Guilherme Coutinho

M00684745

Middlesex University London

Middlesex University

Faculty of Science and Technology

Developing a Convolutional Neural Network for classification of EEG Sleep Staging

Biomedical Engineering Major Project

Supervisor: Prof. Richard Bayford

May 2021

ABSTRACT

The necessity of a good amount of sleep is a crucial activity that has many impacts on the persons health and behaviour, frequently proper sleep is often being neglected. Because of this, there has been attempts in a search of an automatic monitorization for our own sleep quality via new technologies. However, the methods of sleep monitorization are still dependent on raw EEG data, which requires a sleep technologist time and effort to analyse a night of sleep. Machine learning or deep learning is a technique in computer science that has been winning a great interest in the market and academic environment. These tools consist of networks with the inherent ability to self-learn from a variety of features given in the input. From this method is possible to create a network that can classify the sleep stages of an EEG. In this context, this report has as goal to develop a classification algorithm capable of automatically score the sleep stages. To achieve this algorithm the methods used were the usage of a convolutional neural network (CNN) and new EEG visualization techniques, which has the potential to provide the CNN images with distinct features that helps the classifier to be recognise. This idea was collected after numerous research and knowledge found online which resulted in a 17 layered CNN. The development of this model was written in MATLAB and the applications available in toolboxes from deep learning. The model was tested using traditional validation techniques and achieved reasonable values for the training data and validation data. In addition, tests were made of images with unique features to determine the capacity of feature recognition of the CNN and validated its idea in clinics. To conclude, there was noticed how the number of layers and nodes can impact the CNN and its training accuracy as well as the model validation accuracy and given recommendations for further improvement for the model.

Table of Contents

INTRODUCTION

Sleep disorders affects thousands of people every year. According to a cross-national study of sleep problems in European older adults van de Straat and Bracke (2015), an average of 24.2% of the total European population reported to have been bothered by sleep problems. In there, the prevalence rate varied across the countries, however they successfully shown there is a higher prevalence in women, with the odds ratio of 2.014, and less problems in people aged 60 to 69 versus people between 50 and 59 old (0.898). The study also claims that marriage compared to divorce, and higher socio-economic status were associated with less sleep problems. This data was collected based on a survey made to 54 722 men and women all over Europe.

Although is hard to specify the exact definition of a sleep disorder, we know that lack of sleep has negative impacts on the individual health. The most frequent expression of this conditions is in the mental state, with an increased tiredness during the daytime, that can affect the behaviour, the daytime performance or physical health, ultimately impacting the persons mood, work, and social life, and the person may not be fully aware of how this impact his behaviour (Stores 2007).

A research Telzer *et al* (2013) made on forty-six adolescents, which participated in a magnetic imaging (fMRI), showed that adolescents who reported poorer sleep exhibited a greater risk taking, this was associated with a less recruitment of the dorsolateral prefrontal cortex (DLPFC) during control and a greater activation of the reward processing. On another study, it was also investigated that teenager that failed to keep the recommendation of sleeping 8 to 10 hours per night stated on American Academy of Sleep Medicine, would have the cognitive performance impaired, with a deficit in attention span and decline in academic performance. In there was also claimed that there is a relation between sleep quality and the metabolism and obesity (Kansagra 2020).

In the context to these studies, is sufficient to affirm that sleep is an important health factor that should not be neglected and requires more attention and further research. Unfortunately at the moment, there is no simple direct way of scoring the quality levels of sleep and the only resources available are done manually by sleep technologists that can take up to 1 hour time to study a 8-hour EEG sleep (Biswal *et al* 2017). So, in an attempt to help and investigate modern tools it is necessary to ask a research question: Is it possible to facilitate the job of sleep technologists using new computation techniques?

One technique that has been winning popularity in the market and in academic studies is the classification through machine learning. The method of using machine learning to classify EEG wave forms allows a fast, automatic, and inexpensive way of classifying different stages and artifacts. In addition, it can allow a future independence from specialized people.

The objective of this report is to use the tools of machine learning to develop a network capable of learning and automatically classifying three main sleep stages: NREM, REM and WAKE. To achieve this objective, it was used a methodology that followed these main steps:

a) Define the best neural network that should be used for the training of this classification, the platform, and its necessary tools;
b) Choose the right input and design the network layout layers;
c) Develop the code capable of extracting the necessary features, filter them and learn from them, to then perform classification according to the sleep stages;
d) Validate the code, by using images outside from the input.

Furthermore, methods used in this report were based on exploratory research, to clarify the subject, applied research, to answer specific questions of the problem and quantitative research, to search for specific data, such as statistics. The technical procedures used in this final year report about the construction of the EEG sleep staging classifier were the result of the information collected along these three years course.

This report begins by giving all the background information, after that a brief explanation of the brain regions and functioning of the cortex area, the explanation of the EEG machine and the characteristics of sleep, the tools used to visualise the EEG and its detection of the sleep stages. Next is provided all the necessary information for the function of the machine learning classifier. A general view is given of each layer, as well as its interaction. Afterwards, each layer is explained in detail and analysed and afterwards some tests are made to the network with the goal of validating it. Finally comes the discussion and conclusion of this report and some final remarks and recommendations for future reasearch.

1. BACKGROUND INFORMATION

In this chapter is presented all the background theory found in this report, containing important subjects in respect to the EEG sleep classifier, such as concepts of the brain, the medical equipment, the classifier, and its chosen features for the input.

1.1. BRAIN COMPOSITION

According to current knowledge there are about 86 billion neurons in the brain that talk to each other instantly and constantly, sending and receiving information for all types of functions. The neurons are composed by three parts: the dendrite responsible for receiving a signal, the cell body for expressing the signal and the axon that sends out the signal. To exchange signals dendrites connected to each other via axon-dendrite called synapses. When the brain is exchanging information, it uses a molecule called neurotransmitter that binds to receptors on dendrite, opening channels to cause a potential difference in the neurons membrane that travels across the neuron's membrane as a small voltage (Hammond 2015).

According with Seeley, Stephens, and Tate (2003), the brain function can be defined by the conscious perception, the thinking and the conscient motor activity, that by itself can super impose many other systems. These brain functions can be further represented in two main groups: the basal nuclei and the limbic system. The first being responsible for the control of the muscle activity and posture, by limiting the unintentional movement; the second being responsible for the autonomic response to smell, emotions, humour, and other character functions. To understand better the large complexity of the brain and its specific function regions these two main structures require to be divided into other landmarks to explain its functions. Below is listed the four major brain regions that best define and control the human body.

- The cerebral cortex: is the outer layer and the largest part of the brain, its composed by the grey matter that composes 40% of the total brain volume. It is responsible for specific functions of the sensory areas, motor areas and association areas.
- The corpus callosum: is a wide thick nerve tract, mainly composed of white matter, that act as connective pathway transferring motor, sensory and cognitive information

connecting both hemispheres of the brain.

- The diencephalon: is the division of the forebrain that relays sensory information between brain regions, it also controls many autonomic functions of the peripheral nervous system.
- The brainstem: is the part that connects the spinal cord and the cerebrum, is responsible for the automatic control of several important functions such as heartbeat, breathing, blood pressure and many reflexes.

The analyse of these regions gives us a greater inside of the brain's capabilities and functioning, however we can also notice that the brain has a sort of hierarchy to the function of some regions. From today's knowledge we know that the evolution of the brain in mammals has been accompanied by the enlargement of the cerebral cortex. In our specie, it is this region responsible for what separates us from any other animal on the planet. The cerebral cortex as a key role in higher levels of thoughts, attention and perception, language, memory, reasoning and ultimately consciousness (Hofman 1985).

