2022

Evaluation Of Automated Eye Blink Artefact Removal Using Stacked Dense Autoencoder

Matthew Rigney

Evaluation Of Automated Eye Blink Artefact Removal Using Stacked Dense Autoencoder

Matthew Rigney

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

March 2022

Declaration

I certify that this dissertation which I now submit for examination for the award of MSc in Computing (Data Science), 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 text of my work.

This dissertation was prepared according to the regulations for postgraduate study of the Dublin Institute of Technology 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: Matthew Rigney

Date: 21 February 2022

Abstract

The presence of artefacts in Electroencephalograph (EEG) signals can have a considerable impact on the information they portray. In this comparative study, the automated removal of eye blink artefacts using the constrained latent representation of a stacked dense autoencoders (SDAE) and comparing its ability to that of the manual independent component analysis (ICA) approach was evaluated. A comparative evaluation of 5 stacked dense autoencoder architectures lead to a chosen architecture for which the ability to automatically detect and remove eye blink artefacts were both statistically and humanistically evaluated. The ability of the stacked dense autoencoder was statistically evaluated with the manual approach of ICA using the correlation coefficient, a comparative affect on the SNR using both approaches and a humanistic evaluation using visual inspections of the components of the stacked dense autoencoder reconstruction to that of the post ICA reconstruction where an inverse RMSE allowed for a further statistical evaluation of this comparison. It was found that the stacked dense autoencoder was unable to reconstruct random signal segments in any meaningful capacity when compared to that of ICA.

Keywords: Electroencephalography, Deep learning, Stacked Dense Autoencoders, Independent Component Analysis, Artefacts, Eye blinks, Human centred evaluation.

Acknowledgments

I would like to thank Dr. Luca Longo for his guidance, advice, direction and support throughout the course of this dissertation.

I would also like to thank all the staff at Technological University Dublin who have supported me over the past 2 and a half years. Their support and amazing ability to convey the knowledge required has helped me progress and improve. With their guidance, I have been able to achieve my goal.

Lastly I would like to thank family and friends who have supported me throughout my journey.

Contents

List of Figures

List of Tables

List of Acronyms

AE	Autoencoder
CC	Correlation Coefficient
CNN	Convolutional Neural Network
DLN	Deep Learning Network
DEAP	Database of Emotion Analysis using Psychological Signals
EEG	Electroencephalography
ERP	Event Related Potential
ICA	Independent Component Analysis
LSTM	Long Short Term Memory
LSTM-RNN	Long Short Term Memory - Recurrent Neural Network
PCA	Primary Component Analysis
RNN	Recurrent Neural Network
RMSE	Root Mean Square Error
SAE	Stacked Autoencoder
SDAE	Stacked Dense Autoencoder
SNR	Signal to Noise Ratio
VAE	Variational Autoencoder

Chapter 1

Introduction

Electroencephalogram or more commonly known as EEG, are a non invasive means of recording the brains activity using electrical potential to determine stimulation and its recordings allow for a greater understanding of how and where the brain is stimulated. Understanding of EEG and its implications was seen in Binnie and Prior, 1994 where neural activity was refined as a spatiotemporal average of potentials. Brain activity is an insight to a human organism that is not fully understood. It is interpreted rather than definitive. Ability to associate location, temporal characteristics and magnitude with an understanding of Psychology, NeuroScience ,emotions and mental states is truly impressive and a valuable asset for diagnosis and research.

1.1 Background

EEG recordings allow for the brain activity to be analysed and understood. Collection of such activity is achieved using probes that are attached to a subject, these are non invasive and given the non invasive nature means they are inherently more accessible. In Ball et al., 2009, this work showed the performance between invasive and non invasive approaches and highlighted how the invasive approach had better

signal quality but it was subject to the same issues when it came to the presence of noise and artefacts. Invasive approaches require implantation so their use carries risk but still prone to the same humanistic artefacts and to a lesser extent, environmental noise and artefacts. A reference to noise is quite broad and its meaning within research requires clarity as to how noise is understood, what form it takes and how such form is accounted for in research in its processing out or acceptance of presence. Artefacts are a form of noise in that they are an unwanted presences on the desired signals. Origins of which are either the environment or the subject. Such an understanding was stated in Rashid et al., 2020, "There are two sources of EEG artifacts: external or environmental source and physiological source" which highlights these 2 sources. These range from line noise on the probes to muscle twitching of the subject. All of which are inherent and can be suppressed but not eradicated.

The presence of noise in signals is ubiquitous but its affect ranges. In some sectors, the noise is dwarfed by the signal and its affect is minimal and such that, the noise is less impactful. With EEG signals, given that the voltages are in the micro volts range, noise is far more impactful when utilising readings at such a small scale. Removal of such noise is impossible but the removal of some distinct forms of noise is possible. Noise falls into the realm of incoherent and coherent, incoherent noise looks like a distorted haze of amplitudes with no characteristics apart from its randomness where as coherent noise can resemble signal characteristics. These are both termed as noise but this term is only consistent with the incoherent aspect, the coherent noise is seen as signal artefacts but these artefacts may be seen as signals in another area of research if this is what the research seeks to find. In EEG analysis, the coherent noise is muscle twitches, eye movement and eye blinking to name but a few. These are humanistic responses to life and can not be eradicated at source, instead these need to be considered and accounted for. Such accountability comes from the ability to detect and remove such presence. This can be further refined to the two possible means of this, namely manual and automated.

An ability to detect means there is a requirement to learn. A manual approach requires

observation that may vary across observers and there is an inherent ambiguity when trying to classify noise or artefacts from a signal manually. Deep learning allows for internal structural learning using the concept of features and a degree of relevance that can be stipulated in its structure. Leveraging of such can be achieved with the right choice of Deep Learning architecture that is designed to constrain the input data to relevant features that automatically removes aspects such as noise or artefacts.

The aim of this research was to determine whether an automated approach that takes EEG data and reduces it into a smaller feature space to then reconstruct this from its internal representation can in some meaningful way, remove a specific type of artefact while retaining the signal information.

1.2 Research Project/problem

Signal preservation is the core responsibility of any work that tries to remove unwanted elements that affect it's collection. Such an understanding requires that the signals information needs to be retained while the area of artefacts are addressed. The range of artefacts is large and their subjective characteristics match this range so accounting for all artefacts was not possible.. For this reason, a specific artefact was considered, namely the eye blink artefact due to its distinct characteristics and visual profile that makes it visible to an untrained observer.

Removal of such artefacts requires a means of detecting their presence. An ability to determine what is to be retained in the original signal is tied to the ability to determine what is to be removed. Deep learning techniques provided the best means of automatically determining this as these work from the concept of feature extraction and an inferred importance. Autoencoders are a type of Deep Learning architecture that do just this. An ability to constrain data in a internal bottleneck forces a learned representation which is then reconstructed from the constrained representation. Such architecture can be further improved by using Dense layers which accounts for the

interconnect-ability of the brain and means such behaviour can be captured in the internal layers. This can be then heightened by stacking these layers to allow for complex features to be formed in the internal layers as the internal layers constrain the dimensionality of the data. A Stacked Dense Autoencoder allows for the evaluation of an automated eye blink artefact removal technique using no more than its internal structures to extract features and within this, remove eye blink artefacts with the perception of such artefacts failing to be seen as relevant features at a certain point in the latent representation.

Evaluation of such an approach required a baseline to compare with and a means of validating that the desired artefacts were firstly present in the raw data and subsequently removed in the reconstruction. The manual approach of artefact detection allowed for such an understanding, namely ICA. Components were identified and removed manually using the distinctive characteristics of the Eye blink artefacts. ICA is both a baseline and a methodology for detecting eye blinks in the raw data and the stacked dense autoencoder reconstruction.

1.3 Research Objectives

The objective of this research was to determine whether a Stacked Dense Autoencoder architecture can remove Eye blink artefacts from the EEG data without explicitly being trained to do so. This feeds into several objectives that form the top level objective. The first being an understanding in the number of internal layers of the stacked dense autoencoder required to provide meaningful reconstructions within a defined scope of the research. The second being the ability of the stacked dense autoencoder to reconstruct relevant data, in the architectural comparisons this related to its ability to reconstruct the raw data and in the comparative evaluation with the manual approach, its ability to reconstruct a signal that is similar to the manual approach. The third being the extent to which the stacked dense autoencoder can

remove eye blink artefacts via manual inspection of the components that make up the reconstruction.

1.4 Research Methodologies

Methods used seek to ensure accuracy and reliability were core when carrying out the research. To achieve this, an understanding of what the methods were needed to be established. This research was Secondary in nature. The data collected was not that of the researcher and was generated for alternative means by others. This data allowed for research built upon a systematic review of existing research to be achieved. Building on the research element, this was empirical research that was validated using quantitative and qualitative means as the work set out an hypothesis that was tested using the scientific method. The quantitative nature was due to the numerical characteristic of EEG data and how reconstructions were numerically compared for similarity and significance that was determined from these when evaluated while the qualitative nature was in the visual inspection of the components that determined the presence of eye blink artefacts via manual classification against defined profiles. With the generation of a hypothesis guiding the research, this meant that Deductive reasoning was used as the results were inferred. A hypothesis was stipulated and then quantitatively and qualitatively determined. There exists no ground truth with EEG data so there was no perfectly clean signal which meant all results were inferred as their were assumptions required at stages in the research that impact the evaluation.

1.5 Scope and Limitations

The scope of this research was the automated removal of eye blink artefacts using an stacked dense autoencoder architecture on EEG pertaining to 20 participants, 40 videos per participant and 5 time segments per video of the data. The stacked dense

autoencoder architecture was first established using a refined comparative study of 5 defined architectures and then a chosen stacked dense autoencoder architecture has its reconstructions of the raw data evaluated. Evaluations were carried out both statistically and humanistically for the presence of eye blink artefacts while being comparatively evaluated to the manual approach of ICA.

An inherent ability of stacked dense autoencoder's to learn and potentially remove eye blinks while retaining signal information would be advantageous as this could be implemented as a preprocessing step that would remove the need to manually determine the presence of eye blink artefacts and remove the inter and intra variability in distinguishing what was an eye blink and what was not. This can be further exacerbated when consideration was given to the varying types of eye blinks that have different temporal characteristics and varying degrees of amplitude. Eye blinks are non deterministic so profiles were used to try to counter the ambiguity in visual evaluations of the eye blink artefacts.

Data contains both stimuli and non stimuli segments. Stimuli exposed segments relate to the 60 second video clips for which there were 40 for each participant. Within these videos, the data was further segmented as part of this research to look at shorter time segments to see how the stacked dense autoencoder reconstructs across varying lengths. There were 5 segments determined which further segmented the data. With this extensive segmenting of data, it was unattainable to check every reconstruction for the presence of eye blink components due to the exponential growth in the number of segments to be evaluated. Performance metrics for determining the stacked dense autoencoder architecture were depicted using averages of averages while the presence of eye blink artefacts in stacked dense autoencoder reconstructions were determined using a randomizer that checked a random participant, a random video of this participant, a random time segment within this video and a random time slice of this specified segment. This random data segment specified was repeated extensively and used for the evaluation of the manual and automated approaches.

Selection of the stacked dense autoencoder architecture was determined based on variations in its internal structures which related to the amount of neurons in the bottleneck. There were several other elements that play a role in its performance such as hyperparameter optimization but these were not considered for adaption and instead their values were determined based on values found in the existing literature.

EEG data has a fundamental downside when it comes to its utilisation and that is there exists no ground truth. There is no clean signal as all data was acquired with the same susceptibility to both environmental or participant contamination. For this reason, the non stimuli segment that makes up the first 2 minutes of the recordings was assumed to be the baseline noise for the research. Another fundamental consideration with the data was the unquantifiable nature of information loss on EEG signals as there was no definitive point where noise was present on the signal, it was understood that noise is always part of the signal but the degrees of its presence varies. This meant the loss of information through dimensionality reduction can not be quantitatively determined.

Working with participants and EEG probes that were susceptible to noise meant there was an inherent inter and intra subject variability. This makes reproducability across participants of the research more complex as even slight variations can affect the reasoning for inferences made when determining the presence of eye blink artefacts.

The research sought to determine capability and extent in removing eye blinks automatically, however, this comes at a cost of explainability. The manual method of ICA has a mathematical means of determining decomposition's of the signals and the removal of eye blinks was then done using an interpretable template that guided the decisions but the stacked dense autoencoder provided a reconstruction and the reasoning's for each decision made that directed the neural network to generate the reconstruction was unknown. All that was known was the reconstructed signal from the input signal that had a constraint placed on it by the dimensions of the bottleneck. The stacked dense autoencoder returns an output that was the input with some loss, the latent representation was features extracted from the data but reasoning was not

attainable, only inferred.

1.6 Document Outline

1.6.1 Chapter 2 - Review of Existing Literature

In this chapter, the reasoning behind the research was unearthed. Existing literature allowed for the highlighting of what has gone before, what has worked, what has failed and what was to be learned. The literature review for this research looked at existing work in Autoencoders and their implementation in EEG and specifically for artefact removal. Consideration to all facets was evaluated from how approaches determined success, how data was manipulated and what gaps exist or persist.

1.6.2 Chapter 3 - Experiment, Design and Methodology

The definition of the planned experiment, how this was shaped and how this was carried out were detailed in this chapter. This details the steps that were implemented so as to ensure the research question can be answered and the null hypothesis either accepted or rejected.

1.6.3 Chapter 4 - Results, Evaluation and Discussion

This chapter details the findings from the planned experiment. This looked at the underlying reasons for the results and how these were interpreted in the context of this research.

1.6.4 Chapter 5 - Conclusion

The final aspect related to the review. The conclusion detailed the relationship from experiment to outcome, the nuances that make up all facets of this work and the lessons to be learned.

Chapter 2

Review of existing literature

In this chapter, the literature review looked at all aspects of EEG, features, artefact detection, deep learning, source separation and Autoencoders which showed what research has preceded this work in the area of automatic artefact removal.

2.1 Electroencephalogram

2.1.1 EEG Analysis

EEG signals are the manifestation of probing the scalp so as to detect small voltages that are interpreted as brain activity. These signals reflect the human brain and its working state as was highlighted in Safayari and Bolhasani, 2021 where EEG allowed for diagnosis of depression. The nature of EEG signals are that of highly non-linear, non-Gaussian, random, and non-correlated. Such a concept was built upon in the work of Lujan et al., 2021 that looked at the processing of EEG signals using a full scale review of depression diagnosis. This work not only highlighted the time, frequency and spatial domains that hold EEG information but also the susceptibility of the signals. EEG data is non linear with activity that is sporadic and resolution dependent on the

extent of spatial density of the probe placement on the scalp. Their understanding can vary depending on the domain they are analysed in, namely the time, frequency, or space domains.

The utility of these non-invasive signals are that they allow for understanding of brain activities so as to allow for understanding of one's emotional or mental state and diagnosis of conditions that are present or detectable in the brain. To utilise such signals, these first must be acquired and such acquisition is achieved using EEG probes placed on the scalp which record the electrical activity. The electrical activity recorded is not without its challenges. One such challenge is the susceptibility to many forms of noise which is acquired with the desired signal.

With such susceptibility to noise that can be reduced by additional care but not eradicated, models and architectures that utilise EEG signals have to have some means of dealing with such disturbances in a way that deals with the noise without compromising the signal. This is a balancing act and one that needs to be understood as searching for zero signal loss is an unattainable goal but wilfully removing a signal and corrupting the integrity of the EEG signal means that the task is proven wasteful. With this in mind, Bigdely-Shamlo et al., 2015 looked at early stage processing and its goal of noise reduction. The research highlighted the consideration of over processing the data which means it may lose some of the original signal or become too defined for a specific task, that it can not be generalized so it's utility is diminished outside of the single participants use case. The worked carried out in Rakhmatulin, 2020 which looked at data preprocessing and considered the merits of feature extraction and feature selection methods. These will be looked at in depth in a later section but the preprocessing considerations were important as this allowed for the first step in any architecture or methodology that tries to utilise EEG signals, to be built on a solid base with an understanding of how such aspects can remove signals and if this is desired as a means of removing a greater degree of noise via such tools as filtering, then this is done with a balanced understanding of the gains and losses.

The use case of the EEG signal plays a large role in how it is understood. Signals can be represented and analysed in various ways. Analysis can be categorised to the time domain, frequency domain and time-frequency domain which are linear analysis methods. Expanding on this to a fourth category of non linear methods. In Agrawal et al., 2021, the research looked at the complexity of the EEG signal and how representation and analysis was best evaluated in the time-frequency domain. This was further built on when consideration was given to the industrial application as seen in Min et al., 2016, where understanding was given to data resources, interpretability, selection of architectures and hyper-parameters to balance task specific and multi task specific outcomes.

EEG signals are non stationary and random in nature. This is further exacerbated by the fact that these are inherently variable across subjects too, the inter and intra variability makes EEG complex when interpreting their meanings. Such meanings are taken from the patterns detected in the data which come in the form of features and such features are relevant to the use of EEG data in any type of model or architecture developed.

2.1.2 Feature Selection

Building on the concepts of the practical use of EEG, how it is understood and how it is utilised, means that consideration needs to go towards what the EEG contains. What these electrical variations are trying to depict and how it is to be interpreted. This can be seen in Lujan et al., 2021 where feature can be defined as a unique characteristic that allows one to understand the neural activity and assess the state of the brain. Methods exist that allow for EEG signals to have its underlying attributes understood. These are in the way of features selection techniques as they represent information that is non redundant. The process of selecting such features means these features exist and in turn are utilised using statistical or interpretable means to justify their selection. In Musa, 2013, this work looked at several feature selection methods

using statistical analysis on their use to determine their utility. Statistically significant results were displayed in a ranking system which showed the best performing method based on the methodology used to select the features present in the EEG signal. The work seen in Kartika Delimayanti et al., 2020 showed the extent of features that could be present in EEG signals with a high dimensional feature matrix but this was seen as wasteful from a resources standpoint and highlighted the numerous consideration that are required when it comes to implementing a methodology that is not just bound to using as many features as possible, but considers the quality of the features and not just the quantity.

2.1.3 Feature Extraction

Building on the quality consideration of features, leads to how the features are actually extracted. Features are inherent in the data so they are tasked with encapsulating information. This builds on a concept of redundancy, to use less resources without losing information. Features are new data points that are projections of the original data with the goal of information retention but in a lower dimensional representation. The task of feature extraction falls into 2 categories, manual and automated. Manual feature extraction is built on the concept of some form of transformation that is supplemented with domain knowledge so an understanding of the features guides their utility. Automated feature extraction allows the model to infer importance on features so this facilitates a reduction in the required data.

The work seen in Mohammadi and Mahmud, 2019 looked at feature extraction and some of the existing research from the time, frequency and time-frequency domains. Desirable features are to be relevant, interpretable and discriminative. It was shown that the time-frequency domain allowed for the most relevant and discriminative features. In Qu et al., 2020, the work looked at using ICA to obtain pure signals whose energy was extracted as features while Mahdavi-Nasab, 2010 looked at feature extraction using the power spectrum density and statistical methods to reduce the di-

mensionality of the data. Utilising the Welch method for PSD as it produced strong features when evaluated.

Building on the understanding of varying feature extraction methods, Herrera et al., 2013 looked at combining feature extraction methods for feature generation. This successfully leveraged Hjorth features, Wavelet transform and symbolic representation. Some further research which considered other aspects such as spectral information was seen in Windrim et al., 2019 where the research looked at separability of features using spectral measures (SID and spectral angle) to extract features on high dimension and high variable data as opposed to standard square-error loss methods which led to a more discriminative feature representations. Separability was determined using Fisher's discriminant ratio. Building on the concept of feature combinations, W. Liu et al., 2016 looked at EEG and eye signals as bi-modal aspects that are complimentary within the emotion recognition space and leveraged this to extract high level features. It's utility was verified in the research using a confusion matrix to infer a desired outcome in the classification task. Complimentary aspect leveraged the good and bad elements of EEG and eye signals when it comes to conveying emotion. In K. Li et al., 2013, the research looked at effective state recognition. The research used a deep belief network to extract low dimensional features of each channel while retaining the channels characteristics which is the desired outcome for any feature extraction method. Built on a concept that meaningful critical channels can be selected for the task of effective state recognition across different participants.

Features have a direct impact on what can be inferred from the data, how these are generated or restricted impacts the outcome, irrespective of how these are utilised beyond this point.

2.1.4 Noise and Artefacts

The ability to deduce meaning from features generated allows EEG data to have levels of abstraction in it's meaning. Such an understanding indicates that features can be refined from an initial understanding to a more metadata type understanding where data represents an already constructed representation of the original data. This is better thought of as more complex representations of the data where complexity can be extended. Such a concept is important when consideration is given to what the EEG data actually contains. EEG data contains brain activity along with both noise and artefacts which are ubiquitous in EEG data.

Manifestations of noise and artefacts resemble high/low frequency spikes on the signal, image distortions, low frequency drifts or periodic fluctuations. Mitigation can be in the way of ensuring the acquisition of the data is analysed but this can only go so far. The overarching concept is the identification of noise and artefacts which would indicate that there is a means to differentiate what is a signal, what is noise and what are artefacts. This differentiation is further complicated by variability across the brain in relation to its activity. The work seen in Majoros et al., 2019 showed the variability across channels where several orders of magnitude difference in amplitudes for some channels was observed relative to the baseline noise. Channel wise variability shows the susceptibility as such variations are relative to the number of channels used in the collection of the EEG data and their respective location at a given time so its impact varies not only spatially but temporally. Ability to differentiate even with spatial variability in activity has several methods, the first of which is the attenuation methods which was stated in Gramfort et al., 2013 where the specific artefacts are detected and then set to zero before signal reconstruction, the second being the exclusion of contaminated segments where the entire segment was removed and the third being evident in Rashid et al., 2020 where preprocessing techniques allowed for better feature separability and this improves the ability to distinguish between noise, artefacts and signals so a more refined adaption can take place with isolated corrections.

With the susceptibility of EEG data to being corrupted with noise due to its low voltage range when acquiring the signals, separation of these is integral. The importance in being able to differentiate between signal and noise was highlighted in Rashid et al., 2020 where the research stated, "A small SNR and different noise sources are amongst the greatest challenges". By separating the noise and artefacts from the signal allows for their treatment while minimising the modification of the underlying signal.

2.1.5 Source Separation

Separability allows for the ability to distinguish. Distinguish between what is perceived as desired and undesired. Rooted in its utility but key in how signals can be separated to their individual components.

A large amount of research to date has looked at Independent Component Analysis and it's ability to separate signals. In the research carried out by Kachenoura et al., 2012, existing algorithms for ICA were researched and their utility in specific instances for removing noise were explored. This research highlighted how ICA attempts to extract underlying signals that are combined within the EEG signal. An assumption is made that there is a linear mixing of underlying signals in the EEG signal. ICA is a computational method for separating multivariate signals into sub components. Building on separability, Zhang et al., 2015 looked at artefact separation using a combination of Discrete Wavelet Transform and ICA where validity of the method was assessed using correlation analysis and the subsequent performance of the classifier. In Phadikar et al., 2020, a similar decomposition and classification construct was carried out but this time it used the classifier to detect eye blink artifacts which were then fed to an autoencoder to correct. The corrected signal was then reconstructed using interpolation and inverse ICA where the system was evaluated using the degree of similarity between the signals (correlation coefficient), Mutual Information, Normalised Mean Square Error and Structural Similarity Index Measure (SSIM). The research built on a comparative look of the signal before and after some independent

components were removed. In Glass et al., 2004, research tried to use templates of eye blinks so these artefacts could be detected and removed. ICA was again utilised for signal decomposition. Utility of the independent components was also seen in Raduntz et al., 2015, where Linear discrimination Analysis leveraged the feature vectors from the ICA components. ICA has extensive use due to its ability to demix signals but this requires domain knowledge to infer meaning to these independent components. An alternative method to this ability to separate is the ability to infer importance in the features.

The work done in Kumar, 2021 looked at the use of PCA and how it utilised a concept of statistical importance to determine feature importance. This allows for the reduction in the number of features that represent the original data. The concept of importance in features is key but the means of determining importance is the critical factor. When using PCA, the method looks at the variability while with ICA, the method looks at separability. With such a basic understanding in mind, ICA's ability to separate is desired as the detection of artefacts requires the ability to separate it from other components that make up the signal. An ability to manually inspect the signal components allows for the manual detection of artefacts using templates of such artefacts.

Implementations of ICA has been seen with both manual and automated approaches. Evidence can be found in Raduntz et al., 2015 where a direct comparison was used to compare a subjective take on both methods. This work highlighted the ability of both to generate similar outcomes and also highlighted the requirement for domain knowledge experts in the manual implementation of such. Work carried out in Chaumon et al., 2015 highlighted that both methods had an element of variability and there was a limitation on the precision expected. Such limitations in precision were further highlighted in Rashid et al., 2020 where the research focuses on the ICA's ability to disregard both the temporal and spatial relationships between sources which results in relevant information being lost. Disregarding was not subjective to the approach used but the methodology as a whole. Other elements that affect ICA decomposi-

tion's, irrespective of its implementation lay in the data that is fed into the algorithm. In Zakeri et al., 2014, the work showed that the efficiency of the ICA algorithm was affected by the preprocessing steps it implemented and not the approach taken for the ICA decomposition's. A considerable dependency of the preprocessing of the data that manifested in better decomposition's in the algorithm when conveying the components of the input signals was seen in this work.

Building on the affect of preprocessing of the data on the quality of the ICA outcome, the work done in Klug and Gramann, 2020 looked at the affect of movement in EEG experiments, the number of channels and the filtering cut-off's. The research looked at the affect of these and provided recommendations for optimum values to allow for better separations on density specific recording and highlighted use cases, one of which was the use of ICA for eye and muscle artefact detection. The work of Jas et al., 2018, looked at the use of ICA with eye blink and heart beat artefacts, highlighted the prototypical spatial patterns of these artefacts which supports the understanding that these can be noticeably detected using ICA.

