2019

Noise Reduction in EEG Signals using Convolutional Autoencoding Techniques

Conor Hanrahan
Technological University Dublin

Noise Reduction in EEG Signals using Convolutional Autoencoding Techniques

Conor Hanrahan

A dissertation submitted in partial fulfilment of the requirements of
Technological University Dublin for the degree of
M.Sc. in Computer Science (Data Analytics)

September 2019

I certify that this dissertation which I now submit for examination for the award of MSc in Computing (Data Analytics), is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the test of my work.

This dissertation was prepared according to the regulations for postgraduate study of the Technological University Dublin and has not been submitted in whole or part for an award in any other Institute or University.

The work reported on in this dissertation conforms to the principles and requirements of the Institute's guidelines for ethics in research.

Signed: ________________________________

Date: **01 September 2019**

ABSTRACT

The presence of noise in electroencephalography (EEG) signals can significantly reduce the accuracy of the analysis of the signal. This study assesses to what extent stacked autoencoders designed using one-dimensional convolutional neural network layers can reduce noise in EEG signals. The EEG signals, obtained from 81 people, were processed by a two-layer one-dimensional convolutional autoencoder (CAE), whom performed 3 independent button pressing tasks. The signal-to-noise ratios (SNRs) of the signals before and after processing were calculated and the distributions of the SNRs were compared. The performance of the model was compared to noise reduction performance of Principal Component Analysis, with 95% explained variance, by comparing the Harrell-Davis decile differences between the SNR distributions of both methods and the raw signal SNR distribution for each task. It was found that the CAE outperformed PCA for the full dataset across all three tasks, however the CAE did not outperform PCA for the person specific datasets in any of the three tasks. The results indicate that CAEs can perform better than PCA for noise reduction in EEG signals, but performance of the model may be training size dependent.

Key words: *electroencephalography, event-related potential, noise in EEG, signal-to-noise ratio, noise reduction, artifact removal, convolutional autoencoder*

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to Dr. Luca Longo for his assistance and guidance throughout the course of this dissertation.

I would like to thank Alexander Suvorov who, through a series of discussion, aided me with numerous aspects of this undertaking, which allowed me to perform to the best of my ability throughout this process.

I would also like to acknowledge and convey my thanks to the rest of the Technological University Dublin staff that have assisted and supported me throughout my time in the university.

Finally, without the support and encouragement of my family I would not have been able to complete this undertaking. I'm eternally grateful.

TABLE OF CONTENTS

ABSTRACT II

TABLE OF CONTENTS IV

TABLE OF FIGURES VII

TABLE OF TABLES IX

1 INTRODUCTION........ 10

1.1 BACKGROUND 10
1.2 RESEARCH PROJECT 11
1.3 RESEARCH OBJECTIVES 12
1.4 RESEARCH METHODOLOGIES 12
1.5 SCOPE AND LIMITATIONS 14
1.6 DOCUMENT OUTLINE 15

2 REVIEW OF EXISITING LITERATURE 16

2.1 BRAIN ACTIVITY 16
2.1.1 Event Related Potentials 16
2.2 EEG 18
2.3 NOISE IN EEG 22
2.3.1 Definition & Background 22
2.3.2 Causes of noise 22
2.3.3 Signal-to-noise ratio 23
2.4 NOISE REDUCTION TECHNIQUES IN EEG 24
2.4.1 Classical Techniques 24
2.4.2 Machine Learning Techniques 29
2.5 CONVOLUTIONAL AUTOENCODERS 32
2.5.1 Examples of CAE noise reduction 33
2.6 SUMMARY 34
2.6.1 Overview 34

2.6.2 Gaps in Literature ... *35*
2.6.3 Research Question ... *35*

3 EXPERIMENT, DESIGN & METHODOLOGY ... 36

3.1 HYPOTHESIS ... 36
3.2 DATA ... 37
3.2.1 Data Collection ... *37*
3.2.2 Data Preparation ... *38*
3.3 CONVOLUTIONAL AUTOENCODER DESIGN ... 40
3.3.1 Hyperparameter Tuning ... *40*
3.3.2 Final Model Selection ... *42*
3.4 EVALUATION OF RESULTS ... 44
3.4.1 Signal-to-noise ratio ... *44*
3.4.2 Hypothesis Testing ... *45*
3.5 SUMMARY ... 48
3.5.1 Strengths ... *49*
3.5.2 Limitations ... *51*

4 RESULTS, EVALUATION & DISCUSSION ... 54

4.1 DATA EVALUATION ... 54
4.2 CAE MODEL SELECTION ... 55
4.2.1 Hyperparameter Selection ... *55*
4.2.2 Final Model Accuracy, Reliability, Efficiency ... *57*
4.3 EXPERIMENT RESULTS ... 60
4.3.1 Single Stacked Convolutional Autoencoder ... *61*
4.3.2 Multiple Stacked Convolutional Autoencoder ... *68*
4.4 EVALUATION OF RESULTS ... 72
4.4.1 Single Stacked Convolutional Autoencoder ... *73*
4.4.2 Multiple Stacked Convolutional Autoencoder ... *74*
4.5 DISCUSSION ... 75
4.5.1 Strengths ... *75*
4.5.2 Weaknesses/Limitations ... *77*
4.5.3 Potential Causes of Error ... *80*
4.5.4 Potential Improvements ... *82*

5 CONCLUSION 86

5.1 RESEARCH OVERVIEW.. 86

5.2 PROBLEM DEFINITION 86

5.3 DESIGN/EXPERIMENTATION, EVALUATION, RESULTS 88

5.4 CONTRIBUTION & IMPACT.. 90

5.5 FUTURE WORK & RECOMMENDATIONS 91

BIBLIOGRAPHY.. 93

APPENDIX A.. 102

TABLE OF FIGURES

Figure 2.1.1-1 10-20 international electrode placement system (64 + 2 electrodes) 19
Figure 2.4.2-1 Basic structure of a sparse autoencoder 30
Figure 2.4.2-2 Basic structure of a stacked autoencoder 31
Figure 2.5.1-1 Basic structure of stacked convolutional autoencoder 34
Figure 3.2.1-1 Example of EEG signal over 700ms with stimulus presented at 100ms 37
Figure 3.2.2-1 Example of EEG data shape for a single Subject and single Condition before preprocessing 39
Figure 3.2.2-2 Example off EEG data shape for single Subject and single Condition after concatenation 39
Figure 3.4.2-1 Diagram of testing and analysis process 47
Figure 4.2.1-1 Training vs Validation Loss Curve for epoch selection 56
Figure 4.2.2-1 RAW signal vs PCA reconstructed signal – Condition 1 57
Figure 4.2.2-2 RAW signal vs CAE reconstructed signal – Condition 1 57
Figure 4.2.2-3 RAW signal vs PCA reconstructed signal – Condition 2 58
Figure 4.2.2-4 RAW signal vs CAE reconstructed signal – Condition 2 58
Figure 4.2.2-5 RAW signal vs PCA reconstructed signal – Condition 3 59
Figure 4.2.2-6 RAW signal vs CAE reconstructed signal – Condition 3 59
Figure 4.3.1-1 SNR distributions: RAW vs PCA (left), RAW vs CAE (right) - Condition 1 61
Figure 4.3.1-2 Differences in mean SNR values for electrode location on 10-20 system - PCA vs RAW Condition 1 62
Figure 4.3.1-3 Differences in mean SNR values for electrode location on 10-20 system - CAE vs RAW Condition 1 62
Figure 4.3.1-4 SNR distributions: RAW vs PCA (left), RAW vs CAE (right) - Condition 2 63
Figure 4.3.1-5 Differences in mean SNR values for electrode location on 10-20 system - PCA vs RAW Condition 2 64
Figure 4.3.1-6 Differences in mean SNR values for electrode location on 10-20 system - CAE vs RAW Condition 2 64

FIGURE 4.3.1-7 SNR DISTRIBUTION: RAW VS PCA (LEFT), RAW VS CAE (RIGHT) - CONDITION 3 65
FIGURE 4.3.1-8 DIFFERENCES IN MEAN SNR VALUES FOR ELECTRODE LOCATION ON 10-20 SYSTEM - PCA VS RAW CONDITION 3 66
FIGURE 4.3.1-9 DIFFERENCES IN MEAN SNR VALUES FOR ELECTRODE LOCATION ON 10-20 SYSTEM - CAE VS RAW CONDITION 3 66
FIGURE 4.3.2-1 HIGHER MEAN SNR COUNT: PCA VS CAE - CONDITION 1 68
FIGURE 4.3.2-4 HIGHER SNR COUNT: PCA VS CAE - CONDITION 2 69
FIGURE 4.3.2-7 HIGHER SNR COUNT: PCA VS CAE - CONDITION 3 70
FIGURE 4.5.2-1 PCA RECONSTRUCTED EEG SIGNAL FOR EXPLAINED VARIANCE: 99% - 76% 78

TABLE OF TABLES

TABLE 3.3.2-1 FINAL MODEL HYPERPARAMETER TO BE SELECTED 42

TABLE 3.5.2-1 STRENGTHS AND LIMITATIONS FOR EXPERIMENT, DESIGN AND METHODOLOGY 53

TABLE 4.2.1-1 FINAL MODEL HYPERPARAMETERS.. 56

TABLE 4.3.1-1 SIGNAL-TO-NOISE RATIO HARRELL-DAVIS DECILE DIFFERENCES - CONDITION 1.. 61

TABLE 4.3.1-2 SIGNAL-TO-NOISE RATIO HARRELL-DAVIS DECILE DIFFERENCES - CONDITION 2.. 63

TABLE 4.3.1-3 SIGNAL-TO-NOISE RATIO HARRELL-DAVIS DECILE DIFFERENCES - CONDITION 3.. 65

TABLE 4.3.2-1 MEAN HARRELL-DAVIS DECILE VALUES OF SNR DISTRIBUTIONS FOR ALL SUBJECTS - CONDITION 1... 68

TABLE 4.3.2-2 MEAN HARRELL-DAVIS DECILE VALUES OF SNR DISTRIBUTIONS FOR ALL SUBJECTS - CONDITION 2... 69

TABLE 4.3.2-3 MEAN HARRELL-DAVIS DECILE VALUES OF SNR DISTRIBUTIONS FOR ALL SUBJECTS - CONDITION 3... 70

TABLE 4.5.2-1 STRENGTHS VS WEAKNESSES/LIMITATIONS OF DENOISING CAE........... 79

TABLE 4.5.4-1 POTENTIAL CAUSES OF ERROR AND POTENTIAL IMPROVEMENTS OF EXPERIMENT 85

1 INTRODUCTION

Electroencephalography (EEG) is an electrophysiological monitoring technique that detects and records electrical activity in the brain, using electrodes strategically placed on an individual's scalp. It is used to identify and record voltage fluctuations in signals produced by an individual's brain in real-time, having been first conceptualised in 1929 (Berger, 1929). Since then it has been widely utilised and expanded upon in order to gain ever more insights into the workings of the human brain and increase understanding of cognitive processes.

1.1 BACKGROUND

The use of EEG for brain signal analysis has been widely practiced since its creation, highly benefitting the medical and neurological industries.
The technique has been commonly used for the brain activity analysis of individuals in need of constant and accurate brain functionality observations. This includes individuals suffering from epilepsy or patients in medically induced comas. For these individuals, as well as many others, it is extremely important to have accurate EEG signals for analysis in order to ensure the most appropriate diagnosis and prognosis is provided.

There are numerous advantages when using EEG over other methods of brain signal analysis, such as Magnetic Resonance Imaging (MRI) and Computed Tomography (CT), one such advantage being a significantly higher temporal resolution. Brain activity can be recorded to the nearest millisecond using EEG, allowing for very accurate temporal measurements of the signal. Because of this advantage, electroencephalography is still one of the most widely used techniques for brain activity analysis, and therefore it is imperative that the signals obtained for analysis are accurate and understandable. A significant problem that has impeded the accuracy of EEG signal analysis is the issue of *noise*. Noise in EEG signals is caused when undesired artifacts are present in the signal, such as cognitive activity relating to eye blinks or muscle contraction.

The presence of these artifacts, and others, can result in substantial difficulty in distinguishing between the actual signal of the brain activity and random noise present in the brain. This ambiguity can lead to inaccurate and misleading readings that would be deemed inadequate and unusable for analysis. A solution to this problem could be very beneficial to the entire domain of EEG signal analysis, as it could allow for more accurate and concise brain activity analysis, resulting in a more affirmative and effective direction of treatments.

Many approaches have been taken in order to tackle this problem, ranging from classical mathematical transformations, to modern machine learning techniques. There are benefits and drawbacks to all methods used for EEG signal noise reduction. The aim of this paper is to propose a noise reduction technique that can incorporates the advantages of these other methods, while circumventing the various disadvantages associated with each.

A deep learning approach, employing the use of convolutional autoencoders (CAEs), "...an unsupervised learning method for features learning based on CNN (convolutional neural network)" (Wen & Zhang, 2018), is proposed and described in this investigation, as well as the details of the methods utilised in order to fulfil the requirements of an effective denoising model.

1.2 RESEARCH PROJECT

In order to measure the extent of noise present in a signal, the *signal-to-noise ratio* (SNR) can be calculated. This provides a ratio of the power of the total signal to the power of the noise within the signal and is a well-used metric for noise measurement in the field of signal processing. Low signal-to-noise ratios signify less distinguishable signals and a greater presence of noise. In order to tackle the problem of a low signal-to-noise ratio in EEG signals, it is important to remove as much noise as possible from the overall signal with removing as little of the essential information of the original signal as possible. In this investigation, it is explored whether convolutional autoencoding techniques can be used to increase the signal-to-noise ratio of EEG signals. If this is successfully done, a more accurate depiction of the signal can be obtained, resulting in a more precise analysis of the brain's activity.

If this is achievable with the use of CAEs, to what extent can this be achieved, and can a CAE be designed such that it performs better than previously proven noise reduction techniques, such as Principal Component Analysis (PCA)?

1.3 RESEARCH OBJECTIVES

The objective of this research is to investigate whether convolutional autoencoders can be used to increase the signal-to-noise ratio of brain signals obtained by an electroencephalography test with a better performance than previously proven methods, in particular PCA. The performance of a generalised model will be analysed to determine whether a model can be successfully designed to reduce noise in any EEG signal. As well as this, the performance of individual specific models will be analysed. This is to determine whether models would perform better if they were specifically trained and tested for individuals, as brain functionality differs from person to person, or is a generalised model, which can be applied to any individual, more appropriate.

1.4 RESEARCH METHODOLOGIES

In order to address the research objectives, it is important to establish an appropriate and effective methodology to ensure an accurate and reliable outcome is obtained through this investigation.

The research carried out in this investigation is secondary (desk) research, as it is a summary and expansion of existing research. Quantitative methods are used in this exploration, as it is aimed at developing a hypothesis relating to phenomena involving numerical data and allows the quantification of findings and results. As the method used in this research involves gaining knowledge by direct observation, using a hypothesis and predictions that can be tested, it can be categorised as empirical research, using deductive reasoning to conceptualise a theory, formulate a hypothesis, observe the results of a related experiment, and confirm the original theory.

The methodology described above will be undertaken using the following method. The data used in this investigation is real-life EEG signal data from 81 subjects. The subjects were assigned with performing three separate tasks, ~100 times each. The EEG signals were recorded at 64 electrode locations on the scalp for a time of 700ms for each trial. Each 700ms reading from each of the 64 electrodes is used as in input vector, of length 700, for the convolutional autoencoder. The three tasks performed by each subject consisted of pressing a button and hearing a tone with no delay between the button press and the onset of the tone, passively listening to a tone with no button pressing, and pressing a button without a tone being generated, with each of these tasks regarded as Condition 1, Condition 2 and Condition 3 respectively.

The onset of each of these stimuli occurred at 100ms. The EEG readings prior to the onset of stimulus were regarded as the noise of the EEG signal, with any successive reading regarded as the signal combined with the signal noise. Using these values, the SNR can be calculated for each signal recorded at each electrode.

The output of the CAE is a reconstruction of the input vector, with a length of 700. The SNR of the reconstructed signal can be calculated, as done with the input signal, and compared to the original signal. An increase in SNR indicates that the noise of the signal has been reduced.

The distribution of SNRs for all reconstructed signals can be plotted and compared to that of the raw EEG (RAW) signals. The Harrell-Davis decile values are calculated for each distribution and compared, allowing for more precise distribution comparisons. In order to quantify the difference in signal-to-noise ratios, a comparison must be made of the Harrell-Davis decile value differences between the SNR distribution of previously established noise reduction techniques, such as PCA, and the RAW signal SNR distribution. This enables accurate quantification and comparison of techniques, establishing the accuracy of the proposed model.

1.5 SCOPE AND LIMITATIONS

The scope of this limited to the noise reduction of EEG signals, recorded from 81 individuals, 49 of whom were diagnosed with schizophrenia and 32 control participants, whom did not have the condition, with data collected from a single study and an extension of this study. The ratios of participants diagnosed with schizophrenia and control participants for the two investigations were 36:22 and 13:10 respectively. The individuals partook in 3 separate button pressing tasks. These specific tasks were assigned in order to isolate signals from certain cortexes of the brain related to the required tasks. ~100 trials of each task were performed and recorded over a time of 700ms each. 64 electrodes, placed on the individual's scalp, simultaneously recorded the EEG information from each trial.

Examining EEG signals is advantageous for signal analysis as it is simple and cost-effective way to record real-world signals. Recording the electrical activity of the heart is possible using electrocardiography (ECG), which could also be used for signal analysis, although EEG is more versatile, as there are numerous areas of the brain which react differently to various stimuli. This enables complex analysis of the cognitive activity of individuals, whom can be categorised into a number of different classifications, related to the specific research conducted in a particular study. Although some of the data analysed in this particular investigation was obtained from individuals diagnosed with schizophrenia, the context of this research does not limit the analysis to individuals with this diagnosis. This research could be applied to any group of individuals. Ensuring there is a stimulus present at a specific time in each trial, however, allows for specific signal-to-noise ratio analysis, as the signal before the onset of the stimulus can be compared to the signal after the onset of the stimulus. This research described in this paper could not be conducted if there is no stimulus, as the signal-to-noise ratio could not be calculated. The sample size of the data set for analysis as a whole is sufficient, however, if this data is used to compare groups e.g. individuals diagnosed with schizophrenia and control subjects, analysing only 81 subjects to make assumptions on the population as a whole may not be sufficient for low confidence intervals (van Belle, 2008).

1.6 DOCUMENT OUTLINE

The remainder of this dissertation is organised as follows:

Chapter 2 – Review of Existing Literature

This chapter is comprised of a literary review to identify previously explored avenues of research that have investigated the study of brain activity, as well as the various techniques previously used to improve EEG signal analysis. A comparison of previously used techniques for reducing the noise in EEG signals is formulated, and, using this information, an alternative method is proposed, with the expectation of outperforming previous techniques.

Chapter 3 – Experiment, Design and Methodology

This chapter describes the experiment to be conducted in order to fill the gaps in the literature that have been identified in Chapter 2. The design of the experiment is comprehensively explained as well as the methods to be undertaken in order to analyse the performance of the experiment.

Chapter 4 – Results, Evaluation and Discussion

The results obtained from the experiment are described here, and the performance of the model is evaluated. The strengths and weaknesses of the executed model are discussed in detail, along with potential causes of error within the experiment that may have led to inaccurate results, and the areas along the experimental process at which improvements could have been made in order to design a better performing model.

Chapter 5 – Conclusion

Finally, an overview of the entire experiment is provided, describing the results, findings and insights procured during the experimentation process. Further avenues of potential research that could be pursued are proposed, as extensions of the experimentation carried out in this investigation.

2 REVIEW OF EXISITING LITERATURE

2.1 BRAIN ACTIVITY

The study of brain activity has been conducted for centuries, however, with the rapid growth of biological and medical sciences in the 19th century, research in the field of neurology became increasingly popular as the turn of the 20th century approached (Raichle, 2009). Since then, there has been an increasing interest in the study of cognitive processes, with ever improving results as the technological industry exponentially improves and expands.

In order to understand brain activity research in the 21st century, and the advancements being made to obtain more information, more accurately as well as more efficiently, it is imperative to review how brain activity research has progressed over time, and the different techniques that have been employed in order to achieve this.

2.1.1 Event Related Potentials

One focus of brain activity analysis that has been widely analysed throughout the 20th century and into the 21st century is the examination of *event-related potentials* (ERPs). Early researchers sought to characterise variations in the electrical activity of the brain during the performance of simple tasks and sensory processing (Davis, 1939; Walter, 1938). These electrical cognitive processes that occur in the brain when exposed to stimuli, or performing a task, were defined as event-related potentials. As researchers throughout the 20th century developed an increased interest in the activity of the brain, utilising signal averaging techniques became widely popular, with the ERP system becoming a key component in the cognitive neuroscientist analysis methods (Cooper, Winter, Crow & Walter, 1965; Davis H., 1964; Donchin & Cohen, 1967).

Although ERPs were first examined in the early 20th century, the use of ERPs still has a number of advantages over modern neuroimaging techniques, which makes it one of the most popular techniques used for analysing various cognitive functionalities. One of the most prominent reasons for using ERPs for analysing brain activity over other modern techniques involves the temporal resolution available from the use of this method. Using ERPs allows for the measurement of brain activity at a level of milliseconds. As well as this, there is no significant conduction delay between the signal created inside the brain due to a stimulus, and the potential differences recorded on the scalp, which allows for cognitive data to be almost immediately recorded (Nunez & Srinivasan, 2006). Using ERPs provides a direct measurement of the currency of the cognitive system, due to the fact that brain is an electrical conductor, such that not only can cognitive responses to stimuli be identified using ERP data, the response can be quantified in direct relation to the intensity of the stimuli itself.

Early researchers (Adrian & Yamagiwa, 1935; Li, McLennan & Jasper, 1952) proposed a hypothesis that these field potentials were due to "postsynaptic activity of neural ensembles", a hypothesis which is still is widely accepted in the scientific community (Nunez & Srinivasan, 2006; Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). The belief is that the potential defined as ERPs is caused by "electrical potentials generated in the extracellular fluid as ions flow across cell membranes and neurons talk to one another via neurotransmitters" (Woodman, 2010). As a potential field needs to be large enough to transmit through the brain, dura, skin and skull, Cooper et al. (1965) and Ebersole (1997) propose that in order to be able to create an electrical field this with this criterion, $\sim 10^7$ neurons need to be simultaneously active. As well as this, these neurons have to be perpendicularly aligned to the surface of the scalp. Voltages produced by neurons can also be negated if other neurons are simultaneously active but have potentials of opposite orientation (Nunez & Srinivasan, 2006; Luck, 2005). The neurons must be simultaneously active with approximately the same orientation in order to allow the summation of voltages. Due to this, Woodman concludes that the primary generators of ERPs must be "the postsynaptic potentials of cortical pyramidal cells", which are aligned perpendicular to the cortical surface. With knowledge of the "location and orientation of a specific neural generator in the brain", the pattern of the voltage to be observed across the scalp can then be predicted.

The use of ERPs allows for the observation of a sequence of cognitive operations that occur from prior to the delivery of sensory information to the peripheral nervous system until after a behavioural response is rendered. Davis (1939) demonstrated that stereotyped variations in voltage were induced due to the presence of sensory stimuli. This information was particularly useful, as it showed that ERP responses could be quantified. Walter, Cooper, Aldridge, McCallum, and Winter (1964) expanded on this and determined that brain activity associated with task performance preparation could also be measured quantitatively. The “contingent negative variation” was shown to intensify before the onset of the stimulus to which the individuals in the study were required to respond. The peaks and troughs of the ERP signals allow for effective real-time visualisations of cognitive processing.

2.2 EEG

Electroencephalogram (EEG) recordings were the first method developed for direct and non-invasive measurements of brain activity from human subjects (Berger, 1929). The combination of analysing ERPs with EEG recording techniques allowed for much more interpretable information obtained about cognitive processes.

The use of averaging EEG signal voltages time locked to the onset of a stimulus, recorded over numerous trials in order to separate similar potential fluctuations present in all trials from the background noise of the brain activity, was widely used (Donchin & Heffley, 1975; Dawson, 1954). During this time, the *10-20* system was established for standardizing the placement of electrodes which allowed for much replicable studies and comparable ERP results. Jasper (1958) describes the ideal placement of each of the electrodes, determined by “measurement from standard landmarks on the skull” with “adequate coverage of all parts of the head” and standard designated positions. The labelling of the electrodes related to brain areas rather than numbers was to allow effective communication between all individuals, specialists and non-specialists, would be more meaningful. The electrode placement and labels can be seen in **Figure 2.1.1-1.**

Although EEG signals have been widely and effectively used and analysed since the first human trial in 1929, there have been a number of issues that still affect the

accuracy of EEG signal measurements. If these issues are not addressed when performing EEG analysis, it can affect the signal and result in biased outcomes.
Due to the nature of the brain, any stimulus, internally or externally, can evoke a cognitive reaction e.g. blinking of the eye or the occurrence of a sudden external sound. It is therefore imperative that EEG studies are performed ideally in a stimulus free environment to avoid undesirable brain activity.