According to the scientific way of separating the brain functioning, in total the cerebral cortex has 13 functional divisions, however in this chapter we look only at the most important ones.

- The visual cortex, found at the back part of the head, is an area highly specialized in processing information about movement of objects and pattern recognition. It is located at the occipital lobe and receives sensory input from the eyes. It is divided into five areas (V1 to V6).
- The auditory cortex, found at the lateral of the head, is the brain area that performs basic and higher functions from the input from the ears, such as relations in language switching and guessing the components of a sound. This area also has a fundamental role in the spectral temporal, and it is divided into two areas (A1 and A2)
- The motor cortex is the regions of the cerebral cortex involved in planning, control, and execution of movements. Is further divided into five areas (premotor cortex, primary motor cortex, posterior motor cortex and primary somatosensory cortex).
- The somatosensory cortex is an area of the brain that receives and processes all the sensory information from the entire body such as touch, pain, or vibration. This is an important function as it is responsible for the precise localization of each sensation and then the processing to initiate important movements. Is divided into two areas (S1 and S2).

1.2. BRAIN ELETRICAL ACTIVITY AND SLEEP

The electroencephalogram (EEG) is a medical device used to detect the electrical activity from the brain underlying cortical activity, using small metal electrodes that attach to the scalp.

The EEG works due to the activity of the neurons in the cerebral cortex that is somewhat synchronized and occurs in regular firing rhythms (“brain waves”). When electrodes are placed near the scalp, they can pick up variations in the electrical potential that are derived from this regular firing. This analogue action potential signals are then extracted and by using amplification, filtering, and ADC conversion the recorded signal is separated into five waves components of different frequencies that provide specific information about the brain activities. This activity is then converted to an analogue graph relating time and voltage allowing us to read and interpret the different responses of the brain (Teplan, Michal. 2002). The waves are as follows:

- **Alpha 8hz to 12hz, amp 30-50 μV -** Alpha rhythm is seen when the eyes are closed, and the volunteer is relaxed. The greater the brain activation, the lower the alpha activity. Alpha waves are strongest over the occipital cortex (the back of the head) and frontal cortex.
- **Beta 13hz to 30hz, <20 μV -** are the most common rhythm, seen during eyes open, mental activity, and in low frequencies such as 13 to 15 Hz is often related to insomnia, it is reduced in areas of cortical damage and can be accentuated by sedative-hypnotic drugs, such as benzodiazepines
- **Theta 4-8 Hz; < 30 μV -** Normal found in awake children, and during sleep. Typically divided in two components, low theta (4 - 5.45Hz) that correlates to decreased arousal and increase in drowsiness and high theta (6 - 7.45 Hz) claimed to be enhanced during task involving memory.
- **Delta (0.5–4 Hz; up to 100–200 μV) -** This rhythm is dominant during sleep stages three and four (slow-wave sleep) they are the lower frequency and the highest in amplitude, brain gets similar to un-conscious and is where the body heals the most during sleep. Delta waves increase the production of two anti-aging hormones, DHEA and melatonin.
- **Gamma (30-150 Hz) -** Is associated with higher mental activity, including perception and consciousness, it disappears under general anaesthesia and during sleep. Is found during memory formation and perception thinking. Gamma waves are enhanced in Buddhist monks during meditation and are absent in schizophrenics.

The EEG is often used in clinics to detect a variety of neurological diseases such as epilepsy or schizophrenia but can also be used to predict movements, detect sleep stages, emotion, mental load, or seizures. Reading EEG signals is often a difficult task given the tiny voltage produced by the neurons and the impairment imposed by the skull. This often results in waveforms that mimic pathological discharges or false artifacts originated by insecure connections in the electrodes which causes misdiagnosis. According to van Donselaar *et al* (2006), a study made on the accuracy of the diagnosis of epilepsy showed that in British population there is misdiagnosis that varies from a rate of 5% in a prospective childhood study to at least 23%.

Sleep takes up a big role in our daily activities and health, it is estimated that in our lifetime we spend an average of 227 760 hours sleeping, which corresponds to 26 years, almost a quarter of our lifetime sleeping. Sleep is very important, as it provides the body resting and ensures its homeostasis, it promotes mental health and improves physical wellbeing, such as immunity, weight, and fertility (Barkoukis, T.J. and Avidan, Alon. 2007). To ensure a good night sleep the body requires two conditions: a high amount of melatonin, which is synthetized in the pineal gland and regulates the sleep cycles and prolonged REM sleep. Melatonin is a hormone that transmits information regarding the day length, with high levels during the night and low levels during the day (Pandi-Perumal *et al* 2007).

REM (rapid-eye-movement) is one of the phases that occurs during sleep. Sleep consists of a sequence of cycles that alternates between the modes REM and non-REM. Although the function of this modes is not yet well understood, it is known that during REM the body enters in paralysis state and the brain starts producing a form of narrative feature known as dreaming that later can be remembered and recognized as fantasy. During this phase, the body enters in an anabolic state, to help restore the immune, nervous, skeletal, and muscular systems.

The non-REM (NREM) sleep is a period of an average of 90 minutes cycle, that is described by deep sleep or low wave forms. The body, during this phase, decreases the heart rate and the body temperature drops, the brain also uses less energy and become almost unconscious. The America Academy of Medicine (AASM) divides this period into stages N1,

Figure 1 Sleep stages include rapid eye movement (REM) and non-REM (N1-N3)
(Source: Background for EEG Lab, Middlesex University)

N2 and N3 (Figure 1) where this last is the deepest sleep form, where it can be easily found delta activity. The normal cycle occurs in the order N1 to N2 to N3 to N2 to REM. In the beginning of the night there is a big amount of deep sleep (N3) and as the night progresses the amount of deep sleep (N3) decreases in the opposite proportion to the REM sleep, that increases (Parmeggiani *et al* 2005).

Furthermore, other studies found that serotonin also plays a significant role in the control of sleep and behavioural state. However, this is less understood and while some pharmacological studies indicate that serotonin reduces REM sleep via agonist, other indicate that electrical stimulation of the Laterodorsal tegmental nucleus (LDT) (serotonin averse to REM), increases REM sleep. So, serotonin shows to act sometimes as opposite in physiological responses. In the other hand, for the melatonin, was found concrete evidence that its suppression, using the B-adrenoceptor antagonist results in insomnia (Ursin 2002).

1.3. EEG ANALYSIS APPLIED TO SLEEP

The ability to accurately describe changes during the sleep state has been one of the main goals of sleep medicine. This is traditionally done by categorizing sleep EEG into distinct stages through a visual inspection of 30-s epochs. However, often some crucial aspects of the signal might not be fully recognized during visual inspection of the EEG and therefore additional techniques has come into place to clarify the EEG.

One of the first attempts to improve this visual inspection has come in the early 1980s with EEG spectral analysis such as the colour density spectral array (CDSA), density spectral array (DSA) or the hypnospectrogram. Spectral analysis is a powerful method that encompasses analysis of the time-varying oscillations using spectrograms, that plot the signal over time and frequency, allowing for an immediate understanding of the different frequencies at the course of a night sleep. Spectrograms often uses a colour gradient in the y-axis function that indicates the rage of intensity of the EEG frequencies. However these previous methods of spectral analyses had technical limitations such as undefined peaks and noise that obscure important features

Fortunately, there is a method called multitaper spectral estimation, which greatly improves on standard spectral analysis methods giving a clearer and accurate, high-resolution spectral estimation. The main goal of this method is to separate a waveform into different component oscillations based on frequency. The theoretical basis for spectral estimation is the

Fourier analysis, which is a method that consists of decomposing a time domain signal into a series of pure sine waves of different wavelength (Prerau *et al* 2017).

