



ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA

SCUOLA DI INGEGNERIA E ARCHITETTURA

DIPARTIMENTO DI

INGEGNERIA DELL'ENERGIA ELETTRICA E DELL'INFORMAZIONE

"Guglielmo Marconi"

DEI

**CORSO DI LAUREA MAGISTRALE IN
INGEGNERIA ELETTRONICA**

TESI DI LAUREA

in

Hardware-Software Design of IOT Systems M

Smart Glasses as a Sensor Fusion Platform for Acquisition and Processing of ExG and Image Data

Relatore:

Chiar.mo Prof. Davide Rossi

Presentata da:

Andrea Helga Bernardi

Correlatori:

Prof. Simone Benatti

Pierangelo Maria Rapa

Sebastian Frey

Sessione V

Anno Accademico 2022/2023

Alla mia famiglia e ai miei amici, che hanno sempre creduto in me.

Ringraziamenti

Vorrei ringraziare il mio relatore, Prof. Davide Rossi, ed esprimere un sincero ringraziamento al mio correlatore, Prof. Simone Benatti, per avermi nuovamente offerto l'opportunità di lavorare in un campo che mi appassiona profondamente. Grazie anche al Dott. Andrea Cossetini e al Prof. Luca Benini, per l'indimenticabile esperienza a Zurigo.

Un ringraziamento speciale a Pierangelo e Sebastian, per essere stati degli ottimi mentori e per il vostro inestimabile supporto. Grazie anche a Mattia e Victor, per il vostro piccolo ma prezioso aiuto durante questa tesi. Desidero esprimere la mia gratitudine a tutti voi per tutto ciò che mi avete insegnato e per avermi ispirato ancor più passione per questo ambito.

Grazie anche a tutti i ragazzi del Micrel Lab, per le risate e per l'accoglienza. Non vedo l'ora di far parte a tutti gli effetti del gruppo.

Grazie a Fawzia, la migliore amica che abbia mai potuto trovare. Grazie per le tue parole, il tuo affetto e la tua presenza. Grazie ad Aurora, per essere stata la mia guida e una grande amica a Zurigo, senza di te non sarebbe stato lo stesso. Grazie a tutti i nuovi amici zurighesi per i momenti condivisi insieme. E un grazie speciale agli amici di Bologna e ai colleghi dell'università, con i quali ho condiviso due anni pieni di emozioni e di lamentele.

E seppur in questi ultimi due anni siamo stati molto lontani, voglio ringraziare tutta la mia famiglia. Un sincero e profondo ringraziamento a mia mamma Annelise, per tutto il supporto sempre dato. Lo sai, sei una mamma incredibile, la migliore del mondo. Grazie a mio babbo, mio fratello e tutto il resto della famiglia, anche quella acquisita e allargata, per la vostra incondizionata presenza e tutto l'affetto.

Ma soprattutto, desidero ringraziare te, Massimiliano, per essere il mio compagno di vita. Grazie per tutti questi anni trascorsi insieme. La tua presenza costante in tutti i momenti, sia quelli belli che quelli difficili, ha reso la mia vita migliore. Grazie per il tuo sostegno incondizionato, la tua forza e l'amore che mi doni ogni giorno. Sono grata di averti al mio fianco, sei davvero un uomo meraviglioso.

Sommario

L'interazione con oggetti tramite l'analisi dei movimenti delle mani è una importante sfida nel campo delle interfacce uomo-macchina (HMI). Una comprensione più profonda di tali meccanismi è essenziale per migliorare la funzionalità dei dispositivi indossabili che possono essere utilizzati in diverse applicazioni HMI. Ciò richiede un approccio innovativo oltre i metodi tradizionali che possano mimare la complessa interazione dei movimenti delle mani.

Sebbene l'attività elettrica muscolare sia comunemente utilizzata in questo campo, le immagini e i movimenti degli occhi forniscono importanti indicazioni sull'intenzione di una persona nell'esecuzione di gesti. Gli *smart glasses* offrono una piattaforma promettente per la cattura di elettrooculogrammi (EOG) e immagini. Possono anche fungere da gateway di integrazione per una *Body Area Network*, collegandosi ad altri biosensori come, ad esempio, un dispositivo per l'elettromiografia (EMG).

Questa tesi si propone di sviluppare un prototipo di sistema composto da *smart glasses* con sensori EOG, una scheda di acquisizione immagini e un dispositivo EMG. Questi dati vengono trasmessi ad un computer, che ne consente la sincronizzazione e l'elaborazione.

La prima parte della tesi consiste nel determinare il posizionamento ottimale dei sensori EOG sugli occhiali per garantire una corretta acquisizione del segnale rispetto ai metodi tradizionali a base di gel. Per tale scopo, è stato sviluppato un prototipo di occhiali costituito da elettrodi asciutti di Dätwyler, interfacciati con la piattaforma BioGAP. Uno studio esplorativo, basato sulla classificazione di 9 movimenti oculari, in cui è stato confrontato questo prototipo con una configurazione EOG convenzionale, ha mostrato una promettente accuratezza del 95% contro il 96% della configurazione tradizionale. Nella seconda parte della tesi, si impiega la piattaforma GAP9 Shield, dotata di una piccola camera e di un modulo Wi-Fi per l'invio delle immagini. Si procede a stabilire una connessione con BioGAP, assicurando un flusso di comunicazione sincronizzato con un ritardo medio di 120 ms, paragonabile al periodo di frame di 166 ms. Infine, un modello CNN MobileNet V2 è stato adattato per la classificazione binaria di immagini di mano aperta e chiusa, ottenendo un'accuratezza del 97%.

Questa tesi di laurea magistrale è stata parzialmente condotta presso l'ETH di Zurigo.

Abstract

The field of hand movements and grasp control has long faced challenges in replicating the natural movements of the human hand. A deeper understanding of such mechanisms is essential to enhance the functionality of wearable devices that can be used in multiple Human Machine Interface (HMI) applications. This necessitates an innovative approach beyond traditional methods that can mimic the complex interplay in hand movements. Although electrical muscle activity is commonly used in this field, images and eye movements provide important cues about a person's intention during hand gestures by offering visual context.

Smart glasses offer a promising platform for capturing electrooculograms (EOG) and images. They can also serve as an integration gateway for a Body Area Network, connecting with other biosensor nodes such as an electromyogram (EMG) device.

This thesis aims to develop a system consisting of a prototype for smart glasses with EOG sensors, an image acquisition board, and an EMG acquisition device. These data are wirelessly transmitted to a computer, which allows synchronization and processing.

The first part of the thesis consists of determining the optimal placement of EOG sensors on the glasses to balance accuracy and ensure robust signal acquisition with respect to traditional gel-based methods. To achieve this, a glasses prototype integrating dry electrodes from Dätwyler, interfaced with the BioGAP platform, was developed and evaluated. An exploratory study involving 9-task eye movement classification, comparing this prototype with a conventional EOG setup, showed a promising 95% accuracy for the glasses prototype, close to the traditional setup's 96%.

The second part of the thesis involves utilizing the GAP9 Shield, a platform with a small camera and a Wi-Fi shield for image transmission, and establishing a connection with BioGAP, guaranteeing synchronized communication with an average delay of 120 ms, comparable to the frame period of 166 ms.

Finally, a MobileNet V2 CNN model was adapted for binary image classification of open and closed hands, achieving an accuracy of 97%.

This master's thesis has been partially conducted at ETH Zürich.

Contents

List of Acronyms	xv
1 Introduction	1
1.1 HGR and Grasp Control	1
1.2 Contribution and Objective	2
1.3 Thesis Structure	3
2 Background	5
2.1 EOG Theory	5
2.1.1 Eyes Dipole	5
2.1.2 EOG Acquisition methods	6
2.1.3 EOG Signal	7
2.2 EMG Theory	7
2.2.1 Origin of the Myoelectric Signal	7
2.2.2 EMG Acquisition Method	8
2.2.3 sEMG signal	8
2.3 State of the art	9
2.3.1 Eye tracking	9
2.3.2 Smart Glasses	10
2.4 BioGAP	10
2.4.1 GUI	11
2.5 Collaboration with Dätwyler	12
3 The EOG Prototype	15
3.1 Signal Acquisition	15
3.1.1 Electrodes	15
3.1.2 Acquisition Setup	16
3.1.3 Experimental Test	17

Contents

3.2	Data Processing	19
3.2.1	Signal Filtering	19
3.2.2	Signal Segmentation and Feature Extraction	20
3.2.3	Machine Learning	21
4	Camera Module	23
4.1	GAP9 Shield V2	23
4.1.1	Firmware Description	24
4.1.2	Stream and save of images	24
4.2	Classification: MobileNetV2	25
4.2.1	Inverted Residual and Linear Bottleneck Block Structure	26
4.2.2	Transfer Learning via Fine-Tuning	28
5	EMG and Image System	29
5.1	Connection with BioGAP	29
5.2	Acquisition Phase	30
5.3	Offline Processing	30
6	Results	33
6.1	EOG Results	33
6.1.1	Filter Stage	33
6.1.2	Shape Description	35
6.1.3	Amplitude Range	36
6.1.4	Classification Performances	36
6.1.5	K-Fold Cross-Validation Method	38
6.2	Image Classification	39
6.3	Synchronization Image and EMG Acquisition	39
7	Conclusion and Future Work	43
7.1	Conclusion	43
7.2	Future Work	44

List of Figures

2.1	The eye-dipole model with EOG waveform [1]	5
2.2	Electrodes placement for traditional method [2]	6
2.3	Electrodes placement in JINS Meme glasses [3]	7
2.4	EMG Armband with dry electrodes employed in [4]	8
2.5	EMG signal with three contractions over time	9
2.6	BioGAP platform	11
2.7	BioGAP System Diagram	12
2.8	3D-prototype with EEG and EOG dry electrodes presented to Datwyler	13
3.1	Electrodes placement in glasses prototype	16
3.2	Traditional acquisition method	17
3.3	Active gel-electrode used for reference and traditional setup	17
3.4	Hardware Setup for EOG acquisition	18
3.5	Points to look in the projector	18
3.6	Experimental test, using Psychopy, at the projector	19
3.7	Filtering process for EOG	19
3.8	Horizontal EOG from glasses and Trigger Vector	20
4.1	GAP9 Wi-Fi Shield	23
4.2	GAP9 Shield Block Diagram	24
4.3	Image-frames streams in the computer	25
4.4	MobileNet V2 Structure	26
4.5	Data flow in Bottleneck Residual Block [5]	26
4.6	Structure Bottleneck Residual Block MobileNet V2 [5]	27
5.1	Trigger before and after its reconstruction	31
6.1	Baseline Drift in raw EOG with traditional and glasses method	33
6.2	High Pass Filtered Look Down EOG	34

List of Figures

6.3	Low Frequencies Noise Removal Look Down EOG	34
6.4	Comparison raw and filtered EOG from glasses	35
6.5	EOG Signals from glasses during Look Down-Left instruction	35
6.6	Confusion Matrix Traditional Method	37
6.7	Confusion Matrix Glasses Method	37
6.8	3-Fold Cross-Validation in glasses method	38
6.9	First and last frame with close hand to identify start and stop trigger value . . .	39
6.10	Synchronized acquisition of images and EMG	40
6.11	Synchronized acquisition of images and EMG	41
6.12	Acquisition of images and EMG with 126 ms of delay	41

List of Tables

2.1	Measurement configuration of BioGAP AFE	11
6.1	EOG amplitude range with traditional method	36
6.2	EOG amplitude range with glasses method	36

List of Acronyms

ADCAnalog to Digital Converter
AFEAnalog Front-End
BLEBluetooth Low Energy
CNNConvolutional Neural Network
CRPCorneal-Retinal Potential
CSI2Camera Serial Interface 2
DSPDigital Signal Processing
EEGElectroencephalography
EMGElectromyography
EOGElectrooculography
FPSFrames Per Second
GUIGraphical User Interface
HGRHand Gesture Recognition
HMIHuman Machine Interfaces
iEMGintramuscular-Electromyography
IIRInfinite Impulse Response

List of Acronyms

IMUInertial Measurement Unit
KNNK-Near-Neighbors
LPLow Pass Filter
MLMachine Learning
PCBPrinted Circuit Board
PPGPhotoplethysmogram
PULPParallel Ultra Low Power
RMSRoot Mean Square
SDKSoftware Development Kit
sEMGsurface-Electromyography
SoAState of the Art
SoCSystem on Chip
SPISerial Peripheral Interface
TCPTransmission Control Protocol

Introduction

Thanks to the advancement of technologies for biomedical applications and the continuous progress of wearable devices, the development of reliable and non-invasive systems for monitoring human vital functions is continually unfolding [6].

Smart glasses represent an innovative solution capable of acquiring biopotentials, like Electroencephalography (EEG) and electrooculography (EOG) [7], as well as images [8].

These devices open up numerous opportunities for engineering research applications in various fields, including Human Machine Interface (HMI) [9], assistive technologies [10], prosthetics [11], and robotics [12]. Moreover, focusing on hand movements, smart glasses have been used to capture hand gestures [13]. Other approaches explored in these studies involve using electrical muscular activity (Electromyography - EMG) as the primary responsible for hand movements. A drawback of this method lies in its lack of visual context of the surrounding environment. Sensor fusion between camera images and physiological signals offers a comprehensive method for analyzing complex hand movements and grasping actions [14]. Images provide valuable insight into the surrounding environment while incorporating EOG, which enables capturing the directional focus of the user's gaze and contributes to target identification. Furthermore, by integrating an EMG acquisition device to monitor muscular activity, the resulting system mirrors human hand control functionalities more accurately.