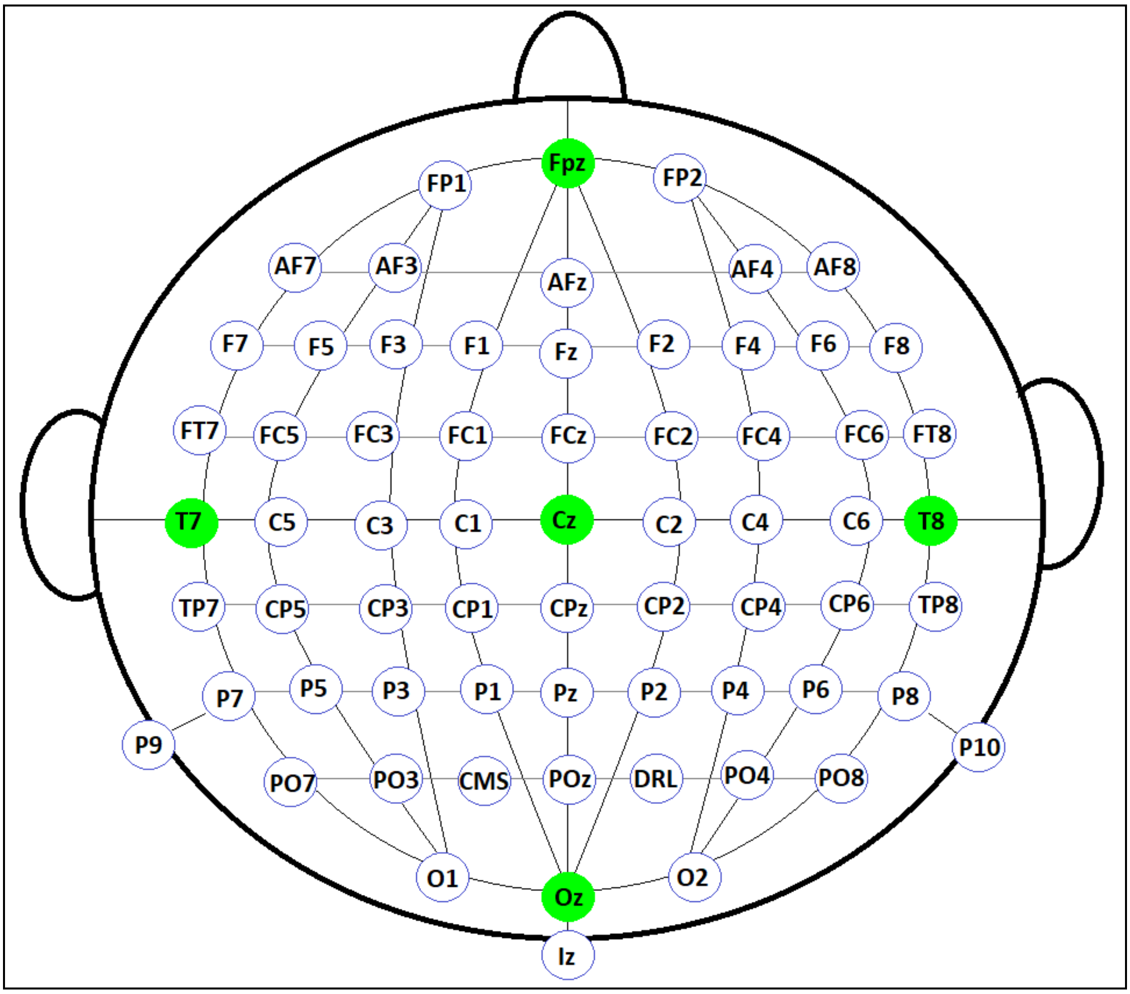


Figure 2.1.1-1 10-20 international electrode placement system (64 + 2 electrodes)

It may not be possible to avoid all stimuli affecting the participants, as something as seemingly insignificant as a brief eye movement can also significantly affect the EEG signal. For this reason, Woodman (2010) describes the importance of performing as "as many trials from each participant" as possible in order to lessen the impact of unintentional cognitive activity. Signal averaging across a large number of trials can significantly reduce the presence of bias in a study. Due to location of the "largest single electrical dipole in the head is the corneoretinal potential, which points from the back of the eye toward the front of the eye", when recording EEG signals and time-locked ERPs, even a small movement of the eye or blink can cause a large potential response. As well as this, eye movements can also be problematic when using auditory stimuli to analyse ERPs, due to visual attention being automatically drawn to the auditory signal source, as described by McDonald, Teder-Sälejärd and Hillyard (2000). The presence of these artifacts in EEG signals is just one of the issues that can affect the EEG signal readings.

Another feature that can negatively affect the accuracy of EEG signals is the concept of delayed stimuli presentation. When gathering this cognitive data from a participant, it is important to note that the waveforms and effects caused by the onset of various stimuli may last for at least a second, if not a number of seconds. If trials are conducted at a time interval of 1 second or less, the baseline period prior to the presence of the stimulus may coincide with the waveforms produced by the previous stimulus. It is therefore necessary to ensure that there are sufficient time intervals present between trials, to avoid the overlapping of brain activity. In addition to this, it may be necessary to retain the signal information of a number of milliseconds before the anticipatory brain activity in order to obtain baseline data to which the signal data can be compared. This is an important concept as it allows researchers to distinguish between the cognitive processes that are related to the stimuli and the background cognitive processes present throughout the brain, otherwise known as *noise*. In addition to this, Woods, Couchesne, Hillyard and Galambos (1980) describe the issue of ERP "refractoriness". The study shows that amplitudes of EEG signals can be reduced when the eliciting stimulus closely follows the preceding stimulus, with auditory stimuli ERP amplitudes reduced even after 1 second had passed between stimuli (Lu, Williamson, & & Kaufman, 1992). Nelson and Lassman (1973) have even suggested that some auditory ERP responses may last tens of seconds. This reinforces the argument that stimuli should not be presented to the participants in rapid succession, as it could lead to inaccurate signal results and biased analysis.

Although these issues can be addressed, and EEG signals can be recorded with as few unwanted environmental influences as possible, one of the most commonly known discrepancies that are present in EEG and ERP analysis is the existence of differences between individuals physiologically. Luck (2005) describes the extent to which differences between ERPs from a number of different individuals can exist. The discrepancies in these signals cannot be due to the noise present in the EEG signals or the process by which the data is gathered, as these differences are seen to be consistent across recording sessions and observers. It is hypothesised that the variances in ERPs between individuals may be due to the underlying cortical folding of the brain of each individual. According to this theory, each person may have similar ERP responses and cognitive processing procedures, yet because the potential pattern observed on the scalp is heavily dependent on the cortical folding pattern, the ultimate position at which certain cortexes end up may vary from person to person. There are a number of other physiological factors that may affect ERP responses from different individuals, such as skull thickness and conductivity. It is theorised that cognitive processes are common across most individuals, and it is the geometric noise in the anatomy of each individual that affects EEG readings.

Different physiological structures are not the only reason for discrepancies of EEG readings across individuals. It is often implicitly assumed that if an EEG signal is at a maximum at a given scalp location e.g. C2 on the 10-20 system, the signal is generated by the cortex located directly below the electrode (Homan, Herman, & Purdy, 1997). However, it is not possible to know the total number of simultaneously active neural network generators that contribute to the relative ERP, therefore it is not possible to determine whether the relevant activity that is measured is solely generated directly below that electrode. With two neurons simultaneously active but with opposite polarity, the net potential produced is 0V. Because of this, it is not possible to say for certain which areas of the brain are active or inactive. Helmholtz (1853) theorised that there are theoretically an infinite number of ways to measure 0 volts, even with significant amounts of potential activity present. However, according to Helmholtz, it is not possible to know where an electrical potential is generated if the number of simultaneously active generators is not known.

The issues mentioned above are important factors of which to be aware and consider when analysing EEG signals. In order to address these issues, an efficient, effective and robust approach must be taken, to ensure as much of the vital information is obtained and retained, while the unnecessary signal information is discarded, so as not to contaminate the entire signal. The signal information to be removed can be regarded as *noise*.

2.3 NOISE IN EEG

2.3.1 Definition & Background

Noise in any signal can be defined as "an unwanted signal that interferes with the communication or measurement of another signal" (Vaseghi, 2008). This definition can be effectively applied to Electroencephalography signals. Due to the nature of brain activity, there can be a large quantity of undesirable signals present in a signal that can affect accurate measurements of ERP data. Any discrepancy caused by the presence of noise in a signal can vastly alter the electrical readings acquired and can, thus, lead to biased analysis of these signals, affecting any further analysis that may be conducted on the signals in the future. Distinguishing between ERPs and noise, however, can be particularly difficult in brain activity analysis.

2.3.2 Causes of noise

There are a number of factors that can produce noise in EEG signals. The largest source of noise in the brain, the alpha band activity, has a frequency range of between 8 – 12 Hz, due to its direct relation to the deployment of spatial attention (Foster, Sutterer, Serences, Vogel, & Awh, 2017). A number of studies demonstrate how alpha-waves are particularly large when individuals are drowsy or bored (Berger, 1929; Pfurtscheller, Stancák, & Neuper, 1996). Because of this, Woodman (2010) proposes a number of methods that can be used in order to ensure participants are alert and engaged in the task required of them. One method suggested involves using short sets of task trials, allowing sufficient time in between each set for participants to re-adjust their focus and engage fully with the task at hand. This can be particularly effective for analysing ERP data as it reduces bias in the data as well as reducing the number of trials needed to acquire sufficient information.

As discussed in **Section 2.2**, almost any mechanical movement of the body will be preceded and superseded by cognitive activity. Actions as seemingly insignificant as a blink or slight eye movement can have a sizeable effect on the EEG signal. Joyce, Gorodnitsky and Kutas (2003) describe how brain signals created by eye movements and blinks "can be orders of magnitude larger than brain-generated electrical potentials". The authors discuss the importance of removing the effects caused by these artifacts, without rejecting all aspects of the contaminated data and signal information from the trial, as this could lead to large quantities of otherwise useful data being lost.

In order to measure the effect of noise on a signal, a sufficient mathematical process is needed. One widely used mathematical process used for quantifying the effect of noise on a signal is the calculation of the *signal-to-noise ratio*.

2.3.3 Signal-to-noise ratio

The signal-to-noise ratio (SNR) of a signal is a measurement that compares the actual signal information to the noise present in the signal. It can be simply defined as the ratio of the power of the signal to the power of the noise in a signal. An optimal SNR of a signal would tend towards infinity as the optimal noise of a signal would tend toward zero.

The signal to noise ratio for a signal is calculated using the formula:

$$\mathrm{SNR} = 20\log_{10}\left(\frac{S}{N}\right)$$

$$S = \sqrt{\frac{\sum(signal)^\wedge 2}{len(signal)}} \qquad N = \sqrt{\frac{\sum(noise)^\wedge 2}{len(noise)}}$$

signal:	voltage amplitude readings of the signal
noise:	voltage amplitude readings of the noise
len:	number of readings

2.4 NOISE REDUCTION TECHNIQUES IN EEG

The denoising of signals and removal of artifacts from EEG signals have been thoroughly investigated from many different perspectives, using numerous techniques, since the conception of electroencephalography in 1929. In recent years, there has been a focus on modern machine learning techniques in efforts to reduce the presence of noise in EEG signals, but before the development of these methods, classical mathematical techniques were relied upon in order to achieve this.

2.4.1 Classical Techniques

2.4.1.1 Common Classical Techniques

A common technique used for the removal of noise from many different types of signals is the use of filters. Ille, Berg and Scherg (2002) examine the use of spatial filters based on artifact and brain signal topographies. The study investigates whether using these spatial filters can remove artifacts completely without any distortion of relevant cognitive activity. The study demonstrated that the technique can prove to be an effective technique for reducing undesired noise in EEG signals. McFarland, McCane, David and Wolpaw (1997) examined the use of spatial features in greater depth, by comparing four alternative spatial filtering techniques, measuring the speed and accuracy at which subjects could control a mouse cursor to move to a target on a video screen. The cursor movement was performed by the subjects through control of the "amplitude and mu-rhythm activity" in electroencephalography signals recorded over the sensorimotor cortex. The four spatial filters that were used in the study included "a standard ear-reference, a common average reference (CAR), a small Laplacian (3 cm to set of surrounding electrodes) and a large Laplacian (6 cm to set of surrounding electrodes)". It was found that the CAR and large Laplacian techniques were the best performing filters, and performed significantly better than the ear-reference technique. The study discusses the importance of establishing appropriate noise reduction techniques in order maximize the signal-to-noise ratio of EEG signals, and therefore improve the speed and accuracy of EEG-based communication.

As effective as the use of filtering proves to be, Woodman (2010) describes how "no filtering settings exist that will remove noise without removing some of the signal itself'. This is due to similar frequency bands in which ERP components and signal noise are found. As well as this, as the intensity in which filtering is increased, the likelihood of amplitude and timing distortion is greatly increased, a concept which was explored by Duncan-Johnson and Donchin (1979). Because of this, it is imperitive to investigate althernative methods for noise reduction in EEG signals, methods which effectively reduce noise in a signal, but retain the vital information of the ERPs.

Other methodologies that have been used since the first recording of human EEG signals have involved specific feature extraction from these signals. Nikulin, Nolte and Curio (2011) have investigated the use of spatio-spectral decomposition for extracting neuronal oscillations from multi-channel EEG, magnetoencephalographic and local field potential recordings. The technique maximizes the power of the signal at a peak frequency, while minimizing the signal power at the surrounding frequency bins, with effective results, outperforming conventional procedures based on independent component analysis in accuracy. The running time of the spatio-spectral decomposition technique was in the range of milliseconds, separating it from other extraction methods, which could take as long as minutes or hours to complete. Haufe, Dahne and Nikulin (2014) propose dimensionality reduction of brain oscillations using spatio-spectral decomposition. The technique was found to particularly advantageous to other supervised techniques, because of its ease of use, absence of supervision and its capability of effective dimensionality reduction of multivariate signal data.

As well as decomposition, specific feature identification and extraction in EEG signals are important factors when addressing the concept of noise reduction. Artifact removal is a highly researched area in electroencephalography, as it can greatly aid in the reduction of signal noise. Maddirala and Shaik (2016) have also investigated artifact removal from EEG signals by combining singular spectrum analysis along with adaptive noise cancellation, in order to remove electrooculographic artifacts from EEG signals which degrade the performance of the brain-computer interface. The technique was found to outperform existing methods in terms of mean absolute error and root mean square error.

Feature extraction has also been effectively executed using various signal transformations. One example of signal transformation that has proven to be extremely effective is the *wavelet transformation/decomposition*. Wavelet transformations were first formulated by Grossman and Morlet (1985). The process is a mathematical procedure for performing signal analysis on time-dependent signals which decomposes a signal into a time-frequency representation, which can then be used for further analysis. These transformations have proven to be particularly useful for noise reduction in EEG. Peng et al. (2013) describe the removal of ocular artifacts from EEG signals using discrete wavelet transformation and adaptive noise cancellation. The technique was found to have superior performance with respect to "the recovery of true EEG signals" as well as an improved tracking performance. Ahmadi and Quiroga (2013) use wavelet transformations as a method for automatically denoising EEG signals, testing simulated, and real, visual and auditory ERP data. The method was shown to perform particularly well for amplitude and latency estimations of the simulated data when compared to *Donoho's thresholding* denoising technique (Donoho & Johnstone, 1994). The method provided a "simple, automatic and fast tool" that accommodated the study of single EEG trial responses and their behaviour.

Mahajan and Morshed (2015) also employ wavelet analysis techniques in order to investigate the use of modified multiscale sample entropy and kurtosis to automatically identify specific features in EEG signals. The study introduced an "unsupervised, robust and computationally fast algorithm" in order to perform biorthogonal wavelet decomposition was used to reduce the noise present in the EEG signal. The performance between the proposed method and independent composed analysis (ICA) decomposition were compared, analysing mutual information, correlation coefficient, and spectral coherence. Improved performance in reconstructed EEG signals using the proposed method was found when compared to the conventional zeroing-ICA and wavelet enhanced ICA artifact removal techniques. Using the proposed procedure negates the need for relying on manual intervention to accurately identify the independent artifactual elements, a method which can be inefficient when compared to the proposed wavelet decomposition method.

As can be seen above, there are benefits to using the techniques described as the primary noise reduction techniques of EEG signals. However, there is a vast body of evidence that supports the use of dimensionality reduction techniques for noise reduction. These techniques have been widely employed for noise reduction, as they provide certain advantages that are unobtainable using other methods.

2.4.1.2 PCA Dimensionality Reduction

Principal Component Analysis (PCA) is another dimensionality reduction technique and is one of the most widely used multivariate techniques in statstics and is commonly used technique for increasing EEG signal-to-noise ratios. PCA was first conceptualised in 1901 by Karl Pearson (Pearson, 1901). The study explored the concept of an "affine space that best fits a series of points, where the fit is measured by the sum of squared orthogonoal distances from each point to the space" (Reris & Brooks, 2015). This concept has been developed and expanded upon since then and is now applied to a vast number of different disciplines, one of which is signal processing.

Garćıa-Laencina, Rodŕıguez-Bermudez and Roca-Dorda (2014) propose the use of Principal Component Analysis, Locality Preserving Projections and Local Fisher Discriminant Analysis to explore dimensionality reduction in EEG signals as a pre-processing method for classification. The techniques were found to be particularly beneficial in the pre-processing stage of data preparation, resulting in high accuracy performance and classification results. Kang and Zhizeng (2012) explore the use of PCA combined with density estimation blind source separation in order to achieve successful de-noising of multi-channel EEG signals. The study showed that the proposed method can successfully eliminate the principal interference of multi-channel EEG signals, both effectively and rapidly, while maintaining stability. Babu and Prasad (2011) compare the performance of PCA dimensionality reduction against wavelet threshold for the removal of ocular artifact from EEG signals. The performance of the techniques was measured by the respective signal-to-noise ratios of the EEG signals after processing by each of these techniques. It was observed that using the PCA method resulted in an increased signal-to-noise ratio when compared to that of the wavelet threshold.

As seen above, the use of PCA for noise reduction in EEG signals can be very effective. However, there are dimensionality reduction techniques that have been employed other than PCA that have also proven to be successful for the reduction of noise in EEG signals.

2.4.1.3 ICA Dimensionality Reduction

Independent Component Analysis was first conceptualised by Hérault and Jutten (1986), named so because of the its similarities with Principal Component Analysis. The study proposed a learning algorithm that had the ability to blindly separate mixtures of independent signals, with the only assumption being that the sources were independent. This idea has been widely expanded upon since then and, similar to PCA, is applied to a number of fields of study today.

Srinivasulu and Sreenath Reddy (2012) adopt the use of ICA for artifact removal and denoising from EEG signals. Each source of electrical potentials, in particular ocular artifacts from which PCA was unable to separate other brain signals, projects a unique topography onto the subject's scalp. The study explores whether it is possible to successfully separate the EEG signal into "mutually independent scalp maps" in order to reduce the noise present in the signal. It was found that the use of ICA could not only protect the useful signals to be analysed, but could also weaken, if not fully remove, unwanted artifacts of the EEG signal. Vorobyov and Cichocki (2002) also explore the potentials of ICA for EEG signal separation as a combination with filtering for extensive noise reduction. Separating the signals using ICA enables the filter to fulfil a much more comprehensive filtering process, as each filter would be individualised for each source. The combination of ICA and other denoising methods have proven extremely successful for EEG noise reduction.

Jung, et al. (1997) discuss how rejecting EEG data contaminated with unwanted artifacts can result in a significant loss of information, a result which may be impractical for clinical data. Because of this, it is extremely important to accurately separate sources in order to reduce EEG signal noise. The study explores using ICA to separate signal sources through blind source separation. It was found that removing a wide variety of EEG artifacts could be successfully achieved to a better extent than the use of regression models. ICA was found to be computationally efficient and can successfully separate EEG and its artifacts without relying on "clean" reference channels to use as baseline data. As successful as ICA has proven to be, Jung et al. note that "the results of ICA are meaningful only when the amount of data and number of channels are large enough". Without this criterion fulfilled, biased results could be produced. The authors suggest conducting an investigation into the minimum data length and number of channels needed for artifact removal. As well as this, ICA requires visual inspection in order to select and remove undesired artifacts manually from the signal.

The benefits of using classical techniques for EEG signal noise reduction are evident and have been proven numerous times to be extremely efficient and effective. However, with recent technological advancements, there has been an increased interest in the use of machine learning techniques to reduce noise in all signal types.

2.4.2 Machine Learning Techniques

With the rapid development of machine learning in recent years, numerous newly proposed strategies have been implemented investigating whether the signal to noise ratio of EEG signals can be significantly increased using machine learning techniques, when compared to classical techniques, such as those mentioned above.

2.4.2.1 Autoencoders

"Autoencoders are neural networks trained to reproduce its input as accurately as possible" (Lauzon, 2012). The aim of these types of neural networks is to capture important variation factors of the input data while negating what is deemed as unnecessary information, and reproduce this reduced data as the output. As a by-product of this, it is hypothesised that the autoencoder will remove what is considered noise by the researcher, leaving only the relevant EEG signal. The implementation of a number of different neural network types has been examined for EEG signal feature extraction and denoising, such as sparse autoencoders, stacked autoencoders, recurrent neural networks and convolutional autoencoders.

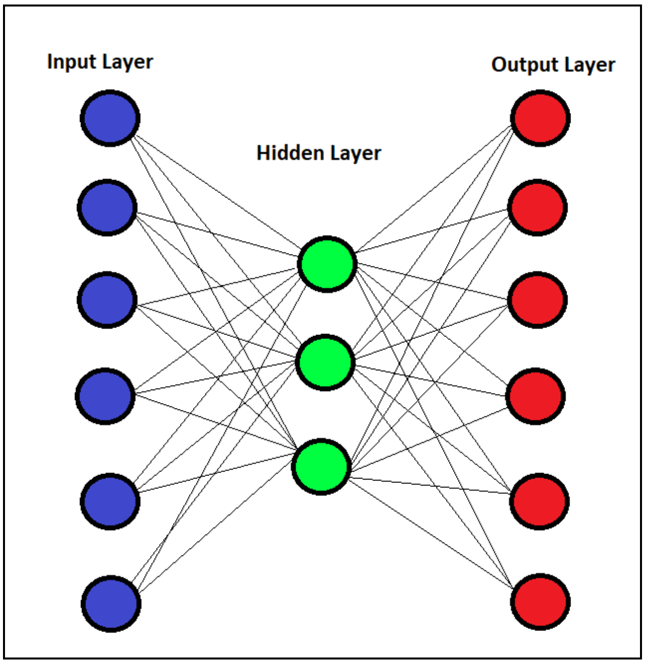


Figure 2.4.2-1 Basic structure of a sparse autoencoder

The potentials of using stacked autoencoders have been investigated for feature identification and extraction in EEG signals. Stacked autoencoders are neural networks comprised of several layers of sparse autoencoders, where each hidden layer output is the input for the successive hidden layer, as described by Vincent, Larochelle, Lajoie, Bengio and Manzagol (2010). Their investigation provides an insight into building deep networks based on these types of autoencoders. This offers an informative foundation for the examination and development of autoencoding algorithms for the denoising of signals. Supratak, Li and Guo (2014) examine the potentials of stacked autoencoders for feature extraction from raw, unlabelled EEG data in order to detect epileptic seizures and Kulasingham, Vibujithan, and De Silva, (2016) discuss the comparison of the classification ability of stacked autoencoders and Deep Belief Networks for EEG signal features. Stacked autoencoders were found to perform particularly well with high accuracy results. It was found, however, that the use of stacked autoencoders included high model complexities as well as long training time durations. For effective and efficient results, it is particularly important to address these issues note when designing a model for EEG signal analysis.

2.4.2.2 Recurrent Neural Networks

As well as autoencoding neural networks, *recurrent* neural networks (RNNs) have also been studied in order to determine to what extent they can be used to denoise EEG signals. Pardede, Turnip, Robinson Manalu and Turnip (2015) employ an adaptive RNN in order to reduce the effect of ocular artifacts in EEG signals. The EEG signals were analysed for three conditions; no ocular input, eye blinks, and closed eyes. The proposed method was successfully able to estimate the cognitive activity according to the given stimulus, and remove the artifacts from all subjects. Selvan and Srinivasan (2000) state the importance of removing ocular artifacts from EEG signals. The study proposes using a combination of adaptive noise cancellation and adaptive signal enhancement in a single recurrent neural network. The RNN was found to successfully remove unwanted artifacts from the signals, increasing the signal-to-noise ratio of the EEG signals.

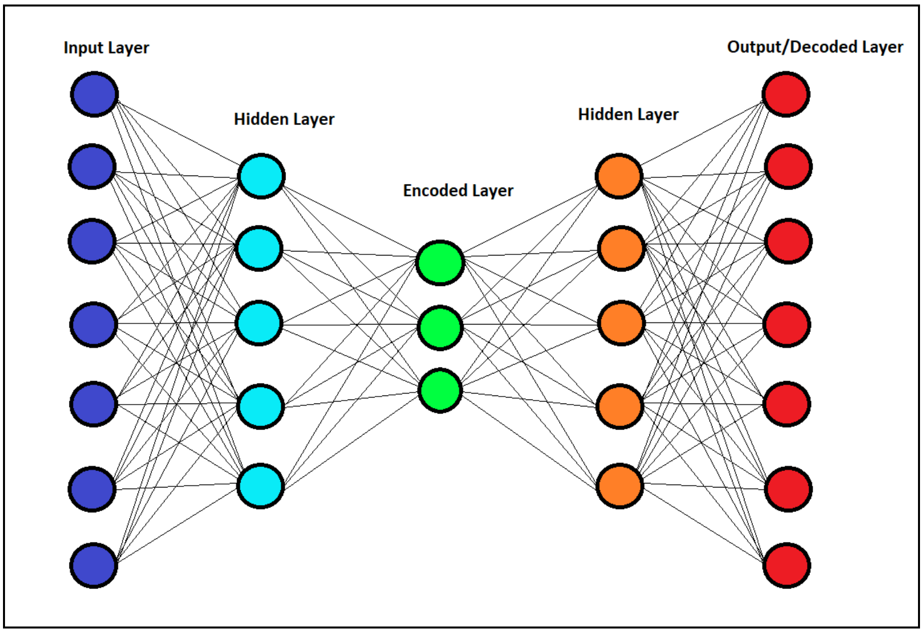


Figure 2.4.2-2 Basic structure of a stacked autoencoder

2.4.2.3 Alternative Machine Learning Techniques

Other than autoencoders, other Machine Learning approaches have been investigated for EEG signals analysis. Support Vector Machines (SVM) have been investigated for EEG signal processing and artifact handling by O’Regan and Marnane (2013). Other neural network configurations have been also been investigated for EEG signal analysis. Nguyen et al. (2012) propose a wavelet neural network for the removal of EOG artifacts from EEG signals. The proposed method successfully removed the artifacts with better results than wavelet thresholding without diminishing important EEG data, while being computationally and more convenient that the tested ICA model.

Although it can be seen that there have been a number of different approaches to EEG signal analysis using various machine learning techniques, there have been a limited number of investigations performed relating to denoising EEG signals specifically using autoencoding Convolutional Neural Networks (CNNs). There is reason to believe that this method could yield more precise results than previous alternative approaches for denoising these signals. In order to investigate this approach, it is important to closely examine the structure and behaviour of CNNs from a number of different perspectives, and from this, asses how they could prove effective at EEG signal denoising.

2.5 CONVOLUTIONAL AUTOENCODERS

Convolutional Neural Networks were first designed by Fukushima (1980). The structure of the CNN was based upon the research carried out by Hubel and Wiesel (1968) which examined the hierarchy structure of the visual nervous system in animals, noting that neurons in the visual cortexes of animals respond individually to small regions of the visual field. The system developed by Fukushima, which was named “neocogitron”, comprised of two layers: a convolution layer and a downsampling (pooling) layer. The convolutional layer initialises the neural network, where the features of the neural network need to be learned using parameter weight sharing, while the downsampling layer summarises regions of a feature map of the convolution layer in a lower dimensional space. This combination of layers produces a summarised version of the data, and has been applied to many fields of research in order to gain further insights previously unobtainable by other methods.

The main advantage of using CNNs over other methods, is such that CNNs can automatically detect what are seen as important features in a system without human supervision. The use of weight sharing efficiently reduces the complexity and number of weights in the network, resulting in a significantly more computationally efficient structure (Neubauer, 1998).

2.5.1 Examples of CAE noise reduction

Denoising autoencoders were first introduced by Vincent et al. (2008) as an extension to classical autoencoders. Autoencoding techniques based on a CNN structure have been implemented in numerous studies since then, in order to investigate the potential for feature extraction and noise reduction in various signals. A basic denoising convolutional autoencoder consists of an input later, with a convolution layer and pooling layer. This result is an *encoded* layer. An upsampling layer and deconvolution layer are added to system in order to restore the data to its original shape.

