EEG Passive Brain Computer Interfaces assess cognitive and emotional condition of the user by means of electric signal acquired from the scalp. In the framework of Industry 4.0, Passive BCI represents a promising monitoring channel to improve human-machine interaction and integration.
In this thesis, the prototyping and characterization of BCI measurement instrumentation to detect basic and complex mental states are presented. Both off-the-shelf instrumentation and CE-marked devices for medical use are exploited to acquire brain signals.
The proposed solutions address the challenge of maximizing hardware wearability (minimizing the number of channels and employing dry electrodes) without penalizing accuracy and latency. To this end, appropriate signal processing strategies based on data-driven approaches are developed.
Semi-custom machine learning algorithms are implemented for feature extraction and classification.
Emotional valence, rehabilitation distraction, learning engagement, and work-related stress are the case studies proposed to experimentally validate the measurement instrumentation.
Databases of EEG signals available online were consulted and experimental campaigns were conducted for a total of more than 200 subjects.
Crucial metrological issues in the measurement instrumentation of passive BCIs are explored: e.g., definition of the measurand and its compatibility with the quantitative approach, experimental reproducibility, as well as cross- and within-subject reproducibility.
The within-subjects accuracy exceeded 92 % and 95 % for distraction and emotional valence, respectively. The cross-subject accuracy reached 99 % in recognition of a stressful condition.

Nicola Moccaldi

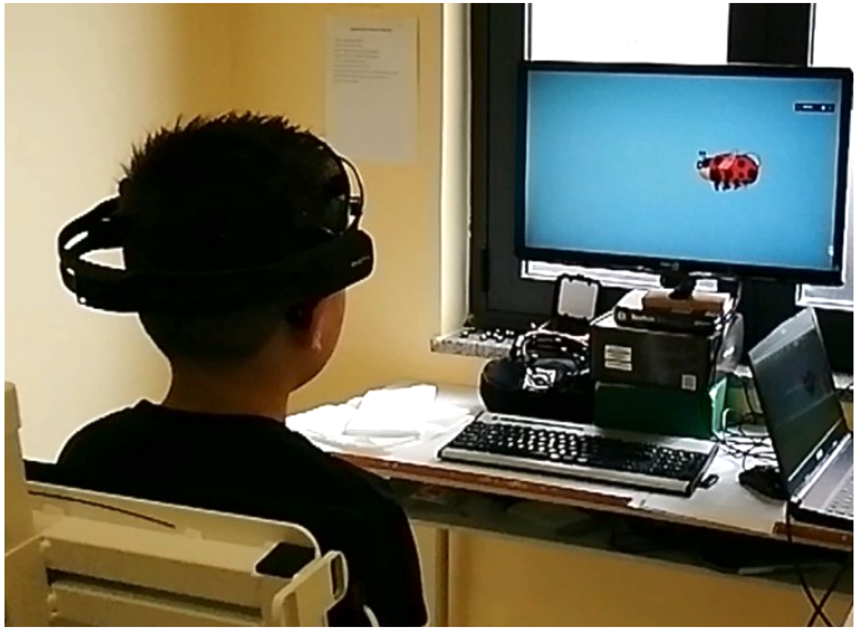
Measurement instrumentation in Passive Brain-Computer Interfaces

2021 - XXXIV

Università degli Studi di Napoli Federico II

Ph.D itee Information Technology Electrical Engineering

Measurement instrumentation in Passive Brain-Computer Interfaces



Nicola Moccaldi

Ph.D. in Information Technology and Electrical Engineering

UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

PH.D. THESIS
IN
INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

MEASUREMENT INSTRUMENTATION IN PASSIVE BRAIN-COMPUTER INTERFACES

NICOLA MOCCALDI

TUTOR: PROF. PASQUALE ARPAIA

COORDINATOR: PROF. DANIELE RICCIO

XXXIV CICLO

SCUOLA POLITECNICA E DELLE SCIENZE DI BASE
DIPARTIMENTO DI INGEGNERIA ELETTRICA E TECNOLOGIE DELL'INFORMAZIONE

UNIVERSITY OF NAPLES FEDERICO II

PH.D. THESIS

IN

INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Measurement instrumentation in Passive Brain-Computer Interfaces

Supervisors:
Prof. Pasquale Arpaia

Candidate:
Nicola Moccaldi

2020 - 2021

"I can do all things in Him who strengtheneth me"
Philippians 4:13

To my family,
to my professor Pasquale Arpaia,
to my colleagues.

Abstract

EEG Passive Brain Computer Interfaces assess cognitive and emotional condition of the user by means of electric signal acquired from the scalp.

In the framework of Industry 4.0, Passive BCI represents a promising monitoring channel to improve human-machine interaction and integration.

In this thesis, the prototyping and characterization of BCI measurement instrumentation to detect basic and complex mental states are presented.

Both off-the-shelf instrumentation and CE-marked devices for medical use are exploited to acquire brain signals.

The proposed solutions address the challenge of maximizing hardware wearability (minimizing the number of channels and employing dry electrodes) without penalizing accuracy and latency. To this end, appropriate signal processing strategies based on data-driven approaches are developed. Semi-custom machine learning algorithms are implemented for feature extraction and classification.

Emotional valence, rehabilitation distraction, learning engagement, and work-related stress are the case studies proposed to experimentally validate the measurement instrumentation. Databases of EEG signals available online were consulted and experimental campaigns were conducted for a total of more than 200 subjects.

Crucial metrological issues in the measurement instrumentation of passive BCIs are explored: e.g., definition of the measurand and its compatibility with the quantitative approach, experimental reproducibility, as well as cross- and within-subject reproducibility.

The within-subjects accuracy exceeded 92 % and 95 % for distraction and emotional valence, respectively. The cross-subject accuracy reached 99 % in recognition of a stressful condition.

Keywords: Brain Computer Interface, Health Monitoring, Wearable Device, Emotions Recognition, Attention Measurement, Stress Assessment, Engagement Detection, Machine Learning, Experimental Reproducibility

Sommario

La Brain Computer Interfaces (BCI) Passiva basata su EEG valuta la condizione cognitiva ed emotiva dell'utilizzatore attraverso il segnale elettrico acquisito dal cuoio capelluto.

Nella cornice dell'Industria 4.0, la BCI passiva rappresenta un promettente canale di monitoraggio per migliorare l'interazione e l'integrazione uomo-macchina.

In questa tesi, vengono presentate la prototipazione e la caratterizzazione della strumentazione di misurazione BCI per rilevare stati mentali di base e complessi.

Per acquisire i segnali cerebrali, sono stati impiegati sia la strumentazione off-the-shelf che i dispositivi a marchio CE per uso medico.

Le soluzioni proposte affrontano la sfida di massimizzare l'indossabilità dell' hardware (minimizzando il numero di canali e impiegando elettrodi dry) senza penalizzare la accuracy e la latenza. A tal fine, vengono sviluppate appropriate strategie di elaborazione del segnale basate su approcci data-driven. In particolare vengono proposti algoritmi di machine learning semi-custom per l'estrazione delle caratteristiche del segnale e la classificazione.

Valenza emotiva, distrazione in riabilitazione, engagement nell'apprendimento e stress legato al lavoro sono i casi di studio proposti per validare sperimentalmente la strumentazione di misura prototipata. Sono stati consultati database di segnali EEG disponibili online e sono state condotte campagne sperimentali per un totale di più di 200 soggetti.

Nella tesi vengono affrontate questioni metrologiche cruciali nella strumentazione di misura delle BCI passive: per esempio, la definizione del misurando e la sua compatibilità con l'approccio di assessment quantitativo, la riproducibilità sperimentale, così come la riproducibilità inter- ed intra-soggettiva.

List of publications

Journal publications:

(i) Apicella, A., Arpaia, P., Frosolone, M., Improta, G., Moccaldi, N., & Pollastro, A.; *"EEG-based Measurement System for Student Engagement Detection in Learning 4.0."* **Scientific Reports**.
Under Revision

(ii) Arpaia, P., Crauso, F., Frosolone, M., Mariconda, M., Minucci, S., & Moccaldi, N; *"A personalized FEM model for reproducible measurement of anti-inflammatory drugs in transdermal administration to knee."* **Scientific Reports**.
In press

(iii) Apicella, A., Arpaia, P., Giugliano, S., Mastrati, G., & Moccaldi, N.; *"High-wearable EEG-Based Transducer for Engagement Detection in Pediatric Rehabilitation."* **Brain Computer Interface** .
In press

(iv) Apicella, A., Arpaia, P., Mastrati, G., & Moccaldi, N; *"EEG-based detection of emotional valence towards a reproducible measurement of emotions."* **Scientific Reports** (November, 2021).
doi:10.1038/s41598-021-00812-7

(v) Arpaia, P., D'Errico, G., De Paolis, L., Moccaldi, N., Nuccetelli, F.; *"A Narrative Review of Mindfulness-Based Interventions Using Virtual Reality."* **Mindfulness** (October, 2021)
https://doi.org/10.1007/s12671-021-01783-6

(vi) Arpaia, P., Bonavolontà, F., Cioffi, A., & Moccaldi, N.; *"Reproducibility Enhancement by Optimized Power Analysis Attacks in Vulnerability Assessment of IoT Transducers."* **IEEE Transactions on Instrumentation and Measurement**, 70 (August, 2021).
doi: 10.1109/TIM.2021.3107610

(vii) Arpaia, P., Bonavolontà, F., Cioffi, A., & Moccaldi, N.; *"Power Measurement-based Vulnerability Assessment of IoT medical devices at varying countermeasures for cybersecurity"*. **IEEE Transactions on Instrumentation and Measurement** 70 (June, 2021).
doi:10.1109/TIM.2021.3088491

(viii) Apicella, A., Arpaia, P., Frosolone, M., & Moccaldi, N.; *"High-wearable EEG-based distraction detection in motor rehabilitation"* **Scientific Reports** 11 (March, 2021).
doi:10.1038/s41598-020-70376-5

(ix) Arpaia, P., Cuocolo, R., Donnarumma, F., Esposito, A., Moccaldi, N., Natalizio, A., & Prevete, R.; *"Conceptual design of a machine learning-based wearable soft sensor for non-invasive cardiovascular risk assessment. Measurement"* **Measurement** 169 (February, 2020).
doi: 10.1016/j.measurement.2020.108551.

(x) Arpaia, P., Cesaro, U., Frosolone, M., Moccaldi, N., & Taglialatela, M.; *"A micro-bioimpedance meter for monitoring insulin bioavailability in personalized diabetes therapy."* **Scientific Reports** 10 (December, 2020).
doi:10.1038/s41598-020-70376-5

(xi) Arpaia, P., Cesaro, U., Gatti, D., & Moccaldi, N.; *"An Ultrasonic Heading Goniometer Intrinsically Robust to Magnetic Interference."* **IEEE Transactions on Instrumentation and Measurement** 69 (November, 2020).
doi:10.1109/TIM.2020.2996785

(xii) Arpaia, P., Moccaldi, N., Prevete, R., Sannino, I., Tedesco, A.; *"A Wearable EEG Instrument for Real-Time Frontal Asymmetry Monitoring in Worker Stress Analysis."* **IEEE Transactions on Instrumentation and Measurement** 69 (October, 2020).
doi:10.1109/TIM.2020.2988744

(xiii) Arpaia, P., Duraccio, L., Moccaldi, N., Rossi, S.; *"Wearable Brain-Computer Interface Instrumentation for Robot-Based Rehabilitation by Augmented Reality."* **IEEE Transactions on Instrumentation and Measurement**, 69 (January, 2020).
doi:10.1109/TIM.2020.2970846

(xiv) Angrisani, L., Arpaia, P., Bonavolontá, F., Moccaldi, N., Schiano Lo Moriello, R.; *"A "learning small enterprise" networked with a FabLab: An academic course 4.0 in instrumentation and measurement."* **Measurement: Journal of the International Measurement Confederation** 150 (jan, 2020).
doi:10.1016/j.measurement.2019.107063

(xv) Angrisani, L., Arpaia, P., Esposito, A., Moccaldi, N.; *"A Wearable Brain-Computer Interface Instrument for Augmented Reality-Based Inspection in Industry 4.0."* **IEEE Transactions on Instrumentation and Measurement** 69 (May, 2019).
doi:10.1109/TIM.2019.2914712

(xvi) Angrisani, L., Arpaia, P., Casinelli, D., Moccaldi, N.; *"A Single-Channel SSVEP-Based Instrument with O!-The-Shelf Components for Trainingless Brain-Computer Interfaces"* **IEEE Transactions on Instrumentation and Measurement** 68 (December, 2019).
doi:10.1109/TIM.2018.2882115

Conference proceedings:

(i) Angrisani, L., Arpaia, P., Esposito, A., Gargiulo, L., Natalizio, A., Mastrati, G., Moccaldi, N., & Parvis, M.; *"Passive and active brain-computer interfaces for rehabilitation in health 4.0."* **Measurement: Sensors** 18 (September, 2021). doi: 10.1016/j.measen.2021.100246

(ii) Arpaia, P., Cuneo, D., Grassini, S., Mancino, F., Minucci, S., Moccaldi, N., & Sannino, I.; *"A finite element model of abdominal human tissue for improving the accuracy in insulin absorption assessment: A feasibility study."* **Measurement: Sensors** 18 (September, 2021).
doi:10.1016/j.measen.2021.100218

(iii) Arpaia, P., Esposito, A., Mancino, F., Moccaldi, N., & Natalizio, A.; *"Active and Passive Brain-Computer Interfaces Integrated with Extended Reality for Applications in Health 4.0".* **In: De Paolis L.T., Arpaia P., Bourdot P. (eds) Augmented Reality, Virtual Reality, and Computer Graphics. AVR 2021. Lecture Notes in Computer Science, vol 12980, Springer** (September, 2021). Available at:
doi:10.1007/978-3-030-87595-4_29

(iv) Paolis, L. T. D., Arpaia, P., D'Errico, G., Gatto, C., Moccaldi, N., & Nuccetelli, F.; *"Immersive VR as a Promising Technology for Computer-Supported Mindfulness".*; **In: De Paolis L.T., Arpaia P., Bourdot P. (eds) Augmented Reality, Virtual Reality, and Computer Graphics. AVR 2021. Lecture Notes in Computer Science, vol 12980, Springer** (September, 2021).
doi:10.1007/978-3-030-87595-4_12.

(v) Arpaia, P., De Benedetto, E., Donato, N., Duraccio, L., & Moccaldi, N.; *"A Wearable SSVEP BCI for AR-Based, Real-Time Monitoring Applications"* **In: 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)**, Lausanne, Switzerland (June, 2021).
doi: 10.1109/MeMeA52024.2021.9478593.

(vi) Angrisani, L., Arpaia, P., De Benedetto, E., Esposito, A., Moccaldi, N., & Parvis, M.; *"Brain-computer Interfaces for Daily-life Applications: a Five-year Experience Report".* **In: 2021 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)** (May, 2021).
doi:10.1109/I2MTC50364.2021.9459844

(vii) Apicella, A., Arpaia, P., Mastrati, G., Moccaldi, N., Prevete, R.; *"Preliminary validation of a measurement system for emotion recognition".* **In: 2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA)** (June, 2020).
doi:10.1109/MeMeA49120.2020.9137353

(viii) Arpaia, P., Crauso, F., Grassini, S., Minucci, S., Moccaldi, N., Sannino, I.; *"Preliminary experimental identification of a FEM human knee model".* **In: 2020**

IEEE International Symposium on Medical Measurements and Applications (MeMeA) (June, 2020).
doi:10.1109/MeMeA49120.2020.9137187

(ix) Arpaia, P., Bravaccio, C., Corrado, G., Duraccio, L., Moccaldi, N., Rossi, S.; *"Robotic Autism Rehabilitation by Wearable Brain-Computer Interface and Augmented Reality".* **In: 2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA)** (June, 2020).
doi:10.1109/MeMeA49120.2020.9137144

(x) Annuzzi, G., Arpaia, P., Cesaro, U., Cuomo, O., Frosolone, M., Grassini, S., Moccaldi, N., Sannino, I.; *"A customized bioimpedance meter for monitoring insulin bioavailability".* **In: 2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)** (May, 2021).
doi:10.1109/I2MTC43012.2020.9128676

(xi) Angrisani, L., Arpaia, P., Donnarumma, F., Esposito, A., Frosolone, M., Improta, G., Moccaldi, N., Natalizio, A., Parvis, M.; *"Instrumentation for motor imagery-based brain computer interfaces relying on dry electrodes: A functional analysis".* **In: 2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)** (May, 2021).
doi:10.1109/I2MTC43012.2020.9129244

(xii) Apicella, A., Arpaia, P., Frosolone, M., Improta, G., Isgrò, F., Moccaldi, N., Natalizio, A.; *"EEG-based attention assessment in motor-rehabilitation".* **In: 24th IMEKO TC4 International Symposium 22nd International Workshop on ADC and DAC Modelling and Testing, IMEKO TC-4 2020)** (September, 2020).
Available at: https://www.imeko.org/publications/tc4-2020/IMEKO-TC4-2020-04.pdf

(xiii) Arpaia, P., Cuocolo, R., Donnarumma, F., D'Andrea, D., Esposito, A., Moccaldi, N., Natalizio, A., Prevete, R.; *"Feasibility of cardiovascular risk assessment through non-invasive measurements".* **In: 2019 II Workshop on Metrology for Industry 4.0 and IoT (MetroInd4.0&IoT))** (June, 2019).
doi:10.1109/METROI4.2019.8792909

(xiv) Angrisani, L., Arpaia, P., Donnarumma, F., Esposito, A., Moccaldi, N., Parvis, M.; *"Metrological performance of a single-channel brain-computer interface based on motor imagery".* **In: 2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)** (May, 2019).
doi:10.1109/I2MTC.2019.8827168

Books:

(i) Arpaia, P., Cesaro, U., Moccaldi, N., Sannino, I. *"Non-Invasive Monitoring of Transdermal Drug Delivery"*. **CRC Press**
In press.

Contents

Abstract vii

Sommario ix

List of publications xi

Introduction 1

I Background 7

1 BCI-detectable mental states in the framework of industry 4.0 9
1.1 Emotions . . . 9
1.2 Distraction . . . 10
1.3 Engagement . . . 11
1.3.1 Engagement in learning 4.0 . . . 11
1.3.2 Engagement and adaptive automated rehabilitation platforms . . . 13
1.4 Stress . . . 14

2 State of art and metrological problems 17
2.1 Measurement of Emotion . . . 17
2.1.1 Related Works . . . 19
2.1.2 Statement of the metrological problem . . . 23
2.2 Measurement of attention . . . 24
2.3 Measurement of engagement . . . 25
2.3.1 Engagement assessment during learning activities . . . 25
2.3.2 Engagement detection in pediatric rehabilitation . . . 26
2.4 Measurement of stress . . . 27
2.4.1 Methods based on handcrafted feature for stress detection . 28
2.4.2 Data-driven methods for stress detection . . . 28

II Proposal 31

3 Basic mental state assessment 33
3.1 Emotional valence detection . . . 33
3.1.1 Basic ideas . . . 33
3.1.2 Architecture . . . 34
3.1.3 Data processing . . . 34

Features extraction and selection 35
Classification 35
3.1.4 Experiments and Results 36
Data acquisition setup 36
Hardware 38
Data processing comparison 39
Experimental results 41
Discussion 44
3.2 Distraction assessment 47
3.2.1 Basic Ideas 47
3.2.2 Method 47
Feature selection and extraction 48
Classification 49
3.2.3 Experimental validation results 49
Experimental Protocol 49
EEG Instrumentation 51
Data Processing 52
Experimental Results 54

4 Complex mental state assessment **57**
4.1 EEG-Based Stress Assessment 57
4.1.1 Design 57
4.1.2 Realization 60
4.1.3 Experimental Setup 61
4.2 Engagement in rehabilitation 67
4.2.1 Basic Ideas 67
4.2.2 Methods 67
4.2.3 Experimental Setup 68
4.2.4 Discussion 71
4.3 Engagement detection in learning 74
4.3.1 Basic Ideas 74
4.3.2 Architecture 74
4.3.3 Processing Framework 75
4.3.4 Experimental Setup 76
4.3.5 Experimental Results 78

Conclusion **82**

References **86**

List of figures **105**

List of tables **109**

List of Abbreviations

ANN	**A**rtificial **N**eural **N**etwork
CPHS	**C**yber **P**hysical **H**umas **S**ystem
CSP	**C**ommon **S**patial **P**attern
EEG	**E**lectro **E**ncefalo **G**ram
FB	**F**ilter **B**ank
k-NN	k **N**earest **N**eighbors
LDA	**L**inear **D**iscriminant **A**nalysis
LR	**L**ogistic **R**regression
PCA	**P**rincipal **C**omponent **A**nalisys
RF	**R**andom **F**orest
SVM	**S**upport **V**ector **M**achine
t-SNE	t-distributed **S**tochastic **N**eighbor **E**mbedding
TFA	**T**ransfer **C**omponent **A**nalysis

Introduction

Cyber-Physical Human Systems (CPHS) integrate the physical and human components into a synthetic hybrid system [1]. In the context of Industry 4.0, human does not just exercise a defined role in an organization, but becomes part of a highly-composite automated system [2, 3]. In industry or in health care, the smart machines, non-human components of CPHS, are more and more connected to the physical environment through sensors of all kinds. Thanks to a distributed intelligence, the non-human actors can elaborate information and make decisions, resulting highly empowered by technology innovation. Also humans benefit from the new technological opportunities: by interacting with new-generation user interfaces, they obtain a strengthening of cognitive, sensorial, and motor skills [4].

Among biosignal-based interfaces, Brain-Computer Interface (BCI) allows both monitoring and control. Humans can send messages or decisions to the CPHS through intentional modulation of brain waves. However, through the same signal, the system (and therefore also the human being part of it) acquires information on the status of the user.

Passive BCI (complementary to active BCI) is the paradigm adopted when the user does not directly and consciously control his electrical brainwaves and, therefore, when the goal is the monitoring of his/her current state. There are many invasive and non-invasive techniques for understanding the brain signals such as PET (Positron Emission Tomography), MEG (Magneto Encephalography), NIRS (Near-infrared Spectroscopy), fMRI (Functional Magnetic Resonance Imaging), EROS (Event-related optical signal), and EEG (Electroencephalogram). Among the mentioned systems, EEG offers a better temporal resolution. Moreover, several portable and wearable EEG solutions are already on the market.

Currently, the wearability improvement of EEG-based BCI instrumentation is a widely shared challenge. Dry electrodes [5, 6] and a low number of channels are promising strategies to enhance user comfort. However, appropriate signal processing algorithms must be developed to compensate for losses in the signal-to-noise ratio and the number of sources.

In the context of Industry 4.0, monitoring of emotions, attention, engagement, and stress are pressing issues in different application domains, both in terms of the production process and product innovation. The ongoing technological transformation introduces new oportunity and risk specifically connected to the new framework of very high human-machine interaction. Passive BCI is a promising channel to improve the adaptivity of cyber-physical system to human. Among those emerging under *Industry 4.0*, five issues will be focused on in this thesis: (i) *Emotional valence in the CPHS*, (ii) *Attention in robotic motor rehabilitation*, *Engagement in adaptive Extended Reality instrumentation for neuro-motor re- habilitation*, *Engagement in learning 4.0*, and *Stress 4.0*.

Emotional valence in the CPHS. Emotion is the response to imaginary or real stimuli characterised by changes in individual's thinking, physiological responses, and behaviour [7]. In the *Circumplex Model* [8] of emotion, *valence* denotes how much an emotion is positive or negative. Discrimination of emotional valence is a broad issue widely addressed in recent decades, affecting the most varied sectors and finding application in multiple domains. Currently, real time monitoring of emotion is proposed in application fields such as: industry [9], health [10, 11], and entertainment [12]. Several biosignals have been studied over the years for emotions recognition: cerebral blood flow[13], electroculographic (EOG) signals[14], electrocardiogram, blood volume pulse, galvanic skin response, respiration, phalanx temperature[15]. In recent years, several studies have focused on the brain signal in particulare based on EEG [16]. There are already some portable EEG solutions on the market. Currently, the measurand definition is a fundamental issue, because the quantity is not univocally identified and many relevant theories are incompatible with the adoption of an interval scale (ordered and proportionated). Moreover, experimental reproducibility, as well as cross-subject and within-subject reproducibility are open challenges.

Attention in robotic motor rehabilitation. The effectiveness of robotic therapy on motor recovery is well assessed in literature [17]. Recently, its impact on cognitive functions was also investigated [18] . In general, neuromotor rehabilitation exercise induces neuronal neuroplasticity and promotes motor recovery [19]. In particular, the repetition of the exercise induces a reorganization of the motor cortex. However, the repetition of the same exercise may induce weariness in the subject and prevent a careful focus on the performance of the exercise. Conversely, completing the exercise, while maintaining the attention focus in a sustained and selective way, promotes neuronal plasticity and motor learning [20, 21]. The attention to the motor task has an enhanced effect on rehabilitation performance[22]. Many studies deal with assessing the attention and its different dimensions through the analysis of the brain signals using the electroencephalography [23]. However, an appropriate approach for clinical application seems to be currently missing [24]. The high number of channels and the use of wet or semi-wet electrodes penalize the wearability, limiting the clinical usability.

Engagement in adaptive Extended Reality instrumentation for neuro-motor rehabilitation. Engagement assessment is fundamental in clinical practice to personalize treatments and improve their effectiveness. Indeed, patients involved in healthcare decision-making tend to perform better and to be healthier.

The standard tools used in clinical practice for engagement assessment are questionnaires or rating scales. Both take into account the patients' awareness of their health and their therapeutic process. Beyond standard tools, biosignals-based measurement methods are emerging. They allow an automated and real-time engagement assessment. In particular, eye-blinking [25], heart rate variability [26], and brain activity [27, 28] were used to detect changes in patient's engagement. Among these, the EEG signal [29] offers good temporal resolution and improves real-time performances. In the rehabilitation field, studies on EEG-based engagement detection were mainly conducted on adults and focused only on cognitive engagement [30]. The reasons could be: (i) the engagement measure in the rehabilitation field has only recently become an object of interest [31], and (ii) EEG-based engagement assessment in pediatric rehabilitation requires the adoption of a respectful clinical protocol to protect the child and his psychophysical integrity (i. e. a non-interventional observational approach). Although such a protocol is more comfortable for the children, it entails a general lack of control over the engagement levels resulting in imbalanced data collections during experimental campaigns.

Engagement in learning 4.0. Man's relationship with knowledge is increasingly mediated by technology. Digital era [32], the period of the pervasive use of information and communication technologies in every area of life, has heavily impacted on *learning* starting from the second half of the last century. Currently, the ongoing fourth industrial revolution (Industry 4.0) expands the role of technology in learning processes even further: automated teaching platforms can real-time adapt to the user skills and the new generation interfaces allow multisensorial interactions with virtual contents.

The 4.0 technologies are strongly impacting on the creation, the conservation, and the transmission of knowledge [33]. In particular, the new immersive eXtended Reality (XR) solutions make possible to achieve *embodied learning* by restoring the role of learning catalyst to bodily activities [34]. Furthermore, wearable transducers and embedded Artificial Intelligence (AI) increase real-time adaptivity in human-machine interaction [35]. In detail, in the Learning 4.0 context, the adaptation is reciprocal: the subject learns to use the human-machine interface, but also the machine adapts to human by learning from her/him [36].

The effectiveness learning process mainly depends on the engagement level of the learner [37]. Therefore, the engagement monitoring is a fundamental aspect allowing the machine to adapt to the user. As concerns the engagement measurability, evaluation grids and self-assessment questionnaires (to be filled out by the observer or by the learner autonomously) are traditionally the most used methods for the behavioral, cognitive, and emotional engagement detection [38]. In recent years, measures based on biosignals are spreading very rapidly. Furthermore, the use of physiological sensors able to detect cognitive and emotional engagement allows the real-time machine adaptive strategies. Among the different physiological biosignals, the EEG appears to be one of the most promising technology thanks to its low cost, low invasiveness, and high temporal resolution. Moreover, the EEG contains a broader range of information about the state of a subject with respect to others biosignals [39]. However, published EEG-based studies still do not take into account the different engagement types (i.e., cognitive, emotional and behavioural) [40].

Stress 4.0. Stress is a psycho-physical pathological response to emotional, cognitive, or social tasks, perceived as excessive by an individual. Many stimuli of different nature (physical, toxic, emotional), external to individuals, could disturb their homeostasis and psychological well-being, bringing to an adaptive or non-adaptive response [41]. In industrial work, stress has negative impact on safety, on the quality of the outcome and, thus, on the cost of the production process as a whole [42]. Technological innovation, indeed, has introduced new sources of stress (*stress 4.0*). Intelligent automated systems in their various configurations, robots or cobots in collaborative meaning [43], interact continuously with individuals in a constant relationship of cooperation and, at the same time, of unconscious competition [44]. In literature, different indicators of stress status are proposed, arising from products of neuroendocrine reactions affecting sympathetic and parasympathetic nervous systems [45]. Some biochemical and biophysical markers are measured usually by invasive methods: (i) Cortisol Concentration in blood or saliva; (ii) Galvanic Skin Response; (iii) Heart Rate; and (iv) Brain Activity. Some use cases of stress recognition based on EEG are given in the literature [46, 47]. Significant is a wearable EEG device for construction workers [48]. High vulnerability characterizes the activities on-site of the workers during a construction process; so, they suffer from load stress. By including an EEG device into their protective helmet, brain waves are monitored and analyzed along the activity, by highlighting possible emotional states and, therefore, actual attention levels [49]. However, state-of-the-art solutions exhibit at least one of the following weaknesses: (i) limitations for daily on-field use, e.g. due to a large number of wet electrodes and use of wired systems; (ii) accuracy less than 90%, even in case of simultaneous ECG and EEG measurements [50]; and (iii) high cost, up to thousands of dollars [51].

This thesis presents feasibility studies on monitoring mental states in typical industry 4.0 contexts, using reproducible measurements of EEG signals, acquired by highly wearable devices. In particular, for emotion assessment, the reference theory adopted allows the measurement of emotions arranging them along interval scales. In the case of engagement, the combined assessment of cognitive and emotional dimension allows a better adaptability of the automated system both in therapy and in learning. In application involving children the instrumentation is prototyped through observational non-interventional experimental activities and a suitable unbalanced data management method is adopted. In case of stress detection a very low-cost prototype is developed.

For all the proposed solutions, several machine learning algorithms were compared to maximize the accuracy in detecting the desired mental state.

The work is divided into two parts: part I, Background, and part II, Proposal. In part I, the definitions of basic and complex mental states are presented, together with the state of the art of their assessment. In part II, the instrumentation prototyping and experimental validation are reported.

The structure of the Chapters is as follows:

- **Chapter 1: BCI detectable mental states**. Emotional valence, Attention, Engagement, and Stress are defined according to the main theories of Psychological Literature.

- **Chapter 2: State of art and metrological problems**. The state of art in measuring the above described mental state is presented. Moreover, some crucial metrological issues are discussed.

- **Chapter 3: Basic mental state assessment**. The prototyping of instrumentation to measure valence emotion and distraction is presented. Then, the experimental validation, with the relative results, is reported.

- **Chapter 4: Complex mental state assessment**. The design and realization of instrumentation to assess engagement and stress are illustrated. Finally, the experiments to validate the proposed solution are reported and results are discussed.

Part I

Background

Chapter 1

BCI-detectable mental states in the framework of industry 4.0

In this chapter, basic and complex mental states relevant in the perspective of industry 4.0 are introduced. In particular, emotions, distraction are discussed in the sessions 1.1 and 1.2, respectively. In this thesis, engagement and stress are labeled as complex mental states because many theories combine at least emotional and congnitive dimensions for their definition. Engagement is discussed in Session 1.3, and Stress in Section 1.4.

1.1 Emotions

Discrimination of emotional valence is a broad issue widely addressed in recent decades, affecting the most varied sectors and finding application in multiple domains. Currently, real time monitoring of emotion is proposed in application fields such as: industry [9], health [10, 11], and entertainment [12]. A huge variety of definitions of the term *emotion* has been provided over the years. According to the highlighted characteristics, definitions can be: (i) affective (i.e., feelings of pleasure/displeasure and excitement/depression), (ii) cognitive (i.e., appraisal processes), (iii) Stimuli-Organism-Response (SOR) based (i.e., effects of external stimuli on physiological mechanisms), (iv) adaptive/ disruptive, (v) multiaspect, (vi) restrictive (i.e., attempt to differentiate emotions from other processes), (vii) motivational, and (viii) skeptical (i.e., the usefulness of the concept of emotion is denied). A definition encompassing all the aforementioned aspects was provided by Kleinginna et al.:"Emotion is a complex set of interactions among subjective and objective factors, mediated by neural-hormonal systems, which can: (i) give rise to affective experiences such as feelings of arousal, pleasure/displeasure; (ii) generate cognitive processes such as emotionally relevant perceptual effects, appraisals, labeling processes; (iii) activate widespread physiological adjustments to the arousing conditions; and (iv) lead to behavior that is often, but not always, expressive, goal-directed, and adaptive" [7].
Also the nature of emotions is a strongly debated issue. The focus is mainly on whether they are discrete or dimensional.
Discrete theories of emotions suggests the existence of few separate emotions, each with specific characteristic patterns. Six basic emotions (i.e., anger, disgust, fear, joy, sadness, and surprise) were proposed by Ekman [52]. Other discrete models

include usually from 2 to 10 emotions. Basic emotions are also called primary emotions. Secondary emotions, instead, result from the combination of the primary ones.
Dimensional theories of emotion propose the existence of underlying affective dimensions common to all emotions. Thus, emotions can be represented in a multidimensional space. In the *Circumplex Model of Affect* [53], proposed by Russel and Mehrabian, emotions are categorized according to valence (i.e., positive/negative affect) and arousal (i.e., low/high level of activation) dimensions. A third dimension, dominance, represents the presence/absence of control over the situation. Further highly accredited theories on emotions are those of *appraisal* [54]. The evaluation an individual makes about a stimulus or a situation determines the elicitation of an emotion. The significance of the event strongly depends on the subject's personal experiences and goals.
Largely studied in the field of emotion recognition are the theories about cortical brain lateralization. The *Theory of Right Hemisphere* claims that each emotional expression and perception takes place in the right hemisphere [55]. The *Theory of Valence* affirms that the right hemisphere is dominant for processing negative emotions and the left hemisphere is dominant for processing positive emotions [56]. Similarly, the *Approach-Withdrawal* model posits the role of the left- and right-anterior regions in processing emotional states in the government of approach and withdrawal behaviors [57]. The *Behavioral Activation System – Behavioral Inhibition System (BAS/BIS)* model states that the left and the right frontal activity reflects the strength of the BAS and BIS systems, respectively [58]. BAS/BIS are the two anatomical paths governing the emotional/motivational systems. The BAS is responsible for the activation of the behavior in response to rewarding stimuli and it associates emotions (which are generally positive, like hope and relief) with these behaviors. The BIS, on the other hand, inhibits behavior in response to stimuli that are new, feared, and adverse. BIS activates with behaviors of passive avoidance and extinction and the related emotions are generally negative (e.g., anxiety, fear).

