

Automated Eye Movement Classification System based on EMG of EOM using Synchrosqueezed Wavelet Transform

B .Tech. Project

Submitted in partial fulfillment of the requirements

of the degree of

Bachelor of Technology

in

Electrical Engineering

By

Rajaputhra Shivani

(Roll No: 200002062)

Under the guidance of

Prof. Ram Bilas Pachori



Department of Electrical Engineering

Indian Institute of Technology Indore

Dec. 2023

Declaration

I hereby declare that the project entitled “Automated Eye Movement Classification System based on EMG of EOM using Synchrosqueezed Wavelet Transform” submitted in partial fulfillment for the award of the degree of Bachelor of Technology in ‘Electrical Engineering’ completed under the supervision of Prof. Ram Bilas Pachori, Department of Electrical Engineering, IIT Indore is an authentic work.

Further, I declare that I have not submitted this work for the award of any other degree elsewhere.

I affirm that this written submission reflects my thoughts expressed in my own words. In instances where I have incorporated the ideas or words of others, I have appropriately cited and referenced the original sources. I further confirm that I have followed all principles of academic honesty and integrity, and there is no misrepresentation, fabrication, or falsification of any idea, data, fact, or source in my submission. I am aware that any infringement of the aforementioned principles may lead to disciplinary measures by the institute and could result in penalties from sources that have not been duly cited or from which proper permission has not been obtained when necessary.

Shivani.R
Rajaputhra Shivani

Date: 06-12-2023

Place: IIT Indore

Preface

This report on “**Automated Eye Movement Classification System based on EMG of EOM using Synchrosqueezed Wavelet Transform**” is prepared under the guidance of Prof. Ram Bilas Pachori.

The current work intends to classify six types of eye movements based on electromyogram (EMG) data from extraocular muscles (EOM). Here, we classify six different types of eye movements based on time-frequency representations of the EMG of EOM signals obtained from synchrosqueezed wavelet transform using convolutional neural network.

Rajaputhra Shivani

B. Tech. IV year

Department of Electrical Engineering

B. Tech. Project Approval Certificate

This is to certify that the dissertation titled "**Automated Eye Movement Classification System based on EMG of EOM using Synchrosqueezed Wavelet Transform**" submitted by , (Roll No. 200002062) is approved for the award of degree of **Bachelor of Technology** in **Electrical Engineering**.



BTP Supervisor(s):
Dr. Ram Bilas Pachori

Date:06.12.2023.....

Designation:Professor, EE, IIT Indore.....

Acknowledgements

I wish to express my sincere gratitude to **Prof. Ram Bilas Pachori**, my project guide, for his invaluable support and encouragement throughout the study. It has been a privilege to undertake this project under his guidance.

I would also like to extend my thanks to **Mr. Vivek Kumar Singh**, Doctoral Research Scholar in the Department of Electrical Engineering, IIT Indore, for his continuous guidance, which greatly contributed to the successful completion of the project.

My appreciation is extended to my parents and close friends for their unwavering support and encouragement, which has been a constant source of motivation. I am grateful to the institute for providing a unique opportunity to engage in healthcare research and for facilitating the necessary resources for the project's fulfillment.

Lastly, I offer my sincere thanks to all individuals who played a role in the completion of this project, even if their names may have inadvertently been missed. Your assistance has been truly appreciated.

Rajaputhra Shivani
B. Tech. IV year
Department of Electrical Engineering

Abstract

This work describes the use of a wavelet method to identify electrooculography (EOG) signals linked to eye movement potentials. Electrodes positioned strategically on the forehead surrounding the eyes are used to record the EOG signals, which makes it easier to record eye movements. To identify the properties of the eye movement waveform, the wavelet features are used. The six categories of activities included in the dataset are blinking, center fixation (no movement), up, down, right, and left. The synchrosqueezed wavelet transform (SWT) is the tool that will be used in the suggested analysis for every ocular signal.

By using a frequency reassignment process, the SWT improves the readability of the time-frequency representation (TFR) by acting as an empirical mode decomposition-like tool. This enables an inverse technique to be used to rebuild nonstationary signals.

An strategy known as transfer learning has been used to deal with this issue. The photos are divided into six classes using a variety of convolutional neural network models that have been trained beforehand. A two-dimensional wavelet transform is used to do this. Remarkably, Resnet50 was able to attain 96% accuracy.

Contents

List of Figures	ix
List of Tables	1
Table of Contents	1
List of Figures	1
1 Introduction	2
1.1 Eye Movements	2
1.1.1 EOG	2
1.2 Motivation for BTP	4
1.3 Dataset used	5
2 Literature Review	6
3 Proposed Method	9
3.1 Wavelet	10
3.1.1 Translation	11
3.1.2 Dilation	11
3.2 Wavelet Transform	12
3.3 Wavelet synchrosqueezed Transform	13
4 Time-Frequency Analysis	14
4.1 Time-Frequency Analysis	15
5 Transfer Learning	22
5.1 What is Transfer Learning?	22
5.2 Different ways of Transfer Learning	24

6 Classification	26
6.1 Introduction	26
6.2 Data split	28
6.3 Training and Testing	29
6.4 Alexnet Pre-trained Network	29
6.5 GoogleNet Pre-trained Network	31
6.6 Resnet 50 Pre-trained Network	32
6.7 Xception Pre-trained Network	34
7 Results and Discussion	37
8 Conclusions and Future Work	41
8.1 Conclusions	41
8.2 Future scope	41
References	42

List of Figures

3.1	Flowchart of the proposed methodology.	9
4.1	EOG signal in normal eye position for (a) left eye. (b) right eye.	16
4.2	TFR of normal eye position for (a) left eye. (b) right eye.	16
4.3	EOG signal in downward eye position for (a) left eye. (b) right eye.	17
4.4	TFR of downward eye position for (a) left eye. (b) right eye.	17
4.5	EOG signal in left eye position for (a) left eye. (b) right eye.	18
4.6	TFR of left eye position for (a) left eye. (b) right eye.	18
4.7	EOG signal in blink eye position for (a) left eye. (b) right eye.	19
4.8	TFR of blink eye position for (a) left eye. (b) right eye.	19
4.9	EOG signal in right eye position for (a) left eye. (b) right eye.	20
4.10	TFR of right eye position for (a) left eye. (b) right eye.	20
4.11	EOG signal in upward eye position for (a) left eye. (b) right eye.	21
4.12	TFR of upward eye position for (a) left eye. (b) right eye.	21
6.1	Training data distribution for all subjects.	29
6.2	Validation data distribution for all subjects.	29
6.3	AlexNet architecture.	30
6.4	AlexNet training progress plot.	30
6.5	AlexNet training loss plot.	31
6.6	GoogleNet architecture.	31
6.7	GoogleNet training progress plot.	32
6.8	GoogleNet training loss plot.	32
6.9	(a) Residual block, (b) ResNet50 architecture.	33
6.10	ResNet 50 training progress plot.	33
6.11	ResNet 50 training loss plot.	34
6.12	Xception architecture.	35
6.13	Xception training progress plot.	35

6.14 Xception training loss plot.	36
7.1 Confusion matrix for AlexNet model.	38
7.2 Confusion matrix for GoogleNet model.	39
7.3 Confusion matrix for ResNet 50 model.	39
7.4 Confusion matrix for Xception model.	40

List of Tables

2.1 Comparision of the proposed method with some existing approaches	7
6.1 Pre-trained neural network model specifications.	28
6.2 Training hyperparameters chosen for the proposed work.	28
7.1 Multiclass test metrics of trained models.	38

Chapter 1

Introduction

1.1 Eye Movements

The human eye, characterized by its spherical structure with a radius of 12mm, plays a pivotal role in detecting signals associated with eye movements, a phenomenon known as electrooculography (EOG). EOG is derived from the cornea retinal potential (CRP), originating within the eyeball as a result of metabolic activities in the retinal epithelium. The generation of CRP is facilitated by the hyperpolarization and depolarization of nervous cells in the retina. This intricate process forms the foundation for EOG, an electrical recording that faithfully corresponds to the various movements of the eyes.

The eye sustains a baseline electrical potential, characterized by a positive charge at the front of the globe and a negative charge at the back. This intriguing electrical polarization was first identified by Emil du Bois-Reymond in 1848 and has since served as the cornerstone for the development of EOG [1].

1.1.1 EOG

EOG recordings are typically obtained using bipolar electrodes positioned externally on the eye. While specific electrode placements may vary, they are commonly situated on the temples or at the far ends of the forehead. When the eyes undergo movement, a distinctive potential difference arises. The magnitudes of the right and left eye movements typically fall within the range of $-75\mu\text{V}$ to $150\mu\text{V}$, respectively.

The polarity of the movement potentials is contingent upon the electrode setup. Specifically, the signal exhibits a positive value when the eyes are moving toward the positive electrode. This relationship underscores the importance of electrode configuration in interpreting EOG signals associated with eye movements [1].

At present, EOG serves as a valuable tool for assessing oculomotor irregularities, including conditions like nystagmus, strabismus, and supranuclear oculomotor dysfunction. To briefly elaborate on supranuclear oculomotor dysfunction, this refers to impairments in eye movement control that originate above the level of the cranial nerve nuclei [2]. EOG's application in evaluating these ocular abnormalities underscores its significance in diagnosing and understanding various disorders affecting eye movements.

EOG serves as a valuable method for measuring the baseline potential of the retina, resulting in the creation of an electro-oculogram signal. This approach is particularly useful in ophthalmological diagnostics and the monitoring of eye movements. The application of EOG is pivotal in the evaluation and diagnosis of diverse eye-related conditions, underscoring its essential contribution to the field of ophthalmology. [3]. The electro-oculogram is a potential generated by the movement of the eyes or eyelids. To comprehend the origin of the EOG signal, one can visualize dipoles situated in the eyes, where the cornea possesses a relatively positive potential compared to the retina [4]. This electrical activity is then captured by a dual-channel signal acquisition system comprising the horizontal (H) and vertical (V) channels.

The horizontal EOG (EOG(H)) involves electrode placement designed to measure horizontal eye movements, while the vertical EOG (EOG(V)) focuses on electrodes strategically positioned to assess vertical eye movements. This arrangement allows for the precise capture and interpretation of the electrical signals associated with the movement of the eyes.

1.2 Motivation for BTP

In recent years, there has been a notable increase in the adoption of eye-tracking technology across a broad spectrum of research domains and practical uses. These applications encompass diverse fields, including human-computer interaction (HCI), augmented reality (AR) and behavioral physiology. The integration of eye-tracking technology has become a pivotal element in advancing immersive technologies, eliciting considerable interest and involvement from the research community.

The motivation behind research on the classification of eye movements stems from its multifaceted applications and its potential to address significant challenges. Beyond its research implications, it holds clinical importance by contributing to the diagnosis and treatment of eye-related disorders. Moreover, its role in enabling individuals with disabilities to interact with technology through eye movements underscores its broader societal impact.

Furthermore, eye-tracking technology contributes significantly to biometrics and security by providing a secure method for user authentication. This research not only propels oculomotor studies, thereby deepening our comprehension of human vision and cognition but also holds implications in neuroscience. It plays a role in illuminating cognitive processes and neurological aspects related to disorders in eye movements. The continuous progress in technology, particularly in signal processing techniques, has broadened the scope for meticulous analysis, fostering interdisciplinary collaboration in these domains.

1.3 Dataset used

In this investigation, we utilized the EMG of EOM dataset, accessible on IEEE data-port [5]. This dataset encompasses six different types of eye movements: up, right, left, down, fixation (steady eyes) and blinking. The myoelectric activity associated with eye movement was captured using two vertical and horizontal electrodes strategically placed on the forehead. To enhance the data quality, the dataset underwent preprocessing with an AD620 instrumentation amplifier.

Each signal in the dataset has a duration of 2 seconds, and the sampling frequency is set at 120 Hz. During the process of recording of these EMG signals, the presence of direct current (DC) noise around 0 Hz was observed. To mitigate the impact of dc noise and adhere to the Nyquist rate, a filtering process was applied within the frequency band of 0.2–40 Hz [6]. The dataset was obtained from ten subjects, each performing 10 pseudorandom repetitions of the aforementioned eye movements.

Chapter 2

Literature Review

In this chapter, we embark on a thorough exploration of existing literature, providing a comprehensive summary of prior research relevant to our project's domain. The objective of our project is elucidated with reference to past works, establishing a contextual framework.

A variety of signal processing and machine learning approaches have been effectively employed for the classification of EOG signals, enabling the differentiation of various types of eye movements [7, 8]. To address the challenges of three-class classification, diverse features have been proposed. These encompass changes in eye behavior [9], integration of eye positional data with hidden Markov models [10], and extraction of temporal features from EOG signals. Temporal features include pulsewidth modulation (PWM) and disparity in the slope of the EOG signal [11], along with the utilization of an extended moving average filter [12]. Additionally, spectral features have played a significant role in categorizing EOG signals into distinct types of eye movements [7, 13].

Notably, Belkacem et al. [14] utilized both electroencephalogram (EEG) and EOG signals for this classification task employing hierarchical classification with thresholding. Lately, the integration of various deep learning approaches has garnered attention for classifying eye movements using EOG signals, resulting in significant enhancements in classification accuracy [15–17].

Table 2.1: Comparision of the proposed method with some existing approaches

S. No	Method	Subjects	Classification problem	CACC(%)
1	STFT-based spectral features with FFNN-based classification [18]	20	11-class	91.40
2	CNN-based feature extraction and classification [19]	2	11-class	90.82
3	Hierarchical classification with thresholding [20]	20	4-class	85.00
4	EMDF with PWM and thresholding-based classification [21]	10	4-class	90.00
5	VOS-based features and thresholding-based classification [22]	NA	8-class	88.75
6	WT-based time-series features and thresholding-based classification [23]	10	6-class	85.20

Lately, extensive research has focused on automated eye movement classification, primarily leveraging EEG and EOG signals. Table 2.1 offers a comparison of our current method with existing approaches concerning the dataset used, the number of subjects involved, the methodology employed, the classification problem addressed, and the corresponding classification accuracy.

