and post-stimulus periods were selected. For this, a

paired t-test with $p < 0.05$ was used. Also, in Fig. 5.2, we can see a graphical and not quantitative example corresponding to subjects 1 and 6. The r^2 coefficient for electrode 47 and subject 1 is shown on the top-right, and the r^2 coefficient for electrode 26 and subject 6 is observed on the bottom-right of the figure. In each graphic, y-axis shows the r^2 coefficient amplitude and x-axis shows experiment time in ms. It can be seen that r^2 amplitude increases after the first 200ms (baseline) and after the visual stimulus is presented (last 200ms).

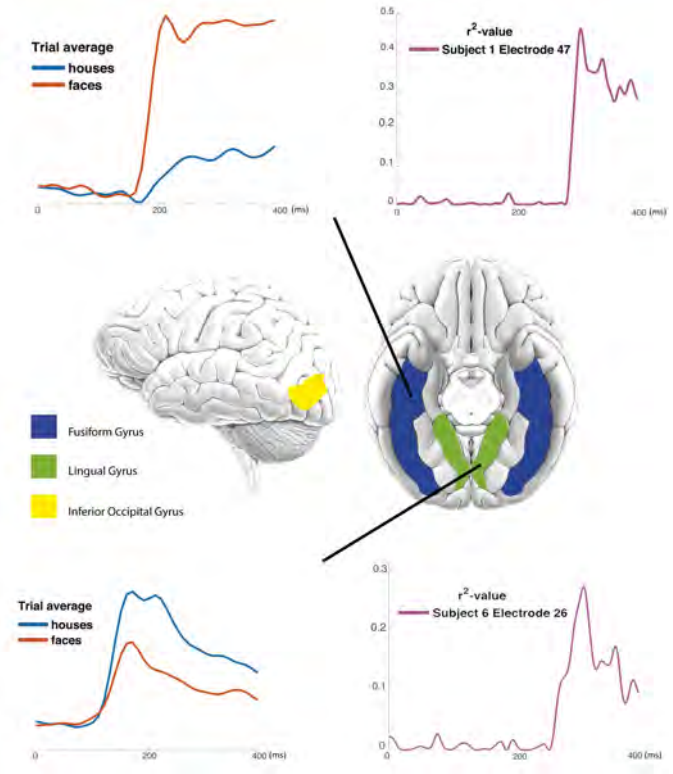

FIGURE 5.2: Brain areas selected for our proposed method. Trial average signal for the two classes and r^2 values for the best discriminative electrode in subject 1 and subject 6.

5.2.3 Classification

Stacked-autoencoders are a specific structure of neural networks that use unsupervised learning (training samples without labels). If $x^1, x^2, x^3, \ldots, x^i$ are i training samples or inputs to the network, an autoencoder uses the same inputs as target values in the backpropagation process, that is $y^i = x^i$. The autoencoder learns then an approximation to the identity function $h_{(W,b)}(x) \approx x$. Autoencoders can estimate a more compact representation of input data if the number of hidden units is lower than the input size [81; 47].

This work proposes a system with two stacked-autoencoders followed by a softmax layer which classifies ECoG signals in two classes of different visual stimuli: houses or faces (see Fig. 5.3). This architecture was first proposed for an example of image classification in [92]. Nevertheless, some architecture modifications were tested until finding the best classification performance in terms of accuracy. The characteristics of the selected classifier are described below.

If the input to the system is of dimension n, the first autoencoder has $n/2$ units and the second autoencoder has $n/4$ units. For instance, if the number of electrodes selected for a subject is 6, the number of inputs n to the first autoencoder is 240 (40 sample points per electrode, representing 400 ms of data). The number of hidden units in the first autoencoder will be 120 ($n/2$) and 60 ($n/4$) for the second autoencoder. In order to show the robustness of the presented method, the results were based on a 10-folds cross-validation procedure.

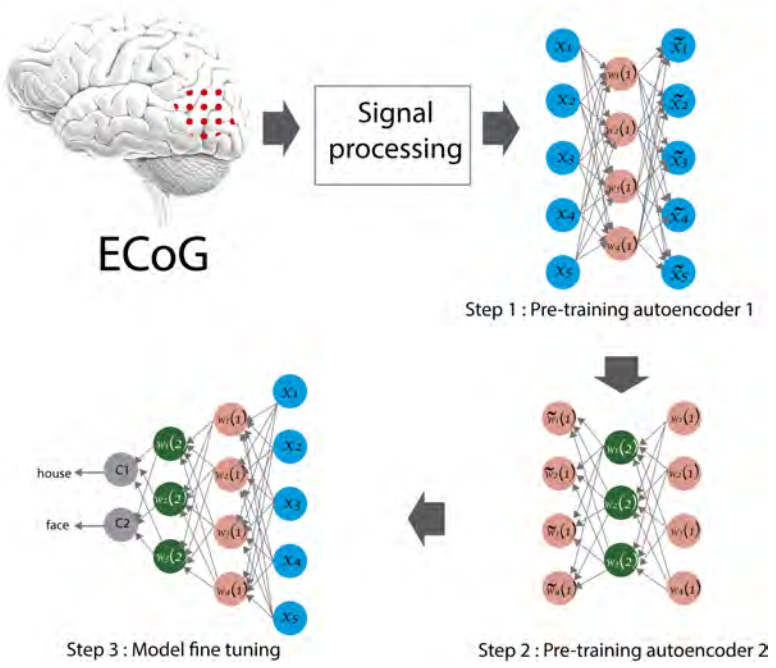

FIGURE 5.3: Network architecture description. Two stacked-autoencoders were trained independently. A softmax layer was added in the model fine-tuning step.

5.3 Results and Discussion

The best average accuracy across subjects obtained was 97% (+/- 0.02%). Nevertheless, this result was similar to state-of-the-art techniques as SVM (see Table 5.2), a big advantage is that a minimum prior information about which features is required for the classification process. Also, the classification accuracy using the different types of spatial filtering in each subject was estimated. The highest result corresponded to the spatial

filtering which works with both LFC and HFC from signals. See Figure 5.4 for these accuracy classification results.

TABLE 5.2: Average accuracy across subjects pf state-or-the-art method SVM and our method

Method	**Accuracy %**
SVM	93.7%
Our method	**97%**

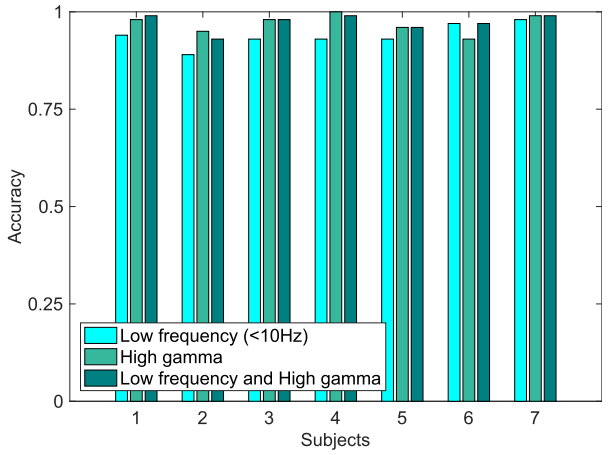

FIGURE 5.4: Accuracy classification results using LFC, HFC or both.

In addition, a confusion matrix was used to estimate other classification performance metrics that could give more specific information about results in each of the two classes. It can be seen in the confusion matrix from the classification using spatial filtering with LFC, that the algorithm classified 4.8% of visual stimuli from house images (class 1) as if they were from face images (class 2). But on the other hand, just a 2.4% of face images are classified wrongly as houses (see Figure 5.5). The same situation is shown in the confusion matrix from the classification using spatial filtering with HFC (see Figure 5.6). This difference in the error percentage between classes could be happening because the human brain has specific areas specialized in face recognition.

When LFC and HFC are used altogether (see Figure 5.7), the algorithm eliminated the error of visual stimuli classification from class 1 (houses). Also, it reduced the error of classifying visual stimuli from class 2 (faces) to 1.9%. The total error of 1.9% was the smallest one. It shows that the combination of LFC and HFC improves global classification performance, but it also shows that the classifier behavior for each class differs when just one of the frequency components is used.

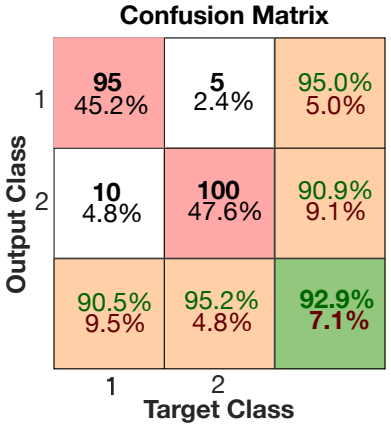

FIGURE 5.5: Confusion matrix using LFC.

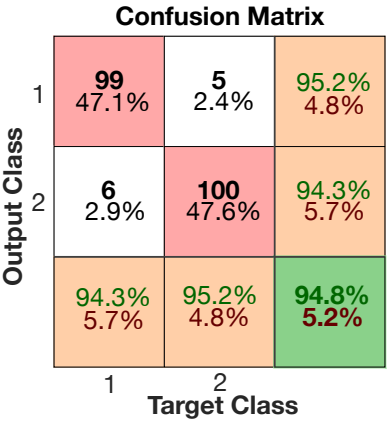

FIGURE 5.6: Confusion matrix using HFC.

5.4 Conclusions

The proposed Deep-learning-based algorithm using stacked-autoencoders with ECoG recordings showed good performance for visual stimuli classification just as state-of-the-art-based techniques. However, this method used minimal prior physiological and an automated statistical technique to select the most discriminative brain signals to train the classifier. However, to the best of our knowledge, little or no effort is spent on the understanding of the parameters learned by the classifier. That was the principal reason to explore this analysis in the next chapter.

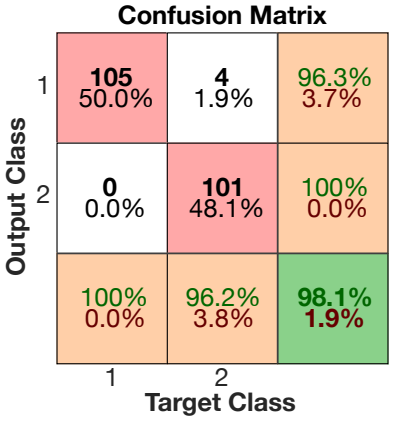

FIGURE 5.7: Confusion matrix using LFC and HFC.

Chapter 6

Application 2, Deep learning-based BCI with visual stimuli: Analysis of Stacked-autoencoders with ECoG

This chapter is based on the research paper titled "Electrocorticographic signals classification for brain computer interfaces using stacked-autoencoders" [79].

6.1 Introduction

The same dataset from the last chapter is used here and most of the proposed methods that were explained already. Nevertheless, this chapter explores the possibility that Deep learning-based BCI can give an insight into what features the algorithm uses to classify and how this is related to basic neurophysiology. An analysis of the classifier's learned parameters is done to achieve this goal and is described in the next sections.

6.2 Method

6.2.1 Dataset Description

A summary of the dataset description is presented as follows.

Electrocorticographic (ECoG) arrays were placed on the subtemporal cortical surface of epileptic patients. Experimental setup characteristics are:

- 7 subjects
- Intervals from 400ms for inter-stimulus and 400ms for visual stimulus
- 31 to 60 electrodes
- 3 trials with 100 visual stimuli
- Random sequence of grayscale images from houses and faces is shown to the subjects

Electrical potentials from the patient's ventral temporal cortical surface were recorded [93]. Electrocorticographic (ECoG) arrays were placed on the subtemporal cortical surface of epileptic patients (see Fig. 6.1).