1.1 HGR and Grasp Control

Hand gestures are one of the most natural ways for humans to express intuitive intention, manipulate objects, and interact with the surrounding environment.

Consequently, Hand Gesture Recognition (HGR) is considered one of the most significant and

active research topic for developing HMIs [15], aiming to interpret human hand gestures using machine learning algorithms.

HGR encompasses various types of gestures, including grasping objects of different shapes, sizes, and textures, enabled by the coordination of the human brain and visual and tactile processes. Understanding this mechanism can be useful across various disciplines inside the HMI, for example:

- Virtual Reality and Gaming, where hand gestures work as the primary input, and a more natural and intuitive user experience is needed [16].
- Robotics and Automation, where the ability to manipulate objects precisely and in various complex manners is essential [17].
- Rehabilitation and Assistive Technologies, where individuals with motor impairments or disabilities can improve their quality of life [18].

Various techniques have been developed in these fields, and current methods primarily rely on visual sensors [19] that interpret gestures by perceiving the environment and hand shape or EMG sensors [4] that directly capture muscle signals. However, each method has limitations: EMG signals vary due to anatomical differences between people, whereas visual-based approaches are affected by changing lighting conditions. This inadequacy underscores the importance of adopting a hybrid approach, which integrates multiple modalities to overcome these challenges by utilizing their combined strengths.

1.2 Contribution and Objective

The user's intention during hand movements and grasping can be deduced from the electrical activity of muscles and images, complemented by an understanding of the target facilitated by eye tracking. At the same time, the primary challenge in HMI lies in developing increasingly precise, minimally invasive, and user-friendly systems [20]. This thesis tackles these challenges by coupling smart glasses for EOG and image acquisitions with an EMG acquisition system, obtaining a system replicating the human capabilities for hand control functionalities. This integration process involves the following main steps:

1. Develop a prototype of smart glasses based on dry electrodes for EOG acquisition, comparing it with a traditional design based on wet electrodes.
2. Incorporate the system with a nano camera, ensuring a wireless connection and transmission of images.
3. Integrate and synchronize the smart glasses with an EMG acquisition system.

1.3 Thesis Structure

The rest of this master thesis is structured as follows:

- Chapter 2 provides an overview of the necessary background for understanding subsequent developments, including explanations of EOG and EMG, State of Art (SoA), the platform utilized for experimentation (BioGAP), and the collaboration with the partnering company.
- Chapter 3 explains the EOG module in detail, focusing on signal acquisition and processing.
- Chapter 4 illustrates the camera module, its firmware, and the employed image classification.
- Chapter 5 describes the procedure for synchronizing the camera module with an EMG acquisition system based on BioGAP.
- Chapter 6 presents the results obtained for each module.
- Chapter 7 exposes the conclusions and outlines the future work.

Background

2.1 EOG Theory

Electrooculography (EOG) is a technique designed to measure the Corneal-Retinal Potential (CRP), that is, the potential between the front (cornea) and back (retina) of the human eye. The CRP results in the observed signal called electrooculogram.

2.1.1 Eyes Dipole

The human eyes act as a dipole. Therefore, the EOG signal varies depending on the eye's direction relative to a specific electrode, resulting in a positive or negative trend.

As illustrated in Figure 2.1, the signal increases when the cornea approaches an electrode and decreases when the cornea moves in the opposite direction. Finally, the EOG reflects changes associated with eye movements.

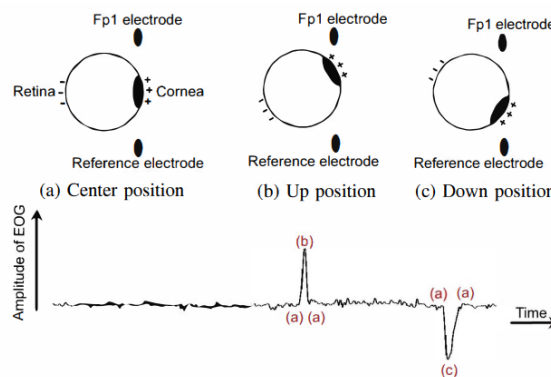


Figure 2.1: The eye-dipole model with EOG waveform [1]

2.1.2 EOG Acquisition methods

As the EOG represents the eye movement, two different signals are typically defined to distinguish the horizontal and vertical movements, computed according to the electrode placement. A standardized acquisition setup has not yet been developed. Following are the two methods used in this thesis: the commonly used approach and an innovative alternative.

Traditional Method

The classical way to acquire EOG is to use gel electrodes attached at strategic points of the face. As reported in [21], multiple possible configurations exist. The most used one has four electrodes placed as described below, and the resulting EOG signals are:

- Vertical signal: differential measurement between one electrode above the eyebrows and one below the eye of the left part of the face.
- Horizontal signal: differential measurement between the two electrodes placed on the distal ends of the forehead by the corner of the eyes.

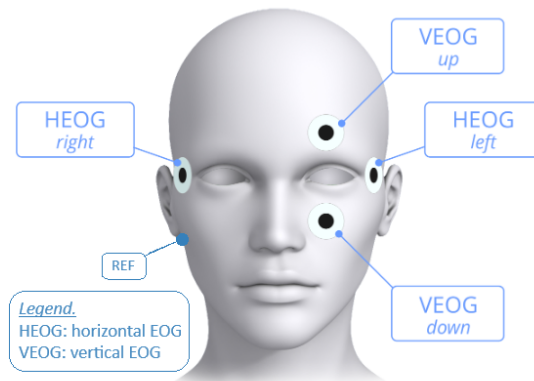


Figure 2.2: Electrodes placement for traditional method [2]

Glasses Method

Researchers tried to find a more comfortable, wearable solution. As introduced in Chapter 1, it is possible to capture EOG using only three electrodes placed in the frame of a pair of glasses: one in the bridge and two in the nose pads (left and right), as shown in Figure 2.3. In such a method, the horizontal signal is given by the contribution of the left and right pads, while the vertical one can be obtained by considering the bridge and both nose pads.

2 Background



Figure 2.3: Electrodes placement in JINS Meme glasses [3]

2.1.3 EOG Signal

The EOG signal, which reflects eye movements, typically exhibits a potential difference with amplitudes ranging from 0 to 100 mV, proportionally to the direction of the eye movements. Additionally, it demonstrates a significant frequency contribution from 0.1 to 40 Hz, as reported in [22]. Like other biomedical signals, the EOG can be affected by multiple noise sources. The most common ones are power line interference (50Hz), environmental noise (due to nearby electronic equipment), cross-talk (the presence of other biosignals nearby), and even motion artifacts (like facial expressions). In terms of signal processing, there's still no standardized procedure, as it strongly depends on the acquisition system used, as mentioned in 2.1.2.

2.2 EMG Theory

Electromyography (EMG) is a functional diagnostic technique for evaluating and detecting muscle electrical activity, called myoelectric signal, during contractions.

2.2.1 Origin of the Myoelectric Signal

Muscles are tissues primarily composed of muscle fibers with contractile capabilities.

The motor unit is the smallest functional unit responsible for muscle contraction, consisting of the α motor neuron and the muscle fibers it innervates. The nerve impulse sent by the motor neuron to the muscle fibers causes an electric signal to occur in them. More precisely, since a motor unit consists of multiple muscle fibers, what is observed is the sum of the amplitudes of all electrical signals of the activated fibers.

2.2.2 EMG Acquisition Method

There are two main types of EMG: intramuscular electromyography (iEMG), which relies on needle electrodes penetrating the skin to reach the muscle of interest and surface electromyography (sEMG) which relies on electrodes placed on the skin above the muscle of interest. Being painless and non-invasive [23], the latter is preferred for developing HMIs.

In sEMG, both dry and wet electrodes can be employed, and they can be used as individual electrodes positioned at specific points of interest or as superficial patches, such as electrode arrays integrated into armbands, as shown in Figure 2.4. This setup allows for recording across a two-dimensional surface, facilitating the acquisition of more comprehensive information from different muscle fibers. However, surface electrodes collect signals from the underlying tissue volume, potentially compromising signal quality [24].

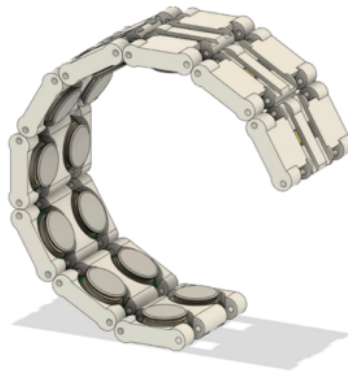


Figure 2.4: EMG Armband with dry electrodes employed in [4]

2.2.3 sEMG signal

The sEMG signal captures the dynamic pattern of muscle activity over time, delineating moments of contraction and rest with zero mean, as shown in Figure 2.5. It offers insights into the movement performed by the muscle under examination, as its amplitude reflects the intensity of muscular activation. However, it is affected by diverse factors such as baseline noise or tissue motion artifacts, as well as the equipment used: electrode type and placement, signal amplification, and filtering techniques [25]. In the realm of signal processing, there are multiple machine learning algorithms and pattern recognition techniques capable of recognizing muscle activity and associating it with a specific gesture despite the variability of the measurement [26].

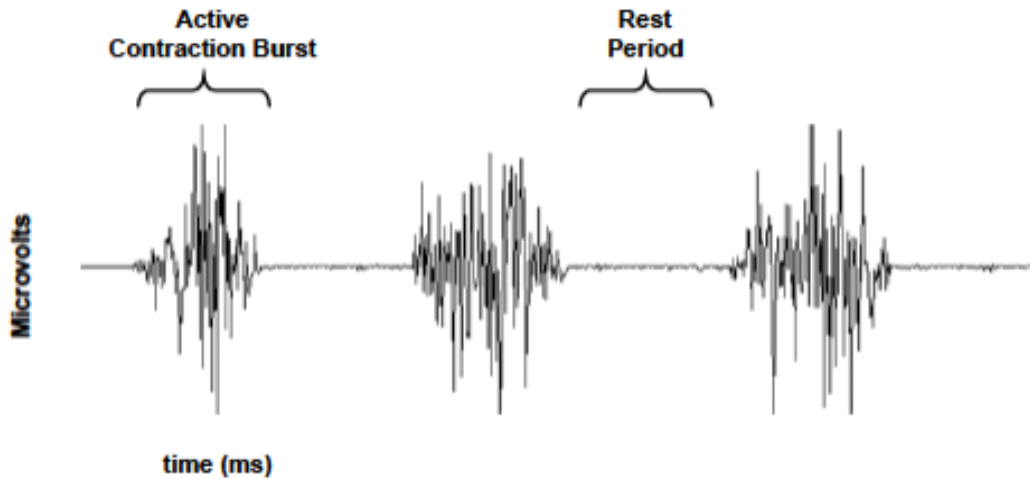


Figure 2.5: EMG signal with three contractions over time

2.3 State of the art

2.3.1 Eye tracking

Eye tracking is a field focused on measuring and analyzing the intricate movements of eyes. Several techniques are used to capture the movement of the eyes. Some commonly used methods are:

- Camera-based Eye Tracking [27]: this is one of the most common methods used today. It involves using specialized cameras to record the movement of the eyes. Algorithms then analyze the video footage to determine the position of the eyes over time.
- Infrared Eye Tracking [28]: this works by emitting infrared light toward the eyes and measuring the reflection using sensors. This method is advantageous in low-light conditions.
- EOG [22]: this technique involves placing electrodes around the eyes to measure the electrical potential difference between the cornea and the retina. By tracking these electrical signals, eye movements can be detected.

Each of these techniques has its advantages and disadvantages, and the choice of method depends on factors such as the required level of accuracy, the intended application, and whether the tracking needs to be done in a laboratory setting or real-world environments.

In this work, EOG has been selected due to its cost-effectiveness and relatively more straightforward setup than other methods.

2.3.2 Smart Glasses

Capture Images

Smart glasses equipped with cameras that fit directly into the frame offer users a seamless and hands-free way to capture images and videos from their perspective.

This provides convenience, allowing users to capture moments without needing to hold a camera [8], and opens up new possibilities for diverse applications. For example, smart glasses have been widely employed for real-time object detection and other computations on edge, thanks to energy-efficient onboard neural network processing [29]. This means that the glasses can recognize objects, faces, or text directly from the captured images without relying on cloud-based processing, ensuring faster response times. These capabilities demonstrate the potential of smart glasses as wearable and society-accepted devices, not just passive recording devices.

EOG

In recent advancements, this technique has led to the creation of wearable devices, such as eyeglasses equipped with dry electrodes. This represents a wearable and non-invasive solution, contrasting with traditional methodologies. JINS MEME glasses [30] represent the primary solution in this domain, equipped with three dry electrodes arranged in the frontal part of the glasses frame: one in the bridge and two in the nose pads. These glasses have undergone testing in multiple studies, confirming the reliability of this electrode placement.

2.4 BioGAP

BioGAP [31] is a novel, compact, and modular biosignal acquisition and processing platform for ExG biosignals. BioGAP integrates:

- nRF52811 System on Chip (SoC) (Nordic Semiconductor) for Bluetooth Low Energy (BLE) connectivity and measurement control;
- GAP9 Parallel Ultra Low Power (PULP) processor (GreenWaves Technologies) for efficient multiprecision data processing (as required for advanced ML and DSP);
- Inertial Measurement Unit (IMU) for movement sensing and a dedicated power management circuit with a rechargeable battery;

2 Background

- a specialized shield, including an analog front-end (AFE) (ADS1298, Texas Instruments) with eight channels, is employed for biopotential measurements. The shield interfaces to the electrodes (both active and passive electrodes are supported) and is configured as detailed in Table 2.1;
- a flexibly placeable photoplethysmograph (PPG) PCB.