In Pontifex, Miskovic, et al., 2017, the work looked at existing manual approaches for ICA and tried to build on this with an automated method that leveraged templates of eye blinks to be used as a means of detecting the artefacts. Although the work was using an automated method, the use of templates to determine the presence of such artefacts was a valid consideration. Within the work, assessments of the spatial features, temporal/spatial features and time series features for the detection of eye blinks were obtained. The scale of amplitude and robustness to noise played a key role in determining the desired approach but could not explain the variability of ICA and how this influences the reconstructions post ICA. Removal of eye blinks may also remove some signal information and appears to be related to the scale of the amplitude in creating a distinctness in the artefact. The work carried out in Pontifex, Gwizdala, et al., 2017 supported such an understanding as it did a repeatability approach to assessing the variability in using ICA but even with 30 iterations, the variability in ICA solutions ultimately influences the EEG signal when eye blink-related activity

were removed. Evaluation of the various algorithms available for implementing ICA was evident in this work. Speed of convergence using FastICA was far better than the likes of Infomax and Second Order Blind Identification (SOBI). FastICA showed a higher degree of variability but was not excessive.

An understanding that there exists a means of distinguishing desired from undesired signal components means that automated approaches can be evaluated using such manual approaches for both its ability in removing eye blink artefacts and as a baseline for statistical evaluation.

2.1.6 Evaluation of Artefacts

To evaluate artefacts, the capability to define them must exist. With such definition, these can then be dealt with and evaluated using the outcome of their utility. A requirement for the type of artefact that is sought must be specified as each artefact has a characteristic that allows for its detection. Such characteristics are non deterministic but stand as a guide for their detection.

The work seen in Zavala et al., 2020 and Pontifex, Miskovic, et al., 2017 looked at manual classification of the eye blink artefacts using defined profiles that convey the characteristics such artefacts. The most compelling research found was done by Roy et al., 2014 where the research set out to utilise the otherwise discarded ocular information that was obtained as part of the non psychological signals. These are additional probes placed to supplement the psychological probes that are placed on the scalp. Such ocular information can be used as a support to evaluate an outcome and in this work, these were used to supplement the understanding garnered by the correlation coefficients. The ability to profile the eye blink artefacts and subsequently support their presence using ocular channels further supports their detection.

By having defined profiles of eye blink artefacts means manual inspection for the presence of eye blink artefacts can be achieved. Automated approaches can utilise this

where their reconstructions can be examined to evaluate the success or lack there of when it comes to automatically removing eye blink artefacts.

2.2 Deep Learning

2.2.1 Deep Learning in EEG

The concept of features representing data means that automated approaches must be able to generate these features. The work seen in Goodfellow et al., 2016 gives great depth to what Deep Learning is and how it leverages experience as part of its learning. No specific formal knowledge is passed and instead, knowledge is learned from hierarchical concepts. Hierarchy can be extended to incorporate further complexity which is seen as layers that extend the depth of the term deep in Deep Learning. The reason for its use is its ability to learn complex hidden patterns and relationships in the data without being explicitly being told this.

Building on the concept of hidden patterns and relationships within the data, means that such a concept is just another way of saying features within the data as features are hidden patterns and relationships that are formed using the data. Features are both high and low level features as there is an inherent relevance with features when used for a specific task. Their impact on the outcome determines their relevance.

It can be seen in G. Li et al., 2016, that deep learning has been around for several years as this review highlighted existing architectures and applications. The research highlighted considerations when using deep learning like data size and its effect on performance along with the susceptibility to noise. Deep Learning requires large swaths of data to fully utilise its performance but this comes at a cost which is paid by computational time and in many cases, the desired data is just not available so reduced performance was inevitable. In Langkvist et al., 2012, the research looked at feature learning using deep belief networks as a means of detecting sleep stage switching.

Variations in layers and neurons were considered and the impact on the performance of the classification. This work struggled to capture correlations between the inputs so its ability to establish relationships between channels was poor. In Sun et al., 2019, the work looked at using a variant of a RNN network called echo state for feature extraction. Leveraging the echo signals of internal neurons to recover signals where the output weights are the EEG features. Evaluation of its experiments were somewhat skewed as the relevance of the design seemed to be tied to a specific experiment that was hard to replicate on similar data. This can be understood better in Jahankhami et al., 2006 where the research looked at feature extraction using Wavelet Transform but used statistical methods to compare epileptic and non epileptic signals that were extreme cases. The distinct nature of extreme cases was not a practical base as the utility of models, are based on its ability to differentiate between the cases that lay close to some boundary.

Looking at the idea of significant improvement, it can be seen in Jia et al., 2014 where the effective state recognition features were extracted using unsupervised learning to determine the features and these were then refined using supervised learning for the features selected. In Cho and Jang, 2020 which looked at different combinations of inputs such as images and network structures for the detection of seizures and in X. Xing et al., 2019 which looked at a framework of a stacked autoencoder for feature extraction and an LSTM-RNN for classification. All these determined significant improvements by using comparative models as baselines. The evaluation of such comparative views were carried out using standard statistical analysis on classification tasks such as accuracy and in the case of Cho and Jang, 2020, statistical significance was determined using the ANOVA and Tukeys post hoc test. It can be seen in X. Xing et al., 2019 with a bit more depth by evaluating each part of the framework, mean square error evaluated the training of the autoencoder, adjusted-R evaluated the testing of the autoencoder and Pearsons Correlation Coefficient determined the inter channel correlations. The comparative consideration was the best means of determining whether the research was successful but this comes with risk as the subjective

use of models means that there may be a meaningful improvement but this may only be in isolation. Again, looking at the work in X. Xing et al., 2019, a Support Vector Machine and LSTM classifier was used to validate the classifier while ICA and SAE validated the SAE. These all have various means of assessing the models but it does not allow for a direct comparisons as some require humanistic interaction to determine relevant features and these may vary from experiment to experiment and researcher to researcher.

The work done by Takahashi et al., 2020, dealt with a concept of patient specific autoencoders to alleviate an inherent issue with binary classification that resulted in false positives that was problematic given the models expected environment. The research looked at creating an additional label that flagged abnormal along with seizure and non seizure. This reduced the models accuracy but had the desired goal of higher precision which highlights the considerations when it comes to determining the success of a model. Similar work seen in Hartmann et al., 2018, looked at spectral features and their internal representations. This work showed higher sensitivity to EEG phase in the early stages of the CNN and higher sensitivity to EEG amplitudes at later stages. While Suryawati et al., 2021 showed the problems with random initialization of Generative Adversarial Networks for feature learning. Such limitations were overcome by using the encoder part of a autoencoder to initialize the neurons. Both studies showed the nuanced nature of determining success, success is subjective and can be manipulated to solve a specific task.

This subjective nature meant that understanding was required which can be refined to the concept of explainability. The work done in Sturm et al., 2016, looked at using heat maps to indicate relevance of data points on decision outcomes while Ellis et al., 2021 looked at explainability of local spectral features to understand how clinical variables such as age and gender affect the outcome and the learned patterns. Explainability is a popular research area in recent times but there is no defined means to explain how such architecture came to the outcomes they relay, this means understanding is limited even if the desired outcome is achieved. A means of assessing the outcome

both though its architectural impact and evaluating an understanding of its outcome is seen as desired. With such evaluation and understanding, the concept of improvement or decline between 2 approaches can be evaluated even if understanding in their construction is not clear and transparent.

2.3 Deep Learning and EEG for Artefact Removal

2.3.1 Autoencoder's

Features represent information in the data. The ability to refine the data to convey a perception of importance in features would allow for an understanding to be drawn as to what is seen as relevant and not relevant so as to allow for the data to be constructed of relevant features and in turn, remove aspects of the data that do not contribute to the representation of the relevant information contained in the data. This leads to the utilization of Autoencoders who use a smaller latent representation to represent the data and it's representation can be determined by its architecture that constrains using a concept of relevance.

Autoencoders are a type of unsupervised artificial neural network. The work seen in Schmidhuber, 2015 looked at Deep Learning in neural networks and how there is an inherent causal affect of interconnections when layers are both interconnected and stacked. Their design consists of an encoder section, a bottleneck section and then a decoder section, all of which can be configurable in their size to manipulate the input data. Their construction is such that they learn a lower dimensional representation of higher dimensional data which was constrained by the size of the bottleneck and its performance determined by how the encoder and decoder are constructed regarding to the type of layer used and the number of such layers. The configuration determines features that are created to represent the input data in a lower dimensional space and this is used by the decoder to reconstruct the data to the original size. This ability

to reconstruct the input from a lower dimensional representation is the crux of the architecture and key to its performance. Autoencoders allow for automatic feature extraction.

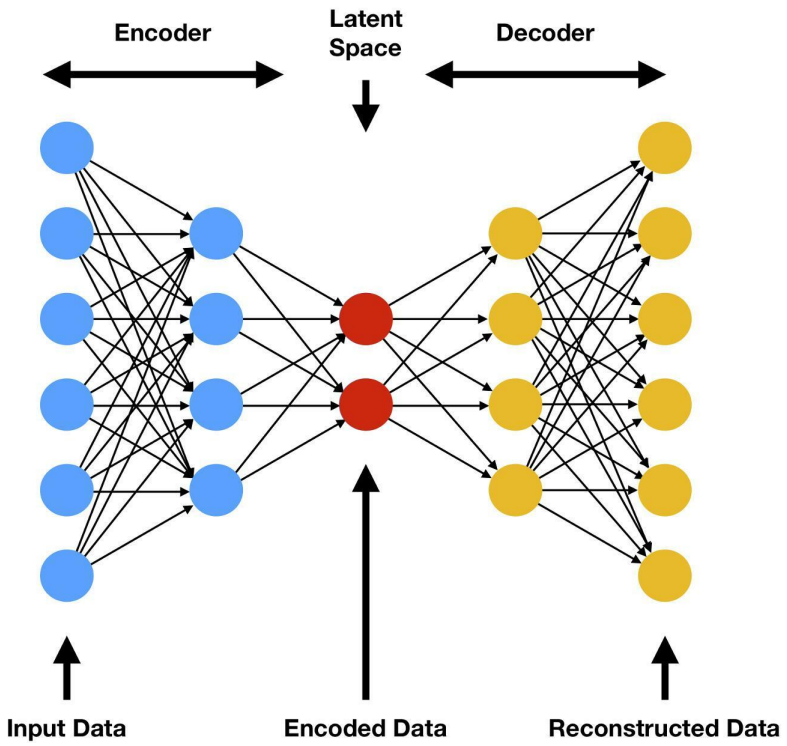

Figure 2.1: Autoencoder

The architecture of an Autoencoder is configurable and the performance assessed using the metrics that compare the input signal to the output signal. The loss within the architecture determines the success of the autoencoder and this is inherently tied to the architectures ability to learn complex relationships that are determined by its depth, type of layers and the bottleneck size as these constraints forces the Autoencoder to reduce its dimensions which comes at the loss of information if its constraint is too small and doesn't learn anything if it is not constrained enough.

The ability to find the optimum bottleneck size is bound to how this is evaluated. Evaluation comes in many forms for reduced dimensional representation. It was seen in

Jahankhami et al., 2006, Kartika Delimayanti et al., 2020, Langkvist et al., 2012, Musa, 2013, Riyanul Islam et al., 2019 that classification metrics were used to determine an architectures ability, these are the ability to successfully classify using accuracy, F1 score, sensitivity, specificity and AUC which were used to support their work. The work of Kumar, 2021 looked at signal to noise ratio as a means of evaluation while Kachenoura et al., 2012 looked at the behaviour of ICA methods as a function of the signal to noise ratio's. The concept of comparing the affect on the signal rather than the specific accuracy of its use case was further highlighted in P. Liu et al., 2019, Lee et al., 2021, Ghosh et al., 2018 and Phadikar et al., 2020 where these all looked at not only a variation of the error in its reconstruction but its correlation to the input and the extent of mutual information and structural similarity present between the input and output as a means of determining the extent of information loss. The lost information was unknown whether this was desired or undesired but its consideration was no less valid. Other methods were also seen in Babiloni et al., 2001 where the Mahalanobis distance was used. This is a distance measurement in multivariate space which aids in the ability to determine the reconstructions similarity to the original signal. Evaluation is dependent on the desired outcome and with this, how the reconstructions are perceived.

Research into the learned representations of autoencoders was seen inCharte et al., 2020, where the means of influencing this representation was evaluated. The learned features were assessed using RMSE, Mean Absolute Error and Mean Absolute Percentage Error. Independent analysis of the features to determine their complexity and separability were evaluated. The ability to comprehend complexity is important as autoencoders are computationally expensive. This was seen in Gogna et al., 2017 where autoencoders were used for simultaneous reconstruction and classification via a learned linear map into a classifier that was dependent on a class label where a reconstruction was provided if no class label was available and a classification was provided if the input data has a class label. The model used a reconstruction benchmark to determine the quality of the reconstruction and classification to determine validity

of the features. This model was an order of magnitude faster than existing methods used via comparisons of computational times. Leading on from the concept of reconstruction, Ghosh et al., 2018 looked at artefact identification and correction using a sliding window approach that was fed to a Support Vector Machine model to identify artefacts which the autoencoder corrects in the signal. Statistical methods were used to determine what signals were contaminated, the model showed higher accuracy but only for distinct eye blink artefacts.

One area where improvements can be achieved is with the stacking of autoencoders as this allows for more complex patterns to be detected as the addition of layers and neurons means the addition of parameters which allow for more complex functions to be fitted. In Jirayucharoensak et al., 2014, the work looked at SAE's within a DLN for feature extraction and PCA for feature reduction with the non-stationarity of the EEG signals being handled by a covariance shift adaption method. This allowed for an improvement in the classification of emotion. The research carried out in Ferri et al., 2020, looked at using SAE's and an alternative method of class specific autoencoders to discriminate between dementia conditions of Nold and ADD using both resting state EEG and sMRI. SAE showed better metrics when assessed comparatively. Another area where SAE's were utilised were in the detection of the P300 component using an ERP experiment as seen in Vareka and Mautner, 2017. The ability of the stacked model to deal with spikes in activity better than other approaches was important as such data is thought of in bursts so the ability to deal with data that exacerbates the non stationarity of EEG data is indicative of a model that generalises well.

Other variations and structures of autoencoders were seen in Riyanul Islam et al., 2019 where convolutional layers were used in the autoencoder for automatic feature extraction and its utility was compared with manual techniques for feature extraction while **Di2020** looked at using Variational AE's to examine a hypothesis in the research that there are variables that each participant inherently have when it comes to emotion. The research showed some interesting results in that neural networks were better at modelling and decoding brain signals than ICA, traditional AE's outperformed VAE's

for data reconstruction using Pearson Correlation Coefficient as the evaluation metric and VAE's outperformed traditional autoencoders in decoding latent factors so it could infer emotional state's using more effective information. The traditional autoencoders outperformed the VAE's in reconstruction which was quite telling as the VAE's makes some assumptions about the data and the complexity involved but this research may not have been the best use of VAE's. The construction of the autoencoder network is integral in its utility and was evident in Wen and Zhang, 2018 where the dimensionality of the feature space was too constrained, highlighting difficulty in extracting effective features from EEG signals.

Building on the concept of effective features in Autoencoder's, the affect of noise on these is one factor that needs to be evaluated. Existing research considered noise reduction by looking at it from different perspectives. In Lee et al., 2021, the research looked at learning the noise and subtracting this from the signal rather than learning the signal and removing the noise. This research showed improvements in signal restoration, symbol demodulation and precise localization's using MSE, SER and localised error respectively. Some research looked at a plethora of denoising techniques, seen in Chatterjee et al., 2020, where the research looked at the categorised denoisng methods of Empirical Mode Decomposition (EMD) based model, statistical based models of Denoising Autoencoder and Wavelet Transform, utility of the sparsity property, Bayesian filters and hybrid methods. The work looked at the impact to the SNR's, and RMSE's to determine the models ability to deal with muscle artefact's, for baseline wander artefact's and for power line interference artefact's. Other methods looked at converting EEG time series to 2D images as seen in J. Liu et al., 2020. This work highlighted the improved processing of the SAE with this conversion.

Stacked denoising Autoencoder's were seen in several areas of research. These ranged from their use in feature extraction in C. Xing et al., 2015 which utilised a sparsity activation function with the SDAE for both high level and highly separable features while in P. Liu et al., 2019, the research looked at using SDAE for load forecasting by leveraging the layer by layer pre-training of stacked AE's that removed the random

initialization seen in previous work for load forecasting and in Yu et al., 2021, SDAE's were used for feature extraction improvements using SNR and causal flow of extracted features between areas of the brain which applied Granger Causality analysis to explore connectivity among areas of the brain. Denoising AE's have a problem with time series EEG data in that how can noise be added to an already noisy signal.

The work in Shin et al., 2013 looked at sparse autoencoders and the ability to further improve this by stacking the sparse layers. The concept of sparsity is one that forces the layers to learn from a limited number of previous layer features that are deemed active as some are conceptually switched off so these do not provide anything to the next internal layer. The extent of sparsity is determined in its construction and this results in the extent of features learned. Research shows extensive work with CNN layers for autoencoders such as that seen in J. Liu et al., 2020 where a combination of CNN's, Stacked Autoencoders, and DNN were implemented and the results were promising with a faster reconstructions without loss to performance. The use of Dense layers which are stacked in a Autoencoder were seen in Wen and Zhang, 2018 which highlighted the utility of the approach and the direct impact of the internal feature size where the bottleneck size directly affects the performance. The interconnections of all Autoencoder's layers allowed for complex features being constructed as second and third order affects were considered without the addition to excessive model complexity.

2.3.2 Stacked Dense Autoencoder's

SDAE's are unsupervised neural networks. The stacked nature means the learned data from the previously layer is fed to the input of the next layer. The depth is configurable and trained using backpropogation to minimize the cost function which in turn, updates the weights and biases as part of its training.

The implementation of such architecture has been seen in the literature, its variations are evident in that the types of layers used may vary but the stacking nature prominent

in the improved performances seen. There has been no general consensus returned on how many layers are required to create an optimum learning environment and each implementation seen in the literature has varying degrees of stacking from shallow to deep architectures. Such variations in depth are also accompanied by the extent to which the bottleneck constrains the representation. This was seen in Tautan et al., 2019 and Said et al., 2017 where the layer wise learning's were assessed when it came to the ability to reconstruct the input signal at the output and how different constraints on the size of the internal layers affected the performance of the model. Consideration on the term 'performance' highlights the subjective nature of what is deemed a successful model which tends to be a comparison with an existing methodology that is seen as a baseline and the aim is to have a significantly better performance than this. The work of Boquet et al., 2021 considered such where the research highlighted the unreliability of the reconstruction error being used as a blanket measurement as this does not take into consideration the implementation of the model. Is the goal a perfect reconstruction or is the goal a near perfect reconstruction with the removal of some artefacts that are deemed as undesirable so the perfect reconstruction with a zero reconstruction error is not always the desired outcome.

2.4 Gaps in the Literature

Deep learning has an extensive body of work when it comes to feature generation and extraction. It's ability in processing data makes it ideal for tasks that require feature engineering and it can detect both linear and non linear relationships. This forms the basis for its use with EEG data as this can be extensive in size, non linear and stochastic in nature. A type of deep learning called autoencoders can be leveraged as its construction is built for dimensionality reduction which is the reduction of features to a lower dimension so as to refine these features for information retention over others. This can be extended by using dense layers in the autoencoder as this type of layer learns non linear combinations of high level features which are stacked to allow

for more complex features to be learned. The output of the SDAE is an automatic attempt of noise reduction and artefact removal as the reduction in latent space forces the reduction of representative features for which eye blink artefacts are perceived as being a feature in some form given their prevalence in EEG data.

Existing research has shown some cases where SDAE's were used and their utility evaluated using a second order classification task. No research was seen which looked at the SDAE's ability to automatically remove eye blink artefacts using a constrained latent representation that was reconstructed and evaluated both statistically to the manual approach of ICA and humanistically to determine the presence of eye blink artefacts in reconstructions on raw EEG data that contained known eye blink artefacts.

To what extent can a Stacked Dense Autoencoder reconstruct raw EEG data segments from the DEAP dataset so that the presence of eye blink artefacts are automatically removed and the reconstructed signal has a strong correlation coefficient with that of the manual approach of ICA to removing eye blink artefacts from the same raw EEG data segment?

Chapter 3

Experiment, Design and Methodology

This chapter looked at the methodology that was used to support the pursuit of fulfilling the research question. This entails the Hypothesis, both Null and Alternate, the means of implementing the research and the tools used to support the research outcome via conceptual understanding and evaluation metrics to statistically determine the significance of the findings.

3.1 Hypothesis

Alternate Hypothesis: If a Stacked Dense Autoencoder is used on raw EEG data segments from the DEAP dataset, then the SDAE will remove eye blink artefacts automatically and the reconstructed signal will have an improved SNR and strong correlation coefficient with a manual ICA implementation on the same segment for which the artefact removal verification will be obtained using a manual check for such components where known artefacts exist.

Null Hypothesis: If a Stacked Dense Autoencoder is used on raw EEG data segments

from the DEAP dataset, then the SDAE will not remove eye blink artefacts automatically and the reconstructed signal will not have an improved SNR and strong correlation coefficient with a manual ICA implementation on the same segment for which the artefact removal verification will be obtained using a manual check for such components where known artefacts exist.

3.2 Data Understanding

The data used for this research was taken from the work done in Koelstra et al., 2011. The DEAP dataset is a publicly available dataset that was created to assess the effective states of its participants. The states were constructed using specific stimuli as a trigger to evoke emotion. The dataset is quite extensive and can allow for a myriad of uses but for this research, the raw data will be used so no preprocessing has taken place from when the data was collected. The data is collected on 32 participants who were 50% male and 50% female with an age range of 19 to 37. The variability of the data collection process was mitigated using a defined proceedure with the presence of an experimenter prior to the start of the recording process. The data collected was in the form of 32 separate Biosemi files.

The data was collected using EEG probes placed on the participants scalp, face and hands. As evident in, Silverman, 2015, the placement of the probes on the scalp were both consistent and coherent in their layout. Found in the American journal of EEG Technology, the 10-20 international standard for probe placement was utilised in the data collection. The same journal looked at its history which was developed in 1947 and standardized the placement of the probes so it could account for aspects such as skull sizes of participants and ensured adequate coverage given the inherent issue with spatial resolution. The standard relates to the placement of the 32 probes on the participants scalp. These probes collect the psychological signals, while an additional 13 peripheral nervous system signals, a further 2 for skin resistance and the stimulus

channel are also acquired to make up 48 channels in total.

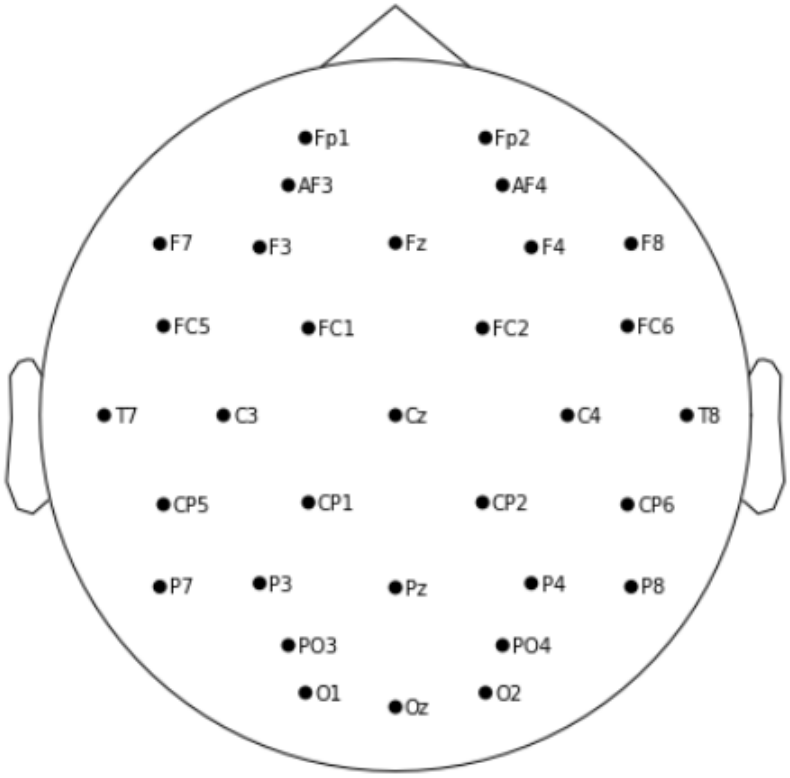

Figure 3.1: EEG 10-20 probe placement - psychological probes

EEG data was recorded with a 2 minute baseline recording prior to any stimuli exposure and from this point, the participant was exposed to a 5 second pre video baseline segment and the music video that played for 1 minute for which the data was deemed under stimuli. From this point the participant had to do a self assessment score for their emotional state and then the cycle repeated with the 5 second baseline again. This repeated for 40 videos in total per participant. The sampling rate of the data was 512Hz for all probes across all participants. The data set was one of a small selection of available datasets that had extensive continuous time series EEG data which was desired for this research.

3.3 Data Preparation

The research sought to input raw data into the SDAE so no preprocessing was desired, however this would impact the ability of the SDAE if ranges were not defined as outliers affect the distinguishable nature of features and the EEG data has frequency bands that contain all relevant information for this work given its general application as opposed to applications such as ERP's who look to garner information from these excessive variations at higher frequencies. To ensure coherent results were achieved, the Makotos preprocessing steps which were seen in "Makoto's preprocessing pipeline", n.d. where the downsampling, low pass filtering and high pass filtering steps were used to create structure to the data reconstructions as EEG signals have frequency bands that define use of the signals as it allows for categorization of the psychological state of a participant. Given the characteristics of eye blinks which were seen in Roy et al., 2014, Matiko et al., 2013 and Zavala et al., 2020 where these have a low frequency profile of about 3 to 4 hertz. Makotos preprocessing entails the downsampling of the signals to 128Hz, low pass filtering at 50Hz and high pass filtering at 1Hz has no known impact on the presence of eye blinks so the perception of raw data being used as the input data still holds as the raw data was refined rather than manipulated.

Preparation of the data involved its segmentation. The first aspect of the data segmentation was to construct the baseline segment which will allow for the research to have a baseline signal to compare with when considering comparisons with what was deemed as noise. Given the inherent noisy nature of EEG data, a ground truth needs to be assumed as no such baseline exists in EEG research that is known to date. This was stated in Leske and Dalal, 2019 where signal distortion was considered when dealing with line noise as the existence of a ground truth for a signal with no line noise did not exist. Highlighting the need to infer what the research deemed as a ground truth. The first 2 minutes of the data was assigned the baseline signal as in this aspect, the participants were monitored but no stimuli was present so this was the assumed baseline noise signal for the research.