In a study conducted by Hema et al. [18], EOG signals from a group of twenty subjects were recorded for the purpose of classifying eleven distinct eye movements. The study employed bandwise energy features extracted from the short-time Fourier transform (STFT), combined with a feedforward neural network (FFNN). The reported maximum classification accuracy achieved in this investigation was 91.40

Thilagaraj et al. [19] conducted a separate study that addressed the challenge of classifying eleven different eye movements, achieving a reported maximum accuracy of 90.82

Hierarchical classification with thresholding was employed by Belkacem et al. [23] to classify 4-class eye movements (left, up, right and down) from EEG and EOG signals, involving 20 subjects and achieving an overall classification accuracy of 85%. Recently Hsieh and Huang [21] proposed a framework for four-class eye movements based on features ex-

tracted using an extended moving difference filter (EMDF) with pulsedwidth modulation (PWM), reporting a maximum classification accuracy of 90% using EEG signals from ten subjects.

Addressing an eight-class eye movement classification problem in [22], Lin et al. utilized EOG signals, achieving an accuracy of 88.75% with a feature extraction method called variation of slope (VOS) and subsequent thresholding. Additionally, Belkacem et al. [23] employed wavelet transform (WT) and time-series-based features to separate EEG signals associated with six sets of eye movements, reporting a maximum entire classification accuracy of 85.2% through thresholding-based classification.

Chapter 3

Proposed Method

The outlined procedure commences with the generation of the dataset, followed by the application of the synchrosqueezed wavelet transform (WSST) for the conversion of 1-D signals into 2-D time–frequency plots. Subsequently dataset preparation, and the classification steps are undertaken. The flowchart illustrating the proposed methodology is depicted in Fig. 3.1.

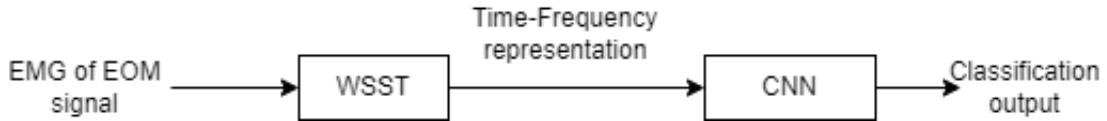


Figure 3.1: Flowchart of the proposed methodology.

The continual and great need for a fascinating new mathematical tool to present both the theory and application for science and engineering seems to have contributed to the creation of wavelet and wavelet transform, which seems more clearly and effectively established. Considering the value and appeal of wavelet analysis in diverse fields, the notion of the wavelet may be expressed through a variety of separate thought processes

3.1 Wavelet

The term wavelet refers to a small wave or a wavelet component. The smallness condition is the requirement that this (window) function be of limited length (compactly support). The situation when this function oscillates is referred to as a “wave.” To represent functions more effectively than Fourier series, wavelet technology was created. The name “wavelet” originated from historical research on time-frequency signal analysis, wave propagation, and sampling theory. Wavelets are made by growing and translating the “mother wavelet,” a single function. Data compression, noise reduction, pattern identification, and rapid calculations are just a few of the many applications that wavelets may be used for. Wavelet theory provides a coherent framework and compelling explanation for many different notions and techniques that have been separately developed in numerous fields. In order to analyse non-stationary data, Jean Morlet and a group of French engineers developed the wavelet idea for the first time in 1982. Wavelet is a set of functions that consists of the translations and dilations of the mother wavelet function [24].

This innovative concept, however, might be considered as a synthesis of a number of concepts from several disciplines, including engineering, physics, and mathematics. In order to analyse a signal, wavelet transforms it into a limited energy function of short duration. The detection of aeroplanes and submarines, data compression, image processing, wave propagation, pattern recognition, computer graphics, and advancements in CAT scans and other medical imaging methods are just a few examples of contemporary uses for wavelet analysis.

An innovative new approach called wavelet analysis has been developed to address difficult issues in engineering, physics, and mathematics. Wavelet allows for the accurate deconstruction of complex information, including speech, music, pictures, and patterns, into basic forms known as the basic foundations at different places and scales. A given continuous time signal is divided into different scale components using a mathematical operation known as a wavelet. Each and every scale component may often be assigned a frequency range. Then, each scale component may be looked at a resolution suitable for that scale [24].

Definition (Wavelet): Wavelets ($\psi_{a,b}(t)$) are functions made up of dilations and translations of a single function, referred to as the mother wavelet ($\psi(t)$).

$$\psi_{a,b}(t) = a^{-1/2}\psi\left(\frac{t-b}{a}\right), a, b \in \mathbb{R}, a \neq 0 \quad (3.1)$$

where b is a parameter of translation that specifies the wavelet's time position and a is a scaling parameter that measures the wavelet's compression degree. The factor $a^{-1/2}$ merely normalised the energy by ensuring that the daughter wavelets, regardless of the scale employed, had the same amount of energy as the mother wavelet.

$\psi_{a,b}(t)$ stretches when $a > 1$ and contracts when $a < 1$

The term “mother” implies that a single major function, or the mother wavelet, is the ancestor of the functions with diverse areas of support used in the transformation process or to put it another way, the mother wavelet is only a template for making the other window functions or wavelets known as daughter wavelets. A daughter wavelet is produced by expanding and interpreting the mother wavelet. Wavelets like the Morlet and mexican hat are two examples. Translation and dilation are the two sorts of alteration that may be applied to wavelets.

3.1.1 Translation

The wavelet is translated when it is moved from one spot to another. In this case, we move the wavelet's centre along the time axis. The wavelet's position with relation to the original signal is represented by translation. The wavelet's distance travelled along the signal is measured by translation. Its symbol is “ b .” [24]

3.1.2 Dilation

Scaling parameters, such as dilation, are used to estimate how much something is compressed or scaled. The letter “ a ” indicates it. The scale factor, or constant “ a ,” adjusts the breadth of the mother wavelet. The inverse of frequency is the scale factor. Mother wavelet is stretched when scale increases and frequency decreases. Mother wavelet is compressed when scale is reduced and frequency is increased. The scale value $a = 1$ is typically used as the starting point for the continuous wavelet transform calculation, which is followed by an increment in the integer values. Depending on the signal being modified, “ a ”s range can vary.

3.2 Wavelet Transform

A wavelet in a wavelet transform describes a function. It helps in providing a representation of signal energy in both time and frequency jointly [25]. Also provides this information are other transforms, including the STFT, the Wigner distribution, etc. The wavelet transform, which may simultaneously give time and frequency information, creates a time-frequency representation of the given signal. It was decided to use the wavelet transform in favour of the STFT. The creation of the wavelet transform addressed the resolution-related problems of the STFT [24].

On the time-frequency plane, it is difficult to determine the time and frequency details of a signal at a specific location. Which spectral component is present at any one moment cannot be determined. The most we can do is to investigate the spectral elements that are active at any particular time. This is a resolution problem, and as a consequence, research has shifted from STFT to wavelet transform. While STFT always offers a fixed resolution, wavelet transform offers variable resolution. We fundamentally need wavelet transform to assess signals whose frequency response fluctuates over time. The Fourier transform should not be used for non-stationary signals. The Fourier transform is inappropriate if the signal is non-stationary or has a frequency that changes over time. If the signal always had a frequency component, the results of the Fourier transform would make sense. A certain frequency component's existence or absence is indicated via the Fourier transform. This information is applicable whenever a component is present. In order to properly apply the Fourier transform to a signal, it is essential to first determine if it is stationary.

The STFT is the most recent version of the Fourier transform. There is not much difference between Fourier transform and STFT. In the STFT, the signal is divided into small enough segments (portions) that they may be taken as stationary. Simply multiplying the signal's Fourier transform by a window function yields the signal's short-term Fourier transform. Wavelet is confined in both time and frequency as opposed to short Fourier transform, which is just focused in frequency [26].

3.3 Wavelet synchrosqueezed Transform

SWT: The continuous wavelet transform (CWT) of the PQD $s(t)$ is given as [27].

$$T(a, b) = \int s(t) b^{-1/2} \psi^* \left(\frac{t-a}{b} \right) dt \quad (3.2)$$

where $s(t)$ is the input signal, $\psi^*(t)$ is the complex conjugate of mother wavelet $\psi(t)$, and $b > 0$ is the scale. In the proposed methodology, The mother wavelet is selected to be the Morlet wavelet, and the CWT is calculated. The Morlet wavelet is a complex-valued wavelet that offers magnitude and phase data. It is more suited for time-frequency analysis of time series data and nonstationary signals because of this feature.

In the synchrosqueezed approach, the IF [28] of the multi component signal is estimated as

$$\widehat{\omega}(a, b) = \frac{-j}{T(a, b)} \frac{\partial T(a, b)}{\partial t} \quad (3.3)$$

We can synthesize the signal because the unequal distribution of energy in CWT is concentrated close to the IF regions of each component of the nonstationary signal $s(t)$. This means that time information is preserved and energy is only reallocated along the frequency axis. [29]. This results in a sharpened TFR and improves the time-frequency resolution, where $T(a, b)$ is the CWT of the input signal and $\widehat{\omega}$ is the IF.

The SWT is computed as [30]

$$T_{sq}(a, b_k) = \Delta\omega^{-1} \sum_{b_k: |\omega(a, b_k) - \omega_m| \leq \frac{\Delta\omega}{2}} T(a, b_k) b_k^{-3/2} (\Delta b)_k \quad (3.4)$$

where $\Delta b_k = b_k - b_{k-1}$ is the scaling step and $\Delta\omega = \omega_m - \omega_{m-1}$ is the frequency bin. The frequency variable ω and scale variable s are discrete values and the CWT, $T(a, b)$, is computed only at discrete values b_k . Likewise, the SWT is computed only at the central frequency ω_b of frequency range $[\omega_m - (\Delta\omega/2), \omega_m + (\Delta\omega/2)]$.

The reconstruction of signal can be achieved by [30]

$$C_\Psi = \int_0^\infty \frac{|\Psi(\omega)|^2}{\omega} d\omega < \infty \quad (3.5)$$

$$S_k(\tau) = \Re \left\{ \frac{1}{C_\Psi} \sum_m T_{sq}(\tau, \omega_m) (\Delta\omega) \right\} \quad (3.6)$$

where C_Ψ is the admissibility condition, $\Psi(\omega)$ is the FT of $\psi(t)$, and $\Re\{\cdot\}$ represents the real part of the function given to it as input.

Chapter 4

Time-Frequency Analysis

Using signal processing we convert and transform data in a form that allows visualization in it that are not possible through direct observation. It will help us to evaluate, optimize, and differentiate between the signals.

Signal is processed on both linear and non-linear systems. In our case the signals are non-linear which means that their values change over time, in uncertain ways. The change in output is not linearly proportional to a change in the input.

A time-domain signal is time-changing, whereas a frequency-domain graph shows the signal amount lying within each given frequency band [31].

1. **Time-domain Signal Processing:** Amplitude points are plotted against time. Signals are examined directly in the time dimension, allowing for the study of amplitude changes, duration, and temporal relationships between different parts of the signal.
2. **Frequency-domain Signal Processing:** Frequencies are plotted against their magnitudes. It employs tools like the Fourier transform to decompose signals into sinusoidal components, revealing their frequency, amplitude, and phase characteristics.

In this study we used a scalogram to get the time-frequency transformation and visualization of the pressure signals in the system and plotted as a function of frequency and time to get the dynamic amplitude and phase shifts of brain oscillations at various frequencies at a particular instance.

Transformation of a signal from the time domain to the frequency domain allows visualization of impossible things through direct observation. It is possible to break down a signal into its component frequencies using the Fourier transform mathematical concept.

4.1 Time-Frequency Analysis

In this study, two-dimensional time-frequency planes are used to characterize phenomena using time-frequency representations. This means that we can extract the local characteristics of the phenomena under study, including its time and frequency information. This results in a more potent analytical tool for managing non-stationary and time-varying characteristics that are present in most complicated physical events. There are essentially two different kinds of time-frequency representations. One covers the time-frequency distributions using a quadratic technique. The second method is a linear one that uses the STFT and the recently developed wavelet transform which we have used here [31, 32].

A mathematical technique called the wavelet transform is frequently employed in signal processing applications. It can unravel unique patterns that are concealed in a sea of data. We require modelling effort to address the prediction problem using time series and neural networks. The capacity of neural networks as a general estimator in the estimation of highly nonlinear systems is constrained. The time-frequency domain wavelet transform has the capacity to simultaneously represent functions and manifest their local properties. The use of these traits enables accurate neural network training to model very nonlinear signals.

These time-frequency analyses are further utilized in the classification of eye movements.