In this method, low power colour is often attributed to the non-REM stages (N1, N2 and N3) where the activity is dominated by high-amplitude and low-frequency slow waves (<4.5Hz), and Spindles (10–16Hz) that are transient events generated in the interplay between neocortical and thalamic circuitries. In the REM sleep, EEG activity is predominantly seen as medium intensity, as this stage is comprised by theta waves (6–9Hz) and gamma oscillations (30–150 Hz) generated by the hippocampus and cortical networks, respectively. In wake state, multiple cortical and subcortical networks show high frequency and low amplitude activity, including theta and gamma oscillations, and varies forms of noise are also seen in the signal due to the persons movement (Adamantidis *et al* 2019).

In Figure 2, we can observe a full night multitaper-spectrogram where it is marked the sleep segmentation periods Wake, REM, and NREM.

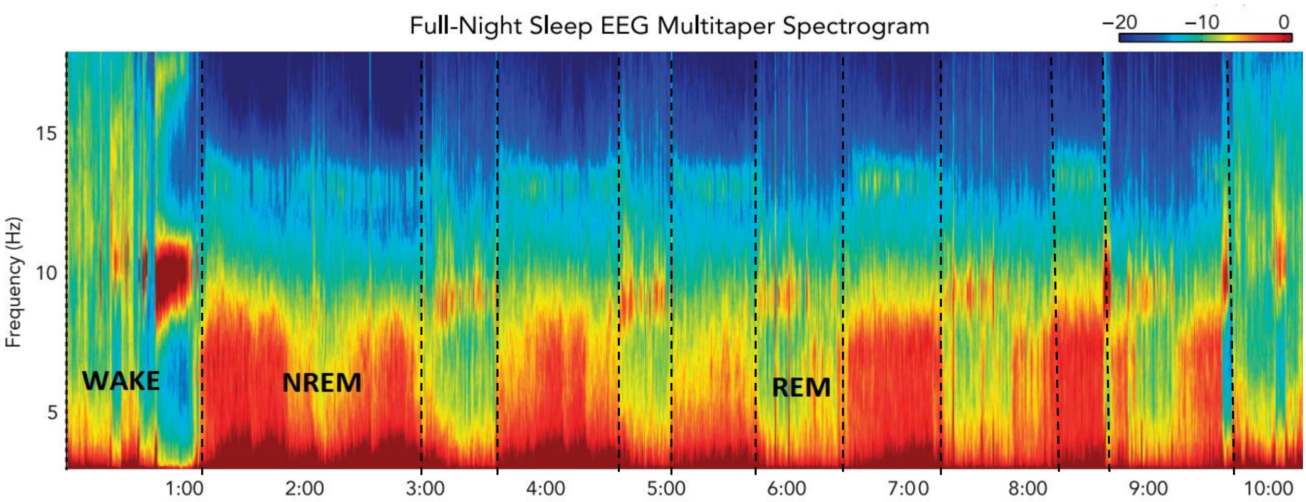


Figure 2: **Full-night of the EEG multitaper spectogram computer-aided method of signal analysis.** This spectrogram has a spectral power based on colour: blue to red (low to high power) as function of time (x-axis) and frequency (y-axis). Furthermore, the multitaper spectrogram provides a good visual representation of the sleeping features sequence of the alternating REM and non-REM modes, allowing the machine learning tools to have distinct features as input for the training data.

2. CNN FOR THE SLEEP STAGING CLASSIFICATION

To reduce misdiagnosis from occurring, medical engineers have researched new ways to pre-processing EEG data before its use in clinics. According to Siuly (2012), there are two main divisions in EEG classification: supervised classification, in which a supervisor instructs the classifier during the construction of the classification model, and unsupervised classification, where the classification model contains some measure of inherent ability. In this report we will study the usage of an unsupervised classification approach.

The selection of the unsupervised classification algorithm for the sleep stage scoring was based on Craik *et al* (2019), that selected and reviewed 90 qualified studies on deep learning application for EEG tasks. In that report there was analysed a series of neural networks architecture and used statistics to provide recommendations for the best architecture related EEG tasks such as motor imagery, mental workload, seizure detection, sleep stage scoring, potential task, and emotion recognition. Unfortunately, in this report, with respect to sleep staging scoring task, due to lack of studies, it was not possible for them to provide an accurate recommendation on a specific deep learning architecture. However, they showed that for this task the most prevalent input formulation tactic was the use of signal values as input, such as spectrogram, that has been proven useful in most of the tasks.

The classification algorithm proposed in this report consists of one popular deep learning architecture commonly used for classification of images, the Convolutional Neural Networks (CNN). This powerful network, first introduced by Yann LeCun in 1988 in a work called LeNet, exhibits an extraordinary performance in a variety of machine learning task achieving frequently an average accuracy of 84% for images and 87% for signal values.

A convolutional neural network is a specialized neural network on computer vision tasks such as recognizing objects, scenes, faces, among other applications. A CNN model usually consists of three layers: an input layer, a hidden and an output layer where the hidden layer often has multiple layers. The basis of CNN is to transform images or signal values into a form of arrays that are feed into the input layer of the neural network. Then the following nodes or hidden layers of the network, perform different manipulations and calculations to create a system that classifies according to the output. Generally, the typical architecture of the hidden layers starts with convolutional layers that, working with a kernel, generate multiple matrices of the image and its features. From there, pooling layers perform the extraction of the features and reduce its size. This is followed with an activation function, such as the ReLU layer, that normalizes the data activating the feature. In CNN, these stacks of layers can be repeated several times depending on the difficulty of features. Finally, the output of this stacked layers goes through a fully connected layer (FC), were the features maps are converted to vectors of arrays try to recognise this features (Koushik 2016). Figure 3 represents an overview of the components involved in the classification of the sleep stages.

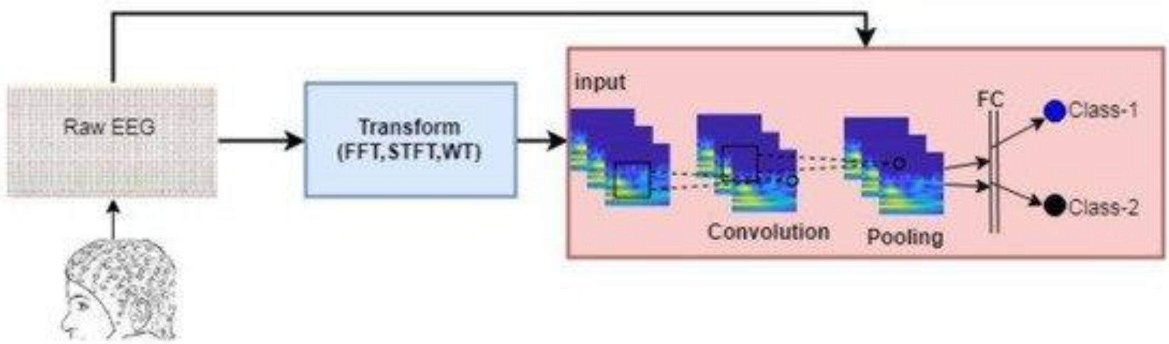


Figure 3: Components involved in the detection of the sleep stages through spectrograms and CNN

2.1. SOFTWARE AND INPUT DATA

To develop this convolutional neural network (CNN), it was used the programming language *MATLAB.* According to Haigh (2008) *MATLAB* is a mathematical software developed by *MathWorks* that focus on matrices, which were initially its only data type. However, today is a multiparadigm software written in C capable to perform matrix manipulations, plotting functions, implementation of algorithms, along with many other functions. *MATLAB* also has the special feature of the addition of many applications available in "toolboxes" that have the capacity to add new packages of functions for specific application areas. In this report to achieve better visualization of the layers and assist in the development of the network design, the application *Deep Network Designer* was installed and used.