The next element to segment was the known stimuli exposed segments which were contained in the video segments. Given the research was seeking to remove eye blinks from relevant psychological signals, the 40 videos were the desired data segments as these contain the stimuli exposure that evokes emotion and brain stimulation. During the data collection process, a stimulus channel was one of the 48 channels available within the data and this stimulus channel had the key event indicators that determined where each element in the recording structure begins and ends. Using the MNE library which was an EEG data processing resource in the Python environment seen in Gramfort et al., 2013, the key event markers were detected and this allowed for defined starting points to segment the videos from.

With the baseline noise signal defined and the 40 video segments created, the next element was the segmenting within the videos themselves. Segmenting allows for a more refined view of the presence of eye blink artefacts. The research evaluated the frequency of eye blinks using existing research such as Abusharha, 2017 which showed an average blinking rate of "19.74 ± 9.12/min" which means that the average was a blink every 3 seconds roughly. With this in mind, the segmenting of the individual video were done by dividing the videos into the following divisions,

- Segment 1 = 1 * 60 second segment
- Segment 2 = 2 * 30 second segments
- Segment 3 = 4 * 15 second segments
- Segment 4 = 8 * 7.5 second segments
- Segment 5 = 16 * 3.75 second segments

This segmenting profile is repeated across all participants within this research. These segments allow for a more granular view as the smaller windows potentially allow for a more distinctive eye blink component to appear and the removal of which can be evaluated when determining the impact on the underlying signal.

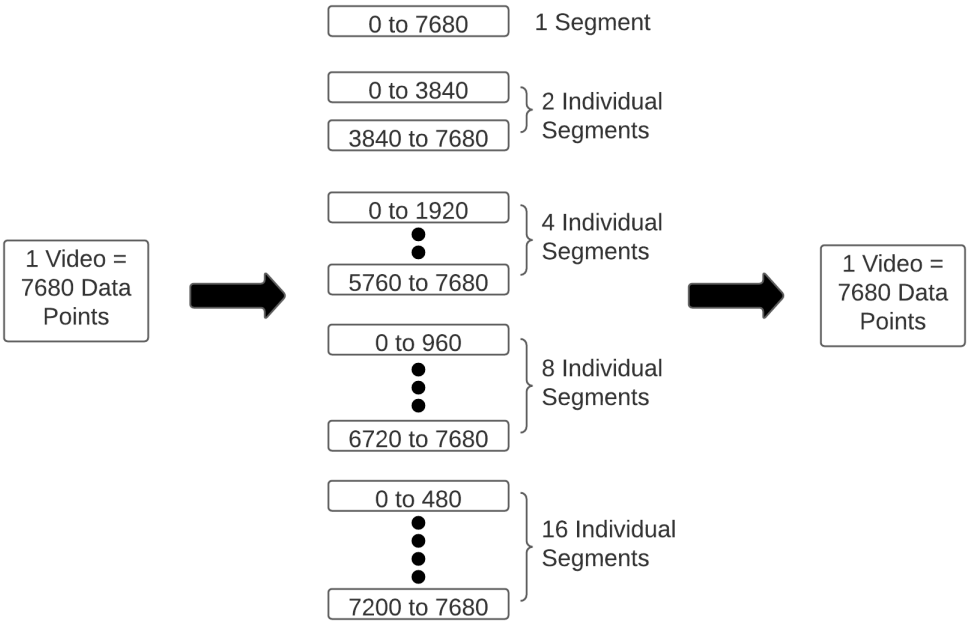

Figure 3.2: Data Segments

3.4 Modelling

SDAE modelling sets out to comparatively evaluate 5 architectures of varying depths statistically. Depth variation comes in the form of the number of internal layers as this directly affects the networks ability to learn. By varying the depth, an optimum depth allows for better performance over an extended number of participants as EEG data is non stationary so subject to high variability which was highlighted in Salazar-Varas and Vazquez, 2019 where robustness of features were effected by the variability of the EEG data. Such variability can be accounted for somewhat by allowing for the architecture to learn complex features while also varying the size of data segments used when trying to generate these representations. This determines the reconstructions which were formed from the features extracted by the SDAE and it is these that were evaluated for eye blink artefacts.

The sequential nature of the research is evident in Figure 3.2 and Figure 3.3 which

details the architecture evaluation stage and the eye blink artefact evaluation stage.

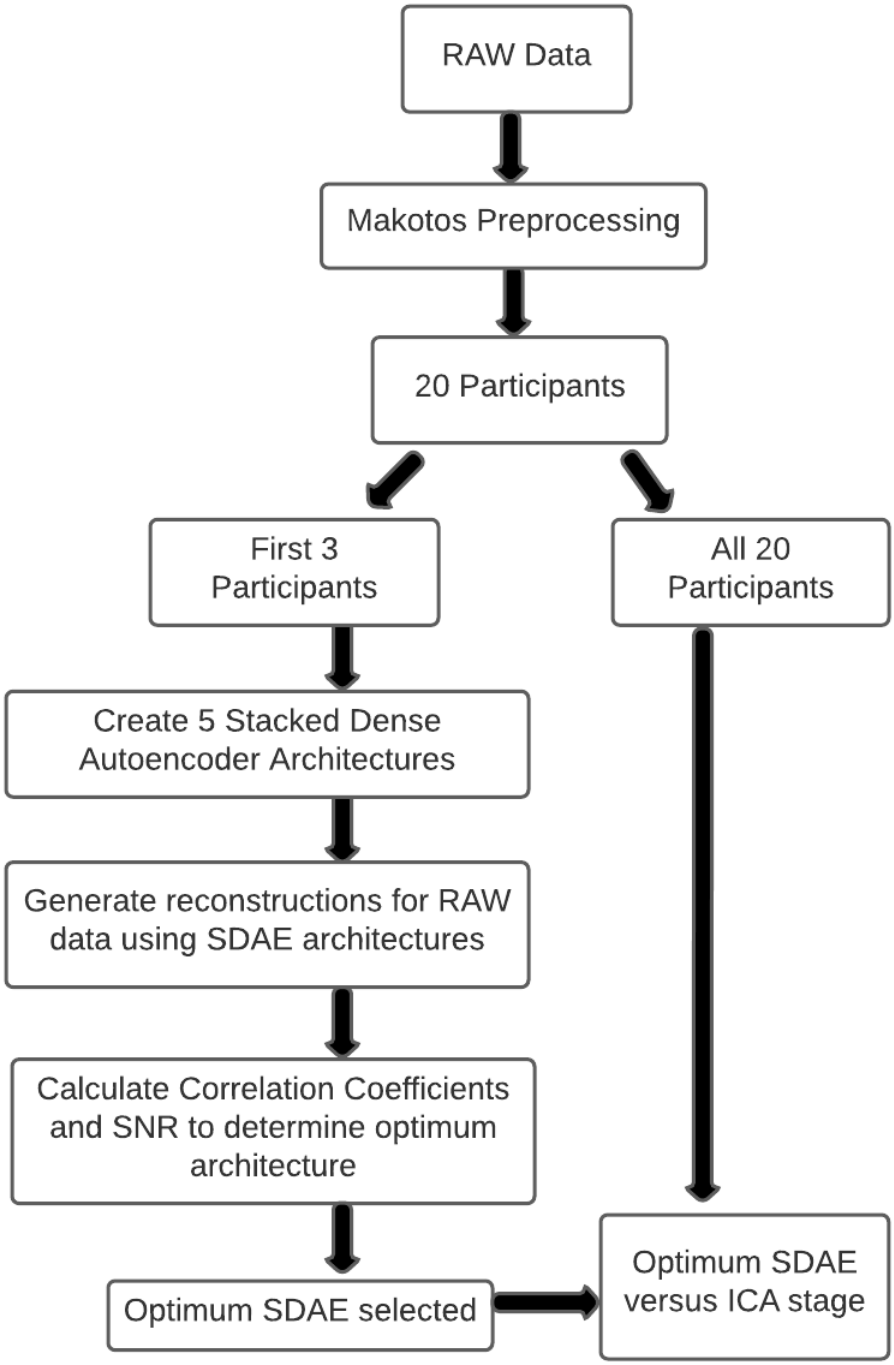

Figure 3.3: Flow chart SDAE architectures

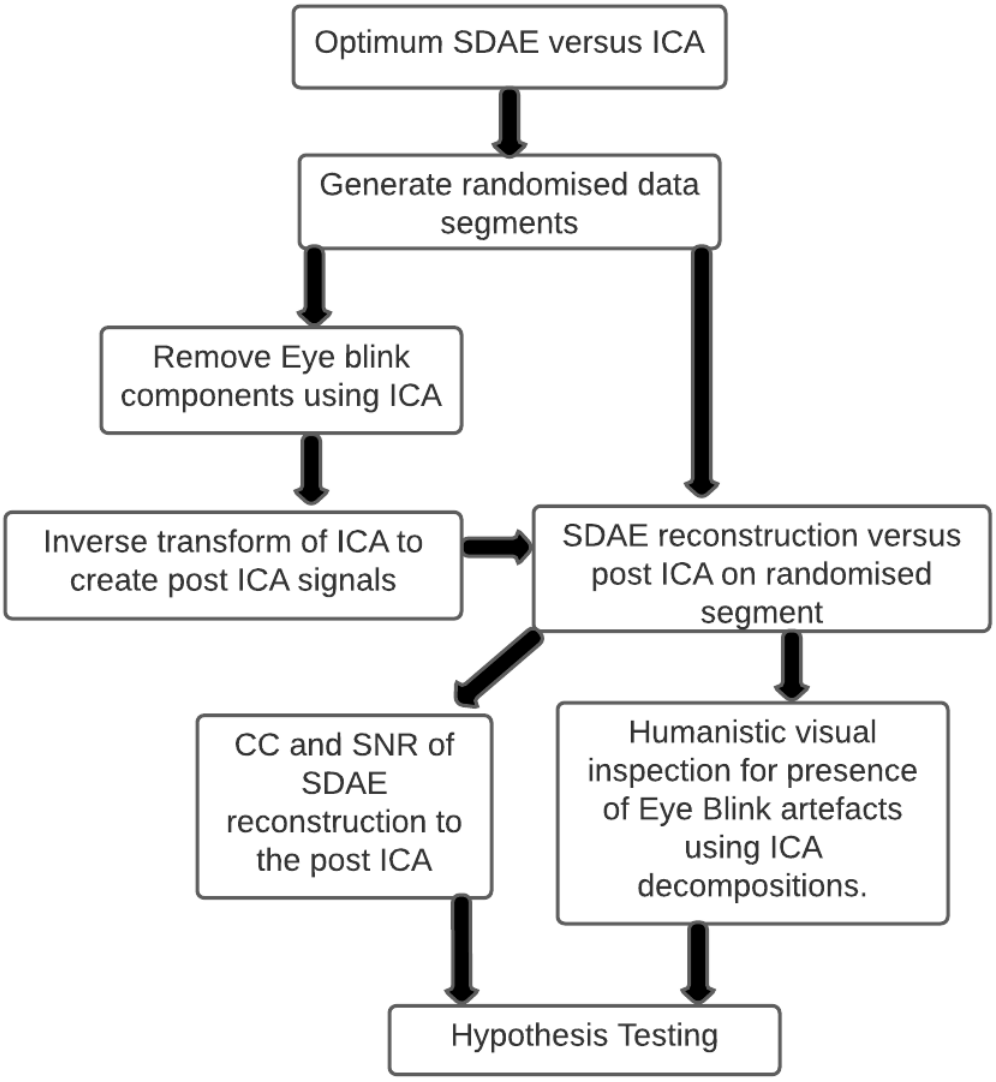

Figure 3.4: Flow chart SDAE versus ICA

3.4.1 Stacked Dense Autoencoder Architectures

Variations using 5 SDAE's were constructed with varying depths to their respective internal structures. Their structures vary but their respective parameters and means of evaluation were all consistent. This allowed for the variations in the latent representation size to be the determining factor in the comparative statistical evaluation. With such variations, eye blink artefacts will appear as relevant features and subse-

quently get removed when constraining the latent representation as eye blink artefacts perpetually exist in EEG and their characteristics are distinct.

Stacked Dense Autoencoder Model 1

SDAE1 has 5 internal layers in the encoder and 5 in the decoder with a bottleneck size of 27 nodes. Reduction in the representation of the input data was forced by constraining the number of nodes in the SDAE as the 27 nodes contains the characteristics to allow the SDAE to reconstruct the input data. Dense layers are utilised to gather inter channel relationships.

Internal layers of the encoder were all constructed so that at each layer, 1 feature was removed so the input data was reduced from 32 features at the input of the encoder to the SDAE, to 31 features, 30 features, 29 features, 28 features and finally the bottleneck of 27 features. Constraining the input data to a lower dimensional representation alleviates some of the potential for the SDAE to learn the identity function as such ability, conceptually is indicative of a desired outcome but such perfect reconstructions were practically undesired. The architecture was evaluated with variations in the data which indicate the architectures requirement to generalise well. This was further enforced with the implementation of Dropout layers between the internal layers of the encoder.

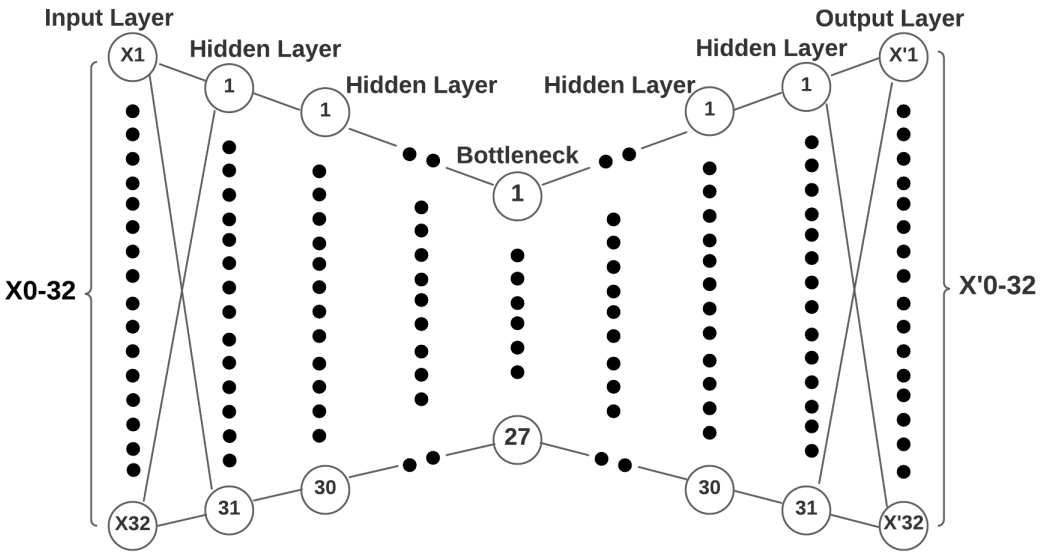

Figure 3.5: SDAE architectures - SDAE1

Stacked Dense Autoencoder Model 2

SDAE2 was constructed with 1 less internal layer to that of SDAE1 for both the encoder and decoder sections and a bottleneck size of 1 additional feature.

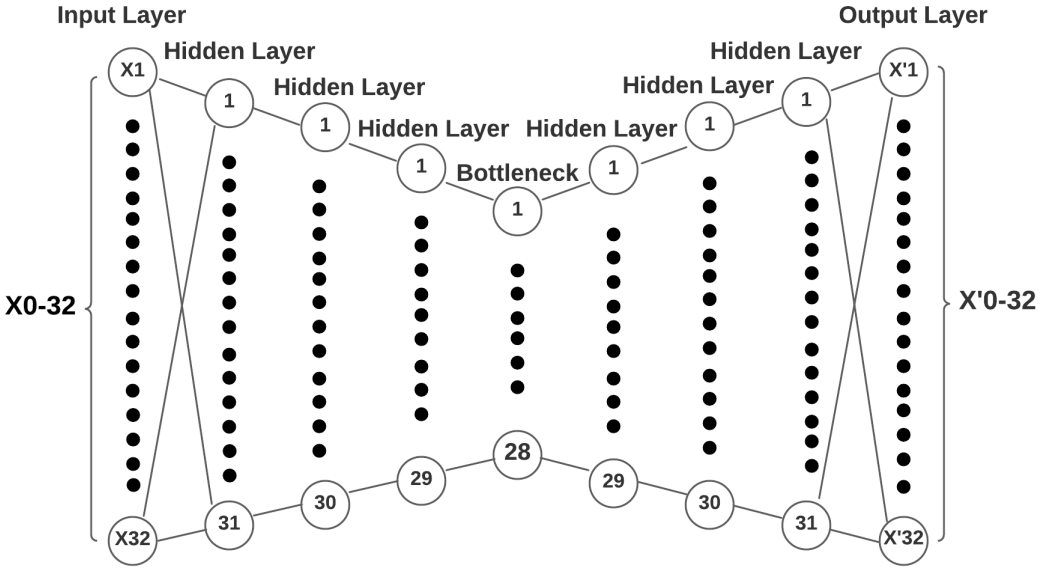

Figure 3.6: SDAE architectures - SDAE2

Stacked Dense Autoencoder Model 3

SDAE3 was constructed with 1 less internal layer to that of SDAE2 for both the encoder and decoder sections and a bottleneck size of 1 additional feature.

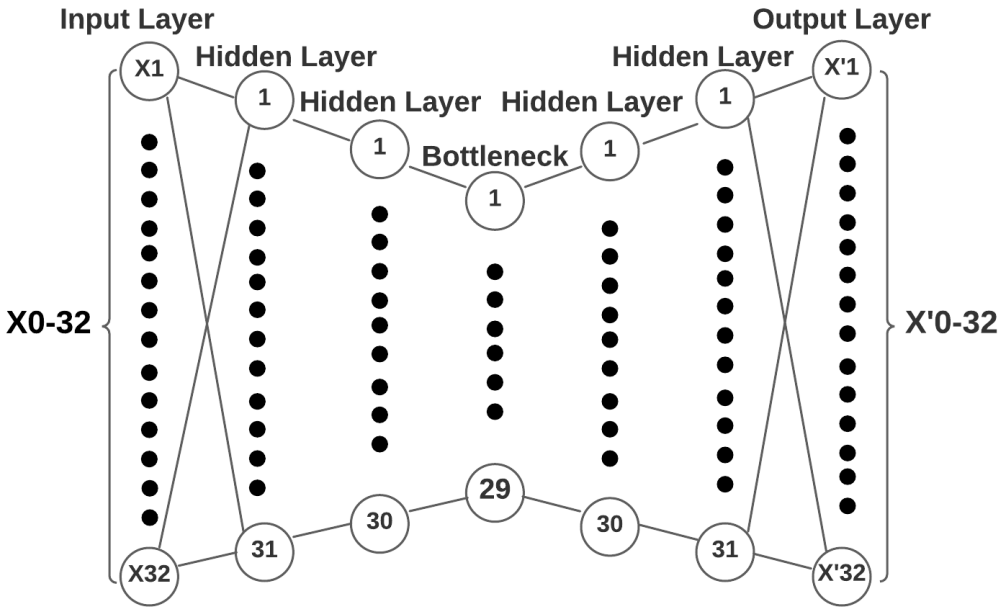

Figure 3.7: SDAE architectures - SDAE3

Stacked Dense Autoencoder Model 4

SDAE4 was constructed with 1 less internal layer to that of SDAE3 for both the encoder and decoder sections and a bottleneck size of 1 additional feature.

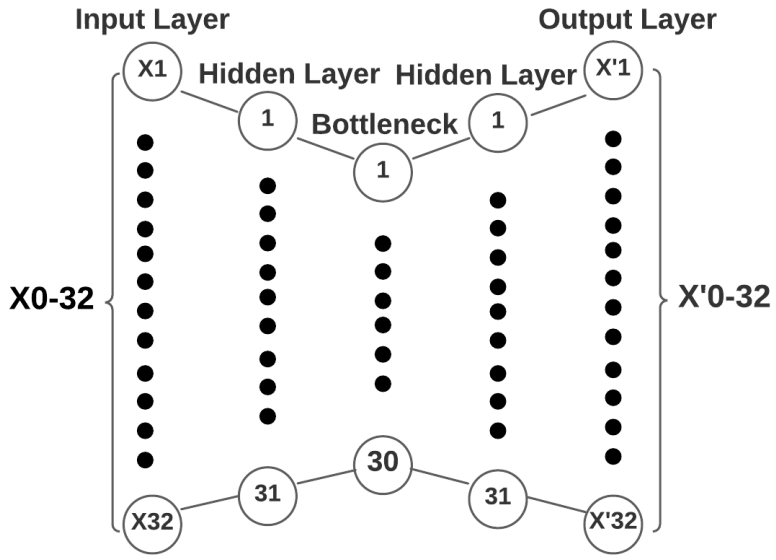

Figure 3.8: SDAE architectures - SDAE4

Stacked Dense Autoencoder Model 5

SDAE5 was constructed with just an encoder, bottleneck and decoder for which there were no internal layers apart from the bottleneck which was constrained at 31 features so the encoder was reduced by just 1 feature in the latent representation. This structure removed the implementation of the Dropout layer.

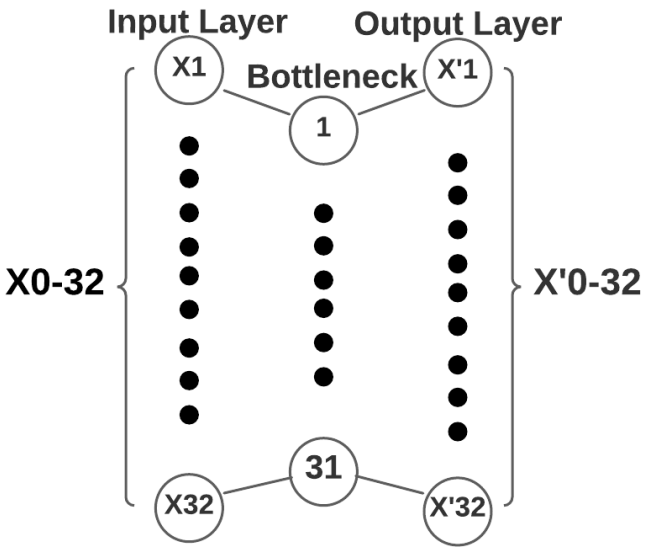

Figure 3.9: SDAE architectures - SDAE5

Stacked Dense Autoencoder Hyperparameters

SDAE's were developed so they could leverage the natural ability of Neural networks to learn features to be harnessed and with varying depths, the learning's would be varying and in turn, allow for an understanding of what depth gave the better learning's for the stochastic nature of EEG signals and their subsequent reconstructions. The architectures allow for an insight to the potential of SDAE's to remove eye blink artefacts automatically and evaluate if this would be achieved in the initial learning's of just the first layer which would be model 5 or it would require a deeper architecture and in turn, require more complex learning's via a model such as model 1 which had 5 internal layers in the encoder. With this in mind, the same hyperparameters were used for each model so as to ensure consistency.

Selection of Hyperparamaters plays a large role in how a model performs. The selection of such pre-learning parameters are integral and are equally as nuanced and complex. Within the scope of this work, the same hyperparameters were selected across all 5

models. This generalising across all 5 models will lead to a reduction in the models potential but given research time constraints, this implementation was justified. The respective values for the hyperparameters were determined from existing research that showed robustness in their performance across various implementations.

- Loss Function = MSE (Mean Square Error). The mean squared error was used as the loss function for the reconstruction. The MSE determines the mean of the squared difference which determined how far apart the original signal is from the reconstructed signal.

- Activation Function = Tanh. The hyperbolic tangent function is a well utilised activation function with EEG data due to its non linear nature. As evident in Aquino-Britez et al., 2021 which highlighted its boundary ranges which allow for more efficient training while Leon et al., 2020 highlighted its zero centred outputs and its ability to be used as considerations such as vanishing gradient which were stated as a disadvantage in their use, were only seen in deep neural networks that far exceed any of the 5 architecture depths being implemented.

- Dropout = 0.1. Configuration of the dropout layer was guided by the work seen in Gogna et al., 2017 and C. Xing et al., 2015 where this technique of disabling random neurons and averaging the error calculations was implemented. As a consequence, it is a form of regularization which aims to ensure the model neither under or overfits.

- Optimiser = Adam. The Adam optimizer default settings have been shown to be robust. Given the research will not be searching for the best hyperparameters, then robust default hyperparameters were desired. As evident in Alzahab et al., 2021 where a literature review of BCI related studies using deep learning stated "From the remaining 76.6% that declared that the optimization algorithm was used, the most used was the ADAptive Momentum (ADAM) optimizer that was used in 55.3% of the cases" and " Empirically, it was shown that ADAM outperformed other optimizers" which further support its use.

- Learning Rate = 0.000001. Work carried out in Leon et al., 2020 gave great depth to the implementation of learning rates across Deep Learning structures like RNN's, CNN's and FFNN's. Selection of the learning rate of 0.000001 was obtained through trials on a single segment of data as it highlights its relevance in that its a fraction of the measured error that is used to correct the weights of the network which directly affects the training of the network and in turn, the ability to learn features from the data.

- Early Stopping = minimum MSE for 20 values (patience). Implementation of Early stopping was utilised to optimize the time required for training the models. At the point of convergence, the model stopped.

- Epochs = 500. The number of forward and backward passes of the training data through the network is determined by the number of Epochs. The value selected was based from existing literature and it's impact was minimal due to convergence happening quicker than originally anticipated.

- Batch Size = 128. Batch size was set based on the sampling rate. Variations were trialed with a size of 128 proving to be the most robust across the varying time segment sizes which determine the amount of data provided to the model.

Stacked Dense Autoencoder Evaluation

Deviations from convention were part of the SDAE architecture structure as the architecture was trained and tested using the same data. Evaluation of its ability to reconstruct data with the removal of eye blink artefacts can only be verified at a later stage and with this, evaluation of the SDAE's comes in the form of convergence. The point of convergence of the SDAE architectures defines the state at which the models stops and no longer attempts to reduce the loss value.