The use of Convolutional Autoencoders (CAE) has been found to be particularly beneficial in the medical industry. Gondara (2016) demonstrates the potentials of using stacked convolutional denoising autoencoders for the reduction of noise in mammogram images and dental radiography images using small sample sizes, as is typical in the medical industry. The technique was found to successfully reduce the noise in the aforementioned medical images, using a sample size of ~300 images. It was found with particularly high noise levels, at which other denoising methods would be unsuccessful, the CAE was able to successfully recover the original signal.

Zhao, Wang, Zhang and Zhang (2015) explore the use of CAE denoising techniques for the improving automatic speech recognition. The study describes the establishment of a CAE designed to learn local musical patterns and remove them from music-embedded speech signals. The model was evaluated using English and Chinese, combined with four types of music: Piano-*Beethoven Moonlight Sonata Chapter 3*, Violin-*Theme from Schindler's List*, Symphony-*Radetzky March*, Rap-*Nunchaku Jay Chow*. The CAE was found to successfully remove musical patterns from the music embedded speech samples. As well as this, the model was found to be consistent across both languages as well as the all four different music types.

However, as denoising autoencoding techniques are relatively recent in their discovery and analysis, there are still many areas of focus to which denoising autoencoders have not been applied, specifically convolutional autoencoders (CAEs).

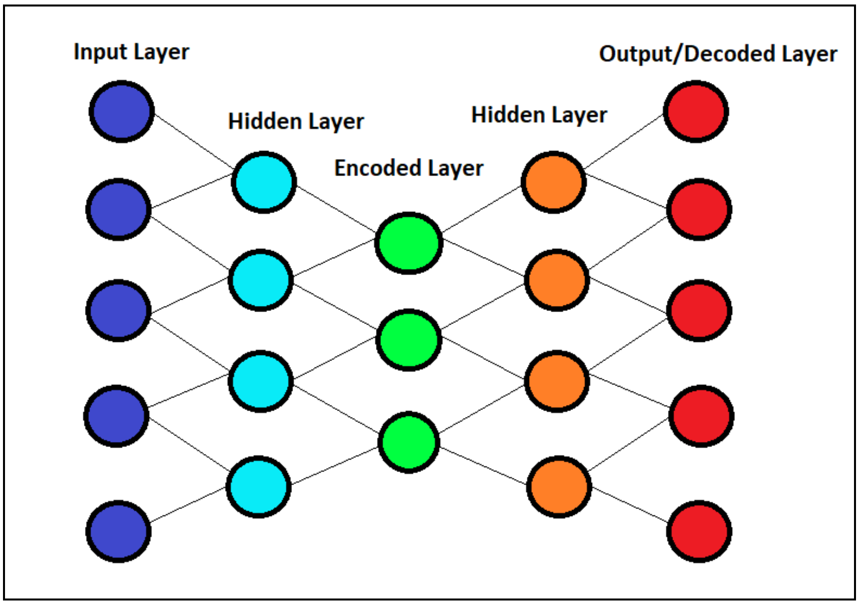


Figure 2.5.1-1 Basic structure of stacked convolutional autoencoder

2.6 SUMMARY

2.6.1 Overview

As can be seen in this chapter, there have been many approaches to feature extraction and noise reduction in EEG signals, both classical approaches as well as machine learning approaches. Classical approaches that have proven to be effective for noise reduction include the use of filters, transformations and dimensionality reduction. Although these methods were found to perform successfully and accurately, there are limitations to the extent at which they can perform. Because of this, recently established machine learning techniques have been employed in order to address the issues which previously eluded classical noise reduction techniques, with a number of avenues of research yet to be pursued.

2.6.2 Gaps in Literature

Although comprehensive research has been conducted into electroencephalography signal feature identification, extraction and noise removal from many aspects, a small number of gaps in this domain have been identified, as mentioned by Yang, Duan, Fan, Hu, and Wang (2018), who describes the potentials of convolutional autoencoders for EEG signal strengthening in future works. From this, similar recommendations, and researching the numerous other methods that have been discussed, the research that will be pursued in this investigation will involve the use of convolutional autoencoders in order to reduce the noise found in EEG signals. There is evidence to suggest that the performance of convolutional autoencoding techniques is better than previously proven effective feature enhancement of signals such as Principal Component Analysis, as described by Helal, Eldawlatly and Taher (2017), and it is expected that in this study, convolutional autoencoders could also prove to be similar to, or more effective than, PCA at reducing noise in electroencephalography signals.

2.6.3 Research Question

These investigations lead to the theory that, as basic autoencoders are seen to be effective at extracting features from EEG signals, convolutional autoencoders could be more precise and efficient at reducing noise in EEG signals.

"To what extent can a stacked autoencoder, designed using one-dimensional convolutional neural network layers, increase the signal-to-noise ratio of electroencephalography signals, and is it possible to outperform previously proven effective noise reduction techniques, such as Principal Component Analysis?"

3 EXPERIMENT, DESIGN & METHODOLOGY

In this project, a convolutional autoencoder will be designed in order to increase the signal-to-noise ratio of Electroencephalogram signals. This chapter describes the design of the experiment that will be implemented in order to decide whether the null hypothesis can be rejected.

The null hypothesis to be tested is described in this chapter, as well as the respective alternate hypothesis. This hypothesis will be tested by designing a convolutional autoencoder, preparing the data to be used appropriately, and tuning the hyperparameters of the convolutional autoencoder to create the best performing model. How the data was originally collected is described in this chapter, as well as the preparation performed on the data to be used in this experiment, its final shape and size, and how it is to be used and processed by the convolutional autoencoder.

As well as this, the details of how the acquired results will be evaluated are described, and how specific statistical techniques can be used to determine the whether the null hypothesis can be rejected. The final model design will then be summarised, outlining the strengths and limitations of the proposed model.

The experimental methodology undertaken in this investigation closely follows the CRISP-DM (**CR**oss-**I**ndustry **S**tandard **P**rocess for **D**ata **M**ining) methodology for the required data mining tasks (Shearar, 2000).

3.1 HYPOTHESIS

H_1: If a stacked autoencoder is designed using a one-dimensional convolutional neural network layer, the signal-to-noise ratio of electroencephalography data can be increased when compared to principal component analysis.

H_0: If a stacked autoencoder is designed using a one-dimensional convolutional neural network layer, the signal-to-noise ratio of electroencephalography data cannot be increased when compared to principal component analysis.

3.2 DATA

3.2.1 Data Collection

The data used in this experiment consists of EEG signals from 81 separate subjects, recorded on 64 individual electrodes, placed at specific locations on each subject's scalp, as described by Ford, Palzes, Roach, and Mathalon (2013). Each subject was exposed to three separate conditions/stimuli, which consisted of individual button pressing tasks. Condition 1 involved the subject pressing a button every 1-2 seconds to deliver a 1000 Hz, 80 dB sound pressure level, tone with no delay between pressing the button and the tone onset. The task was stopped after ~100 audio tones were generated. Condition 2 involved the subject passively listening to a generated tone but not performing any motor task, also conducted ~100 times. Condition 3 consisted of the subject pressing a button at a similar rate as Condition 1, with no tone generated, performed ~100 times.

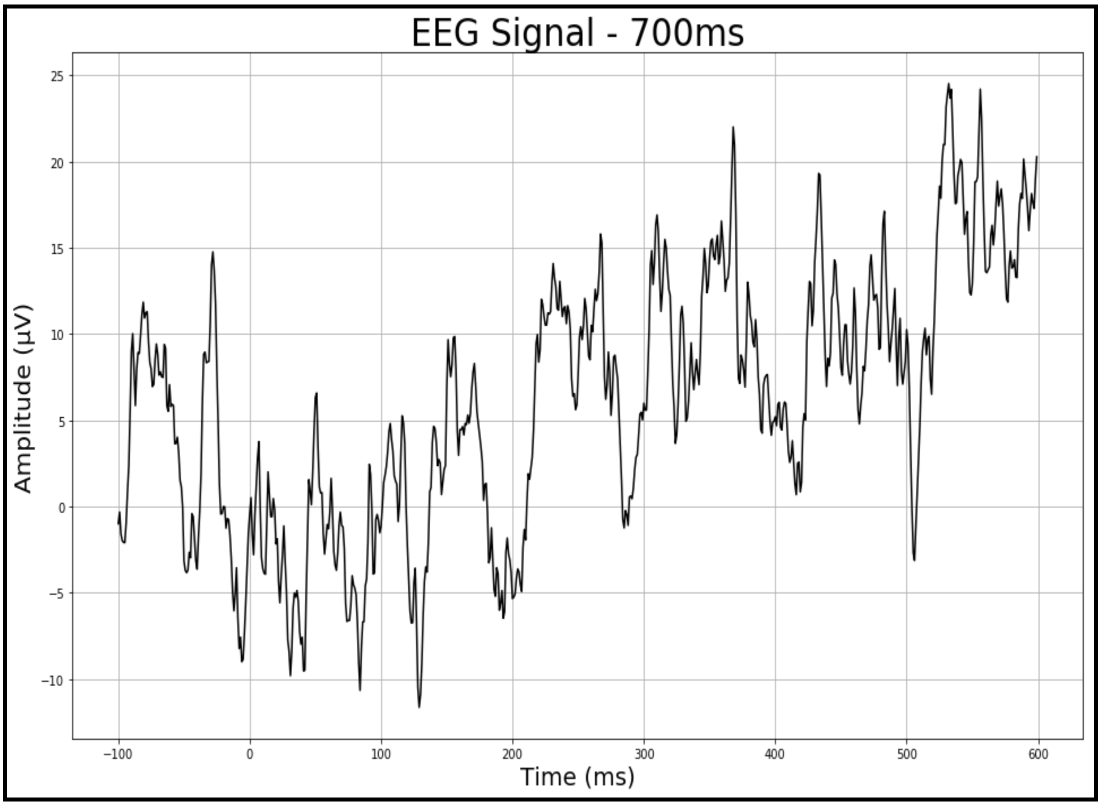


Figure 3.2.1-1 Example of EEG signal over 700ms with stimulus presented at 100ms

The voltage amplitude from each electrode located on the subject's scalp was recorded for each trial over a time of 700ms; 100ms before the onset of the stimulus and 600ms after the onset of the stimulus. The data recorded before the occurrence of the stimulus is regarded as the noise (-100ms to -1ms inclusive). The data recorded after the occurrence of the stimulus is regarded as the signal and noise combined (0ms to 599ms inclusive).

An example of a plot of the obtained signal can be seen in *Figure 3.2.2*.

3.2.2 Data Preparation

The voltage amplitude readings from each of the 64 electrodes were recorded as individual vectors of length 700 for each trial conducted. Three dataframes for each subject were created, one for each condition. The data from each channel and trial was concatenated for each condition resulting in three dataframes for each subject with 700 rows and ~6,000 columns per dataframe. The individual dataframes for each subject was also separately concatenated to produce another dataframe with a shape of 700 rows and ~500,000 columns per condition. Each column is to be used as a 1-dimensional vector of length 700 as an input for the convolutional autoencoder.

For each subject, the vectors from all three condition dataframes were separated into training and test data at a ratio of 70:30 respectively, resulting in 81 training and test datasets to be used for the individual convolutional autoencoder (Multiple CAE) for each subject. Similarly, the overall dataframe was also split into training and test data at a ratio of 70:30 respectively to be used for the single convolutional autoencoder (Single CAE).

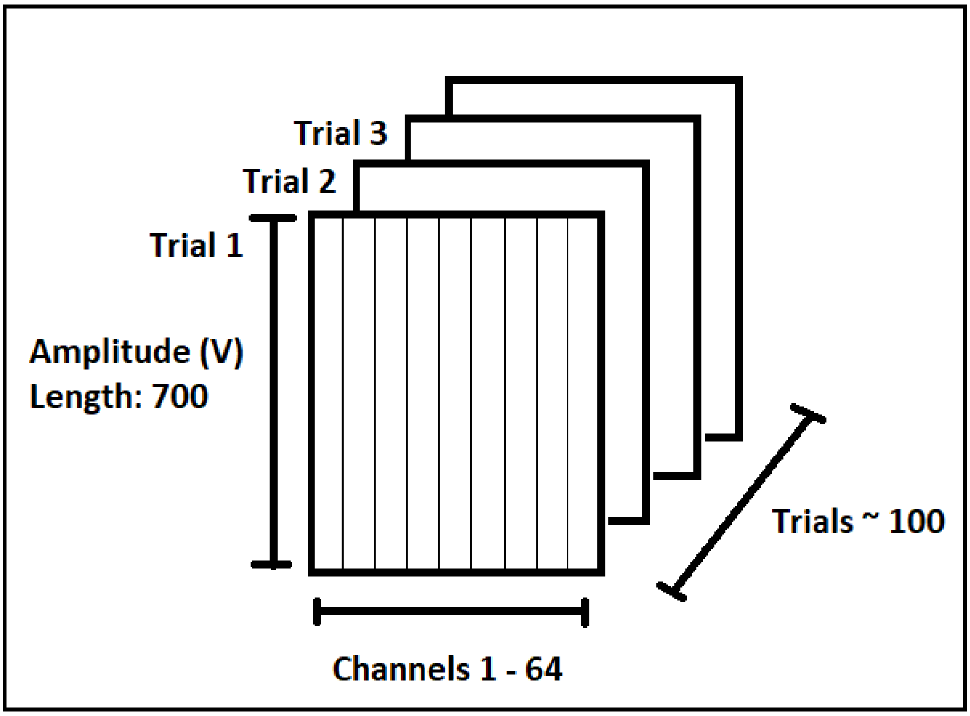


Figure 3.2.2-1 Example of EEG data shape for a single Subject and single Condition before preprocessing

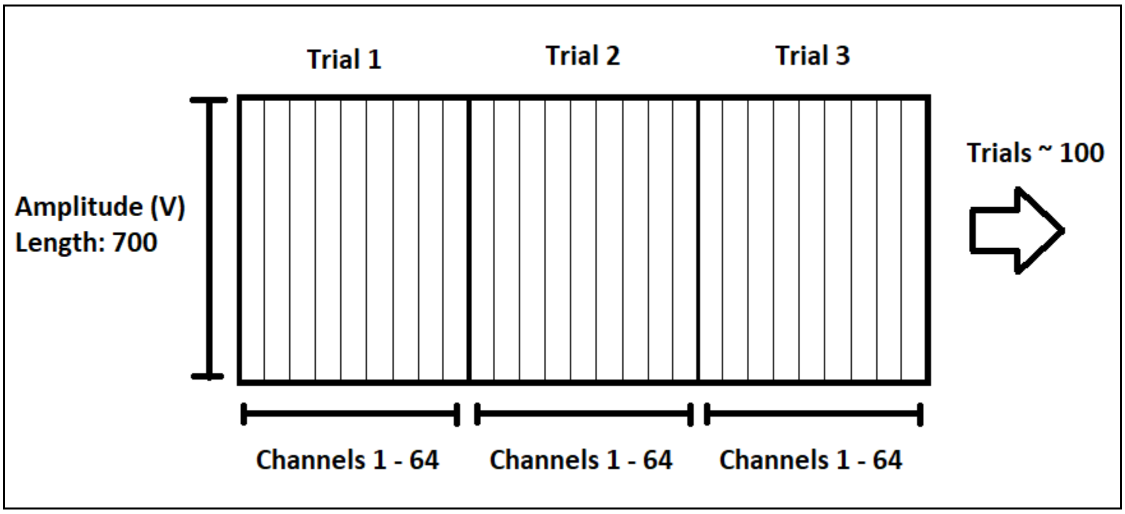


Figure 3.2.2-2 Example off EEG data shape for single Subject and single Condition after concatenation

3.3 CONVOLUTIONAL AUTOENCODER DESIGN

For the designing of a convolutional autoencoder (CAE), there are a number of hyperparameters that must be selected to determine the most efficient and accurate model that can be established. Each column is to be used as a 1-dimensional vector of length 700 as an input for the CAE.

3.3.1 Hyperparameter Tuning

There are a number of hyperparameters that need to be tuned in order to create an effective CAE. The hyperparameters to be addressed include:

1. *number of convolution layers – number of layers that feature in the convolutional autoencoder*
2. *filter size – the dimensionality of the output space of the convolutional neural network*
3. *kernel size – specifies the length/size of the convolutional window that convolves over the data*
4. *pooling strategies – reduces the dimensionality of the vector*
5. *pooling size – the factor by which the dimension is reduced*
6. *activation function – a transformation that is applied to the input signal of a neural network layer, with the output transferred to the next layer*
7. *number of epochs – number of times a dataset is passed forward and backward through the neural network*
8. *batch size – size of the dataset to be passed through the neural network for each epoch*
9. *optimizer – determines the most accurate possible model from the data*
10. *loss function – function used to determine to what extent the algorithm models the dataset*
11. *training/test split – ratio split of data for training and test data*

- **1**. Using one hidden layer in a neural network "…can approximate any function that contains a continuous mapping from one finite space to another", while using two layers "…can represent an arbitrary decision boundary to arbitrary accuracy with rational activation functions and can approximate any smooth mapping to any accuracy" (Heaton, 2018).

- **2/3**. Appropriate kernel and filter sizes must be decided such that the accuracy of the model is unaffected, but the number of parameters is low enough that the model is not overly computationally expensive. Larger kernel sizes result in slower training/test times by the model but allows for complex learning. Smaller kernel sizes have quicker training/test times but do not learn to recognise local features. Smaller values for filter sizes allow for more recognition of more local features in the data whereas larger values are used for more generic feature recognition.

- **4/5**. "A pooling function replaces the output of the net at a certain location with a summary statistic of the nearby outputs" (Goodfellow, Bengio, & Courville, 2016, p. 335). Pooling involves non-linear down-sampling of the data by summarising the data within a window by a single feature. Two commonly used pooling strategies include *MaxPooling* and *AveragePooling*; *MaxPooling* involves summarising the data by using the maximum value within a window, whereas *AveragePooling* summarises the data with the mean value of data within a window. The pooling size is the factor by which each feature map is reduced.

- **6**. Activation functions are linear or non-linear. Non-linear activation functions are generally used for classification problems as the output is always between 0 and 1. Linear activation functions allow for outputs between -infinity and infinity.

- **7/8**. The ideal number of epochs can be determined using a validation curve. There is a trade-off between the batch size and the number of iterations for training the Neural Network. Choosing the correct value for these parameters can combat the issue of overfitting and produce a higher quality model.

- **9**. There are a number of available optimizers to use for CNNs. Stochastic Gradient Descent (SGD), RMSProp, Adagrad, Adadelta, Adam, Adamax, Nadam.

- **10.** An appropriate Loss Function will be chosen based on the objective of the CAE. There are a large number of available Loss Functions available to use.

- **11.** A suitable split for the training and test data is needed to ensure that the model has enough data with which to be trained, but also enough test data in order to accurately evaluate the model performance.

3.3.2 Final Model Selection

The final model to be used will consist of the hyperparameters that will result in the best model performance.

Final Model Selection	
Number of hidden convolution layers	
Kernel size	
Filter size	
Pooling strategy	
Pooling size	
Activation function	
Number of epochs	
Batch size	
Optimizer	
Loss function	
Training/Test Split	

Table 3.3.2-1 Final model hyperparameter to be selected

- **1.** The number of hidden convolution layers to be chosen will be used in order to produce a simple convolutional neural network, reducing computation time but with enough complexity to learn global features in the first layer and learn local features in the second.

- **2/3.** A larger filter size with a larger kernel size will be used in the first layer of the convolutional autoencoder in order to generate an overall interpretation of the signal, learning global features. A smaller filter size and a smaller kernel size will be used in the subsequent layers of the convolutional neural network to learn more complex and local features.

- **4.** As an autoencoder is being designed in order to recreate a signal, it is important to use as much information from the signal as possible. Using *MaxPooling* only uses the maximum value in the window and disregards the other values, whereas using *AveragePooling* takes all information in the window into account.

- **5.** A larger pooling size for the initial hidden layer will be used for the first layer with gradually decreasing with pooling sizes subsequent layers.

- **6.** As the aim of the this autoencoder is to reproduce the original signal, an appropriate activation function must be used. Certain activation functions, such as *RELU* and *tanh*, produce values within a limited range, whereas a *Linear* activation function is not constrained to limits and can reproduce any real input value.

- **7/8.** The epoch number at which the validation curve of the training loss and validation loss of the data converges will be chosen as the ideal epoch number. The batch size to be used will be small enough so as not to overfit the model, but large enough to avoid sample bias.

- **9.** Adam is a widely used optimizer for neural networks as it is “computationally efficient, has little memory requirements has little memory requirements, is invariant to diagonal rescaling of the gradients, and is well suited for problems that are large in terms of data and/or parameters” (Kingma, D. & Ba, J., 2015).

- **10.** As the convolutional autoencoder is used as a regressor, a regression loss function must be used. The Mean Squared Error is a broadly used regressor, as it produces larger errors when larger mistakes are made, compared to smaller errors for smaller mistakes, punishing larger mistakes more harshly. It is calculated by squaring the mean of the squared differences between predicted and actual values, with results always positive.

- **11.** The data was split into training and test data at a ratio of 70:30.

3.4 EVALUATION OF RESULTS

3.4.1 Signal-to-noise ratio

The signal to noise ratio for each signal is calculated using the formula:

$$\mathrm{SNR} = 20\log_{10}\left(\frac{S}{N}\right)$$

$$S = \sqrt{\frac{\sum(signal)\hat{}2}{len(signal)}} \qquad N = \sqrt{\frac{\sum(noise)\hat{}2}{len(noise)}}$$

signal:	voltage amplitude readings 0ms – 600ms
noise:	voltage amplitude readings -100ms – 0ms
len:	number of readings

The voltage amplitude of the signal was recorded over 700ms, recorded at every millisecond. At 100ms, the stimulus was introduced. The signal readings before the stimulus and regarded as the noise of the signal and the signal readings after the stimulus are regarded as the signal. These values of these readings were squared and summed, and then divided by the number of readings taken. The log of the ratio of the square root of these, signal and noise, was calculated and multiplied by a factor of 20 in order to calculate the power signal-to-noise ratio of the signal.

This process is conducted for (1) the test data before being processed by the convolutional autoencoder, (2) the PCA reduced test data (baseline), and (3) the test data after being processed by the convolutional autoencoder.
The distribution of the signal-to-noise ratios for each of processes (1 – 3) are plotted in order to visualise their distributions and visually estimate whether there has been a significant shift in the signal-to-noise ratios before and after being processed by PCA as well as the CAE.

3.4.2 Hypothesis Testing

To determine whether the signal-to-noise ratio has increased when using convolutional autoencoding techniques compared to using Principal Component Analysis, a number of tests will be conducted.

3.4.2.1 Principal Component Analysis

Principal Component Analysis will be used as a baseline test to which the convolutional autoencoder results will be compared in order to quantify the effectiveness of the designed model at increasing the signal-to-noise ratio of the signals.

By deconstructing signals using PCA transformations, the input signal information is summarised in a reduced number of components. By reconstructing a signal using an inverse PCA transformation, the number of components in a signal is increased. When these transformations are combined, a reconstructed signal is produced with summarised information the original signal, but the same number of components. This process results in only vital information being retained, removing what is considered irrelevant information, reducing the noise in the signal, and thus increasing the signal-to-noise ratio.

When using PCA for dimensionality reduction, factors should be stopped when at least 95% of the variance is explained, as recommended by Hair, Anderson, Tatham and

Black (1995). Because of this an explained variance of 95% will be chosen for the PCA reduction.

3.4.2.2 Single Stacked Convolutional Autoencoder

A single stacked convolutional autoencoder was designed for the entire dataset. The distribution of signal-to-noise ratios will be calculated for the raw signals, PCA reduced signals and autoencoded signals for each Condition. The Harrell-Davis estimated (Harrell, F. & Davis, C., 1982) decile values will be calculated for each distribution. This method was chosen to compare distributions as it provides a more informative analysis of the entire distribution rather than single value comparison such as the mean of median.

For each condition, the differences in the respective Harrell-Davis decile values between the raw signal (RAW) SNRs and PCA reduced signal (PCA) SNRs, as well as the differences in the respective Harrell-Davis decile values between the RAW SNRs and autoencoded signal (CAE) SNRs, will be recorded. The mean values of these differences will be compared. If the mean difference of the RAW-CAE SNRs is greater than the RAW-PCA SNRs, there will be evidence to support rejecting the null hypothesis.
To further expand on this, the difference between each Harrell-Davis decile value will also be recorded and examined. This will reveal whether one or more of the decile difference values is an outsider, which could skew the results. If the majority of the deciles, 5 or above, of the RAW-CAE decile differences are greater than the respective RAW-PCA decile differences, this will also indicate that there is evidence to support rejecting the null hypothesis.

3.4.2.3 Multiple Stacked Convolutional Autoencoders

A convolutional autoencoder will also be designed for each Subject of the data using the same hyperparameters as the *Single CAE*. This is to establish whether a generalised model could be designed that could be applied to any individual, rather than being person specific. In order to test the performance of the convolutional autoencoder for each Subject, similar statistical analysis tests as are to be performed on the Single CAE, will be performed on the *Multiple CAE* for each Subject.

The mean Harrell-Davis decile difference calculations will be performed on each of the SNR distributions for each subject. Each of the RAW-CAE results will be compared to their respective RAW-PCA results. If the majority (41 of 81 subjects) of the raw-autoencoder results are greater than their respective raw-PCA results, it can be concluded that there is sufficient evidence to reject the null hypothesis. This will also be performed for each of the 3 Conditions.

The distribution of Harrell-Davis decile differences for RAW-PCA and RAW-CAE SNR values will also be analysed in order to establish whether outliers within the dataset significantly skew the obtained results. This also gives insight into which deciles exhibit the greatest variation and range differences.

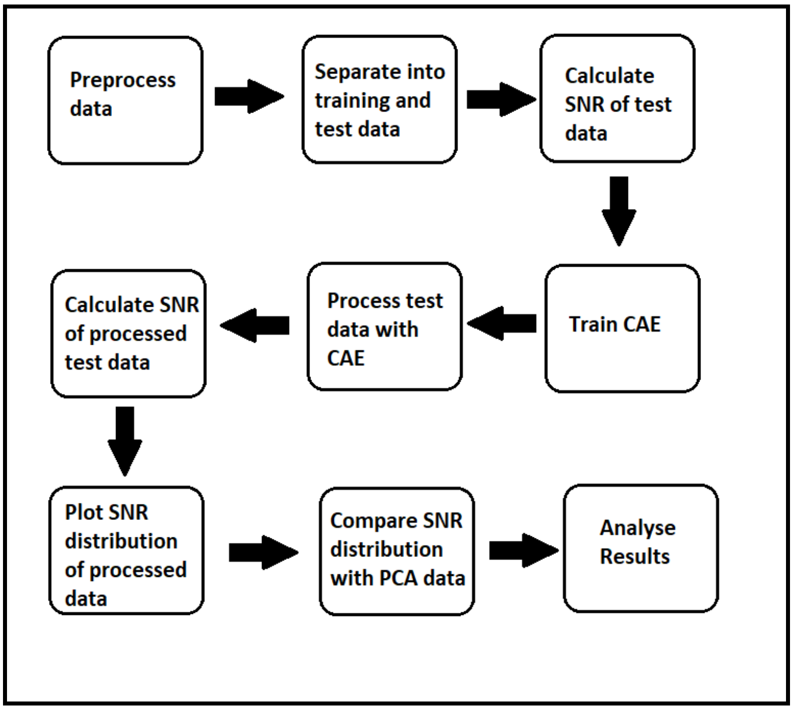


Figure 3.4.2-1 Diagram of testing and analysis process

3.5 SUMMARY

A 1-dimensional convolutional autoencoder was designed in order to increase the signal-to-noise ratio of EEG signals while minimising the loss of important information. The data was obtained using a *10-20* system Electroencephalography cap, which had 64 electrodes, located at specific areas on a Subject's scalp. The voltage amplitude of the brain activity from 81 Subjects was recorded, with each Subject performing three independent tasks, ~100 times for each task. Each trial was recorded over 700ms. The recorded data from each electrode was used as a one-dimensional vector with a length of 700, a single reading every millisecond. The data from each Subject was converted into a dataframe of 700 rows and ~6,000 columns for each Condition, with each column to be used as an input vector to the convolutional autoencoder.