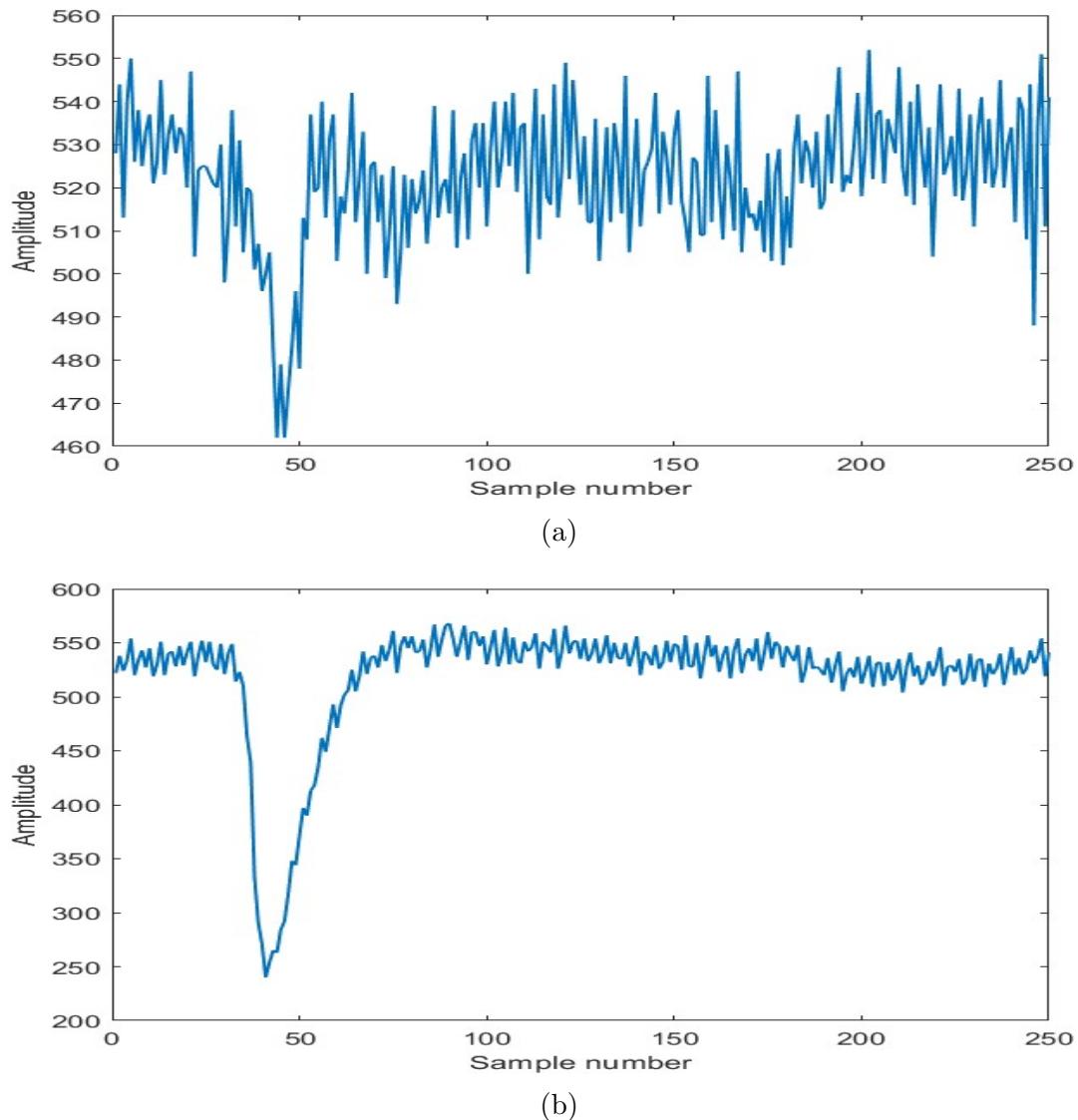


Figure 4.1: EOG signal in normal eye position for (a) left eye. (b) right eye.

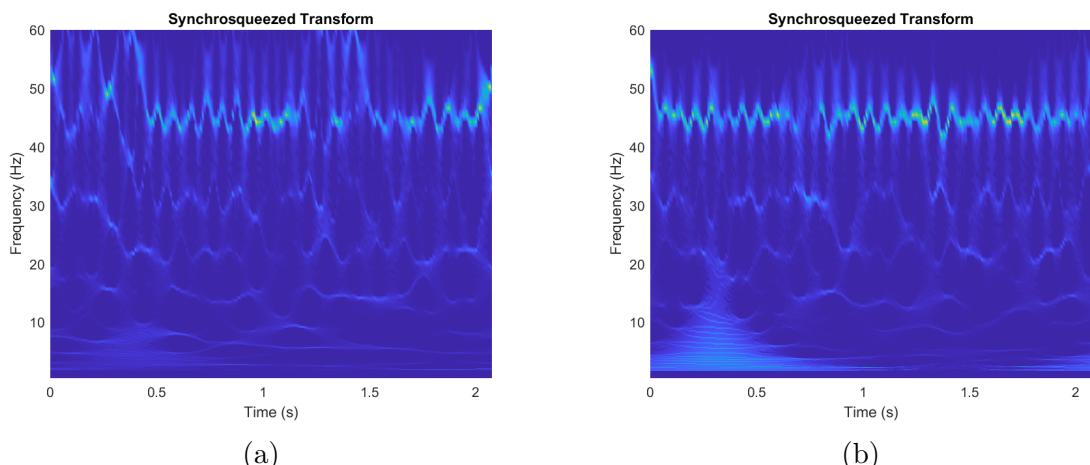


Figure 4.2: TFR of normal eye position for (a) left eye. (b) right eye.

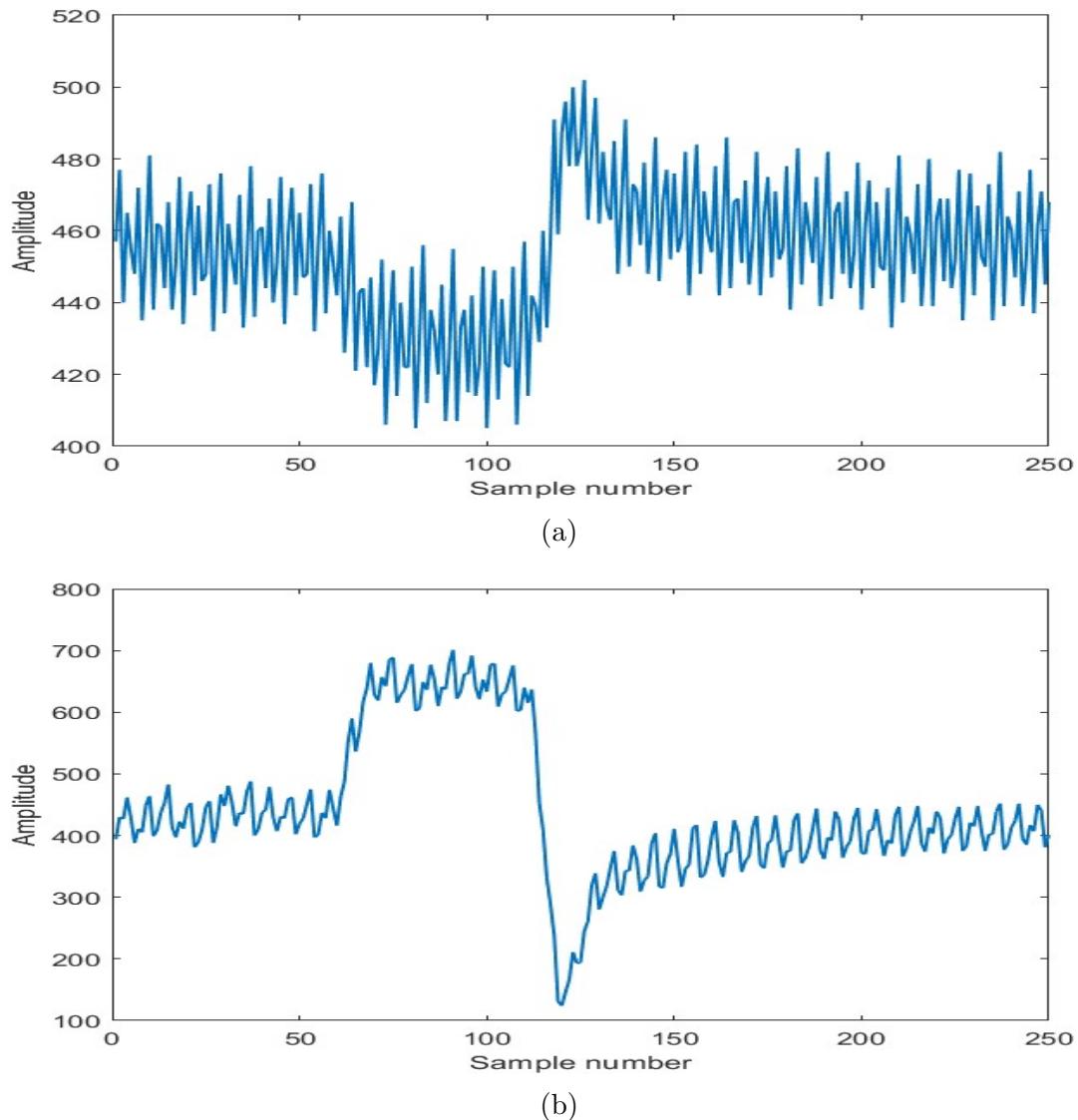


Figure 4.3: EOG signal in downward eye position for (a) left eye. (b) right eye.

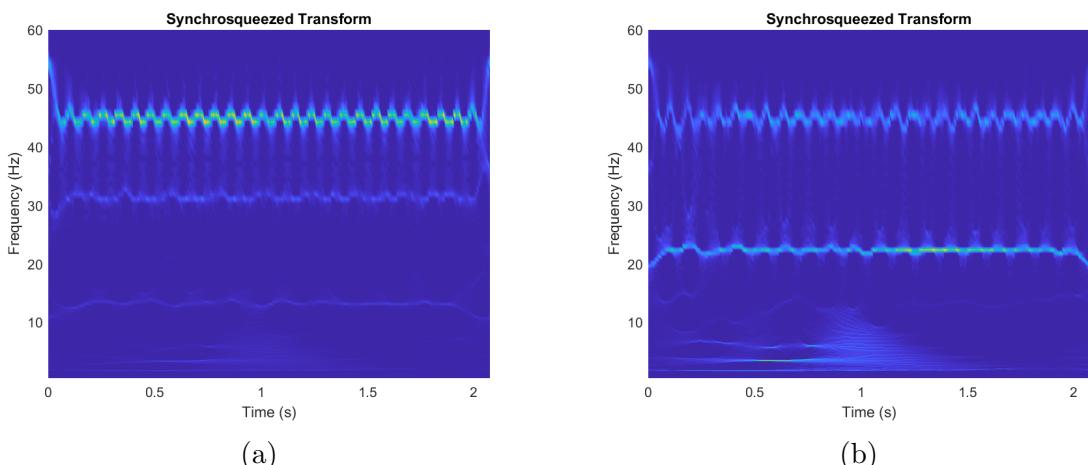


Figure 4.4: TFR of downward eye position for (a) left eye. (b) right eye.

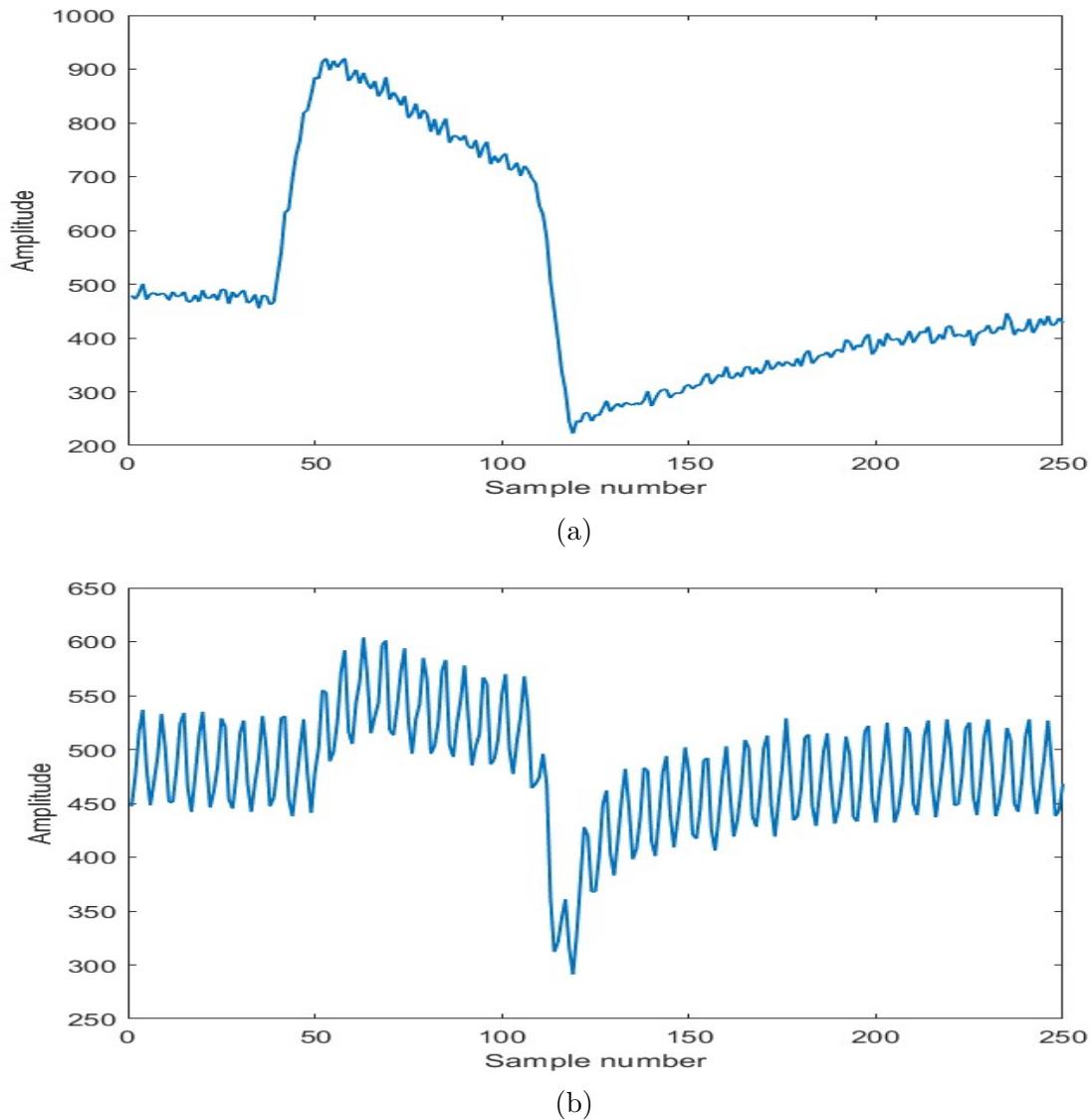


Figure 4.5: EOG signal in left eye position for (a) left eye. (b) right eye.

