UNIVERZA NA PRIMORSKEM
FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN
INFORMACIJSKE TEHNOLOGIJE

Zaključna naloga
(Final project paper)

Naslov zaključne naloge

(Title of final project paper in English)

Ime in priimek: Marko Jankovic
Študijski program: (name of the study program, in Slovene)
Mentor: (with all titles, in Slovene)
Somentor: (with all titles, in Slovene)

Koper, mesec leto

Ključna dokumentacijska informacija

Ime in PRIIMEK: Marko Jankovic

Naslov zaključne naloge:

Kraj:

Leto: Koper

Število listov:	Število slik:	Število tabel:
Število prilog:	Število strani prilog:	Število referenc:

Mentor:

Somentor:

Ključne besede:

Math. Subj. Class. (2010):

Izvleček:
Izvleček predstavlja kratek, a jedrnat prikaz vsebine naloge. V največ 250 besedah nakažemo problem, metode, rezultate, ključne ugotovitve in njihov pomen.

Key words documentation

Name and SURNAME:

Title of final project paper:

Place: Koper

Year:

Number of pages:	Number of figures:	Number of tables:
Number of appendices:	Number of appendix pages:	Number of references:

Mentor: title First Name Last Name, PhD

Co-Mentor:

Keywords:

Math. Subj. Class. (2010):

Abstract:

Acknowledgement

Here we thank all involved with our final project paper, that is, persons or institutions that helped us in our work and/or made it possible. We can also thank the mentor and the co-mentor (if there is one).

Contents

List of Tables

List of Figures

Appendices

List of Abbreviations

i.e. that is
e.g. for example

1 Introduction

Electroencephalogram (EEG) is a noninvasive neuroimaging technique that involves the placement of electrodes on the scalp to record the electrical activity of the brain. This enables researchers to measure and analyze the electrical signals generated by the brain. These signals offer valuable information on the operating mechanisms of the brain, covering the identification of various neurological disorders and the exploration of cognitive processes such as perception, attention, and memory. However, EEG signals are prone to various artifacts, one of the most frequent being eye blinks, which for adults typically occur 15-20 times per minute. Eye closure changes brain activity, so eye-blink tracking of subjects undergoing EEG analysis is relevant for identifying when a subject blinks, falls asleep, or keeps their eyes closed. Eye blink and muscle movement artifacts are usually unwanted signals in the brain that can contaminate EEG data significantly and affect the accuracy of subsequent analysis of brain activity that is being measured with an electroencephalogram (EEG). Traditional methods for detecting and removing eye blink artifacts involve manual inspection, filtering, or statistical techniques like Independent Component Analysis (ICA). While these methods can work, they are very time-consuming and prone to human error, especially in real-time applications. They also fail to capture the complex spatial and temporal relationships in EEG data, hence they vary for different subjects and recording conditions.

1.1 Purpose of the thesis

In recent years CNNs have shown to be very effective in various domains as they can learn hierarchical features from raw data. This thesis proposes a CNN-based approach to detect signals in EEG data automatically, by demonstrating this ability on detection of eye blink artifacts. The aim is to design a model to detect signals in EEG recordings and with that prove that neural networks can be used to enhance EEG recordings and make EEG-based systems perform better.

2 Background and context

2.1 Electroencephalography (EEG)

Electroencephalography (EEG) is the measurement of the electric potentials on the scalp surface generated (in part) by neural activity originating from the brain. The sensitivity of EEG to changes in brain activity on such a millisecond time scale is the major advantage of EEG over other brain imaging modalities such as functional magnetic resonance imaging (fMRI) or near-infrared spectroscopy (NIRS) that operate on time scales in the seconds to minutes range. Over the past 100 years, neuroscientists and clinical neurologists have made use of EEG to obtain insight into cognitive or clinical disease state by applying a variety of signal processing and statistical analyses to EEG time series. More recently there has been growing interest in making use of statistical modeling of EEG signals to directly control physical devices in Brain-Computer Interfaces. EEG power is typically split up into bands which correspond to different spectral peaks that relate to behavior or cognitive state. These bands are typically defined as the delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-20 Hz), and gamma (>20 Hz). [1]

2.2 Neural Networks

2.3 ICA label

ICA was first created to deal with the cocktail party problem, upon which you attempt to isolate a pertinent conversation from the noise of other conversations at a cocktail party (Hyvarinen and Oja 2000). Applying the ICA to the EEG data involves the decomposition of EEG time series data into a set of components. More specifically, EEG data are transformed to a collection of simultaneously recorded outputs of spatial filters applied to the whole multi-channel data, instead of a collection of simultaneously recorded single-channel data records. Thus, ICA is also a source separation technique that attempts to identify independent sources of variance in the EEG data (Anemuller et al. 2003).