1.2 Distraction

In everyday life, many types of distracting effects (visual, auditory, and their combinations) sidetrack attention when performing any task, especially if it requires engagement [59]. Diez et al. identified attention just as the ability to select interesting stimuli, by ignoring other distracting stimuli in the surrounding environment [60]. These distractors play a fundamental role in analyzing the attentional process [61]. Changes in cognitive processes related to attention activate different parts of the brain. Concurrent distracting events deactivate certain brain areas by activating other ones [62].

Ladvas and Berti describe attention as the function that regulates the filtering and organization of the information received from a subject, allowing his/hers adequate responses [63]. Sohlberg and Mateer propose a characterization of attention in four different dimensions [64]: (i) the *Arousal* indicates the activation level and defines the psychophysiological activation allowing the afference of the different stimulations; (ii) the *Selective attention*: points out the ability to focus attention on a specific source or sensory channel; (iii) the *Distributed attention* is the ability to simultaneously process information from multiple sources; and (iv) the *Sustained attention* is the ability to direct and maintain cognitive prolonged activity on a specific stimuli.

The attention monitoring is crucial in neuromotor rehabilitation. Ang et al. prove that a neuromotor rehabilitation exercise induces neuronal neuroplasticity and promotes motor recovery [19]. In particular, the repetition of the exercise induces a reorganization of the motor cortex. However, the repetition of the same exercise may induce weariness in the subject and prevent a careful focus on the performance of the exercise. Conversely, completing the exercise, while maintaining the attention focus in a sustained and selective way, promotes neuronal plasticity and motor learning [20, 21]. The attention to the motor task has an enhanced effect on rehabilitation performance[22].

The use of distracting stimuli during the execution of a motor task, as opposed to the careful concentration, characterizes the experimental set-ups of the studies on the measurement of motor attention [59, 65]. Many studies deal with assessing the attention and its different dimensions through the analysis of the brain signals using the ElectroEncephalography (EEG) [23]. EEG is the most used technique because of its high temporal resolution, non-invasiveness, and low cost. Several studies have shown that the level of attention affects the EEG signal [66, 67]. Therefore, variations in the EEG signal can be used to detect corresponding changes in attention levels [68]. Attention creates a variation in brain signals that can be assessed both in the time and in the frequency domain [69].

1.3 Engagement

The term engagement, derived from the verb "engager", is often used as a synonym for commitment and / or involvement.

Several definitions have been provided over the years because of its multidimensional and heterogeneous nature. In 1990, Kahn based the definition of engagement on three broad dimensions: behavioural, cognitive, and emotional [70]. Behavioral engagement is the set of observable indicators (postures, gestures, actions, etc.) of persistence and participation. Cognitive engagement is the effort to extend one's intellectual commitment beyond the minimum required to complete the task. Finally, emotional engagement is the positive emotional reactions of individuals to a task. In what follows, engagement is analyzed in the framework of learning activities (Subsection 1.3.1) and adaptive automated rehabilitation solutions (Subsection 1.3.2).

1.3.1 Engagement in learning 4.0

Man's relationship with knowledge is increasingly mediated by technology. Digital era [32], the period of the pervasive use of information and communication technologies in every area of life, has heavily impacted on *learning* starting from the second half of the last century. Currently, the ongoing fourth industrial revolution (Industry 4.0) expands the role of technology in learning processes even further: automated teaching platforms can real-time adapt to the user skills and the new generation interfaces allow multi-sensorial interactions with virtual contents.

In the pedagogical domain, the concept of "Learning 4.0" is emerging and it is not just a marketing gimmick [71].

The 4.0 technologies are strongly impacting on the creation, the conservation, and the transmission of knowledge [33]. In particular, the new immersive eXtended Reality (XR) solutions make possible to achieve *embodied learning* by restoring the role of learning catalyst to bodily activities [34]. Furthermore, wearable transducers and embedded Artificial Intelligence (AI) increase real-time adaptivity in human-machine interaction [35]. In detail, in the Learning 4.0 context, the adaptation is reciprocal: the subject learns to use the human-machine interface, but also the machine adapts to human by learning from her/him [36].

Traditionally, learning to use a new technological interface was a once-in-a-lifetime effort as a child. For many people this has occurred with learning to read and write. Recently, the rapidity of technological evolution has been entailing the need to learn how to use several interfaces. The joy-pad, icon, touch/multi-touch screen, speech and gesture recognition are examples of the interface (hardware and software components) evolution of new interfaces.

More specifically, learning to use an interface is an hard task which requires complex cognitive-motor skills. When human beings learned to use the mouse and touchscreen, as well as when they learned to write, read or speak, their minds learned complex cognitive-body patterns [72, 73].

Regarding the human-machine interfaces of older generation, the user was autonomously required to explore the different available resources and learn their use. Currently, the interfaces 4.0 can adapt in real time to the user supporting the learning process.

The effectiveness learning process mainly depends on the engagement level of the learner [37]. Therefore, the engagement monitoring is a fundamental aspect allowing the machine to adapt to the user.

In this context, *Engagement* stands for concentrated attention, commitment, and active involvement in contrast to apathy, lack of interest or superficial participation [74, 75].

In the learning context, Fred Newman, in his report "Student Engagement and Achievement in American Secondary Schools", defines engagement as: "the student's psychological investment in and effort directed toward learning, understanding, or mastering the knowledge, skills, or crafts that academic work is intended to promote" [76, 77].

Moreover, Frederiks defines the student engagement as a meta-construct that includes: behavioral, emotional, and cognitive engagement [78].

In general, learning a new interface can be traced back to a classic learning problem. In the constructivism framework learning consists in the construction of the schemes: units of knowledge, each relating to different aspect of the world, including actions, objects, and abstract concepts [79]. When a subject learns a specific pattern, the *neuroplasticity process* is activated modifying the neural brain structure [80]. Once the process is learned, the brain builds a myelinated axon connection system to automate that. The adjacent neurons fire in unison, and more the experience or operation is repeated, more the synaptic link between neurons becomes strong [81]. The automated use of all mental processes as well as the understanding and use of new technologies occurs through the creation of neural diagrams and maps [82, 83]. During life, humans learn new skills or modify the already learned ones by enriching the existing neural maps. Therefore, the introduction of increasingly innovative technologies requires a continuous brain re-adaptation to new interfaces [84]. This effort is more effective when the learner is engaged. An engaged user actuates learning in an optimal way, avoiding distractions, and increasing the mental performance [85, 86].

In [87], three different types of engagement are proposed:

behavioural, emotional, and cognitive engagements. Behavioral engagement focuses on the observable actions during the learning process [88, 89]. Emotional engagement regards the impact of emotions on the cognitive process effectiveness and the effort sustainability for the users [90]. Cognitive engagement refers to the amount of cognitive resources spent by the user in a specific activity [89, 91].

1.3.2 Engagement and adaptive automated rehabilitation platforms

Engagement assessment is fundamental in clinical practice to personalize treatments and improve their effectiveness. Indeed, patients involved in healthcare decision-making tend to perform better and to be healthier.

In rehabilitation, Graffigna et al. defined patient engagement as a "multidimensional psycho-social process, resulting from the conjoint cognitive, emotional, and behavioral enactment of individuals toward their health condition and management" [92]. The cognitive dimension refers to the meaning given by the patient to the disease, its treatments, its possible developments, and its monitoring. The emotional dimension consists of the emotive reactions of patients in adapting to the onset of the disease and the new living conditions connected to it. The behavioral dimension is connected to all the activities the patient acts out to face the disease and the treatments.

Lequerica et al. defined engagement in rehabilitation as "a deliberate effort and commitment to working toward the goals of rehabilitation interventions, typically demonstrated through active, effortful participation in therapies and cooperation with treatment providers" [93]. Moreover, the authors highlighted the role of motivation in triggering and feeding engagement. Motivation can be intrinsic or extrinsic. Deci and Ryan [94] defined intrinsically motivated behaviours as those "for which the rewards are internal to the person". Conversely, extrinsically motivated behaviours are performed to obtain external reward such as money or praise. According to the authors, intrinsic goals are more powerful motivators than extrinsic or externally imposed goals. Intrinsic motivational factors influencing therapeutic engagement are: (i) perception of the need for treatment; (ii) perception of the probability of a positive outcome; (iii) perception of self-efficacy in completing tasks, and (iv) re-evaluation of beliefs, attitudes and expectations [93].

In pediatric rehabilitation, it is difficult to achieve engagement by relying only on intrinsic motivation. Therefore, the extrinsic motivation is required. Children only react to what is real, concrete, present and immediately satisfying.

A fundamental extrinsic factor in supporting the child's self-esteem and perceived self-efficacy in rehabilitation activities, is the relationship with the therapist. As in the educational field, this process is referred to as scaffolding [95] and it is intended as cognitive and emotional support. Thus, pediatric engagement is a complex construct "involving a connection, a sense of working together, and particular experiences that influence emotions, feelings, and motivation in the therapy process" [96].

1.4 Stress

Stress is a psycho-physical pathological response to emotional, cognitive, or social tasks, perceived as excessive by an individual. Many stimuli of different nature (physical, toxic, emotional), external to individuals, could disturb their homeostasis and psychological well-being, bringing to an adaptive or non-adaptive response [41]. In industrial work, stress has negative impact on safety, on the quality of the outcome and, thus, on the cost of the production process as a whole [42]. Technological innovation, indeed, has introduced new sources of stress (*stress 4.0*). Intelligent automated systems in their various configurations, robots or cobots in collaborative meaning [43], interact continuously with individuals in a constant relationship of cooperation and, at the same time, of unconscious competition.

In literature, different indicators of stress status are proposed, arising from products of neuroendocrine reactions affecting sympathetic and parasympathetic nervous systems [45]. Some biochemical and biophysical markers are measured usually by invasive methods: (i) Cortisol Concentration in blood or saliva; (ii) Galvanic Skin Response; (iii) Heart Rate; and (iv) Brain Activity.

Cortisol is a hormone produced by the adrenal glands with the aim of preserving homeostasis in all conditions tending to alter the normal body balance. Cortisol concentration in blood has been used as the first index of the individual's response to stress. It is measured through repeated blood samples, or through saliva samples, by means of less invasive methods but with less significance [42].

Skin conductance is a further parameter associated to the activation of the sympathetic nervous system and, therefore, to stress. Stress induces an increase in the epidermis moisture and, therefore, a reduction in skin resistance.

Furthermore, stress generates peripheral vasoconstriction that causes a decrease in wave amplitudes of electrocardiogram (ECG) and an increase in the heart rate [45].

Brain activity produces electrical signals as a response to all kind of internal and external stimuli. The signals are recorded either through functional Magnetic Resonance Imaging, Positron Emission Tomography, or electroencephalography

(EEG). All these techniques detect brain activity changes in the limbic system and frontal regions.

EEG is the most widely used because it is easy to implement and little intrusive; moreover, EEG signals can be classified effectively through a frequency analysis. Some use cases of stress recognition based on EEG are given in the literature[46, 47]. Significant is a wearable EEG device for construction workers [48]. High vulnerability characterizes the activities on-site of the workers during a construction process; so, they suffer from load stress. By including an EEG device into their protective helmet, brain waves are monitored and analyzed along the activity, by highlighting possible emotional states and, therefore, actual attention levels [49]. However, state-of-the-art solutions exhibit at least one of the following weaknesses: (i) limitations for daily on-field use, e.g. due to a large number of wet electrodes and use of wired systems; (ii) accuracy less than 90%, even in case of simultaneous ECG and EEG measurements [50]; and (iii) high cost, up to thousands of dollars [51].

Chapter 2

State of art and metrological problems

In this chapter, the state of art of EEG-based measurement of mental states is presented. Furthermore, fundamental metrological issues concerning emotion detection are discussed in Section 2.1. Attention assessment is detailed in Section 2.2, while engagement and stress measurement are presented in Section 2.3 and 2.4, respectively.

2.1 Measurement of Emotion

Several biosignals have been studied over the years for emotions recognition: cerebral blood flow[13], electroculographic (EOG) signals[14], electrocardiogram, blood volume pulse, galvanic skin response, respiration, phalanx temperature[15]. In recent years, several studies have focused on the brain signal. There are many invasive and non-invasive techniques for understanding the brain signals such as: PET (Positron Emission Tomography), MEG (Magneto Encephalography), NIRS (Near-infrared Spectroscopy), fMRI (Functional Magnetic Resonance Imaging), EROS (Event-related optical signal), EEG (Electroencephalogram). Among the mentioned systems, EEG offers a better temporal resolution. There are already some portable EEG solutions on the market. Currently, a scientific challenge is to use dry electrodes [5, 6] and increasingly reduce the number of channels to maximise the user comfort while maintaining high performances.

The measurement of emotions[97] is different from the emotion recognition and it requires certain conditions to be met. The first condition concerns the use of an *interval scale* besides the management of the reproducibility problem. The well-assessed taxonomy given by Stevens [98] provided a fourfold classification scheme of measurement scales: *nominal*, *ordinal*, *interval*, and *ratio* scales. Nominal and ordinal scales represent non-additive quantities and, therefore, cannot be considered for measurements according to the International Vocabulary of Metrology[99]. Studies adopting the theory of discrete emotions [100] employ a nominal scale providing only classifications. Conversely, the Circumplex Model allows the measurement of emotions by arranging them along interval scales.

As concerns the second condition, often, the same stimulus or environmental condition does not induce the same emotion in different subjects (cross-subject reproducibility loss). Furthermore, the same individual exposed to the same stimulus but after a certain period of time, reacts in a different way (within-subject reproducibility loss). In psychology research, suitable sets of stimuli were validated experimentally by using significant samples and are widely used by clinicians and researchers [101]. In particular, several stimuli datasets were produced referring to the Circumplex Model and their scores were arranged along an interval scale. However, the problem of standardizing the induced response remains still open, also considering, for example, the issue of the cross-cultural generality of perceptions. The effectiveness of the emotion induction can be verified by means of self-assessment questionnaires or scales. The use of the validated stimulus rating and the subject's self-assessment can represent an effective strategy towards the construction of a metrological reference for the EEG-based reproducible measurement of emotions [102]. Furthermore, the use of assessment tools during the sample construction can soften possible emotional bias caused by psychiatric disorders.

As concerns the measurement model, older approaches predominantly made use of a priori knowledge. Emotions studies, based on spatial distribution analysis of EEG signal, were principally focused on the asymmetric behaviour of the two cerebral hemispheres[103, 104, 105]. Two theories, in particular, model the relationship between emotions and asymmetry in a different way. The *Theory of Right Hemisphere* posits that the right hemisphere is dominant over the left hemisphere for all forms of emotional expression and perception. Instead, the *Theory of Valence* states that the right hemisphere is dominant (in term of signal amplitude) for negative emotions and the left hemisphere is dominant for positive emotions. In particular the theory of valence focuses on what happens in the two areas of the prefrontal cortex. The prefrontal cortex plays an important role in the control of cognitive functions and in the regulation of the affective system[106]. The EEG asymmetry allows to evaluate the subject's emotional changes and responses and, therefore, it can serve as an individual feature to predict emotional states [107]. The most common frequency index for emotion recognition is the so called *frontal alpha asymmetry* (α_{asim})[108]:

$$\alpha_{asim} = \ln(\alpha_{PSD_L}) - \ln(\alpha_{PSD_R}) \tag{2.1}$$

where the parameters α_{PSD_L} and α_{PSD_R} are the power spectral densities of the left and right hemispheres in the alpha band. Frontal alpha asymmetry could also predict emotion regulation difficulties by resting state EEG recordings. Frontal EEG asymmetry effects are quite robust to individual differences [109]. Several modern Machine Learning systems automatically carry out the feature extraction procedure. Therefore, a very large number of data from different domains (i.e., spatial, spectral or temporal) can be used as input to the classifier without an explicit hand-crafted feature extraction procedure.

Spatial filters usually enhance sensitivity to particular brain sources, to improve source localization, and/or to suppress muscular or ocular artifacts [110]. Two different categories of spatial filters exist: those dependent on data and those not dependent on data. Spatial filters not dependent on data (i.e., Common Average Reference, Surface Laplacian spatial filters) generally use fixed geometric relationships to determine the weights of the transformation matrix. The data-dependent filters, although more complex, allow better results for specific applications because they are derived directly from user's data. They are particularly useful when little is known about specific brain activity or when there are conflicting theories (i.e., theory of valence and theory of the right hemisphere).

2.1.1 Related Works

In this subsection, a State of the Art of the principal works related to emotion detection is reported. All the collected works exhibited at least an experimental sample of 10 subjects. Samples with number of subjects below this threshold were considered not statistically significant. The reported studies are organised in two paragraphs according to the used dataset: public and self-produced. A further paragraph collects analysis on the influencing factors of the experimental setup for the emotion assessment.

Studies based on public datasets. Studies claiming the best accuracy on emotional valence assessment are based on public EEG signal datasets: SEED [111, 112, 113, 114, 115, 116], DEAP [117, 118, 119, 120, 112, 113, 121, 122, 123, 124, 125, 114, 126, 127, 115], and DREAMER [124, 125, 116].

SJTU Emotion EEG Dataset (SEED)[128, 129] is a collection of EEG signals provided by the Center for Brain-like Computing and Machine Intelligence (BCMI laboratory) of the Shanghai Jiao Tong University. EEG data were acquired while 15 participants watched 15 film clips, of about 4 minutes, eliciting positive, neutral, and negative emotions. The videos were selected in order to be understood without explanation, thus an implicit emotion recognition task was employed. The experiment, made of 15 trials, was repeated in 3 different days and EEG signals were recorded through the 62-channel Neuroscan system. Participants filled in the self assessment questionnaire immediately after each trial to report their emotional reactions.
The Dataset for Emotion Analysis using EEG, physiological and video signals (DEAP) [130, 131] is a multimodal dataset developed to analyse human affective states. The EEG and peripheral physiological signals of 32 participants, watching 40 one-minute long music videos were recorded. The EEG signals were acquired through the 32-channel BioSemi device. Participants were informed about the purpose of the experiment, but not further instructions were given, indeed, the emotion recognition task was implicit. Each video was rated in terms of arousal, valence, like/dislike, dominance and familiarity.
The Database for Emotion Recognition through electroencephalogram (EEG) and electrocardiogram (ECG) Signals from Wireless Low-cost Off-the-Shelf Devices (DREAMER) [132, 133] is a multimodal database recorded during emotional elicitation by means of audio-visual stimuli. 18 film clips were employed to elicit: amusement, excitement, happiness, calmness, anger, disgust, fear, sadness, and

surprise. The film clips are long between 65 and 393 s. 23 participants undertook the experiment. Details about the experimental procedure were provided to participants and the rating scales used for emotional assessment were explained. An implicit emotion recognition task was performed since the subjects were not required to get into the target emotional state. Volunteers rated their affective states in terms of valence, arousal, and dominance. EEG signals were captured using the 14-channel Emotiv Epoc+.

A multichannel EEG emotion recognition method based on a Dynamical Graph Convolutional Neural Network (DGCNN) was proposed by Song et al[116]. Experiments were conducted on the 62-channels dataset SEED [134] and on the 14-channels dataset DREAMER [132]. The average accuracies of 90.4 % and 79.95 % were achieved on the SEED dataset for within-subject and cross-subject settings respectively, in a three classes emotion recognition. The average accuracy of 86.23 % was obtained on valence dimension (positive or negative) of the DREAMER dataset in the within-subject configuration.
A Multi-Level Features guided Capsule Network (MLF-CapsNet) was employed by Liu et al. for a multi-channel EEG-based emotion recognition[124]. Valence (positive or negative) was classified with an average accuracy of 97.97 % on the 32-channels DEAP [130] dataset and 94.59 % on the 14-channels DREAMER dataset. Within-subject experiments were performed. Comparable results were obtained by applying an end-to-end Regional-Asymmetric Convolutional Neural Network (RACNN) on the same datasets in a within-subject setup[125].

Studies based on self-produced datasets. EEG signal, acquired through ad hoc experimental activities, are employed in further studies [118, 135, 136]. The main stimuli used to elicit emotions in human subjects are: (i) projection of standardized sets of emotionally arousing images; (ii) viewing audio visuals; (iii) listening to music or sounds; and (iv) recall of autobiographical events. Below, the focus is mainly on studies using standardized image sets (i.e. International Affective Picture System (IAPS)[101], and Geneva Affective Picture Database (GAPED)[137]). The use of a set of normative emotional stimuli (each image is rated according to the valence, arousal and dominance levels) enables to select stimuli eliciting a specific range of emotions.
Mehmood et al. used stimuli from the IAPS dataset to elicit positive or negative valence in 30 subjects[136]. The EEG signals were recorded via an 18 electrolyte gel filled electrodes caps. A feature extraction method, using Hjorth parameters, was implemented. A 70 % cross-subject accuracy was reached using a SVM classifier. Self-assessment tests were not administered to subjects.
More recently, several studies focused on channel reduction for improving the wearability of the emotion detection systems [138, 139, 140, 141, 142, 143, 144, 145, 146, 147].
Marín-Morales at al. designed virtual environments to elicit positive or negative valence[146] . Images from IAPS dataset were used as stimuli. The emotional impact of the stimulus was evaluated using a SAM questionnaire. A set of features, extracted from EEG and ECG signals, was input into a Support Vector Machine classifier obtaining a model's accuracy of 71.21 % along the valence dimension

(binary classification problem). A 10-channel device was used to record the EEG signal from 15 subjects. Sensors' foams were filled with Synapse Conductive Electrode Cream.

The EEG signals of 11 subjects were used to classify valence (positive and negative) by the authors[140]. Pictures from GAPED dataset were used as elicitative stimuli. The accuracy rates of a SVM classifier were 85.41 % and 84.18 % using the whole set of 14 channels and a subset of 10 channels respectively, in the cross-subject setting. EEG signals were acquired through a wet-14 channels device and no self-evaluation questionnaires were used.

Wei et al. proposed a real-time valence emotion detection system based on EEG measurement realized by means of a headband coupled with printed dry electrodes [147]. 12 participants undertook the experiment. Pictures selected from GAPED were used to elicit positive or negative valence. Self-evaluation questionnaires were employed. Two different combinations of 4 channels were tested. In both cases, the cross-subject accuracy was 64.73 %. The highest within-subject accuracy increased to 91.75 % from 86.83 % switching from one configuration to another. The latter two works [140, 147] both proposed the use of standardized stimuli. However, in the first one[140], the concomitant use of self-assessment questionnaires was missing. Moreover, in the second one[147], self-assessment questionnaires were employed but the results were not compared with the scores of the used stimuli. Failure to compare individual reactions with the standardized stimulus scores, negatively impacted on the result of the experiment.

Happy or sad emotions were elicited through images provided by the IAPS, by Ang et al[144]. The EEG signals were acquired through FP1 and FP2 dry electrodes. An Artificial Neural Network (ANN) classifier was fed with discrete wavelet transform coefficients. The best detection accuracy was 81.8 % on 22 subjects. Beyond the use of standardized stimuli, the subjects were also administered self-assessment scales. Moreover it is unclear how the SAM scores were used and whether the approach is within-subject or cross-subject.

Following two studies claiming a single-channel EEG based emotion recognition achieved employing audio-visual stimuli. Ogino et al. developed a model to estimate valence by using a single-channel EEG device[139]. Fast Fourier Transform, Robust Scaling and Support Vector Regression were implemented. EEG signals from 30 subjects were acquired and an average classification accuracy of 72.40 % was reached in the within-subject configuration. Movie clips were used to elicit emotional states and SAMs were administered to the participants for rating the valence score of the stimuli.

A cross-subject emotion recognition system based on Multilayer Perceptron Neural Network was proposed by Pandey et al[143]. An accuracy of 58.5 % was achieved in the recognition of positive or negative valence on DEAP dataset using the F4 channel.

A reduced number of channels implies a low spatial resolution. Traditional strategies for EEG signal feature extraction, combined with a-priori knowledge on spatial and frequency phenomena related to emotions, can be unusable in case of

few electrodes. In a previous work of the Authors, for a single-channel stress detection instrument, a-priori spatial knowledge drove electrodes positioning [148]. However, signal processing was based on innovative and not well-settled strategies. Although proper psychometric tools were adopted for the construction of the experimental sample, the reproducibility of the experiment was adversely affected by the use of not standardized stimuli.
Further not standardized stimuli are personal memories. For example, the study in [145] presents a very interesting data fusion approach for emotion classification based on EEG, ECG, and photoplethysmogram (PPG). The EEG signals were acquired through an 8-channel device. A Convolutional Neural Network (CNN) was used to classify three emotions reaching an average accuracy for the cross-subject case of 76.94 %. However, personal memories of the volunteers were used as stimulus, compromising the reproducibility of the experimental results. Moreover, due to the adoption of the *discrete emotion* model, the study cannot be taken into account for emotion measurement goal.

Influencing factors of the experimental conditions. In the field of emotion recognition, the use of audio-visual stimuli guarantees higher valence intensity (positive or negative) with respect to visual stimuli (pictures) [149]. Therefore, the sensitivity of the measurement system increases and the accuracy in emotion detection can be higher. However, currently there are no standardized audiovisual datasets to employ for eliciting emotions. The only exception is the dataset used by DREAMER, which contains a low number of stimuli (only 18), so penalising their randomic administration and increasing the risk of bias. Not even the most widely used EEG datasets SEED and DEAP employ a standardized stimulus dataset to elicit emotions.
Also the use of explicit rather than implicit tasks affects the effectiveness of the mood induction. Explicit instruction helps participants to get into the target emotional state, but it can be a further source of uncertainty. However, the existing standardized stimuli (IAPS, GAPED, OASIS, etc) are predominantly images characterized in an implicit setup. In order to draw on this resource and make the experiment reproducible, an implicit task, with static images, should therefore be adopted. Among the reported studies, task information is generally omitted.
Another factor that can influence the effectiveness of the emotional state induction is the way of stimuli selection. Referring to the main standardized stimuli datasets, images can be selected by choosing those with higher or lower valence scores. Polarized stimuli could increase the intensity of a certain emotional state with respect to random chosen stimuli.
For all the presented studies (i) type of stimuli, (ii) type of task, (iii) number of channels, (iv) number of participants, (v) classifier, (vi) number of classes, (vii) within-subject accuracy, and (viii) cross-subject accuracy are reported in Table 2.1.
The accuracy values are reported in both the within-subject and cross-subject cases, when available. In the first case, classification was carried out using data of a single subject both for training and test phases, while in the second one, classification was carried out employing the data set as a whole.

TABLE 2.1: Studies on emotion recognition classified according to the employed datasets (i.e. SEED, DEAP, and DREAMER), stimuli (*v*="video", *p*="picture", *m*="memories"), task (*i*="implicit", *e*="explicit", n.a.="not available"), #channels, #participants, #classes, classifiers, and accuracies (n.a.="not available").

Dataset	Study	Stimuli	Task	#channels	#participants	Classifier	#classes	Within-subject accuracy (%)	Cross-subject accuracy (%)
SEED	[111]	*v*	*i*	62	15	SincNet-R	3	94.5	90.0
	[116]	*v*	*i*	62	15	DGCNN	3	90.4	80.0
SEED & DEAP	[112]	*v*	*i*	62	15	DNN	3	n.a.	96.8
		v	*i*	32	32		2	n.a.	89.5
	[113]	*v*	*i*	62	15	SNN	3	n.a.	96.7
		v	*i*	32	32		2	n.a.	78.0
	[123]	*v*	*i*	62	15	SBSSVM	2	n.a.	72.0
		v	*i*	32	32		2	n.a.	89.0
	[114]	*v*	*i*	62	15	CNN	3	90.6	n.a.
		v	*i*	32	32		2	82.8	n.a.
	[115]	*v*	*i*	62	15	CNN	2	n.a.	86.6
		v	*i*	32	32		2	n.a.	72.8
DEAP	[117]	*v*	*i*	32	32	H-ATT-BGRU	2	n.a.	69.3
	[119]	*v*	*i*	32	32	CNN	2	n.a.	77.4
	[120]	*v*	*i*	4	32	LDA	2	n.a.	82.0
	[122]	*v*	*i*	32	32	LSTM-RNN	2	n.a.	81.1
	[126]	*v*	*i*	32	32	Kohonen-NN	2	76.3	n.a.
	[127]	*v*	*i*	32	32	SVM + FCM	2	78.4	n.a.
	[143]	*v*	*i*	1	32	MLP	2	n.a.	58.5
DEAP & DREAMER	[121]	*v*	*i*	32	32	BioCNN	2	83.1	n.a.
		v	*i*	14	23		2	56.0	n.a.
	[124]	*v*	*i*	32	32	MLF-CapsNet	2	98.0	n.a.
		v	*i*	14	23		2	94.6	n.a.
	[125]	*v*	*i*	32	32	RACNN	2	96,7	n.a.
		v	*i*	14	23		2	97,1	n.a.
SELF-PRODUCED	[116]	*v*	*i*	14	23	DGCNN	2	86.2	n.a.
	[118]	*v*	*i*	19	40	MLP, KNN, and SVM	2	n.a.	90.7
	[138]	*v*	*n.a.*	1	20	MC-LS-SVM	2	n.a.	90.6
	[135]	*v*	*n.a.*	14	10	RVM	2	91.2	n.a.
	[139]	*v*	*n.a.*	1	30	SVM	2	72.4	n.a.
	[141]	*v*	*i*	1	19	*k*-NN	3	94.1	n.a.
	[142]	*p*	*e*	3	16	SVM	6	n.a.	83.3
	[140]	*p*	*n.a.*	10	11	SVM	2	n.a.	84.2
	[136]	*p*	*n.a.*	18	30	SVM	2	n.a.	70.0
	[144]	*p*	*n.a.*	2	22	ANN	2	n.a.	81.8
	[146]	*p*	*n.a.*	10	38	SVM	2	n.a.	71.2
	[147]	*p*	*n.a.*	4	12	LDA	2	86.8	64.7
	[145]	*m*	*e*	8	20	CNN	3	n.a.	76.9

2.1.2 Statement of the metrological problem

The path towards the measurability of emotions still remains to be completed. In this study, some important steps are carried out to achieve this goal:

- a theoretical model compatible with emotion measurability was adopted;
- people with high scores on the Patient Health Questionnaire (PHQ) were excluded from the experimental sample in order to soften the bias of depressive disorders;
- standardized stimuli were used jointly with self-assessment questionnaires to reduce the intrinsic uncertainty of the measurand;

Nevertheless, there are still several aspects to continue working on:

- a more complete definition of an emotion model, which incorporates, for example, appropriately adjusted analyses for confounders including the impact of individual personality on the specific emotional response;

- identification of a measurement unit (enhancing the important role played in this direction by biosignals, including the EEG);
- an uncertainty analysis for identifying and weighing the sources in the measurement processes. Just to remember a few: (i) the theoretical model, (ii) the stimulus, (iii) the task, (iv) the specific individual emotional response, (v) the peculiar relationship between the individual emotional response and its manifestation in terms of neurosignal, (vi) the signal acquisition instrument, and (vii) the algorithms for signal classification.

2.2 Measurement of attention

Many studies deal with assessing the attention and its different dimensions through the analysis of the brain signals using the ElectroEncephalography (EEG) [23]. EEG is the most used technique because of its high temporal resolution, non-invasiveness, and low cost. Several studies have shown that the level of attention affects the EEG signal [66, 67]. Therefore, variations in the EEG signal can be used to detect corresponding changes in attention levels [68]. Attention creates a variation in brain signals that can be assessed both in the time and in the frequency domain [69]. In what follows, the analysis focuses on measuring attention during motor rehabilitation activities. In fact, Chapter 3 presents a solution for robot-based adaptive motor rehabilitation.