Moreover, it can work in different modes: streaming, computation on the edge, and sleep mode. In this work, BioGAP is used in the streaming one, and as the name suggests, the acquired ExG data is continuously streamed in real-time to the GUI (Graphical User Interface).



Figure 2.6: BioGAP platform

Parameter	Value
Output data rate	500 SPS
-3dB bandwidth	131 Hz
PGA gain	6
Resolution	24 bit

Table 2.1: Measurement configuration of BioGAP AFE

2.4.1 GUI

A Java GUI on the PC is used as an interface between the user and BioGAP. It is able to switch between the different modes and send commands. Moreover, in the streaming mode case, it can also:

- start and stop measurements;
- display measurements in real-time;
- save parameters (like sampling frequency) and data from each channel into a log file in binary format that can be used for further processing.

The communication between BioGAP and the GUI is established using the Bluetooth low energy (BLE) protocol. A small USB dongle is used as BLE receiver. It's also possible to create a TCP connection with external devices to set the trigger vector of the acquisition. This allows for a specific trigger value to be assigned to each sample of the acquired signal, enabling the identification of particular portions of the signal for specific purposes.

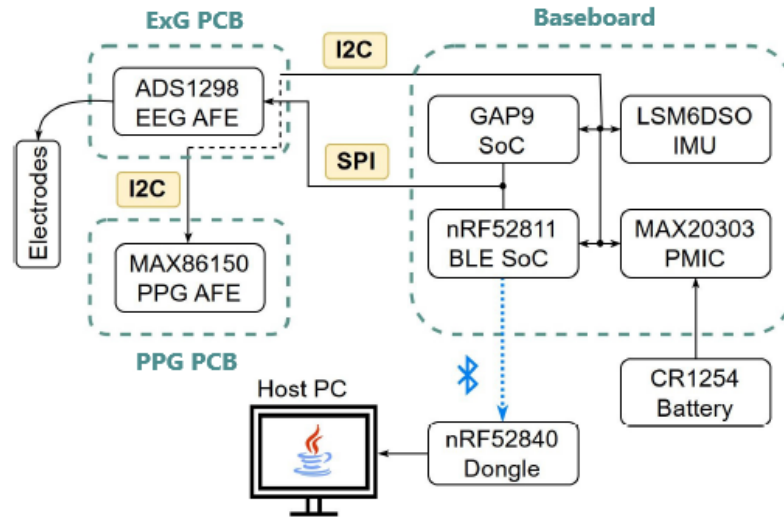


Figure 2.7: BioGAP System Diagram

2.5 Collaboration with Dätwyler

The Integrated Systems Laboratory (ETH Zürich) collaborates with Dätwyler Schweiz AG to develop state-of-the-art biosignal acquisition platforms. Specifically, Dätwyler specializes in the fabrication of soft electrodes for biopotential acquisitions, and the SoftPulse electrode family proved to be the perfect match for PULP’s biosignal systems. In addition, as part of the ETH Future Computing Laboratory (EFCL) framework, the PULP group and Dätwyler joined forces to develop next-generation EEG and EOG smart glasses based on the SoftPulse electrode technology and the BioGAP data acquisition platform. In this context, the envisioned glasses feature:

- the BioGAP acquisition platform;
- a non-stigmatizing design of the glasses frame guaranteeing proper contact of the EEG electrodes to the scalp;
- flex-rigid electronics embedded in the glasses frame for buffering the electrodes’ signal and interfacing to BioGAP;
- multiple EEG channels on the temporal area and behind the ear;
- two connectors for inconspicuous interfacing to in-ear EEG electrodes;
- three EOG electrodes on the frontal part of the frame.

2 Background



Figure 2.8: 3D-prototype with EEG and EOG dry electrodes presented to Datwyler

The EOG Prototype

Before progressing with the smart glasses project in collaboration with Datwyler, a prototype for EOG acquisitions through glasses (without lenses) was developed. This intermediate step was necessary to find out if the dry electrodes employed and their placement in the glasses frame could achieve an EOG signal quality and an eye-movements classification accuracy comparable to the traditional setup.

This part of the project was conducted entirely at ETH Zürich.

3.1 Signal Acquisition

3.1.1 Electrodes

Three dry active electrodes are positioned on the glasses' frame: one on the bridge and one on the nose pads, as illustrated in Figure 3.1. These active electrodes are sheet electrodes wired to a mini-PCB for the signal buffering stage. The sheet-electrode is a 2-layer conductive pad composed of polymer and silver/silver chloride (Ag/AgCl) in direct contact with the skin.

Moreover, these sheet electrodes are derived from the SoftPulse electrode family from Datwyler, a prototype that is not yet commercially available. Part of the project involves evaluating the performance of this electrode prototype. The preference for active electrodes over passive ones stems from the necessity for a more robust signal conditioning, given the minimal amplitude of the signal. Additionally, the power consumption of the active electrode is not significant compared to the overall system.

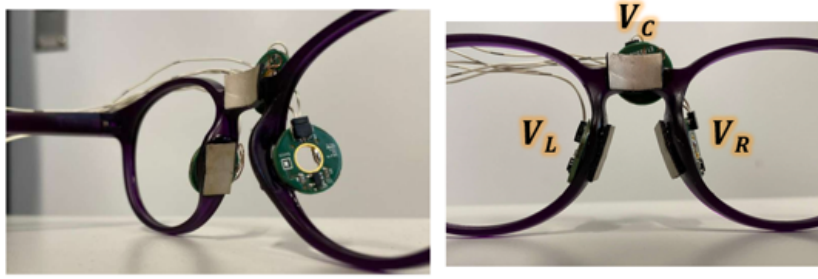


Figure 3.1: Electrodes placement in glasses prototype

3.1.2 Acquisition Setup

These electrodes are connected to three channels of BioGAP.

Denoting voltages acquired from the respective electrodes as V_L , V_R and V_C for the left, right, and central electrodes, respectively, the two signals for eye movements (vertical and horizontal) are computed differentially using the following formulas:

- Vertical signal:

$$V_V = V_C - (V_L + V_R)/2$$

- Horizontal signal:

$$V_H = V_R - V_L$$

The selection of the horizontal signal follows a convention. In contrast, for the vertical signal, we adopt the same approach outlined in [30], a practical solution, utilizing a single signal instead of employing two signals with the same/opposite orientation.

At the same time, to prove the validity of such a system, traditional EOG measurements are also taken. The most used setup is applied, involving four active gel electrodes placed in the face as reported in Chapter 2 and shown in Figure 3.2.

Denoting voltages acquired from the respective electrodes as V_U (up-left), V_D (down-left), V_L (left) and V_R (right), the two signals for eye movements are computed differentially using the following formulas:

- Vertical signal:

$$V_V = V_U - V_D$$

- Horizontal signal:

$$V_H = V_L - V_R$$

3 The EOG Prototype

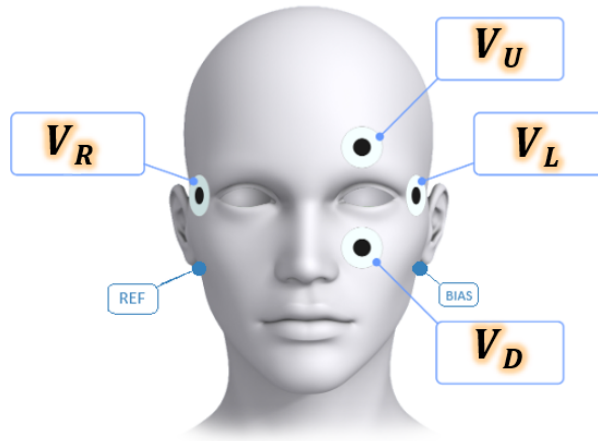


Figure 3.2: Traditional acquisition method

Notice that, for the evaluation of the horizontal signal, the opposite configuration is chosen. This is due only to the practical viewing of the signal: in the traditional method, electrodes are placed in the outer part of the eyes, while with glasses, they're on the inner side. Therefore, to compare the acquired signals, the opposite configuration is needed.

Finally, reference and bias connections are needed:

- Reference electrode: an active electrode attached to the skin through a gel-based electrode placed behind the right ear on the mastoid.
- Bias electrode: a simple passive electrode attached to the skin through a gel-based electrode behind the mastoid's left ear.



Figure 3.3: Active gel-electrode used for reference and traditional setup

3.1.3 Experimental Test

To create a dataset, an experiment was conducted with four subjects (two males and two females, mean age 24) in a room equipped with a projector.

The test duration is approximately three minutes, and it includes three repetitions of different tasks, among which maintaining a steady gaze and performing basic eye movements: looking

3 The EOG Prototype



Figure 3.4: Hardware Setup for EOG acquisition

up, down, left, right and their combination (up-right, down-right, up-left, and down-left) in specific points of the projector, as shown in Figure 3.5.

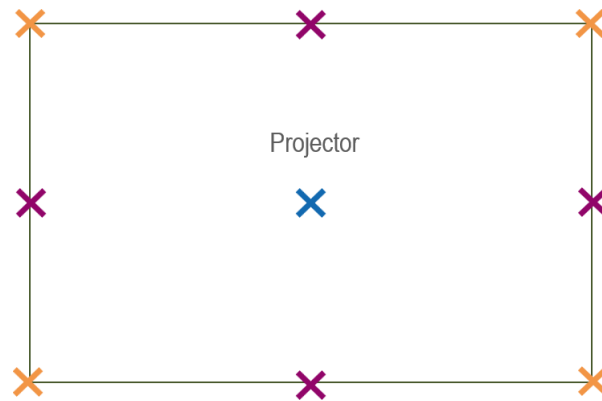


Figure 3.5: Points to look in the projector

During the test, the subject sits down fixed in front of the projector and is asked to move only the eyes, not the head. Moreover, written and image format instructions regarding the task to perform are given to the subject via the projector. Between each task, there's a rest phase in which the users can blink their eyes (to prevent fatigue during the session).

All of this is implemented using a Python script employing the Psychopy package, specifically designed for developing experiments in behavioral science. In addition, a trigger value is set during every task, and through a TCP connection, it is relayed to BioGAP GUI (which attaches the trigger values to the raw data), serving as the gold standard of the test.

The test begins after executing the Python script, initiating the EOG acquisition, and establishing a TCP connection in the BioGAP GUI. BioGAP captures, amplifies, and converts EOG signals and sends them to the GUI via Bluetooth. The raw data are then saved in a binary file provided by the GUI at the end of the acquisition. Data includes waveforms from each channel, time stamps, and a trigger vector, as reported in Chapter 2

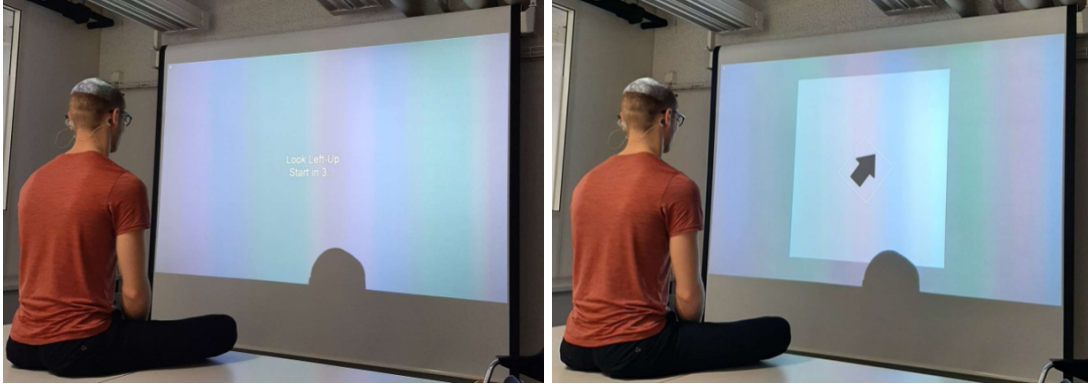


Figure 3.6: Experimental test, using Psychopy, at the projector

3.2 Data Processing

After completion of the experiment, hence the EOG acquisition, the binary file is imported and converted into Jupyter Notebook using a Python Script. All data are saved in a specific Python class, containing attributes related to data from each channel, trigger vector, time stamp, and others. Then, the vertical and horizontal signals of both methods are computed as specified before in Chapter 2.

3.2.1 Signal Filtering

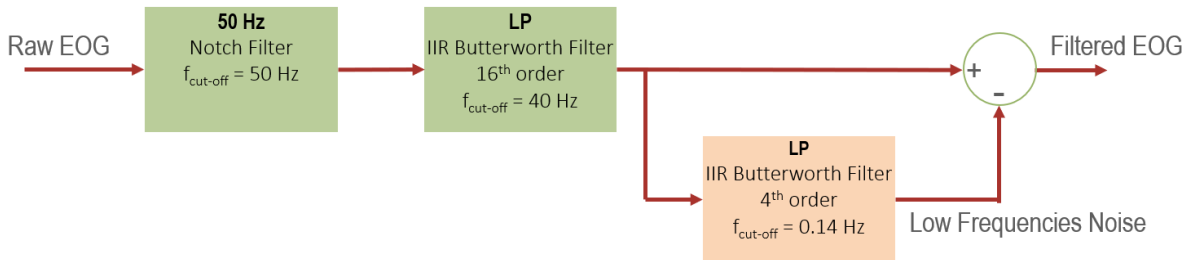


Figure 3.7: Filtering process for EOG

Three steps of filtering are applied to the raw EOG signals, both horizontal and vertical:

- Notch Filter to remove the 50 Hz network interference.
- Low Pass IIR Butterworth Filter of the 16th order with a cut-off frequency of 40 Hz, to remove high frequencies noise according to what was introduced in Chapter 2.
- Low Pass IIR Butterworth Filter of the 4th order with a cut-off frequency of 0.14 Hz to isolate the low frequencies noise and subtract it from the previous filtered signal.