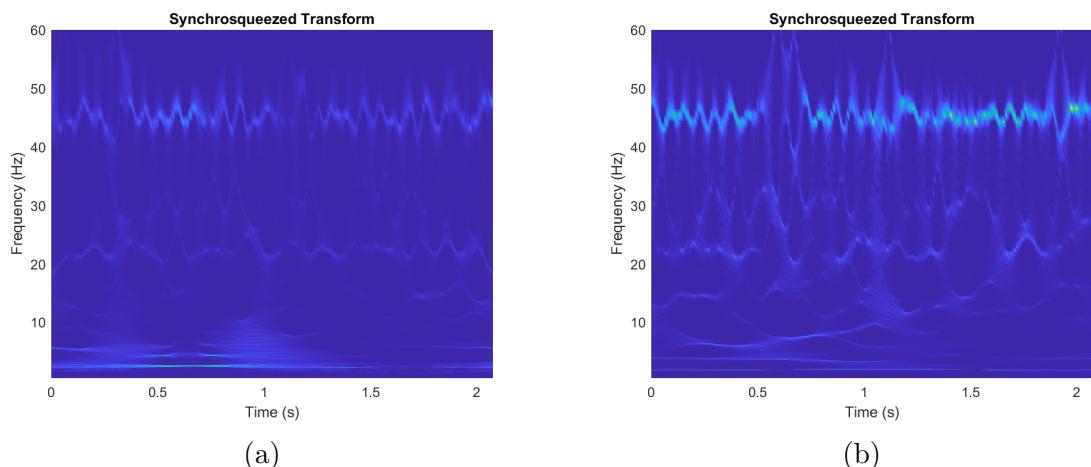


Figure 4.6: TFR of left eye position for (a) left eye. (b) right eye.

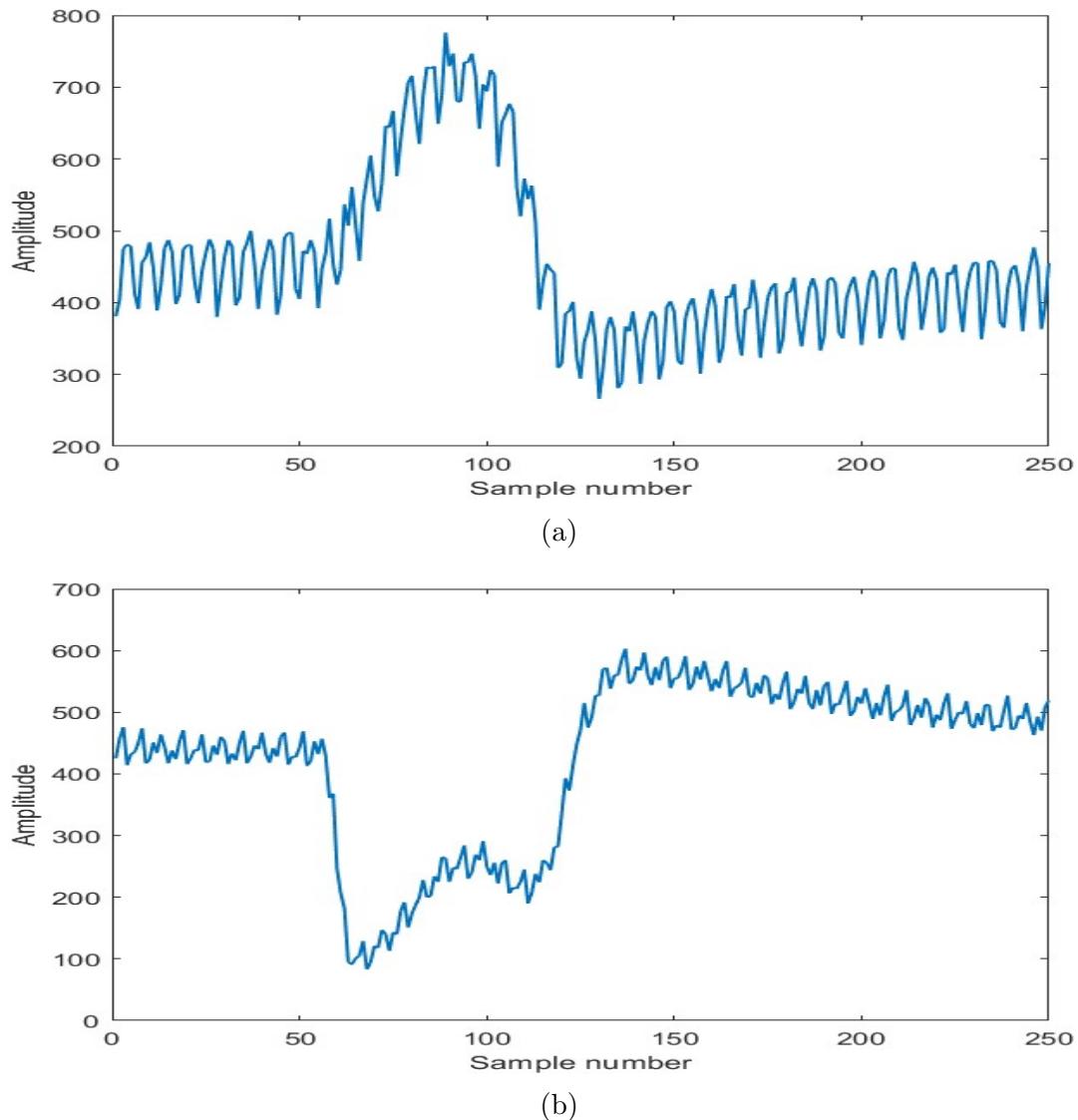


Figure 4.7: EOG signal in blink eye position for (a) left eye. (b) right eye.

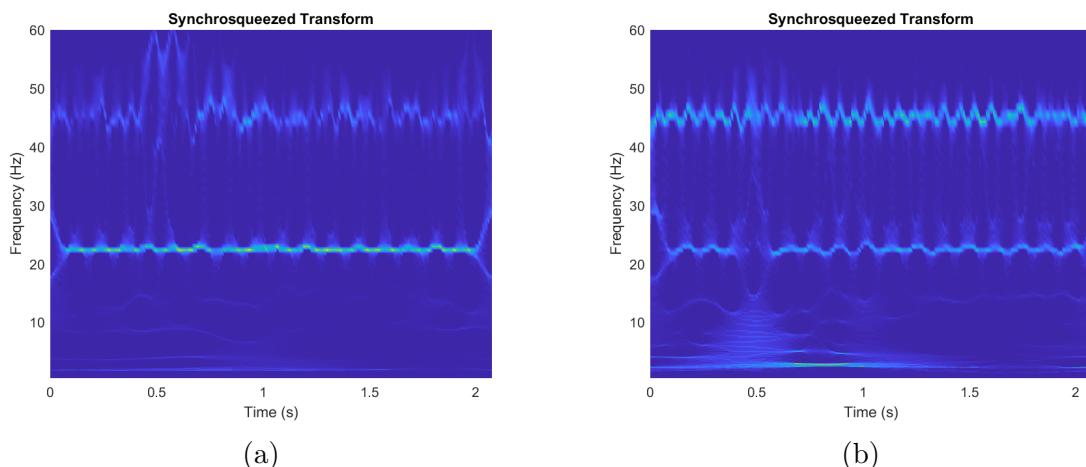


Figure 4.8: TFR of blink eye position for (a) left eye. (b) right eye.

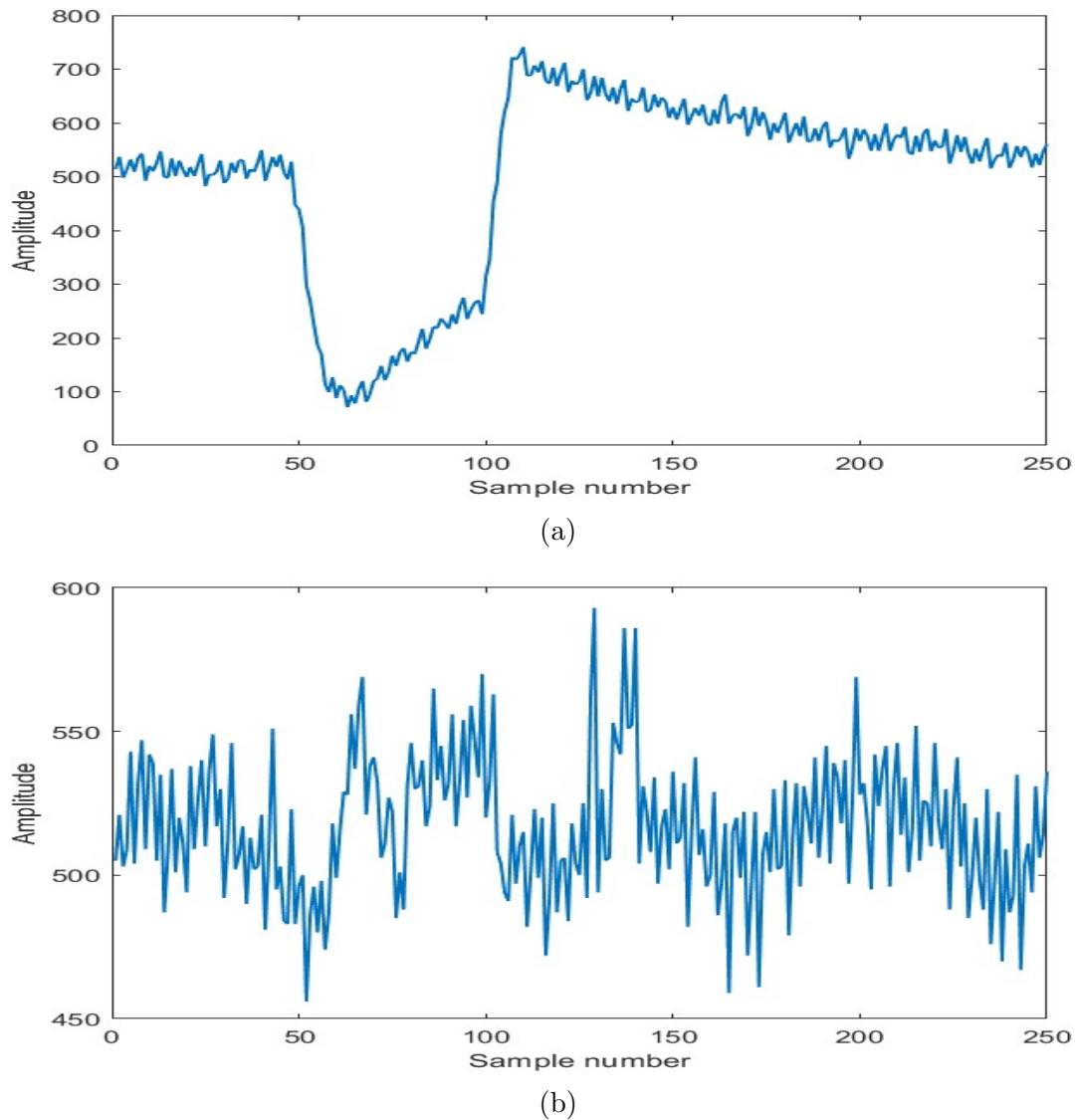


Figure 4.9: EOG signal in right eye position for (a) left eye. (b) right eye.

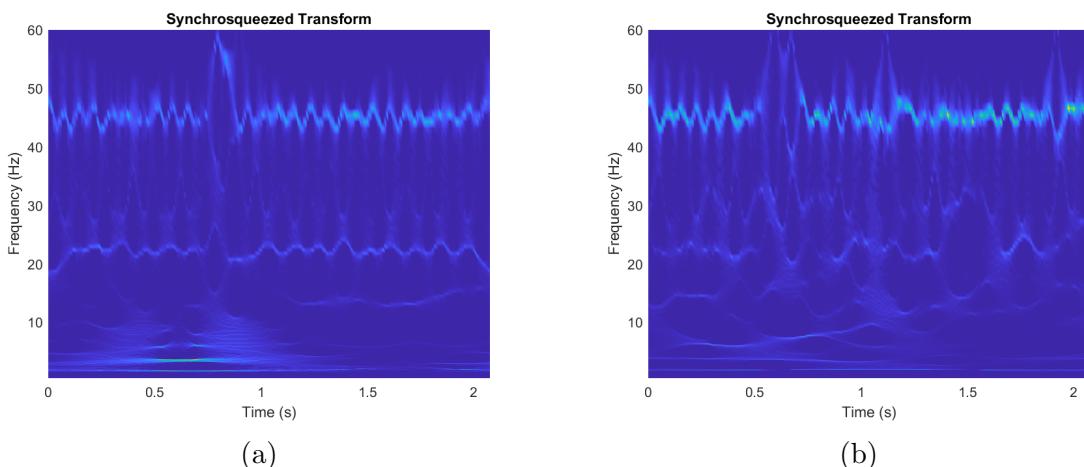


Figure 4.10: TFR of right eye position for (a) left eye. (b) right eye.