Most of the studies in the rehabilitation sector adopted a within subject approach for training the classifiers in distraction detection. Asayb et al. in 2017 [59] proposed to assess the attention during the flexion-extension of the ankle in presence of auditory distractors. Using a 18-channel system and wet electrodes on 12 participants, they obtained an average accuracy of 71 %, by extracting time-frequency features from 1.5-s epochs. Hamadicharef, Brahim, et al. [150] proposed an interesting processing system (already widely used in the EEG field for Motor Imagery) for assessing the attention, during a cognitive task with eyes closed and opened. This processing involves a Filter-Bank in relation to the Common Spatial Pattern. A 15-channel EEG system achieves an average accuracy of 69.2 % on five subjects with a 2-s time window. Antelis, et al. [151] proposed the distraction detection during robot-assisted passive movements of the upper limb. Six patients were connected to a 32-channels EEG by wet electrodes and to the robot's end-effector for assisted passive movements. They got an average accuracy of 76.37 % in classifying 3-s epochs, when mentally count back in threes, starting in a self-selected random three-digit number, assured the distraction condition. In 2019 Asayb et al. [65] proposed an upgrade of their previous work using a 28-channel EEG system and wet electrodes. Three different distractors characterized the experimental set-up. Signal processing was based on spectro-temporal features extracted from 3-s epochs. The obtained average accuracy was 85.8 % by exploiting the motor-related cortical potential. However, in this state of the art, an appropriate approach for clinical application seems to be missing. The high number of channels and the use of wet or semi-wet electrodes penalize the wearability, limiting the clinical usability.

2.3 Measurement of engagement

Two applications field for engagement measurement are presented below. Subsection 2.3.1 focuses the *learning* contex, while in Subsection 2.3.2, the *pediatric neuro-motor rehabilitation* application field is presented.

2.3.1 Engagement assessment during learning activities

In learning activities, evaluation grids and self-assessment questionnaires (to be filled out by the observer or by the learner autonomously) are traditionally the most used methods for the behavioral, cognitive, and emotional engagement detection [38]. In recent years, measures based on biosignals are spreading very rapidly. Furthermore, the use of physiological sensors able to detect cognitive and emotional engagement allows the real-time machine adaptive strategies. Among the different physiological biosignals, the EEG appears to be one of the most promising technology thanks to its low cost, low invasiveness, and high temporal resolution. Moreover, the EEG contains a broader range of information about the state of a subject with respect to others biosignals [39]. In 1995, authors in [152] proposed an engagement index to decide when to use the autopilot and when to switch to the manual one during a fly simulator session. The engagement index was $E = \frac{\beta}{\theta+\alpha}$ where $\alpha, \beta,$ and θ are the EEG frequency bands in (8-13) Hz, (13-22) Hz, and (4-8) Hz respectively.

Several studies used this index as engagement estimator also in learning contexts [153, 38, 154]. However, the proposed index does not take into account the different engagement types (i.e., cognitive, emotional and behavioural) proposed by the theories previously reported in Section 1.3. Different methods for learning engagement detection are proposed in literature [154]. For the behavioral engagement assessment, observation grids (used to support direct observations or video analysis) were proposed [155, 156]. For the cognitive and emotional engagement assessment, self-assessment questionnaires and surveys (compiled autonomously by the user) were developed [157, 158]. In recent years alternative engagement assessment methods based on physiological sensors have established: heart-rate variability, galvanic skin response, and EEG. Among these biosignal, the most promising for engagement assessment is the EEG. As already described, the learning is based on a neurological changes set, and the EEG presents the possibility of studying these the neural modification [38, 29, 159, 31, 160]. The EEG system is low-cost and non-invasive, and provides information on brain activity within milliseconds. It is now commonly used in many application [161, 24] including the cognitive and emotion engagement assessment as well as the detection of the underlying elements: emotions recognition and cognitive load activity assessment respectively [162, 163, 164, 165, 166, 167, 168].

To achieve a correct metrological reference of the EEG-based cognitive and emotional engagement constructs, a reproducibility problem arises. From emotional point of view, when eliciting a specific emotion, the same stimulus does not often induce the same emotion in different subjects. The effectiveness of the induction can be verified by means of self-assessment questionnaires or scales.

The combined use of standardized stimuli and subject's self-assessment ratings can be an effective way to build a metrological reference for a reliable EEG-based emotional engagement detection [102]. From the cognitive point of view, when the subject is learning, the working memory identifies the incoming information and the long-term memory constructs and stores new schemes on the basis of the past ones. While the already built schemes decrease in the working memory load, the construction of new schemes entails its increase [39, 169]. Therefore, increasing difficulty levels allows to induce different cognitive states; the cognitive engagement level grows up according to the proposed exercise difficulty increases.

2.3.2 Engagement detection in pediatric rehabilitation

The standard tools used in clinical practice for engagement assessment are questionnaires or rating scales. Both take into account the patients' awareness of their health and their therapeutic process. In adult rehabilitation, the most used are: Patient Activation Measure (PAM-13) [170] and Patient Health Engagement (PHE) scale [171]. Recently, also in pediatric rehabilitation, engagement assessment scales have been developed and validated. The Pediatric Rehabilitation Intervention Measure of Engagement-Observation (PRIME-O) version [172] and the Pediatric Assessment of Rehabilitation Engagement (PARE) scale [173] were designed to capture signs of emotional, cognitive, and behavioral engagement for clients and service providers and in the client-provider interaction. Beyond standard tools, biosignals-based measurement methods are emerging. They allow an automated and real-time engagement assessment. In particular, eye-blinking [25], heart rate variability [26], and brain activity [27, 28] were used to detect changes in patient's engagement. Among these, the EEG signal [29] offers good temporal resolution and improves real-time performances.

In the rehabilitation field, studies on EEG-based engagement detection were mainly conducted on adults and focused only on cognitive engagement. In [174], a computational framework was proposed for real-time cognitive engagement (CE) recognition using electroencephalography (EEG). A deep Convolutional Neural Network was used to extract task discriminative spatio-temporal features and predict the CE level for two classes: engaged vs. disengaged. Experiments were conducted on 8 subjects performing the Go/No-Go paradigm to induce cognitive fatigue. An average inter-subjective accuracy of 88.13% was reached. In [31], the EEG signals were acquired for monitoring cognitive engagement in stroke patients while they executed active and passive motor tasks. Event-related desynchronization differences between tasks were observed during both initial and post-movement periods. EEG data were used to classify each epoch as involving the active or passive motor task. Average classification accuracy was 80.7 $\pm$ 0.1% for grasping movement and 82.8 $\pm$ 0.1% for supination movement.

Recently, a first study on engagement in pediatric rehabilitation was proposed [175]. Positive/negative engagement of autistic patients was classified starting from EEG signals and gesture recognition. The EEG signals were acquired through the single-channel MindWave; Kinect was instead employed for gesture recognition. Five children (two with autism) undertook the experiment. An intersubjective accuracy of 95.8% was achieved in classifying positive or negative engagement. However, the study does not specify the explored engagement dimensions (i.e. emotional, cognitive, or behavioral). To date, to the best of our knowledge, only one study is present in the literature on this topic. The reasons could be: (i) the engagement measure in the rehabilitation field has only recently become an object of interest [31], and (ii) EEG-based engagement assessment in pediatric rehabilitation requires the adoption of a respectful clinical protocol to protect the child and his psycho-physical integrity (i. e. a non-interventional observational approach). Although such a protocol is more comfortable for the children, it entails a general lack of control over the engagement levels resulting in imbalanced data collections.

2.4 Measurement of stress

Several methods for stress assessment, like self-assessment scale, or questionnaires, follow a psychological approach [176]. As an example, in human-robot interaction, questionnaires to analyze the psychological effect of cycle time on operators [177] highlighted frustration, effort, and a dissatisfaction feeling about own performance.

As more direct and objective tools for stress detection, biosignals have been proposed in several studies [178]. Physiological parameters, as EEG signals, blood volume pulse (BVP), electro-oculogram (EOG), salivary cortisol level (SCL) [179], heart rate variability (HRV) [180], galvanic skin response (GSR), or electromyography (EMG) are assessed [45].

Compared to other biosignals, EEG proved better latency and robustness to artifacts due to physical activity [47] [181]. In Industry 4.0 scenarios, EEG has been widely applied to assess individuals' stress in workplace in order to improve workers' safety, health, well-being, and productivity [48][182] [183]. Thanks to the ease of application and removal, dry electrodes are increasingly used to reliably search human cognitive states in real-life conditions. They guarantee the quality of the EEG signal which approaches the wet sensors, as demonstrated in [5, 6]. Beside that, EEG is regarded as one of the most reliable and effective techniques for identifying fatigue and monitoring stress level in drivers [184, 46, 185]. The strategies for EEG-based stress detection can be divided into two main approaches: *data-driven* and based on *handcrafted features*. In the first approach, the study is linked to the neurophysiological basis of the phenomenon. For data-driven approaches, the brain remains a black-box. They learn the appropriate features of the signal using EEG data according to the criterion of maximizing the accuracy of the classification.

2.4.1 Methods based on handcrafted feature for stress detection

Systematic alterations in frontal EEG asymmetry, in response to specific emotional stimuli, can be exploited to analyze emotional response [109]. In particular, EEG asymmetry proved to be capable of predicting state-related emotional changes and responses. For example, a greater self-reported happiness or positively-valued stimuli might be expected to be associated with greater relative left frontal activity. Therefore, greater relative right frontal activity would be expected in response to negative stimuli [186],[187]. However, fear or happiness response to stimuli may either be attenuated or amplified according to any given individual's trait pattern of frontal EEG asymmetry [186].

Different models were presented: Baron and Kenney linear model may predict individual's response to fear relevant stimuli. According to the relative difference between the left and right hemisphere, the EEG asymmetry may serve both to amplify and attenuate the effect of the fear relevant stimuli. Some individuals show the increase of relative right versus left sided activity in response to negative cues and the increase of left versus right sided activity to positive cues [188]. Coan and Allen [107] presented another linear model to predict emotional experience using emotion type and trait frontal EEG asymmetry. The frontal EEG asymmetry may serve as a useful liability marker also for depression and anxiety [189]. Many works, using EEG caps with a limited number of electrodes, demonstrated that stress causes changes in regions of prefrontal and frontal areas [179] [47] [186].

2.4.2 Data-driven methods for stress detection

Different classification methods try to face the main problems of EEG signals, including the low signal-to-noise ratio, their non stationarity over time within or between users, and the limited amount of training data typically available to calibrate the classifiers[190].

A large number of informative and measurable properties (features) of EEG signals, can be used both in time and frequency domains. Their accurate selection is crucial for the accuracy and the computational cost of classification [191]. Both types of features reap the benefit from being extracted after spatial filtering. Independent Component Analysis and Canonical Correlation Analysis are useful methods for muscle artifact removal in EEG data [192, 193]. Several supervised learning algorithms could be exploited to assess workers stress by using subjects' EEG signals. The classification can be assisted by: linear classifiers, neural networks, non-linear Bayesian classifiers, nearest neighbour classifiers, random forest, naive bayes, and decision tree [194][195]. Linear classifiers are the most popular algorithms for Brain Computer Interface (BCI) applications, such as, Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM). The LDA is used to assess mental fatigue in [196], it divides the data into hyperplanes representing the different classes, with very-low computational burden. A discrimination based on hyperplanes was also used in SVM, with recognition rate of 75.2% to identify three different level of stress out of four, using EEG features and six statistical features in [197]. Meanwhile a better prediction accuracy of

TABLE 2.2: State of art of stress classification

Classifier	Reference	Reported Accuracy%	Acquired Signals	Classes	n° Electrodes	n° Subjects
Artificial Neural Network (ANN)	[184]	76.0%	EEG,ECG,GSR	2 no-stress/stress	14 Wet	22
	[195]	79.2%	EEG,SCL,BVP,PPG	2 levels of stress	5 Wet	15
Cellular Neural Network (CNN)	[184]	92.0%	EEG,ECG,GSR	2 no-stress/stress	14 Wet	22
Decision Tree	[184]	84.0%	EEG,ECG,GSR	2 no-stress/stress	14 Wet	22
Fisher linear discriminant analysis (FLDA)	[185]	90.5%	EEG,EOG	2 alert and fatigue states	32 Wet	8
Gaussian Discriminant Analysis (GDA)	[48]	74.9%	EEG,GSR	2 high or low stress level	14 Wet	11
K-Nearest Neighbors	[48]	65.8%	EEG,GSR	2 high or low stress level	14 Wet	11
(k-NN)	[197]	76.7%	EEG	2 levels of stress	14 Wet	9
Linear discriminant analysis (LDA)	[196]	77.5%	EEG	3 low, medium, high mental fatigue	16 Wet	10
	[47]	86.0%	EEG,ECG,EMG,GSR	3 stress,relax,and neutral	4 Wet	10
Naive Bayes (NB)	[184]	77.0%	EEG,ECG,GSR	2 no-stress/stress	14 Wet	22
	[199]	69.7%	EEG,ECG,GSR	2 mental workload and stress	2 Wet	9
Random Forest (RF)	[200]	79.6%	EEG,EMG,ECG,GSR	4 cognitive states	8 Wet	12
	[194]	84.3%	EEG,ECG,BVP	3 mental stress states	14 Wet	17
Support vector machine (SVM)	[197]	75.2%	EEG	3 levels of stress	14 Wet	9
	[48]	80.3%	EEG, SCL	2 high or low stress level	14 Wet	11
	[140]	85.4%	EEG	2 positive or negative emotion	14 Wet	11
	[50]	87.5%	EEG,ECG,HRV	2 stress and rest	2 Wet	7
	[198]	88.0%	EEG	2 levels of stress	14 Wet	10
	[199]	90.1%	EEG,ECG,GSR	2 mental workload and stress	2 Wet	9

90.5% and 92%, combining different acquired signals, were reported in [184, 185] for drivers. Various supervised machine learning algorithms, using sliding and fixed windowing procedures, were tested in [48]: k-Nearest Neighbors, Gaussian Discriminant Analysis, SVM with different similarity functions (linear, Gaussian, cubic, and quadratic). Among the state-of-the-art classifiers, the SVM yielded the highest classification accuracy of 90.1% [198], using a single-channel EEG. As well as, the highest accuracy of 88.0% was reached by SVM in [199], where individuals' stress was recognized by exploiting only EEG signal as input of the classifier. Tab 2.2 summarizes the reported accuracy of different classifiers, including the numbers and type of EEG electrodes, without reference electrodes, the numbers of different classes according to the acquired bio-signals used as model input.

Part II

Proposal

Chapter 3

Basic mental state assessment

In this chapter methods and experimental validations for emotion and attention assessment are presented

3.1 Emotional valence detection

This Section presents an emotional valence detection method starting from the EEG signal acquired through few dry electrodes. The basic ideas,the architecture, the data processing, and the experimental validation of the proposed approach are presented.

3.1.1 Basic ideas

Below the basic ideas are reported.

- *An EEG-based method for emotional valence detection*: Emotional functions are mediated by specific brain circuits and electrical waveforms. Therefore, the EEG signal varies according to the emotional state of the subject. However, using suitable algorithms, such a state can be recognized.

- *Low number of channels, dry electrodes, wireless connection for a good ergonomics*: An 8 channel-dry electrode device does not require a burdensome installation. The absence of the electrolytic gel eliminates the problem of residues in the hair. The good ergonomics of the instrument is also guaranteed by the absence of connection cables and, therefore, by the wireless transmission of the acquired signals. Both of them simplify the operator's job.

- *Multifactorial metrological reference*: A multifactorial metrological reference was implemented. Images belonging to a statistically validated dataset were used as stimuli for eliciting emotions. Therefore, each image is scored according to the corresponding valence value. The metrological reference of the emotional valence is obtained by combining the scores of the stimuli (statistically founded) with the score of the self-assessment questionnaires (subjective response to the standardized stimulus).

  The Bland-Altman and the Spearman analysis were carried out for comparing Self-assessment questionnaires (SAM) scores and the OASIS dataset scores.

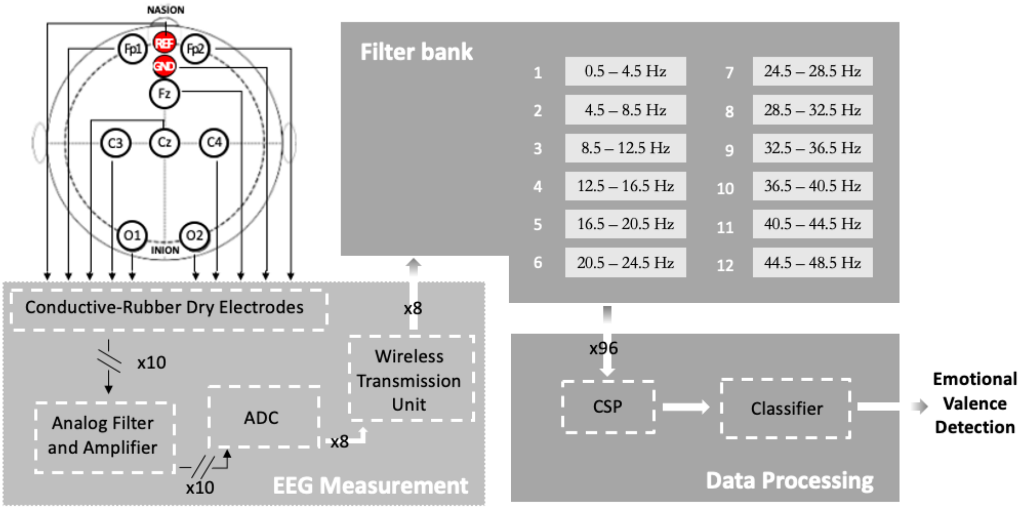


FIGURE 3.1: The proposed valence-detection method (CSP: Common Spatial Pattern algorithm).

- *12–band Filter Bank*: Traditional filtering, employed to extract the information content from the EEG signals, is improved by a 12-band Filter-Bank. Compared to the five typical bands for EEG analysis (alpha, beta, delta, gamma, theta), narrowing the frequency intervals, the features resolution increases.
- *Beyond a priori knowledge*: A supervised spatial filter (namely CSP) guarantees automated feature extraction from spatial and time domains.

3.1.2 Architecture

The architecture of the proposed system is shown in Fig. 3.1. The *conductive-rubber dry electrodes* allow the EEG signals to be sensed directly from the scalp of the subject. Each channel is differential with respect to AFz (REF) and referred to Fpz (GND). Analog signals are conditioned by stages of amplification and filtering (*Analog Filter and Amplifier*). Then, they are digitized by the Analog Digital Converter *ADC* and sent by the *Wireless Transmission Unit* to the *Data Processing* block. A 12-bands *Filter Bank* and a *Common Spatial Pattern* (*CSP*) algorithm carry out the feature extraction. The *Classifier* receives the feature arrays and detects the emotional valence.

3.1.3 Data processing

In this section, the *features extraction and selection* and the *classification* procedures of the proposed method are presented.

Features extraction and selection

Finer-resolution partitions of the traditional EEG bands were proposed for emotion recognition[201, 202]. In the present work, a 12-band Filter Bank version, recently adopted in distraction detection [24], is employed.
Spatial and frequency filtering is applied to the output data of the filter bank. A well-claimed Common Spatial Pattern (CSP), widely used in EEG-based motor imagery classification [203, 204, 205, 206], is used as a spatial filter. For the first time, the FB-CSP pipeline is here proposed in the field of valence emotion detection.
A previous study [207] showed that the CSP spatial filtering method entails the relationship between EEG bands, EEG channels, neural efficiency and emotional stimuli types. It demonstrated that CSP spatial filtering gives significant values on band-channels ($p < 0.004$) combination. Spatial characteristics may provide more relevant information to distinguish different emotional states. A feasibility study demonstrated the CSP capability of applying spatial features to EEG-based emotion recognition reaching average accuracies of 85.85 % and 94.13 % on the self-collected and MAHNOB-HCI datasets. Three emotion tasks were detected with 32 EEG channels[208].
In a binary problem, the CSP computes the covariance matrices of the two classes. By means of a whitening matrix, the input data are transformed in order to have an identity covariance matrix (mainly, all dimensions are statistically independent). Resultant components are sorted on the basis of variance in order: (i) *decreasing*, if the projection matrix is applied to inputs belonging to class 1, and (ii) *ascending*, in case of inputs belonging to class 2. In this way, according to the "variance of each component", data can be more easily separable [209]. The CSP receives as input 3D tensors with dimensions given by the number of channels, filters, and samples.

Classification

In this study, the emotional valence is classified using a k-Nearest Neighbors (k-NN) [210] for cross-subject case and full-connected Artificial Neural Networks (ANNs) [211] for within-subject one. One of the main advantages of the k-NN is that, being non-parametric, it does not require a training phase unlike other Machine Learning methods. In a nutshell, given a set of unlabelled points P to classify, a positive integer k, a distance measure d (e.g., Euclidean) and a set D of already labelled points, for each point $p \in P$, k-NN assigns to p the most frequent class among its k neighbours in D according to the measure d. The number of neighbours k and the distance measure d were set using a cross-validation procedure. Differently from k-NN, ANNs are classification models that require a training procedure. In general, an ANN consists of a set of elements (called *neurons*) arranged together into several layers fully connected between them. Each neuron performs a linear combination of its inputs usually followed by the application of a non-linear function called *activation function*. It was demonstrated [212] that an ANN can approximate arbitrarily complex functions, giving to the model the ability to discriminate between different classes. The number of neurons, the number of layers and the activation functions are hyperparameters given a priori, while the coefficients of each linear combination are learned by the model in a training stage.

3.1.4 Experiments and Results

Data acquisition setup

The experimental protocol was approved by the ethical committee of the University Federico II. Written informed consent was obtained by the subjects before the experiment. All methods were carried out in accordance with relevant guidelines and regulations. Prior informed consent for publication of identifying information and images was obtained by all the participants. Thirty-one volunteers, not suffering from both physical and mental pathologies, were screened by means of the Patient Health Questionnaire (PHQ) for excluding depressive disorders [213]. Six participants were excluded from the experiment owing to their score in PHQ, resulting in twenty five healthy subjects, (52 % male, 48 % female, aged 38 $\pm$ 14). The experiments were conducted in a dark and soundproofed environment to prevent disturbing elements.
The employed Mood Induction Procedure (MIP) was based on the presentation of emotion-inducing material to participants to elicit suitable emotions. The subjects were instructed on the purpose of the experiment. They had to passively gaze at the pictures projected on the screen and, only after, to assess the experienced valence by two classes: negative and positive. Emotional stimuli were presented without explicitly instructing subjects to get into the suggested mood state and regulate their emotions. Nevertheless, the subjects were aware of both the elicitation stimulus and the type of induced emotion (although it was not explicitly stated, they could guess it starting from the self-assessment questionnaire). Thus, the employed task was of a type *implicit-more controlled* [214]. The experiment was made of 26 trials. Each trial lasted 30 s and consisted of: (i) a 5-s white screen, (ii) a 5-s countdown frame employed to relax the subject and separate emotional states mutually, (iii) a 5-s elicitative image projection, and (iv) a 15-s self-assessment (Fig.3.2). The subject was required to express a judgement on the positivity/negativity of his/her valence on a scale from 1 to 5 through the self-assessment manikin (SAM) questionnaire. In each trial, different images were projected, for a total of 26 images. 13 pictures for eliciting negative valence and 13 for eliciting positive valence were employed. Positive and negative tasks were randomly administered to participants in order not to create expectations in the tested subjects.

Images were chosen from the reference database Oasis [215]. Oasis attributes a valence level to each image on a scale from 1.00 to 7.00.

Only Italian volunteers participated the experiment, thus a pre-test on the trans-cultural robustness of the selected images was administered to a different group consisting of 12 subjects. Specifically, suitable pictures were shown and was asked subjects to rate each image using the scale "self assessment manikin" (SAM). Images with a neutral rating from at least 50 % of the subjects were excluded from the experiment. In fact, a stimulus strongly connoted in a specific

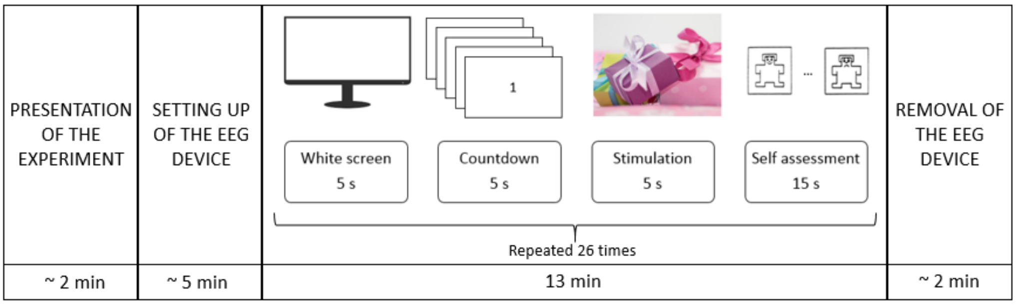


FIGURE 3.2: Experimental protocol.

cultural framework, loses its strength out of that context. An emblematic example are the symbols related to the Ku Klux Klan. Those have a different connotative richness for a citizen of the United States of America compared to European people. The same pre-test revealed very low performances for detecting valence level when the stimuli score was around the the midpoint value of the valence scale. The sensitivity of the system was improved by selecting a suitably polarised subset of Oasis images, as in [136] and [140]. First of all, images with highest and lowest valence score were identified: respectively 6.28 and 1.32. Then, 1.00 was the span chosen to guarantee the trade-off between the maximum image polarization and an adequate quantity of images to build the experiment (>100). Therefore, [1.32, 2.32] and [5.28, 6.28] were adopted as the scoring intervals for negative and positive stimuli valence, respectively and 13 images per group were randomly selected. For each image, the Oasis valence score and the average scores (on all subjects) of the self-assessment are shown in Fig.3.3. The maximum difference between the SAM and the stimuli scores is lower than the average standard deviation (1.00) computed on the Oasis scores.

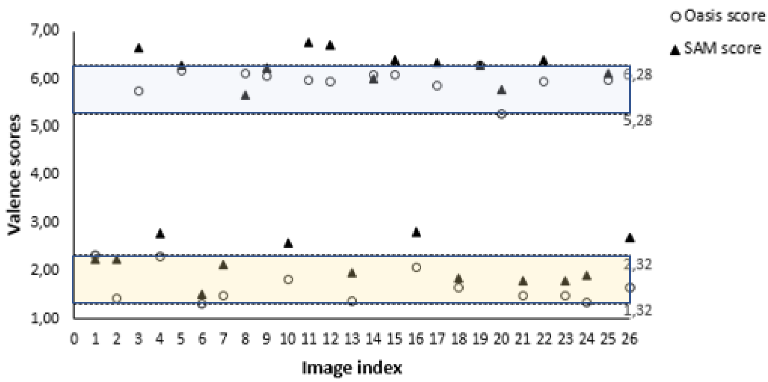


FIGURE 3.3: Oasis valence score and SAM average scores of the 26 images selected for the experiments. The Oasis score intervals used to extract polarized images are identified by dotted lines.

The number of images per class was chosen in order to guarantee a trade-off between the amount of experimental epochs and the user comfort, by minimizing the duration of the experiment simultaneously. In this way, the experiment lasted about 20 min per subject. About 2 min were required for the presentation of the activity to the subject, other 5 min were required for the setting up of the EEG device quality. 13 min were required for the completion of all the 26 trials.

Bland-Altman and Spearman analyzes were carried out to compare the experimental sample with respect to the Oasis experimental sample. The agreement between the measurements expressed by the two samples is verified, as evidenced by a qualitative analysis in Fig. 3.4 and the Spearman correlation index $\rho = 0.799$.

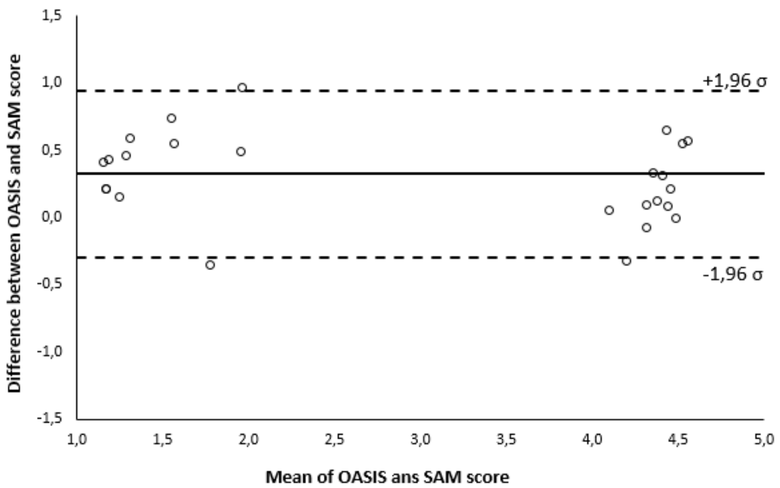


FIGURE 3.4: Bland-Altman analysis on the agreement between stimuli (OASIS) and volunteers perception (SAM)

Hardware

The position of the used channels was chosen by taking into account the well-assessed theories of emotions already presented: frontal asymmetry and right hemisphere asymmetry [103, 108, 104, 105]. The *ab medica Helmate* [216] was found to fit the requirements of the previous mentioned theories because it is equipped with 3 frontal, central, and occipital channels pairs. Indeed, the coverage of almost all areas of the scalp ensured that both frontal and hemispheric asymmetries were recorded, despite the low number of electrodes. The device provided electrodes placed on Fp1, Fp2, Fz, Cz, C3, C4, O1, and O2, according to the 10/20 International Positioning System. The Helmate is Class IIA certified according to Medical Device Regulation (UE) 2017/745 (Fig. 4.6 A). It is provided with a rechargeable battery and is able to transmit the acquired data via Bluetooth, without connection cables. This ultra-light foam helmet is equipped with 10 dry electrodes which 8 acquisition channels (unipolar configuration) and with disposable accessories (under-helmet and under-throat). Electrodes are made of conductive rubber and their endings are coated with Ag/AgCl. They have different shapes to pass through the hair and reach the skin (Fig. 4.6 B).

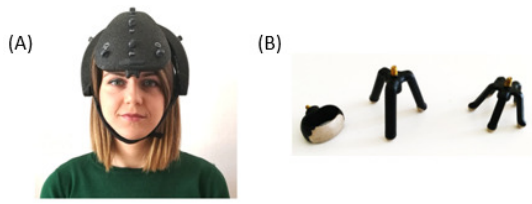


FIGURE 3.5: (A) EEG data acquisition system *Helmate8* and (B) Dry electrodes from *abmedica*

The resulting signals are recorded differentially vs ground (Fpz), and then referenced with respect to AFz, both placed in the frontal region. A dedicated software measures the contact impedance between the electrodes and the scalp. The acquired EEG signal, sampled at 512 Hz, is sent to the Helm8 Software Manager. It allows both to display the signal directly on PC in real time and to apply a large variety of pre-processing filters. The device has an internal μ SD for backup purposes. Helmate incorporates a Texas Instruments analog front-end, the ADS1298 [217]. This is a multichannel, simultaneous sampling, 24-bit, ($\Delta\Sigma$) analog-to-digital converter (ADCs) with built-in programmable gain amplifiers (PGAs), internal reference, and an onboard oscillator. Main features of the ADS1298 are: (i) eight Low-Noise PGAs and Eight High-Resolution ADCs; (ii) input-Referred Noise: 4 μVPP (150 Hz BW, G = 6); (iii) input Bias Current: 200 pA; and, (iv) CMRR: –115 dB.

Data processing comparison

The EEG tracks were acquired at a sampling frequency of 512 Hz and filtered between 0.5 and 48.5 Hz using a zero-phase 4^{th}-order digital Butterworth filter. In the processing stage, the used trials resulted to be 24 for each subject since macroscopic artifacts corrupted one trial of three subjects. So, to keep the dataset balanced, the number of trials was reduced by removing the compromised trial and another one randomly chosen among those of the opposite class. Then, for the remaining subjects, two trials of different classes were randomly removed to guarantee the same amount of data for all the participants. The remaining artifacts were removed from EEG signals using Independent Component Analysis (ICA) by means of the EEGLAB Matlab toolbox version 2019[218]. The recorded EEG signals were divided into 2 s time windows overlapping of 1 s.

The traditional EEG bands delta (0-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (>30 Hz) were extracted. The proposed method was validated by comparing different approaches of features extraction and classification. For EEG features extraction, two different methods were adopted: (i) with and (ii) without a priori spatial-frequency knowledge provided by neurophysiology.

In a-priori spatial knowledge framework, frontal asymmetry feature was chosen, computed by subtracting the left frontal (FP1) from the right (FP2) channel. Moreover, the whole hemispherical asymmetry was also considered and the differences of the three symmetric channel pairs were input to the classifiers. The analysis considered only spatial or both spatial and frequency features, according to the different neurophysiological theories. A-priori frequential knowledge led to the use of a [8-13] Hz (alpha band) pass-band filter (zero-phase 4^{th}-order digital Butterworth filter).

Without a priori knowledge, features were extracted via the PCA and CSP algorithms. For PCA, we used a number of components which explains the 95 % of the data variance. For CSP, all the 96 components returned by the algorithm are used. Also in this case, only the spatial information and the combination of spatial and frequency information were analysed. Input features were 8192 (8 channels * 1024 samples) when PCA and CSP were fed only by spatial information.

The acquired EEG signal was filtered through 12 IIR band-pass filters Chebyshev type 2, with 4 Hz bandwidth, equally spaced from 0.5 to 48.5 Hz. In this way, the traditional five EEG bands (delta, theta, alpha, beta, and gamma) are divided into 12 sub-bands. Therefore, the features resolution is increased by the narrowing of the bands. Thus, features increased to 98304 (12 frequency bands * 8 channels * 1024 samples). The features were then reduced from 98304 to 96 using the CSP algorithm.