Two approaches will be taken to test the proposed hypothesis.

Single CAE: The first approach involves creating a single convolutional autoencoder that uses the data from each Subject, split into their respective Conditions. The entire dataset with ~1.5 million vectors will be used for model training and testing, the data split at a ratio of 70:30 respectively, to promote a generalised convolutional autoencoder design. The test data will be divided into three datasets, each associated with the appropriate Condition. The signal-to-noise ratio will be calculated for the test data before and after processing and recorded. The Harrell-Davis (HD) decile values will be calculated for the SNR distributions of raw EEG signals (RAW), PCA reconstructed signals (PCA) and CAE reconstructed signals (CAE). The HD decile differences between the RAW and PCA SNR distributions, as well as the RAW and CAE SNR distributions, will be calculated.

The RAW-PCA HD decile differences will be subtracted from the respective RAW-CAE HD decile differences, resulting in 9 values.

- The mean of these values will be calculated. A positive calculated value will indicate that the CAE reconstructed signals have higher SNR values than the

PCA reconstructed signals, signifying that the CAE performed better than PCA.

- If 5 of the 9 values are positive, this will also indicate that the CAE reconstructed signals have higher SNR values than the PCA reconstructed signals, signifying that the CAE performed better than PCA.

Multiple CAE: The second approach involves designing an individual convolutional autoencoder for each of the Subjects. This results in 81 independent convolutional autoencoders with ~18,000 vectors, split into training and test datasets, with a ratio of 70:30 respectively. The SNR will be calculated for each vector of the RAW signal, PCA reconstructed signal and CAE reconstructed signal of the test dataset. The HD decile values of the SNR distributions of the RAW data will be calculated and recorded, as well as the HD decile values of the SNR distributions of the PCA and CAE reconstructed signals. This will be performed for each Subject for all three Conditions. The difference in decile values will be calculated for the RAW-PCA distributions and RAW-CAE distributions. The mean of each HD decile difference will be calculated for both distribution differences, for each Subject.

- A greater mean decile difference for the RAW-CAE distributions than the RAW-PCA distributions for a particular Subject indicates that the CAE produces a signal with a higher SNR than PCA. If 41 of the 81 Subjects exhibit higher SNRs for the CAE than PCA according to this criterion, this will signify that the CAE performed better than PCA.

3.5.1 Strengths

Computationally efficient

- Convolutional Neural Networks can be more computationally efficient than other deep neural network types. The complexity of the model is reduced by using weight sharing which can result in a more computationally efficient model (Neubauer, 1998). This can allow for more time focussed on hyperparameter assessment and tuning, enabling more accurate hyperparameter value selection.

Simple hyperparameter selection.

- Simple hyperparameter selection allows better interpretation of the functionality of the model, as well as reducing time needed for hyperparameter selection.

Reproducibility.

- The use of 1-dimensional model inputs provides a general model that can be used with EEG signals recorded by any number of electrodes e.g. 64, 32. This can be particularly effective if only a single area of the brain is to be analysed using a small number of electrodes.

Specific electrode isolation.

- The labelling of the channels enables analysis and performance evaluation of specific electrodes as well as the areas of the brain with which each electrode is associated.

More accurate distribution comparison.

- The use of Harrell-Davis decile values facilitates more accurate comparison of distributions when compared using a single value such as the mean of median.

Generalised and specified analysis.

- The methodology executed in the experiment enables the analysis of a generalised model designed for any individual, as well as models designed for specific individuals. This allows for a broader analysis of the performance of CAEs, comparing generalised models to individualised models**.**

Result interpretability.

- Through visualisation of the SNR distributions, any substantial differences between the kurtosis and skewness of the distributions can be immediately identified. In conjunction with this, the inclusion of HD decile values allows for a more well-rounded analysis and comparison of the SNR distributions.

3.5.2 Limitations

Unknown information loss.

- Unlike Principal Component Analysis, for which the exact amount of explained variance can be decided e.g. 95%, it is not possible to predetermine how much information will be kept using CAEs for EEG signal denoising.

Longer training times.

- Due to the complex nature of Convolutional Autoencoders, although quicker than other Deep Learning techniques, when compared to more classical noise reduction methods, training times and running times are significantly longer than times needed for PCA analysis. As a number of hyperparameter have to be chosen for CAEs, a significant amount of time may also be taken for selecting appropriate values for each hyperparameters.

Less hyperparameters for optimisation.

- Although less hyperparameters results in less time needed for selection, less hyperparameters also yields less opportunities for model modification to acquire better results.

Input constraints.

- The use of a 1-dimensional inputs removes channel dependence, as each channel is treated as an equal input vector for the CAE. However, using a 2-dimensional input that consists of all channel readings from a single trial. This allows for channel specific analysis of the EEG signals, allowing for more accurate evaluations.

Evaluation metric limitations.

- The use of p-value statistics has been highly disputed when analysis brain activity. Because of this, it is difficult to determine whether certain obtained model results are significant, as according to the *p-value* calculated in the analysis signifies no statistical significances (Szucs & Ioannidis, 2017; Calin-Jageman, 2017; Ioannidis, 2005; Button, et al., 2013).

No baseline signals.

- As it is not possible to record a perfectly "clean" EEG signal from a subject due to the complex nature of the brain, the analysis of reconstructed EEG signals cannot be compared to noiseless baseline data. Because of this, it is not possible to definitively determine whether the reconstructed signal has only removed noise, without removing critical signal information. If the RAW signal is denoised too aggressively, it may be a case that the SNR of the signal increases significantly, but too much important signal information is lost, negating the benefit of the high SNR.
- Many studies that analyse the performance EEG noise reduction models use synthetically/artificially created EEG signals with added noise. This allows for an exact analysis of what information is retained or removed from the noisy signal, enabling a more compelling comparison of signals and noise reduction techniques (Hassani & Karami, 2015; Maddirala & Shaik, 2016; Pander, 2019).

Imperfect comparison metrics.

- Although the use of Harrell-Davis decile differences broadens the extent at which the SNR distributions can be compared, using the mean of these values for the comparison of the CAE and PCA performance related to each individual participant could result in decile difference outliers affecting final evaluation of the results.

Linear autoencoder acts similar to PCA.

- An autoencoder that uses a *linear* activation function performs similar dimensionality reduction as PCA, as described by Baldi and Hornik (1989). Although this reduces the extent at which the potential of the autoencoder can be exploited, it does allow for comparison of similar dimensionality reduction methods using different approaches.

Experiment, Design & Methodology: Strengths & Limitations		
Strengths	Experiment	Computationally efficient
		Simple hyperparameter selection
	Design	Reproducibility
		Specific electrode isolation
	Methodology	More accurate distribution comparison
		General and specified analysis
		Result interpretability
Limitations	Experiment	Unknown information loss
	Design	Longer training times
		Less hyperparameters for optimisation
		Input constraints
	Methodology	Evaluation metric limitations
		No baseline signal
		Imperfect comparison metrics
		Linear autoencoder acts similar to PCA

Table 3.5.2-1 Strengths and Limitations for Experiment, Design and Methodology

4 RESULTS, EVALUATION & DISCUSSION

This chapter discusses the implementation of the experimental process, the results obtained from the experiment, and the evaluation of these results when compared to the baseline data. As well as this, the strengths and weaknesses of the conducted experiment are described, along with potential causes of error in the experiment, and stages at which improvements and alterations could have been made during the experimental process.

4.1 DATA EVALUATION

The data used in this experiment consisted of 81 participants, each of whom were required to perform 3 separate button pressing tasks, referred to as Condition 1, Condition 2 and Condition 3 respectively, with each task comprising of up to 100 trials. The cognitive activity of each participant's brain was recorded using 64 electrodes placed on the scalp of each participant using the *10-20 system* electrode placement method. Each trial was recorded over 700ms. The voltage amplitude for each electrode was measured at every millisecond. The onset of the task/stimulus was introduced at 100ms. The reading from each electrode over the 700ms was used as a single vector for length 700.

Single CAE: The entire dataset was concatenated to create a 2-dimensional dataframe, with each column of the dataframe a vector of length 700. The resultant dataframe was of the shape 700 x ~1.5 million. The columns were indexed using the appropriate electrode name of the *10-20 system*. 70% of the columns were chosen randomly to be used as the training dataset for the single CAE with the remaining 30% to be used as the test dataset.
Each column of the test dataset was separated into a new test dataset related to its respective Condition. Each test dataset consisted of ~150,000 vectors of length 700.

Multiple CAE: For the multiple CAEs, Subject 46 was entirely removed from the dataset due to the absence of data for Condition 3. This was to ensure equal readings for all three Conditions. A different convolutional autoencoder (CAE) was designed for each subject with the data of each subject isolated entirely from the other subjects. For each CAE, the data from each subject was concatenated such that the resultant dataframe for each subject was of the shape 700 x ~18,000. As performed with the Single CAE, the data was separated randomly 70%-30% for training and test datasets respectively. The training dataset was used to train the individual CAE of each Subject, and each column of the test dataset separated into a new test dataset related to its appropriate Condition. For each of the 80 CAEs, there were individual training datasets, each of shape 700 x ~12,500, as well as three separate test datasets, each of shape 700 x ~1,800.

The Single CAE and Multiple CAEs were trained as described, using data from all three Conditions combined, whereas each model was separately tested using the data from each Condition separately. This approach allowed the individual examination of each electrode related to the areas of the brain associated with each stimulus type.

Training a Single CAE allowed for a generalised model to be designed, applicable to any individual. Using one CAE for each Subject then allowed for more specialised analysis but only applicable to a single individual.

4.2 CAE MODEL SELECTION

The model that was shown to have the best performance is described in this section.

4.2.1 Hyperparameter Selection

The final hyperparameters used in the are listed in **Table 4.2.1-1**.. This selection of hyperparameters were shown to provide the best performance for EEG noise reduction, with the most accurate signal reconstruction.

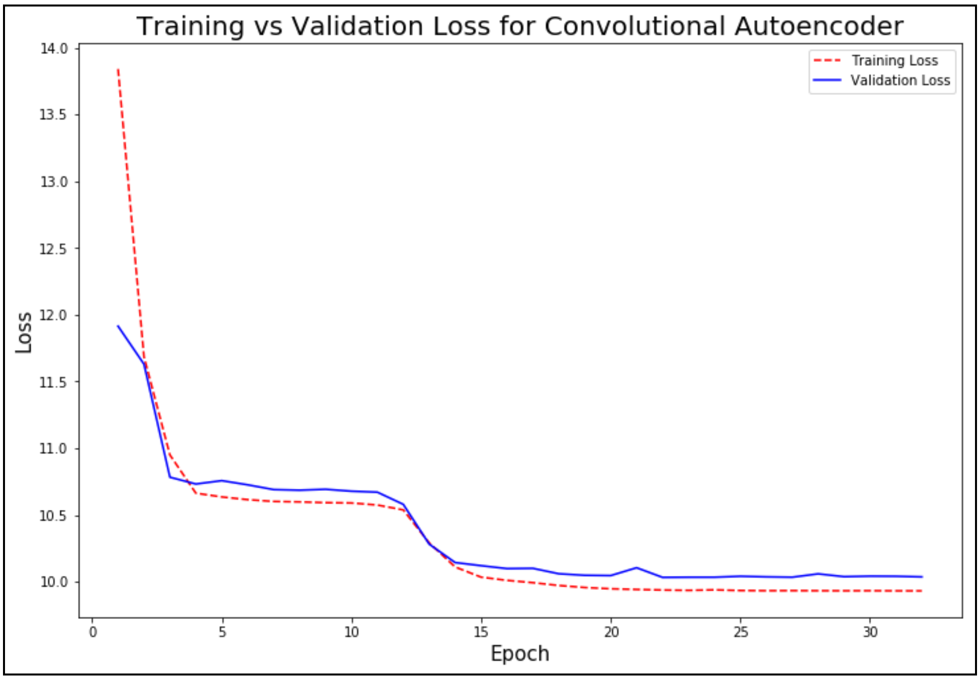


Figure 4.2.1-1 Training vs Validation Loss Curve for epoch selection

Final Model Selection	
Number of hidden convolution layers	2
Kernel size	7,3
Filter size	25, 5
Pooling strategy	AveragePooling
Pooling size	7, 5
Activation function	Linear
Number of epochs	14
Batch size	512
Optimizer	ADAM
Loss function	Mean Squared Error
Training/Test split	70/30

Table 4.2.1-1 Final model hyperparameters

4.2.2 Final Model Accuracy, Reliability, Efficiency

The accuracy of the reconstructed signals can be seen in the figures below. An example from each Condition is shown with the reconstructed signals overlaid on the RAW signals. Accurate signal representation was found for all signals examined.

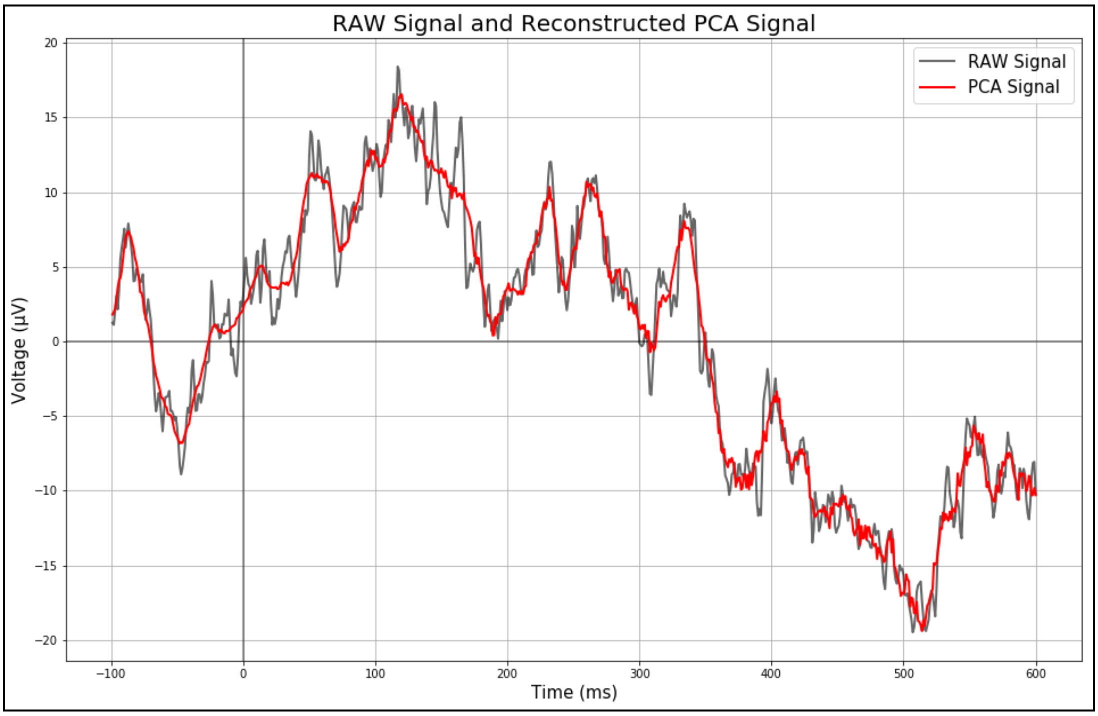


Figure 4.2.2-1 RAW signal vs PCA reconstructed signal – Condition 1

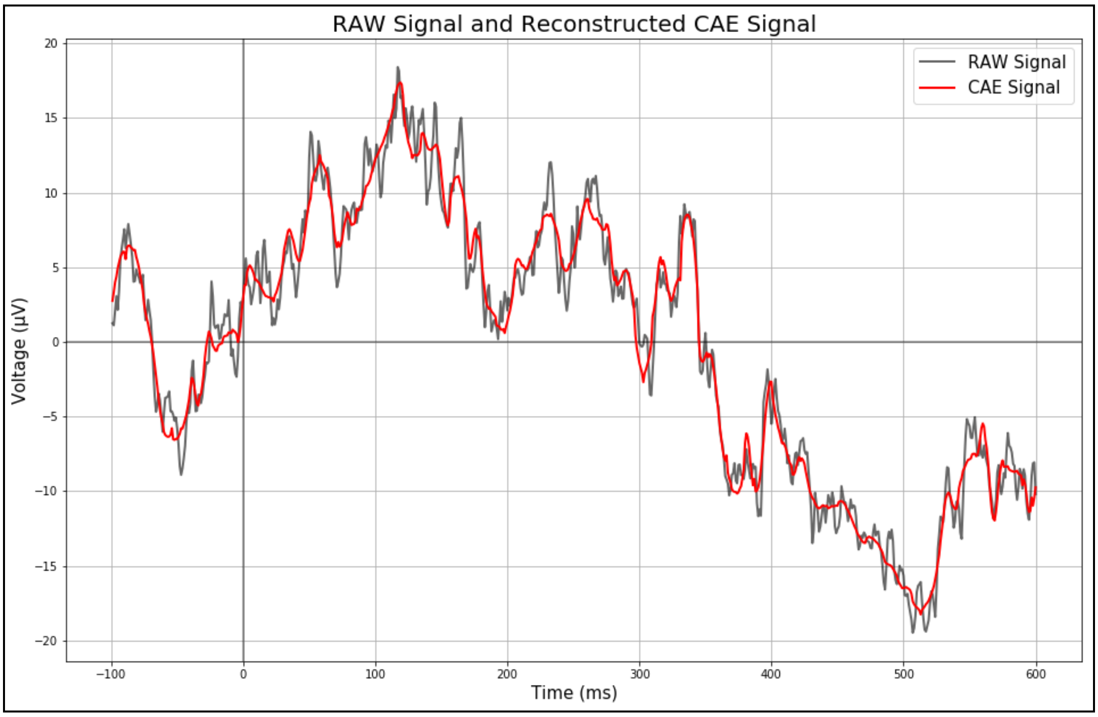


Figure 4.2.2-2 RAW signal vs CAE reconstructed signal – Condition 1

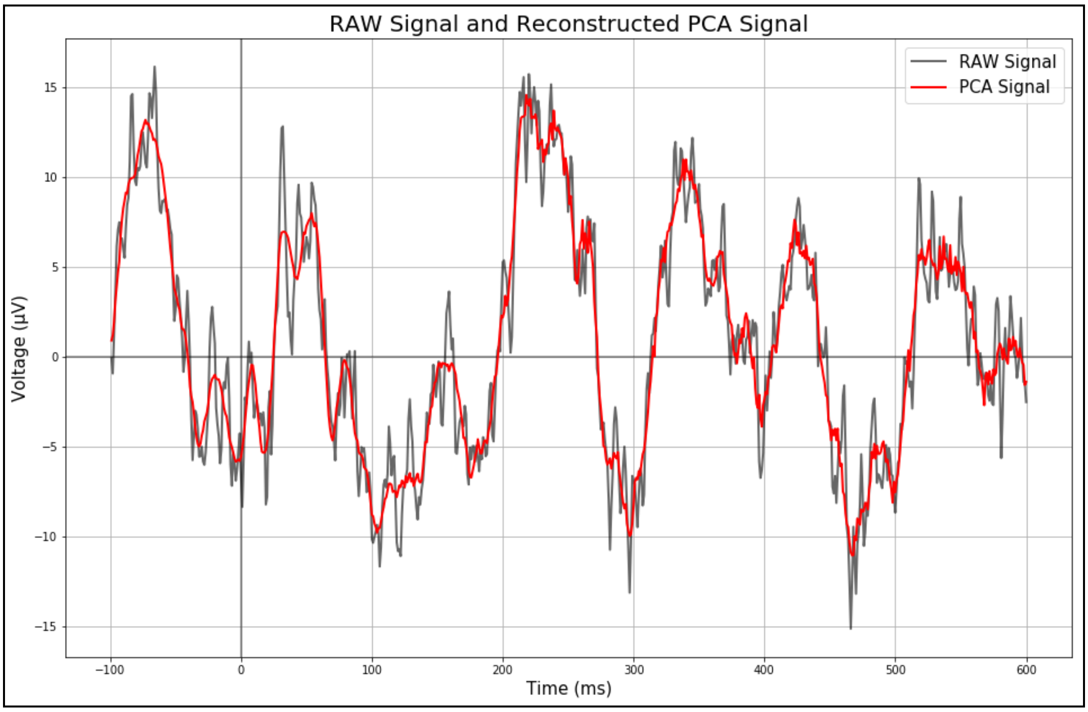


Figure 4.2.2-3 RAW signal vs PCA reconstructed signal – Condition 2

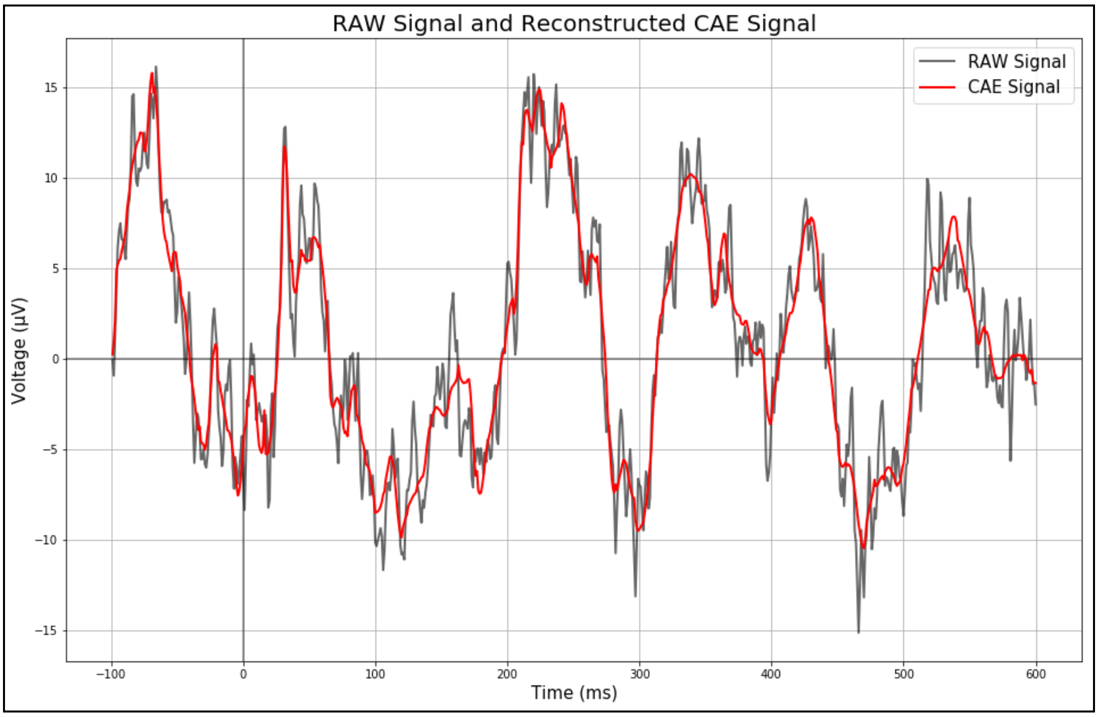


Figure 4.2.2-4 RAW signal vs CAE reconstructed signal – Condition 2

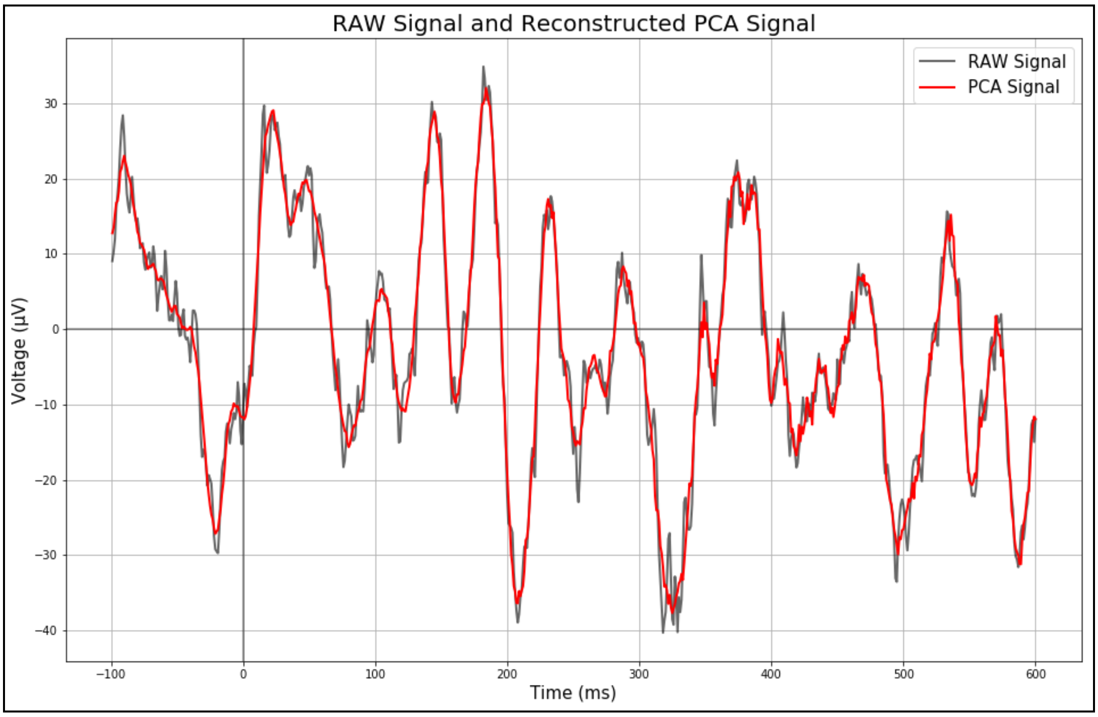


Figure 4.2.2-5 RAW signal vs PCA reconstructed signal – Condition 3

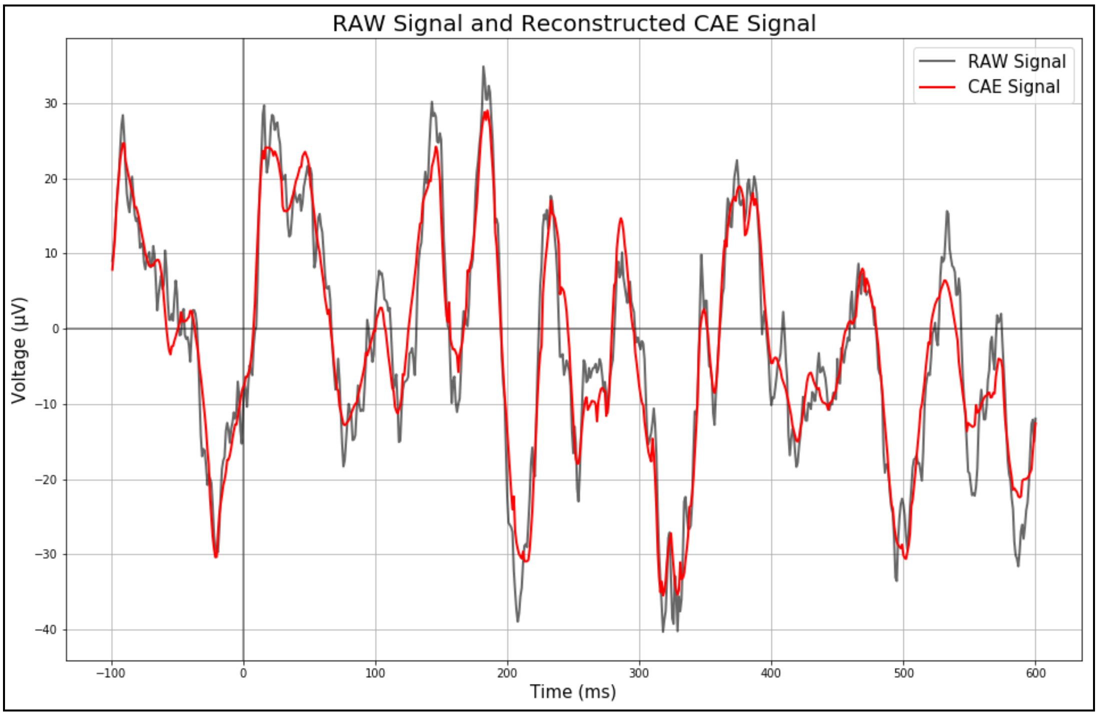


Figure 4.2.2-6 RAW signal vs CAE reconstructed signal – Condition 3

4.3 EXPERIMENT RESULTS

The results obtained from the experiment can be found below. The results are separated into two sections: *Single Stacked Convolutional Autoencoder* and *Multiple Stacked Convolutional Autoencoder*, with each section divided according to each Condition.