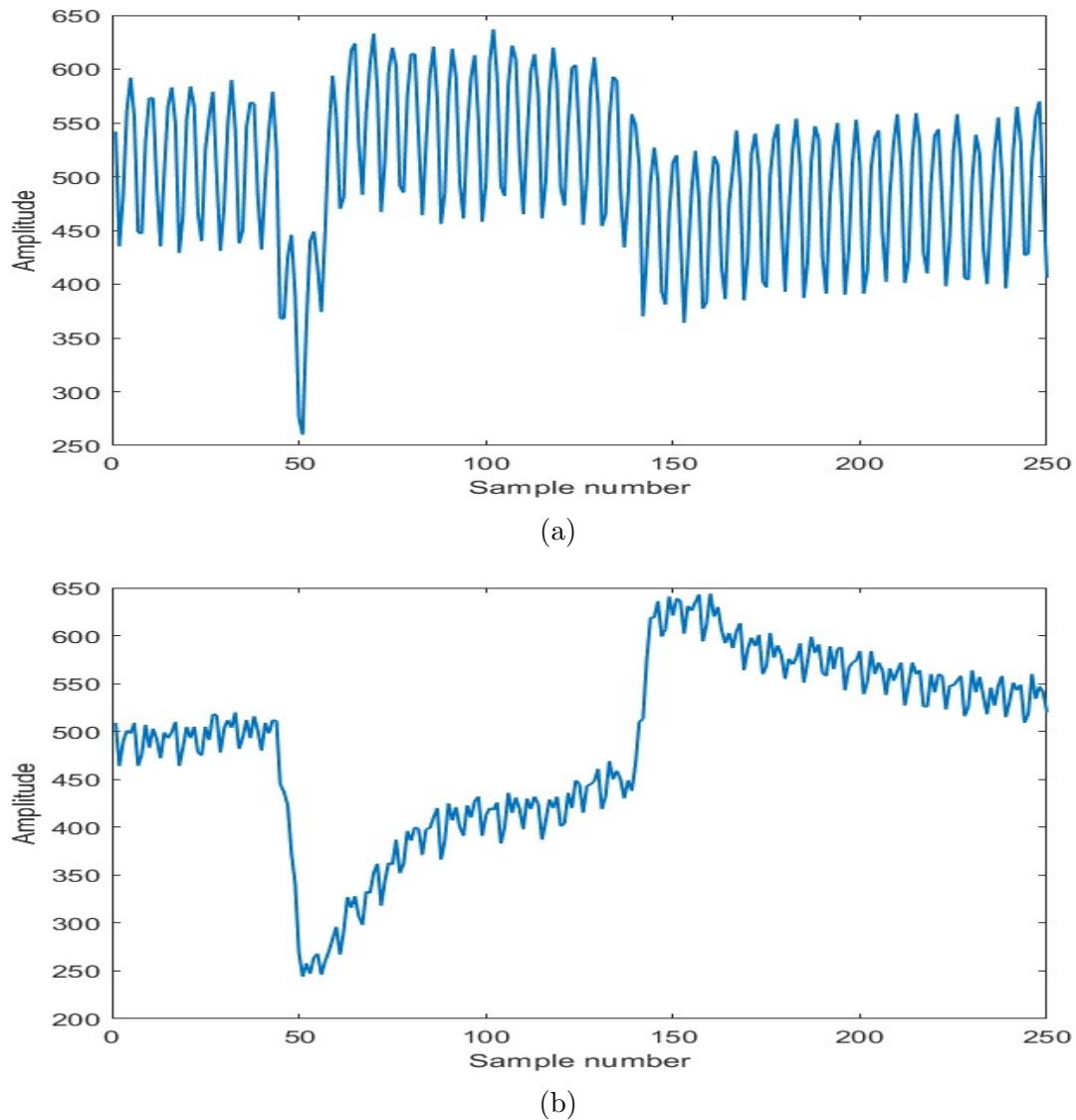


Figure 4.11: EOG signal in upward eye position for (a) left eye. (b) right eye.

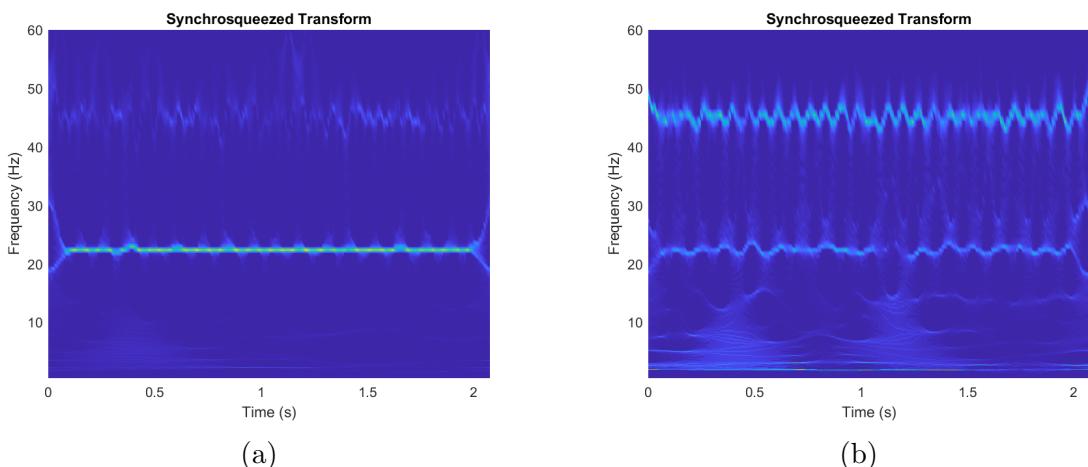


Figure 4.12: TFR of upward eye position for (a) left eye. (b) right eye.

Chapter 5

Transfer Learning

5.1 What is Transfer Learning?

Transfer learning makes use of knowledge from previously executed, related tasks to make learning and developing activity simpler. Utilizing existing knowledge and experience to pick up new, relevant theories more quickly than would otherwise be possible is the goal of transfer learning. The number of training examples required to achieve a target performance on a series of related learning problems, or reduced sample complexity, is a typical way to gauge the usefulness of transfer learning.

This is in contrast to the quantity of training examples required to achieve a target performance on a series of random learning problems. When they are already familiar with related ones, a human learner can often understand a new concept or method with the use of only a few training examples. For instance, learning to drive a van is much easier if we are already skilled at operating a vehicle. Learning French is slightly simpler than learning Chinese if we have learned English, while learning Spanish is slightly simpler than learning Portuguese (vs. German). We are therefore interested in studying the situations in which a learning machine might employ the advanced understanding that results from mastering earlier conceptions in order to help it perform better on future learning assignments. We are also curious about how these developments scale up as the model gets better at picking up a range of related notions [33].

Many practical applications, including speech recognition, computer vision, natural language processing, and cognitive research (such as functional magnetic resonance imaging brain state categorization), to mention a few, may profit from the capacity to transfer

information obtained from past tasks to facilitate learning a new one. Consider training a speech recognition program as an example.

A learning system that has been trained on a large number of speakers can distinguish typical speech patterns like accents or dialects, each of which calls for a distinctive speech recognizer. After that, it can swiftly identify a certain dialect from a small number of carefully picked instances and apply the previously learned recognizer for that dialect when a new individual is given a recognizer to train for. In this circumstance, one might consider the transmitted information to be the common features of each recognizer variation and, in a larger sense, the distribution of speech patterns existing in the population from which these subjects come. The knowledge transfer connected with distribution can help all the applications indicated above and a wide range of others.

Knowing how concepts are distributed among the community may frequently be very useful, assuming that these concepts (such as speech patterns) are independently sampled from a certain population. The target concepts are sampled in accordance with a fixed distribution in a more generic scenario. As we demonstrate below, if we are directly aware of the distribution, we may be able to drastically minimize the number of labeled instances needed to learn a target idea sampled in accordance with the distribution.

However, as we typically do not have direct access to this distribution in real-world learning environments, it would be ideal to be able to learn the distribution in some way based on observations from a series of learning problems with target concepts sampled in line with that distribution [28]. The goal is to reduce the number of labeled examples required to acquire incoming target concepts by employing an estimate of the distribution that was thus calculated, which can be almost as helpful as direct access to the true distribution.

Every formal learning or training program has one overarching goal: to better prepare students for life beyond the classroom. Student's ability to generalize and apply their knowledge from one setting to another, as well as their capacity for adaptation as a whole, should be a primary focus of every teaching-learning interaction.

5.2 Different ways of Transfer Learning

Transfer of learning happens in three different ways [34]:

- **Positive transfer:** Positive transfer is observed when skills or knowledge acquired in one context facilitate learning in another, making the new learning experience more accessible. For instance, the ability to play the violin can make learning the piano easier, and a strong understanding of mathematics can streamline the process of learning physics. Operating a scooter can positively transfer to operating a motorcycle, showcasing how skills acquired in one domain can ease the acquisition of related skills in another context.
- **Negative transfer:** Negative transfer occurs when proficiency in one task hinders the mastery of another job, making the learning process more challenging. An example of negative transfer is evident in language acquisition; speaking Telugu, for instance, may pose difficulties in learning Malayalam. Similarly, driving vehicles with left-hand drives can create challenges when attempting to learn how to operate vehicles with right-hand drives. In these cases, the skills or knowledge from one task interfere with or impede the successful acquisition of skills in a different, related task.
- **Neutral transfer:** Neutral transfer, also referred to as zero transfer, occurs when the learning of one activity has neither a positive nor negative impact on the learning of another task. In this scenario, the skills, knowledge, or experiences gained in one activity do not significantly influence or interfere with the acquisition of skills in a different task. It represents a neutral or independent relationship between the two learning experiences, where the success or difficulty in one task has no notable effect on the other.

The transfer of learning is explained by two significant hypotheses. The term “modern theories” refers to them [34].

The identical element idea was first proposed by E.L. Thorndike. The bulk of transfers, he claims, occur when two circumstances have a high degree of similarity or identity. The hypothesis states that the degree of transfer is proportional to the degree of similarity between the two scenarios. The extent of transfer grows in proportion to the degree of similarity between the components. If you already know how to ride a bike, you won’t have any trouble getting the hang of a moped. This swap can be accomplished in a reasonable amount of time since the two vehicles share so many parts.

Theorem of experience generalisation: Charles Judd created this notion. Because the learner learns a general principle while studying task “A,” theory of generalisation postulates that what is learnt in task “A” transfers to task “B” in some way. We benefit to the extent that our experiences, routines, and knowledge can be generalised and applied to different circumstances.

Chapter 6

Classification

6.1 Introduction

Artificial intelligence has advanced significantly, helping to close the gap between human and machine capabilities [35]. Experts and enthusiasts alike know they need to examine several different dimensions to really excel. One example of such a field is computer vision [36], but there are many others.

The goal of this area of study is to develop methods through which robots may acquire sensory abilities on par with those of humans. This knowledge may subsequently be used by machines for a wide range of applications, such as picture and image categorization, media recreation, recommendation systems, natural language processing, etc. Convolutional neural networks (CNNs) have been the primary method used to create and enhance deep learning-based computer vision advancements throughout time.

CNNs are deep learning systems that are capable of taking an input image, differentiating between different objects and components in the image by applying weights and biases to them. CNN requires a lot less pre-processing than other classification techniques. In contrast to fundamental approaches, CNN may given enough training, acquire such filters and attributes on its own [37]. A CNN is organised similarly to the human brain's connectivity network of neurons since its design was inspired by the visual cortex. Each neuron can only process information from receptive field which is known as a small area of the visual field. All of the visual field is included inside this cluster of regions. By using the proper filters, a CNN may successfully capture the spatial and temporal relationships in a picture. The image dataset is better suited by the architecture since it can reuse

weights and uses fewer parameters. In other words, the network may be taught to better understand how complicated the picture is.

CNNs are composed of interconnected layers, including an input layer, multiple hidden layers, and an output layer. Each node within these layers is assigned a weight and threshold. Activation occurs when the output exceeds the specified threshold, transmitting data to the subsequent layer. CNNs leverage convolution layers with filters to identify distinct features in images, facilitating the recognition of unique characteristics. Through training on image datasets, these networks internally learn relevant features. In deep CNNs, the training process can be time-intensive, attributed to the substantial number of parameters or weights involved [28].

CNN architectures have been very useful in the creation of algorithms that today power AI and will continue to power AI in the near future. The following is a list of some of them [36]:

- SqueezeNet [38]
- AlexNet [39]
- VGGNet [40]
- GoogleNet [41]
- ResNet [42]
- Xception [43]
- NASNet-Large [44]
- DenseNet [45]

The effectiveness of this method is assessed by comparing its performance using four pre-trained networks [46]: AlexNet, GoogLeNet, ResNet50, Xception. This evaluation was conducted within the MATLAB deep network designer environment. The characteristics of the same are shown in the table 6.2

Neural Network	Depth	Parameters (in Millions)	Input-Image Size
AlexNet	8	61.0	227 by 227
GoogLeNet	22	7.0	224 by 224
ResNet50	50	25.6	224 by 224
Xception	71	22.8	299 by 299

Table 6.1: Pre-trained neural network model specifications.

Training Option	Value
Solver	sgdm
InitialLearnRate	0.0003
MinBatchSize	10
MaxEpochs	12
ValidationFrequency	50

Table 6.2: Training hyperparameters chosen for the proposed work.

6.2 Data split

The subsequent step in the procedure involves partitioning the time-frequency plots data into distinct training and test sets. The model then undergoes exclusive training using the allocated training set. The significance of the test set becomes apparent after the training phase, serving as a critical measure to assess the CNN model's performance on new data. This evaluation is crucial for estimating the model's capacity to generalize to unseen data, offering valuable insights into its overall effectiveness.

There are 1200 images in total for all six types of eye movements, of which 200 are for each including left eye and right eye. So, as there are limited images, a randomized split of the dataset, designating 20% of the selected images for the test data and 80% for the training data is performed. Thus, each class in training data comprises 160 images and each class of test data comprises 40 images, respectively.

6.3 Training and Testing

In this study, the MATLAB deep network designer was utilized to customize a pre-trained neural network for classifying the training set of subjects. To enhance model robustness, 20% of the time-frequency plots were randomly selected from the training set (960 plots) to create validation data [47].

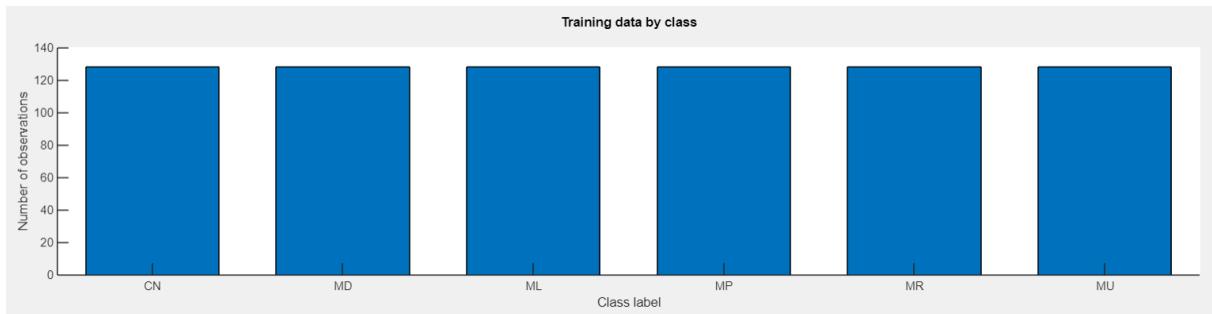


Figure 6.1: Training data distribution for all subjects.