Subsequently, in the classification stage, two types of investigations were carried out: within-subject and cross-subject. In the first case, data of a single subject were employed for training and classification phases, while in the second one, the data set as a whole was employed. In both cases, the proposed method was validated through a stratified 12-fold Cross Validation (CV) procedure. Namely, given a combination of the classifier hyperparameters values, a partition of the data composed of K subsets (folds) is made, preserving the ratio between the samples of different classes. A set T consisting of $K-1$ folds is then used to train the model and, when required, the CSP projection matrix; the remaining fold E to measure the model performances using any metric scores (e.g., accuracy). The whole process is then repeated for all the possible combinations of the K folds. Finally, the average scores on all the test sets are reported. Furthermore, training and test sets are made keeping together the epochs of each trial (consisting of 4 epochs each) in the same set, both in the cross-subject and in the within-subject approach. In this way, the training and the test sets do not include parts of the same trial. Finally, in a 12-fold scheme within-subject setup, 88 epochs for training and 8 epochs for testing are used. Of the 88 epochs used for the training set, 16 are exploited as validation set in the ANNs learning. Instead, in the cross-subject case, considering that the experimental campaign involved 25 subjects, a total of 2400 epochs was used. This, in a 12-fold cross validation scheme, corresponds to 2200 epochs as training test and 200 epochs as test set. In the ANNs learning, 200 epochs are used as the validation set.

k-NN[210] and ANN [211] were compared with other four classifiers: Linear Discriminant Analysis (LDA)[219], Support Vector Machine (SVM)[220], Logistic Regression (LR) [211] and Random Forest (RF) [206]. LDA searches for a linear projection of the data in a lower dimensional space trying to preserve the discriminatory information between the classes contained in the data. A SVM

TABLE 3.1: Classifier optimized hyperparameters and variation range

Classifier	Hyperparameter	Variation Range
k-Nearest Neighbour (k-NN)	Distance (DD)	{cityblock, chebychev, correlation, cosine, euclidean, hamming, jaccard, mahalanobis, minkowski,spearman}
	DistanceWeight (DW)	{equal, inverse, squaredinverse}
	Exponent (E)	[0.5, 3]
	NumNeighbors (NN)	[1, 5]
Support Vector Machine (SVM)	BoxConstraint (BC)	log-scaled in the range [1e-3,1e3]
	KernelFunction (KF)	{gaussian, linear, polynomial}
	KernelScale (KS)	log-scaled in the range [1e-3,1e3]
	PolynomialOrder (PO)	{2,3,4}
Artificial Neural Network (ANN)	Activation Function (AF)	{relu, sigmoid, tanh}
	Hidden Layer nr. of Neurons (HLN)	[25, 200]
Linear Discriminant Analysis (LDA)	Gamma (G)	[0,1]
	Delta (D)	log-scaled in the range [1e-6,1e3]
	DiscrimType (DT)	{linear, quadratic, diagLinear,} {diagQuadratic, pseudoLinear, pseudoQuadratic}
Random Forest (RF)	Depth (D)	[5,20]
	Number of Trees (NT)	[15,100]
	Maximum Depth of the tree	[5,30]
Logistic Regression (LR)	Penalty (P)	{L2, elastic net}
	Inverse of regularization strength (C)	[0.25, 1.0]

defines a separator hyperplane between classes exploiting a subset of the training instances (support vectors). LR is a widely used classification method based on the logistic function. In binary classification, it estimates the probability of a sample $\mathbf{x}$ to belong to a class labelled as $y = 1$ as $P(y|\mathbf{x}) = \frac{\exp(q+\mathbf{wx})}{1+\exp(q+\mathbf{wx})}$ where $\mathbf{w}$ and q are learnable parameters. A RF combines several decision trees to make classifications. The use of several decision tree helps in improving the accuracy. Furthermore, to prevent possible over-fitting, regularization terms in the training procedures were used for SVM learning using the SVM soft-margin formulation [220], and for neural networks learning using a weight decay [221] during the learning algorithm execution. ANNs were trained with the ADAM algorithm. A maximum number of 1000 epochs with a patience of 50 epochs on the validation set was used to train the network models. Figure **??** shows the trend of the accuracy during the first 40 iterations of a learning stage on a single subject model. For all the classifiers, the hyperparameters used during the CV procedure are reported in Table 3.1. Accuracy, precision, and recall are reported to assess the classification output quality. Precision measures result relevancy, while recall how many truly relevant results are returned. The F1 score, combining precision and recall, was computed to assess the classification performance in minimizing false negatives for the first class (negative valence) analysis. Considering many use cases, the minimization of failure in recognizing negative valence is the main issue.

Experimental results

Accuracy was related to the model's ability to correctly differentiate between two valence states. EEG tracks relating to the negative and positive image tasks were associated to the first and the second class, respectively.

The mean of the individual accuracies and standard deviations computed on each subject (within-subject case) and the accuracies and standard deviations computed on all subjects data as a whole (cross-subject case) are showed when a priori spatial-frequency knowledge is used (Table 3.2) or not (Tables 3.3 and 3.4).

TABLE 3.2: Accuracy (mean and standard deviation) considering a priori knowledge i.e. Asymmetry - Within-subject (Within) & Cross-subject (Cross)

CLASSIFIER	**Entire EEG Band**		**αBand**	
	Within	Cross	Within	Cross
k-NN	54.0±4.1	51.0±1.2	53.8±4.0	51.3±0.4
SVM	56.8±3.4	50.8±0.2	56.7±3.0	51.2±0.3
LDA	54.5±3.8	51.2±0.8	53.8±3.5	51.0±1.0
ANN	58.3±3.0	51.8±0.3	58.5±3.0	51.5±1.6
RF	55.7±3.9	50.7± 1.2	54.5± 4.5	50.9±1.3
LR	52.5± 4.1	51.4± 0.2	53.7± 4.3	51.2± 0.7

TABLE 3.3: Accuracy (mean and standard deviation) without considering a priori knowledge i.e. Asymmetry - Within-subject

CLASSIFIER	**Entire EEG Band**			**Filter Bank**		
	No PCA/CSP	**PCA**	**CSP**	**No PCA/CSP**	**PCA**	**CSP**
k-NN	71.0±6.0	67.7±8.4	72.0±8.9	75.6±5.8	66.8±7.2	94.5±3.5
SVM	66.9±8.1	66.3±10.3	73.4±9.5	71.6±8.9	62.0±7.8	95.5±2.8
LDA	63.1±4.9	55.3±4.0	74.0±10.0	62.9±5.3	53.9±3.5	95.0±2.9
ANN	69.7±5.1	66.3±6.2	78.1±8.0	66.7±4.9	65.6± 5.6	**96.1±3.0**
RF	66.4 ± 4.1	58.9 ± 4.2	72.8 ± 9.4	67.4 ± 4.1	59.3 ± 5.0	94.2 ± 2.7
LR	62.7 ± 4.9	52.3 ± 2.9	72.6 ± 9.3	61.0 ± 5.0	51.2 ± 4.0	95.1 ± 2.9

Results are shown at varying the adopted classifier. Better performances are obtained without a-priori knowledge and when features are extracted by combining Filter-Bank and CSP, both in within-subject and cross-subject case. In within-subject analysis, the data subsets are more uniform and all the classifiers provide very high accuracy. In Fig. 3.6 the data of four random subjects projected in the CSP space, with and without the Filter Bank, are compared. The classes, after using the Filter Bank, are easily separable with respect to the use of the only CSP, as highlighted by the results. In Table 3.5, the accuracies in the within-subjects experiments are reported for all the subjects.

In cross-subject analysis, when data from all subjects are merged, variability increases and not all the classifiers give good results. Interestingly, in the cross-subject approach, the k-NN classifier allows to achieve by far the best performance, while the scores degrade using the other classifications setups. This behaviour suggests that the data of similar classes are close together for different

TABLE 3.4: Accuracy (mean and standard deviation) without considering a priori knowledge i.e. Asymmetry - Cross-subject

CLASSIFIER	Entire EEG Band			Filter Bank		
	No PCA/CSP	PCA	CSP	No PCA/CSP	PCA	CSP
k-NN	68.4±0.2	62.1±0.9	56.8±0.5	70.1±1.0	61.1 ± 0.3	**80.2±2.1**
SVM	51.5±0.6	52.1±0.3	61.0±2.0	51.8±1.0	51.2±0.3	71.3±2.0
LDA	53.5±0.7	50.9±0.4	55.4±4.2	52.6±0.1	50.9± 0.2	63.7±2.1
ANN	59.9±1.0	54.5±0.2	58.1±1.1	57.4±0.1	53.7±0.1	63.3±2.7
RF	56.5 ± 0.6	55.3 ± 0.7	59.2 ± 1.9	57.8 ± 1.1	52.5 ± 2.9	65.0 ± 3.8
LR	50.5 ± 1.9	50.6 ± 0.5	55.7 ± 4.9	51.8 ± 0.9	50.9 ± 0.5	58.1 ± 1.5

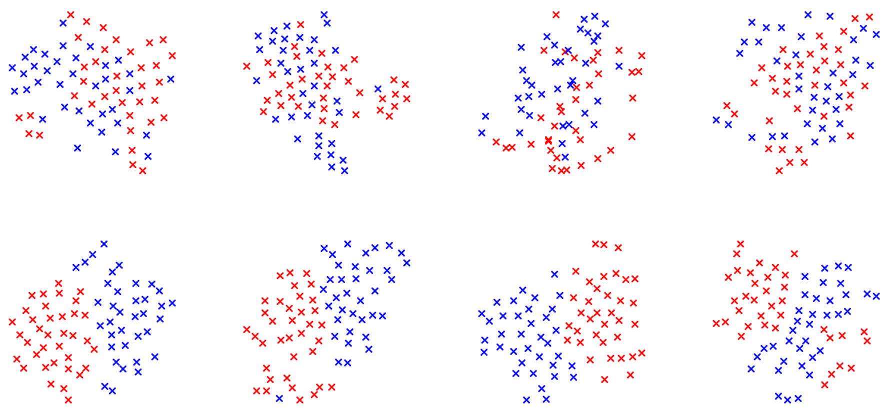


FIGURE 3.6: t-SNE based data comparison of four random subjects projected in the CSP space, without (first row) and with (second row) the Filter Bank. Filter Bank improves the classes (blue and red) separability.

subjects, but that in general they are not easily separable through classical Machine Learning methods. Moreover, a feature selection analysis using the Mutual Information (MI) method, proposed in [150], was made using the best experimental setups of both within-subject and cross-subject approaches. The results reported in Table 3.6 show that just the 12.5 % of the FBCSP features are enough to achieve accuracy performances over the 90 % in the within-subject case. Therefore, the features extracted by the CSP in conjunction with Filter Bank resulted effective in emotional valence recognition.

In conclusion, the proposed solution based on 12-bands Filter-Bank provides the best performances reaching 96.1 % of accuracy with ANN in within-subject analysis and 80.2 % using k-NN with $k = 2$ in cross-subject analysis. In the within-subject case, for the ANN the best top-5 subjects reached the best performances using ANN with one layer with less than 100 neurons equipped with the classical *tanh* activation function, showing that networks with few parameters can be sufficient to address this classification problem as long as a proper set of features is provided. Precision, Recall and F1-score metrics are reported in

TABLE 3.5: Accuracies obtained for each subject in the within-subject experiments when a FC-CSP Pipeline is adopted

Subject	k-NN	SVM	LDA	ANN	RF	LR
#1	95.8	95.8	94.4	95.8	93.3	94.4
#2	95.8	92.2	92.2	93.1	92.2	95.8
#3	94.4	93.6	93.7	94.4	91.1	92.2
#4	95.8	98.6	98.1	99.0	94.4	94.4
#5	91.7	93.8	93.2	94.4	93.1	93.1
#6	97.2	96.2	95.8	97.2	93.9	95.8
#7	95.8	96.1	95.8	98.6	94.4	95.8
#8	97.2	98.6	97.2	99.0	97.2	97.0
#9	98.6	98.6	98.6	98.6	96.2	98.6
#10	92.0	94.6	94.4	97.2	95.8	94.5
#11	95.8	95.0	94.6	97.2	93.6	95.0
#12	94.4	94.7	94.4	97.2	92.3	94.4
#13	98.6	98.6	98.6	99.0	95.8	98.6
#14	95.8	95.7	95.8	95.8	97.2	94.4
#15	85.9	91.2	90.5	91.0	89.9	90.3
#16	95.4	97.2	96.7	98.2	97.2	97.0
#17	86.3	95.0	94.6	93.1	92.7	95.8
#18	93.1	91.4	90.2	92.0	92.7	93.0
#19	97.2	98.6	98.7	99.0	98.6	98.6
#20	94.4	98.5	97.2	95.8	95.4	97.2
#21	97.2	97.2	97.2	98.6	95.8	98.6
#22	98.6	97.9	97.4	99.0	95.4	97.2
#23	89.3	90.0	89.2	89.4	88.1	88.9
#24	95.2	97.9	97.3	98.7	98.6	98.3
#25	90.2	90.1	89.4	90.3	90.2	88.5
Average ± std.	94.4 ± 3.5	95.5 ± 2.8	95.0 ± 2.9	96.1 ± 3.0	94.2 ± 2.7	95.1 ± 2.9

Fig. 3.7.

TABLE 3.6: Accuracy performances of the best processing solutions for both within- and cross-subject approaches at varying the number of input features selected through the Mutual Information strategy.

CLASSIFIER		#Features			
		12	24	50	96
k-NN	Cross	58.7 ± 1.0	65.1 ± 1.8	74.4 ± 0.9	80.2 ± 2.1
ANN	Within	92.8 ± 4.1	93.0 ± 4.1	93.4 ± 1.0	96.1 ± 3.0

Discussion

In the previous Sections the measurability foundation of emotion was discussed. In this thesis, results from the *Self Assessment Manikin* questionnaire confirmed the compatibility of the experimental sample with that of *Oasis* thus improving the reproducibility of the experiment and the generalizability of the outcome. Moreover, the reference theory adopted allows the measurement of emotions arranging them along interval scales. In this framework, the preliminary binary classification of the proposed system could be enhanced by increasing the number of classes. Thus, the number of valence states increase and a higher resolution

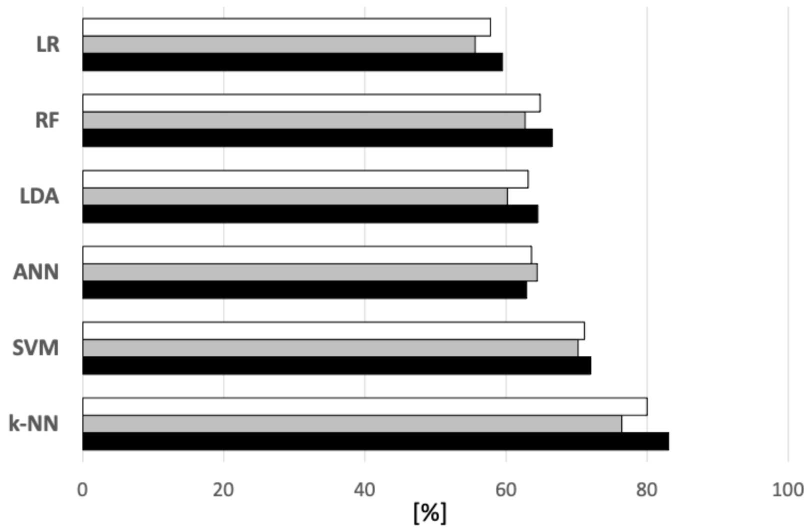


FIGURE 3.7: F1-score (White), Recall (Grey) and Precision (Black) for the best performance of each classifier - Cross-subject

metric scale can be obtained. Therefore, the Circumplex Model is compatible with an upgrade of the proposed binary classification method. It is noteworthy that the number of classes can increase if emotional valence states can be experimentally induced at higher resolution. This is precisely what the standardized stimuli datasets allow because their scores are organised according to an interval scale. The novelty of this research is based on the compliance with different quality parameters. In Table 3.7, this study is compared with the works examined in Section 2.1.1, taking into account the following criteria: (i) classification vs measurement, (ii) standardized stimuli, (iii) self-assessment questionnaires, (iv) number of channels ≤ 10, (v) cross-subject accuracy > 80 % (vi) within-subject accuracy > 90 %. As concerns the first quality parameter, the option between classification and measurement is related to the reference theory adopted (i.e., discrete model or circumplex model).

There are only two studies combining SAM and standardized stimuli ratings for the construction of the metrological reference [144, 146]. Therefore, literature concerning EEG-based emotion detection exhibits a lack of generalizability for the presented results. Among all the examined works, the proposed study is the only one that matches all the aforementioned criteria.

TABLE 3.7: Studies on emotion recognition classified according to metrological approach, number of channels and accuracy (n.a. = "not available", ✔ = "the property is verified". Only for the first line, ✔ = "Measurement")

	[117]	[118]	[111]	[119]	[120]	[112]	[113]	[121]	[122]	[123]	[138]	[124]	[125]	[114]	[126]	[135]	[127]	[115]	[116]	[139]	[140]	[141]	[142]	[143]	[136]	[144]	[145]	[146]	[147]	**Our work**
Measurement vs Classification	✔	✔	✔	✔	✘	✔	✘	✔	✔	✔	✘	✔	✔	✔	✘	✘	✔	✔	✔	✔	✔	✘	✘	✔	✔	✔	✘	✔	✔	✔
Standardized Stimuli	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	✔	✘	✘	✘	✘	✘	✔	✘	✔	✘	✘	✘	✔	✔	✘	✔	✔	✔
Self-assessment Questionaries	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✘	✔	n.a.	✔	✔	✔
#channels$\leq$ 10	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	✘	✘	✘	✘	✘	✘	✘	✘	✔	✔	✔	✔	✔	✘	✔	✔	✔	✔	✔
Cross-subject Accuracy (>80%)	✘	✔	✔	✔	✔	✔	✔	n.a	✔	✔	n.a.	n.a	n.a	n.a	n.a	n.a	n.a	✔	✔	n.a	✔	n.a	✔	✘	✘	✔	✘	✘	✘	✔
Within-subject Accuracy (>90%)	n.a.	n.a.	✔	n.a.	n.a.	n.a.	n.a.	✘	n.a.	n.a.	✔	✔	✔	✔	✘	✔	✘	n.a.	✔	✘	n.a.	✔	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	✔	✔

3.2 Distraction assessment

3.2.1 Basic Ideas

The proposed method for detecting distraction during motor rehabilitation is based on the following key concepts:

- *EEG-based distraction detection*: During a rehabilitation motor task, EEG trend is influenced by the state of the patient attention or distraction to the task itself.
- *Attention vs distraction definition*: Focusing on motor task means imagining, with open eyes, the movement while its execution and trying not to think about anything else. A distracting condition occurs when the patient performs an entirely absorbing cognitive task while continuing to carry out the rehabilitation movement. To the end of evaluating the phenomenon, a rehabilitative motor task is carried out. The assignment is run under conditions of concentration on the action and in the presence of a distractor (auditory, visual, and visuo-auditory) which engages the learner in a concurrent cognitive task analogously as what done in Asayb et al [65].
- *Metrology perspective*: An applied metrological and instrumentation-aimed approach is guaranteed, for the first time, in the EEG based distraction detection.
- *Feature extraction enhancement*: After an artifact removal performed by an Independent Component Analisys (ICA) based algorithm, a multiple band-pass Filter-Bank, in combination with a Common Spatial Pattern algorithm, selects spatial, temporal and frequency features. In particular, a 12-band Filter-Bank is proposed for enhancing, the peculiar contribution of the delta, theta, and alpha bands as fundamental in the analysis of attentional processes [222], compared to previous 9-band approaches [150].
- *High Wearability*: The EEG acquisition system is realized in ultra-light foam. The ergonomic and comfortable device is equipped with a rechargeable battery and transmits the acquired data via Bluetooth. Dry electrodes avoid the inconvenient of electrolytic gel.
- *Clinical applicability*: wearability cannot be a prejudice for accuracies compatible with clinical use. A method with state-of-the-art accuracy (greater than 80 % [65, 150]) is required.
- *Validation based on wide comparison*: Performance of the proposed method are compared with different strategy of EEG feature extraction (including the proposal of Hamadicharef et al. [150]), and different types of classifiers.

3.2.2 Method

The proposed method is depicted in Fig. 3.8. The EEG signals are acquired by *Active Dry Electrodes* from the scalp. Each channel is differential with respect to

AFz (REF), and referred to Fpz (GND), according to 10/20 international system. Analog signals are first transduced by the *Active Dry Electrodes* and then conditioned by the *Analog Front End*. Next, they are digitized by the *Acquisition Unit* and transmitted to the *Data Analysis* stage. Here, after an artifact removal performed by an ICA based algorithm, suitable features are extracted by the chain of a 12-component *Filter Bank* and a *Common Spatial Pattern* (CSP) algorithm. Then, a classifier receives the feature arrays and detects distraction.

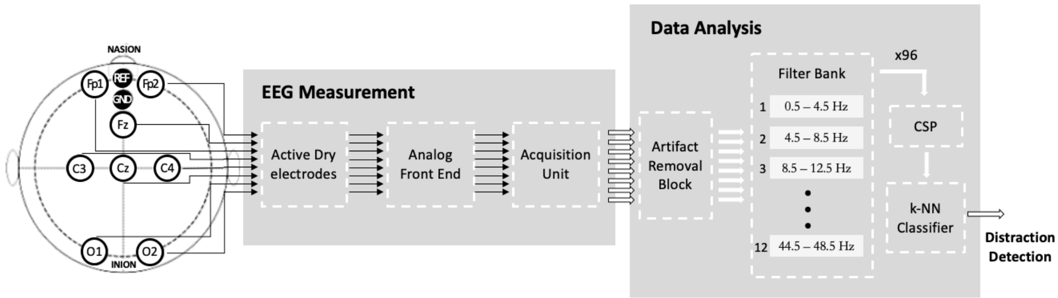


FIGURE 3.8: The proposed distraction-detection method (CSP: Common Spatial Pattern algorithm).

Feature selection and extraction

The EEG signal, acquired through eight channels, was filtered through a 12 IIR band-pass Filter Chebyshev type 2 filter bank, 4 Hz amplitude, equally spaced from 0.5 - 48.5 Hz. In Hamadicharef et al.[150], a filter bank with 9 filters of 8 Hz amplitude equal to [0-40] Hz, with a 4 Hz overlap, was proposed. This solution subdivided the traditional EEG beta and gamma bands into sub-bands, however combining other bands (delta and theta with the first filter between 0 and 8 Hz, as well as theta and alpha with the second filter between 4 and 12 Hz). Considering the relevance of the delta, theta and alpha bands in the analysis of the attention highlighted in Graber et al.[223] and in Coelli et al.[222], the solution proposed in this study allows to enhance their peculiar contribution.

The unit of analysis of the classification activity was identified in time windows of 3 s with an overlap of 1.5 s. Considering a sampling frequency of 256 Sa/s, each of these record is therefore composed of 96 EEG tracks (obtained by applying the 12 filters of the Filter Bank on each of the 8 channels), each one of 1536 samples.

A Common Spatial Pattern (CSP) was used as a spatial filtering algorithm. CSP is one of the most used feature extraction methods for classifying EEG signals [150, 224]. In a binary problem, the CSP acts by calculating the covariance matrices relating to the two classes. These two matrices are simultaneously diagonalized in a way that the eigenvalues of two covariance matrices sum up to 1. Through the subsequent use of a bleaching matrix, a suitable projection matrix is identified in order to reorganize the input into a number of components consistent with the dimensions of the input matrix. In a binary problem, these components are sorted on the basis of variance in order: (i) *decreasing*, if the projection matrix is applied to inputs belonging to class 1, and (ii) *ascending*, in case of inputs belonging to class 2 [225]. In this thesis, the CSP receives the records (epochs) as 3D tensors (channels, filters, and samples). It outputs 2D matrices (channels, filters) reducing the dimensionality of the features by a factor of 1536 (number of sample).

Classification

A k-Nearest Neighbour (k-NN) classifier is used for classifying the CSP output. Compared to other supervised machine learning methods, k-NN is a non-parametric method (i.e., without a priori assumption on the data) which uses the labelled data itself for the classification without any training. The behavior of k-NN in its simplest version can be described as follows: given a set D of labelled points, a distance measure (e.g., Euclidean, Minkowski) and a positive integer k, when a new unlabelled point p is presented, the k-NN algorithm searches in D for the k points nearest to p, so the most present class label along its k neighbors is assigned to p. Thus, the only hyperparameters required to k-NN are a positive integer k and the distance measure to use together with any parameters related to the distance measure if needed. These hyperparameters were set using a cross-validation procedure. k-NN has already been widely used in EEG signal analysis showing interesting results (see for example [226]).

3.2.3 Experimental validation results

In this section, the experimental assessment of our proposal is reported and the results are discussed.

Experimental Protocol

The ethical committee approved the experimental protocol of the University of Naples Federico II. A written informed consent was obtained from each volunteer before the experiment. All experiments were carried out in accordance with relevant guidelines and regulations. A session was based on seventeen volunteers subjects (eleven males and six females, with an average age of 30.76±8.15). All of them had a normal clinical history with normal vision and normal hearing, and no neurological disease. The participants were seated in a comfortable chair with armrests, in a very quiet room, about one meter away from a PC screen. After wearing the EEG-cap, participants were requested to execute a squeeze-ball exercise whenever a start command appeared on the PC screen. Squeeze-ball is one of the most common hand rehabilitation exercises [227]. Following a period of immobilization in plaster, after a surgical intervention or in the presence of inflammatory or degenerative pathologies (e.g., arthrosis, rheumatoid arthritis), hand-ball rehabilitation showed to be important in maintaining or restoring the functional use of the hand [228]. Motor task execution consists of maintaining attention focused only on: (i) the squeeze movement (*attentive-subject trial*), or (ii) a concurrent distractor task (*distracted-subject trial*); in both trials the participant must perform the squeeze-ball movement. An aneroid sphygmomanometers supported the user attention to motor task execution: volunteers were asked to focus the aneroid gauge, while squeezing the bulb and pumping air into the cuff. The distractor task was based on the *Oddball paradigm* [229, 230]: the presentations of sequences of repetitive stimuli, infrequently interrupted by a deviant stimulus. The oddball paradigm is one of the most widely used methods to study the neurophysiology of attention. In the proposed protocol, the volunteer was asked to count the number of certain stimuli sequences. Three types of stimuli sequences were proposed: (i) acoustic, played with a conventional headphone, (ii) visual, displayed on a PC screen, and (iii) and visual-aucoustic combination [231]. Each participant completed one session composed of 30 trials: 15 *attentive-subject trial* and 15 *distracted-subject trial*. The trials sequences were randomly chosen for minimizing the influence of task learning. Each trial consisted of: 2 s task presentation, 9.5 s task execution and 5 s relax. Furthermore, a 15 s baseline was acquired at the beginning of the session. In the following, trial contents are detailed:

- *Attentive-subject trial*
  An Attentive-subject trial notification appears for 2 s on the PC screen. Then, a ball-squeezing image triggers the start of the motor exercise and a new message on the screen asks the subject to focus on the squeezing movement. At the end of the task execution, an image of a relaxing landscape is shown for 5 s.

- *Distracted-subject trial*
  A notification concerning the distractor task (Audio, Visual or Audio-Visual) appears for 2 s on the PC screen. Then, an acoustic message notices the beginning of the motor exercise; a distractor task (based on Oddball paradigm), chosen among the followings, starts:

  – The *Audio Distractor* is based on the auditory oddball paradigm. Eight tones sequences sound through the earbuds. Tones range among three different frequencies: *low*, 500 Hz, *middle*, 1200 Hz, and *high*, 1900 Hz. The tone *low* has 50% probability of occurrence. The occurrence probability of the *middle* and the *high* tones is 25%. The target sequence is the appearance of a diverted tone after the other more frequent one: when the *middle* tone occurs immediately after the *low*, or when the *high* occurs immediately after the *low*. Others combinations are not considered as target occurrences.

  – The *Visual Distractor* task is based on the visual oddball paradigm. Three 2D-Gabor masks were used with different orientation: 90, 60, and 30° (Fig.3.9). The 2D-Gabor mask is a Gaussian kernel function modulated with sinusoidal plane wave. The most probable Gabor (50% of probability) has orientation of 90° , while the diverted Gabor (25% of probability) has 60 or 30° orientation. Eight Gabor sequences occurred on the PC screen. The target sequence was the occurrence of diverted Gabor mask (with orientation of 60 or 30°) after the most frequently with 90° orientation.

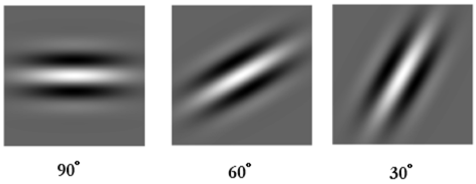


FIGURE 3.9: Visual Distractor task elements based on visual Gabor mask with different orientation: 90°, 60°, and 30°.

– The *Audio-Visual Distractor* task is a combination of the previous oddball paradigms. Eight between tone and Gabor sequences occur randomly. The target sequence is the occurrence of any Gabor mask after the tone. Others combination sequences are not target occurrences.

At the end of the task, a relaxing landscape is presented for 5 s. During the relax period, the subjects are asked to give the number of the observed targets.

EEG Instrumentation

In this thesis, the commercial EEG acquisition system *AB-Medica Helmate* [216] is employed (Fig.3.10 A). The device, composed of ten dry electrodes, guarantees

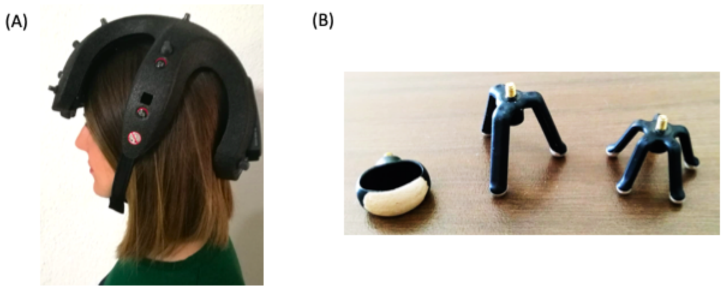


FIGURE 3.10: (A) EEG data acquisition system *Helmate8*, and (B) Different configuration of dry electrodes from *abmedica*. [216].

eight acquisition channels. The EEG signal is acquired by dry electrodes made of conductive rubber with an Ag/AgCl coating at their endings [232]. Three different types of electrodes, with different shapes, are used to pass hair and reach the scalp or join to the hairless areas (Fig.3.10 B). The output signal is recorded as difference between each of 8 channels and the ground electrode (Fpz) [233]. Then, the difference is referenced with respect to the electrode (AFz). A dedicated software (*Helm8 Software Manager*) allows to check the contact impedance between the electrodes and the scalp. EEG signal is acquired with a sampling rate of 512 Sa/s. The acquisition software allows to use several filters (e.g., notch and IIR). This data acquisition system is a certified EEG system Class IIA (according to Medical Device Regulation (EU) 2017/745) with accurate components. A Texas Instruments analog front-end, the ADS1298 [217] with a 24-bit, $\Delta\Sigma$ analog-to-digital converter (ADCs) with built-in programmable gain amplifiers (PGAs), internal reference, and an onboard oscillator, are exploited. The device exhibits the following main metrological performances: (i) CMRR: -115 dB; (ii) eight low-noise PGAs and eight high-resolution ADCs (ADS1298, ADS1298R); (iii) input-referred noise: 4 μVPP (150 Hz BW, G = 6); and (iv) input bias current: 200 pA; joined to the following operating performances: (i) low power: 0.75 mW/channel; and (v) data rate: 250 Sa/s to 32 kSa/s.

Data Processing

During the experiments 4590 epochs composed of 8 channels of 512 samples were acquired. In Table 3.8 number of (i) subjects, (ii) sessions, (iii) trials, (iv) epochs per trial (v) epochs per subject, and (Vi) epochs as a whole are reported.

TABLE 3.8: Data-set composition

# Subjects	# Sessions	# Trials per Session	# Epochs per trial	# Epochs per subject	# Total Epochs
17	3	30	3	270	4590

Half of the epochs were collected during the *attentive-subject trials* and were labeled as belonging to the first class. The remaining part was acquired during the *distracted-subject trials* and was labeled as belonging to the second class. The recorded EEG was divided in 3 s epochs. Each epoch was filtered between 0.5 and 48.5 Hz using a zero-phase 4th-order digital butterworth filter. An independent component analysis (ICA) algorithms - Infomax-ICA[234] - filtered out artifacts from the signal. In particular the version implemented by *Runica* module of *EEGlab* tool was adopted. Feature extraction was implemented either in time domain and frequency domain. For the latter Relative and Absolute Power Spectral Density at varying of frequency bands were considered. Three different frequency bands articulation were examined:

- seven traditional EEG bands: delta [1–4] Hz, theta [4–8] Hz, alpha [8–12] Hz, low beta [12–18] Hz, high beta [18–25] Hz, low gamma [25–35] Hz, and high gamma [35–45] Hz; in this case, the number of features for each epoch was 112 (7 bands * 2 PSD (relative and absolute) * 8 channels);

- nine 8 Hz bands, 4 Hz overlapped, in the range [1-40] Hz; the number of features for each epoch was 144 (9 bands * 2 PSD (relative and absolute) * 8 channels);

- twelve 4 Hz bands, non-overlapped, in the range [0.5-48.5] Hz; the number of features for each epoch was 192 (12 bands * 2 PSD (relative and absolute) * 8 channels);

Regarding time domain, the feature extraction was based on four different approaches:

- only CSP: in this case, the number of features for each epoch was 8 (CSP remaps the input information in a new space with dimensionality equal to the number of channels);

- CSP preceded by different types of Filter-Banks: three different types of Filter-Banks were applied with the same band articulation proposed for the feature extraction in the frequency domain. In these cases CSP remaps the input information in a new space having dimensionality equal to the number of channels (8) multiplied with number of bands, obtaining 56, 72, and 96 number of features respectively.