The filter type and parameters have been chosen according to the linear phase characteristic in the passband without any ripples in the filter's amplitude response and looking at the filtered EOG signal.

All filters are implemented using Python's signal processing toolbox (`scipy.signal`).

3.2.2 Signal Segmentation and Feature Extraction

To proceed with the feature extraction needed for further classification, the EOG signals, both vertical and horizontal, are divided into various single instructions. This is done with a Python script using the trigger vector provided by BioGAP, thanks to the synchronization with the Psychopy test during the acquisition.

Figure 3.8 shows the importance of the trigger, which enables the identification of portions of the signal for the execution of a specific task.

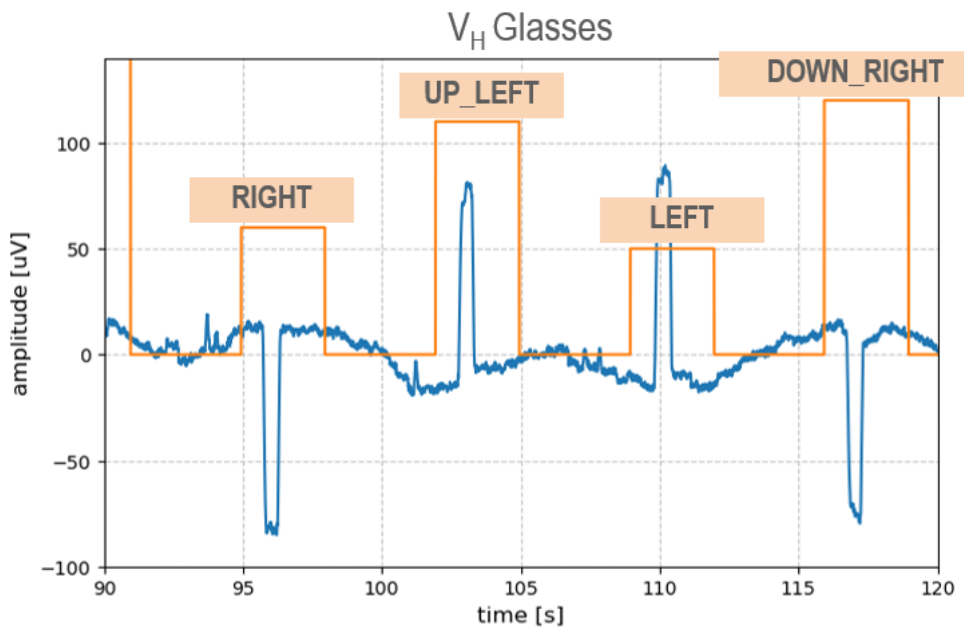


Figure 3.8: Horizontal EOG from glasses and Trigger Vector

Then, time-domain and statistical features are collected for each portion from the vertical and horizontal signals and their first derivatives to create a dataset. Features are mean, standard deviation, maximum, minimum, peak-peak, RMS, Shape Factor, and finally, the Cross-Correlation, which, differently from others, is computed between the horizontal and vertical, both the signals and derivatives.

3.2.3 Machine Learning

Classical machine learning algorithms have been applied to evaluate and compare the performance between the traditional acquisition method and the new prototype based on glasses and also to explore the best processing procedure in this work.

Classification

The first stage of the proposed method is a supervised classifier. The classifier used is K-Nearest Neighbors (KNN) with a number of Neighbors K equal to two.

Before the training phase, the dataset is scaled using Standard Scaler provided by Python, which scales the data to zero mean and unit variance. The training set corresponds to 60% of the data, and the remaining 40% for the testing set. Finally, the following performance metrics were computed: accuracy, F1-score, Recall, and Precision utilizing the K-Fold Cross-Validation method with the number of folds (K) set to three.

Camera Module

This chapter contains the study of the camera module and the CNN employed for binary image classification in HGR. This part of the project has been conducted both at ETH Zürich and at the University of Bologna. This module will be synchronized with the BioGAP GUI to acquire both ExG signals and images.

4.1 GAP9 Shield V2

The GAP9 Shield (version 2), a platform developed by ETH and the University of Bologna, has been adopted to acquire images.



Figure 4.1: GAP9 Wi-Fi Shield

Among the multiple characteristics, the shield integrates:

- The GAP9 SoC (a RISC-V PULP-based) from GreenWaves technologies is the main processing unit with on-chip memories of 128 kB (L1), 1536 kB (L2), and a hardware accelerator (NE16). GAP9 operates at a maximum frequency of 370MHz, controls the attached camera, and manages image processing tasks.

4 Camera Module

- An OmniVision OV5647 camera 5MP-RGB camera, with a resolution set to 320x240 and frame rate about 6 FPS.
- A NINA WiFi module, which includes an ESP32, is used to send the images via Wi-Fi to the computer.

Within the shield, the communication between GAP9 and NINA Modules happens via Serial Peripheral Interface (SPI), where they act as master and slave, respectively. Then, the communication between the camera OV5647 and GAP9 happens via Camera Serial Interface 2 (CSI2), a classical protocol used for efficient transmission of images. Finally, the connection between GAP9 Shield and the host PC happens via Wi-Fi using the TCP protocol, as shown in Figure 4.2.

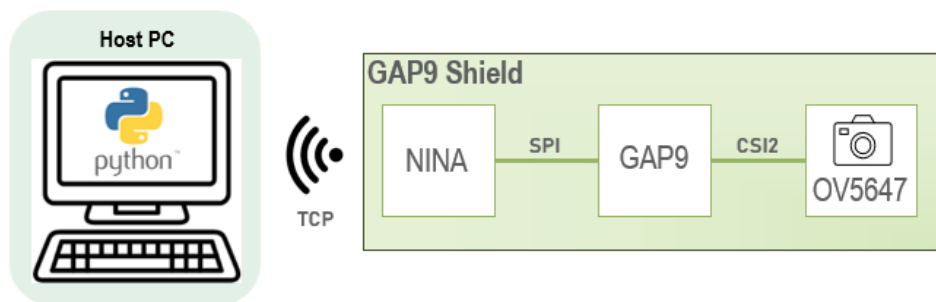


Figure 4.2: GAP9 Shield Block Diagram

4.1.1 Firmware Description

First, the setup for GAP9 Software Development Kit (SDK) and the firmware, needed to configure the board and manage data, have been installed in a virtual machine (Linux-64 bit and Ubuntu 22.04). The firmware for the GAP9 Shield encompasses several functionalities. It deals with the configuration of camera settings and coordinates the transmission via SPI between GAP9 and NINA of the Wi-Fi parameters, including network name, password, and IP port of the computer. Moreover, the firmware incorporates image compression in JPEG format, reducing file sizes for efficient data handling. It further deals with the transmission of data packets, containing images, still via the SPI interface from GAP9 to NINA. Finally, it handles the transmission of these data packets from NINA Wi-Fi module to the host PC using the TCP protocol, completing the data management and communication ecosystem.

4.1.2 Stream and save of images

As a starting point, a Python script for streaming and saving images on the host PC has been studied and tested. This script serves as the central manager for all operations of the Shield. The


```

socketShield.accept()
while(True):
    data = socketShield.receive(data_packet_size)
    header_frame = data[0]
    if(header_frame):
        frame = extract_image(data)
        save_in_folder(frame)
        stream_image(frame)

```

Listing 4.1: Python snippet code to stream and save images using GAP9 Shield

first step involves the establishment of a TCP connection between the host PC and the Shield. Once this connection is established, the camera begins acquiring images and sends them to GAP9 (via CSI2). The frame size is 320x240 bytes, resulting in a total of 76800 bytes per frame. GAP9 compresses the image in JPEG format, obtaining a dimension of 16600 bytes. Then, it sends these data through SPI to the Wi-Fi module (NINA), which finally transmits the data packet via Wi-Fi using the TCP protocol. Subsequently, within a continuous loop, the script receives these data packets containing the image frames. It checks for the presence of the frame header within each received packet. If the header is detected, the script processes the data to extract the individual frame. Finally, the extracted frame is saved in a designated folder on the host PC and streamed in real-time. This seamless process ensures immediate access to the images as they are captured. A pseudo-code representing this process is reported in Listing 4.1, while Figure 4.3 shows an example of streaming frames from GAP9 Shield.

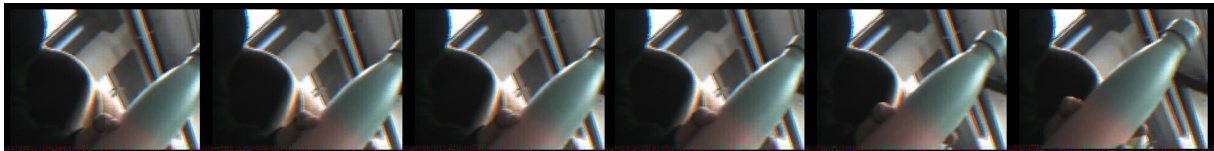


Figure 4.3: Image-frames streams in the computer

4.2 Classification: MobileNetV2

Convolutional Neural Networks (CNNs) are mainly used for image classification [32].

A CNN typically consists of three primary layers: the convolution, pooling, and fully connected (FC) layers. Convolution layers are the main building blocks of CNN models. Different CNN architectures exist, and they differ in the number of layers and the size of trainable parameters.

The MobileNetV2 CNN was chosen as image classifier in this thesis. The choice was due to the fact that it is a small network and it reduces the number of operations and the amount of memory used for computing while maintaining a high predictive accuracy. Therefore, it is suitable for

mobile devices or any devices with low computational power. It is the second updated version of MobileNet, developed by Google and trained on the ImageNet dataset [33].

In Figure 4.4 is reported the structure of the network. The main difference with respect to the previous versions is the Inverted Residual and Linear Bottleneck Block (also known as, Bottleneck Residual Block) in the Convolution Layer. This block will be explored in detail in the next section.

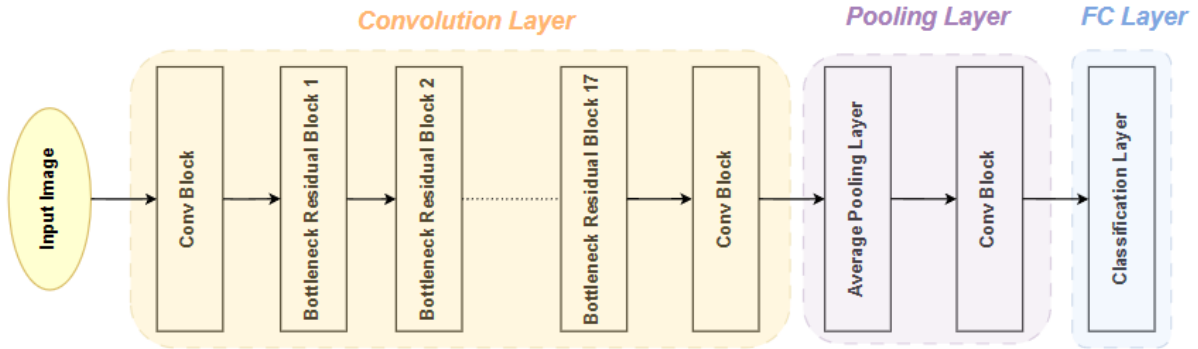


Figure 4.4: MobileNet V2 Structure

4.2.1 Inverted Residual and Linear Bottleneck Block Structure

The main building block of the Convolutional Layer of the MobileNetV2 is the Inverted Residual and Linear Bottleneck Block. It has 3 convolutional layers, and it follows a narrow-wide-narrow approach in terms of the amount of data, as reported in Figure 4.5.

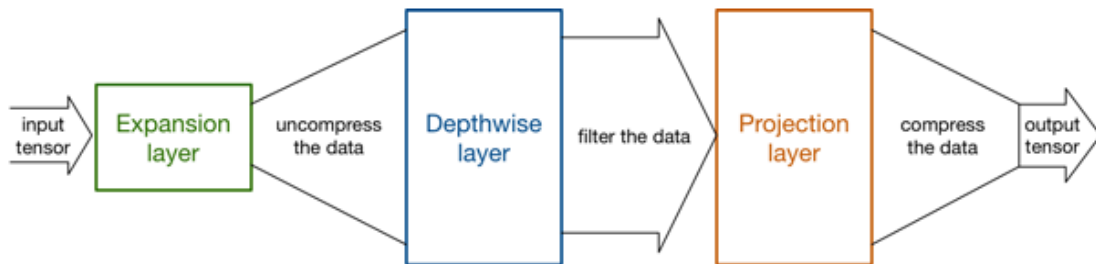


Figure 4.5: Data flow in Bottleneck Residual Block [5]

Figure 4.6 reports in more detail the structure of each layer:

- Expansion Layer: newly incorporated in this network version, it is a 1×1 convolution layer with non-linearity. Functionally, it expands the input dimensions, a crucial step before entering the depthwise convolution phase. Hence, it acts as a decompressor, with an output dimension bigger than the input one.