Figure 6.2: Validation data distribution for all subjects.

The last learnable layer and the final classification layer to was modified to retrain the pre-trained network, adapting it for classifying input images to accommodate the new image data. All these models are trained on the ImageNet dataset with millions of labeled images.

6.4 Alexnet Pre-trained Network

Consists of eight layers: five convolutional layers and three fully connected layers First layer has 96 filters of size 11x11. Second layer has 256 filters of size 5x5. Uses a combination of convolutional and max-pooling layers. rectified linear unit (ReLU) activation is used for introducing non-linearity. Utilizes max-pooling layers to downsample spatial dimensions. Last three layers are fully connected. First fully connected layer has 4096

neurons. Dropout is applied to the first two fully connected layers to prevent overfitting. Final layer uses softmax activation for class probabilities.

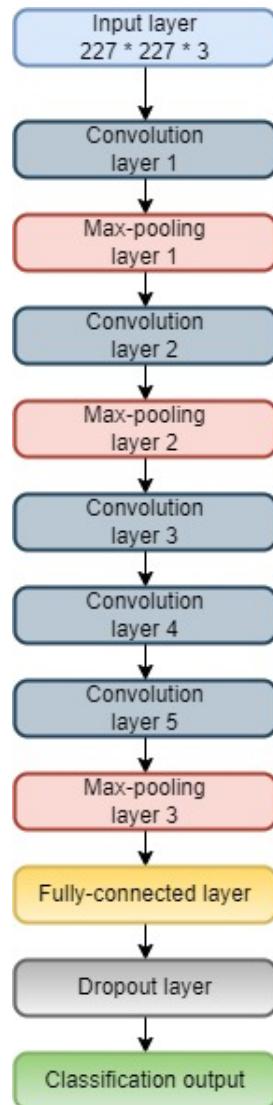


Figure 6.3: AlexNet architecture.

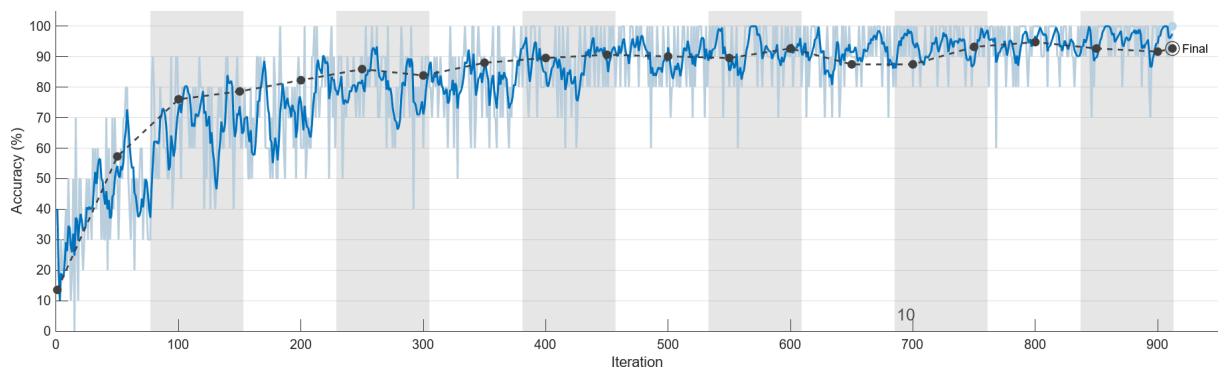


Figure 6.4: AlexNet training progress plot.

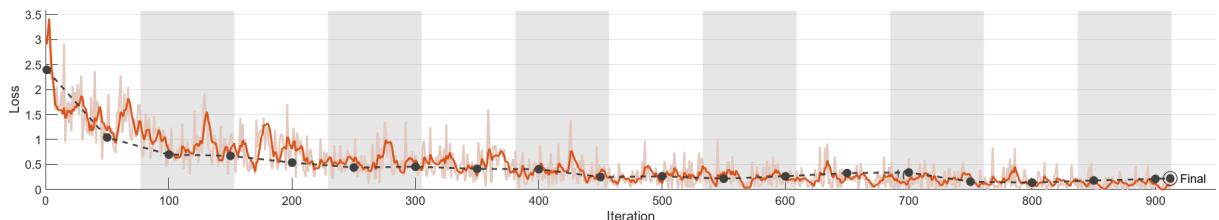


Figure 6.5: AlexNet training loss plot.

6.5 GoogleNet Pre-trained Network

Uses 9 inception modules in total. Employs a mix of 1x1, 3x3, and 5x5 convolutions within the modules. Replaces fully connected layers with global average pooling to reduce overfitting. Uses ReLU activation throughout the network. Applies batch normalization to normalize activations and improve convergence.

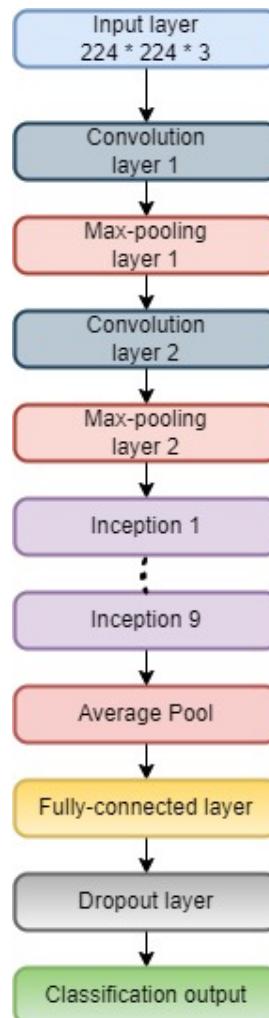


Figure 6.6: GoogleNet architecture.

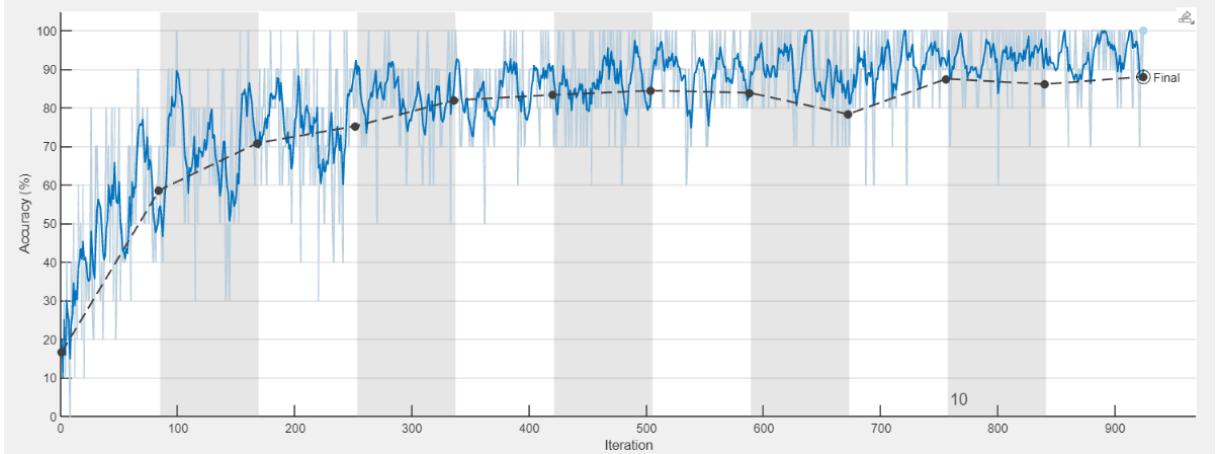


Figure 6.7: GoogleNet training progress plot.

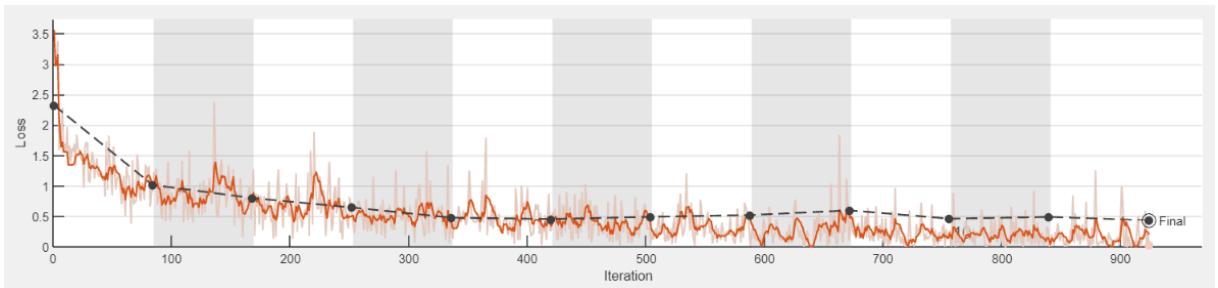


Figure 6.8: GoogleNet training loss plot.

6.6 Resnet 50 Pre-trained Network

Introduces residual blocks with shortcut connections that skip one or more layers, aiding in the training of very deep networks. Uses a bottleneck design in residual blocks, which reduces the computational cost of the network. Employs 3x3 convolutional filters in the residual blocks. The architecture includes bottleneck blocks with 1x1, 3x3, and 1x1 convolutions. Utilizes ReLU activation functions. Applies batch normalization to normalize activations and enhance training stability. The key innovation is skip connections, which help with the flow of gradients during back propagation.

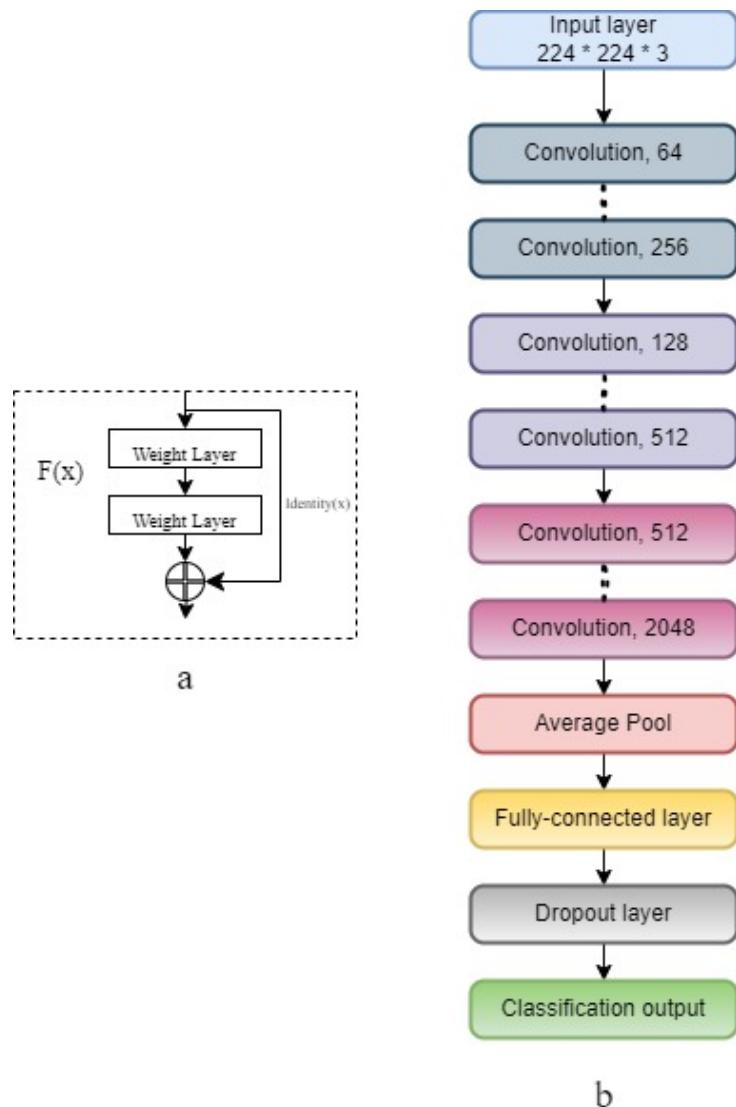


Figure 6.9: (a) Residual block, (b) ResNet50 architecture.

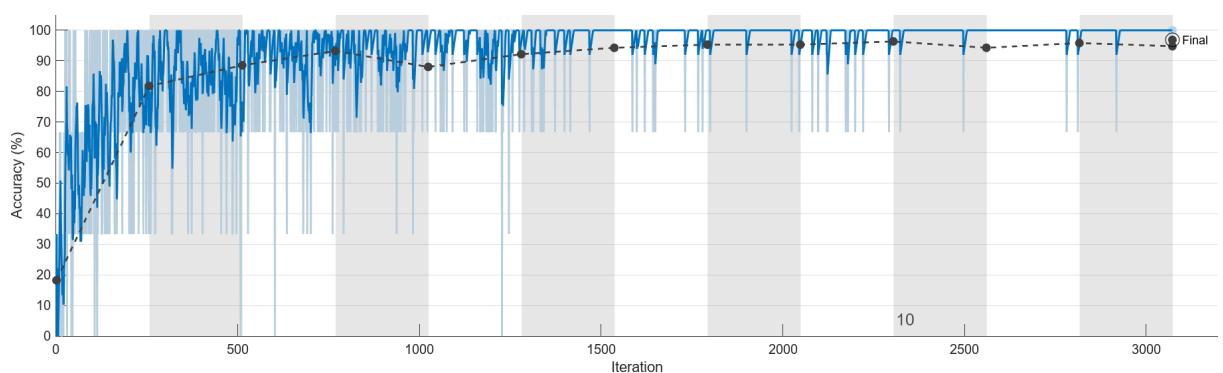


Figure 6.10: ResNet 50 training progress plot.