Five supervised machine learning binary classifiers were used for discriminating between attention or distraction conditions: k-Nearest Neighbour (k-NN), Support Vector Machine (SVM) [211], Artificial Neural Network (ANN) [211], Linear Discriminant Analysis (LDA) [235], and Naive Bayes (NB) [236]. Regularization terms were exploited in the training procedures for neural networks and SVM learning processes, using a weight decay and the soft-margin formulation, respectively. All the classifiers were tested on the seven features types described above. For each subject, the hyperparameters of each classifier were selected by a random search with Nested Cross Validation to mitigate possible bias induced by the low sample size [237]. Differently from the classical k-fold cross validation, Nested CV is composed of two nested k-fold cross validation procedures: the inner one finds the best model hyperparameters, and the outer one estimates the performance of the inner search. Namely, in the classic k-fold CV, given a combination of the hyperparameters values, a set of data is divided into a partition of k subsets (folds). Thus, a set T_I composed of $k-1$ folds is used to train the model and the remaining fold E_I is used for the performance evaluation by computing the appropriate metric scores (e.g., accuracy). This process is repeated for all the combinations of the k folds, by making different pairs of training set T_I and test set E_I at each iteration. In this way, final average metrics scores between all the different test sets E_I are computed. This process is then repeated for each hyperparameters combination, finally returning the best average metrics values together with the related hyperparameters. In this process, the model is evaluated together with the hyperparameters tuning. Instead, in the nested cross validation CV procedure, an outer CV makes a first division of the data into l folds; then, a set T_O composed of $l-1$ folds is used as input to a classical inner k-fold CV procedure, as above described (and therefore further divided into k folds by the inner CV procedure). Then, the returned best hyperparameters values are used to train the model on the T_O set as a whole and tested on the remaining fold, say E_O. This process is repeated for all the combinations of the l folds and the final average metrics on the E_O sets are reported. In this way, the nested CV process avoids a possible bias on the model, due to the use of the same data for the model hyperparameters tuning and the model evaluation. In this thesis, a 10-fold Nested CV was used. In the outer layer, 10% of the data was separated for test and the rest of the data was used to develop a model. In the internal layer, the remaining 90% of the data was used for tuning the hyperparameters. Training and test sets were obtained without separating the trials consisting of 3 epochs each. In this way, the training and the test sets do not include parts of the same trial. The hyperparameters variation range are displayed in Table 3.9.

TABLE 3.9: Classifier optimized Hyperparameters and variation range

Classifier	Hyperparameter	Variation Range
k-Nearest Neighbour (k-NN)	Distance (DD)	{cityblock, chebychev, correlation, cosine, euclidean, hamming, jaccard, mahalanobis, minkowski, seuclidean,spearman}
	DistanceWeight (DW)	{equal, inverse, squaredinverse}
	Exponent (E)	[0.5, 3]
	NumNeighbors (NN)	[1, 5]
Support Vector Machine (SVM)	BoxConstraint (BC)	log-scaled in the range [1e-3,1e3]
	KernelFunction (KF)	{gaussian, linear, polynomial}
	KernelScale (KS)	log-scaled in the range [1e-3,1e3]
	PolynomialOrder (PO)	{1,2,3,4}
Artificial Neural Network (ANN)	Activation Function (AF)	{relu, sigmoid, tanh}
	Hidden Layer nr. of Neurons (HLN)	[25, 200]
Linear Discriminant Analysis (LDA)	Gamma (G)	[0,1]
	Delta (D)	log-scaled in the range [1e-6,1e3]
	DiscrimType (DT)	{linear, quadratic, diagLinear,} {diagQuadratic, pseudoLinear, pseudoQuadratic}
Naive Bayes (NB)	DistributionName (DN)	{normal, kernel}
	Width (W)	log-scaled in the range [1e-4,1e14]
	Kernel (K)	{normal, box, epanechnikov, triangle}

TABLE 3.10: Within-subject accuracy (mean and standard deviation percentage of the 17 subject accuracy) at varying feature and classifier

CLASSIFIER	**FEATURE**						
	FREQUENCY DOMAIN			TIME DOMAIN			
				CSP	Filter-Bank + CSP		
	7 Traditional EEG Bands	9 EEG Bands Proposed in[150]	Proposed 12 EEG Bands		7 Traditional EEG Bands	9 EEG Bands Proposed in[150]	Proposed 12 EEG Bands
$k-NN$	77.5 ± 5.5	76.7 ± 5.5	80.2 ± 5.1	65.9 ± 5.0	87.4 ± 4.1	90.9 ± 3.2	92.8 ± 1.6
SVM	79.9 ± 5.6	76.0 ± 4.0	81.7 ± 6.9	69.2 ± 5.1	86.8 ± 4.5	89.8 ± 3.7	91.1 ± 3.2
LDA	76.7 ± 7.4	75.1 ± 7.2	78.3 ± 6.3	67.7 ± 4.8	82.9 ± 4.5	85.7 ± 6.2	86.6 ± 2.0
ANN	75.6 ± 6.3	73.6 ± 6.7	76.9 ± 6.4	67.2 ± 4.5	81.9 ± 4.5	85.1 ± 5.0	86.3 ± 3.5
NB	64.5 ± 6.2	63.8 ± 5.2	65.3 ± 7.8	65.2 ± 4.9	75.3 ± 7.3	77.0 ± 7.2	78.7 ± 7.5

Experimental Results

A within-subjects approach was realized. The accuracy (mean and standard deviation) for each classifier was assessed at varying the type of input feature. Table 3.10 shows better performances in case of features extracted from the time domain by combining Filter-Bank and CSP.

In particular, the proposed solution based on 12 bandpass Filter-Bank provides the best performances for all classifiers except for LDA. In Table 3.11, the accuracy of the proposed solution is shown for each subject at varying the classifier. In case of k-NN, the mean accuracy reached the maximum value of 92.8±1.6 %. To the best of the authors' knowledge, the accuracy obtained can be considered state-of-the-art when considering a within subjects approach. Regarding rehabilitation goals, the minimization of failure in recognizing distraction is the main issue. Therefore, an F-measure test was carried out to assess the classification performance in minimizing false negatives for the second class (distraction) analysis. Fig. 3.11 shows a k-NN mean Recall higher than 92 %.

TABLE 3.11: Within-subject accuracy of the proposed solution based on the 12 bandpass Filter Bank and Common Spatial Pattern at varying the classifier.

SUBJECT	**CLASSIFIER**				
	k-NN	SVM	LDA	ANN	NB
#1	91.1 ± 5.3	90.3 ± 5.2	88.2 ± 5.3	86.3 ± 7.2	66.0 ± 9.7
#2	92.2 ± 2.2	90.1 ± 5.1	85.1 ± 6.2	83.5 ± 5.3	79,2 ± 9.9
#3	93.3 ± 5.5	92.2 ± 5.1	89.2 ± 7.1	80.3 ± 7.3	87.2 ± 7.3
#4	94.1 ± 4.2	95.0 ± 2.2	89.6 ± 5.5	92.4 ± 6.8	81.3 ± 4.4
#5	90.4 ± 4.3	89.2 ± 6.7	84.3 ± 9.2	84.5 ± 7.6	65.3 ± 9.7
#6	93.3 ± 3.1	96.5 ± 3.8	91.7 ± 6.8	89.7 ± 6.2	74.1 ± 7.3
#7	96.1 ± 3.2	92.3 ± 4.4	87.2 ± 6.8	87.6 ± 8.3	80.0 ± 9.8
#8	93.1 ± 5.2	91.2 ± 6.7	88.4 ± 7.3	87.6 ± 6.1	86.5 ± 6.3
#9	91.2 ± 4.5	89.1 ± 8.8	88.4 ± 9.1	87.6 ± 6.5	82.8 ± 6.2
#10	92.1 ± 4.4	85.2 ± 4.8	80.3 ± 5.7	82.3 ± 6.9	73.2 ± 9.9
#11	91.1 ± 5.3	90.2 ± 6.7	83.5 ± 8.5	82.5 ± 9.1	79.2 ± 7.1
#12	94.8 ± 4.2	93.8 ± 3.3	91.7 ± 6.6	87.6 ± 06	87.3 ± 3.5
#13	93.3 ± 6.2	92.2 ± 7.6	84.2 ± 5.9	86.8 ± 8.4	75.6 ± 8.4
#14	96.6 ± 4.5	96.3 ± 5.3	90.8 ± 5.8	90.4 ± 6.1	86.8 ± 8.2
#15	93.8 ± 6.2	94.1 ± 4.5	88.8 ± 8.1	86.2 ± 6.5	84.4 ± 5.6
#16	93.5 ± 7.3	91.8 ± 5.5	86.6 ± 2.2	87.2 ± 5.5	82.5 ± 5.6
#17	93.2 ± 4.1	84.8 ± 6.5	77.5 ± 1.6	77.8 ± 1.1	66.4 ± 8.0
MEAN	92.8 ± 1.6	91.1 ± 3.2	86.6 ± 2.0	86.3 ± 3.5	78.7 ± 7.5

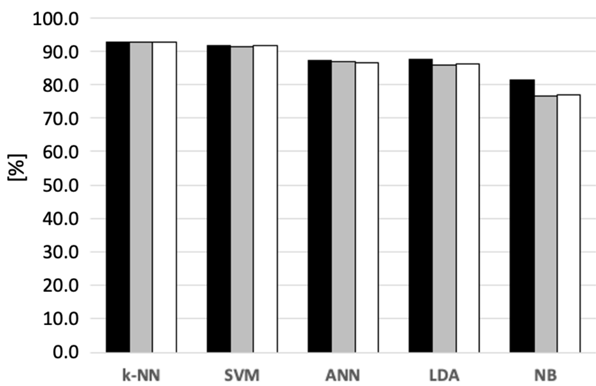


FIGURE 3.11: F-Measure test results for the best performance of each classifier: Precision (black) , Recall (gray), and F1-score (white).

Chapter 4

Complex mental state assessment

In this chapter measurement methods and experimental validation for complex mental state are presented. The word *complex* stands for multidimensional. Both stress and engagement, indeed, are frequently described in terms of basic dimensions, namely: cognitive, emotive, and behavioural. The proposed approach will not address the behavioral dimension because other signals would be more appropriate than the EEG (e.g., video and accelerometric data).

4.1 EEG-Based Stress Assessment

4.1.1 Design

In this Section, (A) the *Basic Ideas*, (B) the *Architecture*, (C) the *Operation*, and (D) the *Feature Extraction and Classification* of the instrument are presented.

Basic Ideas The concept design of the real-time stress monitoring instrument was based on the following main basic ideas.

- *High wearability*: a single differential channel allows the use of only two frontal electrodes in the area FP1 e FP2 according to the EEG International 10–20 system for placement of EEG electrodes on the scalp. The reference electrode is applied to the earlobe. These positions have been used in reports of successful studies on stress [199]. Active dry electrodes avoid the inconvenient of electrolytic gel. A wireless module allows the user to carry out work activity during the EEG acquisition.

- *High accuracy and low latency*: Despite the use of a single differential acquisition channel, a time-domain based machine learning algorithm brings to an accuracy of 98.3 ± 0.4 in stress detection. A time window of 512 samples guarantees a latency of 2 s.

- *Off-the-shelf components*: The measurement of frontal asymmetry by EEg at very-low density (single channel) allows high wearability, maximum accuracy, and low latency by exploiting the lowest cost hardware on the market (< 200 $) [51].

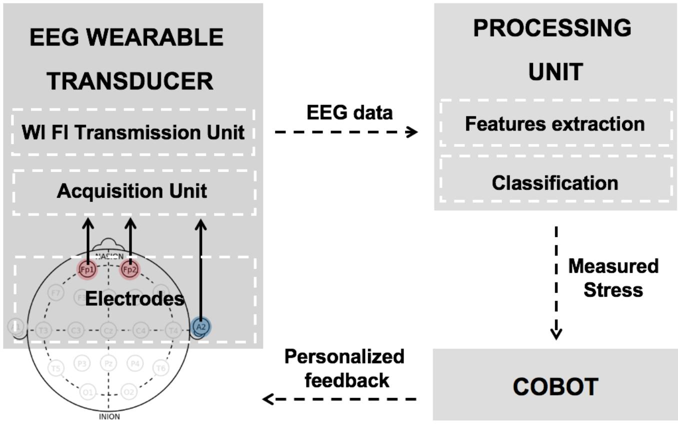


FIGURE 4.1: Architecture of the real-time stress monitoring instrument in Cobot interaction.

Architecture The architecture of the proposed instrument is highlighted in Fig. 4.10, in an example of interaction with a cobot. Prefrontal asymmetry is measured by two *Electrodes* as the difference of brainwaves from position FP1 and FP2, according to 10/20 system. The differential signal is referred to the earlobe. Analog signal is digitized by the *Acquisition Unit* and is sent, via wires, to the *Wi-Fi Transmission Unit*. Digital data arrives at the *Processing Unit* through wireless communication for real-time elaboration. Suitable *features* are extracted from each EEG record to compress data and increase significance. A *Classifier* receives the feature arrays, detects the stress condition, and assess its level. The measured stress is sent to the Cobot.

Operation The instrument allows to detect the onset and to assess the level of the stress arising from the concurrence of high mental load and negative emotional conditions, during the interaction with a Cobot. Once the worker fixes the electrodes on the forehead and on the earlobe, the Processing Unit interface allows to check signal quality both in time and in frequency domain. Subsequently, the stress measurement starts and the acquired data are sent in real time to the Processing Unit, by updating the user condition assessment every 2 s. Measurement results are sent to the Cobot in order to adapt its behavior to the worker stress conditions.

Feature extraction and classification Preliminary experiments in frequency domain highlighted poor accuracy results. Therefore, data analysis was carried out in the time domain. According to the state of the art [188], a EEG time window of 2 s was chosen as the optimal solution.

Feature Extraction was carried out by a standard machine learning technique, the Principal Component Analysis (PCA). This allows to compress data [238] and to approximate signals as a linear combination of a restricted number of orthogonal components. Therefore, data variance is most efficiently explained. Accordingly, a multi-variable signal can be represented as a smaller number of coefficients of the linear combination of the components. PCA also performs a filter function, because it highlights the components with maximum variance (information) of the data. Therefore, selecting only the components with the greatest variability improves signal-to-noise ratio.

For the classifier design, a linear separability test of the data was carried out by an euclidean distance-based K-means algorithm with low computational burden [239]. If a problem is linearly separable, a nonlinear classifier complicates the model unnecessarily and makes the correct learning of the classifier parameters less effective [240]. K-means algorithm estimates k means (centroids) in order to partition data into k clusters where each observation belongs to the cluster with the nearest mean. Then, in case of few outliers, a linear classifier is justified. Therefore, a preliminary analysis was realized.

Preliminary Analysis Ten subjects were divided into two classes with different stress level: (i) *control group*, only cognitive load, and (ii) *experimental group*, cognitive load but with negative emotions. Data were recorded during all the tests with a differential single-channel digitizer, sampling at 256 Sa/s. The signal was elaborated in time domain and without artifact filtering according to [140]. For each volunteer, two EEG tracks of 20 s were processed and divided into 2–s records of 512 samples. The resulting matrix 200x512 was divided in two clusters using the standard K-means algorithm with $K = 2$. In Fig. 4.2, the result of the clustering algorithm is reported. The two experimental groups were separated

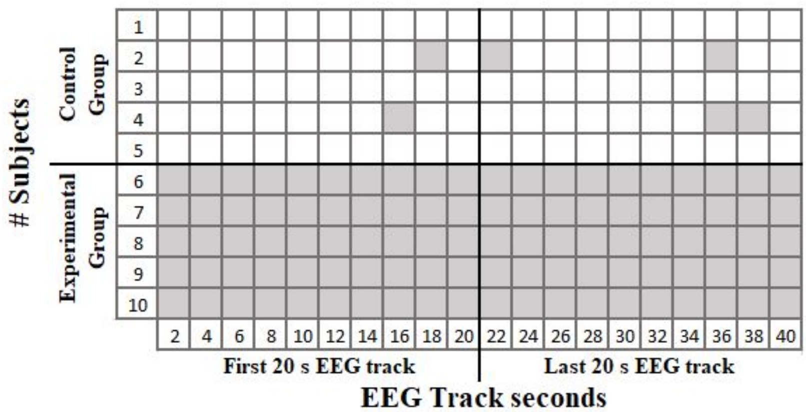


FIGURE 4.2: K-means classification (white: class 1; black: class 2) among the 2 different time phases according to subjects belonging group.

by K-means almost cleanly: on the first five rows, the arrays of the experimental group records, and on the other rows, the ones of the control group. These results suggested that a linear classifier discriminates the points of the two groups adequately. As a consequence, a successful and well-claimed method, the Support Vector Machine (SVM), with linear Kernel was used.

4.1.2 Realization

Hardware

Data Acquisition Unit It is based on the differential single-channel 10-bit digitizer EEG-SMT by Olimex, with maximum sampling rate of 256 Sa/s, an EEG amplifier, and an Atmel ATmega16 Alf and Vegard Reduced Instruction Set Computer processor microcontroller. The gain of the analog-to-digital converter (ADC) of the transducer was set to be 6427 V/V. A right-leg driver [driven rightled (circuit) (DRL)] signal increases the common-mode noise rejection. Universal serial bus is used for both data communication and powering. Moreover, the EEG-SMT has an analog three stages pass-band filter from 0.16 to 59.00 Hz. A previous work proved its suitability for wearable, low-cost, and non-invasive brain activity monitoring, by means of a single differential channel [51].

Dry Electrodes Brain signals are acquired by two dry active electrodes (Olimex EEG-AE), coated with a thin layer of silver chloride to guarantee the best contact impedance. The contact surface is extended by pins of conductive material. In this way, the quality of the acquired signal is preserved even with the electrode on a thick layer of hair. The reference passive dry electrode (Olimex EEG-PE) was applied to the earlobe. The electrodes on the user's forehead are fixed with a tight headband. The electrode on the earlobe is fixed with a clip, to ensure electrical connection.

Transmission unit A Wi-Fi communication channel was implemented to enhance wearability, throughout a Raspberry Pi 3 single-board computer, used as server, connected via UART to the EEG-SMT. The Raspberry Pi 3 uses a BCM43438 wireless chip and operates at ISM frequency bands (2.4 GHz).

Signal processing and classification In time domain, EEG tracks are divided into 2–s records of 512 samples. In this way, raw data are composed of 512 features, i.e. each feature corresponds to just one sample. Then, a feature reduction process is realized by PCA. The first four Principal Components are considered as input in the successive classification step. A trained linear kernel SVM classifier distinguishes records of a stressed or no stressed subject. The length of records determines the latency of 2 s.

Software

Raspberry The EEG signals, digitized by the EEG-SMT Olimex, are acquired by the Raspberry via UART by means of a dedicate software in C and installed on the Raspberry Pi 3. The baud rate is set to 57600 bit/s, with packet size 8, without parity bit. The Raspberry Pi 3 acts also as a Wi-Fi server, receiving from the EEG-SMT the command of start of the acquisition, and sending to the computer station the acquired data. This allows the users to freely move during real life. In view of a stand-alone device release, the computational power guaranteed by the raspberry allows processing to be carried out directly on board.

Processing Unit A specifically designed Matlab graphical user interface (GUI) allows easy interaction with Olimex EEG-SMT, through graphical icons and visual indicators. Moreover, by observing the display windows, EEG signal can be monitored both in time and frequency domain. Meanwhile, a Matlab script implements the linear kernel SVM.

4.1.3 Experimental Setup

Seventeen volunteers underwent an initial screening test administered by the psychologist. Seven participants were excluded from the experiment owing to excess in smoke, high score in anxiety and depression at questionnaires, and low performance at short memory tests. Therefore, ten healthy young volunteers (average age 25 years) of whom five women and five men, participated in the study. The informed consent, containing all the information about the experiment, was provided and signed by the subjects. The protocol was explained by the psychologist. Participants were divided equally into control and experimental groups, to complete a task, which induces mental load, together with (experimental group) or without (control group) negative social feedback. In particular, the Stroop Color and Word Test (SCWT) [241], a neuropsychological test extensively used for both experimental and clinical purposes, aimed to challenge subject using a complex cognitive task. In this test, subjects are required to read as fast as possible color-words printed in an inconsistent color ink, and to name the color of the ink instead of reading the word. This is to be done in a limited time punctuated by the psychologist who also gave information about errors during the performance. Environment was specifically designed in order to stress participants, by means of an attractive prize and an extremely out of range performance. Before and after the Stroop Test, subjects were required to complete two questionnaires: (i) STAI State form [242], to evaluate current anxiety state, and (ii) Rosemberg inventory [243], to assess participants' self-esteem. In them, they had to reflect their emotions in the specific moments during their exercise. Moreover, at the end of experimental tests, participants filled a rating of the experience in the Likert scale. The two groups, experimental and control, were subjected to the same protocols, but only the experimental group was stressed emotionally. During the experiment, the device did not annoy or distract the subject. After each trial, the psychologist asked for feedbacks in order to ensure the safety of participants. They did not experience any discomfort related to the electrode band; after a few minutes, they no longer noticed the device. The most significant 40 s were extracted from each individual test of 180 s. The initial and concluding stages are potentially the most inhomogeneous among them, that is, the most challenging in order to find a regularity, intra individual and even more intra group. The first 10 s of the test, regarded as cognitive warm up, were excluded. Therefore, only the later 20 s were deemed. The final 10 s were discarded, due to observations of the psychologist. The specialist noticed that some subjects showed a renouncing attitude, once realized the impossibility to complete the task. Hence, the previous 20 s were considered. Subsequently, for each subject, the two 20-s EEG tracks were divided in 2-s records. Each record is characterized by 512 time domain features, i.e. the 512 samples contained in 2 s. The total number of records were 200, namely 20 records for 10 subjects of 2 s each. Five subjects were taken from control group and five from experimental (stressed) group. In this way, the total EEG-samples from each group were 51,200. A matrix with 200 records on the rows was obtained by placing the first 100 records referred to the initial and final 20 s stress of control group and subsequently 100 records related to initial and final 20 s of experimental group.

Psychological validation A unique stress index was estimated as sum of normalized indexes to assess the general stress induced to participants. The indexes of performance, anxiety, self-esteem, perceived stress, and motivation, were obtained from parametric STAI and Rosemberg tests, as well as from task performance. One-way ANOVA was used to evaluate stress and motivation indexes on groups with a significance level α=0.05. The experimental group was more stressed compared to the control group, as evidenced by the One-way ANOVA (F=7.49; p=0.026). Instead, any significant difference between gender was noticed. A relevant difference in motivation between groups (F=14.52; p=0.005) showed that control group was more motivated than experimental group at the end of the experiment. Tab. 4.1 shows that, once arranged the stress index in decreasing order, the experimental group is more stressed than control one.

TABLE 4.1: Stress index distribution (descending sort)

Subject	Stress Index	Group
1	1,68	Experimental
2	1,66	Experimental
3	1,52	Experimental
4	1,38	Control
5	1,21	Experimental
6	0,77	Experimental
7	0,69	Control
8	0,54	Control
9	0,14	Control
10	-0,06	Control

Stress Classification Four different machine learning classifiers were used for validating the proposed method, by distinguishing stressed subject signals from no-stressed subject signals: (i) SVM (linear Kernel), (ii) k-nearest neighbors (n_neighbors = 9), (iii) Random Forest (criterion = 'gini', max_depth = 118, min_samples_split = 49) , and (iv) ANN (one hidden layer, activation function for hidden node = hyperbolic tangent, loss function = cross entropy cost, post processing = soft max, training algorithm = Resilient Propagation). In tab 4.2 the optimized iperparameters for each classifier are reported.

TABLE 4.2: Classifier optimized iperparameters and range of variation

Classifier	Iperparameter	Variation range
SVM	Cost parameter (C)	[0.1, 10.1] step = 1.0
Random Forest	n_estimators	{90, 180, 270, 360, 450}
k-NN	n_neighbors	[5, 15] step = 2
ANN	number of internal node	{25, 50, 100, 200}

The behavior of each classifier was also evaluated when the input was preprocessed by PCA.

Importantly, a subject-wise leave-two-out cross-validation evaluation was uniformly conducted in all the experiments in order to build a model capable of generalizing to new subjects. In case of small dataset according to [244], the Leave-p-out cross-validation (LPOCV) guarantees better statistical significance with respect to Leave-one-out cross-validation (LOOCV). Applying LOOCV to our dataset, the cross-validation process is repeated for k = 10 times, i.e. k = n (the number of subjects in the original sample). Instead, LPOCV requires training and validating the model C_p^n times, where C_p^n is the binomial coefficient, n the number of subjects in the original sample, and p is the number of subjects reserved only for the test. In our case (leave-two-out) the two subjects always belong to different groups (experimental vs control). In this way, a higher statistical significance was obtained (k = 25), by keeping training and test datasets balanced concerning the two classes. Therefore, for each iteration, one subject for group was left out from training set and used in the test set.

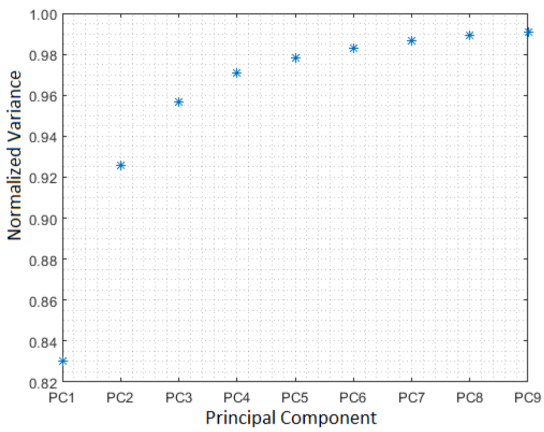


FIGURE 4.3: Cumulative Explained Variance in the PCA.

PCA Analysis Each classifier was fed with both raw data (2-s EEG epoch) and PCA pre-processed data. In particular, for each iteration of the LPOCV method [244] the first p principal components were computed on the training set. Then both training and test set were projected on them. Finally, the reduced representations of both data sets were input to the classifiers. The number of principal

TABLE 4.3: Classifiers accuracy (mean and uncertainty percentage) in Original Data (O.D.) and Principal Components Hyperplanes

	SVM	Random Forest	k-NN	ANN
O.D.	97.5 ± 0.6	98.5 ± 0.3	98.5 ± 0.4	99.2 ± 3.1
PC1	90.5 ± 5.3	$98{,}6 \pm 0.3$	$98{,}9 \pm 0.3$	98.5 ± 3.8
PC2	78.5 ± 7.1	$98{,}8 \pm 0.2$	$98{,}0 \pm 0.5$	98.8 ± 3.9
PC3	93.2 ± 3.4	$98{,}4 \pm 0.5$	$98{,}5 \pm 0.4$	98.7 ± 4.3
PC4	98.3 ± 0.4	$98{,}9 \pm 0.3$	$98{,}5 \pm 0.4$	99.1 ± 2.4
PC5	97.8 ± 0.4	$98{,}8 \pm 0.5$	$98{,}5 \pm 0.4$	99.2 ± 2.8
PC6	97.4 ± 0.6	$98{,}4 \pm 0.5$	$98{,}5 \pm 0.4$	98.9 ± 3.3
PC7	97.8 ± 0.5	99.0 ± 0.4	$98{,}5 \pm 0.4$	98.9 ± 3.6
PC8	97.4 ± 0.6	$98{,}6 \pm 0.5$	$98{,}5 \pm 0.4$	99.0 ± 3.5
PC9	97.9 ± 0.5	$98{,}9 \pm 0.5$	$98{,}5 \pm 0.4$	98.9 ± 4.1

TABLE 4.4: F-measure test results for SVM (mean and uncertainty percentage)

		Precision (%)	Recall (%)
O.D. Hyperplane		$96{,}5 \pm 1{,}0$	$98{,}4 \pm 0{,}7$
P.C. Hyperplanes	PC1	$89{,}2 \pm 5{,}1$	$92{,}2 \pm 5{,}3$
	PC2	$81{,}1 \pm 6{,}2$	$81{,}1 \pm 6{,}7$
	PC3	$96{,}4 \pm 1{,}5$	$93{,}6 \pm 3{,}1$
	PC4	$98{,}2 \pm 0{,}5$	$98{,}5 \pm 0{,}7$
	PC5	$97{,}2 \pm 0{,}5$	$98{,}5 \pm 0{,}7$
	PC6	$96{,}4 \pm 1{,}1$	$98{,}5 \pm 0{,}7$
	PC7	$97{,}2 \pm 0{,}8$	$98{,}5 \pm 0{,}7$
	PC8	$96{,}4 \pm 1{,}1$	$98{,}5 \pm 0{,}7$
	PC9	$97{,}2 \pm 0{,}8$	$98{,}7 \pm 0{,}6$

components p was varied: $p \in \{0, 1, 2, \ldots, 9\}$, where $p = 0$ corresponds to consider original data without PCA. The cumulative explained variance by the first nine components is greater than 99%, when PCA is applied on the dataset as a whole (Fig. 4.3). Therefore, this result highlights an intrinsic dimensionality of the data actually equal to no more than 9 (with respect to 512) and, in this case, the use of PCA for features extraction is validated. The results of the cross-validation strategy are shown in Tab. 4.3, as mean and uncertainty, with and without PCA. The lowest average accuracy for data without PCA is obtained by SVM and is equal to 97.5%. An F-measure test was carried out to assess the classification performance of the worst classifier (SVM). Results are reported in Tab. 4.4.

SVM classification output when p=2 is shown in Fig. 4.4 in PCs space. The PCs plot shows vectors distribution with respect to Support Vector. In that, the diamonds are associated to the control group, while the circles represent the experimental group.

Even with a temporal resolution of 2 s, satisfying results can be obtained in discriminating stress conditions. Generally PCA allows to obtain comparable or better average accuracy when $p > 3$ and, correspondingly, a lower uncertainty. This last result suggests a better noise robustness with PCA.

In bi-dimensional case, PCA highlights that the variance of the control group

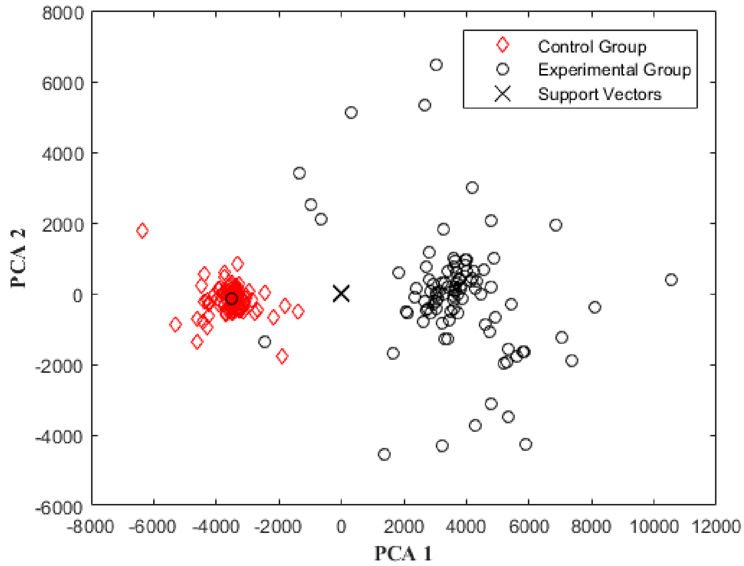


FIGURE 4.4: SVM data distribution in PCs space, p=2, 92,6% of explained variance.

is lower. Among the two groups, a significant difference in dispersion around the mean value as well as amplitudes comes out. Results of Fig. 2 confirm the data separability, even using only the first principal component, capable of explaining almost the 90% of variance. The good correlation between psychometric data and the exposure to different experimental set up (emotionally stressful and not) founds the experimental set up reliability in conditioning participants with regard to the study variable. The high accuracy level suggests that the signal acquired through a single channel preserves the information concerning the frontal asymmetry elicited from an emotional stress condition.

Noise and Bias Robustness Noise robustness was tested on the worst classifier (SVM) in order to assess the robustness of the proposed method. The subject-wise leave-two-out cross-validation strategy was repeated but with a further noise parameter, both (i) to make more generic the proposed method, and (ii) to verify the occurrence of possible bias during acquisition. The second evaluation is aimed to test the noise robustness of classification accuracy after PCA. In particular, two different kinds of noise were considered. In the first test, aimed at generalization, a random Gaussian noise with zero-mean and $\sigma \in \{0.04, 0.08, 0.12, 0.16, 0.20\}$, multiplied by the absolute value of the data maximum, was added. Results are reported in Tab. 4.5.

In the second test, aimed to verify bias, a constant value was added to each subject signal of the test sets. In this way, the signal of each subject was treated with a different random bias. Bias levels were chosen randomly within intervals of increasing amplitude ($\sigma \in [0.04 - 0.20]$, step = 0.04). For this reason, the global effect on the entire data set is noise. Results are reported in Tab. **??**.