For the first section of the code, there was defined all the variables and functions necessary to load the input images to be used in the CNN training. For this, a variable by the name of *DatasetPath* was created to store the directory of the folders that contained the sleep stages images. Then, to manage them, function *imageDatastore* was used to collect and store all of the image's directory, folder, labels, size, and format output. For the network, the loaded images used for training were 12 images sized 300 by 150, with four images for each class: NREM, REM and WAKE, from multitaper spectral analysis. Finally, to give the user the validation of the images used in the input, a randomizer function was used to plot four of the training images.

Figure 4 displays the output of the training images from the first section of the code and Figure 5 the code itself.

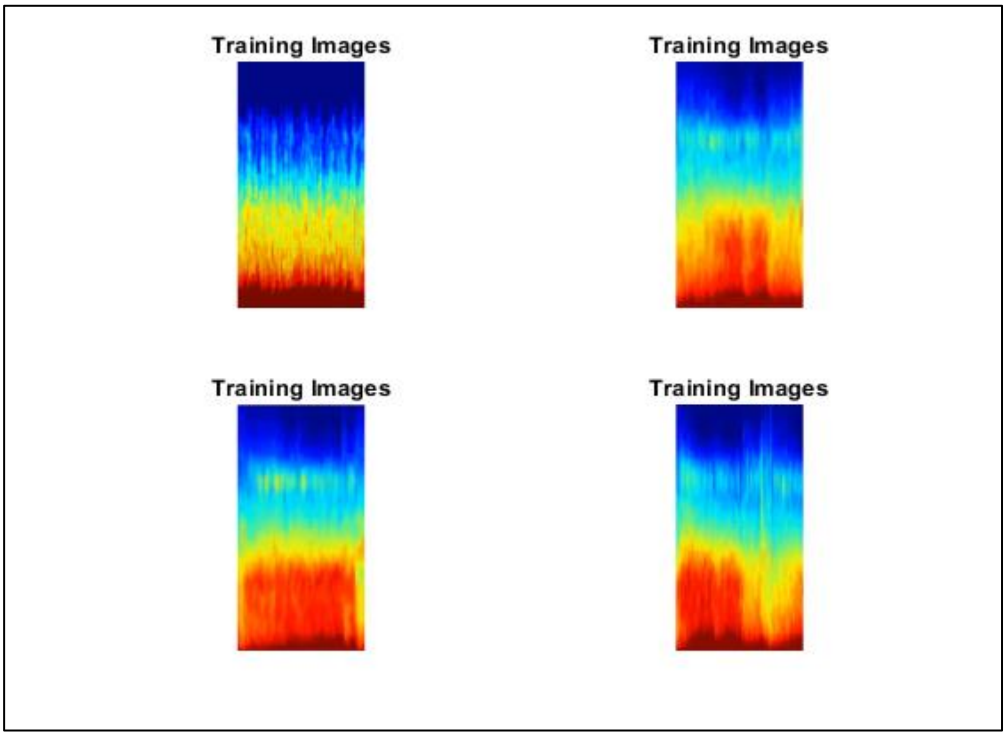


Figure 4: Output of the code first section, where it is displayed a random sample of the sleep stage images load as input for the training.

```
clc
close all

%%dataset management
matlabroot = 'C:\Users\Guilherme\Desktop\CNN\photos\'; % directory

DatasetPath = fullfile(matlabroot, 'Dataset'); % builds a full file specification from the specified folder and its file names

Data = imageDatastore(DatasetPath,'IncludeSubfolders', true, 'LabelSource','foldernames'); % manage a collection of the images

valFiles = 4;
[Train] = splitEachLabel(Data,valFiles,'randomize'); % randomly assigns the specified proportion of files from each label to the new datastores.

%Diplaying Training Images
for i = 1:4
    a = readimage(Train,i);
    figure (1),
    subplot (2,2,i)
    imshow(a)
    title('Training Images')
end
```

Figure 5: Code for the first section of the CNN using a variable to store the directory and a function to manage the input images.

2.2. CONVOLUTIONAL NEURAL NETWORK DESIGN

In the first section of the code all the variables and functions used for storing and managing the image’s directory and its respective labels were used for the loading of the neural network. The following code section contains the layers responsible for the calculations and manipulations used in the recognition of the features and the prediction of the sleep stages.

To design a CNN the most prevalent approach is made using pretrained models for transfer learning. These pretrained models such as GoogLeNet, AlexNet and Inception are proven pretrained networks, that have already been trained with multiple images and learned a large variety of features, which typically makes it easier and faster than training data from scratch. Unfortunately, for this work propose this technique is not convenient since most transfer learning architectures are trained on image recognition tasks and therefore there have a low level of features that can be beneficial for the classification of the EEG spectrograms (Maithra Raghu and Chiyuan Zhang 2019).

Given this unique problem the alternative for designing the CNN was to develop it from scratch from our own data. To accomplish this a variety of research was performed to learn the best configuration for the number of the layers, the hidden units and other hyperparameters. Although, designing a deep neural model has no correct configuration, designing it requires the same level of trial and error and human expertise to achieve a successful network design and avoid errors. Fortunately, to accomplish this there are available online successful models and forum discussions from which it is possible to borrow ideas.

To achieve this best configuration and minimize errors, the development of the CNN begun from a small model with one convolutional layer, a rectified Linear Uni (ReLU), a Polling layer, a fully connected layer, and an output layer SoftMax. Then the model size was gradually expanded by increasing the number of units and the number of layers. That was done stacking on top of each other convolutional layers: ReLU and pooling layers, and FC with dropout layers. This increased of the model resulted in a 17 layered CNN.

As the data goes through each layer the early convolutional layers learn to extract primitive features such as oriented edges. Then, as it moves towards the output layer, the features become more complex responding to more specific concepts. The number of activated neurons and the filter constrains weights and bias used to transmit information between neurons of the CNN can be seen in table 1 of the convolutional layer (CN) and fully connected layers (FC).

Table 1: Number of activation units, weights and bias parameters used in the CN and FC input training of the EEG spectrograms images of the sleep stages

Layer	In	Weights	Bias	Out
CN1	296x146x20	5×5×3×20	1×1x20	148×73×20
CN2	144x69x20	5×5×20×20	1×1x20	72×34×20
CN3	68x30x20	5×5×20×20	1×1x20	34×15×20
FC1	10200	100x10200	100x1	100
FC2	100	3x100	3x1	3
Out	3	-------	-------	1

2.3. CONVOLUTIONAL LAYERS

The convolution layers are the most characteristic layers in the CNN. These layers have the unique characteristics of transforming images using a convolutional operation. Convolution is a mathematical operation that combines two signals using a matrix multiplication.

When an image is set as input the convolutional layer role is to transform each pixel from the images and convert it to a defined number of feature maps. The technique for detecting these features in local pixels is done using a small pre-defined matrix called dimensional filter or kernel, that convolutes around the image matrix recording the results in feature maps to be then used to train the network and recognise the features in the output. The process behind the convolutional operation starts with the filter moving to the right a stride at time and multiplying each element in the filter and summing them. Then the results are stored as a correspondent pixel as seen in Figure 6.