Evaluation was two fold. Firstly the SDAE's were evaluated based on a comparative assessment between the 5 architectures which was achieved using the data of 3 partic-

ipants for which there were 40 videos per participant and 5 segments for each video as outlined previously. Architectures were evaluated based on their ability to reconstruct a correlated signal with the input data and further evaluated on their effect on the signal to noise ratio when compared with that of the original signal. Given the exponential increase in required trials as the participants data was segmented to 40 videos and further segmented to time slices, aggregation of results was used so the values depicted were average of averages so as to ensure legible data is conveyed.

The second element of the evaluation related to the chosen architecture. The chosen SDAE was exposed to an additional 17 participants data that were reconstructed. SDAE reconstructions were evaluated against the manual approach of ICA that detected eye blink artefacts and subsequently removed these from the same time segment of the original signal to that of the reconstruction. With the removal of known eye blink artefacts, the components were inversely transformed so that the post ICA reconstruction have known eye blink artefacts removed. With reconstructions of both the SDAE and post ICA on a specified segment of raw data, a direct comparison between approaches could be achieved. Such comparisons were achieved using both the original signal and EOG channels to confirm the present of eye blink artefacts in the original signal from which the SDAE reconstruction was formed, the SDAE reconstruction and post ICA reconstruction were both decomposed and evaluated in terms of their ability to remove known eye blink artefacts. Both the SDAE reconstruction and post ICA reconstruction were evaluated using a number of randomly specified segments.

3.4.2 ICA Architecture

ICA seeks to establish maximal independence via the transformation of the vectors that were fed into the algorithm which were in the way of time series data segments. The algorithm utilises 2 basic rules, the first being the components that are sought being statistically independent and their distribution being non Gaussian. Its origins can be seen in Hyvärinen and Oja, 2000 where its creation was due to an understanding

of computational and conceptual simplicity.

ICA allowed for the detection of the Eye Blink Artefacts that the research was attempting to determine whether these were removed or retained in the lower dimensional representation and subsequently made up part of the reconstruction via the decoder. ICA decomposition's were influenced by many factors, some of which can be seen in Klug and Gramann, 2020 where factors such as "movement in EEG experiments, the number of channels, and the high-pass filter cutoff during preprocessing" had a ranging influence on the decomposition. This research highlighted the need for high density recordings which can be re-framed to more channels which highlights the relationship between density and the ability to separate components. The influence on the decomposition's can be tolerated as this can not be quantified due to it being experiment related in its attainment and the more data means more computational requirements. Given the time constraints of this work, compromises were required in areas where speed of the attainment of results was given precedence over precision without compromising integrity in that the known variability would have minimal impact.

The FastICA algorithm was utilised as it had quicker convergence to other algorithms. The work done in Pontifex, Gwizdala, et al., 2017 where 32 participants had their respective data ran through different ICA algorithms 30 times so as to create 30 separate decomposition's, showed this ability to quickly converge. Added variability when being utilised was highlighted in the same work. The FastICA showed sub optimal results, "The high degree of variability observed across repeated decomposition's when using the FastICA algorithm was surprising given its popularity to date" highlights such observations. All variants of the ICA algorithm showed variability and given the fast convergence of FastICA, it was used as the algorithm for implementing ICA.

FastICA has a requirement for the data to be both centred and whitened which can be set in the model parameters when using the algorithm. However, these were done manually as the need to remove these steps after the inverse decomposition was re-

quired to allow for the original signals and the post ICA signals to be compared. The "StandardScaler()" function was used to centre the data which was where the scaled data has a mean of zero and a function was created to whiten the data which involves the removal of correlation or some form of dependencies between the data features.

FastICA algorithm was configured the same across all aspects for which it was implemented in. Such configuration can be seen in the following,

- n-components=32. Number of components relates to the number of channels which were seen as the features. It is these that make up the unmixing matrix.
- random-state=0. Sets the reproducibility in the algorithm.
- tol=0.05, The tolerance sets the threshold for the algorithmic convergence.
- max-iter=100. Given the computational requirement need for the larger time slices, a reduced number of iterations was implemented which impacts convergence but did not visibly impact the ability to generate distinct components present in the signals.
- whiten=False. Whitening aspect was set to False as it has a default behaviour of being set to True. Set to False due to the manual whitening implemented.

Implementation of the ICA algorithm via FastICA is not deterministic which means the components created via the unmixing matrix can vary from iteration to iteration. The FastICA algorithm creates components that make up part of the original signal. The components were then plotted to allow for a visual representation of the 32 components that were conveyed. Identification technique for the eye blink artefacts was achieved using the eye blink artefact profiles shown in section 3.4.3. Within the ICA step, once these were identified then these were zeroed out so they were not part of the restoration of the input signal. As evident in Gramfort et al., 2013 where "artifacts can be removed by zeroing out the related independent components before inverse transforming the latent sources back into the measurement space" which highlights the

process of zeroing the undesired components which will be distinct eye blink artefacts in this instance.

3.4.3 Eye Blink Artefact Detection

The components generated by the FastICA algorithm decomposition's allowed for the detection of eye blink artefacts provided these were present within the data segment and distinct enough to be detectable. The method used for such detection were manual in this research. Manual inspection was not without its considerations and one such consideration was the variability it induces across interpretations and a non definitive threshold seen to date in other work. Alternative means were considered but variability was prevalent in the implementation of both manual and automated approaches. As highlighted in Chaumon et al., 2015 where the research stated that both manual and automated approaches showed inherent limitations in their respective precision's when it came to detecting artefacts. Manual inspection allowed for a visual understanding. Visualization of the reconstructed signal components was seen as a key indicator of the ability or lack there of in relation to the SDAE removal of eye blink artefacts.

To allow for comparative evaluation using a manual interpretation which were carried out with limited domain knowledge or experience meant that templates were required as references for determining the presence of eye blink artefacts. These were seen in Matiko et al., 2013, Zavala et al., 2020 and Roy et al., 2014 which characterised the eye blink artefacts using time series analysis and distinct changes in amplitude that accompanied this. In the work seen in Zavala et al., 2020, the research forced the participants to blink which allowed for the profiling of eye blink artefacts and these were stated, "A short-intended blink is detected when the signal drops below -125±15 uV, then rises above 125±15 uV and then goes in a state of normality (between -45 and 45 uV) in a window of 2.5 sec" and "the signals drops below -125±15 uV, rises above 125±15 uV, enters a state of normality, then rises above 125±15 uV and then drops below -125±15 uV again". These papers provided visual depictions of eye

blink artefacts and statistical characteristics to accompany these. Eye blink artefacts were evaluated visually using both the graphical and time series characteristics of the profiles provided which determined the presence of eye blink artefacts.

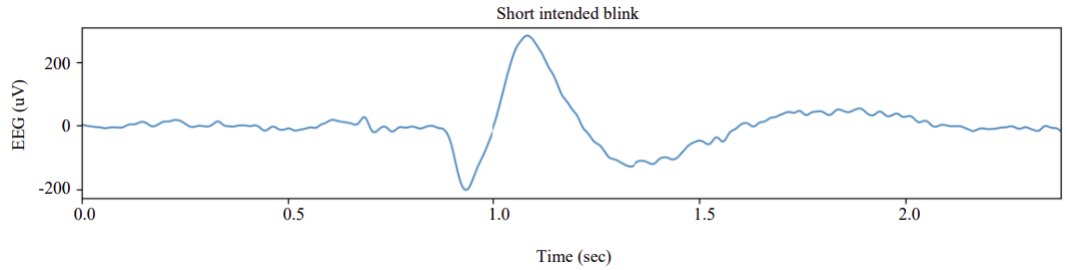

Figure 3.10: Short Eye blink, Zavala et al., 2020

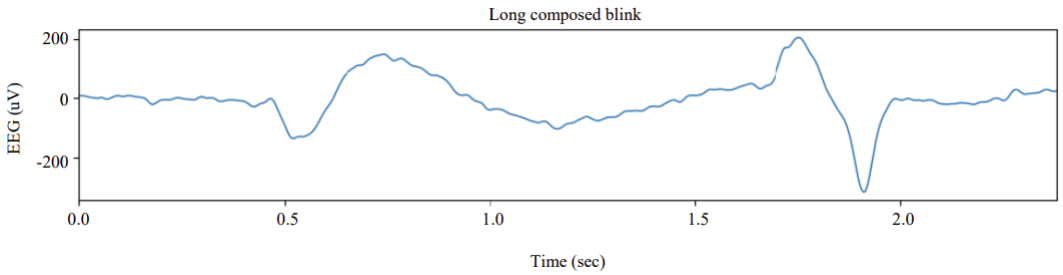

Figure 3.11: Long Eye blink, Zavala et al., 2020

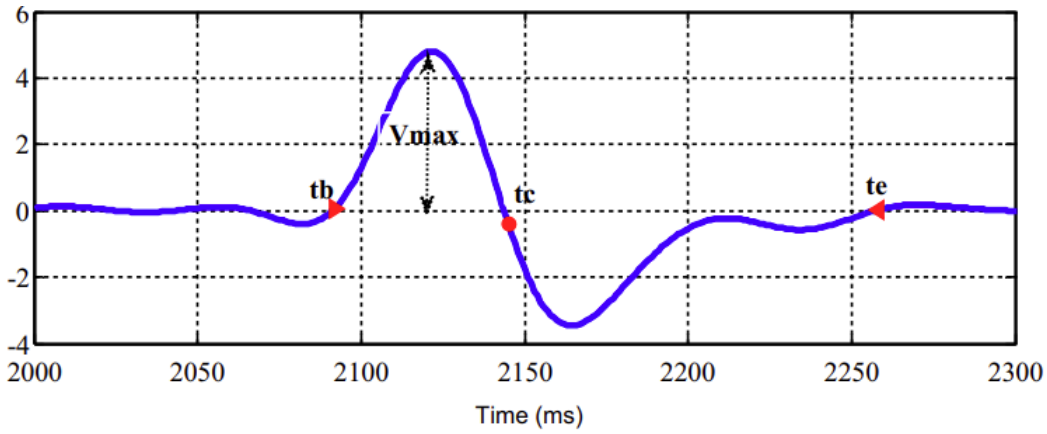

Figure 3.12: Eye blink profile, Matiko et al., 2013

The ability to detect eye blink artefacts allows for their removal. As was evident in Figure 3.13 and Figure 3.14 where eye blink artefacts were detected and subsequently

removed while visually appearing to retain the signals integrity in their respective works.

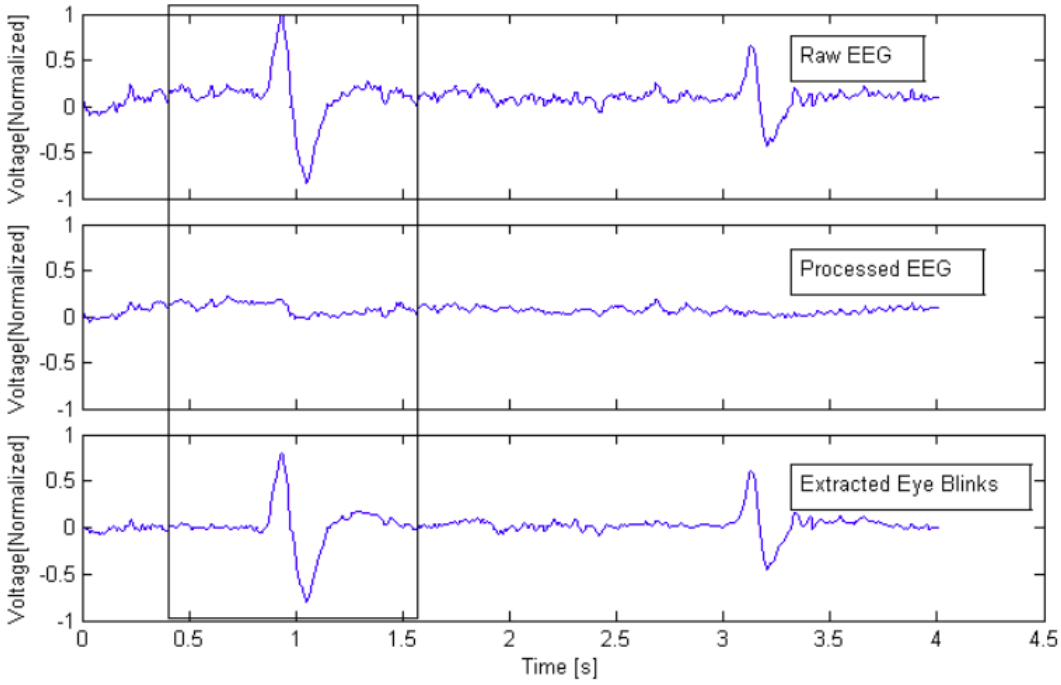

Figure 3.13: Eye blink present in signal, Roy et al., 2014

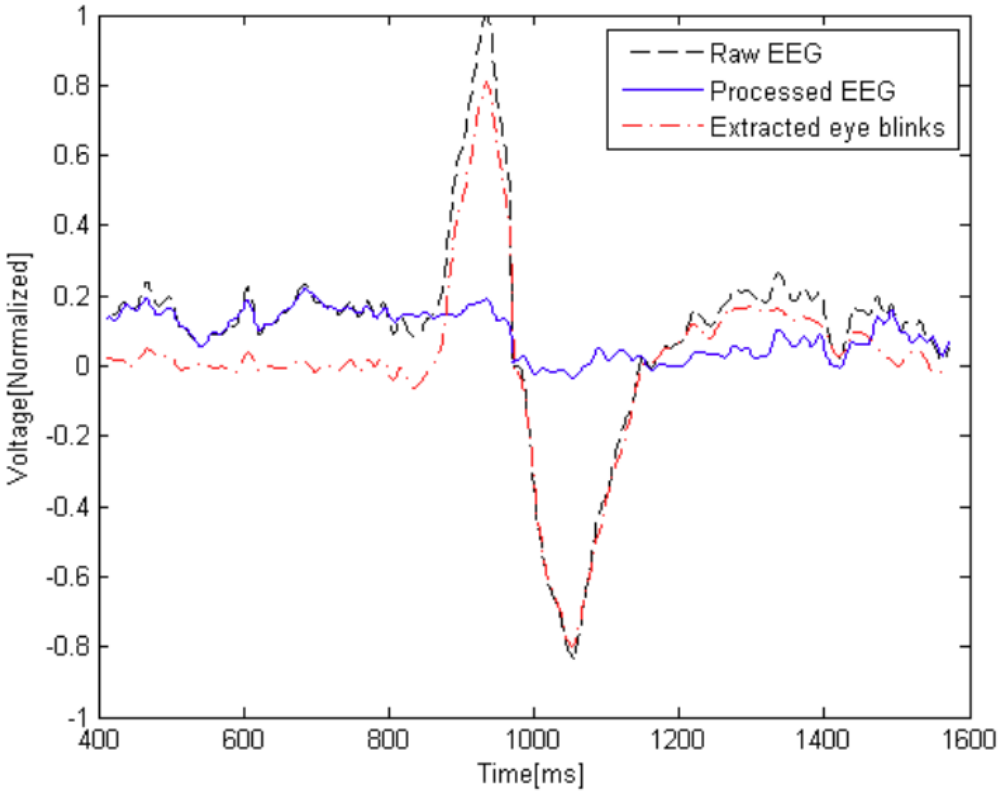

Figure 3.14: Eye blink present and removed in signal, Roy et al., 2014

Support for the detection of eye blinks were found in the additional probes placed on the participants face. These were non psychological probes of "EXG1" and "EXG2" with their placement designed for ocular artefact detection. These allowed for further confirmation of the presence of eye blink artefacts in segments of psychological signals that supported their manual classification as an eye blink artefact.

3.5 Evaluation

A means of evaluating the research was rooted in its ability to determine a reconstruction that both resembles the input signal but not too much as the removal of eye blink artefacts would change the signal from it original form so a perfect reconstruction would indicate non attainment. This indicates that the standard error used for

SDAE's will not suffice for determining a meaningful SDAE performance. The SDAE's were configured to converge based on the loss value but the selection of the optimum architecture was based on alternative metrics.

Given the sequential nature of determining the optimum architecture from the 5 outlined and then its utility in removing eye blink artefacts with that of the manual approach. Averaging was used when conveying results which were gathered across channels, time slices, segments, videos and participants. Reconstructions was generated for all relevant data, not all can be evaluated however given the extent of segmentation across 20 participants and the time intensive nature of running the ICA algorithm when evaluating the SDAE reconstructions with the post ICA reconstruction. All reconstructions of the first 3 participants were evaluated for the SDAE architecture evaluation where as for the comparison with the manual ICA approach, a selection of random reconstructions were evaluated.

3.5.1 Randomised Segment Evaluation Metrics

Given the extensive number of segments to be evaluated meant that a only a cohort of these could be evaluated within the time frame available. Evaluation of the SDAE and post ICA reconstructions can be construed using specific segments that support a narrative. This research wanted to remove any potential bias and with such a consideration in mind, created a random segment generator which specified the participant, video, segment and time slice to be evaluated. Ensuring no specific segments was selected and the randomness supported an unbiased view within the work.

3.5.2 Correlation Coefficient

The Spearman Correlation Coefficient was used as the statistical evaluation metric as it is a non parametric test which fits the nature of EEG data and it looks at the degree of association between variables which the research seeks to quantify. Various

methods were looked at prior to using the Spearman Correlation. As was evident in Bonita et al., 2014 where Pearsons, Spearman and Kendall rank correlations were evaluated comparatively. The research looked at the between channel correlations and highlighted the robustness to noise for these methods. The concept of between channel correlation was suitable to what the research sought to evaluate as the concept was similar, evaluating the correlation between 2 channels but with different distributions, the first being the original signal and the second being the reconstruction of this when it came to the optimum architecture and the ability of the SDAE to remove eye blink artefacts so the SDAE reconstruction with the perceived baseline of the post ICA reconstruction.

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \tag{3.1}$$

Where,

p = Spearman rank correlation coefficient

d = the distances of two ranks of each observation

n = the number of observations.

3.5.3 Signal to Noise Ratio's

Signal to noise ratio can be generalised as a ratio for the power of both the signal and noise which were perceived as having no defined dimensions. The signal is the meaningful information and the noise is unwanted. Noise is induced, its inherent in anything electrical or humanistic. The signal to noise ratios allowed for the research to determine whether there was an increase or decline in the reconstructed signals when compared with the baseline noise which was established using the first 2 minutes of the data recording with the assumption that the lack of stimuli meant the time segment was a reasonable base for the interpretation of the noise of the data. Variations in

the SNR were indicative of an altered state of the original signal so if variations were observed from the original signal to the reconstruction using the same baseline noise, then the reconstruction differs from the original and the extent of such can be implied from the SNR's. The formula's were taken from the work seen in Nagar et al., 2021 and these supplement the correlation values in gaining an insight into how the signal was manipulated.

$$SNR = 20 \log_{10} \frac{S}{N} \tag{3.2}$$

$$S = \sqrt{\frac{\sum(signal)^2}{len(signal)}} \tag{3.3}$$

$$N = \sqrt{\frac{\sum(noise)^2}{len(noise)}} \tag{3.4}$$

Where,

SNR = Signal to noise ratio

S = Signal, the voltage readings.

N = Noise, assumed noise segment of the data

len = length, how many data points long.

With the extensive segmentation and subsequent need for averaging already highlighted for the portrayal of results. More granular evaluations were carried out as part of the SDAE and post ICA evaluation for an understanding of distributions, maximum and minimum values indicate how the channel wise performance affected the averages displayed.

$$\sigma = \sqrt{\frac{\sum (X - \mu)^2}{N}} \tag{3.5}$$

Where,

σ = Population S.D.

X = is individual value

μ = Population Mean

N = Population size

3.5.4 Optimum SDAE's Architecture

Evaluation of the SDAE's was carried out on the first 3 participants using the segmentation's previously stated. Each architecture was subjected to the same data with its configuration stipulated prior to running and was consistent across all architectures. A reconstruction was created for each participant across all videos and segments. These segments were stored in a 4D array that can be set out in the following, (40, 5) which were the (40 videos, 5 time segments) and within this was the arrays that make up the time segment and were depicted as (1, 1, 7680) for the 60 second video for time segment 1 to (5, 16, 480) for time segment 5 which were 16 time slices of the 60 second video. Correlation Coefficients and SNR for every segment were constructed and these were averaged to give a single CC and SNR for each time segment and then further averaged to give a CC and SNR for each video so each participant has an average of averages for the CC and SNR for each video. Comparative evaluation was then implemented to determine the optimum architecture within the scope of the 5 constructed which was used in the evaluation of the SDAE reconstructions for eye blink artefact removal with that of the post ICA reconstructions.

3.5.5 Optimum SDAE Versus ICA

With the chosen SDAE determined, a further 17 participants data were reconstructed. The SDAE architecture had reconstructions for all relevant data segments for the first 20 participants of the DEAP dataset in total.

Evaluation of the optimum SDAE reconstruction to that of the post ICA reconstruction, required a specified time segment to be determined. A means of such was achieved using a randomizer that determined what participant, video and time segment the evaluation would be carried out on. The specified time segment was then taken from the original data and ran through the FastICA algorithm where its eye blink artefacts were removed and the signal restored to give a post ICA reconstruction. The same time segment already had a SDAE reconstruction.

Eye blink artefacts were removed in the post ICA reconstruction using the profiles previously defined. These determined the presence of eye blink artefacts. The formation of the post ICA reconstruction determined if there exists eye blink artefacts in the specified raw data segment and in turn, allowed for the evaluation of the SDAE reconstruction to be done with evidence of known eye blink artefacts present in the original signal for which the SDAE created a reconstruction for.

3.5.6 ICA decomposition's of Original versus SDAE

The removal of the eye blink artefacts can be confirmed by the inspection of the respective components of the SDAE reconstruction using the FastICA algorithm. For confirmation of their removal, confirmation of their existence needed to be established first. This was achieved by using the number of components that were detected as eye blink artefacts in the original signal when generating the post ICA reconstruction and can be supplemented by the ocular probes of "EXG1" and "EXG2" to verify the presence of such artefacts. With the presence of known eye blink artefacts in the original and EOG signals for the random segment, the SDAE reconstruction was evaluated on its ability to remove these within the specified segment.

Evaluation of the SDAE reconstruction used the previously established profiles of eye blink artefacts to determine their presence. The interpretation was not binary and the presence of such artefacts proved inconsistent due to the manual nature of

it's interpretation of what met a visual profile. Their presence determined whether the SDAE was able to remove known eye blink artefacts and to what extent this was achieved. The post ICA reconstruction was also re-evaluated using the FastICA algorithm to evaluate the components that make up its reconstruction to determine whether all eye blink artefacts were removed using the manual approach.

3.5.7 Statistical Evaluation

Statistical evaluation was utilised to determine quantitative evidence that supports conclusions that need to be drawn within the humanistic aspect of eye blink artefact detection. Such statistical evaluation can account for uncertainty in the obtained and observed results. Spearman Correlation Coefficient which is a statistical test, determines significance in the correlations between all channels in the SDAE reconstruction and that of the post ICA construction. A p-value was formed that indicated whether there was a significant correlation. Statistical evaluation allowed for the reconstructions to be comparatively analysed due to the baseline being that of the post ICA reconstruction for which known eye blink artefacts were removed and the SDAE reconstructions were evaluated against this. The ability to reconstruct the original data that was correlated to the post ICA reconstruction and a SNR value that did not vary wildly from that of the same baseline, determined the SDAE's ability in automatically removing eye blink artefacts from RAW EEG data.

3.5.8 Human Centred Evaluation

Human centred evaluation was the manual check for the presence of eye blink artefacts in the SDAE reconstruction from random segments with known eye blink artefacts present. The systematic check allowed for the creation of a table that displayed the known number of eye blink artefacts in a specified segment using the original signal decomposition that was part of the post ICA reconstruction and raw EOG signal

decomposition that form the basis for the known number of eye blink artefacts in a specified segment. The table was then populated with the number of eye blink artefacts detected in both the post ICA reconstruction and the SDAE reconstruction. The formation of such a table allowed for the creation of an Inverse RMSE value to establish a numeric indication of how well the SDAE performed when compared to the manual ICA approach.

Root mean square error is a metric that shows how far apart the predicted values are from the observed values, on average. The predicted value was the number of eye blink artefacts known and the observed value were the number detected in both the post ICA reconstruction and the SDAE reconstruction respectively. It's meaning was inverted in its use for this work as the further from the predicted value was seen as a good performance as such indicated an ability to remove eye blink artefacts.

$$RMSE = \sqrt{\frac{\sum(Pi - Oi)^2}{n}} \tag{3.6}$$

Where,

Pi was the predicted value for the ith observation

Oi was the observed value for the ith observation

n was the sample size

3.6 Summary

Variations in SDAE architecture were proposed to find the optimum size of an internal representation of raw EEG data that allowed for the removal eye blink artefacts. Performance was evaluated on 3 participants data who were exposed to 40 stimuli inducing videos which were reconstructed and evaluated using CC and SNR to the original signal and baseline noise respectively.

An optimum architecture was determined and this reconstructed the data of a further 17 participants. Random segments of the reconstruction were then comparatively evaluated both statistically and humanistically with that of the manual approach of ICA on the same specified segments. The SDAE reconstruction were evaluated using the post ICA as the baseline to which the SDAE needed to achieve a similar reconstruction as the post ICA reconstruction had known eye blink artefacts removed. Statistical evaluation using the CC and SNR was used to evaluate its reconstruction and FastICA was used to evaluate the components that made up the SDAE reconstruction so the presence of eye blinks could be evaluated in random segments that had known eye blink artefacts present.

3.6.1 Strengths of the Design

- Architectures: Assessing 5 different depths allowed for the stacking of layers to model complex relationships for which a definitive depth was not defined in existing literature. Assessing the architectures across 3 different participants supported the merit of a single chosen SDAE.

- Robust Hypothesis testing: Combining both a statistical evaluation and a human centred evaluation meant the research can confirm and quantify the ability of the SDAE to automatically remove eye blink artefacts.

- Eye blink artefact templates: Various research was obtained that showed a similar consensus for the profile of eye blink artefacts.