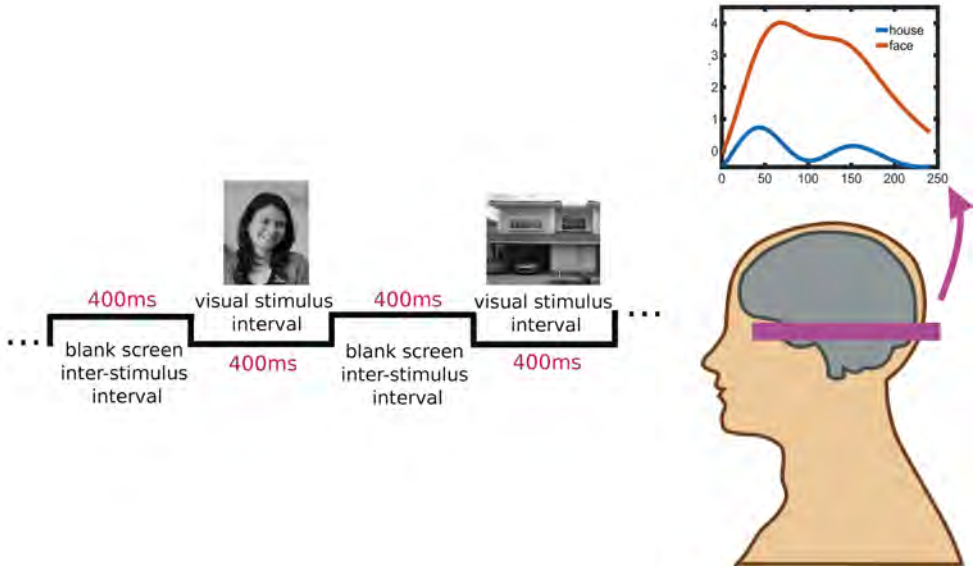

FIGURE 6.1: Experiment description. A random sequence of gray-scale images from faces and houses was shown to patients in intervals of 400ms, in order they could perform a visual stimuli discrimination task. ECoG potentials were recorded.

6.2.2 Signal Processing

Spatial filtering was applied to ECoG signals through common average reference (CAR). CAR is a signal processing technique used to improve signal-to-noise ratio. CAR estimates a signal average across all electrodes and subtracts it from each of them, as follows in Equation (7.1):

$$x(t) = s(t) - \frac{\sum_{i=1}^{N} s_i(t)}{N}, \tag{6.1}$$

where $s(t) = s_1(t), s_2(t), \ldots, s_N(t)$ are the different electrode signals, and N is the number of channels or electrodes.

High Frequency Components HFC were extracted from the signals x(t). A band-pass butterworth filter of 4th order with cut frequencies between 70 Hz and 170 Hz was used to obtain HFB. Then, the signal envelope from HFC was extracted through the calculation of absolute values of the analytic signal $x_{analytic}(t)$, as follows in Equation (6.2):

$$x_{analytic}(t) = x_{bandpass}(t) + jH\{x_{bandpass}(t)\}, \tag{6.2}$$

where H represents the Hilbert transform.

A low-pass butterworth filter of 4th order with a cut frequency of 20Hz was used to smooth HFC signal envelope. Finally, the result signal was downsampled by a factor of 10 which corresponds to a sufficient sample rate according to Nyquist-Shannon sampling theorem.

6.2.3 Electrode Selection

As in the proposed method described in chapter 5, electrodes located in three specific brain regions were used to feed the classifier. Fusiform, Lingual, and Inferior occipital gyrus are brain areas related to face recognition tasks. Also, the electrode selection was based on a t-test (non-paired) between the correlation values in the signal's baseline and the correlations in the time when visual stimuli occurred, to find the most discriminative electrodes and organized them from the most discriminative to the less one.

6.2.4 Network Architecture

As in the last chapter, two stacked autoencoders and a softmax layer were used to classify the two classes of visual stimuli (house and face images) using HFC signal features as inputs. Classifier architecture was selected in such a way that if the input to the system is of dimension n, the first autoencoder has $n/2$ units and the second autoencoder has $n/4$ units. For instance, if the number of electrodes selected for a subject was 6, the number of inputs n to the first autoencoder was 240 (40 sample points per electrode, representing 400 ms of data). The number of hidden units in the first autoencoder might be 120 ($n/2$) and 60 ($n/4$) for the second autoencoder. Also, the results were based on a 10-folds cross-validation procedure.

6.2.5 Neural Network Coefficient Analysis

In order to analyze the features learned by the neural network, the coefficients for each of the autoencoders were observed, finding out that the signal that maximized the activation of each neuron. For this, we extracted for each subject, the most discriminative

electrode in terms of the r^2 value and the first one in the organized list obtained in the previously already explained electrode selection stage. In addition, the autoencoders were trained in the same fashion described above. The set of coefficients that maximally activated each neuron was determined by the following equation (6.3):

$$x_j = \frac{Wij^l}{\sqrt{\sum_{j=1}^{n}(W_{ij}^l)^2}} \tag{6.3}$$

Where x_j is the maximum activation value for input j, W is the weight coefficients matrix in layer l, and i is the number of neuron layer l [47].

6.3 Results

The average responses to faces or houses were calculated for each subject, as well as the r^2 values in the brain regions of interest. Fig. 5.2 shows an example of a representative subject of the activity in the fusiform gyrus and the lingual gyrus. It can be seen that the determination coefficient was higher for the post-stimulus period compared to the baseline ($t < 150$).

The selected electrodes were then used to train the proposed network architecture based on stacked autoencoders. We compared the result to a well-known classifier, SVM (C parameter was calculated using cross-validation and grid search). The results are displayed in Table 6.1.

TABLE 6.1: Comparison between classification accuracy for HFB signal using Neural Network and SVM

Subjects	**NN**	**SVM**
S1	0.99	0.97
S2	0.86	0.81
S3	0.94	0.93
S4	0.99	0.99
S5	0.95	0.95
S6	0.92	0.91
S7	0.99	1.00
p-value		0.172

The proposed method and SVM provided a performance accuracy above the 94%. Even though, the proposed method offered higher accuracy, a paired t-test showed that the differences were not statistically significant at $p < 0.05$ level. However, the proposed method allowed the analysis of the coefficients which shed light on the features that are being learned by the classifier.

Results of the analysis of the network coefficients of the hidden units in the first autoencoder revealed that the signal of maximum activation for each subject is correlated with the average trial response to the stimulus in areas related to visual processing in the brain (see Table 6.2 for a list of the location of the most discriminative electrode in each subject). This suggested that the classifier aimed to find the input sequence with the minimum distance to the average brain activation observed in the training set. The correlation between the class-dependent trial average and the coefficients learned by the classifier (unit with maximum activation) are shown in Table 6.3.

TABLE 6.2: Number, location and brain region of the best discriminative electrode for each subject.

Subjects	**Electrode**	**Brain Region**
S1	47	Fusiform Gyrus
S2	25	Lingual Gyrus
S3	3	Fusiform Gyrus
S4	24	Fusiform Gyrus
S5	16	Fusiform Gyrus
S6	26	Lingual Gyrus
S7	39	Fusiform Gyrus

TABLE 6.3: Biggest absolute value of correlation between trial average from each class and maximum input activation signal for units of first autoencoder in every subject.

Subjects	**Units**	**Correlation**	**P-Value**	**Class**
S1	6	0.62	< 0.0001	houses/faces
S2	2	0.77	< 0.0001	faces
S3	8	0.65	< 0.0001	faces
S4	18	0.79	< 0.0001	faces
S5	8	0.93	< 0.0001	houses
S6	1	0.88	< 0.0001	houses
S7	8	0.63	< 0.0001	faces

Also, results in Table 6.3 show that signal of maximum activation for subjects 2, 3, 4, and 7 is significantly correlated with the average trial from the visual stimulus of faces, and in subjects 5 and 6 the signal of maximum activation is significantly correlated with the average trial from the visual stimulus of houses. Subject 1 shows the same correlation for the two classes of visual stimuli.

In the case of analysis of the weights of the second auto-encoder, results show that the signal that maximally activates all the neurons is highly correlated.

6.4 Discussion

An approach for classifying brain signals using stacked autoencoders was proposed. Physiological information was used to select initially the set of electrodes that were placed in areas related to the visual task executed by the subjects, followed by a second stage of electrode selection based on the coefficient of determination r^2 which could quantify the separability of signals in a particular region when two different types of stimulus were presented. The autoencoders aimed to produce a compact representation of the inputs while preserving as much information as possible. A final layer based on softmax was used to produce a discrete output indicating the class to which the presented stimulus belonged. The proposed approach provided a classification rate of 95% using a 10-fold cross-validation.

A comparison between performance results from the proposed approach and SVM technique was obtained. It is shown that numerical results are in favor of the proposed method, but the results are not statistically significant ($p > 0.05$), meaning that the two approaches perform equally well. However, the proposed method allows the analysis of the weights learned by the network to inquire about what features were being extracted and used to produce the classification output.

The weight analysis is based on the calculation of the input signal that provided the maximum output for each one of the neurons of the first autoencoder. The results showed that the neurons were maximally activated when the input signal corresponded to the trial average of a specific class. This suggests that the task learned by the classifiers was to assign the class label based on the similarity between the input and the average brain response observed in the training set. Furthermore, the results suggested that the classifier showed a bias towards the class that presented lower intra-trial variance. This would explain the fact that for different subjects the signals that maximally activated the neurons were assigned to different classes. The intra-trial variability can be explained since the localization of the electrodes was subject-dependent. Electrodes in the fusiform gyrus could be closer to the face area representation [94] or to the place/object area representation [91] for different subjects.

Finally, results from analysis of the weights of the second autoencoder suggested that the second layer could not compress further the output of the first autoencoder. This type of analysis could be used to determine empirically the deepness required for a particular application.

6.5 Conclusions

In this work, an autoencoder-based method that is capable to decode with high accuracy brain signals acquired during the presentation of two different visual stimuli was

presented. The proposed method also allowed the analysis of the learned weights, to determine the input signal that maximally activated the neurons in the hidden layers. Furthermore, we make emphasis the use of prior information from physiology for the selection of the electrodes of interest, aiming to generate results that were interpretable, rather than combining signals from different areas not related to the task in question.

Future work could be focused on the use of CNN or ConvNets aiming to discover more descriptive features from the brain signals that may help to improve the classification performance and allow for a richer interpretation of the patterns learned by the classifier. Chapter 7 explores this possibility.

Chapter 7

Application 3, Deep learning-based BCI for motor cortex classification of five classes

This chapter is based on the research paper titled "ConvNets for EEG decoding of attempted arm and hand movements of people with spinal cord injury" Submitted and accepted with minor revisions –waiting for publication [95]. Also, this chapter explores the use of pre-trained ConvNets models along with transfer learning to improve performance classification with a moderate computational cost. Results in this final approximation show that a significantly better performance can be achieved compared to state-of-the-art techniques, but also better insights can be gathered from deeper analysis of the ConvNet model layers.