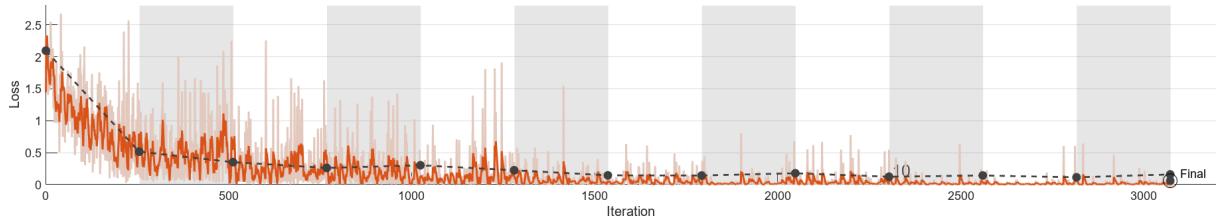


Figure 6.11: ResNet 50 training loss plot.

6.7 Xception Pre-trained Network

It is inspired by the Inception architecture but replaces standard convolutional layers with depthwise separable convolutions. Separates spatial and channel-wise convolutions, reducing computation and improving efficiency. Comprises an entry flow and an exit flow. The entry flow captures fine-grained features, while the exit flow refines the representation. Employs skip connections similar to residual networks, aiding in the flow of gradients during training. Uses ReLU activation functions. Applies batch normalization for normalization of activations and improved training stability. Typically ends with a global average pooling layer for dimensionality reduction.

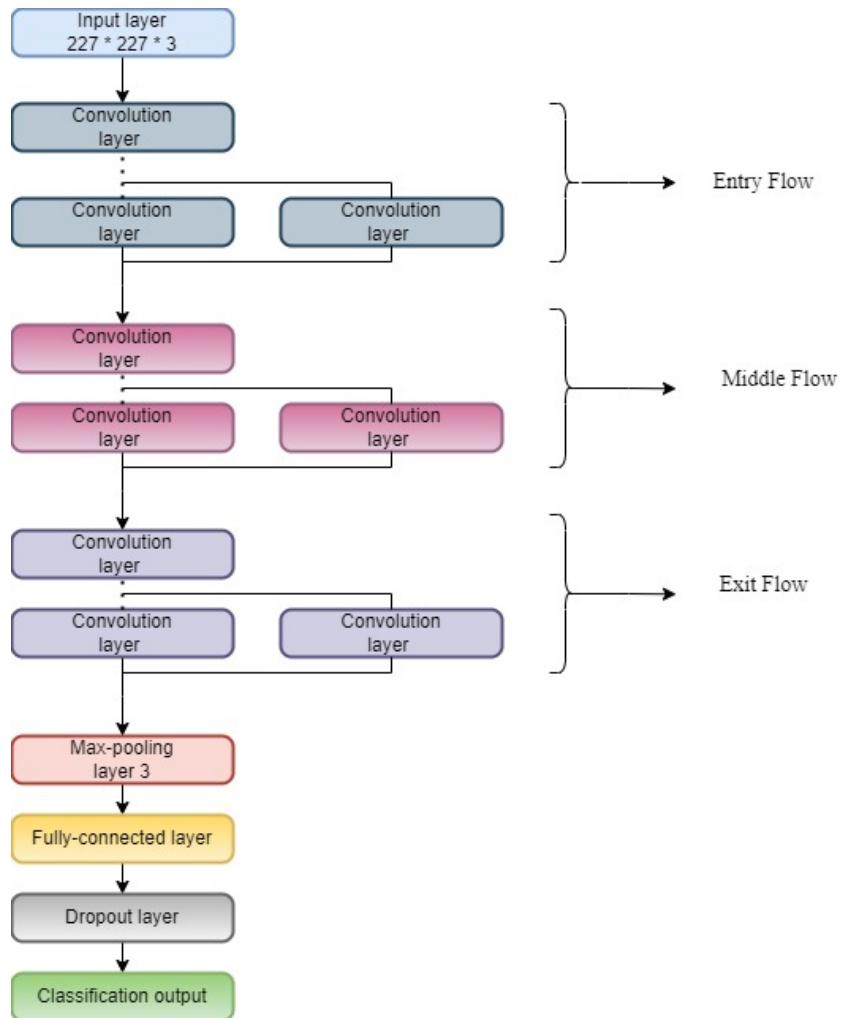


Figure 6.12: Xception architecture.

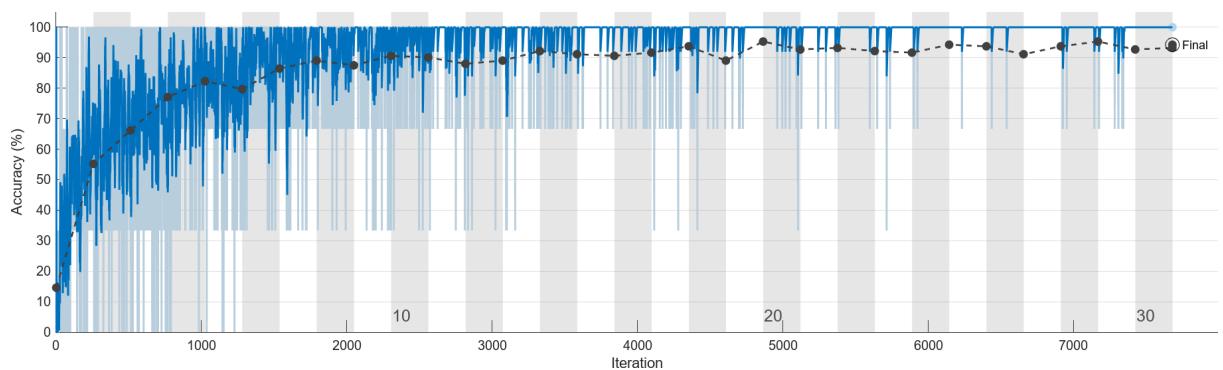


Figure 6.13: Xception training progress plot.

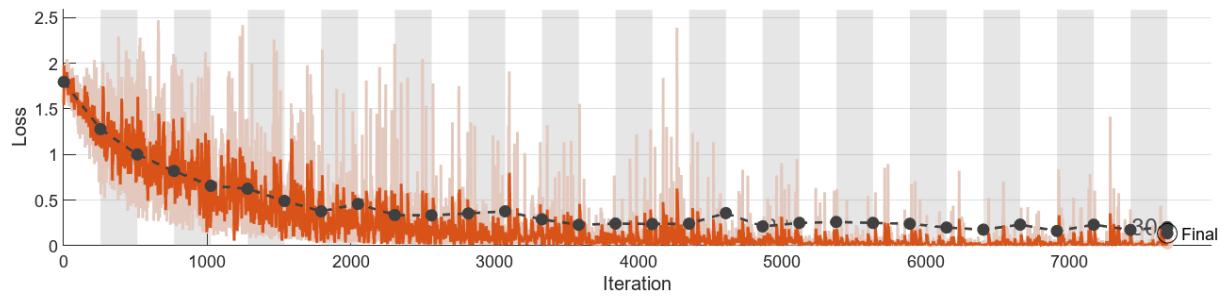


Figure 6.14: Xception training loss plot.

Chapter 7

Results and Discussion

The following chapter contains the results of the proposed methodology applied to the neural networks used in the study, explaining various parameters utilized to evaluate the Classification of Eye Movements. The evaluation involves testing the unseen test data, comprising time-frequency plots, on the trained CNN models. Parameters such as accuracy, precision, recall, specificity, and F1 score [48] are computed for each model across all subjects in the EOG of EOM dataset.

The metrics discussed in this section are determined as follows [49]:

- **True Positive (TP):** Accurately predicted positive values where the actual class is ‘yes,’ and the predicted class is also ‘yes.’
- **True Negative (TN):** Accurately predicted negative values where the actual class is ‘no,’ and the predicted class is also ‘no.’
- **False Positive (FP):** Instances where the actual class is ‘no,’ but the predicted class is ‘yes.’
- **False Negative (FN):** Instances where the actual class is ‘yes,’ but the predicted class is ‘no.’

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (7.1)$$

$$\text{Precision} = \frac{TP}{(TP + FP)} \quad (7.2)$$

$$\text{Recall(Sensitivity)} = \frac{TP}{(TP + FN)} \quad (7.3)$$

$$\text{Specificity} = \frac{TN}{(TN + FP)} \quad (7.4)$$

If we denote Precision by P and Recall by R

$$\text{F1 Score} = \frac{2 \times P \times R}{(P + R)} \quad (7.5)$$

This approach applies a majority voting scheme to the unseen test data of 6 different types of eye movements from the EMG of EOM dataset. The testing accuracy obtained is for all four pretrained networks is listed below 7.1.

S.No	Neural Network	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1 score (%)
1	AlexNet	90.00	91.78	90.00	98.00	89.84
2	GoogLeNet	92.08	92.21	92.08	98.42	92.03
3	Xception	91.67	92.05	91.67	98.33	91.60
4	ResNet50	95.00	95.08	95.00	99.00	94.97

Table 7.1: Multiclass test metrics of trained models.

A confusion matrix provides a summarized representation of predictions in a matrix format [50], illustrating correct and incorrect predictions for each class. This matrix is valuable for gaining insights into which classes the model may be confusing with others, highlighting areas of potential mis-classification [11]. Confusion matrices are obtained for the trained models in the methodology which are shown in the figures, 7.4.

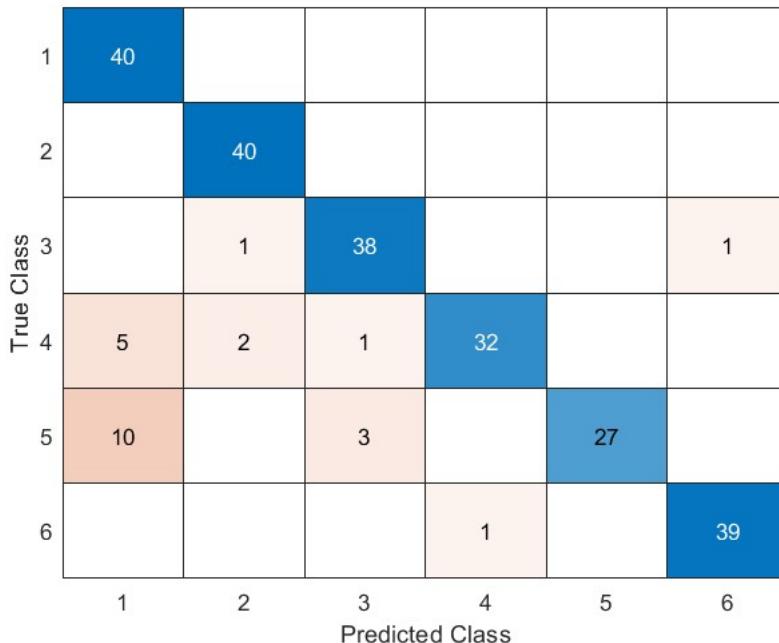


Figure 7.1: Confusion matrix for AlexNet model.

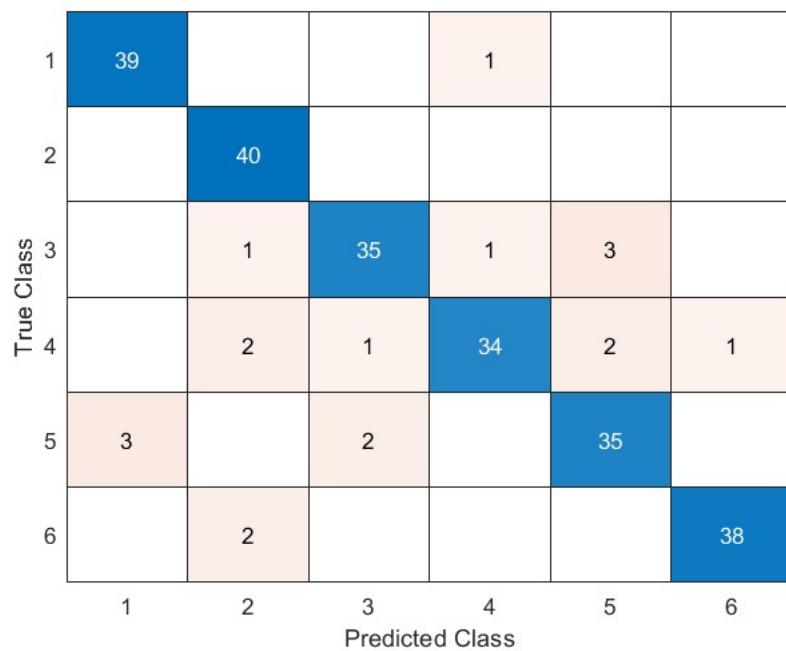


Figure 7.2: Confusion matrix for GoogleNet model.

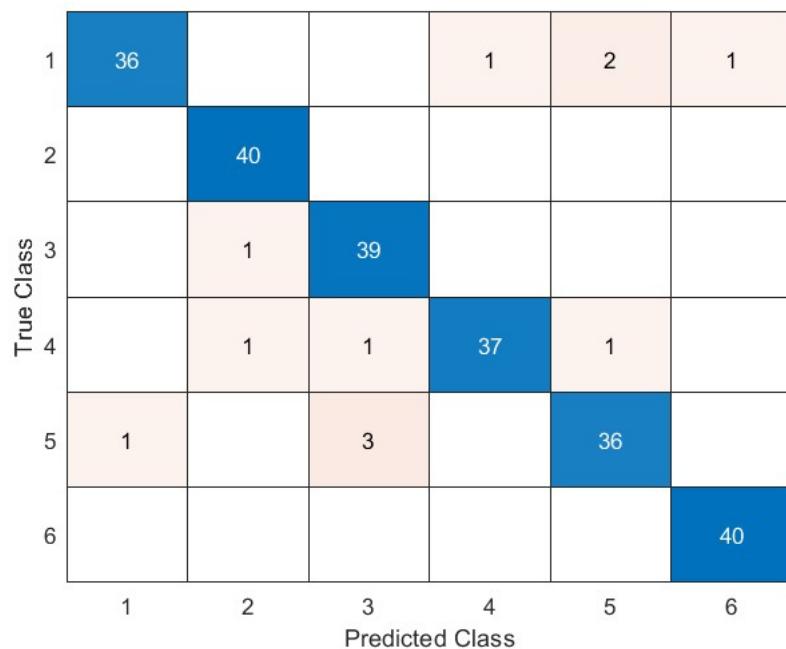


Figure 7.3: Confusion matrix for ResNet 50 model.