Single Stacked Convolutional Autoencoder

- For the Single CAE, the Harrell-Davis decile values of the signal-to-noise ratio distribution for the RAW signals, PCA reconstructed signals, and CAE reconstructed signals were calculated for all three Conditions. The decile differences between the RAW SNR distribution and PCA SNR distribution, as well as the decile differences between the RAW SNR distribution and CAE SNR distribution was recorded and compared.
- The mean SNR for each specific scalp location on the *10-20 system* was calculated for RAW signal, PCA reconstructed signal and CAE reconstructed signal. The mean SNR for each electrode is plotted as on a *10-20 system*, where the individual differences between respective electrodes for RAW, PCA and CAE signals can be seen.

Multiple Stacked Convolutional Autoencoders

- For the Multiple CAEs, the mean difference between the RAW and PCA SNR distributions, and mean difference between RAW and CAE SNR distributions for each subject were calculated. For each subject, the values are compared. The mean Harrell-Davis decile values of the respective RAW SNR distributions for each subject are also calculated, and the mean values related to the PCA reconstructed signal and CAE reconstructed signal are compared.

4.3.1 Single Stacked Convolutional Autoencoder

4.3.1.1 Condition 1

The results for the performance of the single Autoencoder for Condition 1 can be seen in Figure 4.3.1-1.

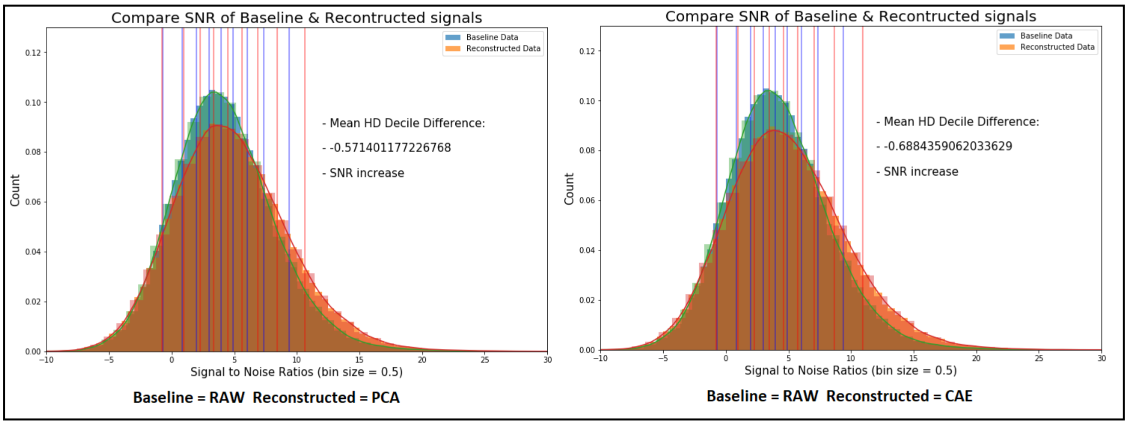


Figure 4.3.1-1 SNR distributions: RAW vs PCA (left), RAW vs CAE (right) - Condition 1

- The mean SNR Harrell Davis decile difference between the RAW signal and PCA reconstructed signals for Condition 1 was calculated to be 0.5714.
- The mean SNR Harrell Davis decile difference between the RAW signal and CAE reconstructed signal SNRs for Condition 1 was calculated to be 0.6884.

The SNR differences for each decile for Condition 1 can be seen in Table 4.3.1-1.

Decile	PCA SNR	CAE SNR	CAE - PCA	CAE - PCA /PCA	% Difference
1	-0.03142476882	-0.07738682117	-0.04596205	-1.462605902	**-146**
2	0.1201771205	0.1288066083	0.008629488	0.07180641196	**+7.2**
3	0.274190522	0.3119674107	0.037776889	0.137776056	**+13.8**
4	0.4025163426	0.4808491347	0.078332792	0.1946077308	**+19.5**
5	0.5473342425	0.6580410602	0.110706818	0.2022654699	**+20.2**
6	0.6874730811	0.8339223299	0.146449249	0.2130254301	**+21.3**
7	0.8494426565	1.039687813	0.190245157	0.2239646848	**+22.4**
8	1.035095757	1.263891817	0.22879606	0.2210385446	**+22.1**
9	1.257805641	1.556143803	0.298338162	0.2371893969	**+23.7**
Mean	**0.5714011772**	**0.6884359062**	**0.117034729**	**0.204820584**	**+20.5**

Table 4.3.1-1 Signal-to-noise ratio Harrell-Davis decile differences - Condition 1

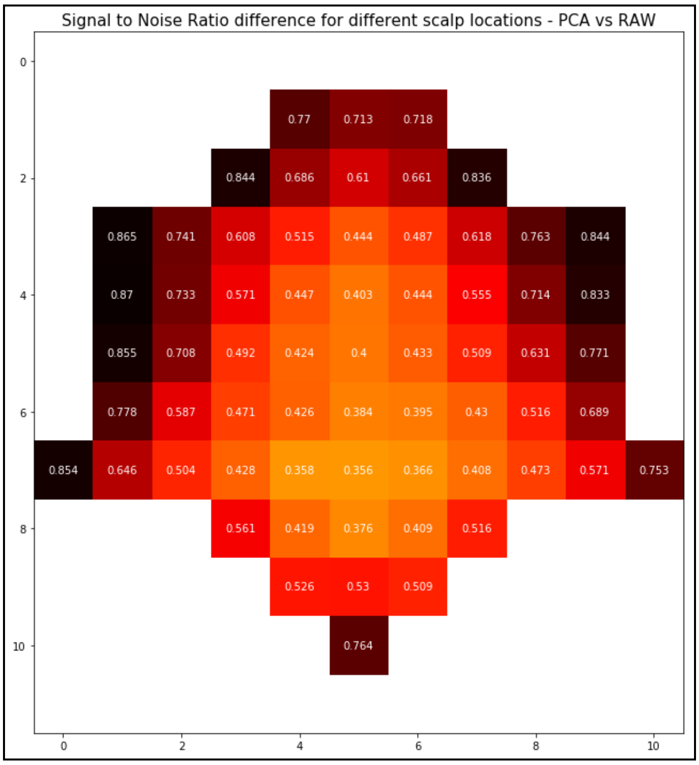


Figure 4.3.1-2 Differences in mean SNR values for electrode location on 10-20 system - PCA vs RAW Condition 1

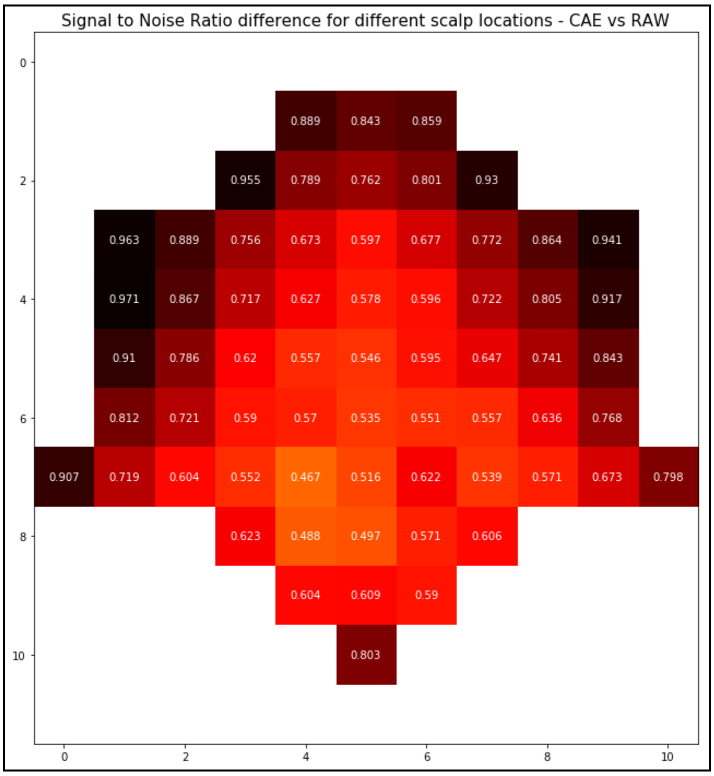


Figure 4.3.1-3 Differences in mean SNR values for electrode location on 10-20 system - CAE vs RAW Condition 1

4.3.1.2 Condition 2

The results for the performance of the single Autoencoder for Condition 2 can be seen in Figure 4.3.1-4.

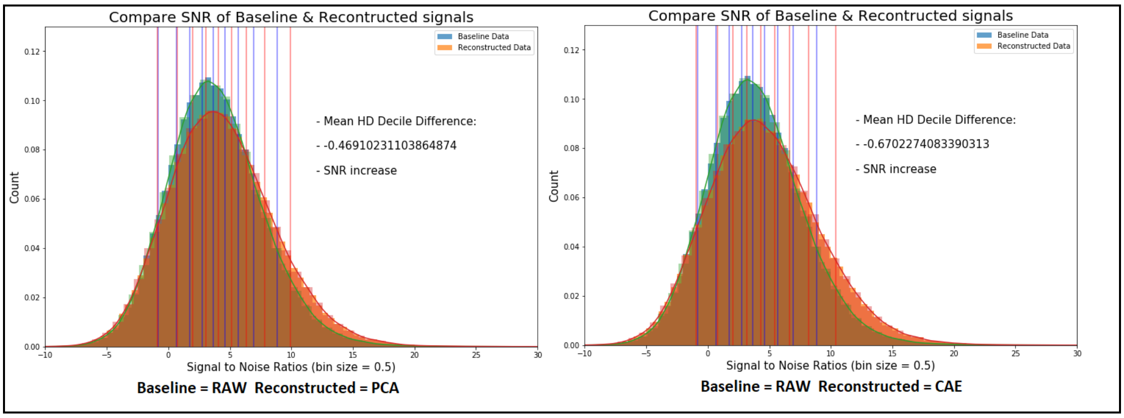


Figure 4.3.1-4 SNR distributions: RAW vs PCA (left), RAW vs CAE (right) - Condition 2

- The mean SNR Harrell Davis decile difference between the RAW signal and PCA reconstructed signals for Condition 2 was calculated to be 0.4691.
- The mean SNR Harrell Davis decile difference between the RAW signal and CAE reconstructed signals for Condition 2 was calculated to be 0.6702.

The SNR differences for each decile for Condition 2 can be seen in Table 4.3.1-2.

Decile	PCA SNR	CAE SNR	CAE - PCA	CAE – PCA /PCA	% Difference
1	-0.05042659403	-0.094220807	-0.04379421	-0.8684745374	-86.8
2	0.08107328304	0.113987496	0.032914213	0.4059810052	+40.6
3	0.2009457197	0.2952191585	0.094273439	0.4691487775	+46.9
4	0.3188029607	0.4691065776	0.150303617	0.4714624248	+47.1
5	0.440214186	0.6301827672	0.189968581	0.4315367091	+43.2
6	0.5616059748	0.8150277499	0.253421775	0.4512447987	+45.1
7	0.7017509903	1.008400572	0.306649582	0.4369777688	+43.7
8	0.8746400409	1.241980364	0.367340323	0.4199902883	+42.0
9	1.093314238	1.552362797	0.459048559	0.41986882	+42.0
Mean	**0.469102311**	**0.6702274083**	**0.201125097**	**0.428744631**	**+42.9**

Table 4.3.1-2 Signal-to-noise ratio Harrell-Davis decile differences - Condition 2

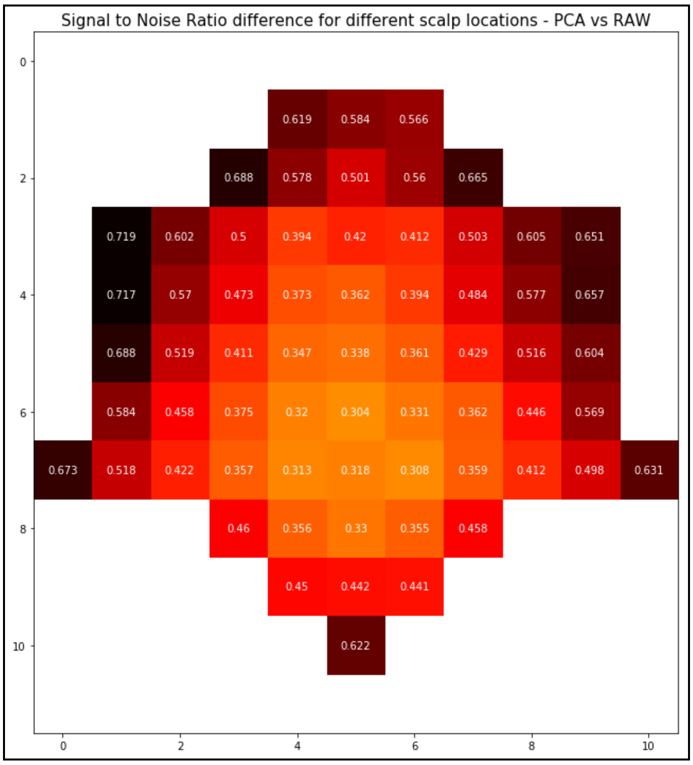


Figure 4.3.1-5 Differences in mean SNR values for electrode location on 10-20 system - PCA vs RAW Condition 2

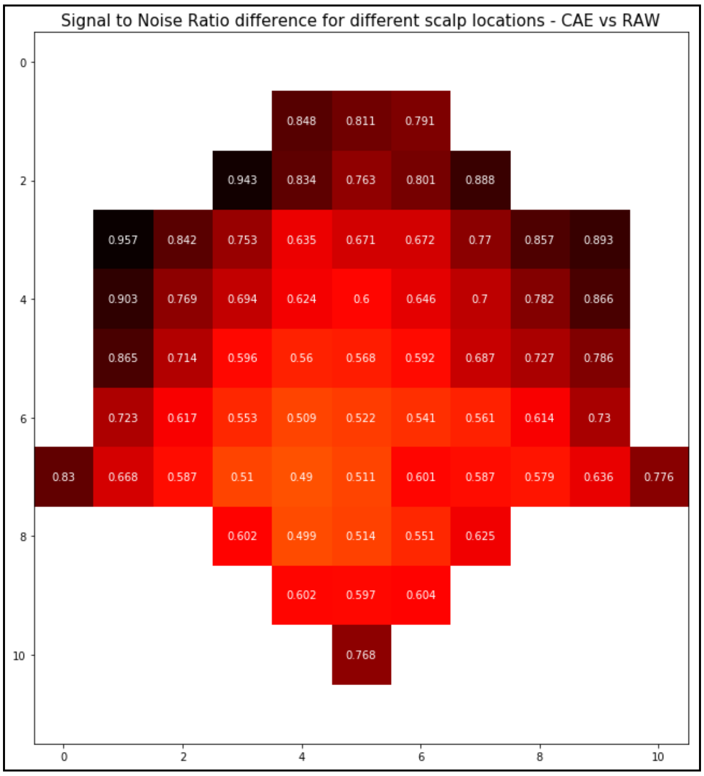


Figure 4.3.1-6 Differences in mean SNR values for electrode location on 10-20 system - CAE vs RAW Condition 2

4.3.1.3 Condition 3

The results for the performance of the single Autoencoder for Condition 3 can be seen in Figure 4.3.1-7.

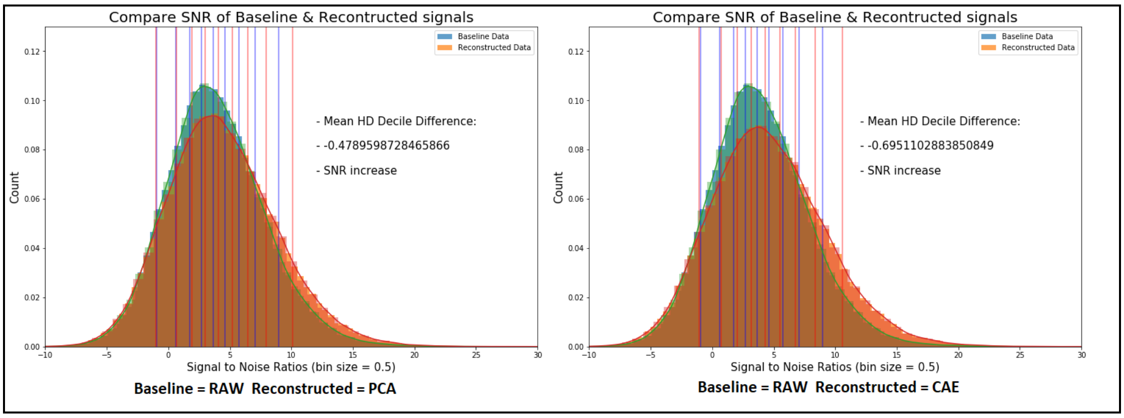


Figure 4.3.1-7 SNR distribution: RAW vs PCA (left), RAW vs CAE (right) - Condition 3

- The mean SNR Harrell Davis decile difference between the RAW signal and PCA reconstructed signals for Condition 3 was calculated to be 0.4790.
- The mean SNR Harrell Davis decile difference between the RAW signal and CAE reconstructed signals for Condition 3 was calculated to be 0.6951.

The SNR differences for each decile for Condition 3 can be seen in

Decile	PCA SNR	CAE SNR	CAE - PCA	CAE – PCA /PCA	% Difference
1	-0.049527944	-0.118527188	-0.06899924	-1.393137676	-139.3
2	0.081396967	0.103816557	0.02241959	0.275435202	27.5
3	0.199245152	0.297605451	0.098360299	0.493664705	49.4
4	0.325068102	0.494535713	0.16946761	0.521329559	52.1
5	0.447053659	0.665415441	0.218361782	0.488446471	48.8
6	0.57171208	0.853070271	0.281358191	0.492132668	49.2
7	0.713745555	1.057585446	0.343839891	0.48174015	48.2
8	0.89637632	1.306574539	0.410198219	0.457618312	45.8
9	1.125568964	1.595916366	0.470347402	0.417875241	41.8
Mean	**0.478959873**	**0.695110288**	**0.216150415**	**0.451291283**	**45.1**

Table 4.3.1-3 Signal-to-noise ratio Harrell-Davis decile differences - Condition 3

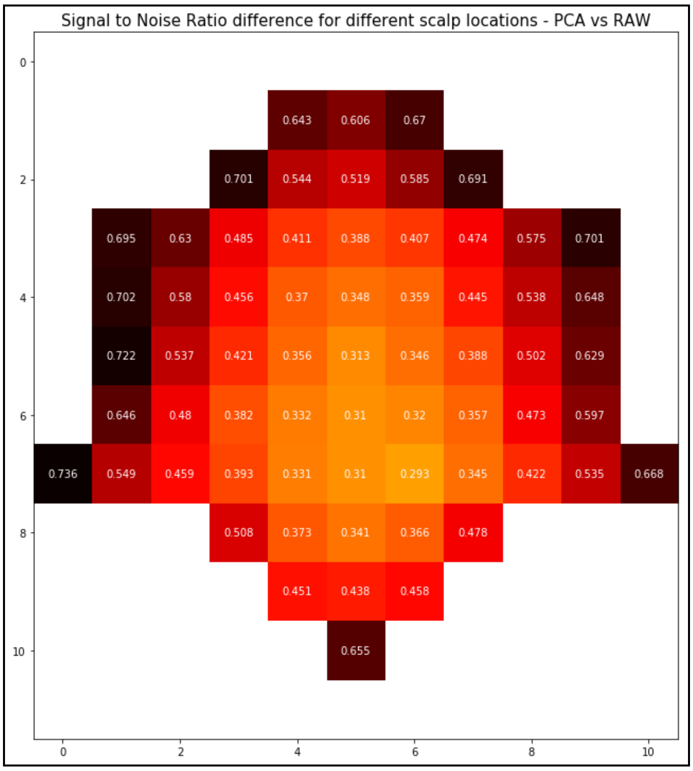


Figure 4.3.1-8 Differences in mean SNR values for electrode location on 10-20 system - PCA vs RAW Condition 3

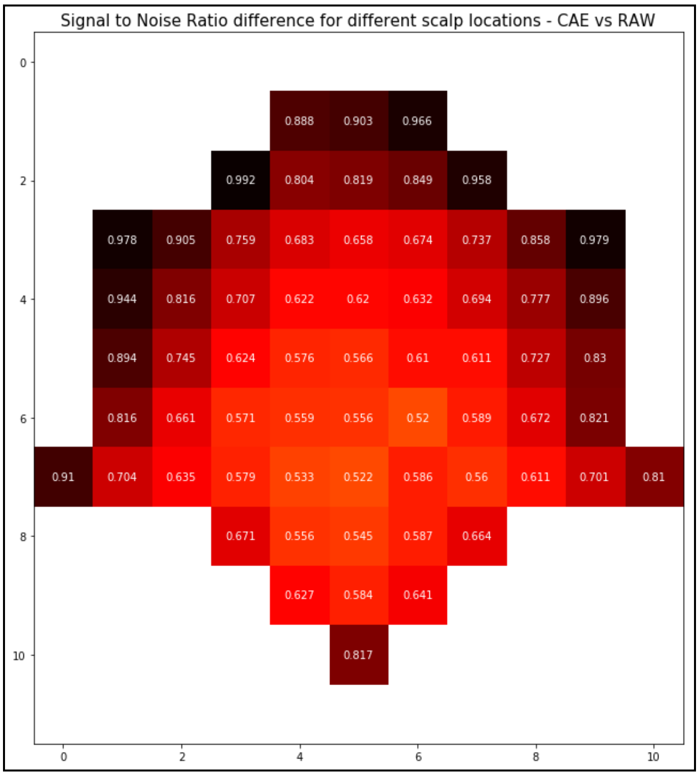


Figure 4.3.1-9 Differences in mean SNR values for electrode location on 10-20 system - CAE vs RAW Condition 3

4.3.1.4 Results Analysis

As can be seen in all three Harrell-Davis decile difference tables, the first decile shows worse performance for the Stacked Convolutional Autoencoder when compared to Principal Component Analysis. As well as this, it can be seen that both PCA and the CAE have a lower first Harrell-Davis decile value than was calculated for the RAW signal. This could be due to a significant amount of noise present in the electrodes which are not associated with the task being performed by the Subject. Because of this, by reducing the apparent “noise” present in the signal for these electrodes, the already low voltage amplitude recorded is reduced to a value lower than the recorded signal prior to the onset of the stimulus. When the signal-to-noise ratio is then calculated for these electrodes, a negative value is obtained, indicating more noise than signal present.

The plots displaying the differences in mean SNR values for electrode location on the 10-20 system reveal the difference in mean signal-to-noise ratios calculated for each electrode located on the scalp. It can be seen that the cortexes more closely associated with audio and motor cognitive activity undergo a greater increase in signal-to-noise ratios than the cortexes not associated with these particular processes.