TABLE 4.5: Accuracy (mean and uncertainty) in Original Data (O.D.) and Principal Components Hyperplanes at varying amplitude of random gaussian noise

	Noise σ percentage value				
	4	8	12	16	20
O. D.	97.9 $\pm$ 0.5	97.0 $\pm$ 0.7	96.1 $\pm$ 0.8	95.2 $\pm$ 0.9	92.2 $\pm$ 1.2
PC1	90.1 $\pm$ 5.3	89.7 $\pm$ 5.2	88.2 $\pm$ 5.2	86.4 $\pm$ 5.1	84.1 $\pm$ 4.8
PC2	79.0 $\pm$ 7.0	77.9 $\pm$ 7.1	77.5 $\pm$ 6.7	74.7 $\pm$ 6.6	74.5 $\pm$ 6.5
PC3	93.2 $\pm$ 3.4	92.4 $\pm$ 3.5	88.7 $\pm$ 3.4	85.7 $\pm$ 3.3	84.4 $\pm$ 3.4
PC4	98.2 $\pm$ 0.4	97.1 $\pm$ 0.7	95.9 $\pm$ 0.7	92.7 $\pm$ 0.9	90.8 $\pm$ 1.1
PC5	97.6 $\pm$ 0.4	97.2 $\pm$ 0.5	94.6 $\pm$ 0.7	91.6 $\pm$ 0.1	89.9 $\pm$ 0.1
PC6	97.7 $\pm$ 0.6	96.6 $\pm$ 0.7	95.1 $\pm$ 1.0	93.3 $\pm$ 1.0	90.4 $\pm$ 1.1
PC7	97.9 $\pm$ 0.5	96.8 $\pm$ 0.7	96.2 $\pm$ 0.8	93.7 $\pm$ 0.9	91.6 $\pm$ 0.1
PC8	97.5 $\pm$ 0.6	96.8 $\pm$ 0.6	96.3 $\pm$ 0.7	93.5 $\pm$ 0.1	90.5 $\pm$ 0.1
PC9	98.1 $\pm$ 0.5	97.2 $\pm$ 0.6	96.6 $\pm$ 0.6	93.7 $\pm$ 0.9	91.9 $\pm$ 0.1

In both the cases, the instrument shows good noise robustness. The classifier with PCA performs better if the noise level is less than 12% of absolute value of the maximum of data. Performance degrades in any case with higher noise levels. In this study, the differential channel and the PCA are exploited to face the problem of artifacts. A differential channel intrinsically rejects the common mode noise. PCA on the differential channel EEG acts like a pass band filter owing to its intrinsic reduction of the signal dimensionality in the PC domain. The combined effect of this two filtering effects improves the signal-to-noise ratio significantly. The experimental analysis of noise robustness validated, ex post, the proposed method.

4.2 Engagement in rehabilitation

4.2.1 Basic Ideas

The aim of this study is to propose an EEG-based engagement detection system in the field of pediatric rehabilitation. The basic ideas of the proposed method are:

- *The use of both the emotional and the cognitive engagement*: an overcoming of the reductionist approach based only on the cognitive dimension, which is particularly unsuitable for children, is proposed.

- *Employment of a low-cost, portable and wireless EEG device*: the goal of engagement assessment was realized using off the shelf components. The wearability was guaranteed by wireless data transmission.

- *Adoption of a subject-dependent approach*: the low inter-individual EEG reproducibility significantly influences the pattern classification in the engagement detection systems [245].

- *Support procedure for user calibration*: the system needs a calibration. To this aim, the user executes a set of rehabilitation sessions on different days. An observational non-interventional protocol is the best choice for maximizing children's comfort. However, this can lead to unbalanced data and a more challenging classifier training phase is required. The recent KMeansSMOTE method [246] is proposed to manage the imbalance of data.

4.2.2 Methods

Architecture The proposed method is sketched in Fig. 4.5. The *semi-wet 14 channel EEG device* allows the EEG signals to be sensed directly from the scalp of the child. Channels are referred to CMS/DRL. Analog signals are conditioned by stages of amplification and filtering (*Analog Filter and Amplifier*). Then, they are digitized by the Analog Digital Converter *ADC* and sent by the *Wireless Transmission Unit* to the *Data Processing* block. The *Classifiers* receive the feature arrays from two trained *Common Spatial Pattern* procedures for detecting the cognitive and emotional engagement.

Data processing In this section, *data preparation, training* and *classification* are presented.

[1] *Data preparation and training*: the EEG tracks are acquired at a sample rate of 128 Sa/s into time windows of 9 s without overlap. EEG signals are filtered through a 4th order Butterworth band-pass filter, between 0.5 Hz and 45 Hz. During the calibration, data are collected and properly labeled by the therapist. Both cognitive and emotional engagements are distinguished in two classes, high and low. Two Common Spatial Pattern procedures (CSP [247]) and two fully-connected feed-forward artificial neural network (ANN) classifiers, are separately trained on cognitive and emotional engagement data.

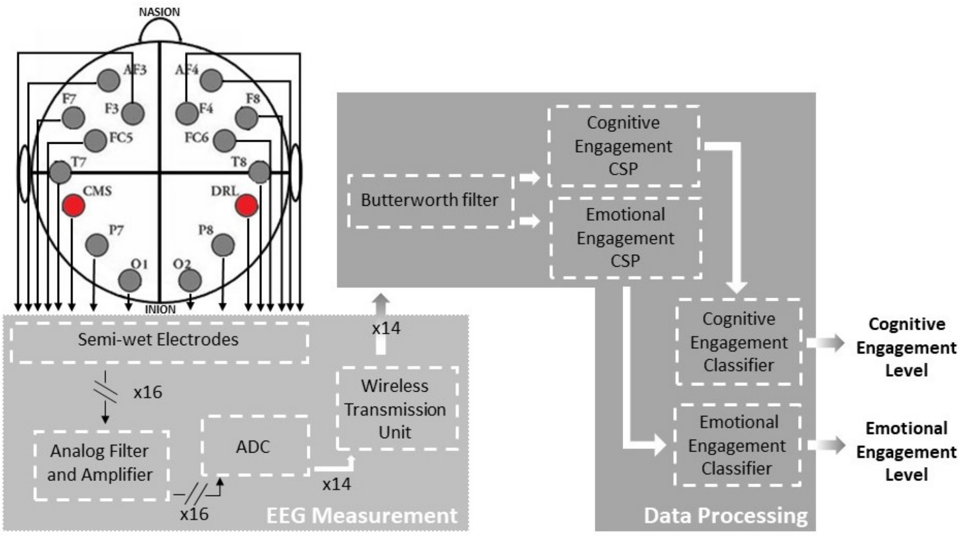


FIGURE 4.5: The proposed cognitive and emotional engagement detection method.

[2] *Classification*: the trained CSPs project multi-channel EEG data belonging to different classes into a new space, where the differences between the variances along the dimensions are maximized. The two trained ANNs for emotional and cognitive engagement classification are fed with the outputs of the previous stage (Fig. 4.5).

4.2.3 Experimental Setup

Sample Four children, three males and one female aged between 5 and 7 years, suffering from disturbances in motor-visual coordination, were selected for the experiment. Each subject was affected at least by one among the following diseases: double hemiplegia, motor skills deficit with dyspraxia, neuropsychomotricity delay, and severe neuropsychomotricity delay in spastic expression from perinatal suffering. Their main symptoms were: lack of strength, motor awkwardness, difficulty in maintaining balance, inadequate postures, spatial disorientation, problems with laterality (right, left confusion), difficulty in managing time, and learning difficulties.

Experimental setup The experimental protocol was approved by the ethical committee of the University Federico II. Families agreed to the experimental activities by releasing a written informed consent before the experiment. Procedures were carried out according to relevant guidelines and regulations [248]. An observational non-interventional protocol maximized the children's comfort. Therefore, part of the ordinary rehabilitation sessions was monitored by EEG for a total of about thirty minutes per week for each subject. The data acquisition took place in a room illuminated by natural light and provided with air exchange.

The adopted therapeutic approach was the Perfetti-Puccini method, also known as Cognitive Therapeutic Exercise [249]. The method aims to recover the injury and activate the brain circuits that govern movement. The child was asked to perform a visual attention exercise while keeping the correct posture of the trunk, neck, and head. An interactive environment [250] was depicted on a screen placed at the eye level of the subject (Fig. 4.6). One of four characters (a bee, a ladybug, a girl, or a little fish) could be chosen to make the game more interesting. The child had to stare at the character on the screen to make it move while maintaining eye contact. Dynamic tracking techniques were employed. The game allowed to set (i) the direction of the character's movement (from right to left and vice versa, or from top to bottom and vice versa), and (ii) the background landscape, to adapt the difficulty level to the patient's needs. A background music was inserted into the game to improve the child engagement. The game provided some features to adapt the therapy to the state of the subject: (i) a simplification of the exercise, (ii) the introduction of elements of novelty, and (iii) a content change.

FIGURE 4.6: Neuromotor rehabilitation session.

Several professional figures contributed to the experimental activity. Physiotherapists explained the exercise to the child (before the first session only), supervised rehabilitation, and helped the child maintaining eye contact and correct posture. A software engineer was responsible for starting the system and saving the data. A biomedical engineer was responsible for the EEG signal acquisition system and, therefore, for the correct setting-up, placement of the device, and electrode-skin quality contact.

Metrological reference Each session was video-recorded by two cameras (front and side framing).

The Pediatric Assessment of Rehabilitation Engagement (PARE) scale was employed for labeling the EEG signals. The emotional, cognitive, and behavioral components of engagement were expressed in terms of: participation, attention, activation, understanding, positive reactions, interest and enthusiasm, posture and movements of the child during the exercise, on a scale from 0 to 4. The PARE scale allowed to assess the rehabilitation session as a whole. The items of the scale were rearranged to be employed in shorter time intervals with the aim of improving the temporal resolution of observations. The behavioral component of engagement cannot be assessed starting from the EEG signal. Therefore, only the cognitive and emotional components of engagement are considered for research purposes. The items referring to the emotional and cognitive spheres were separately grouped. The evaluations were made by a multidisciplinary team while viewing the videos. The evaluators were asked to rate both the components of the engagement on two levels: high/low emotional engagement and high/low cognitive engagement. They also noted the status changes of the emotional and cognitive engagement and the correspondent time instants of occurrence. The consensus among the evaluators was statistically analyzed. The results revealed a total consensus of 95.2 % [251]. Evaluations were used as ground-truth to label the EEG dataset.

Experimental Results In Tables 4.6 and 4.7, the overall averages of the intra-individual balanced accuracies and MCC scores, given by the adopted classifiers, are reported for the cognitive engagement and the emotional engagement, respectively. To better understand to what extend the oversampling strategy can affect the results, the experiments were repeated with or without the application of the oversampling method.

TABLE 4.6: Overall mean of the intra-individual performances on cognitive engagement using three different classifiers: the balanced accuracy (BA) and the Matthews correlation coefficient (MCC) at varying the oversampling methods.

Oversampling	Metric	k-NN	SVM	ANN	Mean
none	BA	67.1	67.4	73.7	69.4 ± 3.0
	MCC	0.31	0.34	0.45	0.36 ± 0.06
SMOTE	BA	68.6	69.8	72.0	70.1 ± 1.4
	MCC	0.33	0.36	0.40	0.36 ± 0.03
BorderlineSMOTE	BA	70.3	70.9	73.6	71.6 ± 1.4
	MCC	0.36	0.38	0.43	0.39 ± 0.03
ADASYN	BA	68.1	68.3	72.5	69.6 ± 2.0
	MCC	0.33	0.33	0.42	0.36 ± 0.04
SVMSMOTE	BA	69.0	69.4	72.9	70.4 ± 1.7
	MCC	0.34	0.36	0.42	0.37 ± 0.03
KMeansSMOTE	BA	69.8	71.1	74.5	$\mathbf{71.8 \pm 1.98}$
	MCC	0.35	0.39	0.46	$\mathbf{0.39 \pm 0.04}$

As regards cognitive engagement, the oversampling method gave a slight improvement; as regards emotional engagement, the oversampling method gave a significant improvement to the performances, especially when the KMeansSMOTE method was employed.

The KmeansSMOTE is less likely to generate minority class data in domain areas predominantly dominated by majority class data. Thus, generated data are closer to the data of the minority class, as showed in Fig. 4.7 where the training data of a highly-unbalanced subject are shown using the t-SNE projection [252]. The data is oversampled with two methods: SMOTE (Fig. 4.7 A) and KMeansSMOTE (Fig. 4.7 B). The latter attenuates the noise thanks to clustering before data interpolation.

TABLE 4.7: Overall mean of the intra-individual performances on emotional engagement using three different classifiers: the balanced accuracy (BA) and the Matthews correlation coefficient (MCC) at varying the oversampling methods.

Oversampling	Metric	k-NN	SVM	ANN	Mean
none	BA	56.3	57.0	61.4	58.2 ± 2.2
	MCC	0.16	0.20	0.26	0.21 ± 0.04
SMOTE	BA	57.6	61.2	67.1	62 ± 3.9
	MCC	0.16	0.24	0.35	0.25 ± 0.08
BorderlineSMOTE	BA	57.3	60.2	66.5	61.3 ± 3.8
	MCC	0.15	0.22	0.34	0.24 ± 0.08
ADASYN	BA	57.0	60.0	67.4	61.5 ± 4.4
	MCC	0.15	0.21	0.36	0.24 ± 0.09
SVMSMOTE	BA	57.3	61.0	64.4	60.9 ± 2.9
	MCC	0.15	0.25	0.31	0.24 ± 0.06
KMeansSMOTE	BA	57.9	63.6	71.2	$\mathbf{64.23 \pm 5.4}$
	MCC	0.18	0.30	0.43	$\mathbf{0.30 \pm 0.10}$

Figures 4.8 and 4.9 show the intra-subjective balanced accuracies obtained both on cognitive and emotional engagement, respectively, using the KMeansSMOTE oversampling method. The ANN classifiers returned the better scores in most subjects, both in the emotional and cognitive engagement.

Furthermore, the MCC and the BA values ensure that the results are not affected by unbalancing bias in the test phase.

4.2.4 Discussion

The results reported in Tabs. 4.6 and 4.7 showed the amount of the improvement given by the oversampling methods in the proposed setup. More in detail, in the emotional engagement classification task, the improvements are more significant (e.g., an increase in accuracy of about 10 %) with respect to the cognitive engagement classification performance. This can be due to the different unbalancing ratios between the classes in the two tasks (i.e., a greater unbalanced data condition in the emotional engagement dataset with respect to the cognitive one). Indeed, in the proposed setup, the SMOTE algorithms generated greater amounts of data in case of strong unbalanced data condition having a greater impact on the classification performances. Therefore, also the cognitive dataset was artificially unbalanced to validate this hypothesis. To this aim, the number of samples was chosen so that the classes distribution was the same as the emotional engagement data. Next, an ANN classification step with and without KMeansSmote was carried out. The resulting performances without any oversampling strategy were 58.73 % and 0.25 for BA and MCC, respectively. Instead, BA and MCC increased to 65.14 % and 0.28, respectively, with KMeansSmote oversampling. The improvement given by KMeansSmote showed that the used oversampling strategy is particularly suitable for this type of data in case of imbalanced condition.
As concerns the data acquisition stage, Emotiv Epoch+ is only partially adaptable to different head sizes. Nevertheless, among the children involved in the experimental activity, the child with the smallest head exhibited an inion-naison

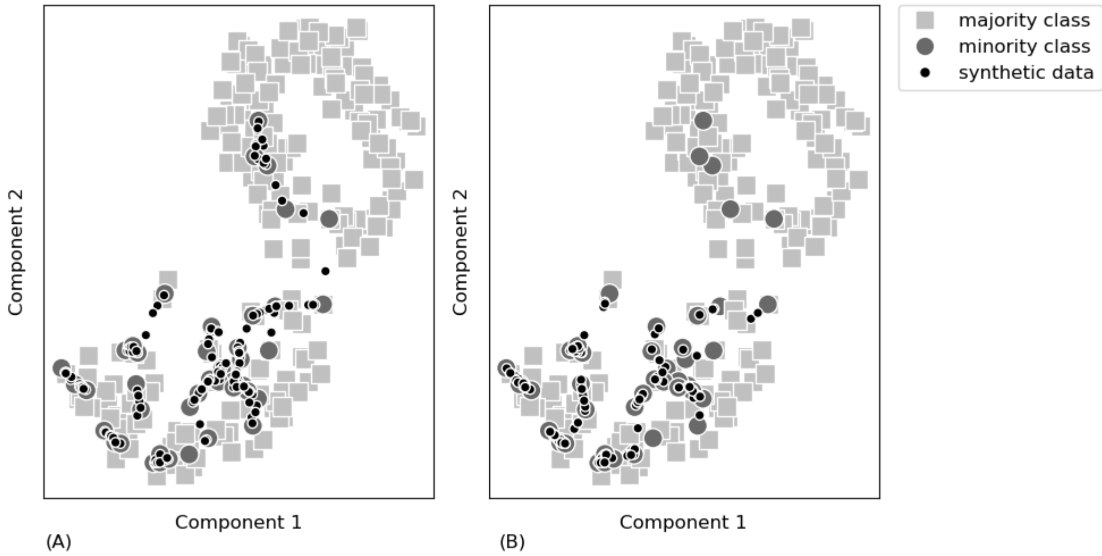


FIGURE 4.7: t-SNE projection of unbalanced EEG data (subject 4) oversampled with two different methods. The SMOTE method (A) randomically interpolates the data of the minority class. The KMeansSMOTE method (B) realizes a clustering before interpolation, attenuating the noise.

distance of 31.0 cm that is within the range of variation in adults of [31,0 - 38,0] cm, well established in literature [253]. By assuming that the manufacturer optimized the product for an average value of the inion-nasion distance of 34.5 cm in adults, in the case of a lower inion-nasion distance of 31.0 cm, the maximum electrode dislocation is about 1.4 cm with respect to the 10-20 International Positioning System. The maximum electrode position shift is appreciated in the frontal area and it gradually decreases until its disappearance, moving from the frontal area to the occipital area of the scalp. Therefore, the distance of each electrode from the reference of the 10-20 International Positioning System is to be considered in order to make reproducible the measurement. Despite the Emotiv Epoc+ device has the largest number of electrodes among the low-cost EEG devices available on the market, it does not guarantee a dense coverage of the parietal area of the scalp. The signal acquired in this area is particularly relevant for the assessment of the spatial attention [254, 255]. However, the device is equipped with 2 electrodes in the parietal areas (i.e., P7 and P8) and the spatial attention is only one component of engagement. Therefore, the engagement was adequately monitored and the measure was significant, as shown by the experimental results.
As regards the implications and applications of the proposed method, adaptivity is currently based on performance monitoring in the rehabilitation field. Characteristics not directly observable (such as patient engagement) are usually not taken into consideration. Conversely, the monitoring and the proper stimulation of patient engagement can strongly improve the effectiveness of the rehabilitation intervention. For example, in the framework of neuromotor rehabilitation, maintaining the attention focus on the exercises promotes neuronal neuroplasticity and motor recovery [22]. Therefore, monitoring cognitive engagement allows

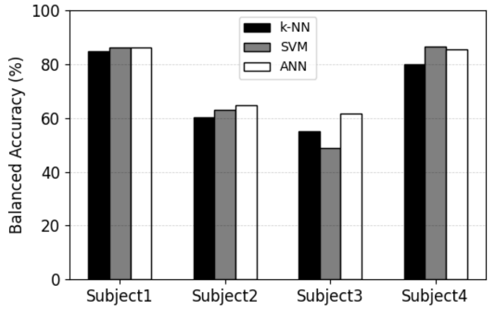


FIGURE 4.8: Cognitive engagement balanced accuracies for each subject based on KMeansSMOTE oversampling technique. Classifier performances are reported.

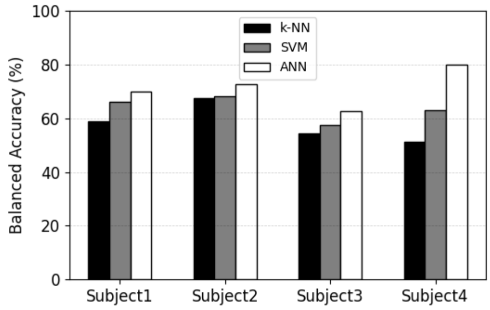


FIGURE 4.9: Emotional engagement balanced accuracies for each subject based on KMeansSMOTE oversampling technique. Classifier performances are reported.

automated systems to adopt appropriate countermeasures when distraction is detected [24].
Rehabilitation performance is also conditioned by the emotional engagement. A low performance may depend, for example, on a state of boredom or worry, rather than on a lack of skills. Chronic health disabilities are often stressors and the stress management is a crucial issue in rehabilitation [256]. The assessment of cognitive and emotional engagement allows to monitor stress levels [257] and to provide the automated rehabilitation platform useful information to better adapt to the user's needs.
Finally, the proposed approach is data driven. Thus, it can be applied flexibly to different targets by identifying *ad-hoc* models suitable for different abled groups.

4.3 Engagement detection in learning

This Section describes an EEG-based cognitive and emotional engagement detection method during a learning task. In this section the *basic ideas*, the *architecture*, and the adopted *processing framework* are outlined.

4.3.1 Basic Ideas

The proposed method is based on the following key concepts:

- *EEG-based subject-adaptative system*: new input channels (EEG) to the Intelligent Teaching Systems enhance the adaptivity to the user in the context of learning 4.0.
- *Cognitive and emotional learning engagement detection*: the assessment of student engagement is realized considering both cognitive and emotional aspects, according to the Frederiks theory [78].
- *Within and cross-subject designs*: both the approaches are experimentally validated in order to pursue accuracy maximization or calibration-time minimization, respectively.
- *Domain Adaptation procedure in cross-subject case*: a Transfer Component Analysis (TCA) [258] allows to use knowledge acquired about other subjects to simplify the system calibration on a new subject.
- *Wearable system*: an ultralight wireless EEG device with few and dry electrodes maximizes the wearability.
- *Multi-factorial metrological reference*: the system is calibrated by using (i) standardized strategies for inducing different levels of cognitive load, and (ii) a public acoustic stimuli dataset to elicit emotions. Moreover, the metrological reference of emotional engagement was confirmed by statistical analysis on the outputs of self-assessment questionnaires.
- *Narrow EEG frequency intervals*: the EEG features resolution is improved by a 12-band Filter-Bank, obtained by sub-dividing the traditional EEG six bands (delta, theta, sigma, alpha, beta, and gamma).

4.3.2 Architecture

The architecture of the proposed system is depicted in Fig. 4.10. The eight *Active Dry Electrodes* acquire the EEG signals directly from the scalp. Each channel is differential with respect to AFz (REF), and referred to Fpz (GND), according to 10/20 international system. After transduction, analog signals are conditioned by the *Analog Front End*. Next, they are digitized by the *Analog Digital Converter* (ADC), and submit an *Artifact removal block* performed by an ICA based algorithm. Then the signals are sent by the wireless Bluetooth transmission to the

Data Processing stage. Here, the suitable feature are extracted by a 12-component *Filter Bank*. The two *Support Vector Machine* (SVM) classifiers receive the features array from two trained *Common Spatial Pattern* (CSP) algorithms for detecting the Cognitive and the Emotional Engagement respectively. Only in the cross-subject case, a baseline removal followed by a TCA procedure is provided during the training stage of the classifier.

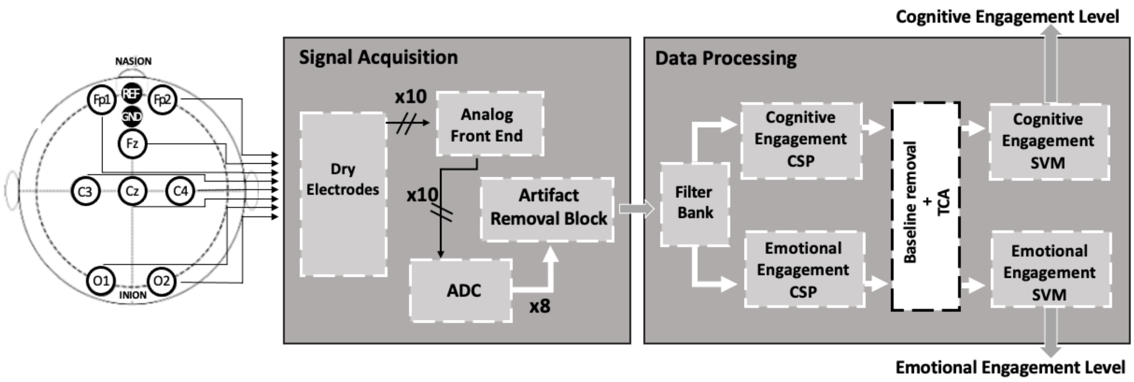


FIGURE 4.10: The architecture of the system for engagement assessment; the white box is active only in the cross-subject case (ADC - Analog Digital Converter, CSP - Common Spatial Pattern, TCA - Transfer Component Analysis, and SVM - Support Vector Machine).

4.3.3 Processing Framework

In this section, (i) the *feature extraction and selection*, the (ii) *baseline removal and Domain Adaptation*, and (iii) the *classification* are detailed.

Feature extraction and selection In this work, a novel Filter Bank version [24] is adopted. EEG signals are acquired by an eight channels device with sample rate of 512 Sa/s.

The acquired signals are then filtered by a filter bank composed of 12 infinite impulse response (IIR) band-pass Chebyshev type 2 filters with 4 Hz amplitude, equally spaced from 0.5 to 48.5 Hz. Then, epochs are extracted using a time window of 3 s with an overlap of 1.5 s.

Then, a Common Spatial Pattern (CSP) [39] is applied. In a binary problem, CSP works by computing the covariance matrices related to the two classes, simultaneously diagonalized such that the eigenvalues of two covariance matrices sum up to 1. Afterwards, a matrix is computed to project the input into a space where the differences between the class variances are maximized. More precisely, in a binary problem, the projected components are sorted by variances in a decreasing or ascending order: the former, when the projection matrix is applied to inputs belonging to the first class, while the latter when inputs belong to the second class [225].

Baseline removal and Domain Adaptation A cross-subject approach has several advantages with respect to a within-subject one, such as the reduction of time for the initial calibration procedure. Unfortunately, the non-stationarity nature of the EEG signal leads to a greater data variability between subjects. This is a well-known problem in the literature, which makes the cross-subject approach a very challenging task [259]. Currently, the Domain Adaptation methods [260] are obtaining a great attention from the scientific community. In this work, the Transfer Component Analisys (TCA) [258] is adopted. TCA is a well-established technique of domain adaptation already used in the EEG signal classification literature with promising results [259]. In a nutshell, TCA searches for a common latent space between data sampled from two different (but related) data distributions by preserving data properties. More in detail, TCA searches for a data projection ϕ that minimizes the *Maximum Mean Discrepancy* (MMD) between the two distributions, that is:

$$||\frac{1}{n_S}\sum_{i=1}^{n_S}\phi(\vec{x}_{Si}) - \frac{1}{n_T}\sum_{i=1}^{n_T}\phi(\vec{x}_{Ti})||^2$$

where n_S and n_T are the numbers of points in the first (*source*) and the second (*target*) domain set respectively, while $\vec{x}_{S_i}$ and $\vec{x}_{Ti}$ are the $i-$th point (epoch) in the two different sets. The data projected in the new latent space are then used as input for the classification pipeline. However, TCA works with only two different domains, differently from a multiple-subject environment, which can lead to a domain composed of several sub-domains generated by the different subjects or sessions. In [259], TCA was tested by considering for the first domain a subset of samples from $N-1$ subjects, where N is the total number of subjects, and with the data of the remaining subject for the other domain. However, this approach does not take into consideration the fact that different subjects may belong to very different domains, leading to poor results. A simple solution consists in subtracting to each subject a baseline signal recorded from the user, for example, in rest condition. However, this last point requires new subject acquisition. Instead, in this work, an average of the signals for each subject is used as baseline, thus avoiding the need for new signal acquisitions.

Classification For the classification stage, Support Vector Machines (SVMs)[261] are implemented. Considering inputs as points in a vector space, SVM is a binary classifier which discriminates data according to a decision hyperplane. Differently from other hyperplane-based classifiers, an SVM finds the hyperplane maximizing the separation between the classes, i.e. the hyperplane having the largest distance from the *margins* of the classes.

4.3.4 Experimental Setup

Twenty-one school age subjects (9 males and 13 females, 23.7 ± 4.1 years) participated in the experiment. The ethical committee of the University of Naples Federico II approved the experimental protocol. All methods were performed in accordance with the relevant guidelines and regulations. Before the experiment, each subject read and signed the informed consent. All volunteers have no neurological diseases. Each subject was seated in a comfortable chair at a distance of 1 m from the computer screen. The location was sanitized before and after of each acquisition as indicated in the COVID-19 academic protocols. Each subject was equipped with a mouse to carry out the experimental test. After wearing the EEG-cap, the contact impedance was assessed to guarantee optimal signal-acquisition conditions. Each subject underwent an experimental session composed by 8 trials. Various stimuli to induce high and low levels of emotive and cognitive engagements were equally distributed among the trials. As stimulus modulating the cognitive engagement level an updated and revised Continuous Performance Test (CPT) [262] was administrated. In particular, a CPT version based on a learning by doing activity on how an interface works was adopted. Whereas, proper background music and social feedback was used to modulate the emotive engagement level . More in details, the three different stimuli are described as follows:

- *Revised CPT:* a red cross and a black circle on the computer screen were presented to the subject. The red cross tends to run out from the circle on the screen in random directions. The subject was asked to keep the cross inside the circle by using the mouse. For each trial, a different difficulty level was set by the experimenter changing the cross speed. The percentage of the time spent by the red cross inside the black circle with respect to the total time was reported to the subject at the end of the trial (Fig. 4.11).

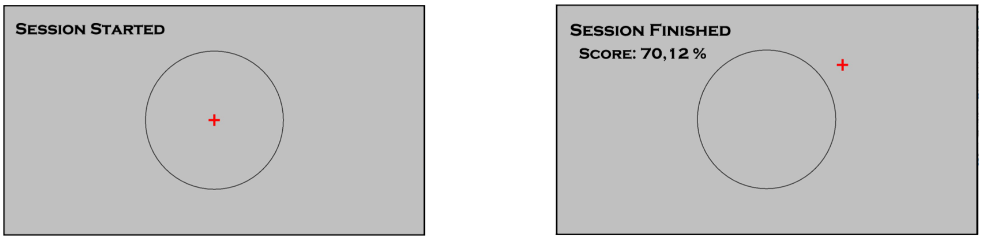
- *Background music:* for each trial, a particular emotive engagement level was favored by proper background music. The music tracks were randomly selected from the MER [263] database where songs are organized according to the 4 quadrants of the emotion Russell's circumplex model [8]. The songs associated with the Q1 and Q4 quadrants (*cheerful music*) were employed in high emotional engagement trials, Q2 and Q3 for the low ones (*sad music*).

- *Social feedbacks:* during each trial, the experimenters gave proper social feedbacks according to the emotive engagement levels under the experimental protocol. The positive and negative social feedbacks consisted of encouraging and disheartening comments respectively, given to subject on his/her ongoing performance. The social feedback effectiveness was improved by the simultaneous music background effects.

A well-founded metrological reference, is ensured by two assessment procedures validating the stimuli effectiveness were used:

- *performance index*: an empirical threshold was used to confirm that an appropriate CPT stimuli response was given by the participant. The threshold changed according to the trial difficulty level.

- *Self Assessment Manikin questionnaire (SAM)*: the emotional engagement level was assessed by a 9-level version of the SAM. The lower emotional engagement level was associated to the SAM score 1, while the greater one to 9.



(A) Session Started (B) Session Finished

FIGURE 4.11: Screen shots from the CPT game. At the beginning of the game (a),the cross starts to run away from the center of the black circumference. Theuser goal is to bring the cross back to the center by using the mouse. At theend of each trial (b), the score indicates the percentage time spent by thecross inside the circumference.

The experimental session started with the administration of the SAM to get information about the initial emotional condition of the subject.

Then, a preliminary CPT training phase to uniform all the participants starting levels was realized. After this preliminary phase, each trial was implemented by a succession of a CPT stage followed by a SAM administration.

Dataset building 45 s acquisition EEG signals were labeled according to two parameters: i) high or low emotional engagement, and ii) high or low cognitive engagement. More in detail, regarding the cognitive engagement, the trials were labeled according to the CPT speed [264, 169], since the higher was the speed the more the cognitive engagement increased [39, 169].

The trials having speed lower than 150 pixels/s were labeled as low_c whereas $high_c$, were assigned to the trials having speed higher than 300 pixels/s.

As concern the emotional engagement, the trials characterized by cheerful/sad music and positive/negative social feedback were labelled as $high_e/low_e$. For each trial, the SAM results (normalized to the initial pre-session value) were consistent with the proposed stimuli. In fact, a one-tailed t-student analysis revealed in the worst case a 0.02 P-value.