More recently computer scientists have also developed specialized filters such as the Sobel filter and the Scharr filter that are specifically built to perform better capturing of specific feature such as horizontal, vertical, and diagonal lines. Nonetheless, the best way of choosing the filter is done when the CNN learns its own filter using back propagation (Kimura *et al* 2019).

In the CNN, three convolutional layers were used where in each of them 20 kernels of size 5 x 5 were used as parameters with automatic padding value yielding 20 features maps at each layer.

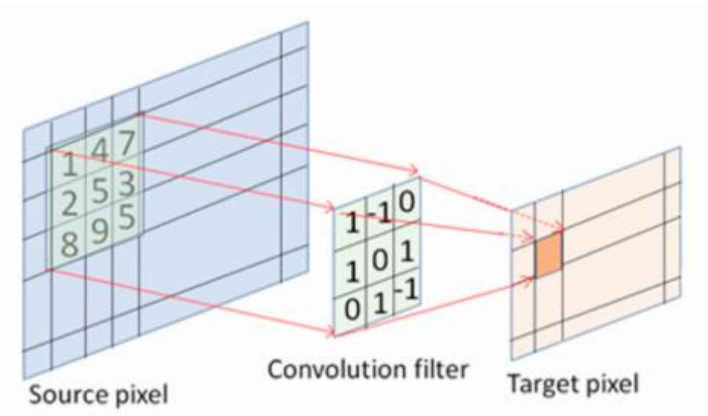


Figure 6: Convolution layer with a filter in Convolutional neural network (CNN)
(Source: https://doi.org/10.3390/w12010096)

2.4. RELU FUNCTION

Conventionally, ReLU is an activation function for hidden layers in deep neural network. The goal of this function is to return all the activation units for each unique feature. The Rectified Linear Units (ReLU) has strong biological and mathematical underpinning. The most common ReLU was demonstrated in 2011, to further improve training of deep neural networks. It works by thresholding the output of the function to 0 if the input is negative and conversely it output the same as input if input is positive. More precisely if input number $x < 0$ it outputs 0, and when $x \geq 0$ the output is a linear function. Figure 7 illustrates the graph of the ReLU function in the network.

The relatively unfavourable findings on ReLU models are due to the dying neurons problem. That is, when a neural network has a large negative bias, the input of the ReLU function might always be 0 and therefore no gradients flow backward through the neurons, and the neurons become stuck, and cannot fire eventually "dying" which impedes the learning progress of a neural network. This problem is addressed in subsequent improvements on ReLU for example Leaky ReLUs, Parametric ReLU or ELU (Agarap 2018).

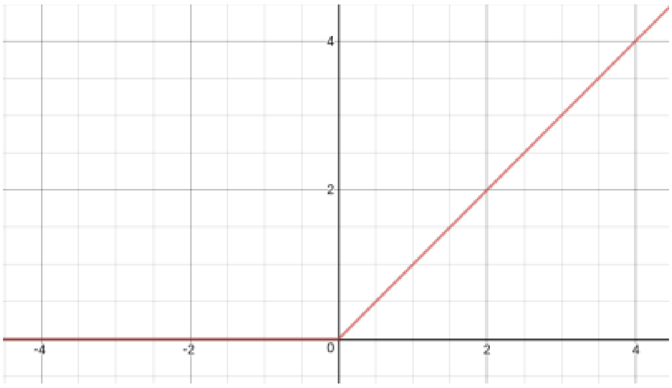


Figure 7: The Rectified Liner Unit (ReLU) activation function produces 0 as an output when x<0, t and linear of slope 1 when x > 0
(Source: arXiv:1803.08375v2)

2.5. POLLING LAYER

Polling layers are a crucial part of CNN architecture to achieve an efficient training process by helping to decrease the size of the feature map and the extracted features. Commonly, the most used CNN pooling techniques are the max pooling and average pooling. In these techniques process, max pooling takes the maximum score in the pooling window, and average pooling selects the average score of the pooling window. However, for our network design it picks the max polling type.

The max polling layer works by down sampling its input, summarizing the input window by picking its maximum value within the kernel. This high pass filter effect contributes for a higher receptive field in the network reducing the output resolution and the number of activations that are feed to the next layer, additionally helping the computational loading performance and reducing overfitting.

On the other hand, the max pooling also has its shortcomings because it ignores all crucial information except the most significant value. That can lead to degrade an opportunity to calculate other informative features in a large and dynamic input data (Jie and Wanda 2020). In our network, the data passes to a max-polling layer of stride 2 with zero-padding (Figure 8).

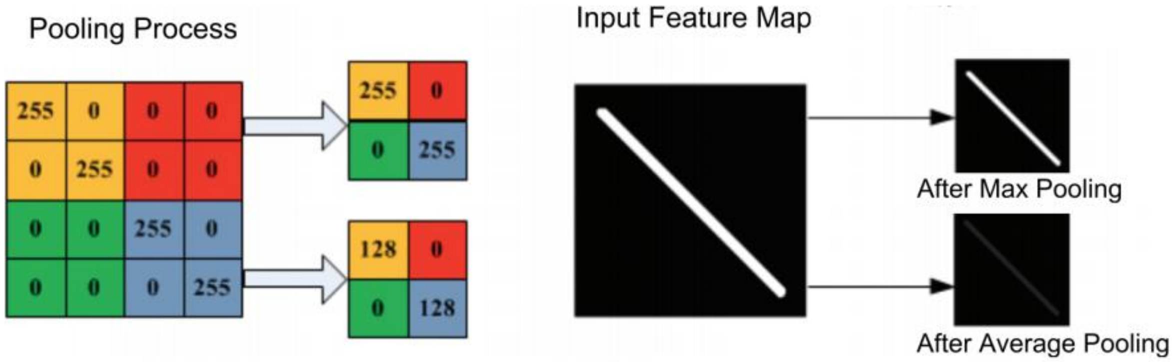


Figure 8: Pooling process of the input feature that illustrates the drawbacks of the max polling layer and average polling (Source: https://doi.org/10.2991/ijcis.d.200120.002)

2.6. DROPOUT LAYER

Overfitting is a serious problem difficult to deal with in deep neural networks. This problem is caused by the combination of many different predictions in the neural net at test time. So, to prevent overfitting from happen, computer scientist developed the Dropout technique for addressing this problem. The Dropout layer is a regularization technique that randomly drops units (along with their connections) from the neural network during training, preventing this way units from co-adapting too much. This practise limits the model capacity to learn features relying on other neurons equalizing the weight of the matrices and encouraging neurons to learn new useful features useful for the classification (Srivastava *et al* 2014).

In our network, it was used this layer two times: before the third convolutional layer and at the end, before the first fully connected layer.

2.7. FULLY CONNECTED LAYERS

Deep convolutional neural networks traditionally consist of a series of convolutional and pooling layers followed by one or more fully connected (FC) layers that combines features into more attributes to perform the final prediction of the categories. However, for a FC to function and learn the individual features from the convolutional layers they require to be converted into a 1D feature vector. This operation is called flattening as it simplifies the output to be used by a dense layer for the final classification (Qian *et al* 2020).

In the present network the data from convolutional layers passes through three fully connected layers. The first FC has 100 nodes, the next has three nodes and for the final classification output layer, another three that are class labels with all the collected information.