4.3.2 Multiple Stacked Convolutional Autoencoder

4.3.2.1 Condition 1

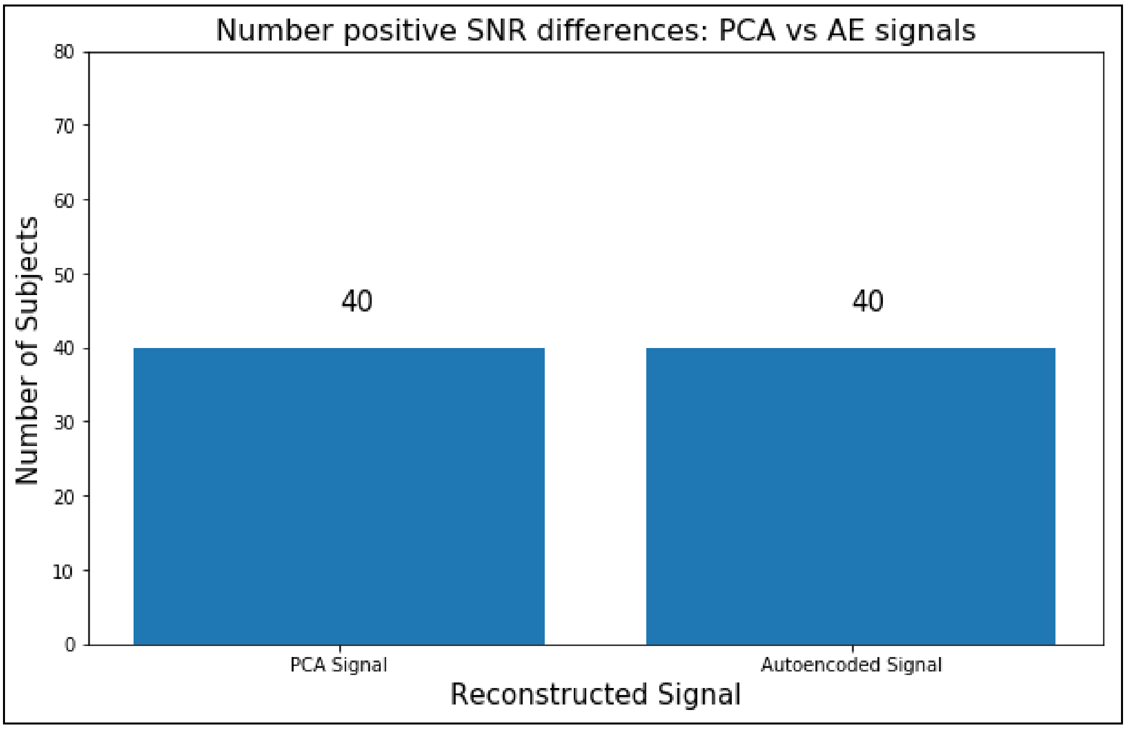


Figure 4.3.2-1 Higher mean SNR count: PCA vs CAE - Condition 1

Decile	RAW vs PCA	RAW vs CAE	CAE- PCA
1	0.0555	-0.0993	-0.1548
2	0.2528	0.1122	-0.1406
3	0.4062	0.2815	-0.1247
4	0.5593	0.4356	-0.1237
5	0.7175	0.5929	-0.1246
6	0.8800	0.7513	-0.1287
7	1.0313	0.9228	-0.1085
8	1.2138	1.1148	-0.099
9	1.4390	1.3711	-0.0679
Mean	**0.7284**	**0.6092**	**-0.1192**

Table 4.3.2-1 Mean Harrell-Davis decile values of SNR distributions for all subjects - Condition 1

4.3.2.2 Condition 2

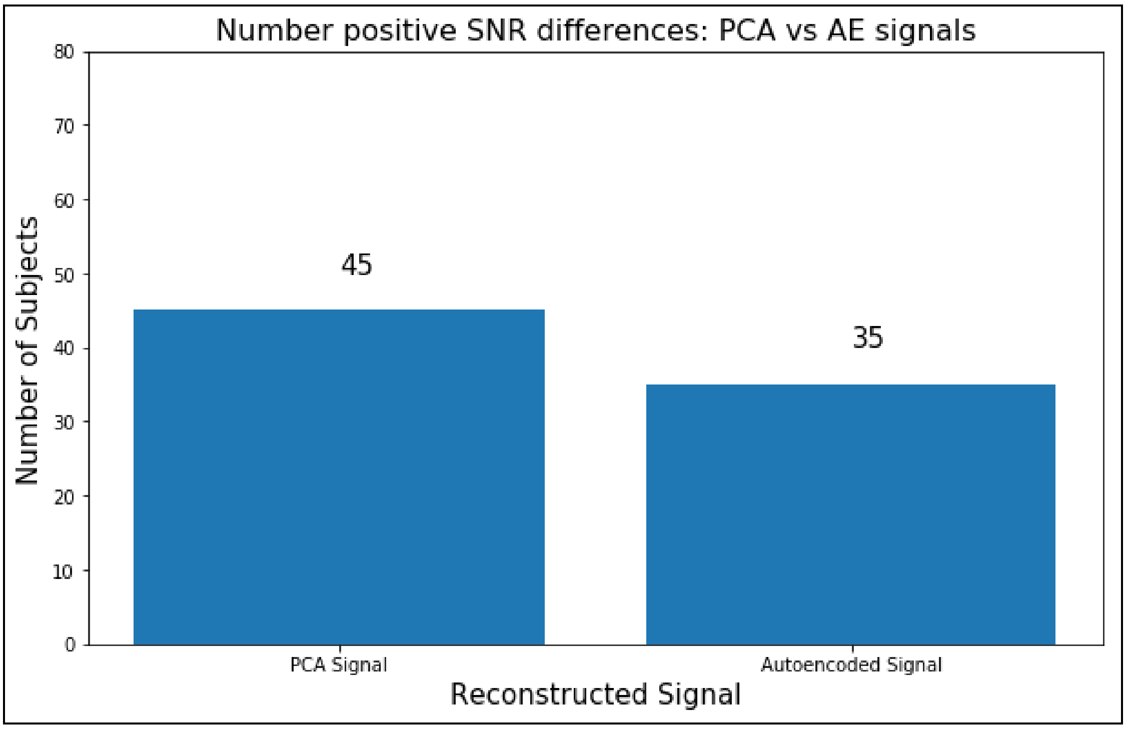


Figure 4.3.2-2 Higher SNR count: PCA vs CAE - Condition 2

Decile	RAW vs PCA	RAW vs CAE	CAE - PCA
1	0.0002	-0.1010	-0.1012
2	0.1632	0.1084	-0.0548
3	0.2905	0.267	-0.0235
4	0.4091	0.4307	0.0216
5	0.5408	0.5885	0.0477
6	0.6753	0.7516	0.0763
7	0.8283	0.9082	0.0799
8	1.0035	1.0912	0.0877
9	1.2549	1.3777	0.1228
Mean	**0.5740**	**0.6025**	**0.0285**

Table 4.3.2-2 Mean Harrell-Davis decile values of SNR distributions for all subjects - Condition 2

4.3.2.3 Condition 3

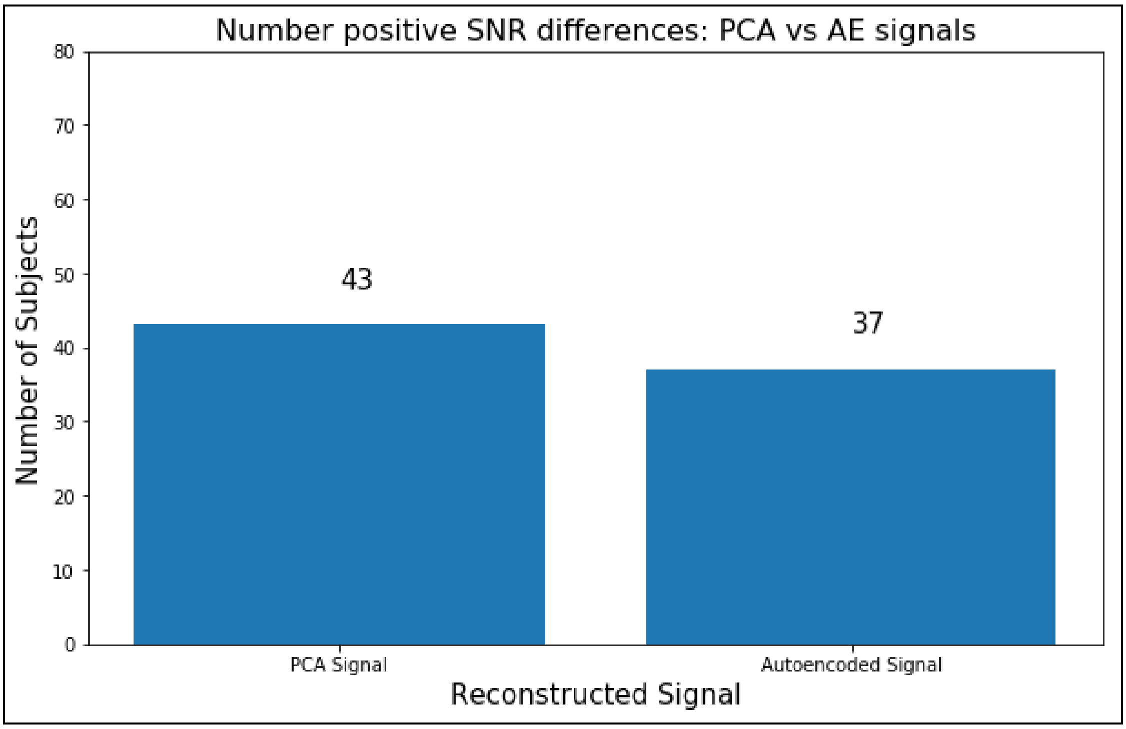


Figure 4.3.2-3 Higher SNR count: PCA vs CAE - Condition 3

Decile	RAW vs PCA	RAW vs CAE	CAE - PCA
1	-0.0261	-0.1256	-0.0995
2	0.1367	0.0854	-0.0513
3	0.2818	0.2660	-0.0158
4	0.4197	0.4435	0.0238
5	0.5508	0.6067	0.0559
6	0.6948	0.7638	0.069
7	0.8421	0.9362	0.0941
8	1.0373	1.1331	0.0958
9	1.274	1.4407	0.1667
Mean	**0.5790**	**0.6166**	**0.0376**

Table 4.3.2-3 Mean Harrell-Davis decile values of SNR distributions for all subjects - Condition 3

4.3.2.4 Results Analysis

The Stacked Convolutional Autoencoder was found to perform better than Principal Component Analysis in not one of the three Conditions. The metric used to compare the techniques was the “mean”. The mean of the Harrell-Davis decile values was used in order to avoid influence caused by extreme outliers, but still acquire an overall effect from the majority of the distribution. However, it can be seen from the associated Harrell-Davis decile difference tables, which calculated the mean value for each decile value distribution, that, using majority voting, CAE performed better than PCA for two of the three tasks performed by the Subjects.

Further investigation is needed into the most appropriate metrics to use when comparing distribution differences in order to perform more accurate analyses.

4.4 EVALUATION OF RESULTS

To evaluate the performance of the Stacked Convolutional Autoencoder in noise reduction of EEG signals, two separate tests were conducted.

For the Single Stacked Convolutional Autoencoder the entire dataset was analysed. The distribution of signal-to-noise ratios for the PCA reconstructed signals and CAE reconstructed signals were plotted and compared to that of the raw EEG (RAW) signals. The Harrell-Davis decile values were calculated for each distribution and the difference in Harrell-Davis decile values of the SNR distributions between the PCA and RAW signal SNR distributions were compared to the difference in Harrell-Davis decile values of the SNR distributions between the CAE and RAW signal SNR distributions. This result in a comparison for each of the 9 deciles. Using majority voting, if 5 of the 9 decile values are greater for the Stacked Convolutional Autoencoder, there would be an indication of evidence to support rejecting the null hypothesis and accepting the alternate hypothesis.

For the Multiple Stacked Convolutional Autoencoders, each Subject was tested independently. The distributions of the RAW, PCA and CAE signals were analysed by calculating the Harrell-Davis decile values for each distribution, and the mean value of the Harrell-Davis decile values was calculated for each distribution. The difference between the calculated value for the RAW signal and calculated values for the PCA signals was compared to the difference between the calculated value for the RAW signal and calculated values for the CAE signals. This process was performed for each Subject. If the CAE was found to perform better than PCA for the majority (41 of 80) of the Subjects analysed, there would be an indication of evidence to support rejecting the null hypothesis and accepting the alternate hypothesis.

4.4.1 Single Stacked Convolutional Autoencoder

4.4.1.1 Condition 1

- Mean SNR difference - RAW vs PCA: 0.5714.
- Mean SNR difference - RAW vs CAE: 0.6884.
- Improvement of 20.4%
- 8 of 9 decile difference values show positive difference for CAE over PCA.
- There is evidence to support rejecting the null hypothesis and accepting the alternate hypothesis.

4.4.1.2 Condition 2

- Mean SNR difference - RAW vs PCA: 0.4691.
- Mean SNR difference - RAW vs CAE: 0.6702.
- Improvement of 42.9%
- 8 of 9 decile difference values show positive difference for CAE over PCA.
- There is evidence to support rejecting the null hypothesis and accepting the alternate hypothesis.

4.4.1.3 Condition 3

- Mean SNR difference - RAW vs PCA: 0.4790.
- Mean SNR difference - RAW vs CAE: 0.6951.
- Improvement of 45.1%
- 8 of 9 decile difference values show positive difference for CAE over PCA.
- There is evidence to support rejecting the null hypothesis and accepting the alternate hypothesis.

SNR was shown to increase in every electrode location on the scalp from RAW signal to PCA and CAE reconstructed signals, with the CAE consistently outperforming PCA across all electrodes for all three Conditions.

4.4.2 Multiple Stacked Convolutional Autoencoder

4.4.2.1 Condition 1

From the 80 subjects analysed:

- Number of subjects with higher mean SNR values using PCA: 40
- Number of subjects with higher mean SNR values using CAE: 40

Distribution of Harrell-Davis decile differences:

- 0 from 9 mean decile values were found to be greater for CAE than PCA signals.

There is not enough significant evidence to support rejecting the null hypothesis.

4.4.2.2 Condition 2

From the 80 subjects analysed:

- Number of subjects with higher mean SNR values using PCA: 45
- Number of subjects with higher mean SNR values using CAE: 35

Distribution of Harrell-Davis decile differences:

- 6 from 9 mean decile values were found to be greater for CAE than PCA signals.

There is not enough significant evidence to support rejecting the null hypothesis.

4.4.2.3 Condition 3

From the 80 subjects analysed:

- Number of subjects with higher mean SNR values using PCA: 43
- Number of subjects with higher mean SNR values using CAE: 37

Distribution of Harrell-Davis decile differences:

- 6 from 9 mean decile values were found to be greater for CAE than PCA signals.

There is not enough significant evidence to support rejecting the null hypothesis.

4.5 DISCUSSION

This section investigates the strengths and weaknesses/limitations of the proposed model as well as potential causes of error for the approach used and improvements that could have been made during the processes.

4.5.1 Strengths

Better performance than PCA.

- Convolutional autoencoders we found to perform better than Principal Component Analysis (95% explained variance) for noise reduction in EEG signals, increasing the SNR.

Computationally efficient.

- The use of CAEs was more computationally efficient than other neural network approaches. As there are significantly fewer memory requirements when running convolutional neural networks, training time for the CNNs were significantly shorter than more dense networks e.g. recurrent neural networks. This reduced the training time of the entire model significantly, which enabled more hyperparameter assessment, allowing for more accurate hyperparameter value selection.

Intuitive and easier to visualise.

- Conceptually, convolutional neural networks are easier to understand, as there are fewer connections within the neural network. This allowed for better hyperparameter selection to support the design of a more accurate model.

Generalised and person specific approaches.

- This approach used in this investigation accounted for a generalised model that could be applied universally, as well as investigating models trained for

Provides another perspective for signal visualisation.

- Due to that fact that it is almost impossible to acquire a “clean” EEG signal, using CAEs to visualise the signal allow for another perspective as to how the actual raw EEG signal may look with no noise. Using these visualisations in conjunction with denoised signals obtained using other techniques allow for a well-rounded estimation of the EEG signal from a number of different perspectives.

Evidence that the technique performs better at higher SNR values.

- From **Error! Reference source not found.** and **Error! Reference source not found.** it can be seen that there is a more significant increase in the SNR values of electrodes related to the appropriate cortexes associated with the required tasks. This is important as it can identify what electrodes are presenting the highest cognitive activity and further investigation can be pursued relating to these specific areas of the brain.

Displays *10-20* system for electrode placement.

- The plotting of the electrode locations of the *10-20* system enables easy interpretation for individuals without backgrounds in neurology or neuroscience. The relative strengths of the SNR values are easily interpreted using the colour – coded heat map for SNR strength. This can allow readers inexperienced with brain activity analysis an intuitive understanding of the different areas of the brain affected by each task.

Use of Harrell-Davis decile value comparisons.

- Comparing the Harrell-Davis decile values of distributions and using the mean value of the difference of these values for different distributions reduced the effect of outliers on the distributions and allowed for a more accurate comparison of the data obtained from the different noise reduction techniques.

No bias in the data.

- The data used in this investigation was all real-world data, obtained from real participants. This ensured there could have been no data in the information, which is not always possible to determine when synthetic data is used for analysis, as assumptions have to be made about the signals.

4.5.2 Weaknesses/Limitations

Longer training times.

- The training of the model can take a significant amount of time when compared to other already highly established and well tested methods such as PCA.

Hyperparameter testing and optimisation time.

- It is only possible to observe the performance of the model once the model has been completely trained. From this, if the various hyperparameters are to be altered, it will take time to re-train the model with the new hyperparameters.

Small SNR increase skews results.

- Although there seems to be significant increase from the SNR of the PCA reconstructed signals and the CAE reconstructed signals, the difference in the SNR values appears more significant than it actually is.

Training data size.

- Although there was a large dataset available for the training of the Single CAE, as there were 81 participants in the study, there was significantly less data available for training each of the Multiple CAEs. Because of this, the Multiple CAE reconstructed signals may be less accurate than the Single CAE reconstructed signals.

Lack of baseline data.

- It is not possible to know whether an EEG signal is "clean" as it not possible to associate all electrical activity recorded at an electrode location with the exact area of the brain located beneath the electrode, as described in S**ection 2.2**.
- Any loss of information due to "denoising" of signal may be substantial for the analyses of the cognitive processes but cannot be concluded definitively.

PCA limited to 95% explained variance.

- This is a similar problem as stated above. Although PCA has been proven to reduce the noise of an EEG signal and increase the SNR, it is not possible to determine what information of the signal is noise and what is related to the cognitive functions of the required tasks. An example of an PCA reconstructed EEG signals with different explained variance values can be seen in **Figure 4.5.2-1**.

Higher SNR does not necessarily mean better signal.

- Expanding on the points above. Although the signals in **Figure 4.5.2-1** show and obvious decrease in noise, with a higher calculated SNR, it is obvious from the plots that the signal created using an explained variance of 76% is vastly different from the signal created using an explained variance of 99%, with reduced amplitudes and distinctive peaks and troughs.

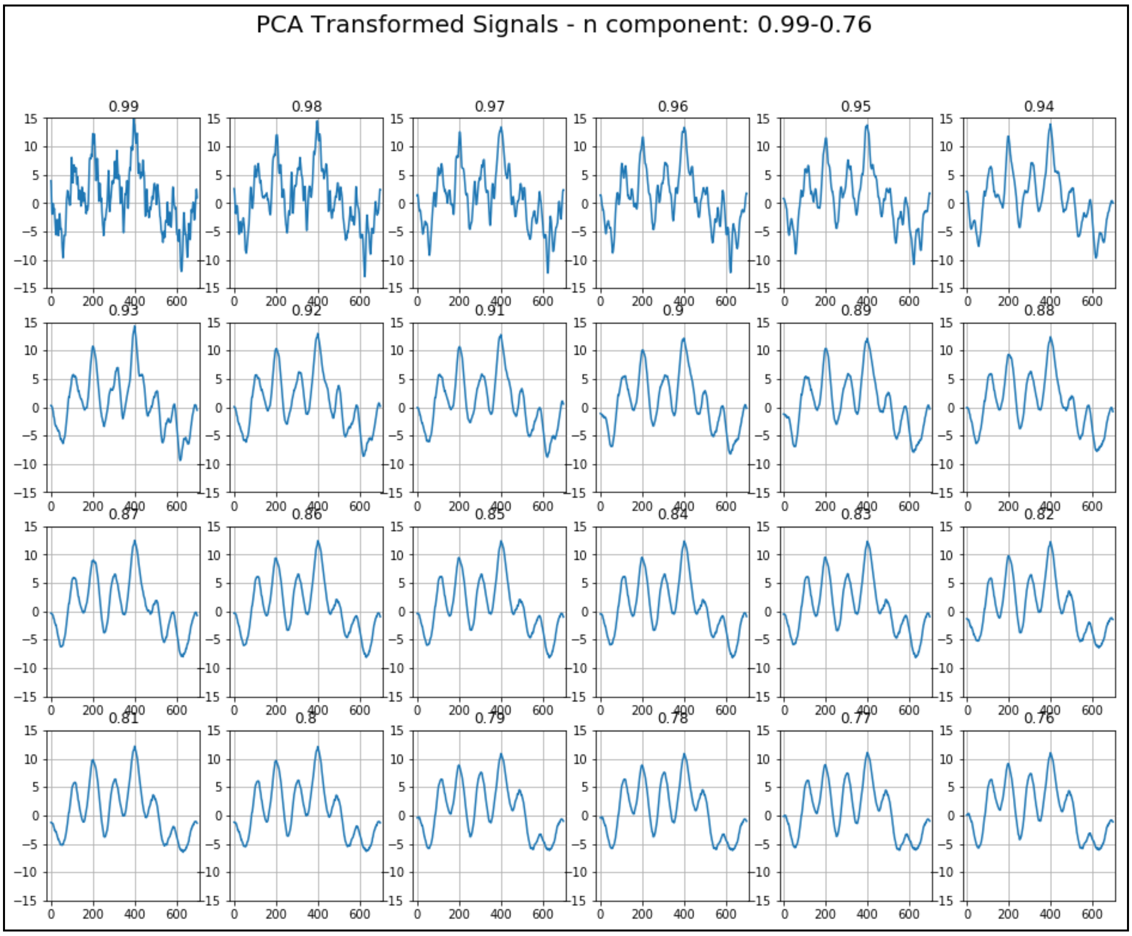


Figure 4.5.2-1 PCA reconstructed EEG signal for explained variance: 99% - 76%

CAE: Strengths & Weaknesses/Limitations	
Strengths	Better performance than PCA
	Computationally efficient
	Intuitive and easier to visualise.
	Generalised and person specific approaches
	Provides another perspective for signal visualisation
	Evidence that the technique performs better at higher SNR values
	Displays *10-20* system for electrode placement
	No bias in the data
Weaknesses/ Limitations	Longer training times than PCA
	Long hyperparameter testing and optimisation time
	Small SNR increase skews results
	Training data size discrepancy between single CAE and multiple CAE
	Lack of baseline data – no clean EEG signal
	PCA limited to 95% explained variance
	Higher SNR does not necessarily mean better signal.

Table 4.5.2-1 Strengths vs Weaknesses/Limitations of denoising CAE

4.5.3 Potential Causes of Error

Hyperparameter sharing.

- The same model was used for both the Single CAE and the Multiple CAEs. Because the model was originally tested on 10% of randomly selected data from the entire dataset, the hyperparameters were tunes and optimised for a general model, and not subject specific. This could result in bias within the multiple CAE models.

Less training data for each subject.

- The lesser performance of the Multiple CAEs compared to the Single CAE may be a result of significant difference in training data available to each model. The Single CAE had more training data than the Multiple CAEs by a factor of 80, which could have heavily hindered the performance and accounted for the performance discrepancies.

Simplicity of autoencoder.

- The autoencoder designed was based on a simple 2-layer convolutional neural network. As a result, there were less hyperparameters to tune, and therefore less factors to adjust in order to obtain more accurate results.

Using unsatisfactory statistical metrics.

- Using the mean values as the SNR comparison metric may not be appropriate as the mean value of an array of numbers can be heavily influenced by outliers within the array.
- For the Multiple CAEs, calculating the mean SNR of all electrodes can lead to biased results. The cortexes not associated with the required tasks have lower signal-to-noise ratios, which reduce the overall mean of the data. Because of this, the impact of the engaged cortexes is reduced. The mean SNR values for each electrode and their location on the *10-20* system can be seen in **Figure 4.3.1-3**.

Discrepancies between subjects.

- The original study associated with this research involved analysing brain activity of individuals diagnosed with schizophrenia. 49 of the subjects were diagnosed with the condition and 22 were used as control subjects. As cognitive processes of individuals with schizophrenia are sometimes different to individuals without the disorder, and therefore the designed model may be construed as biased.

ERP refractoriness.

- The noise recorded for the SNR of a signal is the EEG activity measured 100ms before stimulus, but as described in **Section 2.2**, if trials are conducted in quick succession, the cognitive reaction related to the stimulus may overlap with the noise to be measured of the next stimulus onset, if not with the onset of the stimulus itself. This can result inaccurate analysis of the EEG signal, especially if an individual is predisposed to longer lasting cognitive activity, lessening the impacts of each of their trials.

4.5.4 Potential Improvements

Individualised hyperparameters.

- As each person may process information through different cognitive processes, it may be beneficial to train the CAEs for each subject individually. Future subjects could be classified by a number of medical or physiological attributes and grouped with the subject analysed in this study whom is most similar based on these characteristics. This could lead to more accurate analysis of their EEG signals.

Increase individual training and test data.

- It is always applicable to add more training and test data to a model in order to obtain more accurate results. It may be beneficial to combine this point and the above point, and cluster the 81 individuals based on a number of attributes, which could be more precisely chosen by an expert in the field of neuroscience, resulting in model generalised enough to be applicable to almost any individual, but specific enough to provide more accurate EEG analysis.

Clustering participants.

- As stated in the point above, clustering the participants analysed in this study could be extremely beneficial to the accuracy of the models. The most immediate and obvious clusters that could be used with this data would be used participants diagnosed with schizophrenia and participants whom are not. This could also provide specific brain activity information related to the condition which could be particularly beneficial to the medical industry in the future.

Increasing autoencoder complexity.

- As the autoencoder designed in this study was somewhat simplistic in its design, the use of a more complex model could greatly improve the accuracy of the results. Not only this, but both approaches could be compared in order to analyse whether a simpler approach may be more beneficial for EEG analyses, as it may be more intuitive and less time consuming for hyperparameter optimisation and model training.

Using more appropriate statistical analysis metrics.

- Using the mean can result in inaccurate statistical analysis when comparing distributions. For the Multiple CAE analysis, comparing the mean values of the distributions may have resulted in outlier influence on the final analysis and comparison of noise reduction techniques. Using the mean Harrell-Davis decile differences for technique comparison combatted this discrepancy somewhat but may have still given inaccurate results. Analysing the data by counting the number of positive and negative decile difference per subject when comparing PCA and CAE, as done with the Single CAE, and then counting the number of subjects that exhibited more decile differences in favour of the CAE may have been a more accurate approach.

Isolate electrodes of interest.

- Isolating the electrodes only associated with the cognitive functions required of the tasks could provide more accurate insights and analysis for model comparison. Including electrodes that display mainly noise can affect the outcome of the statistical analysis.

Higher dimensional CNN.

- Using a 2-dimensional CNN instead of the 1-dimensional CNN used in this project could combat the issue raised above, as it isolates each electrode and analyses them appropriately. The input of the 2-D CNN would be a dataframe of shape 700 x 64 (time x channels). Although this would reduce variety of the inputs of the CAE, it would result in a more dynamic analysis, with electrode specific readings.

Train as Condition dependent.

- All models designed in this analysis were trained using data from all three Conditions required of the subjects which provided generalised models for brain activity from a number of different perspectives. However, training a model using data obtained from only a single specific task could result in more accurate analysis. This argument, combined with a 2-D CNN, could greatly improve the accuracy of the model, which could allow for future work in task classification.

Increase number of Conditions.

- As an extension of training the model for Condition independence, analysing cognitive processes related to different stimuli could also be undertaken. Including this data in the training of the model could allow for a higher generalisation for different cognitive analysis, applicable to more individuals and could be more beneficial to the field of neurology.

Compare against other proven methods.

- It could be particularly beneficial to compare the performance of the designed model to other proven noise reduction techniques such as Independent Component Analysis (ICA). This allows for a general overview of the performance of the model and can identify particular issues of the field of study that may prove particularly difficult to overcome.

Combine with other methods.

- As described in **Section** Error! Reference source not found., there are a number of previously proven techniques highly capable of significant noise reduction in EEG signals. Combining the described model in this investigation with a range of other artifact removal and noise reduction techniques could prove extremely effective, resulting with higher performance and more accurate results than would be obtained with this technique alone.