7.1 Introduction

Some BCIs oriented to upper limb disabilities rely on the information that the Movement Related Cortical Potentials (MRCPs) may have. MRCPs are the motor cortex potentials that take place during a movement performance or effort (motor imaginary) and are more notorious in EEG signals from electrodes located in central and middle regions of the head [96]. In their work, Xu et al. [97] tried to classify reach-and-grasp movements: palmar, pinch, push, twist, plug and a final class corresponding to resting state. They used MRCPs and analyzed the difference in cortical EEG features and network structures of the different classes. They projected MRCP into a source space and used its average amplitudes in regions of interest as classification features. In addition, functional connectivity was calculated by means of the phase locking value. Results showed that they could get a similar average accuracy of 49.35%, using source features

with a reduced number of EEG channels. In their research, Wang et al. [98] tried to establish the difference between EEG signals from two different classes of hand movement. Decoding was based on non-linear dynamic parameters of the MRCPs, and classification was achieved using Linear Discriminant Analysis (LDA) model. Results showed significant differences in MRCPs between the different hand movement classes, with an average binary decoding accuracy of 89.5%. Schwarz et al. [99] used EEG recordings from participants who performed self-initiated reach-and-grasp actions toward a glass (palmar grasp) and a spoon (lateral grasp). Their results showed that a multiclass-based decoding approach, including a rest state and MRCPs as inputs to a shrinkage LDA (sLDA) classifier model, could yield a maximum average peak accuracy of 62.3% for a water-based electrode acquisition system.

The EEG and MRCPs decoding process requires an appropriate feature extraction stage along with a classification model, which needs to be trained through an optimization process. Over the last decade, Deep Learning models have been proposed for these tasks due to their outstanding performance [4]. However, the training process of Deep Learning models, such as Convolutional Neural Networks (ConvNets), requires a considerable amount of data due to the number of parameters that need to be estimated. A strategy to deal with this limitation is called Transfer learning [100]. Transfer learning uses a pre-trained network as an efficient approach for small datasets [101]. The learned features from the pre-trained network can be transferred to the training process with the new dataset, requiring only the manipulation of a smaller set of parameters of the model. This manipulation is called "fine-tuning" and requires fewer samples while is less computationally expensive compared to a complete training process. In addition, the fine-tuning of the model may provide good results in terms of generalization.

In a related study, Kumar et al. [102] presented an EEG-based winking signals classification using transfer learning with different architectures for feature extraction in combination with a fine-tuned Random Forest (RF) classifier. Their results showed that the Inception ResNetV2 transfer learning model with the Random Forest classifier had an accuracy of 100% on the training and validation datasets. Also, Sinam et al. [103] proposed a P300 detection-based BCI model using information from a single channel. They used scalogram features from EEG signals to enhance the classification performance. They also used transfer learning with a pre-trained AlexNet as the classifier. Results showed their proposed BCI had a high average information transfer with rates of 13.23 to 26.48 bits/min for disabled people. In their work, Bressan et al. [104] used two datasets with hand movements such as touching, grasping, palmar grasping, and lateral grasping. They used MRCPs as inputs to train a ConvNet model. They compared ConvNet classification performance with an sLDA and an RF model. Results showed that ConvNet had a good performance in both datasets, with an accuracy of 70% and 64%, similar or higher to the LDA and RF models, but with a faster pre-processing. Researchers Khademi et al. [101] proposed hybrid models with pre-trained ConvNets and

Long-Short Term Memory (LSTM) neural networks, for motor imagery classification. Pre-trained ConvNets such as ResNet-50 and Inception-v3 were used to employ more complex features for classification tasks. Transfer learning and data augmentation techniques were adequate for the small dataset they were using, called "BCI Competition IV dataset 2a" [105]. Furthermore, researchers used EEG time-frequency representation obtained from the Continuous Wavelet Transform (CWT) as input images for the ConvNet. The performance results showed a maximum average mean accuracy of 92%. A summary of these results can be seen in Table 7.1.

TABLE 7.1: Related state-or-the-art research works

Research work	Subjects	Features and Classifier	Average Accuracy
Xu et al. [97]	8	MRCP and sLDA	49.35%
Wang et al. [98]	14	MRCP and LDA	89.5%
Schwarz et al. [99]	15	MRCP and sLDA	62.3%
Kumar et al. [102]	5	CWT and RF	100%
Sinam et al. [103]	5 and 4	CWT and ConvNet	92.7% and 93.5%
Bressan et al. [104]	11 and 15	MRCP and ConvNet	70% and 64%
Khademi et al. [101]	9	CWT and ConvNet/LSTM	92%

Another advantage of using ConvNets is the possibility to visualize the learned features of the classification model as a way to understand the decision process of the classifier [100]. Nevertheless, to the best of our knowledge, little effort has been spent on understanding this information in BCI applications.

The pre-trained ConvNet AlexNet [106] is used, in conjunction with a Transfer learning technique, to classify EEG time-frequency information (scalograms) from five different types of hand movements of the BNCI Horizon 2020 database [107]. A novel strategy was proposed, consisting of the MRCPs estimation for its removal from the EEG signals employing an absolute difference; therefore, our approach did not use the MRCPs to classify the different movements. Moreover, we include visualization of the features learned by the model to provide insight into the relevant characteristics for the classification task.

7.2 Materials and Data Processing

7.2.1 Dataset Description

We used BNCI Horizon 2020 database from Ofner et al. [107]. It consists of attempted arm and hand movements in persons with Spinal cord injury (SCI) (accession number

001–2019). It has Electroencephalography (EEG) data from 10 participants with subacute and chronic cervical spinal cord injury in a rehabilitation center (AUVA rehabilitation clinic, Tobelbad, Austria). Participants were asked to execute or attempt different hand movements based on their residual motor abilities. The group of participants was formed by 9 males and 1 female from ages between 20 and 69 years.

EEG signals were recorded using four 16-channel g.USBamps biosignal amplifiers and a g.GAMMAsys/ g.LADYbird active electrode system (g.tec medical engineering GmbH, Austria) with a sampling frequency of 256Hz. The acquisition system preprocessed signals with a band-pass filter from 0.01Hz to 100Hz (8th order Chebyshev filter). Power line interference was eliminated with a notch filter at 50Hz. A total of 61 electrodes covering frontal, central, parietal and temporal areas were used in the signal acquisition. Additionally, electrooculogram (EOG) signals were recorded with 3 electrodes placed above the nasion and below the outer canthi of the eyes. Reference was placed on the left earlobe and ground on AFF2h.

In the experiment, each of the participants sat in front of a computer screen where specific instructions were given. At the trial start, a fixation cross and a beep sound were presented. Participants were asked to focus their gaze on the cross which was displayed during the whole trial period of 5 s to avoid eye movements. The class cue was displayed 2 s after the trial started and, for 3 s until the end of the trial. The class cue could be one of 5 different classes of hand movement: pronation, supination, palmar grasp, lateral grasp, or hand open. Subjects were asked to exclusively execute or attempt the corresponding movement immediately when the class cue was displayed. Finally, there was a break period between trials of 1 s to 3 s. The dataset has recordings of 9 runs with 40 trials per run, and in total 72 trials per class for each of the 10 participants. See Figure 7.1 for the experiment description.

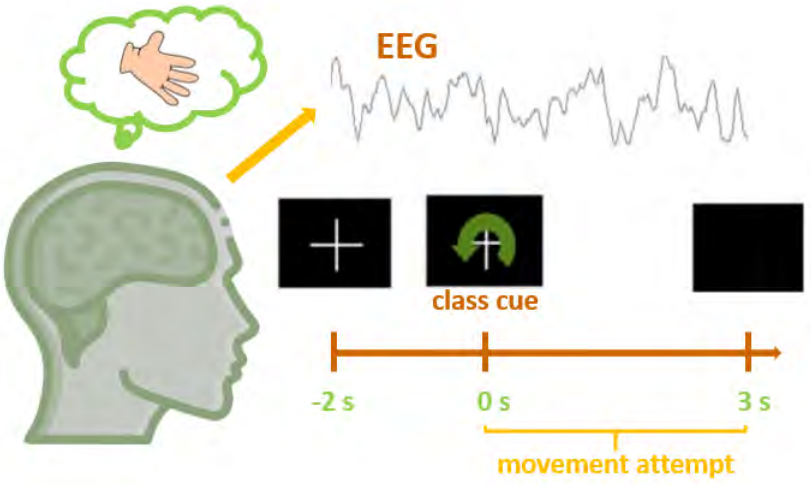

FIGURE 7.1: Experiment description.

7.2.2 Signal Processing

7.2.2.1 Noise Removal

Common Average Reference (CAR) was used for noise removal and improving the signal-to-noise ratio in the 61 EEG channels. CAR estimates a signal average across all electrodes and subtracts it from each of them, as follows in Equation (7.1):

$$y(t) = x(t) - \frac{\sum_{i=1}^{N} x_i(t)}{N}, \tag{7.1}$$

where $y(t)$ is the denoised signal after CAR, $x(t) = x_1(t), x_2(t), \ldots, x_N(t)$ are the different electrode signals, and N is the number of channels or electrodes.

Additionally, EEG signal length was restricted to 2 seconds after the subject executed or attempted the corresponding movement.

7.2.2.2 Movement-related cortical potential MRCP

MRCP has been used as a control signal for BCIs. Previous works have shown that MRCPs can provide meaningful information about different kinds of movements of the same limb (hand open and close, different grasp types, among others). In the specific case of a person with SCI, information on EEG signals from movement attempts or execution could be used by the BCI to control output devices [108]. In their work, Ofner et al. [107] showed that the trial averaged signal on the central electrode Cz presented the typical pattern of MRCPs for movement attempts. In addition, they found that characteristics in MRCPs as the positive and negative peaks could contain discriminative information about the movement class, as well as was found by Zhang et al. research [109]. This fact indicated that information in MRCPs patterns could be used for preparing signals before using them as inputs to the classifier model, to improve its performance. Thus, in our proposed method, we divided the whole dataset into a training set of 80% of data, and a test set of 20% of data. We only used the training set to estimate the average of central electrode Cz signals over trials. In that fashion, we obtained information about MRCPs for each of the 5 classes of attempted movements in each subject. Figure 7.2 is a graphical explanation of the procedure described in this section for the specific subject 2. The average of Cz signals over trials for this subject and for each of the 5 classes is also shown at the bottom-left of this figure. It can be observed that every average signal has MRCP's characteristic peaks no matter to which class it belongs. It is also important to say that the difference between Cz trial average signal amplitudes, for the different classes, is not used as part of the proposed method.

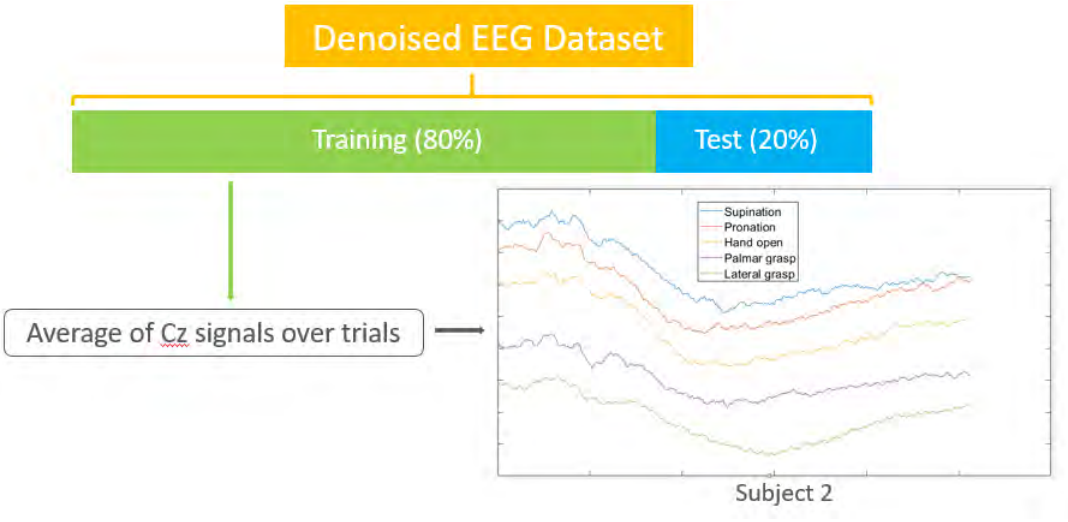

FIGURE 7.2: MRCPs estimation for each class.