2.8. VALIDATION AND RESULTS

To ensure the proposed CNN is properly classifying, some tests and analyses were made in order to verify the precision and accuracy of the neural network, and its proneness to overfitting. The first test performed on the CNN was conducted by a cross-validation technique that quantified the accuracy of images within the trained data and images outside the input. The second test is speculative and cannot be directly measured and can only be based on opinion. The third test can be defined as the ability for the CNN to correctly train the images.

In the first test, to determine the precision provided by the CNN, the accuracy of the model was measured using a common validation technique. The cross-validation is a common practise in machine learning for assessing the performance of the models. This technique works by splitting the input data into two data sets: a training set, that is used to train the model, and the validation set that is used to evaluate the model performance. To perform this validation model was required an application named *experiment manager* available in the *MATLAB* toolboxes. This app was designed as a deep learning neural network laboratory environment that allows to set up experiments for the training data, inspect the parameters and record its observations. Despite this tool, the report acknowledges that this validation method does not guarantee a final result, as a rigorous cross-validation cannot be done with a small input data. This lack of input can cause an unfair distribution of the images in the training data, which potentially can cause bias in the classification. So, to try to compensate these potential bias, four cross-validating tests were made where the division of the input was randomly assigned. The results to this testing was a training accuracy of 69.37% with a standard deviation (std) of 3.16 and for the validation accuracy 43.08% with a std of 3.15.

The second test of the CNN features was made individually also with images outside the input data to verify the recognition of the features. To perform this test six images with similar but challenging characteristics were previously selected to be used by the classification model. Then, using a function named *predict* the output label was displayed with its percentage. However, the results from this function were inconclusive and do not seem to correspond with the expectations. Nonetheless it was a good test to observe the CNN detection capacity. The first two images collected to test the network were from traditional spectral band with the sleeping stages NREM and REM that had similar colours. Then the next two images were from a Single-taper spectrogram that had features with similar shape in the sleep stages Wake and REM. Finally, the last set of two images were from another EEG spectrogram with similar features and colours but difficult to tell apart the sleep stages NREM and REM (Figure 9).

The results from these tests, despite inconclusive, were interesting to observe as the CNN successfully classifies half of the images given, regardless of the group which were taken. According to the *predict* function the images that struggled to be classified were from the spectral band and from the similar EEG spectrogram. Besides this, the CNN accurately predicted the one image from the group single-taper with accuracy of 1. Nonetheless, there was also a false accuracy of 1 in the images of the traditional and Single-taper spectrograms. This study contains some useful information regarding the bias of the CNN and the capacity of recognising the features collected. However, these observations are not reliable since the CNN learns it differently every time it runs. These variations in the outputs are best seen in the challenge classification of the images in the similar EEG spectrogram, where at times the function had an accuracy of 1 and in another run the accuracy was of 0.48.

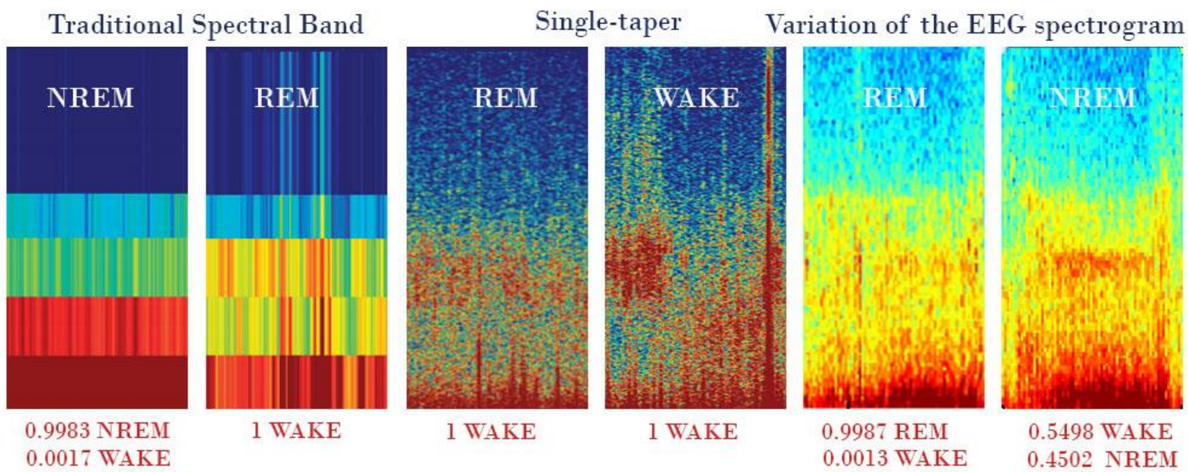


Figure 9: Displays the images used for the second test and the respective accuracy of the classification attributed by the CNN model

In the third test, an evaluation was performed regarding the training progress of the CNN where the difference between the training data and the loss function was analysed. For this the method used to visualise the performance was the functions *trainingOptions* and *TrainNetwork*. These functions creates a set of options for training a network using stochastic gradient descent and the *TrainNetwork* is used to train the CNN. As the training progress starts a window opens with the plot of the training accuracy and its loss. From there, it is possible to see whether or not the model is overfitting. This is seen if the training data accuracy increases, and the loss function decreases (Figure 10).

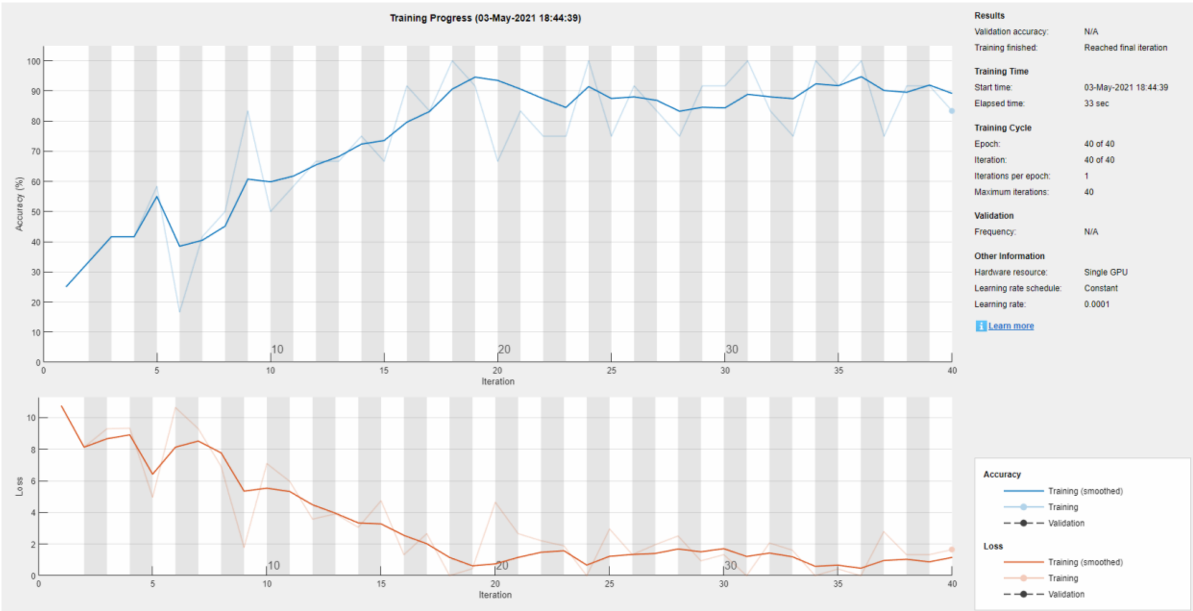


Figure 10: CNN monitor of the training progress, where it is displayed the accuracy and loss plots

3. DISCUSSION

Convolutional Neural Networks demonstrates an effective method of classification for EEG sleep scoring task. In this report, besides the model developed being of small size, it still proved sufficient for the learning of sleep stages features in the EEG multitaper spectrogram. According to the research question the development of this report was made with the end goal of validating new EEG classification methods such the deep learning tools that has been winning popularity in the market and in the academic environment.