CAE: Potential Causes of Error & Potential Improvements	
Potential Causes of Error	Hyperparameter sharing
	Less training data for each subject
	Simplicity of autoencoder
	Using unsatisfactory statistical metrics
	Discrepancies between subjects
	ERP refractoriness
Potential Improvements	Individualised hyperparameters
	Increase individual training and test data
	Clustering participants
	Increasing autoencoder complexity
	Using more appropriate statistical analysis metrics
	Isolate electrodes of interest
	Higher dimensional CNN
	Train as Condition dependent
	Increase number of Conditions
	Compare against other proven methods
	Combine with other methods

Table 4.5.4-1 Potential Causes of Error and Potential Improvements of Experiment

5 CONCLUSION

5.1 RESEARCH OVERVIEW

The research carried out in this investigation explored brain activity analysis, in particular, the use of electroencephalography to analyse cognitive activity. Electrical fluctuations are constantly generated by the brain, when performing any cognitive tasks. Electroencephalography (EEG) is an electrophysiological monitoring technique that detects and records these electrical fluctuations. The technique involves recording voltage fluctuations from different areas of the brain using strategically places electrodes on the scalp of an individual, providing real-time information of the electrical activity.

Using EEG allows for real-time analysis of cognitive activity, with a high temporal resolution. It’s a simple and cost-effective analysis technique when compared to other brain analysis methods, such as Functional Magnetic Resonance Imaging (fMRI) and Computed Tomography (CT) scans. The use of EEG signals has been widely employed on many platforms, greatly benefitting the medical industry in particular.

5.2 PROBLEM DEFINITION

The advantages of using EEG for cognitive analysis are significant, yet as with the other signal processing applications, there come limitations associated with the use of this widely employed technique. Due to the complex nature of the brain, excess electrical fluctuations are constantly present, even in areas to which the task being undertaken by an individual is not related, and as a result, it is not possible to obtain a perfectly “clean” EEG signal. This excess electrical activity can be attributed to the constant stimulation of millions of neurons within the brain. The presence of these constant electrical fluctuations reduces the quality of these EEG signals, and the ability of the researcher to fully analyse and evaluate the cognitive processes that are executed by the brain.

In order to effectively examine the processing procedure of the brain in its entirety, an EEG signal without the presence of these excess voltage fluctuations, as well as other unwanted artifacts, is highly desirable**.** These excess fluctuations and artifacts are known as *noise* in the signal. A robust and efficient method is needed to remove this noise from the signal, and enable the detailed analysis the EEG signal as accurately as can be achieved.

Many approaches have been undertaken in order to address this problem. These methods range from classical mathematical operations and transformations such as filtering and wavelet transformations, as well as dimensionality reduction, such as Principal Component Analysis and Independent Component Analysis, to modern Machine Learning techniques, including Support Vector Machines (SVM) and Recurrent Neural Networks (RNN).

Although these techniques have been thoroughly analysed and tested, and have proven to be vastly effective, various issues have arisen for each technique that have encumbered the noise reduction of EEG signals, and therefore the accuracy of EEG signal analysis. These issues range mathematical to technical, including common frequencies between noise and important signal information, and the need for manual supervision.

The formulation of a technique that can overcome these issues would be very beneficial to the field of EEG signal analysis, as it could allow the improvement of signal quality and, ultimately, enable more accurate analysis of cognitive functionality. In order to achieve these goals and fulfil the conditions required of an effective EEG noise reduction technique, a new idea must be conceptualised, designed and implemented, with the robustness of previously proven techniques, while omitting their respective limitations.

In order to accomplish these goals, a stacked autoencoder designed using a one-dimensional convolutional neural network framework was proposed, with the goal of reducing the noise present in the EEG signals, while maintaining the critical signal information. The desired result of the experiment was to increase signal-to-noise ratios of the EEG signals, with a better performance than previously proven techniques, in particular Principal Component Analysis.

5.3 DESIGN/EXPERIMENTATION, EVALUATION, RESULTS

This investigation was conducted in order to fulfil the alternate hypothesis:

> If a stacked autoencoder is designed using a one-dimensional convolutional neural network layer, the signal-to-noise ratio of electroencephalography data can be increased when compared to principal component analysis.

The 1-dimensional convolutional autoencoder (CAE) was designed in order to increase the signal-to-noise ratio. The data used in the experiment consisted of EEG signal data collected from 81 subjects. Each subject was required to perform three separate tasks, with each task performed ~ 100 times. The tasks consisted of three independent button pressing Conditions, each designed to stimulate different specific cortexes of the brain, or a combination of specific cortexes. The voltage fluctuations during each trial were recorded over a time of 700ms for all three Conditions, with the EEG signals recorded on 64 electrodes placed upon each subject's scalp using the *10-20* location system, a voltage amplitude recording conducted every millisecond. The onset of the Condition occurred at 100ms.

The CAE was designed such that each single reading from each of the electrodes would be stored as a vector of length 700, and used as the input for the model, with the resulting output from the CAE the same length as the input vector. The signal-to-noise ratio was calculated for each individual signal, before and after it was processed by the CAE. The distributions of the SNRs before and after were recorded and compared.

Combined with this, PCA was performed on each of the input vectors to reduce the data to 95% explained variance. The SNR of each reconstructed signal was calculated, and the distribution was analysed and compared to that of the SNR distribution of the original (RAW) signals. To compare the difference in performance between PCA and the CAE, the differences of the signal SNR distribution for each reconstruction method and the raw signal SNR distribution were analysed. This allowed for a quantitative analysis of the CAE performance. In order to compare the distribution differences, the Harrell-Davis (HD) decile values for each distribution were calculated, and the differences between the RAW and PCA HD decile differences and the RAW and CAE HD decile differences were calculated.

Single CAE: In order to determine whether the model was successful at outperforming PCA for EEG noise reduction, and therefore increasing the SNR of the signals, the difference between each HD decile difference for PCA and CAE reconstructed signals were divided by the respective HD decile differences between the RAW and PCA signals. This quantified the extent to which CAE performed in relation to PCA. For evidence to support the rejection of the null hypothesis, 5 of the 9 calculated values must be positive.

Multiple CAE: This process was also conducted for each subject individually, using the data for each participant independently of other participants. However, instead of calculating the 9 HD decile value differences for the SNR distributions of each individual and analysing each value independently, the mean of these HD decile values was compared for each distribution. The mean value of the RAW SNR distribution HD decile differences was subtracted from the mean value of the PCA SNR distribution HD decile differences, as well as the mean value of the CAE SNR distribution SNR differences, for each subject. At least 41 of the 81 subjects must have a greater mean SNR difference between CAE and RAW signals than PCA and RAW signals in order to suggest that there is evidence to support the rejection of the null hypothesis.

The tests were carried out for each of the three Conditions, for both the Single CAE and the Multiple CAEs.

Single CAE:

Condition 1: It was found that 8 of the 9 decile values for the Single CAE were in favour of the CAE, indicating there is evidence to support the rejection of the null hypothesis.

Condition 2: It was found that 8 of the 9 decile values for the Single CAE were in favour of the CAE, indicating there is evidence to support the rejection of the null hypothesis.

Conditions 3: It was found that 8 of the 9 decile values for the Single CAE were in favour of the CAE, indicating there is evidence to support the rejection of the null hypothesis.

Multiple CAE:

Due to missing data for Condition 3 of Subject 46, all information was removed for Subject 46 in order to reduce bias in the results.

Condition 1: It was found that the RAW-CAE SNR distribution HD decile differences were greater than the RAW-PCA SNR distribution HD decile differences for 40 of the 80 subjects. This does not indicate that there is evidence for supporting the rejection of the null hypothesis.

Condition 2: It was found that the CAE-RAW SNR distribution HD decile differences were greater than the RAW-PCA SNR distribution HD decile differences for 35 of the 80 subjects. This does not indicate that there is evidence for supporting the rejection of the null hypothesis.

Condition 3: It was found that the RAW-CAE SNR distribution HD decile differences were greater than the RAW-PCA SNR distribution HD decile differences for 37of the 80 subjects. This does not indicate that there is evidence for supporting the rejection of the null hypothesis.

The information above indicates that a larger generalised model with a large training dataset may be more applicable for noise reduction in EEG signals than individualised, subject dependent models.

5.4 CONTRIBUTION & IMPACT

From the research carried out in this investigation, it was found that it was possible to design and implement a stacked autoencoder based on a 1-dimensional convolutional neural network capable of reducing the noise present in Electroencephalography signals and increase their signal-to-noise ratios.

As well as this, it was established that the convolutional autoencoder could perform better than Principal Component Analysis for noise reduction in EEG signals for a PCA explained variance of 95%.

It was also determined that a general model, trained with EEG signal data from a number of different individuals exhibited a better performance than individualised models trained only using data from a single subject.

This information obtained from the research adds to the large body of work in the field of signal processing and neuroscience that has already been well established. As it is not possible to obtain a "clean" real-world EEG signal, the use of convolutional autoencoders to estimate the shape of an EEG signal provides an alternative approach to approximating EEG signals, which can be used in conjunction with previously used methods, to expand the analysis available to be conducted on the signal.

The concepts proposed and explored in this investigation reinforce the relevance of machine learning and deep learning techniques across numerous industries.

5.5 FUTURE WORK & RECOMMENDATIONS

- Due to the large volume of data processed in this investigation, a large number of further explorations can be conducted into the functionality of individuals' brains, and their cognitive processing characteristics. The other data associated with the subjects analysed in the studies upon which this investigation is based (Ford, Palzes, Roach, & Mathalon, 2013), includes age, gender, number of years in education and whether or not they have been diagnosed with schizophrenia. The availability of this data allows for a wide range of further investigations. These explorations may include differences in cognitive processing for increasing years in education, as well as cortex stimulation differences.

- As the aim of the original study, from which this data was acquired, was to analyse the brain activity of individuals diagnosed with schizophrenia and compared to the brain activity of a number of control subjects, a promising avenue of research which may be beneficial to pursue is using this information to aid classification analyses. With 100 trials and 64 channels for 3 separate conditions for each subject, there is a significant amount of training and test data that can be used in order to classify individuals for a number of different attributes using "clean" EEG signals. Performing transformations, such as Fourier Transformations, on the reconstructed signals could result in more accurate schizophrenia classification, allowing for more accurate diagnosis in the future.

- As each of the channels of the reconstructed data are labelled with their appropriate scalp location, attention could be focused on electrode isolation, and specific electrode behaviour based on particular cognitive tasks.
  This approach could allow for more accurate analysis of specific cortex functionality, and differences in specific cortex behaviour between different groups of individuals.

- As an extension of the proposed research on electrode isolation mentioned above, classification of tasks could also be investigated based on simultaneous readings of all 64 electrodes. As can be seen in the investigation performed in this paper, different cortexes of the brain become stimulated when different cognitive functions are required of an individual. With enough data, it may be possible to use EEG signals in order to classify tasks required of an individual. A wide range of potential possibilities become more attainable with this concept. With successful classification, a brain-computer interface can be established in order to allow successful communication to a machine for the execution of particular tasks. An expansion of this could be the increased functionality of prosthetic limbs.

BIBLIOGRAPHY

Adrian, E. D., & Yamagiwa, K. (1935). The origin of the Berger rhythm. *Brain, 58*, 323-351.

Ahmadi, M., & Quiroga, R. Q. (2013, February). Automatic denoising of single-trial evoked potentials. *NeuroImage*, *66*, 672-680. doi:10.1016/j.neuroimage.2012.10.062

Babu, P. A., & Prasad, K. V. (2011). Removal of Ocular Artifacts from EEG Signals Using Adaptive Threshold PCA and Wavelet Transforms. *International Conference on Communication Systems and Network Technologies* (pp. 572-575). Katra: IEEE. doi:10.1109/CSNT.2011.122

Baldi, P. (2012). Autoencoders, Unsupervised Learning, and Deep Architectures. *Workshop on Unsupervised and Transfer Learning* (pp. 37-50). Journal of Machine Learning Research.

Baldi, P., & Hornik, K. (1989). Neural Networks and Principal Component Analysis: Learning from Examples Without Local Minima. *Neural Networks, 2*(1), 53-58. doi:10.1016/0893-6080(89)90014-2

Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Archiv für Psychiatrie & Nervenkrankheiten,, 87*, 527-570.

Boubchir, L., & Boashash, B. (2013, April). Wavelet Denoising Based on the MAP Estimation Using the BFK Prior With Application to Images and EEG Signals. *IEEE Transactions on Signal Processing, 61*(8), 1880-1894. doi:10.1109/tsp.2013.2245657

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*, 365–376. doi:10.1038/nrn3475

Calin-Jageman, R. J. (2017). After p Values: The New Statistics for Undergraduate Neuroscience Education. *Journal of Undergraduate of Neuroscience Education, 16*(1), E1-E4.

Cooper, R., Winter, A. L., Crow, H. J., & Walter, W. G. (1965). Comparison of subcortical, cortical, and scalp activity using chronically indwelling electrodes in man. *Electroencephalography & Clinical Neurophysiology, 18*, 217-228.

Davis, H. (1964). Enhancement of evoked cortical potential in humans related to a task requiring a decision. *Science, 145,* 182-183.

Davis, P. A. (1939). Effects of acoustic stimuli on the waking human brain. *Journal of Neurophysiology, 2,* 494-499.

Dawson, G. D. (1954). A summation technique for the detection of small evoked potentials. *Electroencephalography & Clinical Neurophysiology, 6*, 64-84.

Donchin, E., & Cohen, L. (1967). Average evoked potentials and intramodality selective attention. *Electroencephalography & Clinical Neurophysiology, 22*, 537-546.

Donchin, E., & Heffley, E. (1975). Minicomputers in the signal-averaging laboratory. *American Psychologist, 30*, 299-312.

Donoho, D., & Johnstone, I. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika, 81*(3), 425-455.

Duncan-Johnson, C. C., & Donchin, E. (1979). The time constant in P300 recording. *Psychophysiology, 16*, 53-55.

Ebersol, J. S. (1997). Defining epileptic foci: Past, present, future. *Journal of Clinical Neurophysiology, 14,* 470-483.

Ford, J. M., Palzes, V. A., Roach, B. J., & Mathalon, D. H. (2013). Did I Do That? Abnormal Predictive Processes in Schizophrenia When Button Pressing to Deliver a Tone. *Schizophrenia Bulletin, 40*(4), 804–812. doi:10.1093/schbul/sbt072

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2017). Alpha-Band Oscillations Enable Spatially and Temporally Resolved Tracking of Covert Spatial Attention. *Psychological science, 28*(7), 929–941. doi:10.1177/0956797617699167

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics, 36*, 193-20.

Garćıa-Laencina, P. J., Rodŕıguez-Bermudez, G., & Roca-Dorda, J. (2014, September). Exploring dimensionality reduction of EEG features in motor imagery task classification. *Expert Systems with Applications, 41*(11), 5285--5295. doi:10.1016/j.eswa.2014.02.043

Gondara, L. (2016). Medical Image Denoising Using Convolutional Denoising Autoencoders. *16th International Conference on Data Mining Workshops (ICDMW)* (p. International Conference on Data Mining Workshops (ICDMW)). Barcelona: IEEE. doi:10.1109/ICDMW.2016.0041

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* MIT Press.

Grossman, A., & Morlet, J. (1985). Decompositions of functions into wavelets of constant shape and related transforms. In L. Streit (Ed.), *Mathematics + Phyics* (pp. 135-165). Singapore: World Scientific.

Haas, L. F. (2003). Hans Berger (1873 - 1941, Richard Caton (1842 - 1926), and electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry, 71*(1), 9. doi:10.1136/jnnp.74.1.9

Hair, J., Anderson, R., Tatham, R., & Black, W. (1995). *Multivariate data analysis* (4 ed.). New Jersey: Prentice-Hall Inc;.

Hassani, M., & Karami, M. R. (2015). Noise Estimation in Electroencephalogram Signal by Using Volterra Series Coefficients. *Journal of Medical Signals and Sensors, 5*(3), 192-200. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/26284176

Haufe, S., Dahne, S., & Nikulin, V. V. (2014, November). Dimensionality reduction for the analysis of brain oscillations. *NeuroImage, 101*, 583-597. doi:10.1016/j.neuroimage.2014.06.073

Helal, M. A., Eldawlatly, S., & Taher, M. (2017). Using Autoencoders for Feature Enhancement in Motor Imagery Brain-Computer Interfaces. *Biomedical Engineering.* ACTAPRESS. doi:10.2316/p.2017.852-052

Helmholtz, H. v. (1853). Über einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern mit Anwendung auf die thierisch-elektrischen Versuche. *Annalen der Physik & Chemie,, 89*, 211-233, 354-377.

Herault, J., & Jutten, C. (1986). Space or time adaptive signal processing by neural network models. *American Institute of Physics Conference Series*, *151*, pp. 206-211. doi:10.1063/1.36258

Homan, R. W., Herman, J., & Purdy, P. (1997). Cerebral location of international 10–20 system electrode placement. *Electroencephalography & Clinical Neurophysiology, 66*, 376-382.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive Fields and Functional Architecture of Monkey Striate Cortex. *The Journal of Physiology, 195*(1), 215–243. doi:10.1113/jphysiol.1968.sp008455

Ille, N., Berg, P., & Scherg, M. (2002). Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *Journal of Clinical Neurophysiology,, 19*, 113-124.

Ioannidis, J. P. (2005). Why Most Published Research Findings Are False. *PLOS Medicine, 2*(8), e124. doi:10.1371/journal.pmed.0020124

Jasper, H. H. (1937). Electrical signs of cortical activity. *Psychological Bulletin, 34*, 411-481.

Jasper, H. H. (1948). Charting the sea of brain waves. *Science, 108*, 343-347.

Jasper, H. H. (1958). The ten-twenty electrode system if the International Federation. *Electroencephalography & Clinical Neurophysiology, 10*, 371-375.

Joyce, C. A., Gorodnitsky, I. F., & Kutas, M. (2003). Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. *Psychophisiology, 41*(2), 313-325. doi:10.1111/j.1469-8986.2003.00141.x

Jung, T.-P., Humphries, C., Lee, T.-W., Makeig, S., McKeown, M. J., Iragui, V., & l., S. T. (1997). Extended lCA Removes Artifacts from Electroencephalographic Recordings. *NIPS'97 Proceedings of the 10th International Conference on Neural Information Processing Systems* (pp. 894-900). Denver: MIT Press.

Kang, D., & Zhizeng, L. (2012). A Method of Denoising Multi-channel EEG Signals Fast Based on PCA and DEBSS Algorithm. *International Conference on Computer Science and Electronics Engineering* (pp. 322-326). Hangzhou: IEEE. doi:10.1109/ICCSEE.2012.105

Kim, J. H. (2018, June). Time Frequency Image and Artificial Neural Network Based Classification of Impact Noise for Machine Fault Diagnosis. *International Journal of Precision Engineering and Manufacturing, 19*(6), 821–827.

Kiran, B. R., Mathew Thomas, D., & Parakka, R. (2018). An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos. *arXiv*. doi:arXiv:1801.03149v2

Koli, M. (2012). Review of Impulse Noise Reduction Techniques. *International Journal on Computer Science and Engineering, 4*(2), 184-196.

Kulasingham, J. P., Vibujithan, V., & De Silva, A. C. (2016). Deep belief networks and stacked autoencoders for the P300 Guilty Knowledge Test. *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences IECBES* (pp. 127-132). Kuala Lumpur: IEEE. doi:10.1109/IECBES.2016.7843428

Lauzon, F. Q. (2012). An introduction to deep learning. *International Conference on Information Science, Signal Processing and their Applications (ISSPA)* (pp. 1438-1439). Montreal: IEEE. doi:10.1109/ISSPA.2012.6310529

Li, C.-L., McLennan, H., & Jasper, H. (1952). Brain waves and unit discharge in cerebral cortex. *Science, 116*, 656-657.

Logothetis, N., Pauls, J., Augath, M. A., Trinath, T., & Oeltermann, A. (2001). Neurophysical investigation of the basis of the fMRI signal. *Nature, 412*, 150-157.

Lu, X., Tsao, Y., Matsuda, S., & Hori, C. (2013). Speech Enhancement Based on Deep Denoising Autoencoder. *INTERSPEECH-2013* (pp. 436-440). Lyon: International Speech Communication Association.

Lu, Z.-L., Williamson, S. J., & & Kaufman, L. (1992). Behavioral lifetime of human auditory sensory memory predicted by physiological measures. *Science, 258*, 1668-1670.

Luck, S. J. (2005). *An introduction to the event-related potential.* Cambridge: MIT Press.

Maddirala, A. K., & Shaik, R. A. (2016). Removal of EOG Artifacts From Single Channel EEG Signals Using Combined Singular Spectrum Analysis and Adaptive Noise Canceler. *IEEE Sensors Journal, 16*(23), 8279-8287. doi:10.1109/JSEN.2016.2560219

Maddirala, A., & Shaik, R. A. (2016). Removal of EOG Artifacts from Single Channel EEG Signals using Combined Singular Spectrum Analysis and Adaptive Noise Canceler. *IEEE Sensors Journal*, 1. doi:10.1109/jsen.2016.2560219

Mahajan, R., & Morshed, B. I. (2015, January). Unsupervised Eye Blink Artifact Denoising of EEG Data with Modified Multiscale Sample Entropy, Kurtosis, and Wavelet-ICA. *IEEE Journal of Biomedical and Health Informatics, 19*(1), 158-165. doi:10.1109/jbhi.2014.2333010

Maria Navin, J., & Balaji, K. (2016). Performance Analysis of Neural Networks and Support Vector Machines using Confusion Matrix. *International Journal of Advanced Research in Science, Engineering and Technology*, 2106 - 2019.

McDonald, J. J., Teder-Sälejärvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature, 407*, 906-908.

McFarland, D. J., McCane, L. M., David, S. V., & Wolpaw, J. R. (1997). Spatial filter selection for EEG-based communication. *Electroencephalography and Clinical Neurophysiology, 103*(3), 386-394.

Nelson, D. A., & Lassman, F. M. (1973). Combined effects of recovery period and stimulus intensity on the human auditory evoked vertex response. *Journal of Speech & Hearing Research, 16*, 297-308.

Neubauer, C. (1998). Evaluation of Convolutional Neural Networks for Visual Recognition. *IEEE TRANSACTIONS ON NEURAL NETWORKS, 9*(4), 685-696. doi:10.1109/72.701181

Nguyen, H.-A. T., Musson, J., Li, F., Wang, W., Zhang, G., Xu, R., . . . Li, J. (2012). EOG artifact removal using a wavelet neural network. *Neurocomputing, 97*(1), 374-389. doi:10.1016/j.neucom.2012.04.016

Nikulin, V. V., Nolte, G., & Curio, G. (2011, April). A novel method for reliable and fast extraction of neuronal EEG/MEG oscillations on the basis of spatio-spectral decomposition. *NeuroImage, 55*(4), 1528-1535. doi:10.1016/j.neuroimage.2011.01.057

Novakovic, J. D., Veljovic, A., Ilic, S. S., Papic, Z., & Tomovic, M. (2017). Evaluation of Classification Models in Machine Learning. *Theory and Applications of Mathematics & Computer Science*, 39-46.

Nunez, P., & Srinivasan, R. (2006). *Electrical fields of the brain. The neurophysics of EEG* (2 ed.). Oxford: Oxford University Press.

O’Regan, S., & Marnane, W. (2013). Multimodal detection of head-movement artefacts in EEG. *Journal of Neuroscience Methods, 218*(1), 110-120.

Pander, T. (2019). EEG signal improvement with cascaded filter based on OWA operator. *Signal, Image and Video Processing, 13*(6), 1165-1171. doi:10.1007/s11760-019-01458-9

Pardede, J., Turnip, M., Robinson Manalu, D., & Turnip, A. (2015). Adaptive Recurrent Neural Network for Reduction of Noise and Estimation of Source from Recorded EEG Signals. *ARPN Journal of Engineering and Applied Sciences, 10*(3), 993-997.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series 6, 2*(11), pp. 559-572.

Peng, H., Hu, B., Shi, Q., Ratcliffe, M., Zhao, Q., Qi, Y., & Gao, G. (2013). Removal of ocular artifacts in EEG - an improved approach combining DWT and ANC for portable applications. *IEEE Journal of Biomedical and Health Informatics, 17*(3), 600-307.

Pfurtscheller, G., Stancák, A. J., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band—an electrophysiological correlate of cortical idling: A review. *International Journal of Psychophysiology, 24*, 39-46.

Qiu, Y., Zhou, W., Yu, N., & Du, P. (2018, September). Denoising Sparse Autoencoder-Based Ictal EEG Classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26*(9), 1717-1726. doi:10.1109/TNSRE.2018.2864306

Raichle, M. E. (2009). A brief history of human brain mapping. *Trends in Neurosciences, 32*(2), 118-126. doi:10.1016/j.tins.2008.11.001

Reris, R., & Brooks, J. P. (2015). Principal Component Analysis and Optimization: a tutorial. *INFORMS Computing Society Conference*, (pp. 212-225). Richmond. doi:10.1287/ics.2015.0016

Saruwatari, H., Kurita, S., Takeda, K., Itakura, F., Nishikawa, T., & Shikano, K. (2003). Blind source separation combining frequency-domain ICA and beamforming. *EURASIP Journal on Advances in Signal Processing, 2003*(11), 1135–1146.

Selvan, S., & Srinivasan, R. (2000). Recurrent Neural Network based Efficient Adaptive Filtering Technique for the Removal of Ocular Artefacts from EEG. *IETE Technical Review, 17*, 73-78.

Shearar, C. (2000). The CRISP-DM Model: The New Blueprint for Data Mining. *Journal of Data Warehousing, 5*(4), 13-2.

Srinivasulu, A., & Sreenath Reddy, M. (2012). Artifacts Removing From EEG Signals by ICA Algorithms,. *IOSR Journal of Electrical and Electronics Engineering, 2*(4), 11-16.

Supratak, A., Li, L., & Guo, Y. (2014). Feature extraction with stacked autoencoders for epileptic seizure detection. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 4184-4187). Chicago: IEEE. doi:10.1109/EMBC.2014.6944546

Szucs, D., & Ioannidis, J. P. (2017). When Null Hypothesis Significance Testing Is Unsuitable for Research: A Reassessment. *Frontier in Human Neuroscience, 11*, 390. doi:10.3389/fnhum.2017.00390

van Belle, G. (2008). *Statistical Rules of Thumb* (2 ed.). John Wiley & Sons, Inc. All rights reserved. doi:10.1002/9780470377963

Vaseghi, S. V. (2008). *Advanced Digital Signal Processing and Noise Reduction* (4 ed.). Chichester: John Wiley & Sons Ltd.

Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P.-A. (2008). Extracting and composing robust features with denoising autoencoders. *Proceedings of the 25th international conference on Machine learning - ICML.* ACM Press. doi:10.1145/1390156.1390294

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., & Manzagol, P.-A. (2010, December). Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. *J. Mach. Learn. Res., 11*, 3371--3408. Retrieved from http://dl.acm.org/citation.cfm?id=1756006.1953039

Vorobyov, S., & Cichocki, A. (2002). Blind noise reduction for multisensory signals using ICA and subspace filtering, with application to EEG analysis. *Biological Cybernetics, 86*(4), 293-303. doi:10.1007/s00422-001-0298-6

Walter, W. G. (1938). Critical review: The technique and application of electro-encephalography. *Journal of Neurology & Psychiatry, 1*, 359-385.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electrical sign of sensorimotor association and expectance in the human brain. *Nature, 203*, 380-384.