2.3.1 ICA label based artifact removal

3 Methodology

In this section, the proposed methodology is described stepwise. First, a description of the data is given, and then the data pre-processing is explained, followed by the architecture of the model.

3.1 Datasets

3.1.1 Datasets description

This dataset examines the effects of caffeine on brain stimulation and consists of EEG recordings of two categories of participants: the first group which had an intake of regular coffee, and the other which had decaffeinated coffee. The primary difference between the two types of participants was the drinks and hence caffeine content. Participants from both groups were given the chance to add different amounts of sugar to their drinks, which causes an issue with the data but this is not the concern of this analysis.

In relation to our project, this dataset will be used for dealing with artifact identification and detection in EEG data, and while the original research was focused on caffeine stimulation of the brain, we are interested only in artifacts present in EEG signals, such as blinking, movements of muscles, and other activities nearby that interfere with the EEG readings.

3.1.2 Data preprocessing

While EEG recordings tend to contain noise and artifacts such as eye blinking or movement, EEG signals measured from the scalp do not necessarily accurately represent signals originating from the brain. Therefore, it is very essential to apply preprocessing and denoising to the recorded EEG data. Generally, preprocessing steps include the transformations or reorganizations of the recorded EEG data by removing bad or artifact-ridden data without changing clean data (transformation) and segmenting continuous raw signals without change of the data (reorganizations).

Our preprocessing of the data comes in 6 steps:

1. Filter 1 to 50Hz:

We apply a bandpass filter to the EEG data, keeping frequencies between 1 Hz (lower limit) and 50 Hz (upper limit).

1 Hz: This high-pass filter removes very slow components (below 1 Hz) that may correspond to artifacts like slow drifts in signals.

50 Hz: This low-pass filter removes fast components above 50 Hz, such as muscle artifacts and electrical noise.

2. Cutaway first 1000 samples

   We remove the first 1000 samples, which might correspond to initial noise or artifacts at the start of the recording.

3. Re-reference to average

   This step re-references the EEG signals by subtracting the average of all electrodes from each electrode. This is commonly done to minimize noise and make the signals more comparable across channels.

   The assumption of average reference is: the sum of the electric field values recorded at all scalp electrodes (sufficiently dense and evenly distributed) is always 0, and the current passing through the base of the skull to the neck and body is negligible. Since our EEG recording system has enough even channels using average reference makes sense as the overall activity averages to 0.

4. Resample from 600 to 300 Hz

   We resample the data from a sampling rate of 600 Hz to 300 Hz. Resampling reduces the data size while retaining enough frequency resolution for EEG analysis.

5. Check dataset integrity

   This function checks the integrity of the EEG dataset after the preprocessing steps to ensure there are no inconsistencies or issues.

6. Run ICA

   In this step, Independent Component Analysis (ICA) is applied to separate independent sources (e.g., eye blinks, muscle noise, and brain activity) in the EEG signals. This helps in identifying and later removing artifacts, which is a common step after the basic preprocessing (filtering, re-referencing, etc.) has been completed.

3.2 Model selection and description

4 Results

In this chapter, we report the performance of our model using key metrics such as accuracy, precision, and recall. These results show how well the model identifies and classifies artifacts in our test dataset. We will examine the results of our thorough testing process. The outcomes of our model's tests, both qualitative and quantitative, are presented here in our analysis. In addition to highlighting the model's potential, this assessment points out the drawbacks or areas in need of more development.

4.1 Result of training

4.2 Result of testing the model

4.3 Alternative model comparisson

5 Discussion and conclusion

In few sentences we briefly summarize the content of the project paper. This is also the place where we can add some further references for the interested reader.

6 Literature

7 Povzetek naloge v slovenskem jeziku

This chapter contains a longer summary of the final project paper in Slovene, in total length between 4.000 and 10.000 characters (spaces included).

8 Bibliography

[1] Michael Nunez, Paul Nunez, and Ramesh Srinivasan, *Electroencephalography (EEG): Neurophysics, Experimental Methods, and Signal Processing*, Jan. 1, 2016, pp. 175–197, ISBN: 978-1-4822-2097-1, DOI: 10.13140/RG.2.2.12706.63687. *(Cited on page 2.)*

Appendices

A Title of First Appendix

Here we add the first appendix.

B Title of Second Appendix

Here we add the second appendix.