7.2.2.3 Independent Component Analysis ICA

Independent Component Analysis (ICA) uses a generative model that describes how the measured signals are produced. In the specific case of EEG, electrode signals or EEG channels are the measured signals. ICA model represents EEG electrode signals as a linear combination of the independent sources of meaningful specific neural activity [110]. ICA model is described in Equation (7.2):

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} = A \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_n(t) \end{bmatrix} \tag{7.2}$$

where **s** is the source signals vector, **A** is the mixing matrix composed by constant elements, and **x** is the electrode signals vector.

In this case, the number of EEG channels was the same number of independent sources. However, there are also some cases in which the number of channels is bigger than the number of sources. In those cases, techniques that reduce dimensionality as Principal Components Analysis (PCA) can be applied before using ICA.

In this work, 10 independent components were extracted from the training and test set of the 61 EEG electrode signals already preprocessed with CAR. The number of independent components was experimentally estimated, meaning that more and less than 10 independent components were tested, but the highest classification performance in terms of accuracy was obtained with 10 components. After that, the absolute difference between the estimated 10 independent components and trial averaged signal of Cz was performed for each class and for each subject, to enhance the information not related

to MRCPs in the independent components. Figure 7.3 is a graphical explanation of the procedure described in this section for the specific subject 2 and for the specific classes of palmar grasp and hand open. The difference between the average of Cz signals over trials for this subject and for the two specific classes, is also shown at the bottom of this figure. Also, the differences between the signal amplitudes, for the two specific classes, are not used as part of the proposed method.

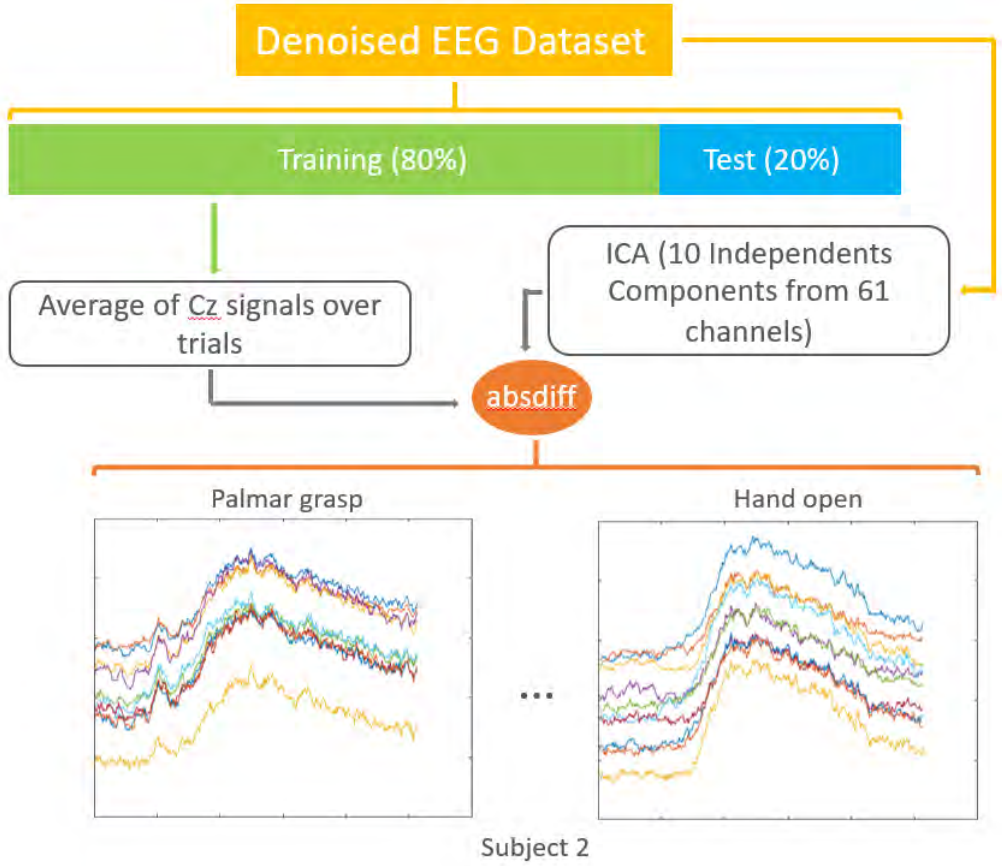

FIGURE 7.3: Preprocessing with ICA and MRCPs.

7.2.2.4 Continuous Wavelet Transform CWT

Continuous Wavelet Transform (CWT) provides a signal representation in the time-frequency domain called scalogram. CWT is widely used in biomedical signal processing because of its advantages in comparison to other time-frequency domain representations as the Short-Time Fourier Transform (STFT). CWT uses time-limited basis functions to decompose and analyze the time-limited events of a signal, rather than using periodic time-unlimited functions as in the STFT, a characteristic that can be useful for the non-stationary nature of biomedical signals as the EEG [111].

In the present investigation, CWT was used to obtain the scalograms of the 10 pre-processed independent components already estimated in the previous step. We used

the Morlet mother wavelet, as it has been widely used for biomedical signal processing [102][112][113]. Also, we established the frequency limits between 0.5 and 100Hz. The 10 scalograms were concatenated over their horizontal axis and saved into a normalized grayscale image. In Figure 7.4, it is shown a graphical explanation of the previous description of CWT stage for the specific subject 2 and specific supination class.

Then, the grayscale image was modified to have a squared size and an RGB color scheme, which were the image characteristics that the selected ConvNet used as input. That was how we obtained a training set and a test set of scalogram images to be used in the next step.

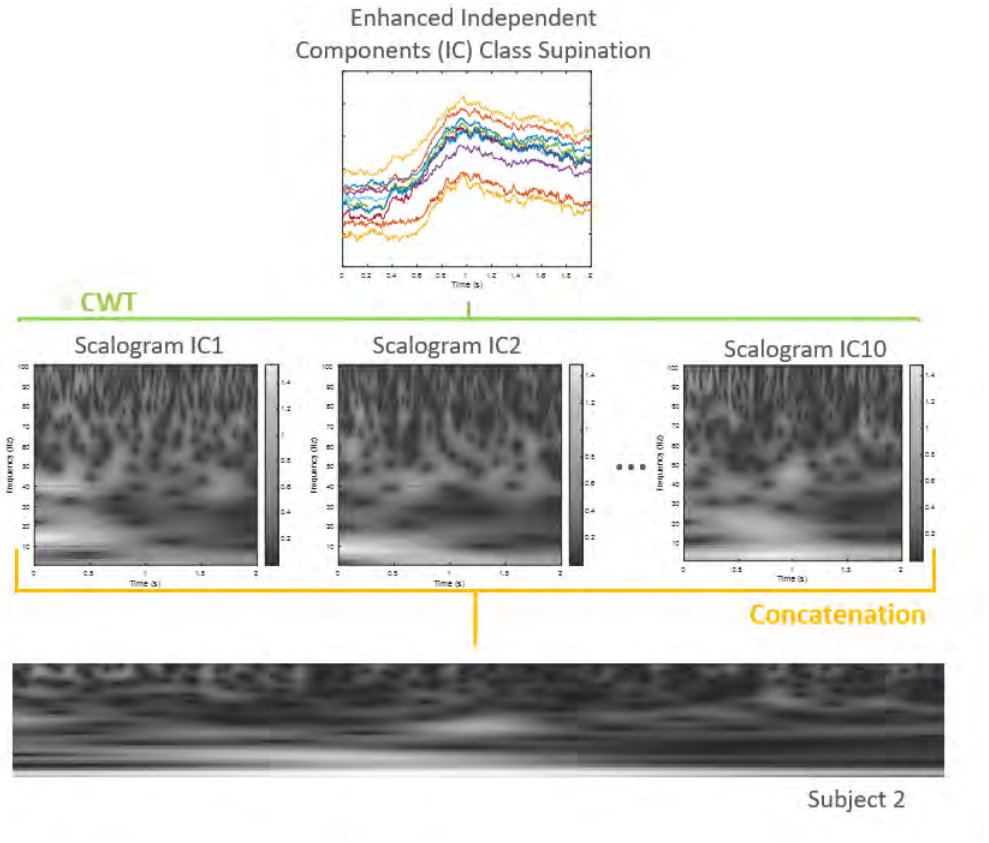

FIGURE 7.4: Concatenated scalogram image from the 10 preprocessed Independent Components.

7.2.3 Feature extraction and Classifier: Transfer Learning

In our method, Transfer learning using AlexNet was done. AlexNet is a pre-trained ConvNet with a training set the of over a million images classified into 1000 classes. AlexNet contains 5 convolutional layers and 3 fully connected or dense layers customized for classifying the original 1000 categories [106]. Through extensive training, Alexnet convolutional layers have learned a big amount of image feature representations. On the other hand, the dense layers have learned image representations that are specific to the 1000 classes of the original model [100]. In our work, a fine-tuning of AlexNet was

done using the training set of scalogram images, in order to modify the dense layers and fit the new classification task of the 5 categories of arm and hand movements. AlexNet network architecture can be seen in Figure 7.5.

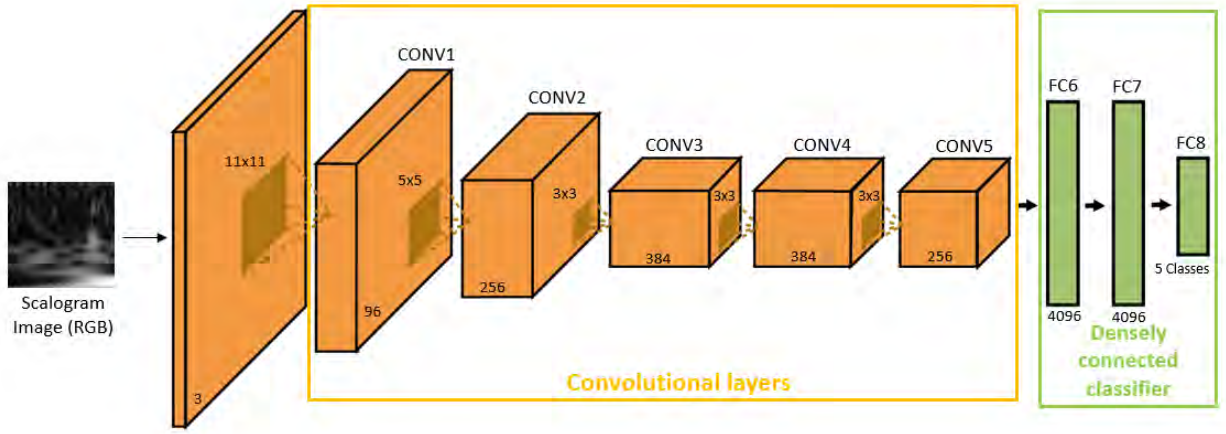

FIGURE 7.5: AlexNet network for classification of 5 types of arm and hand movements.

7.2.4 ConvNet learning visualization

Contrary to other Deep Learning techniques, ConvNets have the advantage of allowing visualization and interpretation of learned image representations. Three techniques were used:

1. **Visualization of ConvNet filters:** shows a visual pattern each layer filter in a ConvNet is maximally responsive to.
2. **Visualization of layer activations:** shows the layer activation or output with a specific input image. It provides information about the input decomposition in different feature maps that are the result of applying the layer filters [100].
3. **Gradient-weighted Class Activation Mapping (Grad-CAM):** uses the gradient of the classification score regarding the ConvNet learned features, to produce a heatmap with the regions of an input image that led to a specific classification result [114].