4 Camera Module

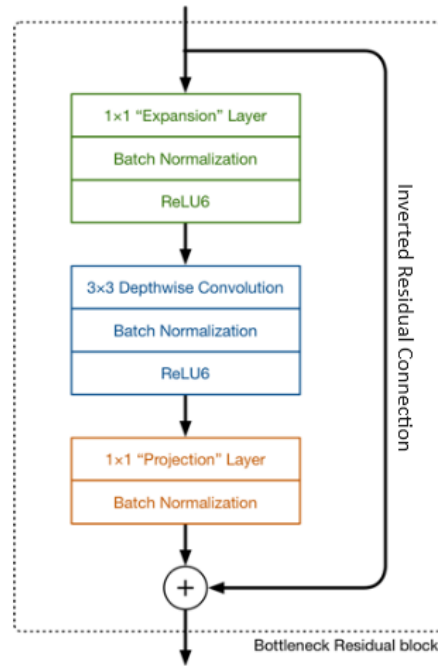


Figure 4.6: Structure Bottleneck Residual Block MobileNet V2 [5]

- Depthwise Layer: operating with a depthwise 3x3 convolution, this layer serves as a filter for the input data.
- Projection Layer: this layer compresses the data to make it small again. It is another 1×1 convolution layer without non-linearity. Unlike the Expansion Layer, it projects data with a high number of dimensions in input into a tensor with a smaller dimension. This kind of layer is also called a bottleneck layer because it reduces the amount of data that flows through the network. This is where the “bottleneck residual block” gets its name from.
- Inverted Residual Connection: it concatenates the entire block’s input and output, allowing the gradient to flow more easily during training and helps mitigate the vanishing gradient problem. The attribute "inverted" is because the connection exists between narrow parts of the network, which is the opposite of how an original residual connection works.

Moreover, as is common in modern architecture, each layer has batch normalization, and the activation function is Rectified Linear Unit (ReLU6). However, it’s important to note that the projection layer is an exception where no activation function is applied to its output. This approach is adopted because the projection layer generates low-dimensional data, and applying a non-linearity at this stage can lead to the loss of valuable information, negatively impacting performance.

4.2.2 Transfer Learning via Fine-Tuning

One advantage of CNNs is their ability to be reused for related problems instead of necessitating training from scratch. Transfer learning is a technique enabling the utilization of pre-trained models with slight modifications, thereby reducing training time. In this work, transfer learning was implemented via fine-tuning: the pre-trained model is fine-tuned on a smaller and more specific dataset. The implementation was done in PyTorch, an open-source Python library for developing deep learning models. A MobileNetV2 model pre-trained on the ImageNet dataset was loaded for this purpose.

During fine-tuning, layers may be selectively frozen, indicating that the layer should not undergo further training. The classification layer is the only one to be retrained. Additionally, it was modified to classify into two classes (binary classification problem), as the original network provides 1000 output classes.

A dataset collected with the Gap9Shield was employed to train the fine-tuned model. The dataset was split into training and testing sets in an 80 to 20 proportion. Since MobileNetV2 requires input images of size 224x224 pixels while ours were 320x240 pixels, resizing was necessary.

Prior to model training, two important parameters needed to be defined:

- **Loss function:** this function is crucial for model improvement, estimating the loss by comparing predicted output with actual output and calculating the loss value. Binary Cross Entropy (BCE) WithLogitsLoss function was employed in our model, which is commonly used for binary classification problems where the model output is a probability value between 0 and 1.
- **Optimizer:** an optimizer updates the network weights based on the output of the loss function. The Adam optimizer with a learning rate of $1e-5$ was utilized.

The training process commenced once the model was fine-tuned and the dataset prepared. During training, an input image was processed, a prediction was made, and a loss function was used to estimate the error. Subsequently, the network weights were updated to minimize the error and improve prediction accuracy.

Once the network is trained, resulting in good performance in terms of low loss and high accuracy, the model is saved and subsequently utilized for the testing phase.

EMG and Image System

This chapter outlines the procedure for testing and assessing the synchronized connection between the GAP9 Shield and the BioGAP GUI. The objective is to develop a system capable of simultaneously acquiring both images and ExG data. Specifically, EMG is being considered for this purpose. The synchronization is done thanks to the trigger vector.

This part of the project has been entirely conducted at the University of Bologna.

5.1 Connection with BioGAP

As introduced in Chapter 2, BioGAP is interfaced via BLE with its GUI, and it's possible to use the TCP protocol to set up communication with other devices. Therefore, to establish a connection between BioGAP and GAP9 Shield, the host PC is used as a central communication handler and creates a socket for BioGAP. For such purpose, a Python script similar to the one introduced in Chapter 4 has been developed, and its workflow is represented by the pseudo-code reported in Listing 5.1. Firstly, TCP connections are established between the host PC and both BioGAP, managed by the GUI, and the Shield. After that, as explained in Chapter 4, the shield starts acquiring images and sends them via Wi-Fi to the host PC. Then, within a while-loop, the script receives the data packets containing images and writes them in a binary file for further offline elaboration. After writing, it checks for the presence of the packet's header, and if detected, a trigger is sent to the BioGAP GUI, as this means that a frame is effectively acquired. Therefore, the trigger is used to keep track of each acquired image: each frame is identified by the corresponding value of the trigger (the first frame, with trigger = 1, the second one with trigger = 2, etc). Triggers are also used to verify the synchronization during offline processing, as reported in the next section.

```
trigger_frame = 0
file_images = open("images_file.bin", "wb")
socketBioGAP.accept()
socketShield.accept()
while(True):
    data = socket shield.receive(data_packet_size)
    file_images.write(data)
    header_frame = data[0]
    if(header_frame):
        trigger_frame += 1
        socketBioGAP.send(trigger_frame)
```

Listing 5.1: Python snippet code for synchronization between the Shield and the GUI

5.2 Acquisition Phase

Simultaneous acquisition of both EMG and images is performed to test the implemented code. The EMG acquisition involves a simple setup using gel-based electrodes (same employed in the EOG module, reported in Figure 3.1): a single channel of BioGAP using an active electrode put in the center of the forearm, then an active electrode in the wrist as reference and the bias electrode on the elbow. The subject under test performs the opening and closing of the hand in front of the camera five times. These kinds of movement are chosen because of their clear differences in the EMG signal and images.

5.3 Offline Processing

A Python script is developed for the following post-acquisition steps:

- **Trigger Reconstruction:** as the trigger is transmitted to the GUI as a character, it may saturate due to the number of images acquired. Consequently, when the trigger reaches a maximum representable value, it is reset to zero. According to this, a simple function is implemented to re-adjust the trigger in a more continuous and linear behavior, see Figure 5.1, and also to have the correct correspondence with frames.
- **Images Saving:** all acquired images are reconstructed from the binary file written in acquisition time, saved in the host PC, and named with the corresponding number of images.
- **EMG Data Conversion:** as usual, the signal acquired by BioGAP is saved in a binary file during acquisition time and then converted for visualization.

5 EMG and Image System

- **EMG Filtering:** the EMG signals are filtered using a Band-Pass IIR filter of the 4th order, with 10 and 500 Hz cut-off frequencies. This is done using Python's signal processing toolbox (`scipy.signal`).
- **Verifying Synchronization:** from the saved images, the first and last frames with a closed hand are manually identified; using these frame numbers, the corresponding trigger values are determined. Therefore, these values are plotted with the EMG and compared in time with the increasing intensity of the EMG, which corresponds to the hand's closing.

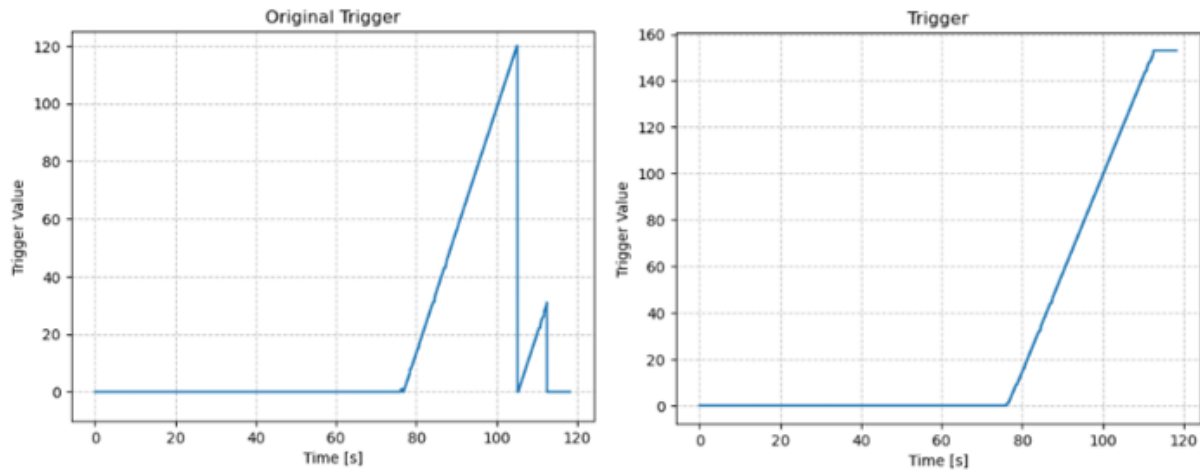


Figure 5.1: Trigger before and after its reconstruction

The trigger vector plays a crucial role in synchronizing data captured from different sources. This synchronization enables the training of new algorithms to explore innovative sensor fusion techniques.

Results

This chapter presents the results obtained in this thesis. Specifically, they are categorized into different sections corresponding to the respective modules. The initial section discusses the outcomes related to the EOG module, followed by the section on image classification, and concludes with the section on the synchronized acquisition of images and EMG.

6.1 EOG Results

This section examines and compares the results derived from the integrated EOG-acquisition system within the glasses prototype with those obtained through the traditional acquisition method (Chapter 3). The comparison primarily evaluates signal amplitude, shape, and classification performance.

6.1.1 Filter Stage

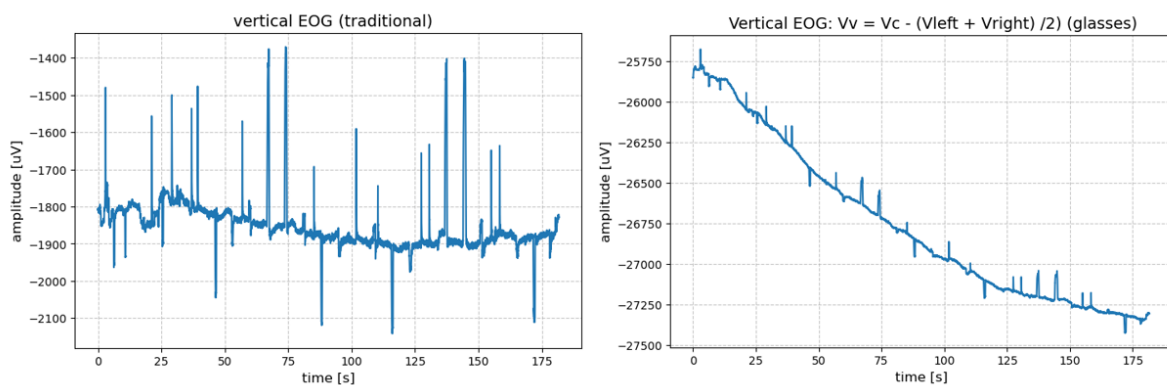


Figure 6.1: Baseline Drift in raw EOG with traditional and glasses method

6 Results

Figure 6.1, displays the raw vertical EOG acquired using traditional and glasses methods. In both cases, low-frequency noise is evident, manifesting as a drift effect. As expected, this effect is more pronounced in the glasses method. This is primarily attributed to the distinct nature of the electrodes: gel-based electrodes offer higher signal quality compared to dry electrodes, not only due to the presence of the gel but also because of better adherence to the skin. Consequently, the signal captured from the glasses method exhibits more noise.

To address and mitigate such noise, two potential solutions have been examined:

1. High Pass Chebychev type I Filter of the 4th order with a cut-off frequency of 0.34 Hz.
2. Low Pass IIR Butterworth Filter of the 4th order with a cut-off frequency of 0.14 Hz designed to isolate the low-frequency noise, which is then subtracted from the previous filtered signal, as reported in Figure 3.7.

Figures 6.2, 6.3 show the results of these two different approaches.