- Data segmenting: By segmenting the data, allowed for an understanding of the impact of segment sizes which was relevant due to the potential number of eye blink artefacts that were present in the signal. The ubiquitous nature of eye blinks in EEG data means larger segments will have a large number of eye blinks present so the impact of several being removed can be assessed against the shorter time segments for which 1 or even non were present, allowing for an understanding of the impact on potential signal information loss.

- Random participant: The random selection of a participant, video and time segment removes any form of selective biasing in the results conveyed. The randomization allowed for a fair and balanced view of the SDAE's ability in removing eye blink artefacts as results that convey a predefined narrative can not be selected.

3.6.2 Limitations of the Design

- Baseline noise: There exists no ground truth for a clean EEG signal in any recording so a baseline noise was assumed. Such an assumption used the first 2 minutes of the data recording as the baseline noise as the participant had no exposure to stimuli within the experiment. This does not consider the potential for inherent brain stimulation so the participant may not be exposed to direct stimuli but their brain may be stimulated due to their surroundings.

- Information loss: There was no way to determine signal loss. The inherent presence of noise meant that any variations in signal quality can not determine whether original signal information was lost. SDAE reconstruction was similar, the loss of information couldn't be calculated as no metric exists for such.

- Eye blinks: The use of templates meant that their was a subjective element added to the research. The point at which some appears large enough to be determined an eye blink may vary across researchers and given the time constraints, statistical characterization was not implemented to confirm the presence of eye blinks so their was an element of human variability.

- Hyperparameters: The hyperparameters of the SDAE were not optimised. The full potential of the SDAE architecture was not utilised.

- Limited participants: 20 participants were evaluated, 32 were available in the data but given the extensive computational requirements for reconstructing data for the SDAE's, a more refined number of participants was used.

- Dataset: A single dataset was utilised which meant the SDAE architectures ability to generalise could not be evaluated.
- Aggregation of results: The conveying of results was portrayed as an average of averages so nuance of inter channel impact was not portrayed. Aggregation of results generalises the SDAE's ability which masked segments in which it may have showed potential.
- Training of model: The training of the model focused on convergence but did not consider the level at which such a convergence happened.
- Explainability: How the SDAE made the decisions it did to generate the reconstructions it returned were unknown and could only be theoreticized as it is a form of deep learning.
- Domain knowledge: Manual detection required an understanding of what the signals conveyed. The nuances of eye blink characterization needed to be understood as their was no defined length or amplitude for eye blinks, just a general characterization of a narrowed view.
- Segmenting of the data: For the shortest data segment, there were 16 of such per video. Given there was no overlapping between segments meant eye blinks may have been missed if they occurred across any of the points of segmentation.

3.6.3 Equipment and Tools

The ability to answer the research question required the ability to carry out the steps needed to implement the empirical research. The Python programming language was used as the language was readable and has extensive libraries available which reduced the work load as existing libraries are leveraged to carry out tasks. The Python language was ran on Google Colab as the environment allows for access to GPU's which aid in the processing times that tend be excessive when implementing Deep Learning models.

Chapter 4

Results, Evaluation and Discussion

This chapter display the results that were sequentially obtained for both the SDAE architecture comparisons and subsequent evaluation of its ability to that of the ICA approach. These results were conveyed using aggregation of the CC's and SNR's for readability on the comparative evaluation of the SDAE architectures while random segments were used to compare the optimum SDAE reconstruction with the post ICA reconstruction. The results for which were not aggregated and an in depth comparison of the ability to remove eye blink artefacts was achieved using both statistical and human centred evaluation methods on the respective reconstructions.

4.1 Modelling

The optimum SDAE architecture was established prior to determining the SDAE's ability to remove eye blink artefacts. For this reason, the results across all 5 architectures were obtained and form there, these were evaluated to determine the chosen architecture which allowed for the next stage of the results to be generated.

4.1.1 Modelling - SDAE Architectures

A comparison of the 5 SDAE architectures across 3 participants where the data were segmented was completed. With the segmentation of the data already established in the design section, the metrics were aggregated so that the channel wise values were averaged to form a single value to represent all 32 channels, these were then averaged for each time segment that was created so there were 5 such averages for each video and these were further averaged so there was a singular value for each video. With a value formed for each video for which there were 40 and repeated for each of the 3 participants on all architectures, such value aggregation allowed for further averaging by forming a single value for the 40 videos so each architecture had a singular value for each participant. The singular value was formed for both the CC's and SNR's.

The values displayed in table 4.1 show the averaged value for the specified participant.

	C.C	**Pvalue**	**SNR**
SDAE1-P1	0.005387	0.036437	27.990769
SDAE1-P2	-0.001210	0.036437	26.898129
SDAE1-P3	0.004858	0.036164	27.220164
SDAE2-P1	-0.001151	0.0371221	25.626538
SDAE2-P2	0.004811	0.037508	24.328116
SDAE2-P3	0.002960	0.040853	25.642930
SDAE3-P1	0.003348	0.036164	21.872959
SDAE3-P2	0.002369	0.041290	21.312047
SDAE3-P3	-0.000009	0.036712	22.072206
SDAE4-P1	-0.001046	0.036551	17.125710
SDAE4-P2	0.003410	0.036829	17.652729
SDAE4-P3	0.003364	0.037672	18.294193
SDAE5-P1	0.000556	0.032596	9.359257
SDAE5-P2	0.003449	0.032458	9.812392
SDAE5-P3	-0.000097	0.033985	11.250644
Original-P1	-	-	-3.9718
Original-P2	-	-	-2.5043
Original-P3	-	-	-0.7516

Table 4.1: SDAE architecture statistical metrics

What was evident in the results generated for the SDAE architectures, was the increased SNR which was measured using the db scale. This was calculated using the first 2 minute segment of the data as the baseline noise which was further refined to the first minute to mirror the data size of the reconstructions. The refined 1 minute segment was utilised as the noise aspect when calculating the SNR. The SNR of all 5 architectures was noticeably higher than that of the SNR for the original signal. The higher SNR's indicated a noticeably different signal reconstructed to that of the original signal for which the SDAE's were exposed to. Such differences in the SNR's

were further exacerbated as the SDAE depths increased which was evident in Figure 4.1.

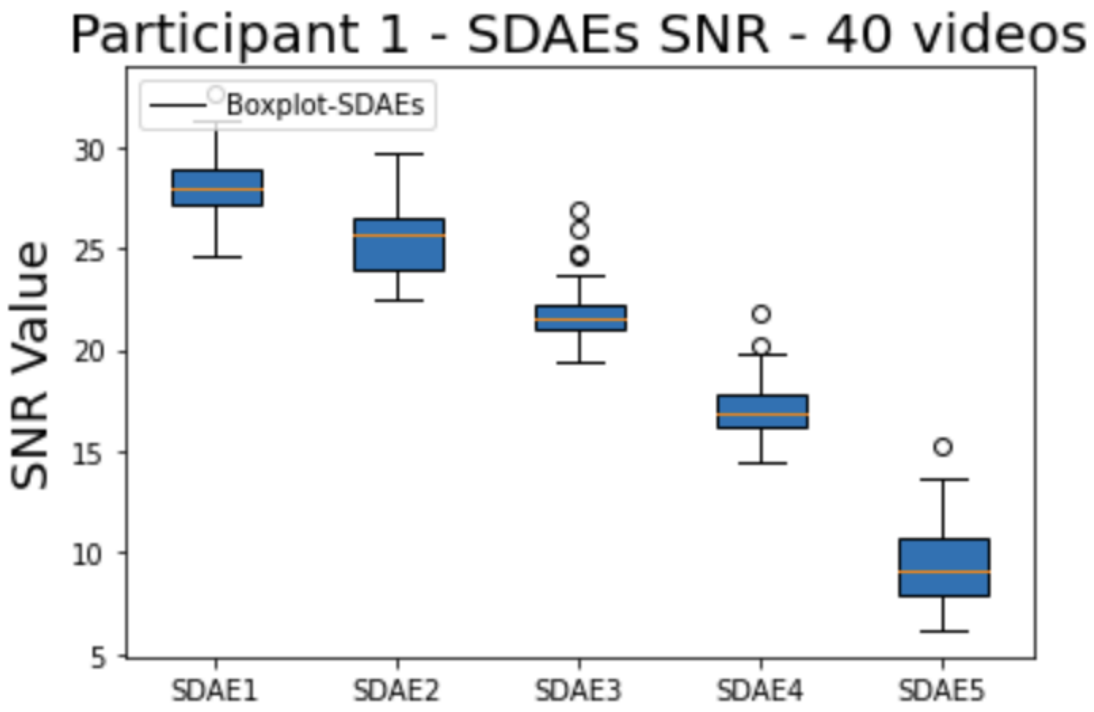

Figure 4.1: Participant1-SDAEs-SNR

The core metric to determine a successful reconstruction using the SDAE architectures was the ability to obtain a significant correlation with the original signal. Correlation Coefficients across all channels were calculated and averaged along with the P value for each that determined the significance of the results obtained. The Spearman correlation coefficients showed association between channels of the original signal and that of the corresponding SDAE reconstructed signal. The statistical test numerically determined the extent to which changes were mirrored and the direction was determined by the sign that supported the metric, whether it be positive or negative. The scale used was -1 to +1 so either end of the spectrum indicated a strong correlation in that the changes were mirrored but the direction was dependent on the sign returned so as one increased, the other either increased or decreased. The results seen in table 4.1 indicate no correlation exists across all SDAE architectures when exposed to the Spearman Rank Correlation test. A strong correlation would be above the +0.5 or

below -0.5 thresholds but neither of these were satisfied in the values returned. It must be noted that the averaging removed the nuanced nature of a channel wise understanding but as a whole, there existed no relationship between the original signal and the reconstruction.

The SDAE chosen was that of SDAE4. This was determined as the chosen architecture within the scope of the research as its performance albeit underwhelming over 3 participants, had a level of stability not seen in other models where the CC was similar across the 3 participants and a similar behaviour was noted for its SNR. The stability was better than that seen on other architectures, namely SDAE5 and SDAE3 where participant 3 had a CC that varied wildly to the CC returned on participant 1 and 2. Variation in the SNR which was indicative of a changing signal which would mean an excessively large SNR compared to that of the original signal was not desired if correlated reconstructions were to be obtained. Given all SDAE architectures returned higher SNR than that of the original signal meant no meaningful information could be taken from these comparisons. The stable under performance of SDAE4 allowed for the potential impact of more data improving the correlation and significance of the results seen. The architecture of SDAE4 provided the best avenue to obtain a reconstruction that was correlated to the original signal as the depth appeared to affect its ability of reconstructing the original signal so a shallower architecture was preferred.

4.1.2 Modelling - SDAE Architecture Versus ICA

Building on the selection of SDAE4 as the optimum architecture within the framework of this research, a further 17 participants data were reconstructed using the SDAE4 architecture which supplemented the 3 participants already reconstructed by the architecture. This allowed the randomizer to select any time segment within the 20 participants as there existed a reconstruction for the random stimuli exposed segment. The randomizer selected specific segment that were taken from the original data, fed into the ICA algorithm where the signal was decomposed and its components evalu-

ated for the presence of eye blink artefacts which were subsequently zeroed out and the then inversely transformed to form the post ICA reconstruction. At this point, there was the original signal, the post ICA reconstruction and the SDAE reconstruction all available for evaluation. The SDAE reconstruction was evaluated with both the original signal from which it was constructed and the post ICA reconstruction that it was to be evaluated against.

Looking at the original signal versus the SDAE reconstruction using the random time segment. Sections of the original signal appeared mirrored in the SDAE reconstruction but these were present at lower amplitudes, infrequent and there existed a noticeable reduction in amplitudes overall.

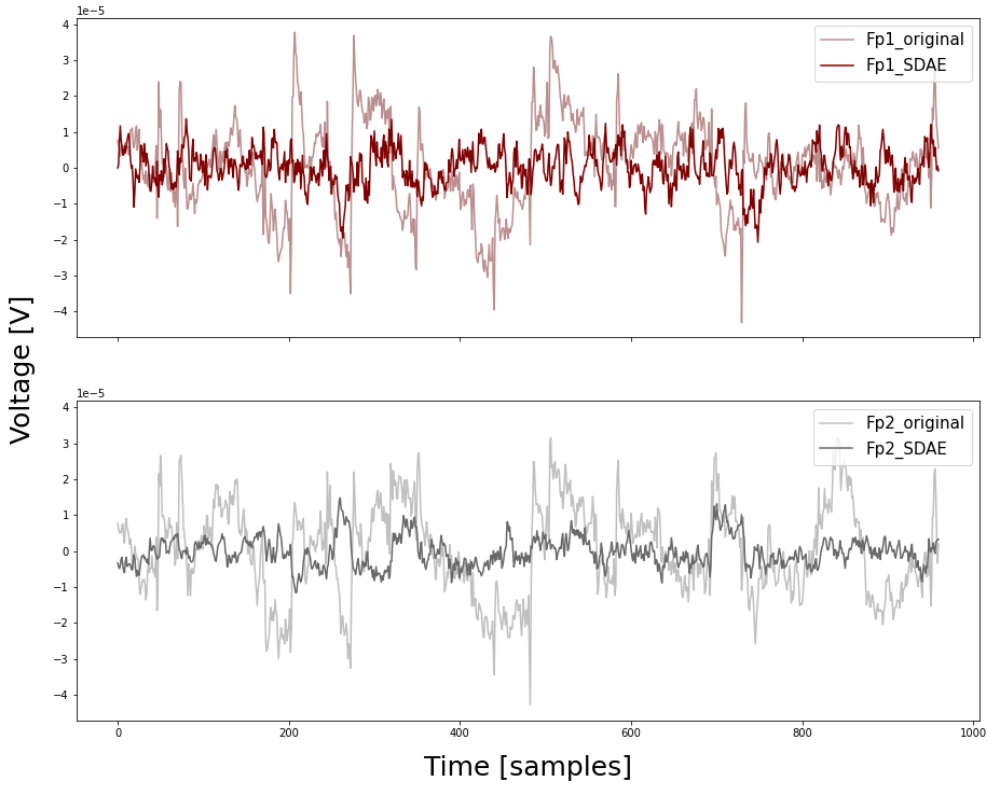

Figure 4.2: P3-V26-S1-T1-Original-vs-SDAE plots Fp1-Fp2

The reduction in amplitude was potentially attributed to the removal of eye blink artefacts as such artefacts were shown to have increased amplitudes as was evident in their profiles. If the reduction in amplitudes while the lower amplitude mirroring of the original signal persisted, then the initial observations would be promising for the SDAE's ability to remove eye blink artefacts albeit it may not manifest into a strong correlation, it may be able to remove eye blink artefacts which distort it to such an extent that it doesn't resemble the original signal. This however, would appear promising in it's ability to correlate to the post ICA reconstruction which removes said eye blink artefacts so its resemblance to the reconstruction may be stronger than that of the original signal. Further random segments were evaluated and these did not support this initial observation, amplitudes were preserved in the SDAE reconstructions when plotted against that of the original signal for these random segments as seen in Figure 4.3 and Figure 4.4.

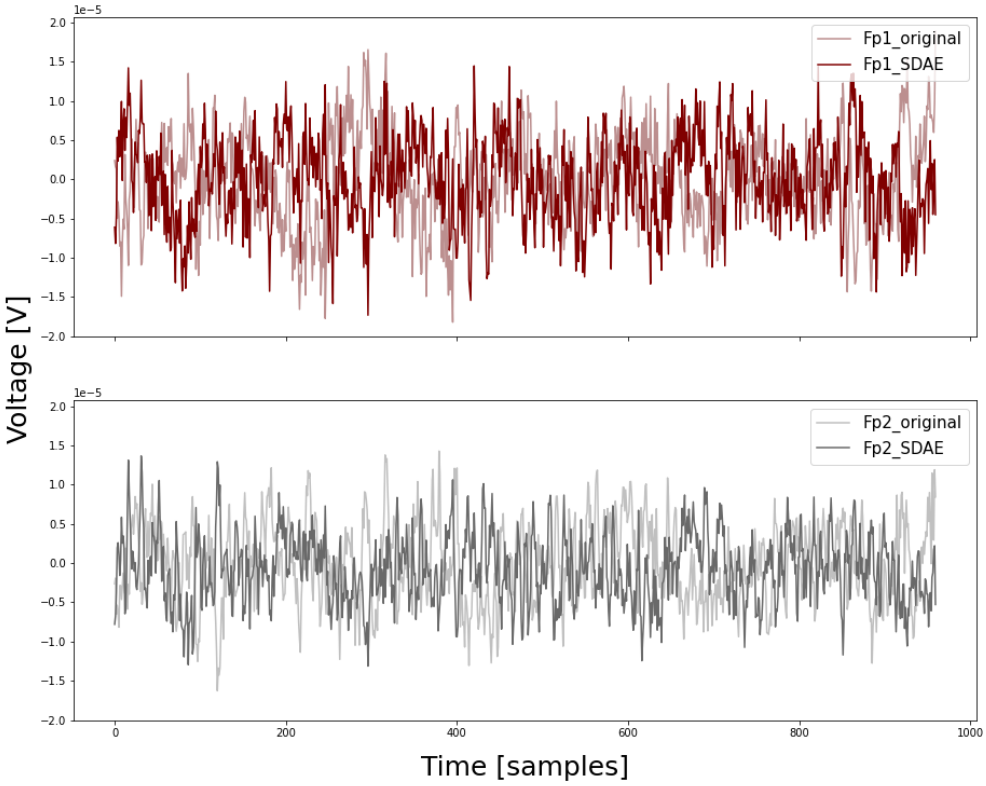

Figure 4.3: P1-V8-S3-T2-Original-vs-SDAE plots Fp1-Fp2

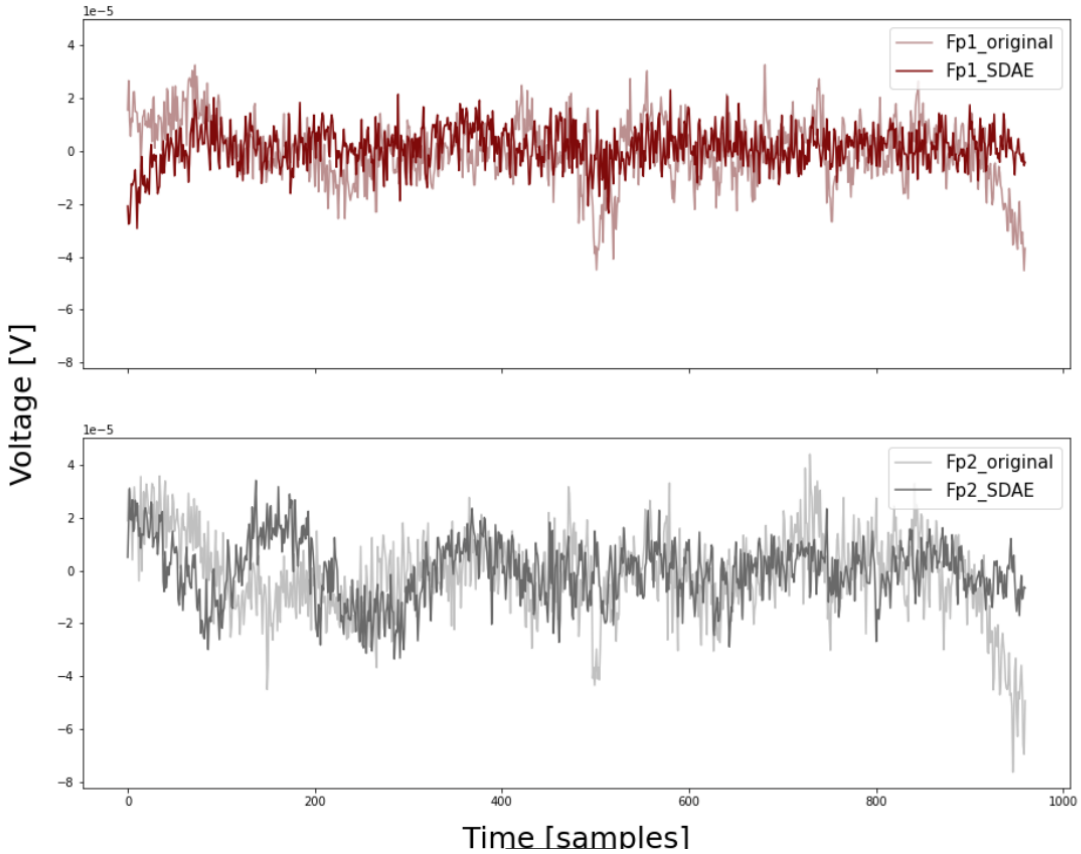

Figure 4.4: P16-V34-S4-T3-Original-vs-SDAE plots Fp1-Fp2

Given the random segment generator specifies what reconstructions were to be evaluated, this required the same segment in the original data to be used for the formation of the post ICA reconstruction. The original signal was used with the Fast ICA algorithm to determine the presence of eye blink artefacts. The 32 channels of the random segment made up the input to the ICA algorithm which meant there were 32 components determined, the ability to graphically depict all 32 together with clarity was unattainable so snippets were chosen to highlight the presence of distinct eye blink artefacts that were subsequently removed in the inverse decomposition which formed the post ICA reconstruction.

In Figure 4.3, ICA component 29 was a distinct eye blink artefact.

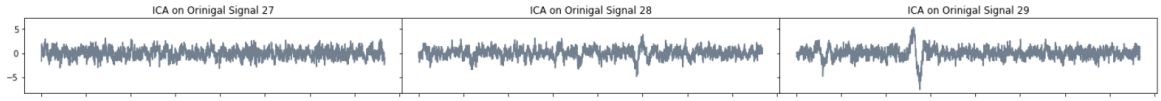

Figure 4.5: P19-V32-S1-T1-ICA-Original

In Figure 4.4, ICA component 17 was a distinct eye blink artefact.

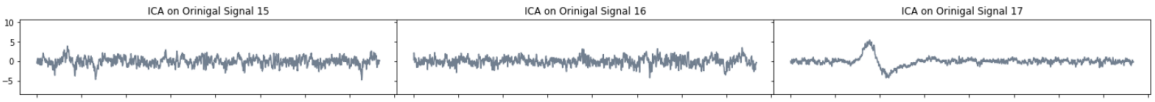

Figure 4.6: P17-V8-S2-T2-ICA-Original

In Figure 4.5, ICA component 18, 21 and 23 were distinct eye blink artefacts.

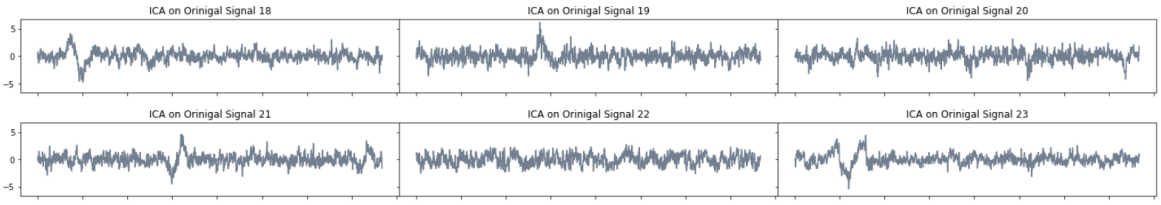

Figure 4.7: P1-V8-S3-T2-ICA-Original

In Figure 4.6, ICA component 3, 8, 11 and 14 were distinct eye blink artefacts.

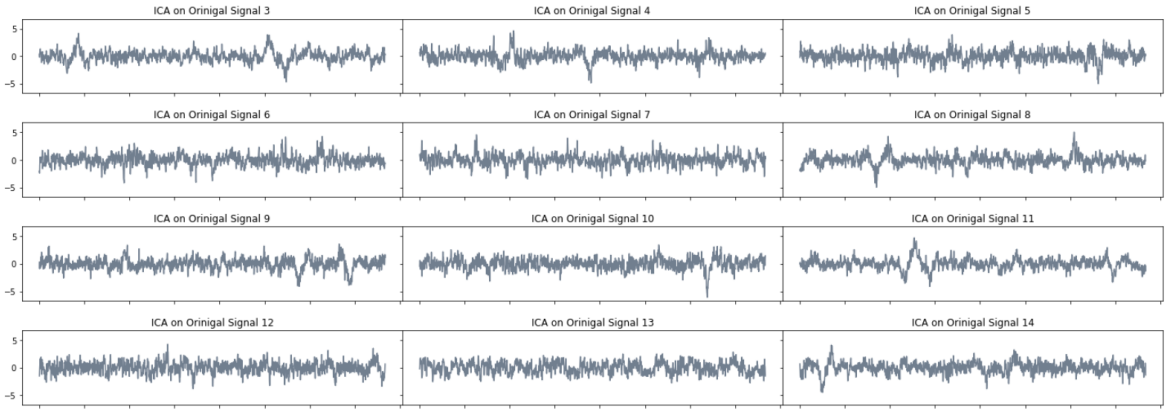

Figure 4.8: P1-V8-S3-T2-ICA-Original

The ocular probes were utilised to confirm the presence of eye blink artefacts. As evident in Figure 4.8 where an eye blink was seen at roughly sample 350 to sample 400. The refined check in a narrowed window allowed for the confirmation of the presence of eye blink artefacts to support those seen in the raw signal. Similar was seen in figure 4.9 where sample 100 to sample 200 where a distinct eye blink artefact was detected. These ocular channels were originally envisaged as being a good source for distinct eye blink artefacts but in reality, they appeared to have a high level of noise present. This may originate from their ability to not only detect eye blinks, but also eye movement which was another ocular characteristic and affects the distinctness of the eye blink artefacts in these channels when 2 were being utilised.