7.3 Experimental Results and Discussion

7.3.1 Movement Related Cortical Potential MRCP

MRCP characteristic peaks can be observed in the trial averaged signal on the central electrode Cz for each of the five classes. Around the first 0.5 ms after the cue class, the

trial averaged signal shows a positive peak, followed by a negative peak at around 1 s. This morphology was described by Ofner et al. [107]. Figure 7.6 shows the trial averaged signal in Cz electrode and across subjects for each of the five-movement classes. It can be observed in this figure, that these signals exhibit the already mentioned characteristics of MRCPs, which were used as part of the processing in the Independent Component stage. It is important to mention, that the information of statistical signal differences between classes, is not employed in the proposed method.

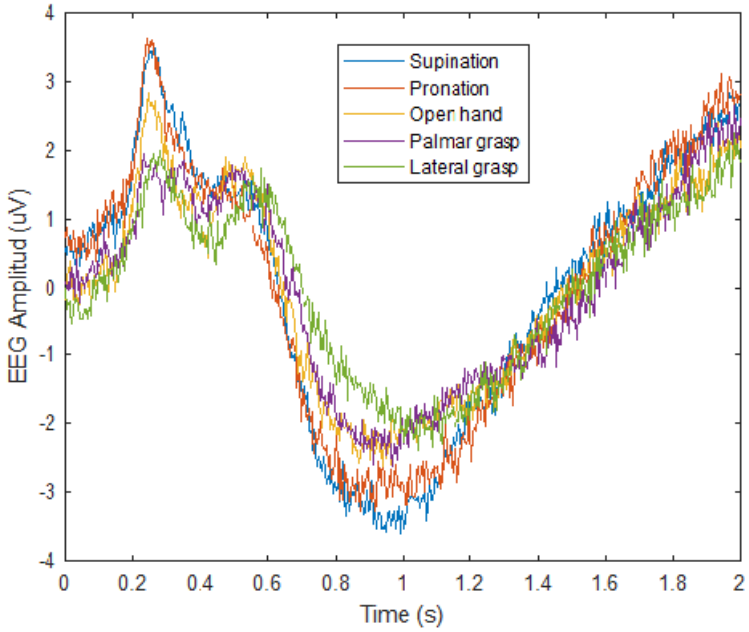

FIGURE 7.6: Grand averages of electrical potentials in Cz electrode for each class.

7.3.2 Performance Evaluation

For the performance evaluation of the classification model, the scalogram test set was used. The computation of the classification accuracy was carried out. It indicates the ratio between the number of accurately predicted observations and the total number of observations. In addition, we used a confusion matrix, which provided two more performance metrics: precision and recall (sensitivity). The precision metric measures the ratio between true positive observations and the total number of predicted positive observations. The recall or sensitivity measures the ratio between true positive observations and the sum of true positive and negative observations [115].

7.3.2.1 Movement classification

The classification of movements by the proposed method showed a maximum mean accuracy of 76%. This is a significant improvement over the results obtained by Ofner

et al. [107], which had a maximum average accuracy of 45.5%. In addition, a better classification performance in each of the five classes was accomplished (see confusion matrix in Figure 7.7). For more details about the performance comparison between the two methods for each class, see Table 7.2.

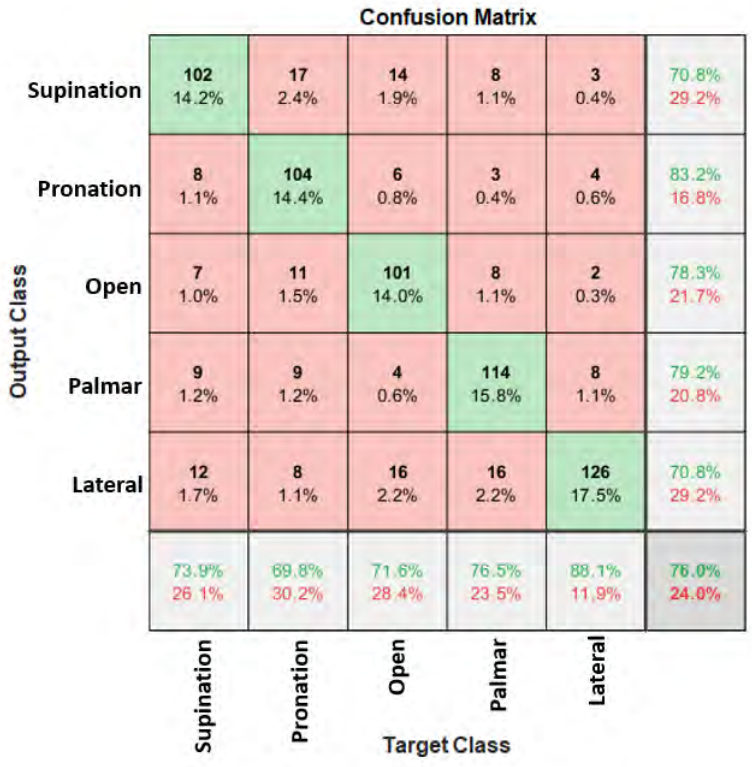

FIGURE 7.7: Confusion matrix using the method proposed for classifications of 5 types of arm and hand movements.

TABLE 7.2: Performance comparison between state-or-the-art method and our method proposed for the five arm and hand movement classes

Precision %					
Method	**Pronation**	**Supination**	**Open**	**Palmar**	**Lateral**
sLDA [107]	47%	42%	47%	43%	45%
Our method	**74%**	**70%**	**72%**	**76%**	**88%**
Recall (Sensitivity) %					
sLDA [107]	47%	42%	46%	43%	46%
Our method	**71%**	**83%**	**78%**	**79%**	**71%**

Other researchers used different datasets related to movement attempts. Their results can be seen in Table 7.3.

TABLE 7.3: Characteristics and average accuracy from other state-or-the-art methods that used different datasets

Research work	Signals	Classes	Method	Accuracy
Aly H. et al. [116]	EEG/EMG	5 wrist and hand mov.	LSTM	95.2%
Bressan et al. [104]	EEG	2 datasets - 4 hand mov.	ConvNet	70% and 64%
Wang et al. [117]	EEG	7 upper limbs motor imagery	VS-LSTM	76.2%
Xu et al. [97]	EEG	6 hand mov.	sLDA	49.4%
Our method	EEG	5 hand mov.	ConvNet	76.0%

7.3.3 ConvNet Learning Visualization

7.3.3.1 Visualization of ConvNet filters

It was possible to visualize the filter patterns of the convolutional layers of our model. This visualization shows how a ConvNet layer can decompose its inputs as a combination of learned filters. Additionally, filter patterns were more complex as one went deeper into the model. Thus, filters from the first layer can encode edges in different directions. Deeper layers learn filters that encode textures made from combinations of edges. Finally, the deepest layer will have the task of encoding more complex textures that can characterize the whole scalogram. Figure 7.8 shows the first 16 filter patterns for layers one, three, and five.

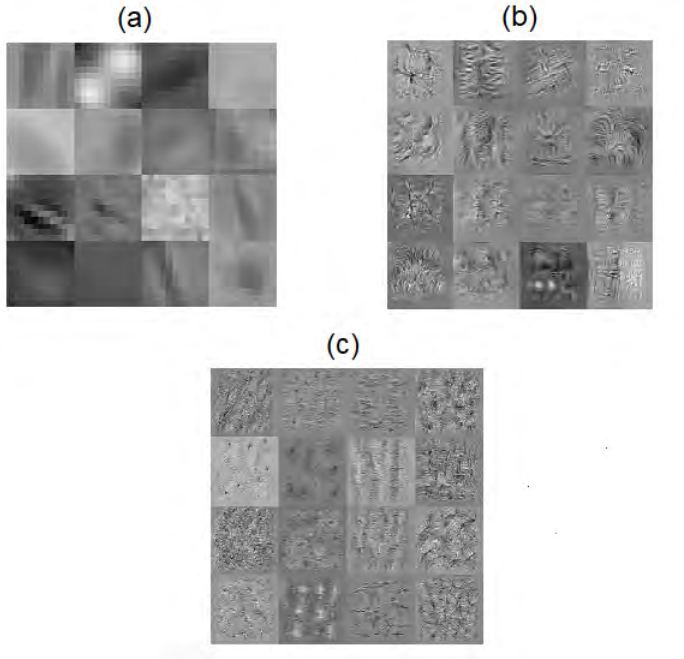

FIGURE 7.8: (a) First 16 filter patterns for first convolutional layer, (b) filter patterns for third layer and, (c) filter patterns for fifth layer.

7.3.3.2 Visualization of layer activations

With a specific scalogram image input, the display of its decomposition into feature maps can be accomplished in the different convolutional and pooling layers of the ConvNet. Figure 7.9 shows some feature maps for the convolutional layers one, three, and five. The selected input is the scalogram number 72 from subject 10 for the supination class. In addition, Figure 7.10 presents the maximum activation for the same scalogram in the first, third, and fifth convolutional layers.

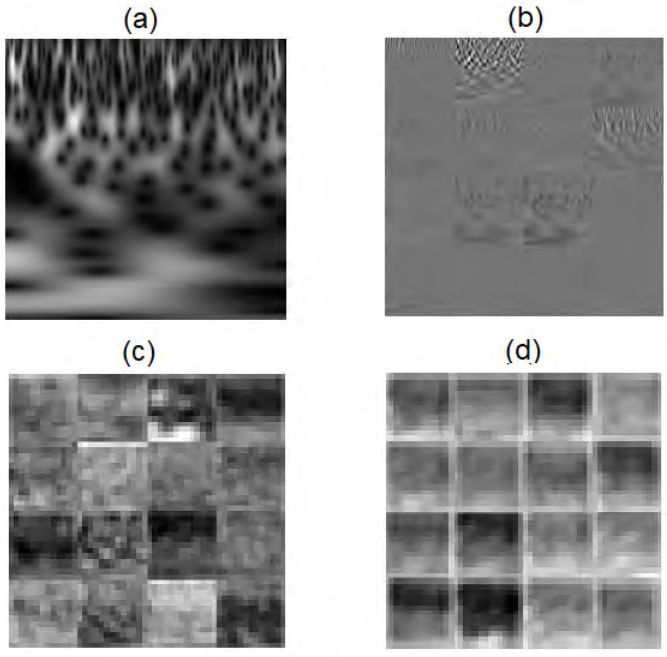

FIGURE 7.9: (a) Input image: scalogram 76 from subject 10 for the supination class, (b) first 16 feature maps for convolutional layer one, (c) third and (d) fifth layer.

7.3.3.3 Gradient-weighted Class Activation Mapping (Grad-CAM)

We used Grad-CAM visualization to make heat maps of class activation over specific input scalograms. Grad-CAM was helpful for understanding which regions of a scalogram are relevant to the model. In Figure 7.11, we can observe the original scalograms and Grad-CAM visualization for pronation and palmar grasp in the same subject. Additionally, in Figure 7.12, we can see the original scalograms and Grad-CAM visualization for supination in two different participants.