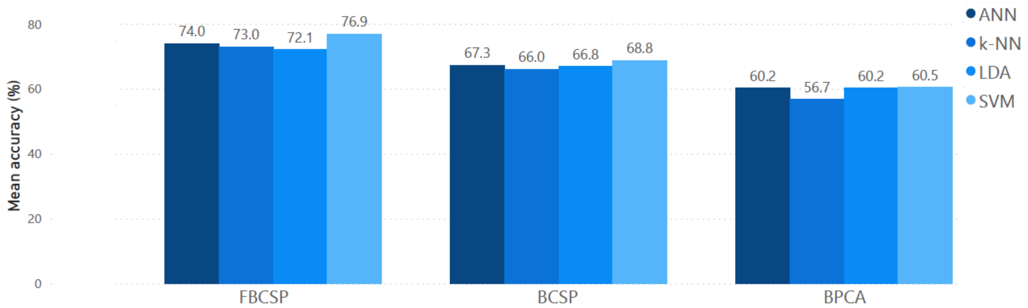
4.3.5 Experimental Results

In this section, the experimental results obtained in within- and cross-subject cases are reported. Firstly, to make a comparison with the classical literature approach, the engagement index proposed in [152] was used as feature for a classification of the cognitive engagement. Unfortunately, as highlighted by the results reported in Tab. 4.8 accuracy performances were not optimal. In fact, this feature is mainly used in non-predictive applications (e.g., [154]).

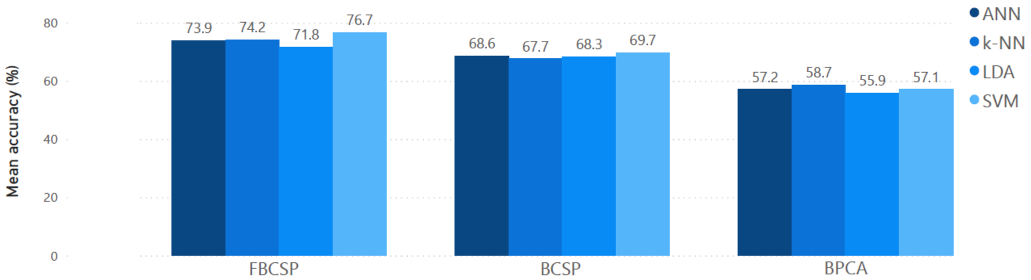
Instead, the best results both on cognitive and emotional engagements (Fig. 4.12) were achieved using features extracted by Filter-Bank and CSP.
Quantitative results related to the use of Filter Bank and CSP for each classifier

TABLE 4.8: Within-subject experimental results. Classification accuracies using the *Engagement Index* [152] for cognitive engagement classifications are reported.

Method	Cognitive Engagement
SVM	54.8 ± 4.9
k-NN	53.7 ± 5.7
ANN	53.1 ± 5.4
LDA	50.7 ± 6.2



(A) Cognitive engagement



(B) Emotional engagement

FIGURE 4.12: Within-subject performances of the compared processing techniques in (a) cognitive engagement and (b) emotional engagement detection. Each bar describes the average accuracy over all the subjects.

can be observed in Tab. 4.9: among the different classifiers, SVM stands out with a better performance than the others, reaching its best mean accuracies of 76.9 ± 10.2 on cognitive engagement classification and of 76.7 ± 10.0 on emotional engagement. Results are computed as the average accuracy over all the subjects.

The results reported in Fig. 4.12b show that the Filter Bank improves the classification performance in a significative way. This can be due to the use of several sub-bands which highlight the signal main characteristics, allowing the CSP computation to project the subject data in a more discriminative common space. In Fig. 4.13, BCSP and FBCSP are compared through t-SNE [265] on the subject data transformed using the two different methods. The figure shows that, for several subjects, CSP applied after FB projects the data in a space where they are easily separable with respect to the BCSP case. A t-SNE plot of the data first and after removing the average value of each subject is shown in Fig. 4.14. The data without for-subject average removal (Fig. 6a) are disposed in several clusters

TABLE 4.9: Within-subject experimental results. Accuracies are reported on data preprocessed using Filter Bank and CSP for cognitive engagement and emotional engagement classifications. The best performance average values are highlighted in bold.

Method	Cognitive Engagement (proposed)	Emotional Engagement (proposed)
SVM	**76.9 $\pm$ 10.2**	**76.7 $\pm$ 10.0**
k-NN	73.0 $\pm$ 9.7	74.2 $\pm$ 10.3
ANN	74.0 $\pm$ 9.2	73.9 $\pm$ 9.1
LDA	72.1 $\pm$ 11.4	71.6 $\pm$ 9.3

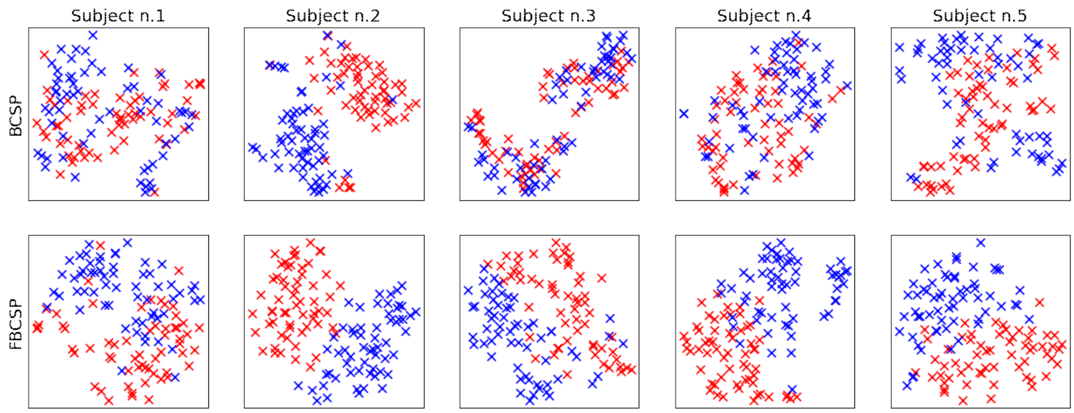
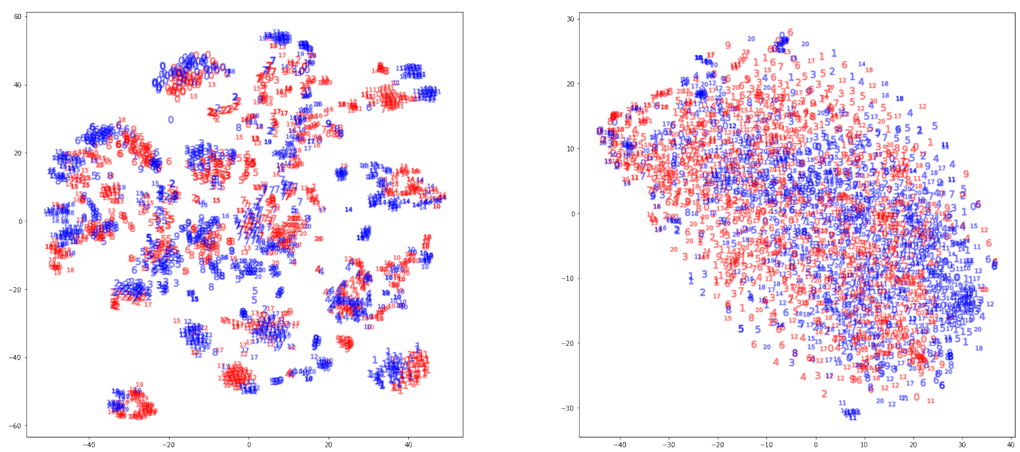


FIGURE 4.13: Filter Bank impact on the class (red and blue points) separability. t-SNE-based features plot of five subjects randomly sampled (first row: without Filter Bank; second row: with Filter Bank).

over the t-SNE space, exhibiting a fragmentation tendency. Instead, after the for-subject average removal (Fig. 6b), the data result more homogeneous, enhancing the model generalizability. A comparison using TCA with and without the for-subject average removal is made and the resulting performances are reported in Tab. 4.10. The results show that removing the for-subject average from each subject boosts the performance with respect to using TCA alone (more than 3 % of improvement in almost all classifiers, especially in Cognitive Engagement case).

Method	With For-Subject Average Removal		Without For-Subject Average Removal	
	Cognitive	Emotional	Cognitive	Emotional
SVM	**72.8 $\pm$ 0.11**	**66.2 $\pm$ 0.14**	64.0 $\pm$ 0.11	61.7 $\pm$ 0.10
k-NN	69.6 $\pm$ 0.11	61.9 $\pm$ 0.09	57.1 $\pm$ 0.09	56.9 $\pm$ 0.10
ANN	72.6 $\pm$ 0.12	65.7 $\pm$ 0.14	69.7 $\pm$ 0.12	65.8 $\pm$ 0.15
LDA	69.5 $\pm$ 0.12	65.3 $\pm$ 0.14	69.6 $\pm$ 0.13	64.6 $\pm$ 0.13

TABLE 4.10: Cross-subject experimental results using FBCSP followed by TCA. Accuracies are reported with and without for-subject average removal for cognitive and emotional engagement detection. The best performance values are highlighted in bold.



(A) (B)

FIGURE 4.14: A comparison using t-SNE of the FBCSP data first (a) and after (b) removing the average value of each subject, in the cross-subject approach.

Conclusions

Different passive BCI solutions were described from the design phase to the experimental validation phase. Wearability was always ensured by a low channel count and dry, or semi-wet electrodes. The main goals achieved are reported below for emotional valence, attention, engagement, and stress detection, respectively.

Emotional valence. The EEG-based system proposed for emotional-*valence* detection exhibited an accuracy of 96.1 % and 80.2 % in within-subject and cross-subject analysis, respectively. Important steps towards the measurability of emotions were proposed: Firstly, the Valence detection occurs along the interval scale theorized by the Circumplex Model of emotions. Thus, the current binary choice, positive valence vs negative valence, could represent a first step towards the adoption of a metric scale with a finer resolution. Secondly, the experimental sample was collected by managing the bias of depressive disorders. Finally, results from the *Self Assessment Manikin* questionnaire confirmed the compatibility of the experimental sample with that of *Oasis*. Hence, a metrological reference was built taking into account both the statistical strength of the data set OASIS and the collected data about the subject perception. The OASIS dataset was also subjected to a cross-cultural validity check.

A priori information is not needed using algorithms capable of extracting features from data through an appropriate spatial and frequency filtering. Classification is carried out with a time window of 2 s. The achieved performances are due to the combined use of a custom 12-band Filter Bank with CSP spatial filtering algorithm. This approach is widely used in the motor imagery field and was proven to be valid also for emotion recognition. The high ergonomics and accuracy are compatible with the principal applications of emotional valence recognition.

Attention-Distraction. The method for detecting a state of attention and distraction during the execution of a motor act shows experimentally a state-of-the-art mean accuracy of 92.8±1.6 % and a mean recall of 92.6%. Attention status classification is carried out on 3 s epochs. The level of performance achieved also arise from the use of a 12-filter custom Filter Bank which enhances the contributions of the significant EEG bands for attention analysis. The method turns out to be immediately usable in rehabilitation for offering to therapists: (i) a tool capable of assessing patients' attention levels towards the proposed exercises; and (ii) the possibility to implement strategies that, through the recovery of attention, increase the rehabilitation effectiveness.

Engagement in pediatric rehabilitation. A low cost EEG-based engagement (cognitive and emotional) detection system is proposed for pediatric rehabilitation. A subject-dependent approach is adopted and a specific easy calibration is provided for personalized medicine. Wearability is guaranteed by a wireless cap with semi-wet electrodes and 14 data acquisition channels. The proposed method, based on KMeansSMOTE and ANN, showed experimentally a mean balanced accuracy of $71.2\ \%$ and $74.5\ \%$ for the emotional and cognitive engagement, respectively. Furthermore, a comparison between several oversampling strategies was made, showing that the KMeansSMOTE can be a promising oversampling method for unbalanced EEG engagement datasets. Effective management of unbalanced dataset allows the implementation of observational non-interventional protocol. The KMeansSMOTE method is the core of the proposed calibration procedure, but also a promising technique for researchers focused on the observation of the spontaneous children behavior. The distance of each electrode from the reference of the 10-20 International Positioning System was noted to make the measurement reproducible, being reproducibility a quality parameter of the measurement itself.

Engagement in learning. The proposed system can be used in the context of Learning 4.0 as a new input channel of an adaptive automated teaching platform to improve the learning effectiveness. The system is validated on students during a training stage involving cognitive and motor skills and aimed to learn how to use a human-machine interface. Standard stimuli, performance indicator, and self assessment questionnaires were employed to guarantee a well founded metrologically reference. The proposed method, based on Filter Bank, CSP and SVM, experimentally showed the best performance. In particular, in the cross-subject case, an average accuracy of $72.8\ \%$ and $66.2\ \%$ was reached for the cognitive engagement and emotional engagement respectively by using TCA and for-subject average removal. Instead, in the within-subject case, an accuracy of $76.9\ \%$ and $76.7\ \%$ was reached for the cognitive engagement and emotional engagement, respectively.

Stress from interaction with cobots. A method to assess stress condition in real time trough a high-wearable EEG-based device was proposed. EEG signal amplitudes variations between prefrontal right and left zone were acquired through a single differential channel. The induced stress status was verified by a psychologist through (i) questionnaires administered before and after the stress test, and (ii) performance assessment. Time domain features were used in the classification procedure. Four standard machine learning classifiers (SVM, k-NN, Random Forest, and ANN) reached more than 90% accuracy in distinguishing each 2-s epoch of EEG. Generally, PCA allows to obtain a better noise robustness. The results show the adequacy of the proposed solution based on a single-acquisition channel and time domain-based feature selection. In the worst case, the SVM Linear -Kernel classifier succeeded in discriminating stress conditions with an accuracy of 97.5 $\pm$ 0.6% and a latency of 2 s. For latency above 4 s the accuracy reaches 100%. Noise robustness was tested in order to exclude the impact of bias during signal acquisition and to empower generality to the results. The proposed method gives a new way to detect prefrontal asymmetry traditionally associated to emotional stress condition.

Future developments of the research will be: (i) the development of the metrological foundation of mental states measurement (theoretical model, measurement unity, uncertainty analysis); (ii) a resolution improvement of the metric scale; (iii) combined use of different biosignals (besides EEG); (iv) a deep analysis on interactions among the number of electrodes, classifiers, and the accuracy; and (v) experiments on different processing strategies: in this thesis, the binary nature of the problems enhanced the classification performances of certain classifier. In future works aimed at increasing the metric scale resolution, other methods may result more effective (full-connected neural networks, Convolutional Neural Networks [145] etc.) for example in a regression-based perspective. In future works, new measurement solutions will be tested to guarantee more adaptivity to children's head size and more dense coverage of selected scalp area. Further future experimental activity with a larger number of subjects are necessary to consolidate the statistical significance of these preliminary results.). Electrode scalp locations used in this study, FP1 and FP2, are considered as sensitive to ocular artifacts. However, our experiments did not highlight this problem. In any case, further experimental campaigns will be carried out on new areas of the scalp. In this way, the impact on the classifier of the information produced by both the EEG signals and the eye movements will be deepen.

References

[1] Sulayman K Sowe et al. "Cyber-physical-human systems: Putting people in the loop". In: *IT professional* 18.1 (2016), pp. 10–13.

[2] Bogdan-Constantin Pirvu, Constantin-Bala Zamfirescu, and Dominic Gorecky. "Engineering insights from an anthropocentric cyber-physical system: A case study for an assembly station". In: *Mechatronics* 34 (2016), pp. 147–159.

[3] Gunar Schirner et al. "The future of human-in-the-loop cyber-physical systems". In: *Computer* 46.1 (2013), pp. 36–45.

[4] Iveta Zolotová et al. "Smart and cognitive solutions for Operator 4.0: Laboratory H-CPPS case studies". In: *Computers & Industrial Engineering* 139 (2020), p. 105471.

[5] Lun-De Liao et al. "A novel 16-channel wireless system for electroencephalography measurements with dry spring-loaded sensors". In: *IEEE Transactions on Instrumentation and Measurement* 63.6 (2014), pp. 1545–1555.

[6] Yu-Chun Chen, Bor-Shyh Lin, and Jeng-Shyang Pan. "Novel noncontact dry electrode with adaptive mechanical design for measuring EEG in a hairy site". In: *IEEE Transactions on Instrumentation and Measurement* 64.12 (2015), pp. 3361–3368.

[7] Paul R Kleinginna and Anne M Kleinginna. "A categorized list of emotion definitions, with suggestions for a consensual definition". In: *Motivation and emotion* 5.4 (1981), pp. 345–379.

[8] James A Russell. "A circumplex model of affect." In: *Journal of personality and social psychology* 39.6 (1980), p. 1161.

[9] Maya Golan, Yuval Cohen, and Gonen Singer. "A framework for operator–workstation interaction in Industry 4.0". In: *International Journal of Production Research* 58.8 (2020), pp. 2421–2432.

[10] Yisi Liu, Olga Sourina, and Minh Khoa Nguyen. "Real-time EEG-based emotion recognition and its applications". In: *Transactions on computational science XII*. Springer, 2011, pp. 256–277.

[11] Cristina Anamaria Pop et al. "Can the social robot Probo help children with autism to identify situation-based emotions? A series of single case experiments". In: *International Journal of Humanoid Robotics* 10.03 (2013), p. 1350025.

[12] Christian Jones and Jamie Sutherland. "Acoustic emotion recognition for affective computer gaming". In: *Affect and emotion in human-computer interaction*. Springer, 2008, pp. 209–219.

[13] Sergio Paradiso et al. "Cerebral blood flow changes associated with attribution of emotional valence to pleasant, unpleasant, and neutral visual stimuli in a PET study of normal subjects". In: *American Journal of Psychiatry* 156.10 (1999), pp. 1618–1629.

[14] Joao Perdiz, Gabriel Pires, and Urbano J Nunes. "Emotional state detection based on EMG and EOG biosignals: A short survey". In: *2017 IEEE 5th Portuguese Meeting on Bioengineering (ENBENG)*. IEEE. 2017, pp. 1–4.

[15] Mitchel Benovoy, Jeremy R Cooperstock, and Jordan Deitcher. "Biosignals analysis and its application in a performance setting". In: *Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing*. 2008, pp. 253–258.

[16] Andrea Apicella et al. "EEG-Based Detection of Emotional Valence towards a Reproducible Measurement of Emotions". In: *Scientific Reports* 11.21615 (2021).

[17] Roger Gassert and Volker Dietz. "Rehabilitation robots for the treatment of sensorimotor deficits: a neurophysiological perspective". In: *Journal of neuroengineering and rehabilitation* 15.1 (2018), pp. 1–15.

[18] Irene Aprile et al. "Robotic rehabilitation: an opportunity to improve cognitive functions in subjects with stroke. An explorative study". In: *Frontiers in Neurology* 11 (2020), p. 1498.

[19] Kai Keng Ang and Cuntai Guan. "Brain-Computer Interface in Stroke Rehabilitation". In: *Journal of Computing Science and Engineering* 7.2 (2013), pp. 139–146.

[20] Steven C Cramer et al. "Harnessing neuroplasticity for clinical applications". In: *Brain* 134.6 (2011), pp. 1591–1609.

[21] Genane Loheswaran et al. "Impairment of neuroplasticity in the dorsolateral prefrontal cortex by alcohol". In: *Scientific reports* 7.1 (2017), pp. 1–8.

[22] Keng Peng Tee et al. "Augmenting cognitive processes in robot-assisted motor rehabilitation". In: *2008 2nd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics*. IEEE. 2008, pp. 698–703.

[23] Karl Schweizer and Helfried Moosbrugger. "Attention and working memory as predictors of intelligence". In: *Intelligence* 32.4 (2004), pp. 329–347.

[24] Andrea Apicella et al. "High-wearable EEG-based distraction detection in motor rehabilitation". In: *Scientific Reports* 11.1 (2021), pp. 1–9.

[25] Carolyn Ranti et al. "Blink Rate patterns provide a Reliable Measure of individual engagement with Scene content". In: *Scientific Reports* 10.1 (2020), pp. 1–10.

[26] Guido HE Gendolla. "Self-relevance of performance, task difficulty, and task engagement assessed as cardiovascular response". In: *Motivation and Emotion* 23.1 (1999), pp. 45–66.

[27] Angela R Harrivel et al. "Monitoring attentional state with fNIRS". In: *Frontiers in human neuroscience* 7 (2013), p. 861.

[28] Ehsan T Esfahani et al. "Adaptation of rehabilitation system based on user's mental engagement". In: *ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection. 2015.

[29] Chris Berka et al. "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks". In: *Aviation, space, and environmental medicine* 78.5 (2007), B231–B244.

[30] Andrea Apicella et al. "High-wearable EEG-Based transducer for Engagement Detection in Pediatric Rehabilitation". In: *Brain Computer Intereface* 11 (2021), p. 1498.

[31] Wanjoo Park et al. "Assessment of cognitive engagement in stroke patients from single-trial EEG during motor rehabilitation". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 23.3 (2014), pp. 351–362.

[32] Antonio M Battro and Kurt W Fischer. "Mind, brain, and education in the digital era". In: *Mind, Brain, and Education* 6.1 (2012), pp. 49–50.

[33] Harisa Mardiana and Haris Kaisar Daniels. "Technological Determinism, New Literacies and Learning Process and the Impact towards Future Learning." In: *Online Submission* 5.3 (2019), pp. 219–229.

[34] Robb Lindgren and Mina Johnson-Glenberg. "Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality". In: *Educational researcher* 42.8 (2013), pp. 445–452.

[35] Lana Plumanns Anjarichert et al. "Learning 4.0: Virtual immersive engineering education". In: *Digit. Univ* 2 (2016), p. 51.

[36] Daniela Janssen et al. "Virtual environments in higher education-Immersion as a key construct for learning 4.0." In: *iJAC* 9.2 (2016), pp. 20–26.

[37] Richard Gross. *Psychology: The science of mind and behaviour 7th edition*. Hodder Education, 2015.

[38] Asma Ben Khedher, Imène Jraidi, and Claude Frasson. "Tracking students' mental engagement using EEG signals during an interaction with a virtual learning environment". In: *Journal of Intelligent Learning Systems and Applications* 11.1 (2019), pp. 1–14.

[39] Lena M Andreessen et al. "Toward neuroadaptive support technologies for improving digital reading: A passive BCI-based assessment of mental workload imposed by text difficulty and presentation speed during reading". In: *User Modeling and User-Adapted Interaction* 31.1 (2021), pp. 75–104.

[40] Andrea Apicella et al. "EEG-based Measurement System for Student Engagement Detection in Learning 4.0". In: (2021).

[41] Michael Poole and Malcolm Warner. *The IEBM handbook of human resource management*. International Thomson Business, 1998.

[42] JE Fischer et al. "Objectifying psychomental stress in the workplace–an example". In: *International archives of occupational and environmental health* 73.1 (2000), S46–S52.

[43] Bruno Siciliano and Oussama Khatib. *Springer handbook of robotics*. Springer, 2016.

[44] Marlena R. Fraune et al. "Is Human-Robot Interaction More Competitive Between Groups Than Between Individuals?" In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 2019, pp. 104–113. DOI: 10.1109/HRI.2019.8673241.

[45] Nandita Sharma and Tom Gedeon. "Objective measures, sensors and computational techniques for stress recognition and classification: A survey". In: *Computer methods and programs in biomedicine* 108.3 (2012), pp. 1287–1301.

[46] Minho Choi et al. "Wearable device-based system to monitor a driver's stress, fatigue, and drowsiness". In: *IEEE Transactions on Instrumentation and Measurement* 67.3 (2017), pp. 634–645.

[47] Jesus Minguillon et al. "Portable System for Real-Time Detection of Stress Level". In: *Sensors* 18.8 (2018), p. 2504.

[48] Houtan Jebelli, Sungjoo Hwang, and SangHyun Lee. "EEG-based workers' stress recognition at construction sites". In: *Automation in Construction* 93 (2018), pp. 315–324.

[49] Cornelia Setz et al. "Discriminating stress from cognitive load using a wearable EDA device". In: *IEEE Transactions on information technology in biomedicine* 14.2 (2009), pp. 410–417.

[50] Joong Woo Ahn, Yunseo Ku, and Hee Chan Kim. "A Novel Wearable EEG and ECG Recording System for Stress Assessment". In: *Sensors* 19.9 (2019), p. 1991.

[51] Leopoldo Angrisani et al. "A Single-Channel SSVEP-Based Instrument With Off-the-Shelf Components for Trainingless Brain-Computer Interfaces". In: *IEEE Transactions on Instrumentation and Measurement* (2018).

[52] Paul Ekman. "Basic emotions". In: *Handbook of cognition and emotion* 98.45-60 (1999), p. 16.

[53] Albert Mehrabian and James A Russell. "The basic emotional impact of environments". In: *Perceptual and motor skills* 38.1 (1974), pp. 283–301.

[54] Nico H Frijda. "Emotion, cognitive structure, and action tendency". In: *Cognition and emotion* 1.2 (1987), pp. 115–143.

[55] Joan C Borod et al. "Right hemisphere emotional perception: evidence across multiple channels." In: *Neuropsychology* 12.3 (1998), p. 446.

[56] Geoffrey L Ahern and Gary E Schwartz. "Differential lateralization for positive versus negative emotion". In: *Neuropsychologia* 17.6 (1979), pp. 693–698.

[57] Richard J Davidson et al. "Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology: I." In: *Journal of personality and social psychology* 58.2 (1990), p. 330.

[58] Steven K Sutton and Richard J Davidson. "Prefrontal brain asymmetry: A biological substrate of the behavioral approach and inhibition systems". In: *Psychological science* 8.3 (1997), pp. 204–210.

[59] Susan Aliakbaryhosseinabadi et al. "Classification of EEG signals to identify variations in attention during motor task execution". In: *Journal of neuroscience methods* 284 (2017), pp. 27–34.

[60] Pablo F Diez et al. "Attention-level transitory response: a novel hybrid BCI approach". In: *Journal of neural engineering* 12.5 (2015), p. 056007.

[61] Mairon Noam et al. "Behavioral and EEG Measures Show no Amplifying Effects of Shared Attention on Attention or Memory". In: *Scientific Reports (Nature Publisher Group)* 10.1 (2020).

[62] NJ Hill and Bernhard Schölkopf. "An online brain–computer interface based on shifting attention to concurrent streams of auditory stimuli". In: *Journal of neural engineering* 9.2 (2012), p. 026011.

[63] Laurel J Buxbaum et al. "Hemispatial neglect: Subtypes, neuroanatomy, and disability". In: *Neurology* 62.5 (2004), pp. 749–756.

[64] M. Sohlberg. "Theory and remediation of attention disorders". In: *Introduction to Cognitive Rehabilitation Theory & Practice* (1989), pp. 110–135.

[65] Susan Aliakbaryhosseinabadi et al. "Classification of Movement Preparation Between Attended and Distracted Self-Paced Motor Tasks". In: *IEEE Transactions on Biomedical Engineering* 66.11 (2019), pp. 3060–3071.

[66] Natalie Mrachacz-Kersting et al. "An associative brain-computer-interface for acute stroke patients". In: *Converging Clinical and Engineering Research on Neurorehabilitation II*. Springer, 2017, pp. 841–845.

[67] Jiaojiao Yang et al. "Classification of children with attention deficit hyperactivity disorder using PCA and K-nearest neighbors during interference control task". In: *Advances in Cognitive Neurodynamics (V)*. Springer, 2016, pp. 447–453.

[68] Yoritaka Akimoto et al. "High-gamma activity in an attention network predicts individual differences in elderly adults' behavioral performance". In: *Neuroimage* 100 (2014), pp. 290–300.

[69] Leandro da Silva-Sauer et al. "Concentration on performance with P300-based BCI systems: A matter of interface features". In: *Applied ergonomics* 52 (2016), pp. 325–332.

[70] William A Kahn. "Psychological conditions of personal engagement and disengagement at work". In: *Academy of management journal* 33.4 (1990), pp. 692–724.

[71] Marco Klopp and Jörg Abke. "'Learning 4.0': A Conceptual Discussion". In: *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*. IEEE. 2018, pp. 871–876.

[72] Daniel B Willingham. "A neuropsychological theory of motor skill learning." In: *Psychological review* 105.3 (1998), p. 558.

[73] Uta Sailer, J Randall Flanagan, and Roland S Johansson. “Eye–hand coordination during learning of a novel visuomotor task”. In: *Journal of Neuroscience* 25.39 (2005), pp. 8833–8842.

[74] So-Young Park. “Student engagement and classroom variables in improving mathematics achievement”. In: *Asia Pacific Education Review* 6.1 (2005), pp. 87–97.

[75] Susie Lamborn, Fred Newmann, and Gary Wehlage. “The significance and sources of student engagement”. In: *Student engagement and achievement in American secondary schools* (1992), pp. 11–39.

[76] Antoine Lutz et al. “Attention regulation and monitoring in meditation”. In: *Trends in cognitive sciences* 12.4 (2008), pp. 163–169.

[77] James P Connell and James G Wellborn. “Competence, autonomy, and relatedness: A motivational analysis of self-system processes.” In: *Cultural processes in child development: The Minnesota symposia on child psychology*. Vol. 23. Psychology Press. 1991, pp. 43–78.

[78] Jennifer A Fredricks, Phyllis C Blumenfeld, and Alison H Paris. “School engagement: Potential of the concept, state of the evidence”. In: *Review of educational research* 74.1 (2004), pp. 59–109.

[79] Saul McLeod. “Jean Piaget’s theory of cognitive development”. In: *Simply Psychology* (2018), pp. 1–9.

[80] Eric L Garland and Matthew Owen Howard. “Neuroplasticity, psychosocial genomics, and the biopsychosocial paradigm in the 21st century”. In: *Health & social work* 34.3 (2009), pp. 191–199.

[81] Jeffrey A Kleim and Theresa A Jones. “Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage”. In: (2008).

[82] Larry D Rosen, Nancy Cheever, and L Mark Carrier. *The Wiley handbook of psychology, technology, and society*. John Wiley & Sons, 2015.

[83] Mark Carrier. *From Smartphones to Social Media: How Technology Affects Our Brains and Behavior*. ABC-CLIO, 2018.

[84] Norman Doidge. *The brain that changes itself: Stories of personal triumph from the frontiers of brain science*. Penguin, 2007.

[85] Richard VanDeWeghe. *Engaged learning*. Corwin Press, 2009.

[86] Hyungshim Jang, Johnmarshall Reeve, and Edward L Deci. “Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure.” In: *Journal of educational psychology* 102.3 (2010), p. 588.

[87] Oqab Alrashidi, Huy P Phan, and Bing H Ngu. “Academic Engagement: An Overview of Its Definitions, Dimensions, and Major Conceptualisations.” In: *International Education Studies* 9.12 (2016), pp. 41–52.

[88] Elise Cappella et al. "Classroom peer relationships and behavioral engagement in elementary school: The role of social network equity". In: *American journal of community psychology* 52.3-4 (2013), pp. 367–379.

[89] Maura Pilotti et al. "Factors Related to Cognitive, Emotional, and Behavioral Engagement in the Online Asynchronous Classroom." In: *International Journal of Teaching and Learning in Higher Education* 29.1 (2017), pp. 145–153.

[90] Arlindo Silva and Ricardo Simoes. *Handbook of Research on Trends in Product Design and Development: Technological and Organizational Perspectives: Technological and Organizational Perspectives*. IGI Global, 2010.

[91] Jerome I Rotgans and Henk G Schmidt. "Cognitive engagement in the problem-based learning classroom". In: *Advances in health sciences education* 16.4 (2011), pp. 465–479.

[92] Serena Barello et al. "eHealth for patient engagement: a systematic review". In: *Frontiers in psychology* 6 (2016), p. 2013.

[93] Anthony H Lequerica and Kathleen Kortte. "Therapeutic engagement: a proposed model of engagement in medical rehabilitation". In: *American journal of physical medicine & rehabilitation* 89.5 (2010), pp. 415–422.

[94] Edward L Deci and Richard M Ryan. "Conceptualizations of intrinsic motivation and self-determination". In: *Intrinsic motivation and self-determination in human behavior*. Springer, 1985, pp. 11–40.

[95] Pauline Gibbons. *Scaffolding language, scaffolding learning*. Portsmouth, NH: Heinemann, 2002.

[96] Gillian King et al. "The nature, value, and experience of engagement in pediatric rehabilitation: perspectives of youth, caregivers, and service providers". In: *Developmental neurorehabilitation* 23.1 (2020), pp. 18–30.

[97] Giovanni B Rossi and Birgitta Berglund. "Measurement involving human perception and interpretation". In: *Measurement* 44.5 (2011), pp. 815–822.

[98] Stanley Smith Stevens. "The direct estimation of sensory magnitudes: Loudness". In: *The American journal of psychology* 69.1 (1956), pp. 1–25.

[99] Paul De Bièvre. "The 2012 International Vocabulary of Metrology:"VIM"". In: *Accreditation and Quality Assurance* 17.2 (2012), pp. 231–232.

[100] Paul Ed Ekman and Richard J Davidson. *The nature of emotion: Fundamental questions.* Oxford University Press, 1994.

[101] Peter J Lang. "International affective picture system (IAPS): Affective ratings of pictures and instruction manual". In: *Technical report* (2005).

[102] Robert Jenke, Angelika Peer, and Martin Buss. "Feature extraction and selection for emotion recognition from EEG". In: *IEEE Transactions on Affective computing* 5.3 (2014), pp. 327–339.

[103] Heath A Demaree et al. "Brain lateralization of emotional processing: historical roots and a future incorporating "dominance"". In: *Behavioral and cognitive neuroscience reviews* 4.1 (2005), pp. 3–20.

[104] James A Coan and John JB Allen. "The state and trait nature of frontal EEG asymmetry in emotion." In: (2003).

[105] Richard J Davidson. "Hemispheric asymmetry and emotion". In: *Approaches to emotion* 2 (1984), pp. 39–57.