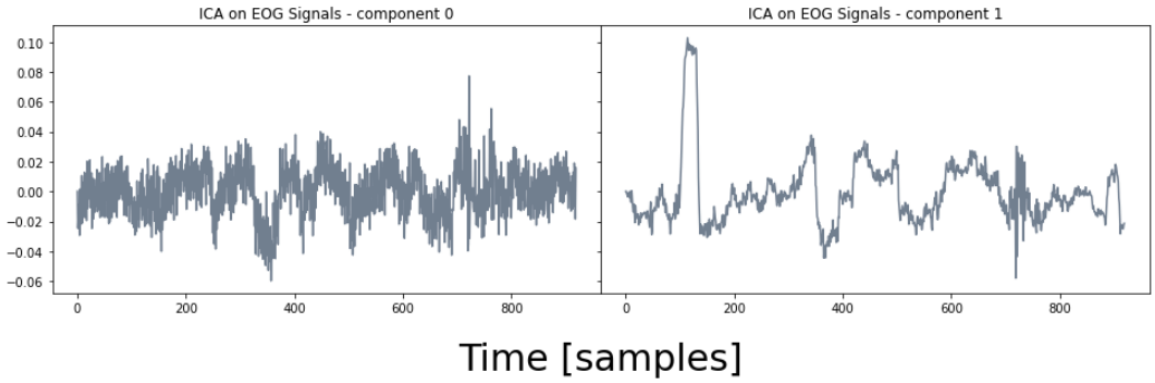

Figure 4.9: P1-V3-S2-T1-EOG

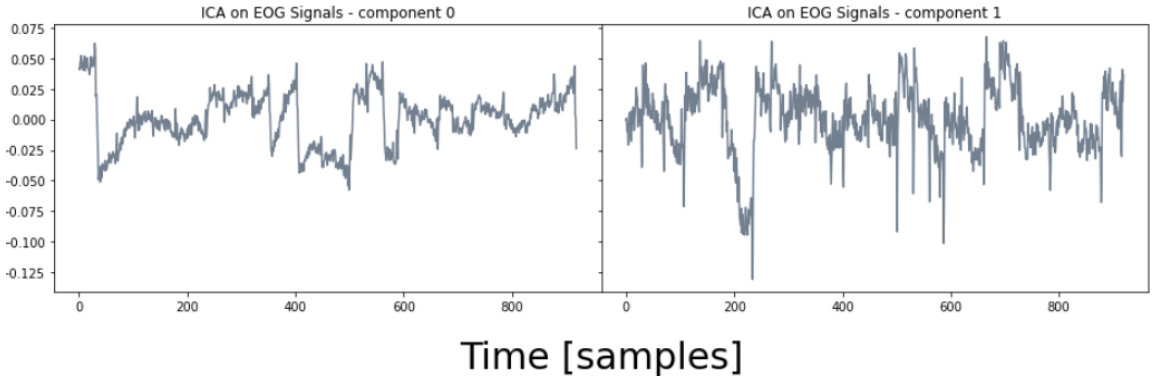

Figure 4.10: P8-V13-S3-T4-EOG

SDAE reconstructions were evaluated on their ability to remove eye blink artefacts.

For this to be achieved, its components needed to be evaluated. The use of ICA was both to establish a reconstruction of the original data that the SDAE reconstruction could be evaluated against and a means of evaluating the SDAE reconstruction for the presence of eye blink artefacts. ICA decomposed the SDAE reconstruction. The components that made up the SDAE reconstructions were evaluated using the defined profiles that were outlined in section 3.4.3 and which were used in the formation of the post ICA reconstruction.

Distinct eye blink artefacts were seen in the SDAE reconstructions when random time segments were evaluated. These were evident in the distinct artefacts of component 15 in Figure 4.7.

Figure 4.11: P19-V32-S1-T1-SDAE4

In Figure 4.8, ICA component 30 was a distinct eye blink artefacts.

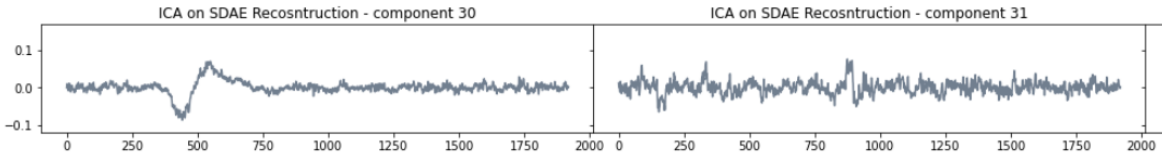

Figure 4.12: P17-V8-S2-T2-SDAE4

In Figure 4.9, ICA component 11, 12 and 14 were distinct eye blink artefacts.

Figure 4.13: P1-V8-S3-T2-SDAE4

Distinctiveness of eye blink artefacts allowed for their detection, the more prominent they were in their appearance allowed for their manual classification as an eye blink artefact. Such distinctiveness was not a given. This was seen in Figure 4.10 where the profile of an eye blink was seen in components 21, 26 and 30, however their amplitude was not increased enough to make such an assertion deterministic. This creates variability in the interpretation.

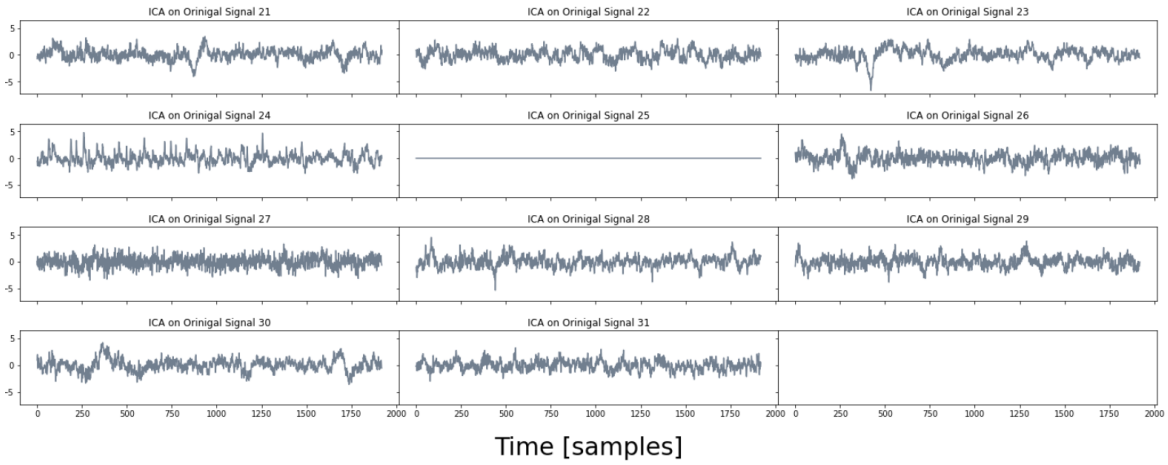

Figure 4.14: P3-V26-S1-T1-ICA-Original

Another instance of similar lack of distinctiveness can be seen in Figure 4.11 where component 0 has the profile of an eye blink in some segments but the amplitudes were not increased enough to be distinct. The absence of distinctiveness means variability manifests as the random segment from the original signal provided, were fed to the ICA algorithm from which the components were evaluated and a post ICA reconstruction was formed. The ocular channels were also fed to the ICA algorithm to confirm the presence of eye blink artefacts that supplement the detection in the original signal. The SDAE reconstruction were then evaluated using the ICA algorithm to determine the presence of eye blink artefacts and finally, the post ICA reconstruction was fed to the ICA algorithm to ensure all eye blink artefacts were removed. This equates to 4 instances where the detection of eye blink artefacts were to be evaluated using the manual classification of their presence in the components conveyed. The potential variability in interpretation as the presence of eye blink artefacts were not always

distinct as was highlighted here.

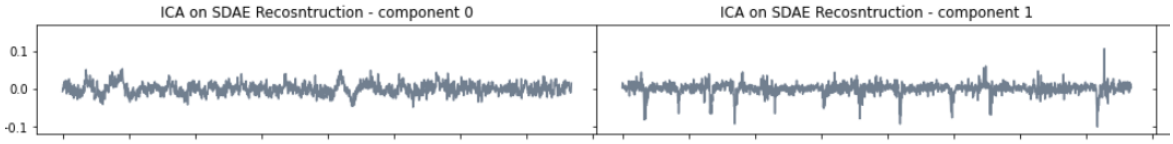

Figure 4.15: P17-V8-S2-T2-SDAE4

Visual evaluation of the SDAE reconstruction to that of the post ICA reconstruction allowed for an initial understanding of their similarities. The ability of the SDAE reconstruction to mirror the post ICA reconstruction for which eye blink artefacts were visually confirmed and subsequently removed meant that the SDAE achieved a similar feet to that of the manual ICA approach. Evaluation of such visuals required a more refined view and with this, a narrowed view using relevant channels was implemented for such plots. Such plots can be seen in Figure 4.12 where random segments were selected and the channels of "AF3"/"AF4" and "Fp1"/"Fp2" were chosen given their proximity to the eyes in relation to the scalp. These plots showed some segments within the depicted time period were mirrored with a degree of accuracy where as other segments were not. The initial observation showed some signs of promise as the reconstructions mirrored the characteristics of the signals in some instances but given the stochastic nature of EEG data, statistical evaluation later determines the merits of the initial observation as this may be random in nature or not refined enough in these plots which would not support the SDAE's ability.

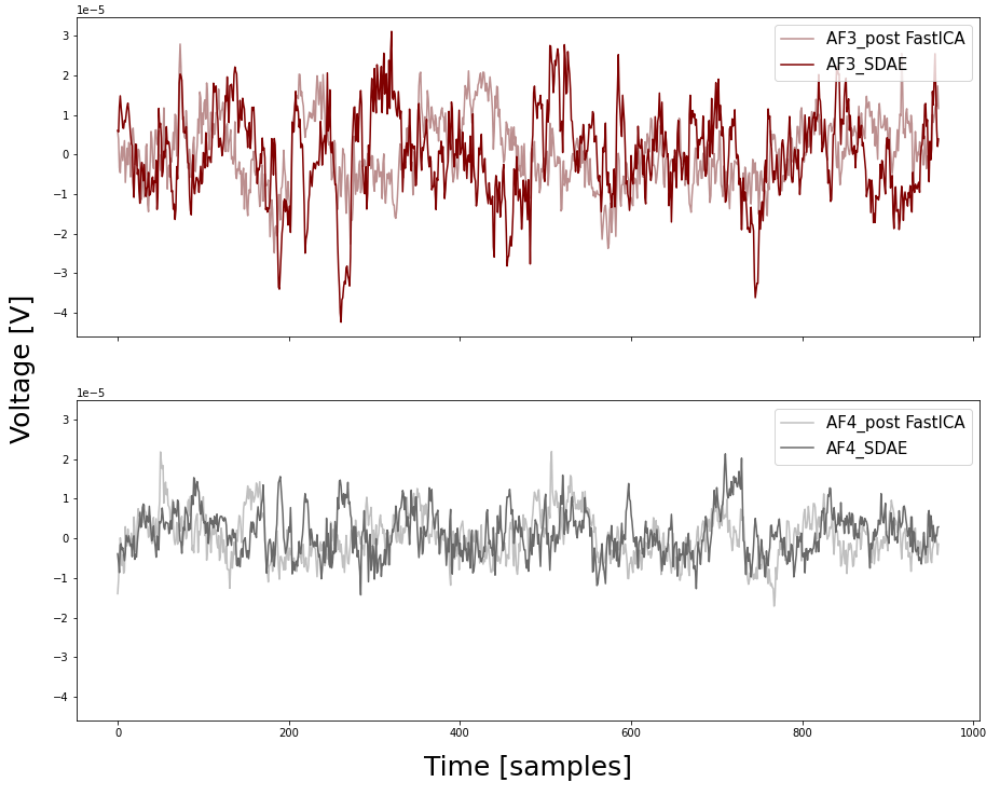

Figure 4.16: P3-V26-S1-T1-postICA-vs-SDAE plots AF3-AF4

Similar observations were seen in Figure 4.13, Figure 4.14 and Figure 4.15. The reconstructions could not mirror the post ICA construction consistently but did achieve elements in extended time segments where the lower frequency elements were reconstructed with reasonable visual accuracy.

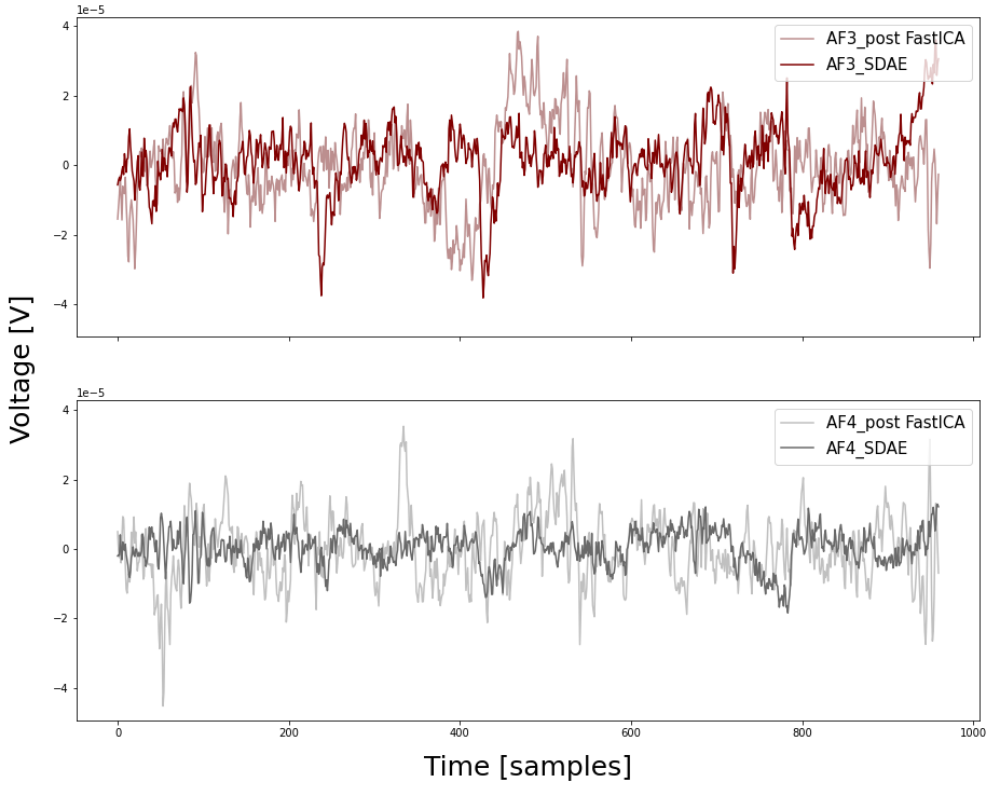

Figure 4.17: P17-V8-S2-T2-postICA-vs-SDAE plots AF3-AF4

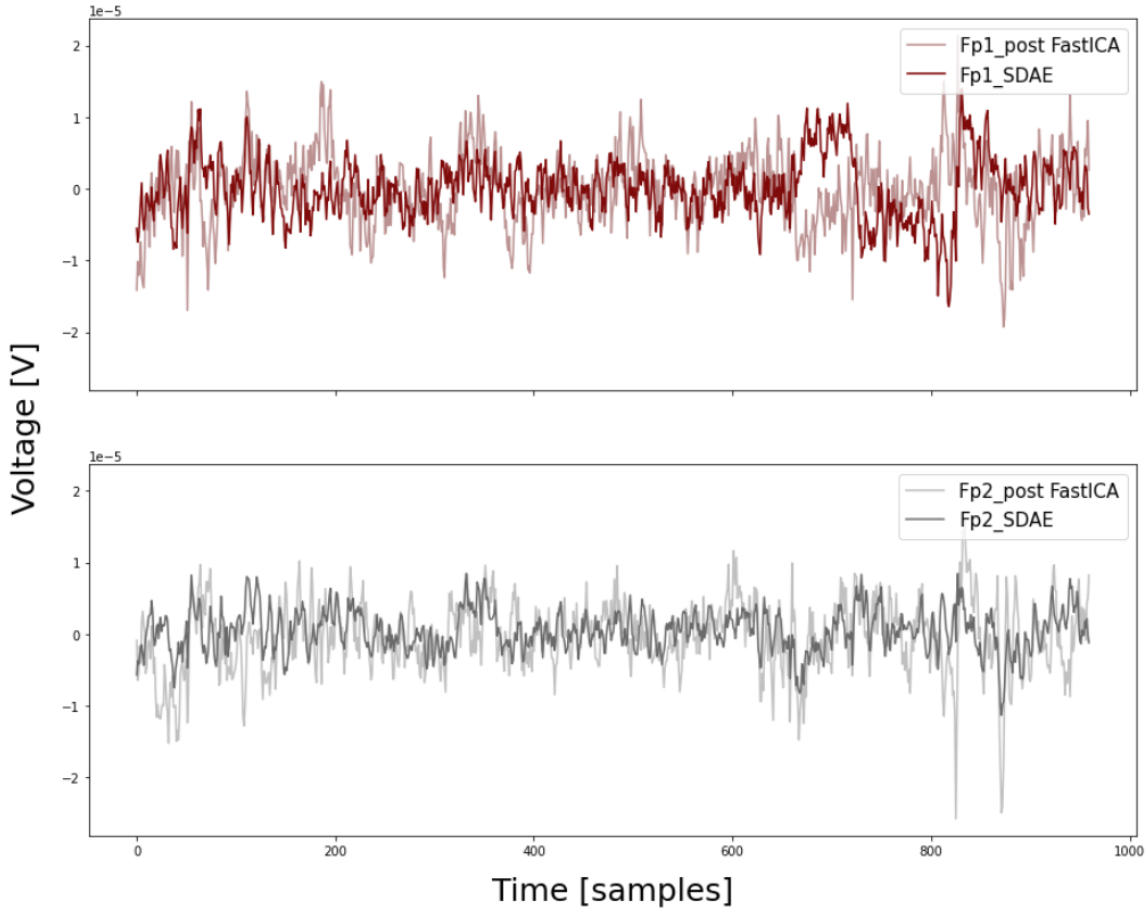

Figure 4.18: P1-V3-S2-T1-postICA-vs-SDAE plots Fp1-Fp2

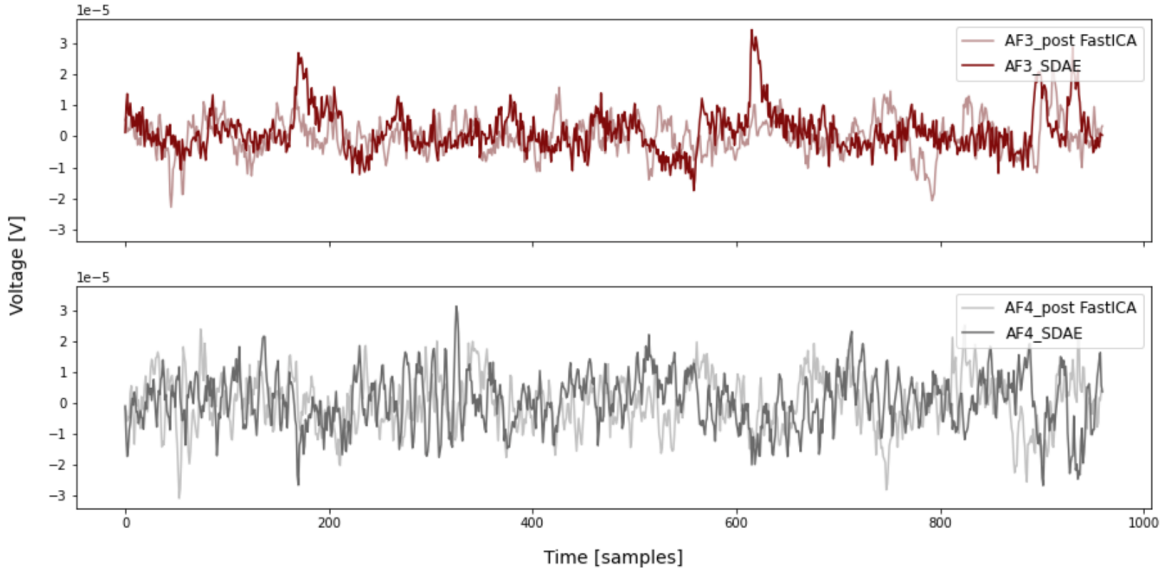

Figure 4.19: P4-V39-S4-T2-postICA-vs-SDAE plots AF3-AF4

Another visualization tool utilised to visually inspect the similarities in the reconstructions were seen in the scatter plots obtained. Such a plot would indicate an initial understanding of the potential for correlation between the reconstructions if such a relationship existed. This required a channel wise comparison between the reconstructions using "Fp1" and "Fp2" respectively. In Figure 4.20, showed signs of correlation on "Fp2" between the reconstructions where as Figure 4.21 showed no sign of correlation on the neighbouring channel of "Fp1" for the same segment. This initial view on a whole showed a visual lack of any form of correlation between the reconstruction, however, Figure 4.20 and Figure 4.21 showed the nuanced nature that may be lost in the averaging of channels.

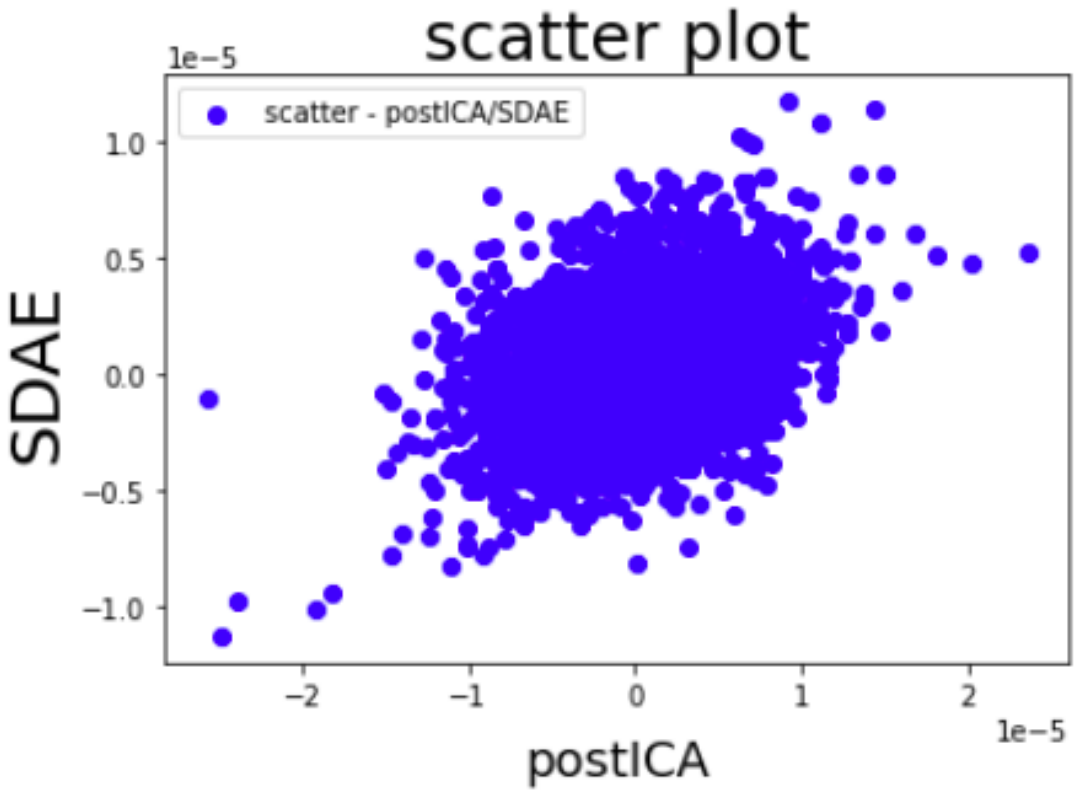

Figure 4.20: P1-V3-S2-T1-postICA-vs-SDAE scatter plots Fp2

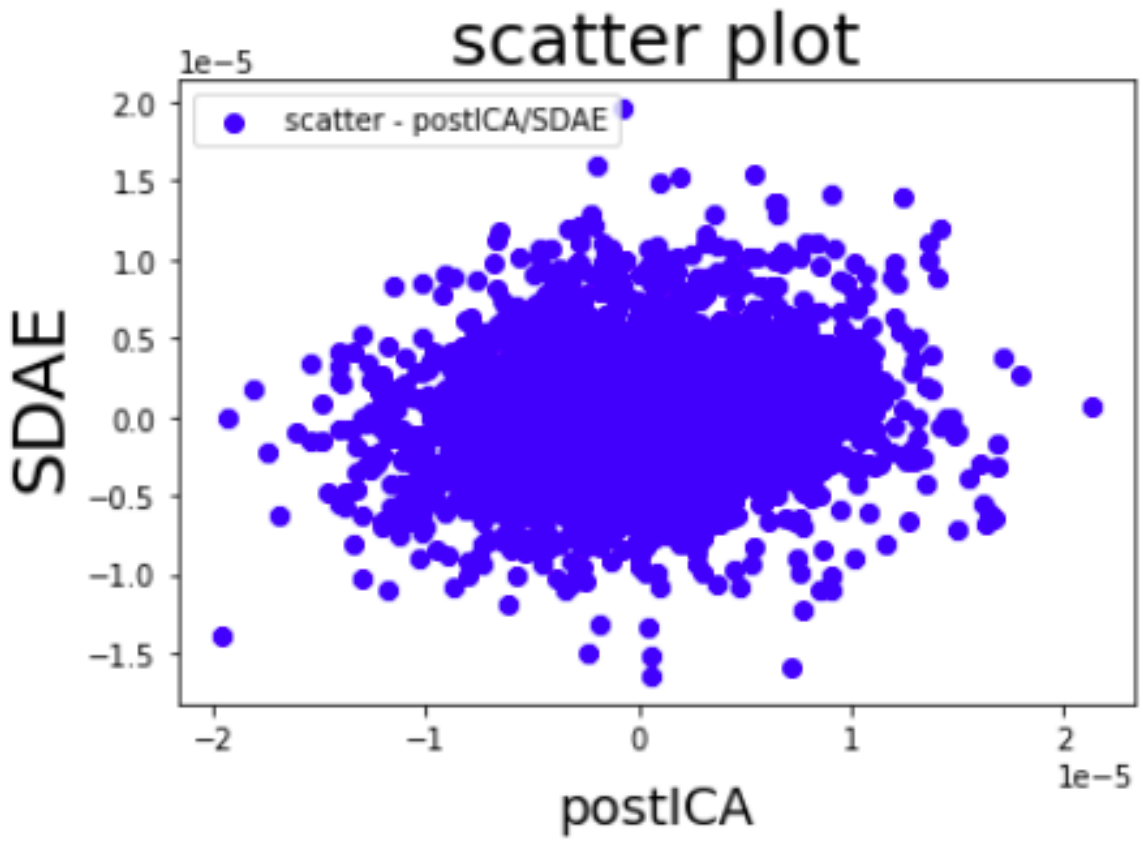

Figure 4.21: P1-V3-S2-T1-postICA-vs-SDAE scatter plots Fp1

The detection of eye blink artefacts allowed for a valid comparison as a manual reconstruction with the eye blink artefacts removed was utilised as a somewhat desired signal for which the SDAE reconstruction should have a strong correlation with. Such a correlation indicates a relationship but does not infer which reconstruction was perceived as better or correct. It merely established a relationship between a known reconstruction with eye blink artefacts removed and a reconstruction that tried to leverage it's internal layers to remove redundant features for which eye blink artefacts would fall under if the SDAE showed potential via a strong correlation with the manual approach. This was further supported by the detection of eye blink artefacts in the SDAE reconstruction that were decomposed to evaluate their ability to remove known eye blink artefacts present in the original signal from which the SDAE reconstruction was formed from. The visuals displayed from Figure 4.2 to Figure 4.21 highlighted

just a snapshot of a wide range of evaluations on random segments that were best conveyed using these examples to support the findings.