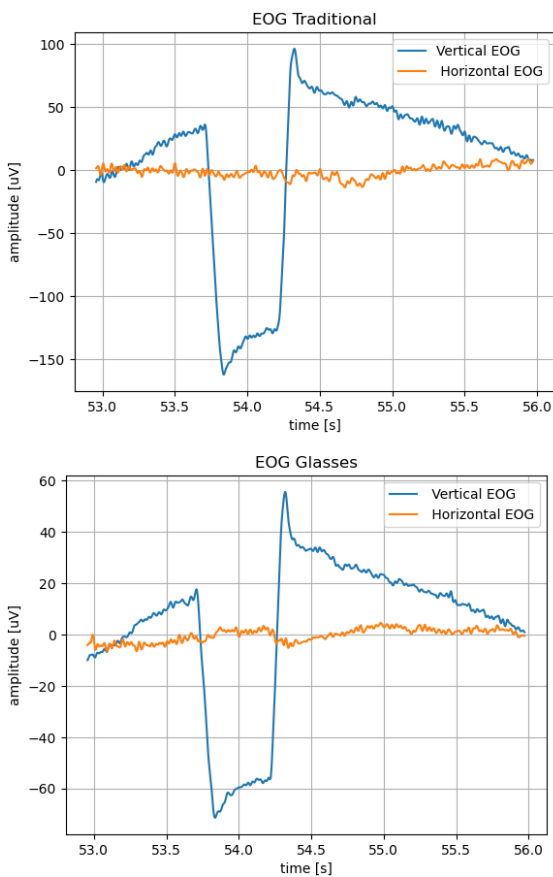


Figure 6.2: High Pass Filtered Look Down EOG

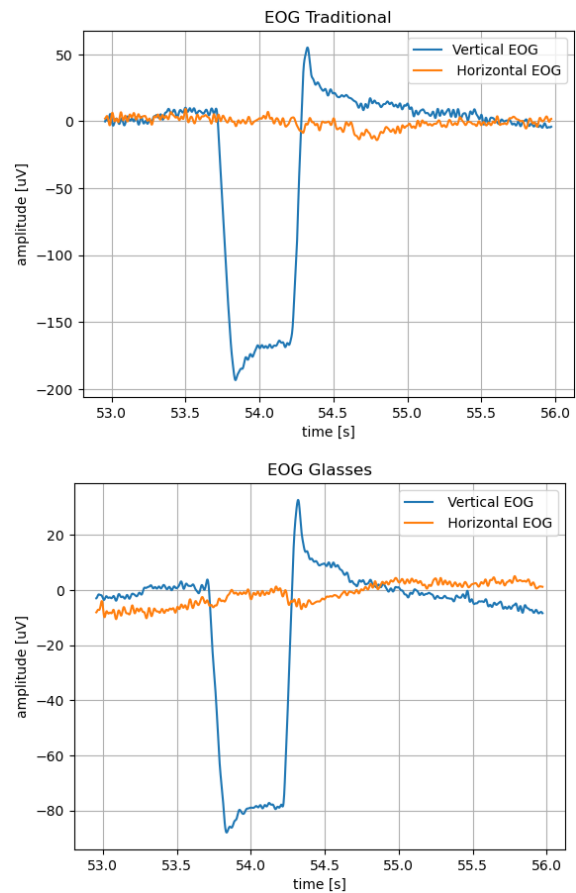


Figure 6.3: Low Frequencies Noise Removal Look Down EOG

The second solution appears more coherent based on the literature. Consequently, this approach has been adopted in the filtering process. Figure 6.4 illustrates the raw EOG signals, both vertical and horizontal, before and after the filtering process.

6 Results

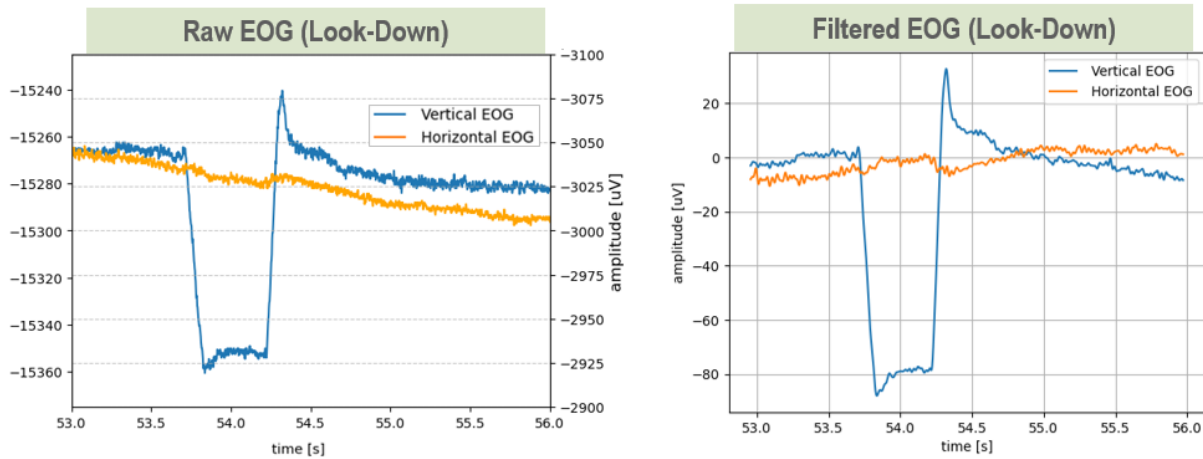


Figure 6.4: Comparison raw and filtered EOG from glasses

6.1.2 Shape Description

As introduced in Chapter 2, the EOG signal varies depending on the direction of the eye relative to a specific electrode. When the cornea approaches the electrode, the signal exhibits a positive trend, decreasing when the cornea moves in the opposite direction. The formulas used to compute both methods' vertical and horizontal signals are detailed in Chapter 3. Figure 6.5 displays the EOG signals captured during a 'looking down-left' movement with the glasses. The vertical signal corresponds to the contribution of vertical movement (looking down), exhibiting a negative trend as the cornea moves away from the central main electrode. Similarly, the horizontal signal is associated with horizontal movement (looking left), demonstrating a positive trend. This is because, during a leftward gaze, the contribution for the inner right electrode (the first element in the formula) is positive, resulting in a higher value for the signal.

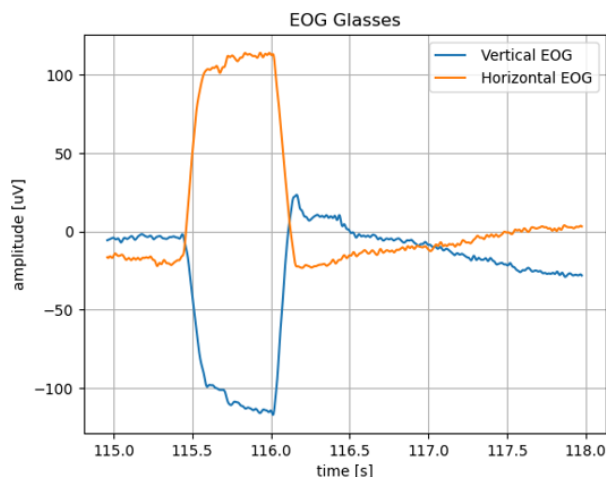


Figure 6.5: EOG Signals from glasses during Look Down-Left instruction

6.1.3 Amplitude Range

Table 6.1, 6.2 illustrate the amplitude range values of the EOG obtained through both traditional and glasses methods while executing various instructions. Notably, higher values are observed for horizontal movements compared to vertical ones when focusing only on the direction of movements. This observation aligns with the test setup, which utilized a projector (rectangular shape), resulting in a greater emphasis on horizontal shapes than vertical ones. This finding confirms that EOG accurately represents eye movement, with an amplitude proportional to the angle of movements.

Left	Right	Up	Down
300	300	150	150

Table 6.1: EOG amplitude range with traditional method

Left	Right	Up	Down
100	100	60	60

Table 6.2: EOG amplitude range with glasses method

When comparing the methods, it becomes evident that signals acquired through glasses exhibit smaller values. This discrepancy can be attributed to the different nature of the electrodes utilized in the two methods, as mentioned also in section 6.1.1. Another crucial factor is the difference in electrode-to-electrode distance, which is significantly higher in the traditional method compared to the glasses method, resulting in a larger signal amplitude. Despite these slight variations, it is noteworthy that both methods maintain the consistent shape of each eye movement.

6.1.4 Classification Performances

This section presents the performance results derived from a classification using a K-Nearest Neighbors (KNN) classifier with a specified parameter setting of $k = 2$. Figure 6.6 and 6.7 report the confusion matrix from the prediction of the classification and the correct label. Additionally, accuracy metrics are provided to assess the classifier's performance quantitatively: the accuracy is computed from the confusion matrix, obtaining 95-96% for the glasses prototype and 96-97% for the traditional setup.

6 Results

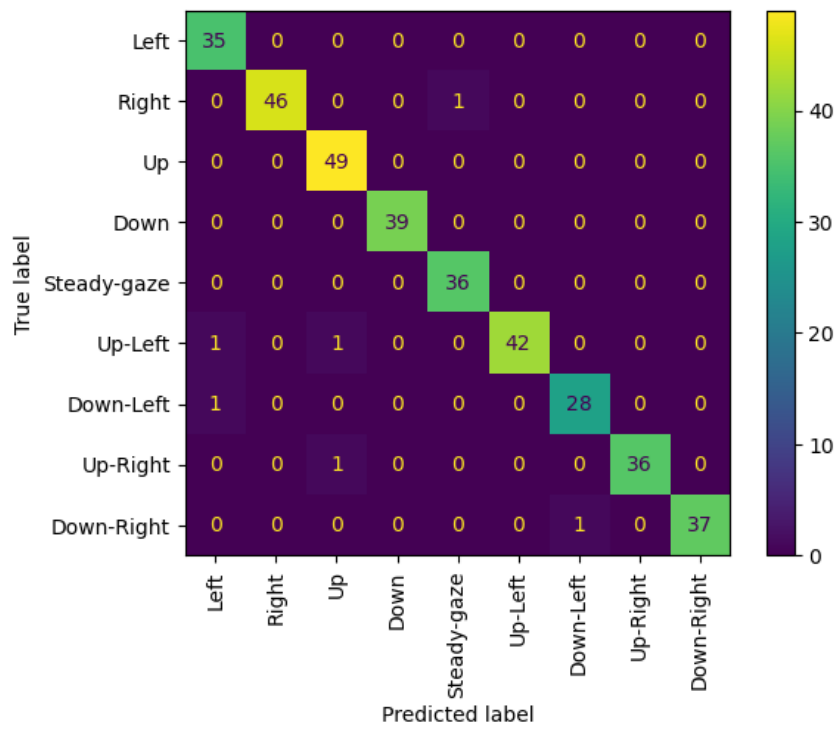


Figure 6.6: Confusion Matrix Traditional Method

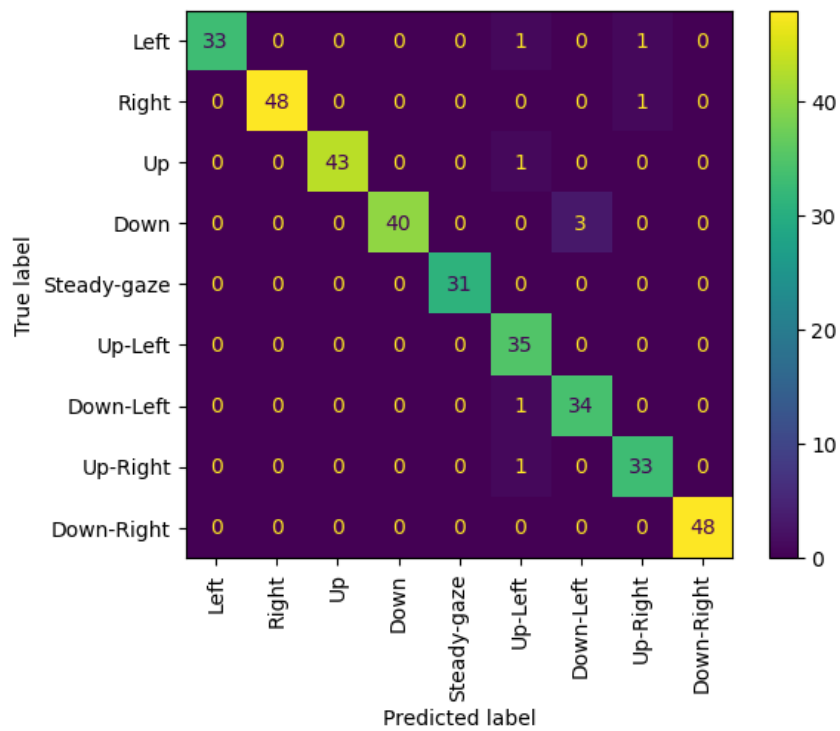


Figure 6.7: Confusion Matrix Glasses Method

6.1.5 K-Fold Cross-Validation Method

This section reports the accuracy, precision, recall, and F1-score performance for EOG acquisition through glasses, using the k-fold cross-validation method with several folds, $k = 3$.