Focusing on the development of this classification algorithm, CNN is a deep learning architecture characterized by the extraction of features through images, and by its ability to predict according to those images (Lecun *et al* 1998). In addition to this neural network, an

EEG visualisation technique named multitaper spectral analysis was used to achieve better results due to this method that makes EEG data clearer and more accurate than any other traditional approaches multitaper spectrogram facilitating the recognition of features and the classification accuracy.

While developing this CNN there was a better understanding of the role of each layer: that the first input layer collected only images of size 150 by 300. Next, that the convolutional layers were used to create a matrix of the images that holds the features. Also, that the size of the kernel filter involved a high number of tentative and research where there were tried filters of smaller and larger size (3 x 3 and 7 x 7). Here however, the filter symmetry was important to not cause distortions across the layers, as the smaller filters increased the learning rate, but the bigger ones caused distortions. So the choice of 20 filters 5 x 5 gives us the best test-set performance. Next, it was better understood the function of the polling layers that extracted the sleeping features using a stride, in this case, two, to locally summarize the images matrices by the maximum number surrounding. This was followed by a ReLU activation function that selected these features and passed them to fully connected layers were a flattening of these maps were transformed into vectors that made the final prediction of the sleep stages. However, to prevent overfitting, a common problem in neural networks, dropout layers were added. Despite the optimised performance for the hyperparameter and layers, in the case of expansion of the CNN model or the input size, it is recommended a revaluation of the hyperparameters and test them multiple times to achieve better accuracy.

Aside from the construction of the CNN, one large difficulty was found in the utilization of the right images for the input. The use of multitaper-spectrograms rather than raw data from EEG reduced the difficulty of accurately predict the sleep stages, however it also limited the range of its utilization since this method is not common and there is a large variation of other spectrogram methods. The use of other EEG visualization methods in this model is obviously no viable and although having similar characters, they are probable to be misclassified. This was seen in the second test in the validation of the network, were the traditional spectral band, single-taper, and EEG spectrogram variation, as expected, did not provide an accurate result on the images. However, they were useful to challenge the recognition of the features since the difference in colours or shape were not significant.

The CNN model currently uses ten of the 12 available images from the multitaper spectrogram. The other two images came from Stokes and Prerau (2020). The remaining images used in the validation are also from the multitaper spectrogram, except two of the EEG spectrogram variation (UZH -Institute of Pharmacology and Toxicology - Sleep EEG and Sleep Regulation, 2021). There is still much space for future research concerning creating a universal

EEG classifier or improving the CNN model accuracy by using more input for the training of the images.

The application of this CNN to real medical applications is difficult to answer due to our limited CNN model. To recognise the features with high precision represents a far great problem that can potentially have two outcomes. There are two hypotheses of the effect of the CNN: (1) The increasing of the network model and its input achieves a high validation accuracy, (2) The CNN will always have the potential to misclassification. The results of this study support the later by showing the unpredictability of the CNN classification seen in the second test validation, where the model classifies the last image as NREM, sometimes with accuracy of 1 and other times around 0.5.

To measure the network accuracy a cross-validation study was done to assess the performance on the classification of images. This assessment showed that the CNN accuracy of the training data were of 69.37% and for the validation data 43.08% on average, which ultimately could never be sufficient to be used in a real application as it is likely to lead the user to wrong conclusions and beliefs that can potentially be harmful. Nonetheless, this model proves the concept, however the action required that could be taken to increase the network accuracy is unknown, but possible to be identified in further investigation.

Additional steps in the design of the network such as utilization of more layers, skip connection, and larger set of quality input data may increase the network performance. For this we recommend the ResNet architectures that have been shown to achieve high accuracy results. This method uses skip connections that give the layers a reference point so that adding more layers would not worse the performance. This also creates an additional path for the gradient to flow back along making easier to optimize the earlier layers so further investigation into the best methodologies for network design is needed.

In support to the potential improvements, other designs of automatic sleep staging scoring using deep learning have been developed to try to achieve a higher classification accuracy. One approach to this is the software *Sleepnet*, that uses an EEG neural network based on a hybrid formulation of CNN and a Recurrent neural network (RNN) to deploy annotations for sleep staging. In that software, the CNN used 2D spectrograms size of 29 by 257 for the input layer and for the hidden layers they used a 3x3 kernel in the convolution layer, followed by a max-poling layer and a ReLU for the identification of features. However, after this, they passed the features collected to a RNN model, that was tailored towards modelling sequential data and based on a Long short-term memory (LSTM). This model then learned the temporal dependencies of the features extracted by the CNN and finally passed to a SoftMax function to generate the final sleep stage prediction (Biswal et al., 2017b). So, the usage of these two deep learning properties for the classification of sleep staging in EEG spectrogram within the same network may ensure higher fidelity in the prediction output which can be potential solution to a successful EEG classification model that avoids misclassification. So furthermore research into techniques and designs can be made to improve this model.

CONCLUSION

During the development of project, it was presented the background theory in respect to sleep staging, as well as the EEG machine and some visual inspection methods. In addition, there was also presented a detailed description of the network and its layers for the construction of the CNN sleep staging classification.

Using sleep images of a multitaper spectrogram, we have developed an automated sleep stage classification algorithm, which trained a deep neural network to automatically output a label of 3 EEG sleep stages. We evaluated the recognition using different combinations of features and performed a cross-validation that achieved a validation of near 50%, expected by the model size and number of input images. The results achieved and observations suggested that with further development CNN has great potential to be incorporated into clinical environments or development of a self-monitorization device. Additionally, this report can also contribute as an inspiring reference and verification for further studies on areas such as the nature of cognition, the sleep medicine field, and the future applications of deep neural structures in brain machine interface.

REFERENCES

Adamantidis A R, Gutierrez Herrera C and Gent T C 2019 Oscillating circuitries in the sleeping brain *Nat. Rev. Neurosci.* **20** 746–62

Agarap A F 2018 Deep Learning using Rectified Linear Units (ReLU) Online: http://arxiv.org/abs/1803.08375

Biswal S, Kulas J, Sun H, Goparaju B, Westover M B, Bianchi M T and Sun J 2017 SLEEPNET: Automated Sleep Staging System via Deep Learning Online: http://arxiv.org/abs/1707.08262

Craik A, He Y and Contreras-Vidal J L 2019 Deep learning for electroencephalogram (EEG) classification tasks: A review *J. Neural Eng.* **16** 031001 Online: https://doi.org/10.1088/1741-2552/ab0ab5

van Donselaar C A, Stroink H and Arts W-F 2006 How Confident Are We of the Diagnosis of Epilepsy? *Epilepsia* **47** 9–13 Online: http://doi.wiley.com/10.1111/j.1528-1167.2006.00653.x

Haigh T 2008 Cleve Moler: Mathematical Software Pioneer and Creator of Matlab *IEEE Ann. Hist. Comput.* **30** 87–91

Hammond C 2015 Neurons *Cellular and Molecular Neurophysiology: Fourth Edition* (Elsevier Inc.) pp 3–23