7.3.4 Discussion

MRCPs were extracted, and the information contained in the signals has some discriminative patterns, according to Ofner et al. [107]; however, these patterns cannot

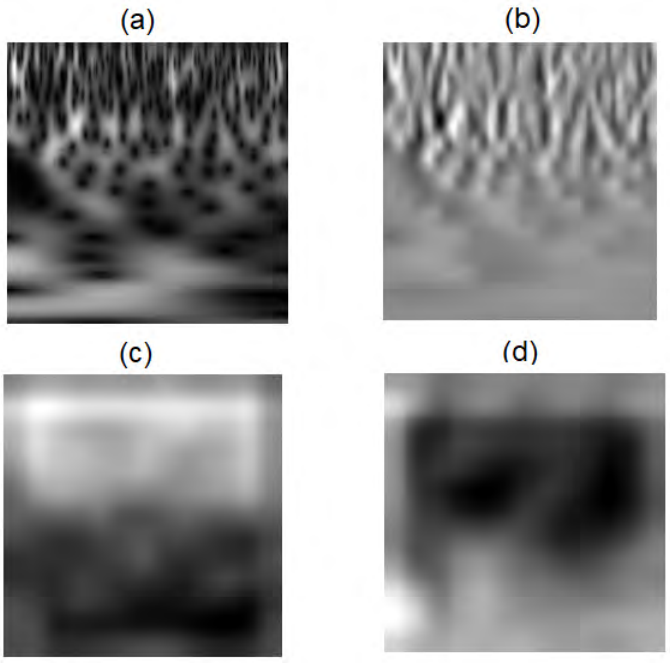

FIGURE 7.10: (a) Input image: scalogram 76 from subject 10 for the supination class, (b) maximun activation for the input image in first, (c) third and, (d) fifth convolutional layer.

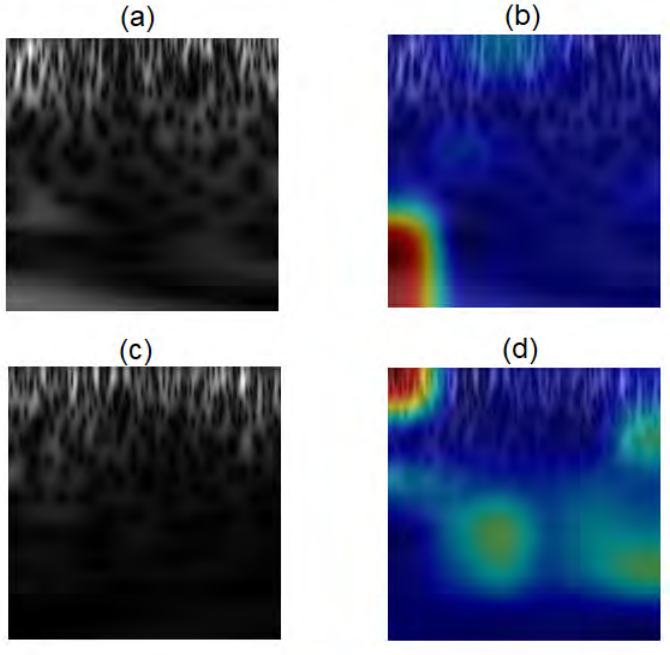

FIGURE 7.11: (a) Scalogram 1 corresponding to pronation in subject 5, (b) Grad-CAM visualization of scalogram in (a), (c) Scalogram 1 corresponding to palmar grasp in subject 5, (d) Grad-CAM visualization of scalogram in (c).

be accurately estimated unless a large set of samples is available. On the other hand, Zhang et al. [109] performed a statistical analysis and found, in some cases, a significant

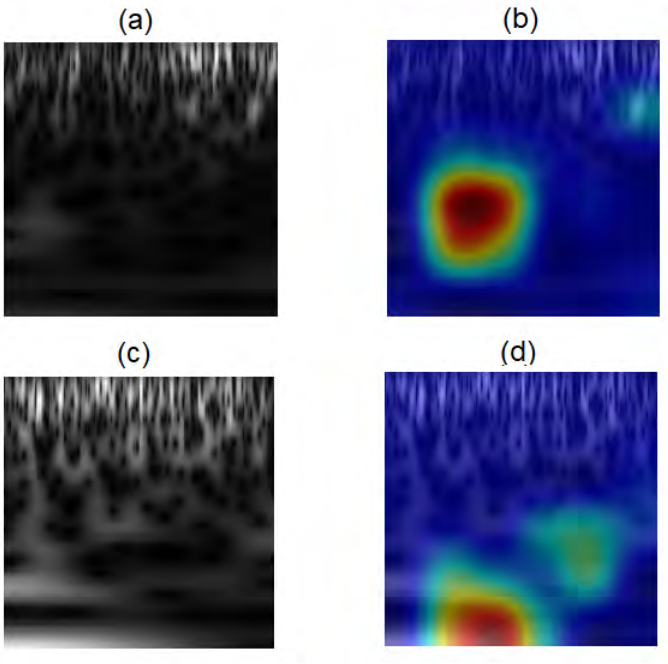

FIGURE 7.12: (a) Scalogram 10 for supination in subject 5, (b) Grad-CAM visualization of scalogram in (a), (c) Scalogram 1 for supination in subject 10, (b) Grad-CAM visualization of scalogram in(c)

difference for the negative peak amplitude, but not the same for the positive peak amplitude for each motion. Therefore, an additional processing step is needed before passing the data to the classification model to improve the performance. The method proposed in this paper shows a different approach, where the MRCP information is disregarded (using the absolute difference), and only the remaining information is analyzed. The results exhibited an increase of more than 25% in average accuracy, compared to Ofner et al. [107] demonstrating that the most relevant features in the EEG, for hand movement commands, might not be present in the MRCPs.

Similar approaches to arm and hand movement or motor imagery decoding were conducted by Bressan et al.[104], Wang et al.[117], Xu et al. [**?**], and Aly et al. [116]. The performance of their models was similar to or inferior to the one presented in this paper for the studies using only the EEG signal. When a combination of EEG and electromyography (EMG) is used, accuracy improved up to 95.2%. Additionally, Wang et al. [117] showed an accuracy of 96.6% for a binary classification problem (motor imagery vs. rest); however, for a multiclass problem, they achieved an acceptable average performance of 78% for the upper limbs motor imagery detection using MRCPs (elbow flexion/extension and forearm supination/pronation) while for the hand imagery (hand open/close), the average accuracy was 72.5%. These results indicate two things: the first one is that binary classification can be achieved successfully from EEG features, but when the decoding of specific hand commands is desired, the same features might not provide the necessary information. Contrary to the other research works shown

in Tables 7.2 and 7.3, we think that MRCPs might be interfering negatively in the multiclass classification process, as our results showed.

Filters in the ConvNet might be related to specific frequency components for the first layers (see Figure 7.8 (a)). However, a change over the x-axis can be observed: there are no horizontal patterns, only diagonal and vertical patterns. For deeper layers (Figure 7.8 (b) and (c)), more complex patterns are observed, where time locations and narrower frequency bands are more relevant. A direct comparison of these patterns and the MRCPs cannot be accomplished, due to the ICA process and the concatenation of the scalograms to form the images that feed the ConvNet. Further analysis might provide insightful information about the physiological process but, to the best of our knowledge, there are no other related works to compare with.

Maximum activation of a specific scalogram for the first convolutional layer showed some vertical edges (Figure 7.10 (a) and (b)), but it changed for the third layer, showing the maximum activation relied on the high frequencies of the independent components (Figure 7.10 (c)), and in the case of the fifth layer is related to the low frequencies (Figure 7.10 (d)). Additionally, these characteristics remained all along the 10 independent components.

From the Grad-CAM results, we can infer that the time-frequency areas in the scalograms vary with the type of movement and participant. Figure 7.11 indicates that in the case of subject 5, for pronation, the low-frequency components at the beginning of the scalogram were the most important in the classification process, while for the palmar grasp, the medium-frequency and high-frequency components were the most important. Something similar is presented in Figure 7.12, in this case with the two different subjects, for the same class: Grad-CAM indicates that for supination, scalograms from subjects 5 and 10 have different regions of interest in the classification problem.

7.4 Conclusions and Future work

Using time-frequency representation along with pre-trained ConvNets as Alexnet, could be a better tool to classify brain signals related to different hand movements. Scalograms, an image representation of EEG, assisted the ConvNet to learn significant spatial patterns useful for the classification process. Also, a priori information related to MRCPs as part of signal processing helped to extract relevant information that led to a significant improvement in the classification performance.

On the other hand, Grad-CAM revealed that independent components and their region of interest changed for every subject and every hand movement class.

For future research work, we would like to use other pre-trained ConvNets architectures to compare classification performance and find the best fit for our method. In addition,

we would like to use our method in a dataset of hand movements attempted by healthy subjects, so that we can compare it with the dataset of subjects with SCI. Furthermore, we want to use information coming from Grad-cam and activations, to accomplish a more detailed analysis from a physiological point of view.

Chapter 8

General Conclusions and Future Work

The main objective of this thesis was to establish that BCI systems based on Deep Learning techniques, that make use of prior physiological information and proper signal processing techniques could provide better performance than state-of-the-art approaches in terms of classification accuracy. Also, to perform a classifier analysis based on its ability to provide information about relevant learned brain signal features and their relations to the processes of interest.

To address this objective, this thesis presented the use of an autoencoder-based method that was capable to decode with high accuracy brain signals acquired during the presentation of two different visual stimuli. Also, it used a few prior physiological and correlation information to select a reduced number of discriminative electrodes to train the classifier.

In the next step, this work showed the analysis of the weights learned by the stacked-autoencoders, to determine the input signal that maximally activated the neurons in the hidden layers. Furthermore, it was shown that the use of prior information from physiology for the selection of the electrodes of interest, aimed to generate results that are interpretable, rather than combining signals from different areas not related to the studied tasks.

Finally, it was shown that using time-frequency signal representation as scalograms and pre-trained ConvNets as Alexnet could accurately classify EEG signals related to different hand movements. Moreover, this image representation of EEG, helped the CNN or ConvNet to learn significant spatial patterns useful for the classification process. Furthermore, MRCPs information used as part of the signal processing led to significantly improved classification performance.

For future research, we would like to use other pre-trained CNN or ConvNets architectures and other signal processing techniques to evaluate their classification accuracy and their potential to allow for a richer interpretation of the patterns learned by the classifier. Also, we want to use information coming from Grad-cam and activations, to accomplish a more detailed analysis from a physiological point of view. Additionally, it would be worth trying other deep-learning-based techniques such as RNN to include temporal features for classification tasks followed by network analysis.

Bibliography

[1] Kaur, M., Ahmed, P., and Rafiq, M. Q., "Technology development for unblessed people using bci: A survey," *International Journal of Computer Applications* **40**, 18–24 (2012).

[2] Scherer, R., Leeb, R., Friedman, D., Gernot, R. M., Slater, M., and Pfurtscheller, G., "Self-paced (asynchronous) bci control of a wheelchair in virtual environments : A case study with a tetraplegic," *Computational Intelligence and Neuroscience* **2007** (2007).

[3] Leeb, R., Sagha, H., Chavarriaga, R., and Mill, R., "A hybrid brain – computer interface based on the fusion of electroencephalographic and electromyographic activities," *Journal of Neural Engineering* **8** (2011).

[4] Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., and Yger, F., "A review of classification algorithms for eeg-based brain-computer interfaces: A 10-year update," *Journal of Neural Engineering* (2018).

[5] Saa, J. F. D., Pesters, A. D., and Mcfarland, D., "Word-level language modeling for p300 spellers based on discriminative graphical models," *Journal of Neural Engineering* **12**, 1–16 (2016).

[6] McFarland, D. J. and Wolpaw, J. R., "Brain-computer interfaces for communication and control," *Communications of the ACM* **54**, 60–66 (2011).