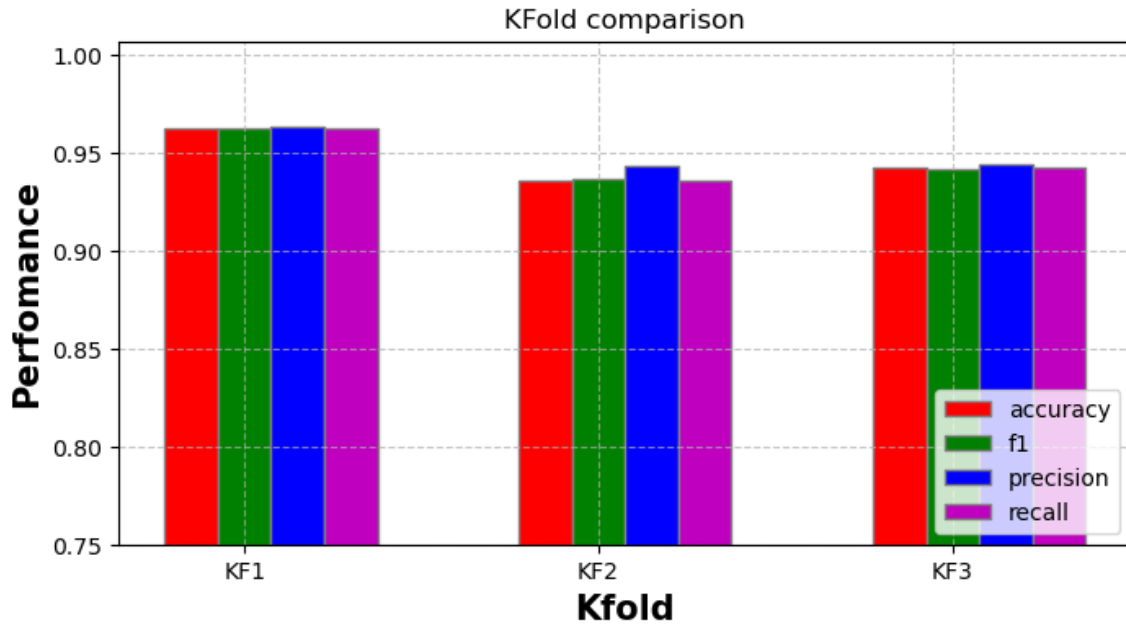


Figure 6.8: 3-Fold Cross-Validation in glasses method

6.2 Image Classification

This section presents the results obtained from Chapter 4, specifically focusing on the classification of images using a pre-trained CNN. The model utilized is the MobileNet V2, wherein fine-tuning is conducted by training only the classification layer. Furthermore, the model is modified to classify only two classes: open and closed hand. Consequently, the training dataset comprises images of open and closed hands captured with the GAP9 Shield.

Following the training process, the classification accuracy on the testing portion of the dataset reaches approximately 96-97%.

6.3 Synchronization Image and EMG Acquisition

This section reports considerations and results about the synchronization between the GAP9 Shield and the BioGAP GUI for simultaneous acquisition of images and EMG signals.

According to the offline procedure introduced in Chapter 5, Figure 6.9 illustrates how, from the saved images, the first and last frames featuring a closed hand are manually identified. Using these frame numbers, the corresponding trigger indexes are determined and then plotted with the EMG signal, facilitating the identification of these time frames in the signal. For instance, in the given example, the onset frame is identified as number 37, and the end frame as number 45. From the trigger vector of the EMG acquisition, the indices of the first sample with a trigger value of 37 and the last sample with a trigger value of 45 are extracted. Thus ensuring accurate capture of the hand closure period, from the onset to the end.

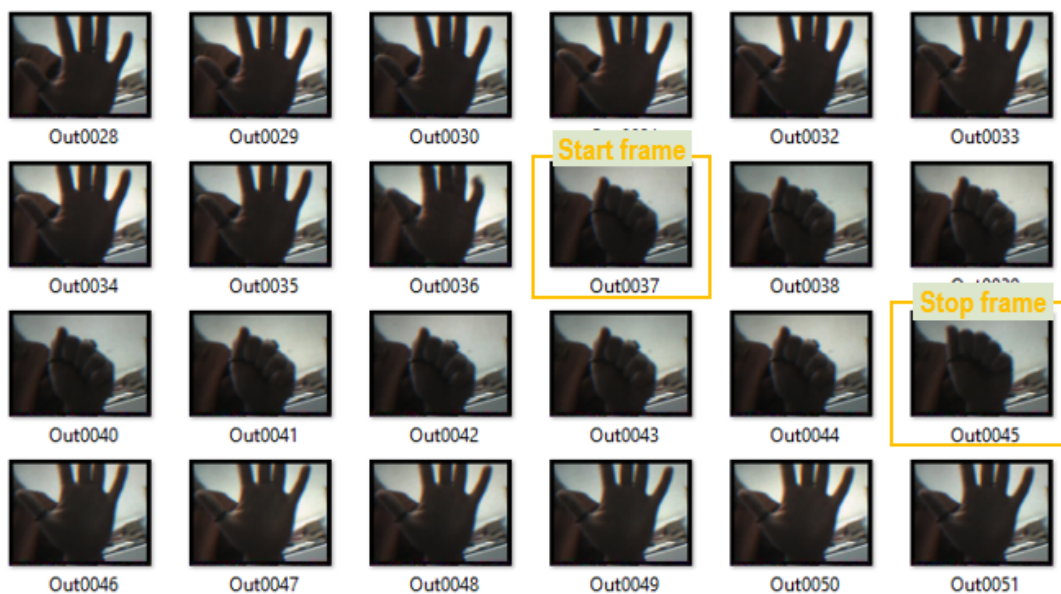


Figure 6.9: First and last frame with close hand to identify start and stop trigger value

6 Results

In Figure 6.10 is reported a synchronized example with trigger values and images to give a better understanding of the system.

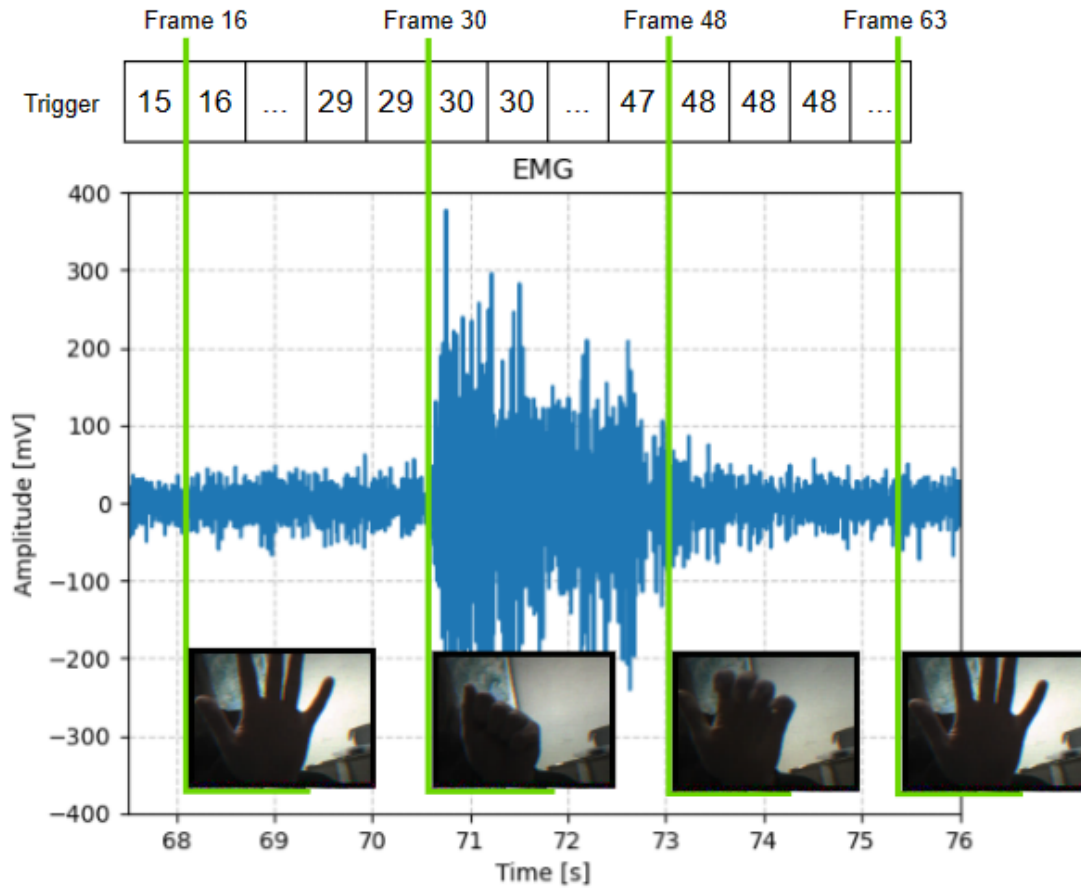


Figure 6.10: Synchronized acquisition of images and EMG

Comparing the index value and the contraction in the EMG signal (which corresponds to the hand's closing) over time, different results in terms of delay are observed. In some cases, like in Figure 6.11, the trigger of the first frame aligns perfectly with the onset of the contraction, and the same is true for the end before re-opening the hand. However, in other cases, like in Figure 6.12, a noticeable delay is present.

In this example, the EMG contraction is identified to have started from 76.75 s, while the starting time, given by the frame, is 76.624 s. Therefore, the resulting delay is given by:

$$delay = EMGcontraction - TimeStartFrame = 76.75 - 76.624 = 126ms$$

By calculating these delays, an average delay of approximately 120 ms is determined. This average value is considered acceptable, considering the camera's frame rate is about 6 FPS; thus, a delay exceeding 166 ms would misalign the corresponding frame. These different delays could be attributed to the quality of the Wi-Fi connection.

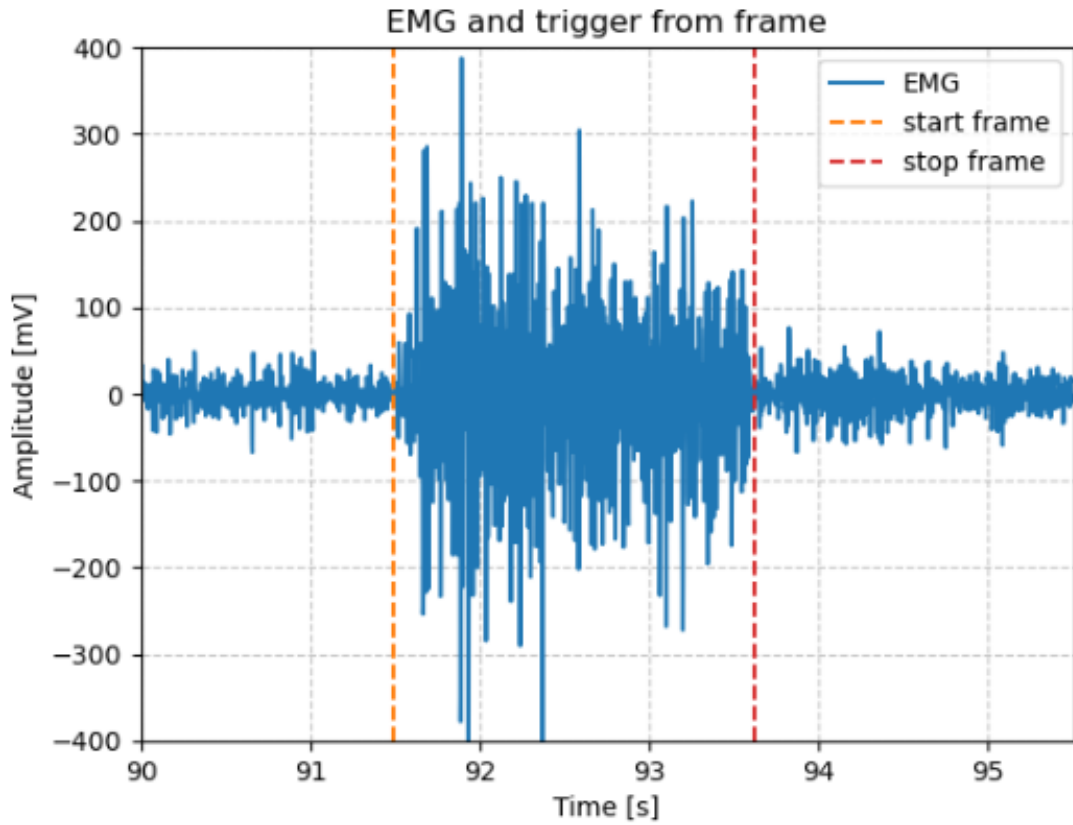


Figure 6.11: Synchronized acquisition of images and EMG

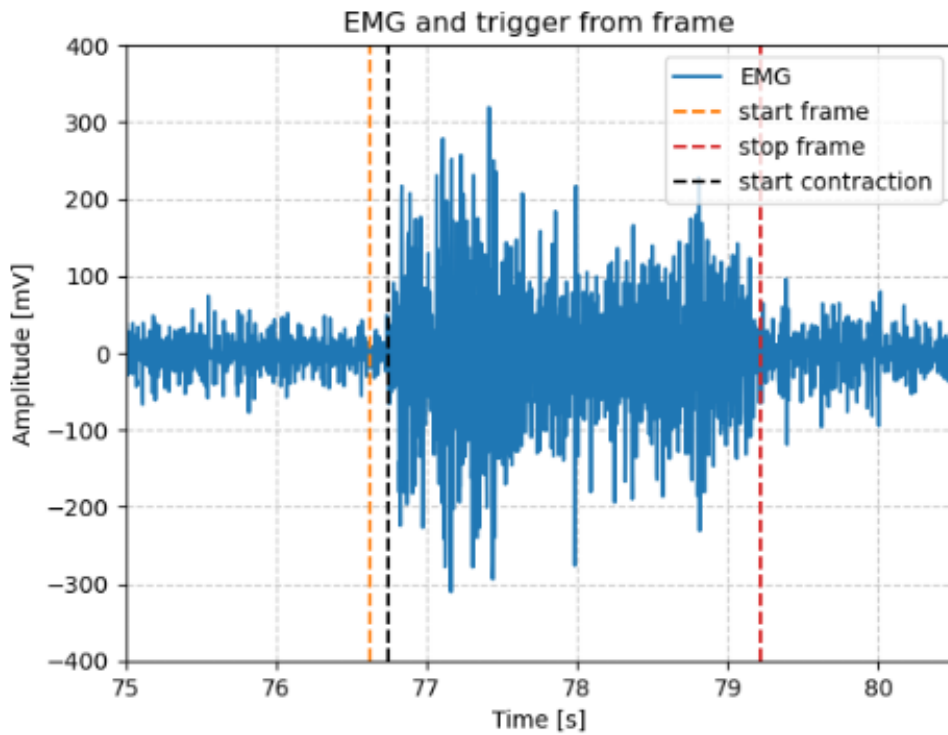


Figure 6.12: Acquisition of images and EMG with 126 ms of delay

Conclusion and Future Work

7.1 Conclusion

In this thesis, a smart glasses prototype designed for simultaneous ExG measurement and image capture was developed and evaluated.