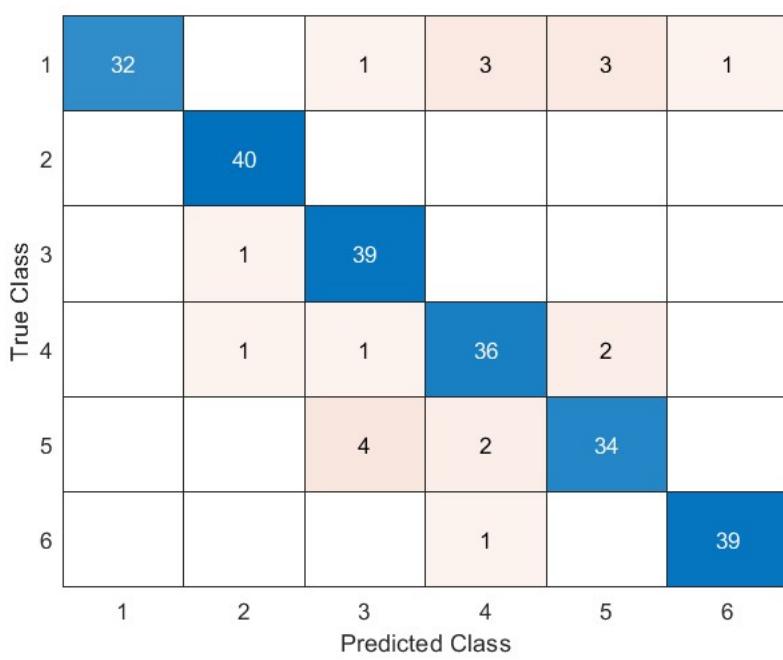


Figure 7.4: Confusion matrix for Xception model.

Chapter 8

Conclusions and Future Work

This chapter concludes the result obtained in previous chapter and mentions the future scope of the study.

8.1 Conclusions

This study presents an innovative approach to classify the six Eye Movements using EOG signals. Optimal results are attained by employing the ResNet50 pre-trained CNN model to classify time-frequency plots generated through WSST on the EMG of EOM dataset. Leveraging the capabilities of WSST ensures a thorough examination of dynamic patterns inherent in electrooculogram signals. WSST's sophisticated analysis enables a comprehensive exploration of the intricate temporal dynamics present in EOG data, providing valuable insights into the underlying patterns and facilitating a more nuanced understanding of Eye Movement activity. The integration of blind segmentation and transfer learning substantially reduces operational complexity. The methodology's effectiveness is assessed through the computation of various multi-class test metrics for all trained models, offering a thorough analysis of the implemented approach.

8.2 Future scope

- Although there is potential for extending the application of this method to other biomedical signals within the realm of human-computer interaction, the current study has been limited to a specific dataset. To establish robustness and generalizability, experiments can be conducted with an out-of-sample dataset in the future.
- Developing a customized CNN architecture for classifying the time-frequency rep-

resentations can increase test accuracy.

- Modifying the proposed methodology for real-time applications like continuous authentication or monitoring is crucial for practical deployment in various sectors, including healthcare and security.

References

- [1] J. Malmivuo and R. Plonsey, *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press, USA, 1995.
- [2] K. Shinomiya, N. Itsuki, M. Kubo, and H. Shiota, “Analyses of the characteristics of potential and cross-talk at each electrode in electro-oculogram,” *The Journal of Medical Investigation*, vol. 55, no. 1, 2, pp. 120–126, 2008.
- [3] F. Jagla, M. Jergelova, and I. Riečanský, “Saccadic eye movement related potentials.,” *Physiological research*, vol. 56, no. 6, 2007.
- [4] S. H. Nam, J. Y. Lee, and J. Y. Kim, “Biological-signal-based user-interface system for virtual-reality applications for healthcare,” *Journal of Sensors*, vol. 2018, 2018.
- [5] V. Asanza, “Electromyography (EMG) of the extraocular muscles (EOM),” 2021.
- [6] V. Asanza, E. Peláez, F. Loayza, I. Mesa, J. Díaz, and E. Valarezo, “Emg signal processing with clustering algorithms for motor gesture tasks,” in *2018 IEEE Third Ecuador Technical Chapters Meeting (ETCM)*, pp. 1–6, IEEE, 2018.
- [7] M. A. Fikri, P. I. Santosa, and S. Wibirama, “A review on opportunities and challenges of machine learning and deep learning for eye movements classification,” in *2021 IEEE International Biomedical Instrumentation and Technology Conference (IBITeC)*, pp. 65–70, IEEE, 2021.
- [8] A. N. Belkacem, D. Shin, H. Kambara, N. Yoshimura, and Y. Koike, “Online classification algorithm for eye-movement-based communication systems using two temporal EEG sensors,” *Biomedical Signal Processing and Control*, vol. 16, pp. 40–47, 2015.
- [9] X.-S. Li, Z.-Z. Fan, Y.-Y. Ren, X.-L. Zheng, and R. Yang, “Classification of eye movement and its application in driving based on a refined pre-processing and machine learning algorithm,” *IEEE Access*, vol. 9, pp. 136164–136181, 2021.

- [10] Y. Zhu, Y. Yan, and O. Komogortsev, “Hierarchical HMM for eye movement classification,” in *Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16*, pp. 544–554, Springer, 2020.
- [11] S. Aungsakun, A. Phinyomark, P. Phukpattaranont, and C. Limsakul, “Robust eye movement recognition using EOG signal for human-computer interface,” in *Software Engineering and Computer Systems: Second International Conference, ICSECS 2011, Kuantan, Pahang, Malaysia, June 27-29, 2011, Proceedings, Part II 2*, pp. 714–723, Springer, 2011.
- [12] C.-T. Lin, J.-T. King, P. Bharadwaj, C.-H. Chen, A. Gupta, W. Ding, and M. Prasad, “EOG-based eye movement classification and application on HCI baseball game,” *IEEE Access*, vol. 7, pp. 96166–96176, 2019.
- [13] H. CR and P. MP, “Classification of eye movements using electrooculography and neural networks,” *Int. J. Hum. Comput. Interact.(IJHCI)*, vol. 5, no. 4, p. 51, 2014.
- [14] A. N. Belkacem, H. Hirose, N. Yoshimura, D. Shin, and Y. Koike, “Classification of four eye directions from EEG signals for eye-movement-based communication systems,” *life*, vol. 1, p. 3, 2014.
- [15] J. Goltz, M. Grossberg, and R. Etemadpour, “Exploring simple neural network architectures for eye movement classification,” in *Proceedings of the 11th ACM symposium on eye tracking research & Applications*, pp. 1–5, 2019.
- [16] M. Thilagaraj, B. Dwarakanath, S. Ramkumar, K. Karthikeyan, A. Prabhu, G. Saravananakumar, M. P. Rajasekaran, N. Arunkumar, *et al.*, “Eye movement signal classification for developing human-computer interface using electrooculogram,” *Journal of Healthcare Engineering*, vol. 2021, 2021.
- [17] S. Ramkumar, G. Emayavaramban, J. M. A. Navamani, R. R. Devi, A. Prema, B. Booba, and P. Sriramakrishnan, “Human computer interface for neurodegenerative patients using machine learning algorithms,” in *Advances in Computerized Analysis in Clinical and Medical Imaging*, pp. 51–66, Chapman and Hall/CRC, 2019.
- [18] H. CR and P. MP, “Classification of eye movements using electrooculography and neural networks,” *Int. J. Hum. Comput. Interact.(IJHCI)*, vol. 5, no. 4, p. 51, 2014.

- [19] M. Thilagaraj, B. Dwarakanath, S. Ramkumar, K. Karthikeyan, A. Prabhu, G. Saravananakumar, M. P. Rajasekaran, N. Arunkumar, *et al.*, “Eye movement signal classification for developing human-computer interface using electrooculogram,” *Journal of Healthcare Engineering*, vol. 2021, 2021.
- [20] A. N. Belkacem, H. Hirose, N. Yoshimura, D. Shin, and Y. Koike, “Classification of four eye directions from EEG signals for eye-movement-based communication systems,” *life*, vol. 1, p. 3, 2014.
- [21] C.-H. Hsieh and Y.-H. Huang, “Low-complexity EEG-based eye movement classification using extended moving difference filter and pulse width demodulation,” in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 7238–7241, IEEE, 2015.
- [22] C.-T. Lin, J.-T. King, P. Bharadwaj, C.-H. Chen, A. Gupta, W. Ding, and M. Prasad, “EOG-based eye movement classification and application on HCI baseball game,” *IEEE Access*, vol. 7, pp. 96166–96176, 2019.
- [23] A. N. Belkacem, D. Shin, H. Kambara, N. Yoshimura, and Y. Koike, “Online classification algorithm for eye-movement-based communication systems using two temporal EEG sensors,” *Biomedical Signal Processing and Control*, vol. 16, pp. 40–47, 2015.
- [24] I. Daubechies, *Ten lectures on wavelets*. SIAM, 1992.
- [25] A. N. Akansu and R. A. Haddad, *Multiresolution signal decomposition: transforms, subbands, and wavelets*. Academic press, 2001.
- [26] P. Yadav, “A brief description of wavelet and wavelet transforms and their applications,” *International Journal of Statistics and Applied Mathematics*, vol. 3, no. 1, pp. 266–271, 2018.
- [27] S. Madhavan, R. K. Tripathy, and R. B. Pachori, “Time-frequency domain deep convolutional neural network for the classification of focal and non-focal EEG signals,” *IEEE Sensors Journal*, vol. 20, no. 6, pp. 3078–3086, 2019.
- [28] Y. U. Vishwanath, S. Esakkirajan, B. Keerthiveena, and R. B. Pachori, “A generalized classification framework for power quality disturbances based on synchrosqueezed wavelet transform and convolutional neural networks,” *IEEE Transactions on Instrumentation and Measurement*, 2023.

- [29] I. Daubechies, J. Lu, and H.-T. Wu, “Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool,” *Applied and computational harmonic analysis*, vol. 30, no. 2, pp. 243–261, 2011.
- [30] T. Oberlin, S. Meignen, and V. Perrier, “The Fourier-based synchrosqueezing transform,” in *2014 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 315–319, IEEE, 2014.
- [31] I. Selesnick, “Wavelet transform with tunable Q-factor IEEE transactions on signal processing. 2011 aug; 8 (59) pp: 3560-75. doi: 10.1109/TSP.2011.
- [32] R. B. Pachori, *Time-frequency analysis techniques and their applications*. CRC Press, 2023.
- [33] K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” *Journal of Big data*, vol. 3, no. 1, pp. 1–40, 2016.
- [34] E. S. Olivas, J. D. M. Guerrero, M. Martinez-Sober, J. R. Magdalena-Benedito, L. Serrano, *et al.*, *Handbook of research on machine learning applications and trends: Algorithms, methods, and techniques: Algorithms, methods, and techniques*. IGI global, 2009.
- [35] P. G. Buta, “A few preliminary considerations on infringement of copyright by artificial intelligence,” *Rom. J. Intell. Prop. L.*, p. 41, 2023.
- [36] S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” in *2017 international conference on engineering and technology (ICET)*, pp. 1–6, IEEE, 2017.
- [37] K. O’Shea and R. Nash, “An introduction to convolutional neural networks,” *arXiv preprint arXiv:1511.08458*, 2015.
- [38] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, “SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 mb model size,” *arXiv preprint arXiv:1602.07360*, 2016.
- [39] S. Samir, E. Emry, K. El-Sayed, and H. Onsi, “Optimization of a pre-trained AlexNet model for detecting and localizing image forgeries,” *Information*, vol. 11, no. 5, p. 275, 2020.

- [40] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.
- [41] J. Rajasegaran, V. Jayasundara, S. Jayasekara, H. Jayasekara, S. Seneviratne, and R. Rodrigo, “Deepcaps: Going deeper with capsule networks,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10725–10733, 2019.
- [42] M. Shafiq and Z. Gu, “Deep residual learning for image recognition: A survey,” *Applied Sciences*, vol. 12, no. 18, p. 8972, 2022.
- [43] J. Carreira, H. Madeira, and J. G. Silva, “Xception: A technique for the experimental evaluation of dependability in modern computers,” *IEEE Transactions on Software Engineering*, vol. 24, no. 2, pp. 125–136, 1998.
- [44] İ. İlhan, E. Bahı, and M. Karaköse, “An improved deepfake detection approach with NASNetlarge CNN,” in *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*, pp. 598–602, IEEE, 2022.
- [45] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.
- [46] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Advances in neural information processing systems*, vol. 25, 2012.
- [47] T. Cody, E. Lanus, D. D. Doyle, and L. Freeman, “Systematic training and testing for machine learning using combinatorial interaction testing,” in *2022 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pp. 102–109, IEEE, 2022.
- [48] M. J. J. Ghrabat, G. Ma, I. Y. Malood, S. S. Alresheedi, and Z. A. Abduljabbar, “An effective image retrieval based on optimized genetic algorithm utilized a novel SVM-based convolutional neural network classifier,” *Human-centric Computing and Information Sciences*, vol. 9, pp. 1–29, 2019.
- [49] D. van Ravenzwaaij and J. P. Ioannidis, “True and false positive rates for different criteria of evaluating statistical evidence from clinical trials,” *BMC medical research methodology*, vol. 19, pp. 1–10, 2019.

- [50] H. Turki, M. A. H. Taieb, and M. B. Aouicha, “Knowledge-based construction of confusion matrices for multi-label classification algorithms using semantic similarity measures,” *arXiv preprint arXiv:2011.00109*, 2020.