[7] Mak, J. N. and Wolpaw, J. R., "Clinical applications of brain – computer interfaces : Current state and future prospects," *IEEE Reviews in Biomedical Engineering* **2**, 187–199 (2009).

[8] Wolpaw, J., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M., "Brain computer interfaces for communication and control," *Clinical Neurophysiology* **113**, 767–791 (2002).

[9] Allison, B. Z., Mcfarland, D. J., Schalk, G., Dong, S., Moore, M., and Wolpaw, J. R., "Towards an independent brain – computer interface using steady state visual evoked potentials," *Clinical Neurophysiology 119 (2008)* **119**, 399–408 (2008).

[10] Gürkök, H., Nijholt, A., Gürkök, H., and Nijholt, A., "Brain – computer interfaces for multimodal interaction : A survey and principles," *International Journal of Human-Computer Interaction* **7318** (2015).

[11] Vashisht, V., Prasad, T. V., and Prasad, S. V. A. V., "Techonology boon : Eeg based brain computer iinterface - a survey," *International Journal of Computer Science and Mobile Computing* **2**, 447–454 (2013).

[12] Wolpaw, J. R., Loeb, G. E., Allison, B. Z., Donchin, E., Feix, O., Heetderks, W. J., Nijboer, F., Shain, W. G., and Turner, J. N., "Bci meeting 2005 — workshop on signals and recording methods," *IEEE Reviews in Biomedical Engineering* **14**, 138–141 (2006).

[13] Sanei, S. and Chambers, J., [*EEG Signal Processing*], john wiley ed. (2007).

[14] Wolpaw, J. R. and Mcfarland, D. J., "Control of a two-dimensional movement signal by a noninvasive brain – computer interface in humans," *Proceedings of the National Academy of Sciences of the United States of America* **101** (2004).

[15] Cristancho, J. and Delgado-Saa, J. F., "A bootstrapping method for improving the classification performance of the p300 speller," *Revista Mexicana de Ingeniería Biomédica* **Vol. 41** (2020).

[16] Saa, J. F. D. and Cetin, M., "Bayesian nonparametric models for synchronous brain-computer interfaces," *Computer Vision and Pattern Recognition* , 1–7 (2016).

[17] Saa, J. F. and Çetin, M., "A latent discriminative model-based approach for classification of imaginary motor tasks from eeg data," *Journal of Neural Engineering* **9** (2012).

[18] Mcfarland, D. J. and Wolpaw, J. R., "Eeg-based brain – computer interfaces," *Curr Opin Biomed Eng.* , 194–200 (2017).

[19] Schalk, G. and Leuthardt, E. C., "Brain-computer interfaces using electrocorticographic signals," *IEEE Reviews in Biomedical Engineering* **4**, 140–154 (2011).

[20] Miller, K. J., Schalk, G., Hermes, D., Ojemann, J. G., Rajesh, P., and Rao, N., "Spontaneous decoding of the timing and content of human object perception from cortical surface recordings reveals complementary information in the event- related potential and broadband spectral change," *PLOS Computational Biology* , 1–20 (2016).

[21] Crone, N. E., Sinai, A., and Korzeniewska, A., "High-frequency gamma oscillations and human brain mapping with electrocorticography," *Progress in Brain Research* (2006).

[22] Sturm, I., Blankertz, B., Potes, C., Schalk, G., and Curio, G., "Ecog high gamma activity reveals distinct cortical representations of lyrics passages , harmonic and timbre-related changes in a rock song," *Frontiers in human neuroscience* **8**, 1–14 (2014).

[23] Potes, C., Brunner, P., Gunduz, A., Knight, R. T., and Schalk, G., "Neuroimage spatial and temporal relationships of electrocorticographic alpha and gamma activity during auditory processing," *NeuroImage 97* **97**, 188–195 (2014).

[24] Krusienski, D. J., Sellers, E. W., McFarland, D. J., Vaughan, T. M., and Wolpaw, J., "Toward enhanced p300 speller performance," *Journal of Neuroscience Methods* **167**, 15–21 (2009).

[25] Mcfarland, D. J. and Wolpaw, J. R., "Sensorimotor rhythm-based brain-computer interface (bci): model order selection for autoregressive spectral analysis," *Journal of Neural Engineering* **5**, 155–162 (2009).

[26] Shirly, G. and Jerritta, S., "Time domain analysis of electroencephalogram (eeg) signals for word level comprehension in deaf graduates with congenital and acquired hearing loss," *IOP Conference Series: Materials Science and Engineering* **1070**, 012083 (2 2021).

[27] Harpale, V. K. and Bairagi, V. K., "Time and frequency domain analysis of eeg signals for seizure detection: A review," *International Conference on Microelectronics, Computing and Communication, MicroCom 2016* , Institute of Electrical and Electronics Engineers Inc. (7 2016).

[28] Vélez-Lora, H. J., Méndez-Vásquez, D. J., and Delgado-Saa, J. F., "Classification of imaginary motor task from electroencephalographic signals: A comparison of feature selection methods and classification algorithms," *Revista Mexicana de Ingeniería Biomédica* **39**, 95–104 (2018).

[29] Saa, J. F. D. and Cetin, M., "Discriminative methods for classification of asynchronous imaginary motor tasks from eeg data," *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **21**, 716–724 (2013).

[30] Klimesch, W., "Alpha-band oscillations, attention, and controlled access to stored information," *Trends in Cognitive Sciences* **16**, 606–617 (2012).

[31] Cervenka, M. C., Franaszczuk, P. J., Crone, N. E., Hong, B., Caffo, B. S., Bhatt, P., Lenz, F. A., and Boatman-Reich, D., "Reliability of early cortical auditory gamma-band responses," *Clinical Neurophysiology* **124**, 70–82 (2013).

[32] Saa, J. F. D., Pesters, A. D., and Cetin, M., "Asynchronous decoding of finger movements from ecog signals using long-range dependencies conditional random fields," *Journal of Neural Engineering* **13**, 1–8 (2016).

[33] Lotte, F., Congedo, M., Anatole, L., Lamarche, F., and Arnaldi, B., "A review of classification algorithms for eeg-based brain – computer interfaces," *Journal of Neural Engineering* **4**, 24 (2007).

[34] Saa, J. F. D., "Probabilistic graphical models for brain computer interfaces," (2014).

[35] Scherer, R., Müller, G. R., Neuper, C., Graimann, B., and Pfurtscheller, G., "An asynchronously controlled eeg-based virtual keyboard : Improvement of the spelling rate," *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING* **51**, 979–984 (2004).

[36] Garrett, D., Peterson, D. A., Anderson, C. W., and Thaut, M. H., "Comparison of linear, nonlinear, and feature selection methods for eeg signal classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **11**, 141–144 (2003).

[37] Bostanov, V., "Bci competition 2003 — data sets ib and iib : Feature extraction from event-related brain potentials with the continuous wavelet transform and the t -value scalogram," *IEEE Transactions on Biomedical Engineering* **51**, 1057–1061 (2004).

[38] Bi, L., an Fan, X., Jie, K., Teng, T., Ding, H., and Liu, Y., "Using a head-up display-based steady-state visually evoked potential brain – computer interface to control a simulated vehicle," *IEEE Transactions on Intelligent Transportation Systems* **15**, 959–966 (2014).

[39] Hastie, T., Tibshinari, R., and Friedman, J., [*The Elements of Statistical Learning Data mining, Inference and Prediction*], Springer, 2nd ed. (2009).

[40] Garcia, G. N., Ebrahimi, T., and Vesin, J.-M., "Support vector eeg classification in the fourier and time-frequency correlation domains," *Proceedings of the 1st International IEEE EMBS* , 591–594 (2003).

[41] Blankertz, B., Curio, G., and Müller, K.-R., "Classifying single trial eeg: Towards brain computer interfacing," *Advances in Neural Information Processing Systems* **1**, 157–164 (2002).

[42] Rakotomamonjy, A., Guigue, V., Mallet, G., and Alvarado, V., "Ensemble of svms for improving brain computer interface p300 speller performances," *International Conference on Artificial Neural Networks* (2005).

[43] Nielsen, M., "Neural networks and deep learning," (2016).

[44] Palaniappan, R., "Brain computer interface design using band powers extracted during mental tasks," *Proceedings of the 2nd International IEEE EMBS* , 321–324 (2005).

[45] Chai, R., Ling, S. H., Hunter, G. P., Tran, Y., and Nguyen, H. T., "Brain – computer interface classifier for wheelchair commands using neural network with fuzzy particle swarm optimization," *IEEE Journal of Biomedical and Health Informatics* **18**, 1614–1624 (2014).

[46] Forslund, P., "A neural network based brain – computer interface for classification of movement related eeg," (2003).

[47] Andrew, N., Jiquan, N., Yu, F. C., Yifan, M., and Caroline, S., "Ufldl tutorial," (2013).

[48] Heaton, J., [*Artificial Intelligence for Humans, Volume 3: Deep Learning and Neural Networks*], Heaton Research, Inc. (2015).

[49] Lei, T., Barzilay, R., and Jaakkola, T., "Rationalizing neural predictions," *Conference in Empirical Methods in Natural Language Processing* (2016).

[50] Ribeiro, M. T. and Guestrin, C., "" why should i trust you ?" explaining the predictions of any classifier," *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016).

[51] Donchin, E., Spencer, K. M., and Wijesinghe, R., "The mental prosthesis : Assessing the speed of a p300-based brain–computer interface," *IEEE Transactions on Rehabilitation Engineering* **8**, 174–179 (2000).

[52] Kayagil, T. A., Bai, O., Henriquez, C. S., Lin, P., Furlani, S. J., Vorbach, S., and Hallett, M., "A binary method for simple and accurate two-dimensional cursor control from eeg with minimal subject training," *Journal of NeuroEngineering and Rehabilitation* **16**, 1–16 (2009).

[53] Galan, F., Nuttin, M., Lew, E., Ferrez, P. W., Vanacker, G., Philips, J., and del R. Millan, J., "A brain-actuated wheelchair : Asynchronous and non-invasive brain-computer interfaces for continuous control of robots," *Clinical Neurophysiology* **119**, 2159–2169 (2008).

[54] Muller-Putz, G. R., Scherer, R., Pfurtscheller, G., and Rupp, R., "Eeg-based neuroprosthesis control : A step towards clinical practice," *Neuroscience Letters* **382**, 169–174 (2005).

[55] Yu, D. and Deng, L., "Deep learning and its applications to signal and information processing," *IEEE SIGNAL PROCESSING MAGAZINE* **154**, 145–150 (2011).

[56] Längkvist, M., Karlsson, L., and Loutfi, A., "A review of unsupervised feature learning and deep learning for time-series modeling," *Pattern Recognition Letters* **42**, 11–24 (2014).

[57] Richardson, F., Reynolds, D., and Dehak, N., "Deep neural network approaches to speaker and language recognition," *IIEEE SIGNAL PROCESSING LETTERS* **22**, 1671–1675 (2015).

[58] Lee, H., Largman, Y., Pham, P., and Ng, A. Y., "Unsupervised feature learning for audio classification using convolutional deep belief networks," *Advances in Neural Information Processing Systems 22* , 1–9 (2009).

[59] Humphrey, E. J., Bello, J. P., and LeCun, Y., "Feature learning and deep architectures : New directions for music informatics," *Journal of Intelligent Information Systems* **41** (2013).

[60] Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., and Fei-Fei, L., "Large-scale video classification with convolutional neural networks," *Conference on Neural Information Processing Systems (NIPS)* (2014).

[61] Zeiler, M. D. and Fergus, R., "Visualizing and understanding convolutional networks," *European Conference on Computer Vision* (2012).