4.2 Evaluation

The means of evaluation were in 2 forms, the first of these being the statistical evaluation approach using the Spearman Rank Correlation Coefficient and its supporting P value to infer significance of the correlation along with the signal to noise ratio's and the second being the human centred evaluation that utilised eye blink artefact profiles that allowed for the inspection of the SDAE reconstructions to determine it's ability to remove eye blink artefacts compared to that of a post ICA reconstruction using an inverse RMSE value to provide some numeric indication of success across several random segments.

4.2.1 Randomised Segment Evaluation Metrics

Using the random segment generator that directed the model to a specific reconstruction within a respective time slice, segment, video and participant meant that this was extracted and used as the signal for evaluation. The same segment was referenced in the original raw data where it was fed into the Fast ICA algorithm where its components were evaluated against defined eye blink artefact profiles and subsequently zeroed out if deemed to be indicative of an eye blink artefact so as to create the post ICA reconstruction. The values seen in Table 4.2 show the Spearman Rank Correlation Coefficient and subsequent P value for the specified SDAE reconstruction to that of the post ICA which was formulated using a channel wise comparison and averaged across all 32 channels for this random segment. Formation of the SNR was done with a similar channel wise comparison averaging to that of the baseline noise for both the SDAE and post ICA reconstructions. This was repeated across several random segments as seen in this table.

	CC	**Pvalue**	**SNR-ICA**	**SNR-SDAE**
P9-V17-S3-T2	-0.00298	0.20002	-3.098	10.394
P16-V25-S3-T1	0.003146	0.269869	-9.173	7.747
P16-V34-S4-T3	0.000184	0.261623	-9.849	3.378
P8-V21-S2-T2	-0.000717	0.267415	-15.619	5.656
P5-V24-S5-T1	0.021091	0.291013	-9.810	2.619
P19-V29-S1-T1	-0.001611	0.406250	1.992	22.323
P13-V37-S3-T3	-0.007164	0.272461	-1.193	20.430
P12-V5-S2-T2	0.004216	0.234576	1.332	26.718
P2-V30-S2-T2	-0.008426	0.077936	-1.336	20.174
P3-V24-S5-T15	-0.019680	0.126283	2.489	2.831
P8-V13-S3-T4	0.003847	0.218728	-10.055	6.254
P1-V3-S2-T1	0.02976	0.09588	-5.062	18.496
P20-V23-S1-T1	-0.006380	0.128978	-7.188	23.722
P18-V21-S5-T5	-0.001948	0.196853	4.738	-0.210
P7-V10-S4-T6	0.007227	0.377391	-9.930	7.012
P6-V8-S2-T1	-0.001707	0.330812	1.675	8.747
P4-V39-S4-T2	0.004880	0.215398	-6.867	6.490
P11-V14-S3-T1	0.008390	0.40400	-11.238	4.556
P3-V26-S1-T1	-0.011285	0.027735	-4.835	22.912
P15-V24-S4-T1	0.019257	0.207268	-1.840	6.123
P9-V8-S2-T1	-0.003797	0.234361	-0.969	11.713
P19-V32-S1-T1	-0.002633	0.185613	-7.323	21.076
P14-V16-S5-T1	0.0155508	0.255328	-6.822	5.884
P17-V8-S2-T2	-0.002808	0.143782	-8.478	11.650
P1-V8-S3-T2	-0.000362	0.046056	-12.722	4.259

Table 4.2: Evaluation metrics

4.2.2 Correlation Coefficient

The core statistical evaluation was that of the Spearman Rank Correlation Coefficient. Information in Table 4.2 showed aggregated correlations of the 32 channels between the SDAE and post ICA reconstructions. The averaging was consistent across time slices used which range from 7680 data points in "P19-V32-S1-T1" to 480 in "P18-V21-S5-T5". There exists no correlation between the segments of the SDAE reconstructions and that of the post ICA construction when statistically evaluated using the Spearman Rank Correlation test. All random segments shown in Table 4.2 have correlation values below 0.03 which indicates no correlation exists in any of the segments when comparing the SDAE reconstruction to that of the post ICA. Reasons for such could be due to the averaging of the channel wise correlations which masks the nuance of a channel wise performance where more stimulated channels perform worse due to the stochastic nature of EEG data being harder to form a correlation between two different types of reconstructions of the same data. Other nuances that can be seen relate to the smaller data segments having a better average Correlation Coefficient. The shorter segments do appear to perform better as seen with "P5-V24-S5-T1 = 0.02" and "P14-V16-S5-T1 = 0.015" on average which were both 480 data points while the larger segments, "P20-V23-S1-T1 = -0.006" has a worse performance however one instance of a larger segment appeared to show some weak correlation "P1-V3-S2-T1 = 0.029" but this was a one of case and not consistent with similar segment sizes. This allowed for an interesting consideration to be raised, the smaller segments are exposed to less eye blink artefacts which meant there were less components removed and in turn, the post ICA construction more closely resembles the original signal from which the SDAE was reconstructed from.

The averaging affect can be seen in Figure 4.22 for some random segments, these appear to indicate a small level of averaging masking some potential correlation as evident in the box plots. Even though on a whole, there exists no correlation, some instances show small levels of correlations.

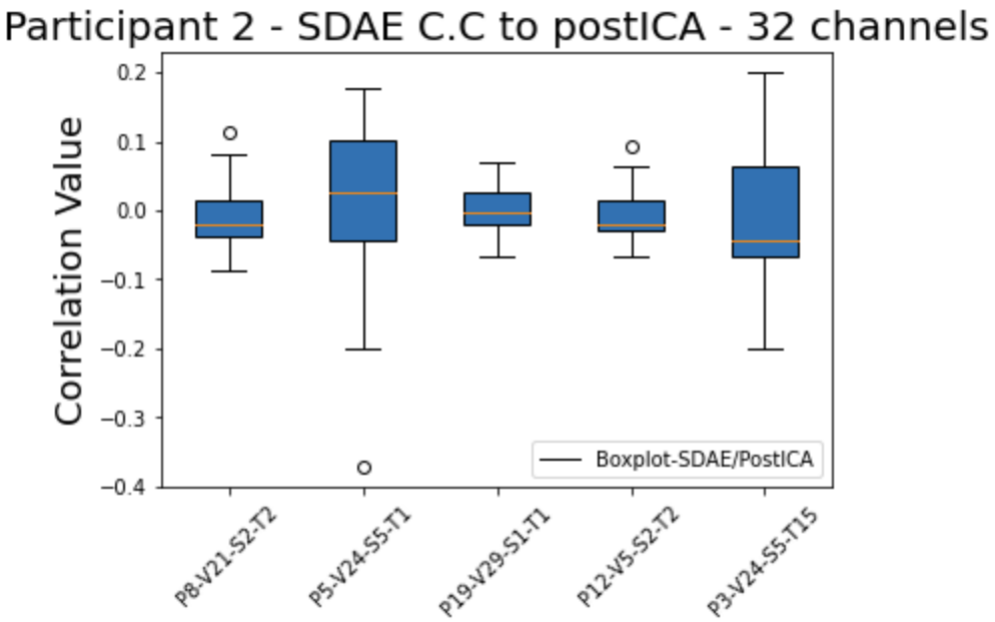

Figure 4.22: CC of SDAE and postICA - Channel wise plots on random segments

4.2.3 Signal to Noise Ratio

The signal to noise ratio looked at the concept of the desired signal and how it differs from the perceived background noise that is a undesired signal. Both the signal and noise can change sharply and in orders of magnitude in EEG data as its stochastic in nature. The SNR used the db scale which looked at the ratio of the signals, greater than zero indicates more signal than noise while less than zero indicates more noise than signal. This ratio was impacted by the numeric scale that the data was utilising, the micro volts of the EEG data means larger separations were still small voltages that makes it more difficult to distinguish brain activity as the impact of noise was more pronounced at lower scales and the assumption of the baseline noise being free of brain stimulation may be a unstable baseline assumption as it was unquantifiable.

Looking at table 4.2, it was distinct the difference in polarity between the post ICA SNR's and that of the SDAE SNR's at initial viewing. Given that the SNR decreasing indicates a loss of signal then one theory was that the ICA process of removing

components was degrading the quality of the signal in its reconstruction. The reason for this can be rooted in the method for eye blink removal in which components with eye blink artefacts present were zeroed out. Signal information that was in this component was removed as was evident when examining the previous plots of the eye blinks in their respective components, these took up a small time sample in the larger sample conveyed so their removal was leading to a loss of signal information. This was further supported when considering the extent of eye blink artefacts removed and the respective SNR returned. P1-V8-S3-T2 had 12 artefacts removed and showed an SNR of -12.722db while P20-V23-S1-T1 had 8 artefacts removed and showed a SNR of -7.188db. These highlight the affect the eye blink removal has on the signal when assessed against the baseline noise. This brings up a second possible theory in that the baseline noise has inferred that the baseline noise signal only contains noise as there was no stimuli present at its recording but doesn't account for brain stimulation even without stimuli which could be triggered by humanistic traits such as nervousness and apprehension given the experiment they were subjected to. This theory was harder to support as there exists no quantifiable means of determining a segment of the data that was noise free and the variability across participants affects any possible generalising when evaluating in relation to the SNR impact.

The SDAE's SNR vary over the random segments conveyed in table 4.2 with a high of 26.718db for P12-V5-S2-T2 and a low of -0.210db for P18-V21-S5-T5. The higher signal ratio may be due to an increased amplitude in the reconstruction. This can be checked when assessing the Figures that convey the plots of both the post ICA and that of the SDAE's for random segments. Looking at Figure 4.14 and Figure 4.15, there doesn't appear to be any noticeable difference in amplitude, in fact, the SDAE appears to have a reduced amplitude on average when compared to that of the post ICA reconstruction so the higher SNR can not be contributed to a higher amplitude in the reconstruction directly. Another potential reason for the increased SNR laid in the architecture of the SDAE where it has a constrained representation of the original signal so a defined number of features were used to represent the data

which removes aspects perceived as unimportant when it comes to reconstructing the original signal. This in turn, removes unwanted noise from the original signal in the reconstruction which improves the SNR. This concept was further supported when looking at the SNR's returned for the 5 SDAE architectures where the SNR's increased as the architecture got deeper. The constrained representation using a perception of relevance for features that make up its bottleneck improved the SNR but was likely to result in a loss of signal information if the depth was excessive which couldn't be numerically represented.

With no definitive understanding or statistical reasoning observed from table 4.2. Table 4.3 and table 4.4 showed the standard deviations and min-max of both the post ICA reconstruction and the SDAE reconstruction channel wise on random segments. The marked difference in the dispersion channel wise was noticeable which may highlight why the SNR was higher. Over 32 channels, the average value returned may be affected by outliers that distort the value's shown in table 4.2 or potentially there exists large instability in the SDAE's reconstructions which was not just isolated to 1 or 2 outliers.

	S.D-pICA	**S.D-SDAE**
P1-V8-S3-T2	12.3439	13.5293
P17-V8-S2-T2	10.3772	9.4559
P14-V16-S5-T1	15.7521	11.2397
P19-V32-S1-T1	13.7667	11.5098
P9-V8-S2-T1	15.6757	15.8102
P15-V24-S4-T1	21.0326	15.7915

Table 4.3: Standard deviations of the SNR's across channels

	MIN-pICA	**MIN-SDAE**	**MAX-pICA**	**MAX-SDAE**
P1-V8-S3-T2	-42.460	-26.601	34.324	59.710
P17-V8-S2-T2	-33.101	-7.375	11.107	34.435
P14-V16-S5-T1	-41.747	-13.971	34.094	32.172
P19-V32-S1-T1	-34.638	-7.599	23.784	42.664
P9-V8-S2-T1	-25.918	-11.231	28.289	44.152
P15-V24-S4-T1	-81.428	-39.004	29.550	29.619

Table 4.4: Min/Max of the SNR's across channels

The high standard deviations show a high dispersion across the channels which was aggregated to form the values shown in table 4.2. The standard deviation of some random segments shown in table 4.3 highlight the extent of the variation. The range using the minimum and maximum values from Table 4.4 showed the post ICA reconstructions varying wildly while the SDAE had a reduced range in comparison.

4.2.4 Selection of Optimum SDAE

The selection of the optimum architecture was a necessary step in determining a reasonable representation of a SDAE that had some potential to inherently remove eye blink artefacts. Its selection was a comparative study against other model variations using the original signal and their respective reconstructions. The variations in architectural depth were seen as a key means of determining to what extent, does the internal dense layers create complex features that allow for information retention and the removal of eye blink artefacts. It was evident that the optimum architecture was seen as optimum from a comparative study with other variations but it has been shown to be sub optimum in regards to its reconstructions when evaluated. The perceived under performance meant that the selection of the optimum architecture was achieved using a perception of potential when exposed to more data and a closer SNR to that

of the original signal. This didn't detract from the SDAE architecture's utility in the research.

4.2.5 ICA decomposition's of postICA Versus SDAE

Ability to determine the relevance of the decomposition's were tied to the subjective nature of the interpretation of the eye blink artefact characteristics. The variations in the length of the time slices ranging from 3.75 seconds and 480 data points to 60 seconds and 7680 data points meant that an eye blink artefacts detected in the lower range of the time slice segments lengths had less of an impact on the reconstruction when removing these component as the eye blink artefacts were of a certain profile and their characteristics were short in duration so the impact was less in relation to the additional loss of signal information in that component. However, on the higher end of the time slice segment lengths, the impact was exacerbated. The short duration of eye blink artefacts detected on a component spanning 60 seconds meant the potential for loss of signal information was far higher. The extent of this loss was unquantifiable but it was seen in the impact on the SNR's of random segments in table 2.4.

Ability to detect eye blink artefacts using ICA decomposition's allowed for the post ICA reconstruction and SDAE reconstruction components to be evaluated for the presence of eye blink artefacts. Both the post ICA and SDAE reconstructions encounter the same limitations with the eye blink artefact profiles at different stages. The post ICA reconstruction was directly impacted in its formation as these were used to remove unwanted components while the SDAE was evaluated using the same profiles so the variability highlighted affects the ability to properly determine the SDAE's inherent capability to remove eye blink artefacts.

4.2.6 Statistical Evaluation

Statistical evaluation was rooted in the correlations between the post ICA and SDAE reconstructions and the respective SNR's to a perceived baseline noise signal that was assumed. These allowed for numeric representation of the respective capabilities in dealing with eye blink artefacts.

Such an evaluation required an understanding of the context in which they were being used. Correlation looks at the relationship between variables and how they changed, such means of evaluation allowed for the quantification of the relationship between the SDAE reconstruction and that of the post ICA reconstruction. The stochastic nature of EEG data does not impede the ability to analyse its patterns, however, it's nature was that of randomness and uncertainty which makes the ability to determine signal segments free of noise impossible. The baseline noise signal used had a direct impact on a key indicator of determining a meaningful reconstruction which was the SNR. Questions arose about the SNR's returned as these appeared to indicate a loss of signal information in the post ICA reconstructions while these showed a positive affect for the SDAE reconstructions. At initial viewing, it would appear as desired but given that the humanistic evaluation which did not support this, the assumption of noise appeared to have a drastic impact on the statistical evaluation metric of the SNR. With the distinct lack of correlation to that of the post ICA, the statistical evaluation did not support the SDAE's ability to remove eye blink artefacts automatically.

4.2.7 Human Centred Evaluation

The humanistic element was not deterministic as it was a visual inspection using research based profiles obtained to evaluate the presence of eye blink artefacts in the components conveyed using the Fast ICA algorithm which estimated the independent sources from a mixed source signal. A randomised signal generator determined which segments of the data were evaluated using this human centred analysis. Evaluations

of the reconstructions looked at, allowed for the creation of an Inverse RMSE value as there was an expected number of eye blink artefacts that were detected in the raw signal which was then used as the baseline to assess the ability of both the SDAE and ICA to remove eye blinks in their reconstruction's.

The results conveyed in table 4.3 indicate the segments evaluated and the respective values generated. It was evident, that the ICA outperformed the SDAE in nearly all cases. The SDAE showed some potential in the short segments but as the segments were larger, the performance began to really diverge from that of the ICA. The SDAE did remove some eye blink artefacts at the larger segment sizes but this was random and inconsistent when compared to that of the manual approach of ICA.

	RAW	**ICA reconstruction**	**SDAE reconstruction**
P8-V21-S2-T2	2	0	0
P13-V37-S3-T3	6	2	6
P12-V5-S2-T2	6	1	3
P2-V30-S2-T2	2	0	2
P3-V24-S5-T15	1	0	0
P8-V13-S3-T4	6	0	3
P1-V3-S2-T1	4	0	3
P20-V23-S1-T1	8	1	6
P18-V21-S5-T5	1	0	0
P7-V10-S4-T6	1	0	0
P6-V8-S2-T1	6	1	5
P4-V39-S4-T2	4	0	4
P11-V14-S3-T1	2	0	3
P3-V26-S1-T1	9	1	3
P15-V24-S4-T1	3	0	1
P9-V8-S2-T1	7	0	5
P19-V32-S1-T1	8	0	4
P14-V16-S5-T1	1	0	1
P17-V8-S2-T2	4	0	5
P1-V8-S3-T2	12	0	12

Table 4.5: Eye blink components detected

The RMSE value for the post ICA signal against the original signal was 5.205 while the RMSE value for the SDAE against the original was 2.121. This indicated numerically how close to the predicted distribution which was the number of eye blink artefacts detected in each random segment referenced in Table 4.3, the post ICA and SDAE reconstructions were when assessing the presence of eye blink artefacts in their respective reconstructions. Given a larger values indicated a distribution further from

the predicted line meant that this indicated which performed better and in this case, the post ICA outperformed the SDAE by a considerable margin.

4.3 Summary

Variations in architectures were proposed for the stacking of dense layers in an Autoencoder. The architectural variations were evaluated on the first 3 participants where they were tasked with reconstructing the raw EEG segments. These were evaluated comparatively using the C.C to the original signal and the SNR using the assumed baseline noise. SDAE4 was seen as the optimum architecture from the 5 evaluated.

The chosen architecture of SADE4 was then comparatively evaluated with the manual ICA approach for eye blink artefact removal. Ability to remove eye blink artefacts was based on the understanding that such artefacts would make up part of the components that can be separated from the reconstructed signal using the ICA algorithm. Formation of the post ICA reconstruction was achieved using the manual detection of eye blink artefacts which were classified as such using visual profiles supported from other research. SDAE reconstructions were formed with the potential for eye blink artefact removal using its constrained latent representation from which the reconstructions were created. Statistical evaluation of both reconstructions was achieved using the Spearman Correlation Coefficient test and the SNR's. Presence of eye blink artefacts was determined humanistically using the ICA algorithm to convey the components that made up both the SDAE reconstruction and the post ICA reconstruction. The presence of eye blinks were confirmed from the raw signals and EOG channel signals, this allowed for the formation of a numeric indication of the ability to remove eye blink artefacts as this formed a known baseline and the values returned showed the ability to account for these in their respective reconstructions.

To test the hypothesis, both the statistical test and the human centred evaluation were used. The Spearman Rank Correlation Coefficient test returned an average correlation

value across a range of random segments of 0.00184 with a P value of 0.21902 across the same range. This indicated no correlation existed between the post ICA reconstruction and that of the SDAE reconstruction. The Inverse RMSE generated from a random cohort of segments seen in table 4.4 showed the post ICA reconstruction performing far better than that of the SDAE when it came to removing known eye blink artefacts present. This displayed the inability of the SDAE to remove known eye blink artefacts present in random segments evaluated. With both the statistical and human centred evaluation elements considered, it must be concluded that there does not exist enough evidence to support rejecting the null hypothesis.

4.3.1 Strengths of the Results

Visual interpretation: The ability of ICA to allow for a visual representation of the components that formed the reconstructions meant a level of interpretability of the work could be established. Neural networks have an unexplainability characteristic to them so the fact that the reconstruction can be interpreted in its use visually allowed for an understanding to how the architecture was affecting the data in its reconstructions. This was further emphasised in the variations in the architecture that were statistically evaluated and which was supplemented with an extensive evaluation of a single chosen architecture so as to provide the best potential for determining the extent of the results.

Raw data: The use of raw data with minimal preprocessing that was defined meant the work had a generality about it. There was nothing assumed about the data in its reconstruction and only its use was where assumptions were made. This allowed for repeatability to some extent as the process of reconstructing the original signal was not altered but the means of interpreting the reconstructions in the determination of the presence of eye blink artefacts was where repeatability was not present or applicable.

Randomization: The randomization of what segments were evaluated meant the work

displays a fair and balanced view of the capability of the SDAE model. Specific segments were not chosen and were instead generated by a randomizer which dictated what segments were to be evaluated.

4.3.2 Limitations of the Results

Baseline noise: The assumption that the first 2 minutes of the recordings was seen as the baseline noise has a direct affect on the interpretation of the results. Given the fact that there exists no ground truth for a noise free EEG signal then this assumption of stimuli free signal being the noise was valid but it doesn't account for a humanistic element of brain stimulation even without external stimuli. This directly impacted the SNR which was a key numeric indicator for the comparisons of both the post ICA reconstruction and that of the SDAE reconstruction.

Potentially poor architecture: The comparisons of the 5 SDAE architectures was achieved by evaluating their respective ability in reconstructing the original signal. The approach meant that these were evaluated against one another but there was nothing to say that the evaluation resulted in a good architecture, the results showed all architectures returned meaningless reconstructions within the scope of this work which meant an underwhelming model was selected for evaluation against the ICA approach.

Feature formation: The SDAE's ability to remove eye blink artefacts was based on an understanding the such artefacts would appear as features within the latent representation and subsequently be removed as the latent representation became constrained. This was dependent on the data and given the random nature of eye blink artefacts which are non cyclic and non periodic, the randomness of their presence affects the ability of the SDAE to determine such as features.

Eye blink profiles: The detection of eye blink artefacts was achieved using ICA. The ability to detect required an ability to distinguish which proved non deterministic. This

created variability as the profile's used allow for the detection of distinct artefacts but there were numerous eye blink artefacts that were not distinct and so a humanistic interpretation was required.

Spatial resolution: Eye blink artefacts were to be detected within stimuli induced data recordings, given the 32 channels used for the recordings meant that 32 components could be detected using the ICA algorithm. This limits the ability of the ICA as was supported in Klug and Gramann, 2020 where the work stated "it can be stated that a higher scalp electrode density generally leads to a better ICA decomposition". The spatial resolution of EEG probes plays a role when looking for a specific type of artefact as such artefacts are generally more pronounced in a defined region of the scalp which goes against the ability of ICA which is grounded in the concept of a global artefact which indicates that the artefact sought must be present on several channels and not just isolated to 1 or 2 channels. Any limitation in ICA affects the perceived ability of the SDAE as it is this that it is being evaluated against.

Extent of evaluation: Given the extensive reconstructions and time intensive nature of running the ICA algorithm on segments of data, randomly selected segments were evaluated for the presence of eye blink artefacts. Such randomisation removes potential bias but requires extensive testing to ensure the performance is adequately portrayed as the use of a small cohort of results has the potential to inaccurately convey a strong or weak performing model.

Uncontrolled convergence: The SDAE was trained using the concept of convergence of the MSE loss, the performance was not thresholded so it may converge at a poor point. No evaluation of its error was considered outside of the convergence as the work looked at the inherent ability to reconstruct a signal for who's merits were assessed based on the presence of eye blink artefacts.

Chapter 5

Conclusion

5.1 Research Overview

The ability to automatically detect and remove eye blink artefacts in raw EEG data would allow for the removal of the manual step that still exists in artefact removal when using EEG data. This not only streamlines the process but removes the ambiguity in understanding what was seen as an eye blink artefact and what was not. If the automatic method could show results that match or exceed that of the manual implementation then its implementation was valid as it required no time intensive ICA decomposition, manual detection, removal and inverse transformation to form a reconstructed post ICA signal with these artefacts removed. This would be more time effective and allow for the eradication of ambiguity in interpretations.

5.2 Problem Definition

Looking at the high level consideration of manual versus automated approaches to preprocessing data, the ability to automate with minimal sacrifice to performance is thought of as desired. This work falls under such instances. Eye blink artefacts are

distinct in nature and a known human characteristic so their presence are prevalent in all EEG recording that exceed a few seconds. The ability to establish an extent to which these were removed without the manual intervention was where this research originated.

In order to determine whether such an undertaking could be achieved, a method that allowed for such automated removal was required. Given the variable and non cyclic nature of eye blink artefacts along with the spatio-temporal relation of EEG channels, the ability to learn was required and such learning's could not be in isolation, these required interconnections to mirror the human brain so to develop the concept of what was relevant in the EEG data provided. This led to the SDAE being the proposed methodology with varying depths to locate the most promising depth that determined the amount of learning that allowed for signal information retention while removing eye blink artefacts and from this point, the chosen architecture was then evaluated using a comparative evaluation against a manual implementation that manually tried to achieve the same goal of eye blink artefact removal.