Hofman M A 1985 Size and shape of the cerebral cortex in mammals: I. The cortical surface *Brain. Behav. Evol.* **27** 28–40 Online: https://pubmed.ncbi.nlm.nih.gov/3836731/

Jie H J and Wanda P 2020 Runpool: A dynamic pooling layer for convolution neural network *Int. J. Comput. Intell. Syst.* **13** 66–76 Online: https://www.atlantis-press.com/journals/ijcis/125932844

Kansagra S 2020 Sleep disorders in adolescents *Pediatrics* **145** S204–9 Online: https://doi.org/10.1542/peds.2019-2056I

Kimura N, Yoshinaga I, Sekijima K, Azechi I and Baba D 2019 Convolutional Neural Network Coupled with a Transfer-Learning Approach for Time-Series Flood Predictions *Water* **12** 96 Online: https://www.mdpi.com/2073-4441/12/1/96

Koushik J 2016 Understanding Convolutional Neural Networks Online: http://arxiv.org/abs/1605.09081

Lecun Y, Bottou E, Bengio Y and Haffner P 1998 *Gradient-Based Learning Applied to Document Recognition*

Maithra Raghu and Chiyuan Zhang 2019 Google AI Blog: Understanding Transfer Learning for Medical Imaging Online: https://ai.googleblog.com/2019/12/understanding-transfer-learning-for.html

Pandi-Perumal S R, Srinivasan V, Spence D W and Cardinali D P 2007 Role of the melatonin system in the control of sleep: Therapeutic implications *CNS Drugs* **21** 995–1018 Online: https://pubmed.ncbi.nlm.nih.gov/18020480/

Parmeggiani P L, Velluti R A and Peraita-Adrados R 2005 ELECTROENCEPHALOGRAPHY, POLYSOMNOGRAPHY, AND OTHER SLEEP RECORDING SYSTEMS *The Physiologic Nature of Sleep* (PUBLISHED BY IMPERIAL COLLEGE PRESS AND DISTRIBUTED BY WORLD SCIENTIFIC PUBLISHING CO.) pp 103–22

Pharma.uzh.ch. 2021. *UZH -Institute of Pharmacology and Toxicology - Sleep EEG and Sleep Regulation*. online: https://www.pharma.uzh.ch/en/research/Alumni/sleepbiology0/sleepeeg.html

Prerau M J, Brown R E, Bianchi M T, Ellenbogen J M and Purdon P L 2017 Sleep neurophysiological dynamics through the lens of multitaper spectral analysis *Physiology* **32** 60–92 Online: https://pubmed.ncbi.nlm.nih.gov/27927806/

Qian Z, Hayes T L, Kafle K and Kanan C 2020 Do We Need Fully Connected Output Layers in Convolutional Networks? *arXiv* Online: http://arxiv.org/abs/2004.13587

Siuly 2012 *ANALYSIS AND CLASSIFICATION OF EEG SIGNALS*

Srivastava N, Hinton G, Krizhevsky A and Salakhutdinov R 2014 *Dropout: A Simple Way to Prevent Neural Networks from Overfitting* vol 15

Stokes P A and Prerau M J 2020 Estimation of Time-Varying Spectral Peaks and Decomposition of EEG Spectrograms *IEEE Access* **8** 218257–78

Stores G 2007 Clinical diagnosis and misdiagnosis of sleep disorders *J. Neurol. Neurosurg. Psychiatry* **78** 1293–7 Online: https://pubmed.ncbi.nlm.nih.gov/18024690/

Seeley, R., Stephens, T. and Tate, P., 2003. Anatomy and Physiology. 6th ed. McGraw-Hill Science Engineering, pp.451-580

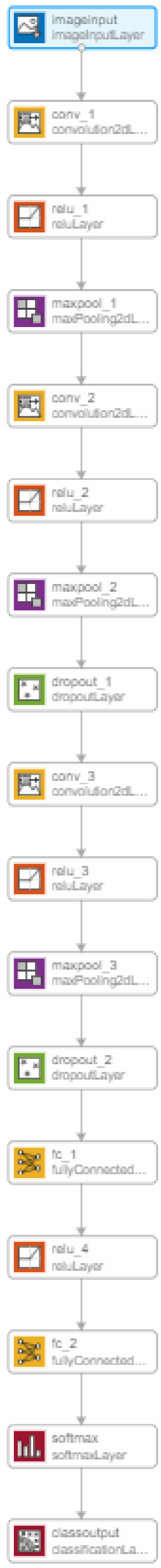
van de Straat V and Bracke P 2015 How well does Europe sleep? A cross-national study of sleep problems in European older adults *Int. J. Public Health* **60** 643–50

Telzer E H, Fuligni A J, Lieberman M D and Galván A 2013 The effects of poor quality sleep on brain function and risk taking in adolescence *Neuroimage* **71** 275–83 Online: https://pubmed.ncbi.nlm.nih.gov/23376698/

Ursin R 2002 Serotonin and sleep *Sleep Med. Rev.* **6** 55–67

APPENDIX

APPENDIX A: SLEEP STAGING CNN LAYERS



APPENDIX B: SLEEP STAGING CNN CODE

```
clc
close all

%%dataset management
matlabroot = 'C:\Users\Guilherme\Desktop\CNN\photos\'; % directory

DatasetPath = fullfile(matlabroot, 'Dataset'); % builds a full file specification from the specified folder and its images

Data = imageDatastore(DatasetPath,'IncludeSubfolders', true, 'LabelSource','foldernames'); % manage the collection of the images

valFiles = 4;
[Train] = splitEachLabel(Data,valFiles,'randomize'); % randomly assigns the specified proportion of files from each label to the new datastores.

%Diplaying Training Images
for i = 1:4
    a = readimage(Train,i);
    figure (1),
    subplot (2,2,i)
    imshow(a)
    title('Training Images')
end
```

%% Define The Network Layers

```
%CNN layers, architecture.

layers1 = [imageInputLayer([300 150 3])      % returns an image input layer and specifies the InputSize

        convolution2dLayer(5,20)             % applies sliding convolutional filters to the input
        reluLayer                            % Rectified Linear
```

```
Unit (ReLu) sets threshold and minimizes small input to 0
        maxPooling2dLayer(2,'Stride',2)          %  Down-sampling  by
dividing the input into rectangular pooling regions

        convolution2dLayer(5,20)
        reluLayer
        maxPooling2dLayer(2,'Stride',2)

        dropoutLayer                            % Prevents overfitting
on the training data

        convolution2dLayer(5,20)
        reluLayer
        maxPooling2dLayer(2,'Stride',2)

        dropoutLayer

        fullyConnectedLayer(100)           % Multiplies the input by
a weight matrix

        reluLayer

        fullyConnectedLayer(3)             % outputsize

        softmaxLayer                    % multi-class classification
        classificationLayer()];

option = trainingOptions('adam','MaxEpochs',40, ...        % Monitors
Deep Learning Training Progress
    'InitialLearnRate',0.0001,'Plots','training-progress');

% training
convnet = trainNetwork(Data,layers1,option);        %ADAM optimizer
```

```
%%Classify the Images in Test Data and Compute Accuracy
%testing
allclass=[];

valFiles = 2;
[imdsValidation] = splitEachLabel(Data,valFiles,'randomize');

%loading of the validation images
inp1=imread(C:\Users\Guilherme\Desktop\CNN\photos\VALDIATION_IMG\ban
d_REM.jpg');
figure,imshow(inp1); %plot the images


out=classify(convnet,inp1); %classification

labelW = predict(convnet,inp1) %Compute the accuracy



msgbox(char(out)) %plot the sleep stages classification
```