[62] Zhou, B., Lapedriza, A., Xiao, J., Torralba, A., and Oliva, A., "Learning deep features for scene recognition using places database," *Advances in Neural Information Processing Systems* (2014).

[63] Hosseini-asl, E., Zurada, J. M., and Nasraoui, O., "Deep learning of part-based representation of data using sparse autoencoders with nonnegativity constraints," *IEEE Transactions on Neural Networks and Learning Systems* **27**, 2486–2498 (2016).

[64] Lee, H., Grosse, R., Ranganath, R., and Ng, A. Y., "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations," *Proceedings of the 26th Annual International Conference on Machine Learning* , 609–616 (2009).

[65] Hermans, M. and Schrauwen, B., "Training and analyzing deep recurrent neural networks," *Advances in Neural Information Processing Systems 26 (NIPS)* , 1–9 (2013).

[66] Karpathy, A. and Fei-Fei, L., "Deep visual-semantic alignments for generating image descriptions," *IEEE Conference on Computer Vision and Pattern Recognition* (2015).

[67] Cecotti, H. and Graser, A., "Convolutional neural networks for p300 detection with application to brain-computer interfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**, 433–445 (2011).

[68] Sturm, I., Lapuschkin, S., Samek, W., and robert Müller, K., "Interpretable deep neural networks for single-trial eeg classification," *Journal of Neuroscience Methods* **274**, 141–145 (2016).

[69] Bach, S., Binder, A., Montavon, G., Klauschen, F., Muller, K.-R., and Samek, W., "On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation," *PLoS ONE* **10**, 1–46 (2015).

[70] Mirowski, P. W., Lecun, Y., Madhavan, D., and Kuzniecky, R., "Comparing svm and convolutional networks for epileptic seizure prediction from intracranial eeg," *IEEE Workshop on Machine Learning for Signal Processing,* (2008). NULL.

[71] Jingwei, L. I. U., Yin, C., and Weidong, Z., "Deep learning eeg response representation for brain computer interface," *Proceedings of the 34th Chinese Control Conference* , 3518–3523 (2015). NULL.

[72] Drouin-Picaro, A. and Falk, T. H., "Using deep neural networks for natural saccade classification from electroencephalograms," *IEEE EMBS International Student Conference (ISC)* (2016).

[73] Wang, Z., Lyu, S., Schalk, G., and Ji, Q., "Deep feature learning using target priors with applications in ecog signal decoding for bci," *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence* , 1785–1791 (2013).

[74] Martin, L., Karlsson, L., and Loutfi, A., "Sleep stage classification using unsupervised feature learning," *Advances in Artificial Neural Systems* **2012** (2012).

[75] Wulsin, D. F., Gupta, J. R., Mani, R., Blanco, J. A., and Litt, B., "Modeling electroencephalography waveforms with semi-supervised deep belief nets : fast classification and anomaly measurement," *JOURNAL OF NEURAL ENGINEERING* **036015** (2011).

[76] Ma, T., Li, H., Yang, H., Lv, X., Lia, P., Liua, T., Yao, D., and Xu, P., "The extraction of motion-onset vep bci features based on deep learning and compressed sensing," *Journal of Neuroscience Methods* **accepted m** (2016). NULL.

[77] Shih, J. J., Krusienski, D. J., and Wolpaw, J. R., "Brain-computer interfaces in medicine," *Mayo Clinic Proceedings* **87**, 268–279 (2012).

[78] Cancino, J. F. D. S. S. L., "Classification of visual stimuli from electrocorticographic recordings using stacked autoencoders," *Proceedings of Neuroscience 2016* (2016).

[79] Cancino, S. and Saa, J. F. D., "Electrocorticographic signals classification for brain computer interfaces using stacked-autoencoders," 18, SPIE-Intl Soc Optical Eng (8 2020).

[80] Kao, J. C., Stavisky, S. D., Sussillo, D., Nuyujukian, P., and Shenoy, K. V., "Information systems opportunities in brain – machine interface decoders," *Proceedings of the IEEE* **102**, 666–682 (2014).

[81] Vareka, L. and Mautner, P., "Stacked autoencoders for the p300 component detection," *frontiers in Neuroscience* **11**, 1–9 (2017).

[82] Jirayucharoensak, S., Pan-Ngum, S., Israsena, P., Jirayucharoensak, S., Pan-Ngum, S., and Israsena, P., "Eeg-based emotion recognition using deep learning network with principal component based covariate shift adaptation," *The Scientific World Journal* **2014**, 10 (2014).

[83] sang Kwak, N., robert Mu, K., and whan Lee, S., "A convolutional neural network for steady state visual evoked potential classification under ambulatory environment," *PLoS ONE* , 1–20 (2017).

[84] Xie, Z., Schwartz, O., and Prasad, A., "Decoding of finger trajectory from ecog using deep learning," (2017).

[85] Tabar, Y. R. and Halici, U., "A novel deep learning approach for classification of eeg motor imagery signals," *Journal of Neural Engineering* **14** (2017).

[86] Maji, P. and Mullins, R., "On the reduction of computational complexity of deep convolutional neural networks," *Entropy* **20** (4 2018).

[87] Verma, A., Meenpal, T., and Acharya, B., "Computational cost reduction of convolution neural networks by insignificant filter removal," (2022).

[88] Rahman, T., Chowdhury, M. E., Khandakar, A., Islam, K. R., Islam, K. F., Mahbub, Z. B., Kadir, M. A., and Kashem, S., "Transfer learning with deep convolutional neural network (cnn) for pneumonia detection using chest x-ray," *Applied Sciences (Switzerland)* **10** (5 2020).

[89] Estomih, M., Gruener, G., Dockerty, P., and Fitzgerald, M., [*FitzGerald Neuroanatomía clínica y neurociencia*], Elsevier (2017).

[90] Afifi, K., Bergman, R., Orizaga, J., and Sandoval, J., [*Neuroanatomía funcional : Texto y atlas*], México : McGraw-Hill Interamericana (2006).

[91] Mechelli, A., Humphreys, G. W., Mayall, K., Olson, A., and Price, C. J., "Differential effects of word length and visual contrast in the fusiform and lingual gyri during," *Proceedings of the Royal Society B: Biological Sciences* **267**, 1909–1913 (2000).

[92] Mathworks, "Train stacked autoencoders for image classification."

[93] Miller, K. J. and Ojemann, J. G., "Data and analyses for "spontaneous decoding of the timing and content of human object perception from cortical surface recordings reveals complementary information in the event-related potential and broadband spectral change"," *Stanford Digital Repository* (2015).

[94] Rossion, B., Schiltz, C., and Crommelinck, M., "The functionally defined right occipital and fusiform "face areas" discriminate novel from visually familiar faces," *NeuroImage* **19**, 877–883 (2003).

[95] Cancino, S., López, J. M., Delgado, J. F., Norelli, S., Cancino, S. S., and Schettini, N., "Convnets for eeg decoding of attempted arm and hand move-ments of people with spinal cord injury." Submited in february 20th 2023 to Advanced Intelligent Systems Journal.

[96] Mirzabagherian, H., Sardari, M. A., Menhaj, M. B., and Suratgar, A. A., "Classification of raw spinal cord injury eeg data based on the temporal-spatial inception deep convolutional neural network," *9th RSI International Conference on Robotics and Mechatronics, ICRoM 2021* , 43–50 (2021).

[97] Xu, B., Deng, L., Zhang, D., Xue, M., Li, H., Zeng, H., and Song, A., "Electroencephalogram source imaging and brain network based natural grasps decoding," *Frontiers in Neuroscience* **15** (11 2021).

[98] Wang, J., Bi, L., and Fei, W., "Using non-linear dynamics of eeg signals to classify primary hand movement intent under opposite hand movement," *Frontiers in Neurorobotics* **16** (4 2022).

[99] Schwarz, A., Escolano, C., Montesano, L., and Müller-Putz, G. R., "Analyzing and decoding natural reach-and-grasp actions using gel, water and dry eeg systems," *Frontiers in Neuroscience* **14** (8 2020).

[100] Chollet, F., [*Deep Learning with Python*], Manning (2017).

[101] Khademi, Z., Ebrahimi, F., and Kordy, H. M., "A transfer learning-based cnn and lstm hybrid deep learning model to classify motor imagery eeg signals," *Computers in Biology and Medicine* **143** (4 2022).

[102] Kumar, J. L. M., Rashid, M., Musa, R. M., Razman, M. A. M., Sulaiman, N., Jailani, R., and Majeed, A. P. A., "The classification of eeg-based winking signals: A transfer learning and random forest pipeline," *PeerJ* **9** (3 2021).

[103] Singh, S. A., Meitei, T. G., Devi, N. D., and Majumder, S., "A deep neural network approach for p300 detection-based bci using single-channel eeg scalogram images," *Physical and Engineering Sciences in Medicine* (2021).

[104] Bressan, G., Cisotto, G., Müller-Putz, G. R., and Wriessnegger, S. C., "Deep learning-based classification of fine hand movements from low frequency eeg," *Future Internet* **13** (5 2021).

[105] Brunner, C., Leeb, R., Müller-Putz, G. R., Schlögl, A., and Pfurtscheller, G., "Bci competition 2008-graz data set a experimental paradigm," *BCI Competition 2008-Graz data set A* (2008).

[106] Krizhevsky, A., Sutskever, I., and Hinton, G. E., "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems 25 (NIPS 2012)* (2012).

[107] Ofner, P., Schwarz, A., Pereira, J., Wyss, D., Wildburger, R., and Müller-Putz, G. R., "Attempted arm and hand movements can be decoded from low-frequency eeg from persons with spinal cord injury," *Scientific Reports* **9** (12 2019).

[108] Gu, Y., Dremstrup, K., and Farina, D., "Single-trial discrimination of type and speed of wrist movements from eeg recordings," *Clinical Neurophysiology* , 1596–1600 (2009).

[109] Zhang, M., Wu, J., Song, J., Fu, R., Ma, R., chuan Jiang, Y., and Chen, Y. F., "Decoding coordinated directions of bimanual movements from eeg signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (2022).

[110] Semmlow, J. L., [*Biosignal and Medical Image Processing*], CRC Press (10 2008).

[111] Najarian, K. and Splinter, R., [*Biomedical Signal and Image Processing*], CRC Press, 2 ed. (4 2016).

[112] Shukla, R., Kumar, B., Gaurav, G., Singh, G., and Sahani, A. K., "Epileptic seizure detection using continuous wavelet transform and deep neural networks," *Lecture Notes in Electrical Engineering* **886**, 291–300 (2022).

[113] Aslan, Z. and Akin, M., "A deep learning approach in automated detection of schizophrenia using scalogram images of eeg signals," *Physical and Engineering Sciences in Medicine* **45**, 83–96 (3 2022).

[114] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D., "Grad-cam: Visual explanations from deep networks via gradient-based localization," *IEEE International Conference on Computer Vision (ICCV)* (10 2016).

[115] Sokolova, M. and Lapalme, G., "A systematic analysis of performance measures for classification tasks," *Information Processing and Management* **45**, 427–437 (7 2009).

[116] Aly, H. and Youssef, S. M., "Bio-signal based motion control system using deep learning models: a deep learning approach for motion classification using eeg and emg signal fusion," *Journal of Ambient Intelligence and Humanized Computing* (2 2021).

[117] Wang, P., Wang, M., Zhou, Y., Xu, Z., and Zhang, D., "Multiband decomposition and spectral discriminative analysis for motor imagery bci via deep neural network," *Frontiers of Computer Science* **16** (10 2022).