The initial part, conducted at ETH Zürich, involved the realization of a pair of glasses equipped with dry sheet electrodes supplied by Dätwyler. These electrodes were interfaced with the BioGAP platform to acquire EOG signals, and their placement was done to optimize signal acquisition, ensuring seamless integration into the glasses frame for minimal intrusiveness.

An exploratory study using a 9-task eye movement classification scheme compared the performance of the glasses-based prototype with a conventional EOG setup. The results were promising, showing a comparable accuracy (95%) for the smart glasses compared to the traditional setup (96%) validating the effectiveness of the EOG prototype.

The subsequent part, conducted at the University of Bologna, includes the integration of the GAP9 Shield, used for capturing images, focused on integrating the GAP9 Shield for image capture with the BioGAP platform for general ExG acquisition. This integration ensured synchronized data acquisition, with an average delay of 120 ms, comparable to the frame period of 166 ms. Additionally, MobileNet V2, a CNN for embedded devices, was fine-tuned and utilized for binary image classification of hand gestures, achieving an accuracy of up to 97%.

The results obtained in this thesis build the basis for the development of authentic smart glasses capable of sensor fusion, including EOG, EMG, and images, based on the BioGAP platform and GAP9 Shield. This offers a promising tool for a more natural study and understanding of hand-eye coordination during hand gestures and grasp control mechanisms.

7.2 Future Work

For future work, it is suggested to explore the realization of a PCB integrating the BioGAP platform and GAP9 Shield, with a focus on seamless integration into the glass frame for both the board and the camera itself. In order to improve the system's autonomy, moving to onboard data processing holds significant progress, obtaining a real-time analysis and interpretation of data.

Additionally, there is a need for further exploration of EOG literature, presenting opportunities for additional investigation. This could involve the development of more efficient processing and preprocess algorithms or the exploration of regression tasks.

Finally, there is a potential to build a comprehensive dataset from diverse modalities, analyzing intricate hand gestures, including grasping different objects. Integration of data on eye movements, visual cues, and muscle movements could deepen the understanding of grasp control. This approach could significantly advance prosthesis control, leading to more natural and responsive prosthetics and improving human-machine interfaces overall.

Bibliography

- [1] S. N. Abbas and M. Abo-Zahhad, *Eye Blinking EOG Signals as Biometrics*. Springer International Publishing, Dec. 2016, p. 121–140. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-47301-7_5
- [2] EOG Traditional Recording. [Online]. Available: https://mitsar-eeg.com/2020/12/09/biosignals_eog_ecg_emg/
- [3] H. Robert, K. Kise, and O. Augereau, “Real-time wordometer demonstration using commercial eog glasses,” in *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, ser. UbiComp '17. ACM, Sep. 2017. [Online]. Available: <http://dx.doi.org/10.1145/3123024.3123183>
- [4] M. Orlandi, P. M. Rapa, M. Zanghieri, S. Frey, V. Kartsch, L. Benini, and S. Benatti, “An adaptive dynamic mixing model for semg real-time ica on an ultra-low power processor,” in *2023 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2023, pp. 1–5.
- [5] NN MobilenetV2. [Online]. Available: <https://machinethink.net/blog/mobilenet-v2>
- [6] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, “A review of wearable sensors and systems with application in rehabilitation,” *Journal of NeuroEngineering and Rehabilitation*, vol. 9, no. 1, Apr. 2012. [Online]. Available: <http://dx.doi.org/10.1186/1743-0003-9-21>
- [7] N. Kosmyna, C. Morris, T. Nguyen, S. Zepf, J. Hernandez, and P. Maes, “Attentivu: Designing eeg and eog compatible glasses for physiological sensing and feedback in the car,” in *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ser. AutomotiveUI '19. ACM, Sep. 2019. [Online]. Available: <http://dx.doi.org/10.1145/3342197.3344516>

Bibliography

- [8] Ray-Ban Meta Smart Glasses. [Online]. Available: <https://www.ray-ban.com/italy/ray-ban-meta-smart-glasses>
- [9] O. Danielsson, M. Holm, and A. Syberfeldt, “Augmented reality smart glasses in industrial assembly: Current status and future challenges,” *Journal of Industrial Information Integration*, vol. 20, p. 100175, Dec. 2020. [Online]. Available: <http://dx.doi.org/10.1016/j.jii.2020.100175>
- [10] E. Waisberg, J. Ong, M. Masalkhi, N. Zaman, P. Sarker, A. G. Lee, and A. Tavakkoli, “Meta smart glasses—large language models and the future for assistive glasses for individuals with vision impairments,” *Eye*, Dec. 2023. [Online]. Available: <http://dx.doi.org/10.1038/s41433-023-02842-z>
- [11] M. Marković, H. Karnal, B. Graimann, D. Farina, and S. Dosen, “Glimpse: Google glass interface for sensory feedback in myoelectric hand prostheses,” *Journal of neural engineering*, vol. 14, p. 036007, 03 2017.
- [12] X. S. Zheng, C. Foucault, P. Matos da Silva, S. Dasari, T. Yang, and S. Goose, “Eye-wearable technology for machine maintenance: Effects of display position and hands-free operation,” in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15. ACM, Apr. 2015. [Online]. Available: <http://dx.doi.org/10.1145/2702123.2702305>
- [13] Z. Huang, W. Li, and P. Hui, “Ubii: Towards seamless interaction between digital and physical worlds,” in *Proceedings of the 23rd ACM international conference on Multimedia*, ser. MM '15. ACM, Oct. 2015. [Online]. Available: <http://dx.doi.org/10.1145/2733373.2806266>
- [14] J. Zhang, B. Wang, C. Zhang, Y. Xiao, and M. Y. Wang, “An eeg/emg/eog-based multimodal human-machine interface to real-time control of a soft robot hand,” *Frontiers in Neurorobotics*, vol. 13, Mar. 2019. [Online]. Available: <http://dx.doi.org/10.3389/fnbot.2019.00007>
- [15] L. Chen, F. Wang, H. Deng, and K. Ji, “A survey on hand gesture recognition,” in *2013 International Conference on Computer Sciences and Applications*, 2013, pp. 313–316.
- [16] A. Clark and D. Moodley, “A system for a hand gesture-manipulated virtual reality environment,” in *Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists*, ser. SAICSIT '16. ACM, Sep. 2016. [Online]. Available: <http://dx.doi.org/10.1145/2987491.2987511>

Bibliography

- [17] R. Meattini, S. Benatti, U. Scarcia, L. Benini, and C. Melchiorri, “Experimental evaluation of a semg-based human-robot interface for human-like grasping tasks,” in *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2015, pp. 1030–1035.
- [18] M. R. Ahsan, M. Ibrahimy, and O. Khalifa, “Advances in electromyogram signal classification to improve the quality of life for the disabled and aged people,” *Journal of Computer Science*, vol. 6, pp. 706–715, 02 2010.
- [19] Y. Fang, K. Wang, J. Cheng, and H. Lu, “A real-time hand gesture recognition method,” in *2007 IEEE International Conference on Multimedia and Expo*, 2007, pp. 995–998.
- [20] M. Krey, *Wearable Device Technology in Healthcare—Exploring Constraining and Enabling Factors*. Springer Singapore, Dec. 2019, p. 1–13. [Online]. Available: http://dx.doi.org/10.1007/978-981-15-0637-6_1
- [21] W.-D. Chang, “Electrooculograms for human–computer interaction: A review,” *Sensors*, vol. 19, no. 12, p. 2690, Jun. 2019. [Online]. Available: <http://dx.doi.org/10.3390/s19122690>
- [22] G. de Melo and S. L. Vieira, *Electrooculography: A Proposed Methodology for Sensing Human Eye Movement*. Springer International Publishing, 2022, p. 763–769. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-70601-2_116
- [23] F. Negro, S. Muceli, A. M. Castronovo, A. Holobar, and D. Farina, “Multi-channel intramuscular and surface emg decomposition by convolutive blind source separation,” *Journal of Neural Engineering*, vol. 13, no. 2, p. 026027, Feb. 2016. [Online]. Available: <http://dx.doi.org/10.1088/1741-2560/13/2/026027>
- [24] E. Stålberg, H. van Dijk, B. Falck, J. Kimura, C. Neuwirth, M. Pitt, S. Podnar, D. I. Rubin, S. Rutkove, D. B. Sanders, M. Sonoo, H. Tankisi, and M. Zwartz, “Standards for quantification of emg and neurography,” *Clinical Neurophysiology*, vol. 130, no. 9, p. 1688–1729, Sep. 2019. [Online]. Available: <http://dx.doi.org/10.1016/j.clinph.2019.05.008>
- [25] M. Barbero, R. Merletti, and A. Rainoldi, *Atlas of Muscle Innervation Zones: Understanding Surface Electromyography and Its Applications*. Springer Milan, 2012. [Online]. Available: <https://books.google.it/books?id=rWF-8HE9kf8C>
- [26] M. Orlandi, M. Zanghieri, V. J. K. Morinigo, F. Conti, D. Schiavone, L. Benini, and S. Benatti, “semg neural spikes reconstruction for gesture recognition on a low-power multicore processor,” in *2022 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2022, pp. 704–708.

Bibliography

- [27] S. Park, E. Aksan, X. Zhang, and O. Hilliges, *Towards End-to-End Video-Based Eye-Tracking*. Springer International Publishing, 2020, p. 747–763. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-58610-2_44
- [28] M. Mokatren, T. Kuflik, and I. Shimshoni, “Calibration-free mobile eye-tracking using corneal imaging,” *Sensors*, vol. 24, no. 4, 2024. [Online]. Available: <https://www.mdpi.com/1424-8220/24/4/1237>
- [29] J. Moosmann, P. Bonazzi, Y. Li, S. Bian, P. Mayer, L. Benini, and M. Magno, “Ultra-efficient on-device object detection on ai-integrated smart glasses with tinyissimoyolo,” 2023. [Online]. Available: <https://arxiv.org/abs/2311.01057>
- [30] JINS MEME Smart Glasses. [Online]. Available: <https://jinsmeme.com>
- [31] S. Frey, M. Guermandi, S. Benatti, V. Kartsch, A. Cossettini, and L. Benini, “Biogap: a 10-core fp-capable ultra-low power iot processor, with medical-grade afe and ble connectivity for wearable biosignal processing,” in *2023 IEEE International Conference on Omni-layer Intelligent Systems (COINS)*, 2023, pp. 1–7.
- [32] U. Seidaliyeva, D. Akhmetov, L. Ilipbayeva, and E. T. Matson, “Real-time and accurate drone detection in a video with a static background,” *Sensors*, vol. 20, no. 14, p. 3856, Jul. 2020. [Online]. Available: <http://dx.doi.org/10.3390/s20143856>
- [33] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.
- [34] S. Imbesi and M. Corzani, “Multisensory cues for gait rehabilitation with smart glasses: Methodology, design, and results of a preliminary pilot,” *Sensors*, vol. 23, no. 2, p. 874, Jan. 2023. [Online]. Available: <http://dx.doi.org/10.3390/s23020874>
- [35] S. Frey, V. Kartsch, C. Leitner, A. Cossettini, S. Vostrikov, S. Benatti, and L. Benini, “A wearable ultra-low-power semg-triggered ultrasound system for long-term muscle activity monitoring,” in *2023 IEEE International Ultrasonics Symposium (IUS)*, 2023, pp. 1–4.
- [36] V. J. K. Morinigo, S. Benatti, and L. Benini, “A high snr, low-latency dry emg acquisition system for unobtrusive hmi devices,” in *2022 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2022, pp. 544–548.
- [37] W.-J. Chang, L.-B. Chen, C.-H. Hsu, J.-H. Chen, T.-C. Yang, and C.-P. Lin, “Medglasses: A wearable smart-glasses-based drug pill recognition system using deep learning for visually impaired chronic patients,” *IEEE Access*, vol. 8, 2020.

Bibliography

- [38] N. Gospodinov and G. Krastev, “Cyber-physical system for smoke detection and recognition with smart glasses,” in *2023 31st National Conference with International Participation (TELECOM)*. IEEE, Nov. 2023. [Online]. Available: <http://dx.doi.org/10.1109/TELECOM59629.2023.10409706>
- [39] A. Huynh, Z. Marchant, I. Singh, and F. Aveta, “Assistive reading device for blind individuals,” in *2023 IEEE 14th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, 2023, pp. 0606–0612.
- [40] A. Shunaran, H. Nofal, M. Aldabe, S. Alsadi, T. Foqha, A. Al Wahhabi, and K. Al-Qawasmi, *Smart Glass for Visually Impaired Person*. Cham: Springer Nature Switzerland, 2024, pp. 627–633. [Online]. Available: https://doi.org/10.1007/978-3-031-36895-0_52
- [41] H. Kim, E. Kim, C. Choi, and W.-H. Yeo, “Advances in soft and dry electrodes for wearable health monitoring devices,” *Micromachines*, vol. 13, no. 4, 2022. [Online]. Available: <https://www.mdpi.com/2072-666X/13/4/629>