5.3 Design/Experimentation, Evaluation & Results

The DEAP dataset provided continuous data that was segmented to 60 second videos of known stimuli exposure. Non stimuli relevant data was present also which allowed for an assumption of a noise signal as part of this work. Further segmenting of the data allowed for shorter duration's within the 60 second videos to be assessed to create a more refined view which allowed for the impact of the number of artefacts being detected on the statistical evaluations to be understood and these could in turn, be restructured back to the 60 second videos. The first 20 participants out of the 32 available were utilised and were further refined to the first 3 participants when it came to evaluating the optimum architecture within the constrains of this research. This involved comparative evaluation of 5 architectures of varying depths on their ability

to reconstruct the original signal for which it was trained on but with the nuance of its intended use which was not a perfect reconstruction as the known eye blink artefacts somewhere in the original signals get removed through the internal learning's of the SDAE when constructing its internal representation at a lower dimension. Both Correlation to the original signal and SNR using the assumed noise signal of a known non stimuli segment were used to determine the chosen architecture.

With the chosen architecture, the remaining 17 participants data were reconstructed using the architecture. Given the exponential growth of segments to be evaluated across 20 participants, 40 videos per participant and then the further segmenting within these videos, random segments were chosen and these were both statistically and humanistically evaluated against a manual approach of using ICA to detect for eye blink components within specified segments and in turn, remove these before being reconstructed. This allowed for an original signal segment, a post ICA reconstruction of this segment and a SDAE reconstruction of the same segment to be comparatively evaluated. Statistical evaluation using Spearman Correlation Coefficient showed no significant relationship existed across various segments between the post ICA reconstruction and that of the SDAE reconstruction while the SNR showed the potential for signal information loss when it cam to extensive eye blink artefact removal. A humanistic evaluation showed minimal potential for the SDAE to remove the eye blink artefacts with the performance being poor when evaluated using an inverse interpretation of the RMSE which conveyed large statistical difference between the post ICA reconstruction count of eye blink artefacts present to that of the SDAE reconstruction eye blink artefacts count across various segments.

5.4 Contributions and Impact

This research evaluated the applicability of SDAE's as an automated means of eye blink artefact removal. Through such an evaluation, it was determined that the SDAE was

unable to achieve this in any meaningful way and further to this, showed no signs of potential for any further exploration in its utility to do so. If the artefact that was sought to be removed was periodic and defined then the ability of the neural network to detect such an artefact increases as it uses mathematically means to look for patterns that manifest into features. A repetitive and periodic signal would appear as a feature and with internal constraints of a bottleneck, at some stage in its lowering of its dimensions, it would remove the feature and in turn, achieve the desired task. This was not the case however as the artefact in question were eye blinks that are human characteristics and are not periodic and cyclic, they are sporadic and varying within stochastic EEG signals which adds complexity. This diminished the likelihood of the SDAE being able to achieve meaningful results and even prevents the research from finding at what point can eye blinks be removed via the constraining of the latent representation which was the motivation for the varying depths for the SDAE architectural comparisons.

5.5 Future Work & Recommendations

Future work for the automated removal of eye blink artefacts lays outside of the use of SDAE's. The lack of hyperparamater optimization was a constraint for the model as this affected its ability to learn, however, any potential gain from this would not be extensive enough to form any further merit for the exploration of such a method for the removal of eye blink artefacts.

The concept of decomposition's is a key component to any means of eye blink artefact removal. This research looked at the temporal features of eye blink artefacts when evaluating their presence and this guided the evaluation of the work. Expanding on this in any future work, should leverage spatial features. Such features show the isolated areas of the scalp for which are stimulated in a specified point in time. Leveraging the spatial and temporal features together allows for future work to automate the

removal of eye blink artefacts using ICA for decomposition's, Identification using both spatial and temporal features as ICA disregards spatial and temporal relationships in its decomposition's, automated classification and subsequent removal. This can be achieved using more conventional Machine Learning approaches to that of Deep learning as it is guided by the concept of identification and removal rather than the ability to inherently learn and remove.

If future work sought to continue down the Deep Learning path, such a path should lead towards CNN's. These have a distinct ability to learn important features in the data. Their widespread use with image and object recognition supports their use with such a task and given the highlighted spatial consideration that has huge merit for its inclusion in future work, CNN's provide a promising approach for further exploration. Recognition is another form of identification and this has proven paramount as the first step in removing eye blink artefacts. Such an approach could build a library/lexicon or extension for eye blink artefacts that are developed in recorded data and such a formation could then be leveraged for live recordings similar to the utility of lexicon's in Natural Language Processing which evolves the work beyond the recorded time series data that this work utilised.

Bibliography

Abusharha, A. (2017). Changes in blink rate and ocular symptoms during different reading tasks. *2017 Clinical Optometry Volume 9*, 133–138. https://doi.org/10.2147/OPTO.S142718.

Agrawal, J., Gupta, M., & Garg, H. (2021). Early stress detection and analysis using eeg signals in machine learning framework. *2021 IOP Conference Series Materials Science and Engineering 1116(1):012134*, 1–12. https://doi.org/doi:10.1088/1757-899X/1116/1/012134

Alzahab, N., Apollonio, L., Di Lorio, A., Alshalak, M., Larlori, S., Ferracuti, F., Monteriu, A., & Porcaro, C. (2021). Hybrid deep learning (hdl)-based brain-computer interface (bci) systems: A systematic review. *2021 Brain Sciences 11(1):75*, 1–37. https://doi.org/https://doi.org/10.3390/brainsci11010075

Aquino-Britez, D., Ortiz, A., Ortega, J., Leon, J., Formoso, M., Gan, J., & Escobar, J. (2021). Optimization of deep architectures for eeg signal classification: An automl approach using evolutionary algorithms. *2021 Sensors (Basel) (6):2096*, 1–21. https://doi.org/https://doi.org/10.3390/s21062096

Babiloni, F., Bianchi, L., Semeraro, F., Millan, J., Mourino, J., Cattini, A., Salinari, S., Marciani, M., & Cincotti, F. (2001). Mahalanobis distance-based classifiers are able to recognize eeg patterns by using few eeg electrodes. *2001 Proceedings of the 23rd Annual International Conference of the IEEEVolume: 1*, 1–6. https://doi.org/10.1109/IEMBS.2001.1019019

Ball, T., Kern, M., Mutschler, I., Andreas, A., & Shulz-Bomhage, A. (2009). Signal quality of simultaneously recorded invasive and non-invasive eeg. *2009 Neu-*

roimage 1;46(3), 708–716. https://doi.org/10.1016/j.neuroimage.2009.02.028.

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The prep pipeline: Standardized preprocessing for large-scale eeg analysis. *Front. Neuroinform. Volume 9 Article 16*, 1–20. https://doi.org/https://doi.org/10.3389/fninf.2015.00016

Binnie, C., & Prior, P. (1994). Electroencephalography. *1994 Journal of Neurology, Neurosurgery Psychiatry, 57(11)*, 1308–1319. https://doi.org/http://dx.doi.org/10.1136/jnnp.57.11.1308

Bonita, J., Ambolode, L., Rsoenberg, B., Cellucci, C., Watanabe, T., Rapp, P., & Albano, A. (2014). Time domain measures of inter-channel eeg correlations: A comparison of linear, nonparametric and nonlinear measures. *2014 Cognitive Neurodynamics Volume 8,1*, 1–15. https://doi.org/10.1007/s11571-013-9267-8

Boquet, G., Macias, E., Morell, A., Serrano, J., & Vicario, J. (2021). Theoretical tuning of the autoencoder bottleneck layer dimension: A mutual information-based algorithm. *2020 28th European Signal Processing Conference (EUSIPCO)*, 1512–1516. https://doi.org/10.23919/Eusipco47968.2020.9287226

Charte, D., Charte, F., Del Jesus, M., & Herrera, F. (2020). An analysis on the use of autoencoders for representation learning: Fundamentals, learning task case studies, explainability and challenges. *2020 Neurocomputing Volume 404*, 93–107. https://doi.org/10.1016/j.neucom.2020.04.057

Chatterjee, S., Simita Thakur, R., Narayan Yadav, R., Gupta, L., & Kumar, D. R. (2020). Review of noise removal techniques in ecg signals. *2020 IET Signal Processing. Volume 14, Issue 9.*, 569–590. https://doi.org/https://doi.org/10.1049/iet-spr.2020.0104

Chaumon, M., Bishop, D., & Busch, N. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *2015 Journal of Neuroscience Methods 250*, 1–50. https://doi.org/10.1016/j.jneumeth.2015.02.025

Cho, K.-o., & Jang, H.-J. (2020). Comparison of different input modalities and network structures for deep learning-based seizure detection. *2020 Sci Rep 10, 122*, 1–11. https://doi.org/https://doi.org/10.1038/s41598-019-56958-y

Ellis, C., Sendi, M., Miller, R., & Calhoun, V. (2021). A gradient-based spectral explainability method for eeg deep learning classifiers. *2021 bioRxiv*, 1–6. https://doi.org/10.1101/2021.07.14.452360

Ferri, R., Babilomi, C., Karami, V., Ivano Triggiani, A., Carducci, F., Noce, G., Lizio, R., Pascarelli, M., Soricelli, A., Amenta, F., Bozzao, A., Romano, A., Guibilei, F., Del Percio, C., Stocchi, F., Frisomi, G., Nobili, F., Patane, L., & Arena, P. (2020). Stacked autoencoders as new models for an accurate alzheimer's disease classification support using resting-state eeg and mri measurements. *2021 Clinical Neurophysiology 132*, 232–245. https://doi.org/https://doi.org/10.1016/j.clinph.2020.09.015

Ghosh, R., Biswas, S. K., & Sinha, N. (2018). Automated eye blink artifact removal from eeg using support vector machine and autoencoder. *2018 IET Signal Processing Volume 13.*, 2–9. https://doi.org/10.1049/iet-spr.2018.5111

Glass, K., Frishkoff, G., Frank, R., Davey, C., Dien, J., Malony, A., & Tucker, D. (2004). A framework for evaluating ica methods of artifact removal from multichannel eeg. *2004 Conference: Independent Component Analysis and Blind Signal Separation, Fifth International Conference, ICA, Proceedings*, 2–9. https://doi.org/10.1007/978-3-540-30110-3_130

Gogna, A., Majumdar, A., & Ward, R. (2017). Semi-supervised stacked label consistent autoencoder for reconstruction and analysis of biomedical signals. *2017 IEEE Transactions on Biomedical Engineering. Volume 64*, 2196–2205. https://doi.org/10.1109/TBME.2016.2631620

Goodfellow, I., Benigio, Y., & Courville, A. (2016). *Deep learning* [https://www.deeplearningbook.org/]. MIT Press.

Gramfort, A., Luessi, M., Larson, E., Engemann, D., Stohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hamalainen, M. (2013). Meg and eeg

data analysis with mne-python. *2013 Frontiers in Neuroscience 7(7):267*, 1–14. https://doi.org/10.3389/fnins.2013.00267

Hartmann, K. G., Schirmeister, R. T., & Ball, T. (2018). Hierarchical internal representation of spectral features in deep convolutional networks trained for eeg decoding. *2018 Conference: 2018 6th International Conference on Brain and Computer Interface (BCI)*. https://doi.org/10.1109/IWW-BCI.2018.8311493

Herrera, L., Fernandes, C., Mora, A., & Migotina, D. (2013). Combination of heterogeneous eeg feature extraction methods and stacked sequential learning for sleep stage classification. *2013 International Journal of Neural Systems, Volume 23, Issue 3*, 2–19. https://doi.org/10.1142/S0129065713500123

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: Algorithms and applications. *2000 Neural Networks Volume 13*, 411–430. https://doi.org/https://doi.org/10.1016/S0893-6080(00)00026-5

Jahankhami, P., Kodogiannis, V., & Revett, K. (2006). Eeg signal classification using wavelet feature extraction and neural networks. *2006 IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing (JVA'06)*, 2–6. https://doi.org/https://doi.org/10.1109/JVA.2006.17

Jas, M., Larson, E., Engeman, D., Leppakangas, J., Taulu, S., Hamalainen, M., & Gramfort, A. (2018). A reproducible meg/eeg group study with the mne software: Recommendations, quality assessments, and good practices. *2018 Front. Neurosci. 12:530.*, 1–18. https://doi.org/https://doi.org/10.3389/fnins.2018.00530

Jia, X., Li, K., Li, X., & Zhang, A. (2014). A novel semi-supervised deep learning framework for affective state recognition on eeg signals. *2014 IEEE International Conference on Bioinformatics and Bioengineering*, 30–37. https://doi.org/10.1109/BIBE.2014.26

Jirayucharoensak, S., Pan-Ngum, S., & Israena, P. (2014). Eeg-based emotion recognition using deep learning network with principal component based covariate shift adaptation. *2014 The Scientific World Journal*, 1–10. https://doi.org/10.1155/2014/627892

Kachenoura, A., Karfoul, A., Comon, P., & Wednling, F. (2012). Ica-based eeg denoising: A comparative analysis of fifteen methods. *2012 Bulletin of the Polish Acadamey of Sciences Technical Sciences, Vol. 60, No. 3,* 2–13. https://doi.org/10.2478/v10175-012-0052-3

Kartika Delimayanti, M., Purnama, B., Giang Nguyen, N., Reza Faisal, M., Robiatul Mahmudah, K., Fatma, K., Indrianiand Mamoru, & Satou, K. (2020). Classification of brainwaves for sleep stages by high-dimensional fft features from eeg signals. *2020 Applied Sciences*, 1–12. https://doi.org/10.3390/app10051797

Klug, M., & Gramann, K. (2020). Identifying key factors for improving ica-based decomposition of eeg data in mobile and stationary experiments. *2020 European Journal of Neuroscience*, 2–16. https://doi.org/https://doi.org/10.1111/ejn.14992

Koelstra, S., Muhl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., & Patras, I. (2011). Deap: A database for emotion analysis using physiological signals. *2011 IEEE Transactions on Affective Computing 3*, 18–31. https://doi.org/10.1109/T-AFFC.2011.15

Kumar, K. (2021). Estimation of number of levels of scaling the principal components in denoising eeg signals. *2021 Biomedical Pharmacology Journal Volume 14(1)*, 425–433. https://doi.org/10.13005/bpj/2142

Langkvist, M., Karlsson, L., & Loutfi, A. (2012). Sleep stage classification using unsupervised feature learning. *2012 Advances in Artificial Neural Systems*, 2–10. https://doi.org/10.1155/2012/107046

Lee, W.-H., Ozger, M., Challita, U., & Sung, K. W. (2021). Noise learning based denoising autoencoder. *2021 IEEE Communications Letters PP(99):1-1*, 1–5. https://doi.org/10.1109/LCOMM.2021.3091800

Leon, J., Escobar, J. J., Ortiz, A., Ortega, J., Gonzalez, J., Martin-Smith, P., Gan, J., & Damas, M. (2020). Deep learning for eeg-based motor imagery classification: Accuracy-cost trade-off. *2020 PLoS ONE 15(6)*, 1–30. https://doi.org/https://doi.org/10.1371/journal.pone.0234178

Leske, S., & Dalal, S. (2019). Reducing power line noise in eeg and meg data via spectrum interpolation. *2019 NeuroImage, 189*, 763–776. https://doi.org/https://doi.org/10.1016/j.neuroimage.2019.01.026

Li, G., Ha Lee, C., Jung, J., Chul Youn, Y., & Camacho, D. (2016). Deep learning for eeg data analytics: A survey. *2019 Concurrency and Computation Practice and Experience*, 1–15. https://doi.org/10.1002/cpe.5199

Li, K., Li, X., Zhang, Y., & Zhang, A. (2013). Affective state recognition from eeg with deep belief networks. *2013 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2–7. https://doi.org/10.1109/BIBM.2013.6732507

Liu, J., Wu, G., Luo, Y., Qiu, S., Yang, S., Li, W., & Bi, Y. (2020). Eeg-based emotion classification using a deep neural network and sparse autoencoder. *2020 Front. Syst. Neurosci. 14:43*, 1–14. https://doi.org/https://doi.org/10.3389/fnsys.2020.00043

Liu, P., Zheng, P., & Chen, Z. (2019). Deep learning with stacked denoising autoencoder for short-term electric load forecasting. *2019 Energies*, 1–15. https://doi.org/https://doi.org/10.3390/en12122445

Liu, W., Zheng, W.-L., & Lu, B.-L. (2016). Emotion recognition using multimodal deep learning. *2016 International Conference on Neural Information Processing*, 1–9. https://doi.org/10.1007/978-3-319-46672-958

Lujan, M., Jimeno, M. V., Sotos, J. M., Ricarte, J. J., & Borja, A. (2021). A survey on eeg signal processing techniques and machine learning: Applications to the neurofeedback of autobiographical memory deficits in schizophrenia. *2021 Electronics 10(23):3037*, 1–19. https://doi.org/https://doi.org/10.3390/electronics10233037

Mahdavi-Nasab, H. (2010). Analysis and classification of eeg signals using spectral analysis and recurrent neural networks. *2010 17th Iranian Conference of Biomedical Engineering ICBME*, 2–5. https://doi.org/10.1109/ICBME.2010.5704931

Majoros, T., Ujvari, B., & Oniga, S. (2019). Eeg data processing with neural network. *2019 Carpathian Journal of Electronic and Computer Engineering 12(2)*, 33–36. https://doi.org/10.2478/cjece-2019-0014

Makoto's preprocessing pipeline [Accessed: 2022-01-01]. (n.d.).

Matiko, J., Beeby, S., & Tudor, J. (2013). Real time eye blink noise removal from eeg signals using morphological component analysis. *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 13–16. https://doi.org/10.1109/EMBC.2013.6609425

Min, S., Lee, B., & Yoon, S. (2016). Deep learning in bioinformatics. *2017 Briefings in Bioinformatics, 18(5)*, 851–869. https://doi.org/https://doi.org/10.1093/bib/bbw068

Mohammadi, M., & Mahmud, H. (2019). The state of the art in feature extraction methods for eeg classification. *2019 UHD Journal of Science and Technology 3(2):16*, 2–9. https://doi.org/10.21928/uhdjst.v3n2y2019.pp16-23

Musa, A. B. (2013). A comparison of 1-regularizion, pca, kpca and ica for dimensionality reduction in logistic regression. *2013 International Journal of Machine Learning and Cybernetics*, 1–19. https://doi.org/10.1007/s13042-013-0171-7

Nagar, S., Kumar, A., & Swamy, M. (2021). Orthogonal features-based eeg signal denoising using fractionally compressed autoencoder. *2021 Signal Processing Volume 188*, 1–9. https://doi.org/10.1016/j.sigpro.2021.108225

Phadikar, S., Sinha, N., & Ghosh, R. (2020). Automatic eeg eyeblink artefact identification and removal technique using independent component analysis in combination with support vector machines and denoising autoencoder. *2020 IET Signal Process. Vol. 14 Iss. 6*, 396–405. https://doi.org/https://doi.org/10.1049/iet-spr.2020.0025

Pontifex, M., Gwizdala, K., Parks, A., Billinger, M., & Brunner, C. (2017). Variability of ica decomposition may impact eeg signals when used to remove eyeblink artifacts. *2017 Psychophysiology. 54(3)*, 386–398. https://doi.org/10.1111/psyp.12804

Pontifex, M., Miskovic, V., & Laszlo, S. (2017). Evaluating the efficacy of fully automated approaches for the selection of eyeblink ica components. *2017 Psychophysiology 54(5)*, 4–19. https://doi.org/10.1111/psyp.12827

Qu, H., Shan, Y., Liu, Y., Fan, Z., Zhang, J., & Wanyan, X. (2020). Mental workload classification method based on eeg independent component features. *2020 Applied Sciences 10(9):3036*, 1–15. https://doi.org/10.3390/app10093036

Raduntz, T., Scouten, J., Hochmuth, O., & Meffert, B. (2015). Eeg artifact elimination by extraction of ica-component features using image processing algorithms. *2015 Journal of Neuroscience Methods 243*, 84–93. https://doi.org/https://doi.org/10.1016/j.jneumeth.2015.01.030

Rakhmatulin, I. (2020). Review of eeg feature selection by neural networks. *2020 International Journal of Science and Business. Volume 4, Issue 9.*, 101–112. https://doi.org/10.5281/zenodo.3987894

Rashid, M., Sulaiman, N., Majeed, A., Musa, R., Ab, A., Bari, B., & Khatun, S. (2020). Current status, challenges, and possible solutions of eeg-based brain-computer interface: A comprehensive review. *2020 Front Neurorobot 14 25*, 1–35. https://doi.org/10.3389/fnbot.2020.00025

Riyanul Islam, M., Barua, S., Uddin Ahmed, M., Begum, S., & DiFlumeri, G. (2019). Deep learning for automatic eeg feature extraction: An application in drivers' mental workload classification. *2019 In book: Human Mental Workload: Models and Applications, Third International Symposium, H-WORKLOAD, Proceedings*, 2–16. https://doi.org/10.1007/978-3-030-32423-0_8

Roy, R., Charbonnier, S., & Bonnet, S. (2014). Eye blink characterization from frontal eeg electrodes using source separation and pattern recognition algorithms. *2014 Biomedical Signal Processing and Control 14*, 256–264. https://doi.org/https://doi.org/10.1016/j.bspc.2014.08.007

Safayari, A., & Bolhasani, H. (2021). Depression diagnosis by deep learning using eeg signals: A systematic review. *2021 Medicine in Novel Technology and Devices Volume 12*, 1–16. https://doi.org/ttps://doi.org/10.1016/j.medntd.2021.100102

Said, A. B., Mohamed, A., Elfouly, T., Harras, K., & Wang, J. (2017). Multimodal deep learning approach for joint eeg-emg data compression and classification. *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 1–6. https://doi.org/10.1109/WCNC.2017.7925709

Salazar-Varas, R., & Vazquez, R. (2019). Facing high eeg signals variability during classification using fractal dimension and different cutoff frequencies. *2019 Comput Intell Neurosci. 9174307*, 1–12. https://doi.org/10.1155/2019/9174307

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *2015 Neural Networks Volume 61*, 1–33. https://doi.org/https://doi.org/10.1016/j.neunet.2014.09.003

Shin, H.-C., Orton, M., Collins, D., Doran, S., & Leach, M. (2013). Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4d patient data. *2013 IEEE Transactions on Pattern Analysis and Machine Intelligence 35(8)*, 1930–1943. https://doi.org/10.1109/TPAMI.2012.277

Silverman, D. (2015). The rationale and history of the 10-20 system of the international federation. *2015 American Journal of EEG Technology Volume 3*, 17–22. https://doi.org/https://doi.org/10.1080/00029238.1963.11080602

Sturm, I., Bach, S., Samek, W., & Muller, K.-R. (2016). Interpretable deep neural networks for single-trial eeg classification. *2016 Journal of Neuroscience Methods 274(2)*, 141–145. https://doi.org/https://doi.org/10.1016/j.jneumeth.2016.10.008

Sun, L., Jin, B., Yang, H., Tong, J., Liu, C., & Xiong, H. (2019). Unsupervised eeg feature extraction based on echo state network. *2019 Information Sciences 475*, 1–17. https://doi.org/10.1016/j.ins.2018.09.057

Suryawati, E., Pardede, H., Zilvan, V., Ramdan, A., Krisnandi, D., Heryana, A., Yuwana, S., Suryo Kasumo, B., Arisal, A., & Afif, A. S. (2021). Unsupervised feature learning-based encoder and adversarial networks. *2021 Journal of Big Data 8, 118*, 1–17. https://doi.org/https://doi.org/10.1186/s40537-021-00508-9

Takahashi, H., Emami, A., Shinozaki, T., Kunli, N., Matsuo, T., & Kawai, K. (2020). Convolutional neural network with autoencoder-assisted multiclass labelling for seizure detection based on scalp electroencephalography. *2020 Computers in Biology and Medicine 125 104016*, 1–9. https://doi.org/https://doi.org/10.1016/j.compbiomed.2020.104016

Tautan, A.-M., Dogariu, M., & Lonescu, B. (2019). Detection of epileptic seizures using unsupervised learning techniques for feature extraction. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2377–2381. https://doi.org/10.1109/EMBC.2019.8856315

Vareka, L., & Mautner, P. (2017). Stacked autoencoders for the p300 component detection. *2017 Front. Neurosci. 11:302*, 1–9. https://doi.org/https://doi.org/10.3389/fnins.2017.00302

Wen, T., & Zhang, Z. (2018). Deep convolution neural network and autoencoders-based unsupervised feature learning of eeg signals. *2018 IEEE Access PP(99):1-1*, 1–12. https://doi.org/10.1109/ACCESS.2018.2833746

Windrim, L., Ramakrishnan, R., Melkumyan, A., Murphy, R., & Chilingaryan, A. (2019). Unsupervised feature-learning for hyperspectral data with autoencoders. *2019 Remote Sensing*, 1–19. https://doi.org/https://doi.org/10.3390/rs11070864

Xing, C., Ma, L., & Yang, X. (2015). Stacked denoise autoencoder based feature extraction and classification for hyperspectral images. *2016 Journal of Sensors*, 1–10. https://doi.org/https://doi.org/10.1155/2016/3632943

Xing, X., Li, Z., Xu, T., Shu, L., Hu, B., & Xu, X. (2019). Sae+lstm: A new framework for emotion recognition from multi-channel eeg. *2019 Frontiers in Neurorobotics Volume 13*, 1–14. https://doi.org/https://doi.org/10.3389/fnbot.2019.00037

Yu, Z., Li, L., Zhang, W., Lv, H., Liu, Y., & Khalique, U. (2021). An adaptive eeg feature extraction method based on stacked denoising autoencoder for mental fatigue connectivity. *2021 Neural Plasticity*, 1–12. https://doi.org/10.1155/2021/3965385

Zakeri, Z., Assecondi, S., Bagshaw, A., & Arvanitis, T. (2014). Influence of signal preprocessing on ica-based eeg decomposition. *2014 IFMBE proceedings 41*, 734–737. https://doi.org/10.1007/978-3-319-00846-2_182

Zavala, S., Chicaiza, K. O., Lopez, J. L., Ramirez, J. C., & Yoo, S. G. (2020). Eeg signal processing model for eye blink detection. *Journal of Engineering and Applied Sciences 15 (7): 1671-1675, 2020*, 1671–1675. https://doi.org/10.36478/jeasci.2020.1671.1675

Zhang, C., Tong, L., Zeng, Y., Jiang, J., Bu, H., Yan, B., & Li, J. (2015). Automatic artifact removal from electroencephalogram data based on a priori artifact information. *2015 BioMed Research International Volume 2015, Article ID 720450*, 1–9. https://doi.org/https://doi.org/10.1155/2015/720450