Wen, T., & Zhang, Z. (2018). Deep convolutional neural network and autoencoder-based unsupervised feature learning of EEG signals. *IEEE Access, 6*, 25399-25410.

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining* (pp. 29-39). Citeseer.

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics, 72*(8), 2031-2046. doi:10.3758/APP.72.8.2031

Woods, D. L., Hillyard, S. A., Courchesne, E., & Galambos, R. (1980). Electrophysiological signs of split-second decision-making. *Science*, *207*, 655-657.

Yang, B., Duan, K., Fan, C., Hu, C., & Wang, J. (2018, May). Automatic ocular artifacts removal in EEG using deep learning. *Biomedical Signal Processing and Control, 43,* 148--158. doi:10.1016/j.bspc.2018.02.021

Yin, Z., & Zhang, J. (2016). Recognition of Cognitive Task Load levels using single channel {EEG} and Stacked Denoising Autoencoder. *2016 35th Chinese Control Conference (CCC).* IEEE. doi:10.1109/chicc.2016.7553961

Zhao, M., Wang, D., Zhang, Z., & Zhang, X. (2015). Music removal by convolutional denoising autoencoder in speech recognition. *2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)* (pp. 338-341). Hong Kong: IEEE. doi:10.1109/APSIPA.2015.7415289

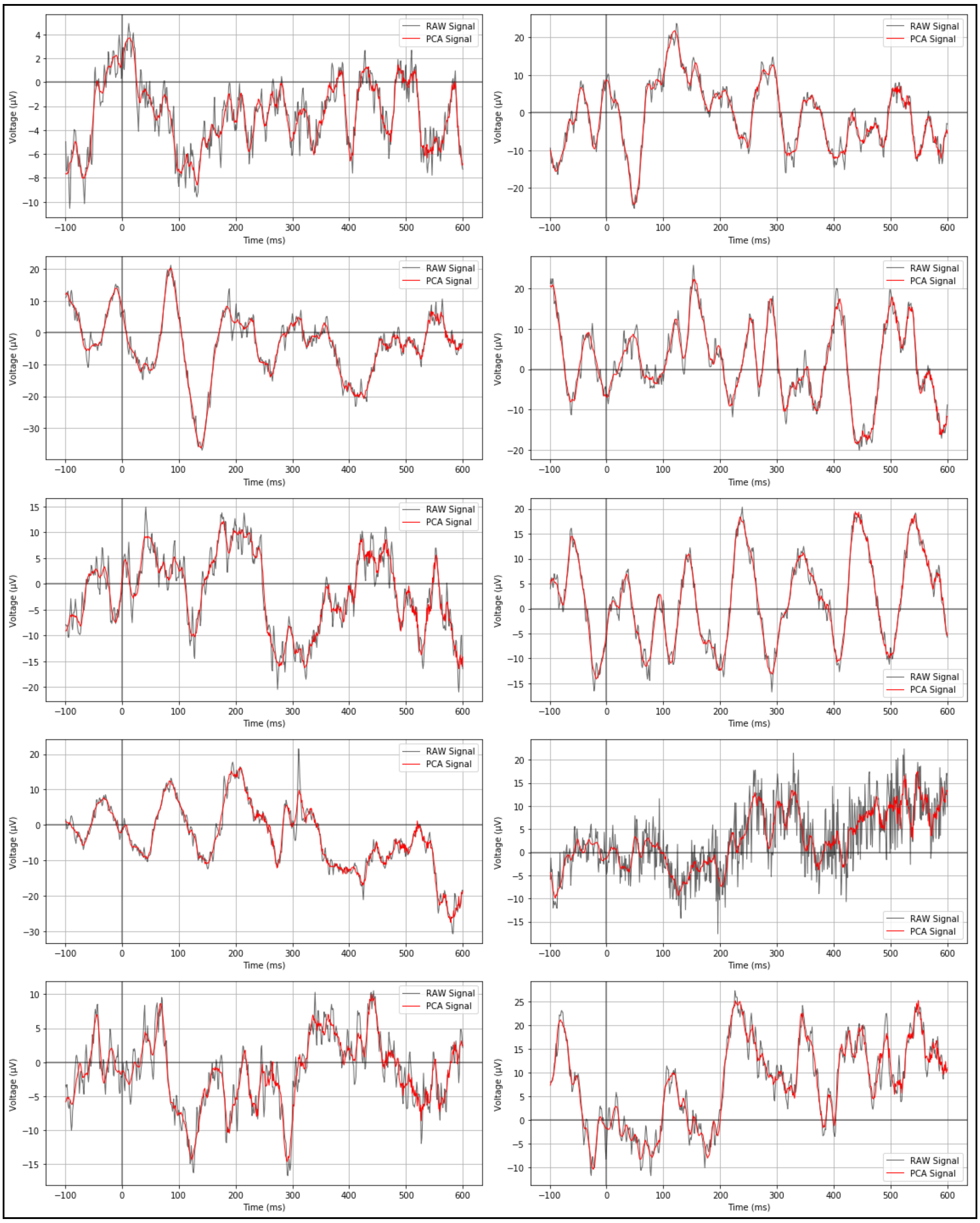
APPENDIX A

The data used in this investigation was acquired from:
https://www.kaggle.com/broach/button-tone-sz.

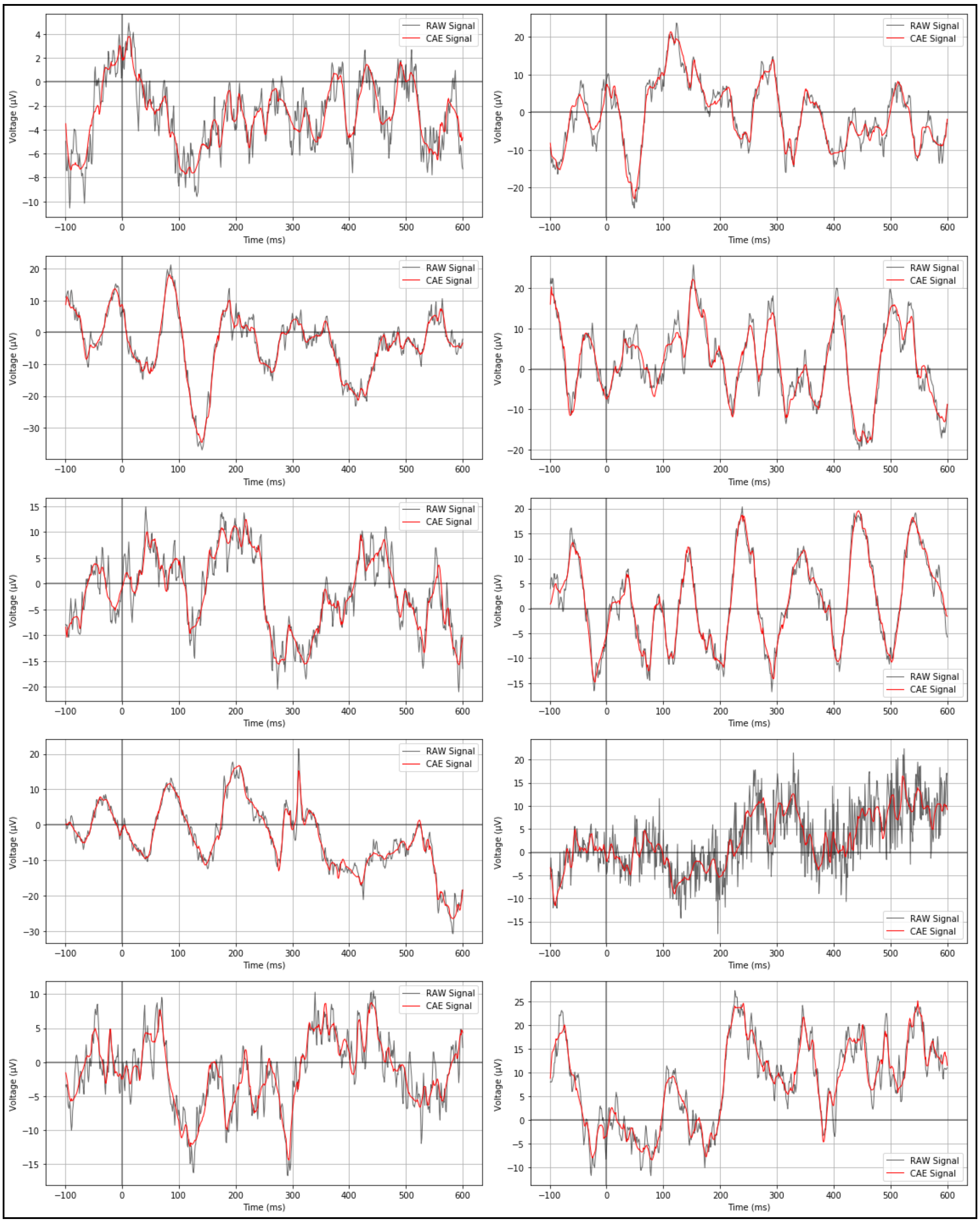
It may be important to note that a significant amount of time was spent focusing on the use of Fourier Transformations (FTs) of the EEG signals in order to investigate whether they could be used to train the convolutional autoencoder to produce more accurate results. FTs transform a signal from the time domain into the frequency domain.

- It was hypothesized that an FT could be performed on the signal, the output of this transformation could be processed by the CAE, and an inverse FT could be applied to the output of the CAE in order to produce the original signal, but with less noise.
- It was found not to be possible due to the fact that the mathematical processes of FTs produce a complex number as the output, a number with both real and imaginary parts. It is not possible to process imaginary numbers with neural networks (NNs). The imaginary part of the transformation is automatically removed before processing, leaving only the real part as the NN input. Although the output of the NN will be a real number in the frequency domain, if this value is to be returned back to a time domain signal using an inverse FT, it will be assumed by the algorithm that the value of the imaginary part of the signal has a value of 0, which is not equivalent to the NN input. The inverse FT then produces a signal in the time domain accounting for this information, thus producing a significantly different signal than was first used.
- Often for FT analysis, the absolute value of the FT output is used instead of the original output, as it ensures the imaginary part of the output is used, but can still be processed and visualised etc. However, this still results in information loss. If an inverse FT is applied to this value, the algorithm again assumed the imaginary value is 0, and therefore produces a different output than the original FT input.

- Although the use of FTs is limited for autoencoding, these processes can be, and have successfully been, used for classification purposes and have been found to perform very well.



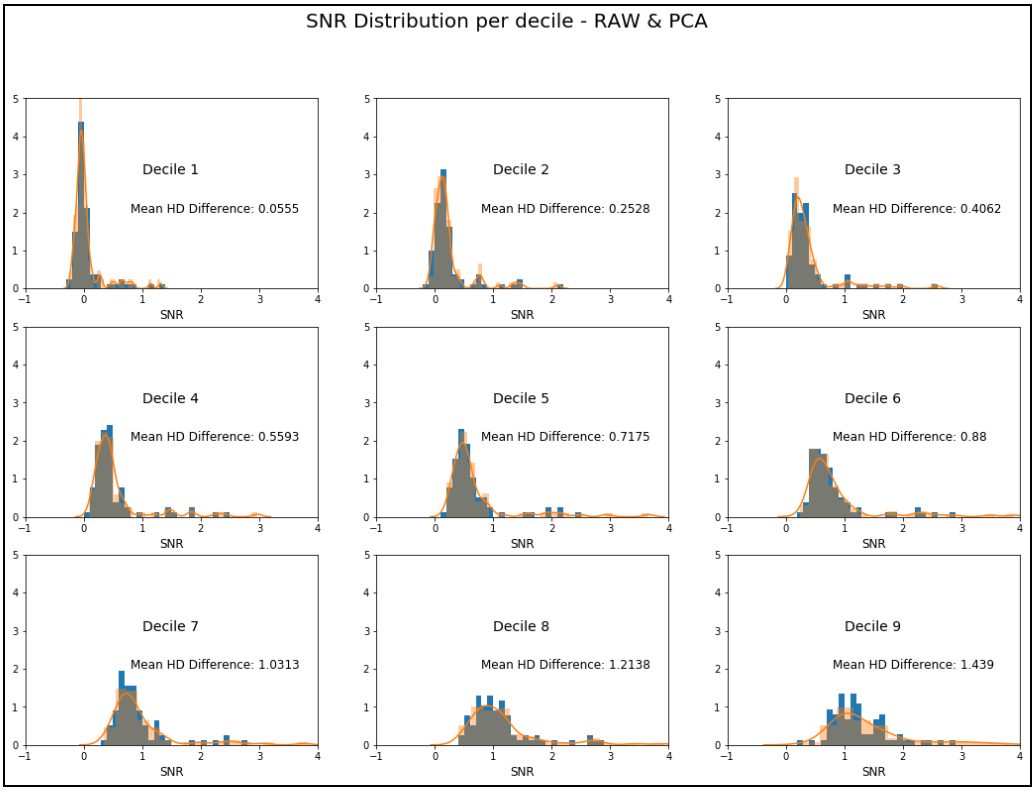
10 random RAW signals with PCA reconstructed signals overlaid



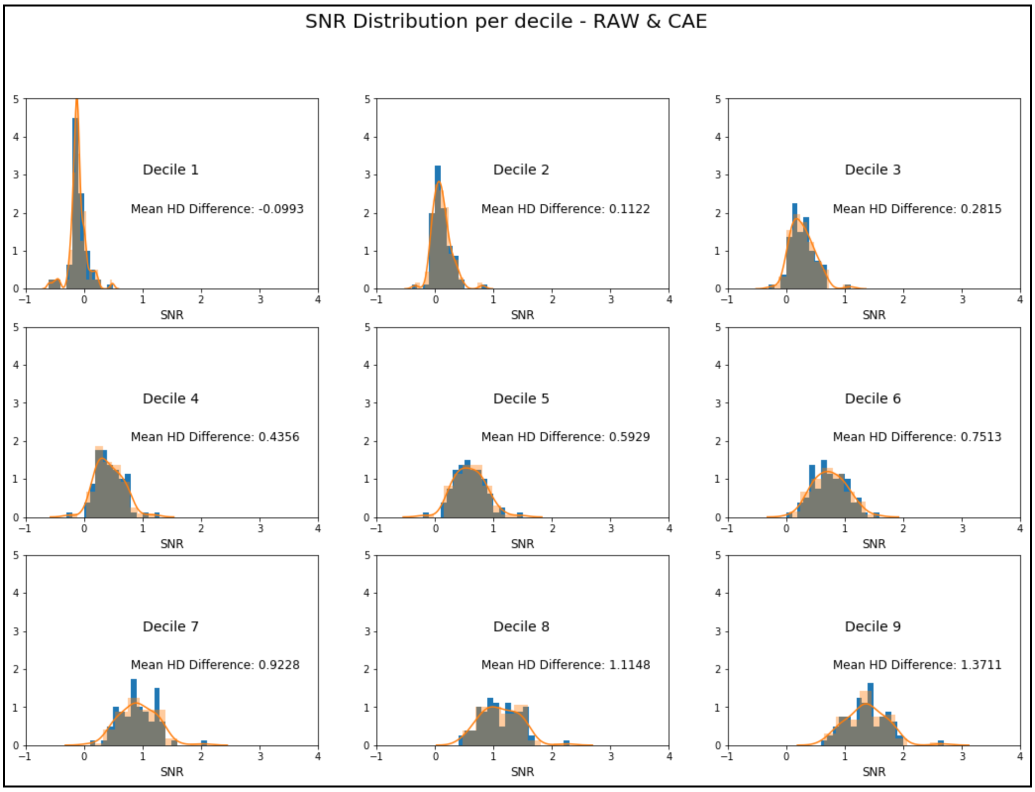
10 random RAW signals with CAE reconstructed signals overlaid

The figures below allow for the comparison of Harrell-Davis decile difference distributions between the SNR values for RAW, PCA and CAE signals. It can be seen in all figures that there is a larger range for higher number deciles than lower. This could prove beneficial for future investigations in EEG SNR distributions, as it can isolate particular deciles, and associate these with particular electrodes of the 10-20 system.

Condition 1:

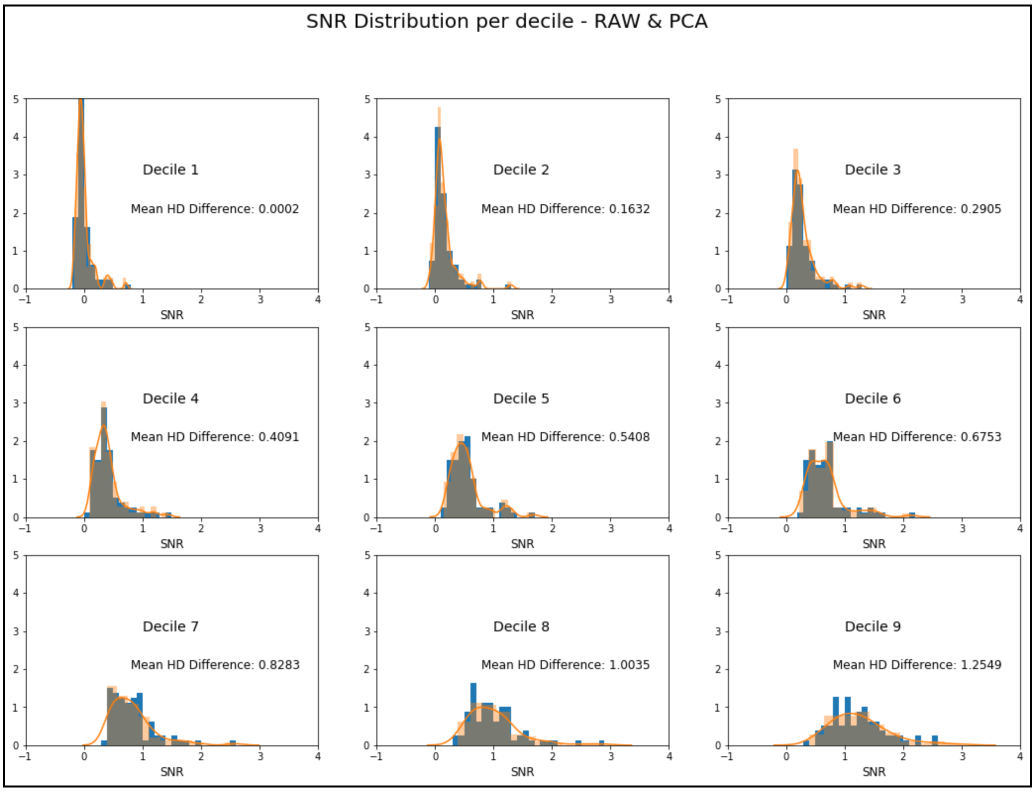


Distributions of Harrell-Davis decile differences RAW vs PCA – Condition 1

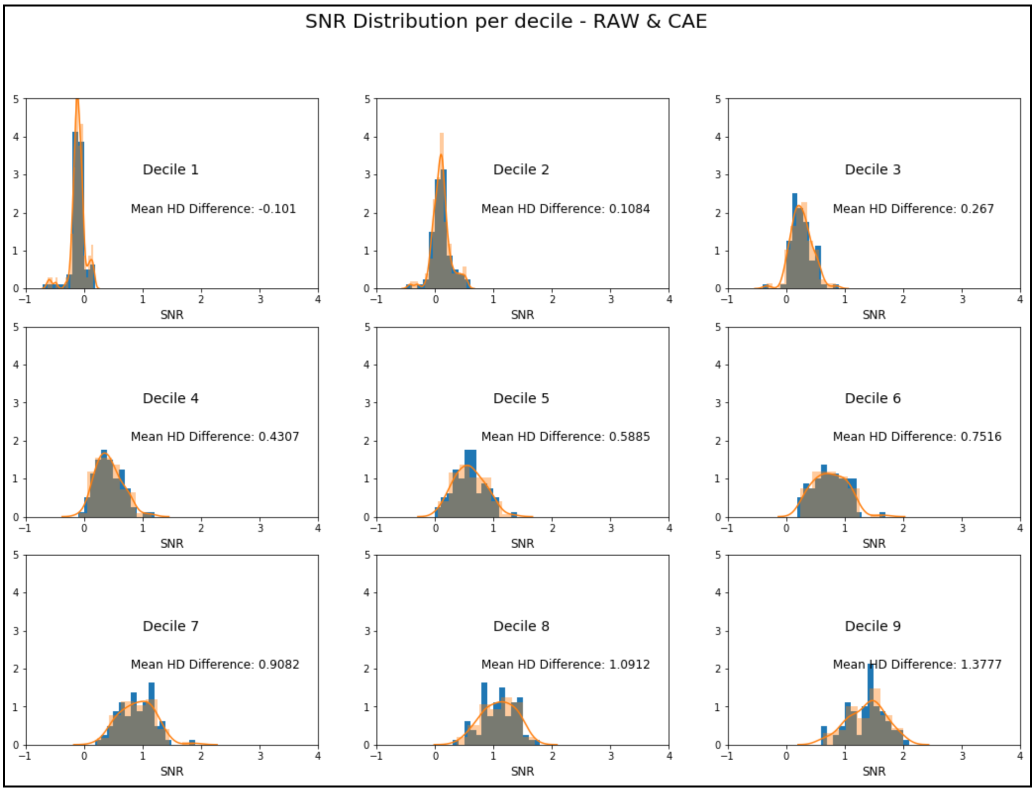


Distributions of Harrell-Davis decile differences RAW vs CAE – Condition 1

Condition 2:

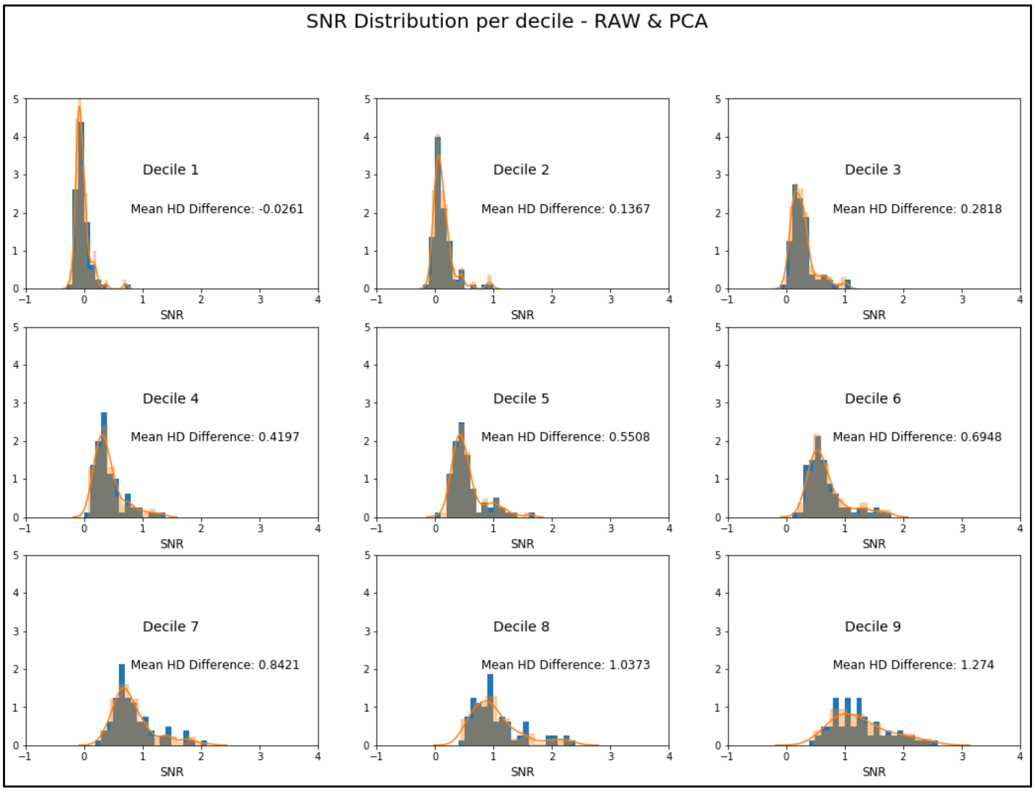


Distributions of Harrell-Davis decile differences RAW vs PCA – Condition 2

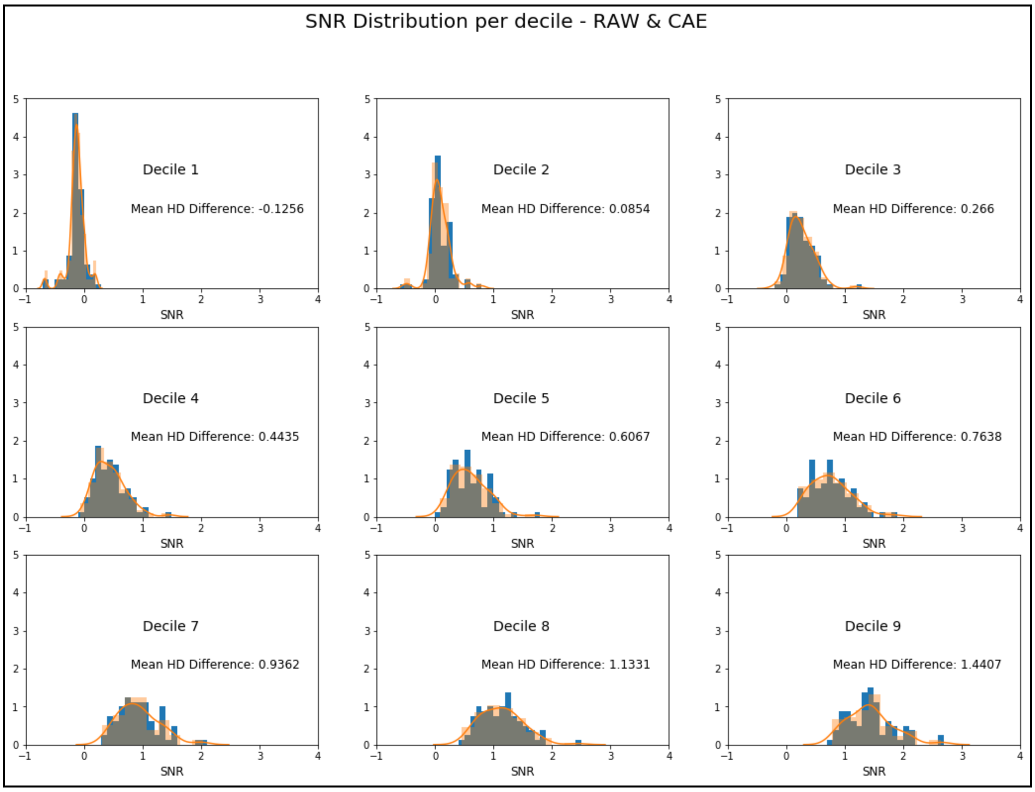


Distributions of Harrell-Davis decile differences RAW vs CAE – Condition 2

Condition 3:



Distributions of Harrell-Davis decile differences RAW vs PCA – Condition 3



Distributions of Harrell-Davis decile differences RAW vs CAE – Condition 3