[106] Antoine Bechara et al. "Insensitivity to future consequences following damage to human prefrontal cortex". In: *Cognition* 50 (1994), pp. 1–3.

[107] James A Coan and John JB Allen. "Frontal EEG asymmetry and the behavioral activation and inhibition systems". In: *Psychophysiology* 40.1 (2003), pp. 106–114.

[108] Dirk Hagemann et al. "Frontal brain asymmetry and affective style: A conceptual replication". In: *Psychophysiology* 35.4 (1998), pp. 372–388.

[109] James A Coan and John JB Allen. "Frontal EEG asymmetry as a moderator and mediator of emotion". In: *Biological psychology* 67.1-2 (2004), pp. 7–50.

[110] Jonathan Wolpaw and Elizabeth Winter Wolpaw. *Brain-computer interfaces: principles and practice*. OUP USA, 2012.

[111] Hong Zeng et al. "EEG emotion classification using an improved SincNet-based deep learning model". In: *Brain sciences* 9.11 (2019), p. 326.

[112] Yuling Luo et al. "EEG-based Emotion Classification Using Deep Neural Network and Sparse Autoencoder". In: *Frontiers in Systems Neuroscience* 14 (2020), p. 43.

[113] Yuling Luo et al. "EEG-Based Emotion Classification Using Spiking Neural Networks". In: *IEEE Access* 8 (2020), pp. 46007–46016.

[114] Fei Wang et al. "Emotion recognition with convolutional neural network and EEG-based EFDMs". In: *Neuropsychologia* (2020), p. 107506.

[115] Yucel Cimtay and Erhan Ekmekcioglu. "Investigating the use of pretrained convolutional neural network on cross-subject and cross-dataset EEG emotion recognition". In: *Sensors* 20.7 (2020), p. 2034.

[116] Tengfei Song et al. "EEG emotion recognition using dynamical graph convolutional neural networks". In: *IEEE Transactions on Affective Computing* (2018).

[117] JX Chen, DM Jiang, and YN Zhang. "A hierarchical bidirectional GRU model with attention for EEG-based emotion classification". In: *IEEE Access* 7 (2019), pp. 118530–118540.

[118] Morteza Zangeneh Soroush et al. "Emotion classification through nonlinear EEG analysis using machine learning methods". In: *International Clinical Neuroscience Journal* 5.4 (2018), p. 135.

[119] Habib Ullah et al. "Internal emotion classification using EEG signal with sparse discriminative ensemble". In: *IEEE Access* 7 (2019), pp. 40144–40153.

[120] Debashis Das Chakladar and Sanjay Chakraborty. "EEG based emotion classification using "Correlation Based Subset Selection"". In: *Biologically inspired cognitive architectures* 24 (2018), pp. 98–106.

[121] Hector A Gonzalez et al. "BioCNN: A hardware inference engine for EEG-based emotion detection". In: *IEEE Access* 8 (2020), pp. 140896–140914.

[122] Xiaofen Xing et al. "SAE+ LSTM: A New framework for emotion recognition from multi-channel EEG". In: *Frontiers in Neurorobotics* 13 (2019), p. 37.

[123] Fu Yang et al. "Cross-subject emotion recognition using multi-method fusion from high-dimensional features". In: *Frontiers in Computational Neuroscience* 13 (2019), p. 53.

[124] Yu Liu et al. "Multi-channel EEG-based emotion recognition via a multi-level features guided capsule network". In: *Computers in Biology and Medicine* 123 (2020), p. 103927.

[125] Heng Cui et al. "EEG-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network". In: *Knowledge-Based Systems* 205 (2020), p. 106243.

[126] D Jude Hemanth. "EEG signal based Modified Kohonen Neural Networks for Classification of Human Mental Emotions". In: *Journal of Artificial Intelligence and Systems* 2 (2020), pp. 1–13.

[127] Kairui Guo et al. "A hybrid fuzzy cognitive map/support vector machine approach for EEG-based emotion classification using compressed sensing". In: *International Journal of Fuzzy Systems* 21.1 (2019), pp. 263–273.

[128] Wei-Long Zheng and Bao-Liang Lu. "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks". In: *IEEE Transactions on Autonomous Mental Development* 7.3 (2015), pp. 162–175.

[129] *SEED-dataset*. https://bcmi.sjtu.edu.cn/home/seed/seed.html. 2021.

[130] Sander Koelstra et al. "Deap: A database for emotion analysis; using physiological signals". In: *IEEE transactions on affective computing* 3.1 (2011), pp. 18–31.

[131] *DEAP-dataset*. http://www.eecs.qmul.ac.uk/mmv/datasets/deap/readme.html. 2021.

[132] Stamos Katsigiannis and Naeem Ramzan. "DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices". In: *IEEE journal of biomedical and health informatics* 22.1 (2017), pp. 98–107.

[133] *DREAMER-dataset*. https://zenodo.org/record/546113#.YMNN6NUzbIU. 2021.

[134] Ruo-Nan Duan, Jia-Yi Zhu, and Bao-Liang Lu. "Differential entropy feature for EEG-based emotion classification". In: *6th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE. 2013, pp. 81–84.

[135] Qiang Gao et al. "EEG based emotion recognition using fusion feature extraction method". In: *Multimedia Tools and Applications* 79.37 (2020), pp. 27057–27074.

[136] Raja Majid Mehmood and Hyo Jong Lee. "EEG based emotion recognition from human brain using Hjorth parameters and SVM". In: *International Journal of Bio-Science and Bio-Technology* 7.3 (2015), pp. 23–32.

[137] Elise S Dan-Glauser and Klaus R Scherer. "The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance". In: *Behavior research methods* 43.2 (2011), p. 468.

[138] Sachin Taran and Varun Bajaj. "Emotion recognition from single-channel EEG signals using a two-stage correlation and instantaneous frequency-based filtering method". In: *Computer Methods and Programs in Biomedicine* 173 (2019), pp. 157–165.

[139] Mikito Ogino and Yasue Mitsukura. "A Mobile Application for Estimating Emotional Valence Using a Single-Channel EEG Device". In: *2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*. IEEE. 2018, pp. 1043–1048.

[140] Noppadon Jatupaiboon, Setha Pan-ngum, and Pasin Israsena. "Emotion classification using minimal EEG channels and frequency bands". In: *The 2013 10th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. IEEE. 2013, pp. 21–24.

[141] Amir Jalilifard, Ednaldo Brigante Pizzolato, and Md Kafiul Islam. "Emotion classification using single-channel scalp-EEG recording". In: *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2016, pp. 845–849.

[142] Panagiotis C Petrantonakis and Leontios J Hadjileontiadis. "Emotion recognition from EEG using higher order crossings". In: *IEEE Transactions on information Technology in Biomedicine* 14.2 (2009), pp. 186–197.

[143] Pallavi Pandey and KR Seeja. "Emotional state recognition with EEG signals using subject independent approach". In: *Data Science and Big Data Analytics*. Springer, 2019, pp. 117–124.

[144] Adrian Qi-Xiang Ang, Yi Qi Yeong, and Wee Wee. "Emotion classification from EEG signals using time-frequency-DWT features and ANN". In: *Journal of Computer and Communications* 5.3 (2017), pp. 75–79.

[145] Cheng-Jie Yang et al. "An AI-Edge Platform with Multimodal Wearable Physiological Signals Monitoring Sensors for Affective Computing Applications". In: *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE. 2020, pp. 1–5.

[146] Javier Marín-Morales et al. "Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors". In: *Scientific reports* 8.1 (2018), pp. 1–15.

[147] Yang Wei, Yue Wu, and John Tudor. "A real-time wearable emotion detection headband based on EEG measurement". In: *Sensors and Actuators A: Physical* 263 (2017), pp. 614–621.

[148] Pasquale Arpaia et al. “A wearable EEG instrument for real-time frontal asymmetry monitoring in worker stress analysis”. In: *IEEE Transactions on Instrumentation and Measurement* (2020).

[149] Murugappan Murugappan et al. “An Investigation on visual and audiovisual stimulus based emotion recognition using EEG”. In: *International Journal of Medical Engineering and Informatics* 1.3 (2009), pp. 342–356.

[150] Brahim Hamadicharef et al. “Learning EEG-based spectral-spatial patterns for attention level measurement”. In: *2009 IEEE International Symposium on Circuits and Systems*. IEEE. 2009, pp. 1465–1468.

[151] Javier M Antelis et al. “Detection of movements with attention or distraction to the motor task during robot-assisted passive movements of the upper limb”. In: *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2012, pp. 6410–6413.

[152] Alan T Pope, Edward H Bogart, and Debbie S Bartolome. “Biocybernetic system evaluates indices of operator engagement in automated task”. In: *Biological psychology* 40.1-2 (1995), pp. 187–195.

[153] Atef Eldenfria and Hosam Al-Samarraie. “Towards an online continuous adaptation mechanism (OCAM) for enhanced engagement: An EEG study”. In: *International Journal of Human–Computer Interaction* 35.20 (2019), pp. 1960–1974.

[154] Nataliya Kosmyna and Pattie Maes. “AttentivU: an EEG-based closed-loop biofeedback system for real-time monitoring and improvement of engagement for personalized learning”. In: *Sensors* 19.23 (2019), p. 5200.

[155] Allan Wigfield et al. “Role of reading engagement in mediating effects of reading comprehension instruction on reading outcomes”. In: *Psychology in the Schools* 45.5 (2008), pp. 432–445.

[156] Sue Helme and David Clarke. “Identifying cognitive engagement in the mathematics classroom”. In: *Mathematics Education Research Journal* 13.2 (2001), pp. 133–153.

[157] Pu-Shih Daniel Chen, Amber D Lambert, and Kevin R Guidry. “Engaging online learners: The impact of Web-based learning technology on college student engagement”. In: *Computers & Education* 54.4 (2010), pp. 1222–1232.

[158] Shuhaimi Jaafar, Nur Suriana Awaludin, and Nor Suhaily Bakar. In: *E-proceeding of the Conference on Management and Muamalah*. 2014, pp. 128–135.

[159] Neelesh Kumar and Konstantinos P Michmizos. “Machine Learning for Motor Learning: EEG-based Continuous Assessment of Cognitive Engagement for Adaptive Rehabilitation Robots”. In: *arXiv preprint arXiv:2002.07541* (2020).

[160] Asma Ben Khedher, Imène Jraidi, and Claude Frasson. "Exploring students' eye movements to assess learning performance in a serious game". In: *EdMedia+ Innovate Learning*. Association for the Advancement of Computing in Education (AACE). 2018, pp. 394–401.

[161] Leopoldo Angrisani et al. "Instrumentation for Motor Imagery-based Brain Computer Interfaces relying on dry electrodes: a functional analysis". In: *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE. 2020, pp. 1–6.

[162] Mohamed S Benlamine et al. "BARGAIN: behavioral affective rule-based games adaptation interface–towards emotionally intelligent games: application on a virtual reality environment for socio-moral development". In: *User Modeling and User-Adapted Interaction* (2021), pp. 1–35.

[163] Xiao-Wei Wang, Dan Nie, and Bao-Liang Lu. "Emotional state classification from EEG data using machine learning approach". In: *Neurocomputing* 129 (2014), pp. 94–106.

[164] Mohammad Soleymani et al. "Analysis of EEG signals and facial expressions for continuous emotion detection". In: *IEEE Transactions on Affective Computing* 7.1 (2015), pp. 17–28.

[165] Ning Zhuang et al. "Emotion recognition from EEG signals using multidimensional information in EMD domain". In: *BioMed research international* 2017 (2017).

[166] Imène Jraidi, Maher Chaouachi, and Claude Frasson. "A hierarchical probabilistic framework for recognizing learners' interaction experience trends and emotions". In: *Advances in Human-Computer Interaction* 2014 (2014).

[167] P Aricò et al. "A passive brain–computer interface application for the mental workload assessment on professional air traffic controllers during realistic air traffic control tasks". In: *Progress in brain research* 228 (2016), pp. 295–328.

[168] Shouyi Wang, Jacek Gwizdka, and W Art Chaovalitwongse. "Using wireless EEG signals to assess memory workload in the n-back task". In: *IEEE Transactions on Human-Machine Systems* 46.3 (2015), pp. 424–435.

[169] Fred Paas et al. "Cognitive load measurement as a means to advance cognitive load theory". In: *Educational psychologist* 38.1 (2003), pp. 63–71.

[170] Judith H Hibbard et al. "Development and testing of a short form of the patient activation measure". In: *Health services research* 40.6p1 (2005), pp. 1918–1930.

[171] Guendalina Graffigna et al. "Measuring patient engagement: development and psychometric properties of the Patient Health Engagement (PHE) Scale". In: *Frontiers in psychology* 6 (2015), p. 274.

[172] Gillian King et al. "Development of an observational measure of therapy engagement for pediatric rehabilitation". In: *Disability and Rehabilitation* 41.1 (2019), pp. 86–97.

[173] Elena Dell'Aquila et al. "A Preparatory Study for Measuring Engagement in Pediatric Virtual and Robotics Rehabilitation Settings". In: *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 2020, pp. 183–185.

[174] N. Kumar and K. P. Michmizos. "Machine Learning for Motor Learning: EEG-based Continuous Assessment of Cognitive Engagement for Adaptive Rehabilitation Robots". In: *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*. 2020, pp. 521–526. DOI: 10.1109/BioRob49111.2020.9224368.

[175] Xin Dang, Ran Wei, and Guohui Li. "An efficient movement and mental classification for children with autism based on motion and EEG features". In: *Journal of Ambient Intelligence and Humanized Computing* 8.6 (2017), pp. 907–912.

[176] Gerald Matthews et al. "Validation of a comprehensive stress state questionnaire: Towards a state big three". In: *Personality psychology in Europe* 7 (1999), pp. 335–350.

[177] Sarah L Mïller-Abdelrazeq et al. "Perceived Effects of Cycle Time in Human-Robot-Interaction". In: *2018 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*. IEEE. 2018, pp. 25–30.

[178] Jacqueline Wijsman et al. "Towards mental stress detection using wearable physiological sensors". In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2011, pp. 1798–1801.

[179] Anne-Marie Brouwer et al. "EEG alpha asymmetry, heart rate variability and cortisol in response to virtual reality induced stress". In: *Journal of Cybertherapy & Rehabilitation* 4.1 (2011), pp. 21–34.

[180] Suresh GR et al. "Development of four stress levels in group stroop colour word test using HRV analysis." In: *Biomedical Research (0970-938X)* 28.1 (2017).

[181] Abdullah Mohammed and Lihui Wang. "Brainwaves driven human-robot collaborative assembly". In: *CIRP annals* 67.1 (2018), pp. 13–16.

[182] Bruce Wallace et al. "EEG/ERP: within episodic assessment framework for cognition". In: *IEEE Transactions on Instrumentation and Measurement* 66.10 (2017), pp. 2525–2534.

[183] Aimé Lay-Ekuakille et al. "Entropy index in quantitative EEG measurement for diagnosis accuracy". In: *IEEE Transactions on Instrumentation and Measurement* 63.6 (2013), pp. 1440–1450.

[184] Stephan Mühlbacher-Karrer et al. "A driver state detection system—Combining a capacitive hand detection sensor with physiological sensors". In: *IEEE Transactions on Instrumentation and Measurement* 66.4 (2017), pp. 624–636.

[185] Zhongke Gao et al. "Relative wavelet entropy complex network for improving EEG-based fatigue driving classification". In: *IEEE Transactions on Instrumentation and Measurement* 99 (2018), pp. 1–7.

[186] Ilona Papousek et al. “Prefrontal EEG alpha asymmetry changes while observing disaster happening to other people: cardiac correlates and prediction of emotional impact”. In: *Biological psychology* 103 (2014), pp. 184–194.

[187] Ronald N Goodman et al. “Stress, emotion regulation and cognitive performance: The predictive contributions of trait and state relative frontal EEG alpha asymmetry”. In: *International Journal of Psychophysiology* 87.2 (2013), pp. 115–123.

[188] Jesus Minguillon, Miguel A Lopez-Gordo, and Francisco Pelayo. “Stress assessment by prefrontal relative gamma”. In: *Frontiers in computational neuroscience* 10 (2016), p. 101.

[189] Richard J Davidson. “Anterior electrophysiological asymmetries, emotion, and depression: Conceptual and methodological conundrums”. In: *Psychophysiology* 35.5 (1998), pp. 607–614.

[190] Dean J Krusienski et al. “Critical issues in state-of-the-art brain–computer interface signal processing”. In: *Journal of neural engineering* 8.2 (2011), p. 025002.

[191] Fabien Lotte et al. “A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update”. In: *Journal of neural engineering* 15.3 (2018), p. 031005.

[192] Xun Chen et al. “Independent vector analysis applied to remove muscle artifacts in EEG data”. In: *IEEE Transactions on Instrumentation and Measurement* 66.7 (2017), pp. 1770–1779.

[193] Xun Chen et al. “The use of multivariate EMD and CCA for denoising muscle artifacts from few-channel EEG recordings”. In: *IEEE transactions on instrumentation and measurement* 67.2 (2017), pp. 359–370.

[194] Matteo Zanetti et al. “Multilevel assessment of mental stress via network physiology paradigm using consumer wearable devices”. In: *Journal of Ambient Intelligence and Humanized Computing* (2019), pp. 1–10.

[195] Seyyed Abed Hosseini and Mohammad Ali Khalilzadeh. “Emotional stress recognition system using EEG and psychophysiological signals: Using new labelling process of EEG signals in emotional stress state”. In: *2010 international conference on biomedical engineering and computer science*. IEEE. 2010, pp. 1–6.

[196] Pengbo Zhang et al. “EEG feature selection based on weighted-normalized mutual information for mental fatigue classification”. In: *2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings*. IEEE. 2016, pp. 1–6.

[197] Xiyuan Hou et al. “EEG based stress monitoring”. In: *2015 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE. 2015, pp. 3110–3115.

[198] Guo Jun and Kavallur Gopi Smitha. “EEG based stress level identification”. In: *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE. 2016, pp. 003270–003274.

[199] A Secerbegovic et al. "Mental workload vs. stress differentiation using single-channel EEG". In: *CMBEBIH 2017*. Springer, 2017, pp. 511–515.

[200] Dayi Bian et al. "Physiology-based Affect Recognition During Driving in Virtual Environment for Autism Intervention." In: *PhyCS*. 2015, pp. 137–145.

[201] Daniela Sammler et al. "Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music". In: *Psychophysiology* 44.2 (2007), pp. 293–304.

[202] Lindsay Brown, Bernard Grundlehner, and Julien Penders. "Towards wireless emotional valence detection from EEG". In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2011, pp. 2188–2191.

[203] Shiu Kumar et al. "A deep learning approach for motor imagery EEG signal classification". In: *2016 3rd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*. IEEE. 2016, pp. 34–39.

[204] Zheng Yang Chin et al. "Multi-class filter bank common spatial pattern for four-class motor imagery BCI". In: *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2009, pp. 571–574.

[205] Kavitha P Thomas et al. "A new discriminative common spatial pattern method for motor imagery brain–computer interfaces". In: *IEEE Transactions on Biomedical Engineering* 56.11 (2009), pp. 2730–2733.

[206] Maouia Bentlemsan et al. "Random forest and filter bank common spatial patterns for EEG-based motor imagery classification". In: *2014 5th International conference on intelligent systems, modelling and simulation*. IEEE. 2014, pp. 235–238.

[207] Merve Dogruyol Basar, Adil Deniz Duru, and Aydin Akan. "Emotional state detection based on common spatial patterns of EEG". In: *Signal, Image and Video Processing* (2019), pp. 1–9.

[208] Mengmeng Yan et al. "An improved common spatial pattern combined with channel-selection strategy for electroencephalography-based emotion recognition". In: *Medical Engineering & Physics* (2020).

[209] Shiu Kumar et al. "Decimation filter with common spatial pattern and fishers Discriminant analysis for motor imagery classification". In: *2016 international joint conference on neural networks (IJCNN)*. IEEE. 2016, pp. 2090–2095.

[210] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.

[211] Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006.

[212] G. Cybenko. "Approximation by superpositions of a sigmoidal function". In: *Mathematics of Control, Signals, and Systems (MCSS)* 2.4 (Dec. 1989), pp. 303–314.

[213] Kurt Kroenke and Robert L Spitzer. "The PHQ-9: a new depression diagnostic and severity measure". In: *Psychiatric annals* 32.9 (2002), pp. 509–515.

[214] Laura Martin Braunstein, James J Gross, and Kevin N Ochsner. "Explicit and implicit emotion regulation: a multi-level framework". In: *Social cognitive and affective neuroscience* 12.10 (2017), pp. 1545–1557.

[215] Benedek Kurdi, Shayn Lozano, and Mahzarin R Banaji. "Introducing the open affective standardized image set (OASIS)". In: *Behavior research methods* 49.2 (2017), pp. 457–470.

[216] *AB-Medica S.P.A.* https://www.abmedica.it/. 2020.

[217] *Texasinstrument-ADS1298*. https://www.ti.com/lit/ds/symlink/ads1296r.pdf. 2020-02-28.

[218] Thea Radüntz et al. "EEG artifact elimination by extraction of ICA-component features using image processing algorithms". In: *Journal of neuroscience methods* 243 (2015), pp. 84–93.

[219] Fabien Lotte et al. "A review of classification algorithms for EEG-based brain–computer interfaces". In: *Journal of neural engineering* 4.2 (2007), R1.

[220] Corinna Cortes and Vladimir Vapnik. "Support-vector networks". In: *Machine learning* 20.3 (1995), pp. 273–297.

[221] Anders Krogh and John Hertz. "A Simple Weight Decay Can Improve Generalization". In: *Advances in Neural Information Processing Systems*. Ed. by J. Moody, S. Hanson, and R. P. Lippmann. Vol. 4. Morgan-Kaufmann, 1992. URL: https://proceedings.neurips.cc/paper/1991/file/8eefcfdf5990e441f0fb6f3fad709e21-Paper.pdf.

[222] Stefania Coelli et al. "EEG indices correlate with sustained attention performance in patients affected by diffuse axonal injury". In: *Medical & biological engineering & computing* 56.6 (2018), pp. 991–1001.

[223] Emily Graber and Takako Fujioka. "Induced Beta Power Modulations during Isochronous Auditory Beats Reflect Intentional Anticipation before Gradual Tempo Changes". In: *Scientific Reports* 10.1 (2020), pp. 1–12.

[224] Choon Guan Lim et al. "A brain-computer interface based attention training program for treating attention deficit hyperactivity disorder". In: *PloS one* 7.10 (2012).

[225] Javier Asensio-Cubero, John Q Gan, and Ramaswamy Palaniappan. "Multiresolution analysis over graphs for a motor imagery based online BCI game". In: *Computers in biology and medicine* 68 (2016), pp. 21–26.

[226] Bin Hu et al. "Attention recognition in EEG-based affective learning research using CFS+ KNN algorithm". In: *IEEE/ACM transactions on computational biology and bioinformatics* 15.1 (2016), pp. 38–45.

[227] Joana Pereira, Andreea Ioana Sburlea, and Gernot R Müller-Putz. "EEG patterns of self-paced movement imaginations towards externally-cued and internally-selected targets". In: *Scientific reports* 8.1 (2018), pp. 1–15.

[228] John Z Wu et al. "An evaluation of the contact forces on the fingers when squeezing a spherical rehabilitation ball". In: *Bio-medical materials and engineering* 29.5 (2018), pp. 629–639.

[229] Peng Ye et al. "Comparison of DP3 signals evoked by comfortable 3D images and 2D images—an event-related potential study using an oddball task". In: *Scientific reports* 7 (2017), p. 43110.

[230] John Polich and Catherine Margala. "P300 and probability: comparison of oddball and single-stimulus paradigms". In: *International Journal of Psychophysiology* 25.2 (1997), pp. 169–176.

[231] Scott A Huettel and Gregory McCarthy. "What is odd in the oddball task?: Prefrontal cortex is activated by dynamic changes in response strategy". In: *Neuropsychologia* 42.3 (2004), pp. 379–386.

[232] Hermann Hinrichs et al. "Comparison between a wireless dry electrode EEG system with a conventional wired wet electrode EEG system for clinical applications". In: *Scientific Reports* 10.1 (2020), pp. 1–14.

[233] Gaetano Gargiulo et al. "A mobile EEG system with dry electrodes". In: *2008 IEEE Biomedical Circuits and Systems Conference*. IEEE. 2008, pp. 273–276.

[234] Arnaud Delorme, Terrence Sejnowski, and Scott Makeig. "Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis". In: *Neuroimage* 34.4 (2007), pp. 1443–1449.

[235] Keinosuke Fukunaga. *Introduction to statistical pattern recognition*. Elsevier, 2013.

[236] Pedro Domingos and Michael Pazzani. "On the optimality of the simple Bayesian classifier under zero-one loss". In: *Machine learning* 29.2-3 (1997), pp. 103–130.

[237] Sudhir Varma and Richard Simon. "Bias in error estimation when using cross-validation for model selection". In: *BMC bioinformatics* 7.1 (2006), p. 91.

[238] Ian Jolliffe. *Principal component analysis*. Springer, 2011.

[239] En-Chi Chang, Shian-Chang Huang, and Hsin-Hung Wu. "Using K-means method and spectral clustering technique in an outfitter's value analysis". In: *Quality & Quantity* 44.4 (2010), pp. 807–815.

[240] Christopher M Bishop et al. *Neural networks for pattern recognition*. Oxford university press, 1995.

[241] Patrice Renaud and Jean-Pierre Blondin. "The stress of Stroop performance: Physiological and emotional responses to color–word interference, task pacing, and pacing speed". In: *International Journal of Psychophysiology* 27.2 (1997), pp. 87–97.

[242] Charles D Spielberger. "Manual for the State-Trait Anxiety Inventory STAI (form Y)(" self-evaluation questionnaire")". In: (1983).

[243] Jim Blascovich and Joe Tomaka. "Measures of self-esteem". In: *Measures of personality and social psychological attitudes* 1 (1991), pp. 115–160.

[244] Alain Celisse and Stéphane Robin. "Nonparametric density estimation by exact leave-p-out cross-validation". In: *Computational Statistics & Data Analysis* 52.5 (2008), pp. 2350–2368.

[245] Bo-Qun Ma et al. "Reducing the subject variability of eeg signals with adversarial domain generalization". In: *International Conference on Neural Information Processing*. Springer. 2019, pp. 30–42.

[246] Georgios Douzas, Fernando Bacao, and Felix Last. "Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE". In: *Information Sciences* 465 (2018), 1–20. ISSN: 0020-0255. DOI: 10.1016/j.ins.2018.06.056.

[247] Zoltan J Koles, Michael S Lazar, and Steven Z Zhou. "Spatial patterns underlying population differences in the background EEG". In: *Brain topography* 2.4 (1990), pp. 275–284.

[248] *Society for Research in Child Development Ethical Principles and Standards for Developmental Scientists*. 2021. URL: https://www.srcd.org/sites/default/files/file-attachments/SRCDethicalprinciplesstandardsfordevsci_3.2021.pdf (visited on 03/31/2021).

[249] Ratanapat Chanubol et al. "A randomized controlled trial of Cognitive Sensory Motor Training Therapy on the recovery of arm function in acute stroke patients". In: *Clinical Rehabilitation* 26.12 (2012). PMID: 22649162, pp. 1096–1104. DOI: 10.1177/0269215512444631.

[250] Sofiane Boucenna et al. "Interactive technologies for autistic children: A review". In: *Cognitive Computation* 6.4 (2014), pp. 722–740.

[251] Steve W Kozlowski and Keith Hattrup. "A disagreement about within-group agreement: Disentangling issues of consistency versus consensus." In: *Journal of applied psychology* 77.2 (1992), p. 161.

[252] Laurens Van der Maaten and Geoffrey Hinton. "Visualizing data using t-SNE." In: *Journal of machine learning research* 9.11 (2008).

[253] Michael S Myslobodsky and Jacob Bar-Ziv. "Locations of occipital EEG electrodes verified by computed tomography". In: *Electroencephalography and clinical neurophysiology* 72.4 (1989), pp. 362–366.

[254] J-M Hopf and George R Mangun. "Shifting visual attention in space: an electrophysiological analysis using high spatial resolution mapping". In: *Clinical neurophysiology* 111.7 (2000), pp. 1241–1257.

[255] Dixiu Liu et al. "The time course of spatial attention shifts in elementary arithmetic". In: *Scientific reports* 7.1 (2017), pp. 1–8.

[256] AR Mandel and SM Keller. "Stress management in rehabilitation." In: *Archives of physical medicine and rehabilitation* 67.6 (1986), pp. 375–379.

[257] Yisi Liu et al. "EEG-based mental workload and stress recognition of crew members in maritime virtual simulator: a case study". In: *2017 International Conference on Cyberworlds (CW)*. IEEE. 2017, pp. 64–71.

[258] Sinno Jialin Pan et al. "Domain adaptation via transfer component analysis". In: *IEEE Transactions on Neural Networks* 22.2 (2010), pp. 199–210.

[259] Wei-Long Zheng et al. "Transfer components between subjects for EEG-based emotion recognition". In: *2015 international conference on affective computing and intelligent interaction (ACII)*. IEEE. 2015, pp. 917–922.

[260] Sinno Jialin Pan and Qiang Yang. "A survey on transfer learning". In: *IEEE Transactions on knowledge and data engineering* 22.10 (2009), pp. 1345–1359.

[261] William S Noble. "What is a support vector machine?" In: *Nature biotechnology* 24.12 (2006), pp. 1565–1567.

[262] Antoine Gaume, Gérard Dreyfus, and François-Benoît Vialatte. "A cognitive brain–computer interface monitoring sustained attentional variations during a continuous task". In: *Cognitive neurodynamics* 13.3 (2019), pp. 257–269.

[263] Renato Panda, Ricardo Malheiro, and Rui Pedro Paiva. "Novel audio features for music emotion recognition". In: *IEEE Transactions on Affective Computing* 11.4 (2018), pp. 614–626.

[264] Pavlo Antonenko et al. "Using electroencephalography to measure cognitive load". In: *Educational Psychology Review* 22.4 (2010), pp. 425–438.

[265] Laurens Van der Maaten and Geoffrey Hinton. "Visualizing data using t-SNE." In: *Journal of machine learning research* 9.11 (2008).

List of Figures

3.1 The proposed valence-detection method (CSP: Common Spatial Pattern algorithm). 34
3.2 Experimental protocol. 37
3.3 Oasis valence score and SAM average scores of the 26 images selected for the experiments. The Oasis score intervals used to extract polarized images are identified by dotted lines. 37
3.4 Bland-Altman analysis on the agreement between stimuli (OASIS) and volunteers perception (SAM) 38
3.5 (A) EEG data acquisition system *Helmate8* and (B) Dry electrodes from *abmedica* 39
3.6 t-SNE based data comparison of four random subjects projected in the CSP space, without (first row) and with (second row) the Filter Bank. Filter Bank improves the classes (blue and red) separability. . 43
3.7 F1-score (White), Recall (Grey) and Precision (Black) for the best performance of each classifier - Cross-subject 45
3.8 The proposed distraction-detection method (CSP: Common Spatial Pattern algorithm). 48
3.9 Visual Distractor task elements based on visual Gabor mask with different orientation: 90°, 60°, and 30°. 51
3.10 (A) EEG data acquisition system *Helmate8*, and (B) Different configuration of dry electrodes from *abmedica*. [216]. 51
3.11 F-Measure test results for the best performance of each classifier: Precision (black) , Recall (gray), and F1-score (white). 55

4.1 Architecture of the real-time stress monitoring instrument in Cobot interaction. 58
4.2 K-means classification (white: class 1; black: class 2) among the 2 different time phases according to subjects belonging group. 59
4.3 Cumulative Explained Variance in the PCA. 63
4.4 SVM data distribution in PCs space, p=2, 92,6% of explained variance. 65
4.5 The proposed cognitive and emotional engagement detection method. 68
4.6 Neuromotor rehabilitation session. 69
4.7 t-SNE projection of unbalanced EEG data (subject 4) oversampled with two different methods. The SMOTE method (A) randomically interpolates the data of the minority class. The KMeansSMOTE method (B) realizes a clustering before interpolation, attenuating the noise. 72

4.8 Cognitive engagement balanced accuracies for each subject based on KMeansSMOTE oversampling technique. Classifier performances are reported. 73
4.9 Emotional engagement balanced accuracies for each subject based on KMeansSMOTE oversampling technique. Classifier performances are reported. 73
4.10 The architecture of the system for engagement assessment; the white box is active only in the cross-subject case (ADC - Analog Digital Converter, CSP - Common Spatial Pattern, TCA - Transfer Component Analysis, and SVM - Support Vector Machine). 75
4.11 Screen shots from the CPT game. At the beginning of the game (a),the cross starts to run away from the center of the black circumference. Theuser goal is to bring the cross back to the center by using the mouse. At theend of each trial (b), the score indicates the percentage time spent by thecross inside the circumference. 78
4.12 Within-subject performances of the compared processing techniques in (a) cognitive engagement and (b) emotional engagement detection. Each bar describes the average accuracy over all the subjects. 79
4.13 Filter Bank impact on the class (red and blue points) separability. t-SNE-based features plot of five subjects randomly sampled (first row: without Filter Bank; second row: with Filter Bank). 80
4.14 A comparison using t-SNE of the FBCSP data first (a) and after (b) removing the average value of each subject, in the cross-subject approach. 81

List of Tables

2.1 Studies on emotion recognition classified according to the employed datasets (i.e. SEED, DEAP, and DREAMER), stimuli (*v*="video", *p*="picture", *m*="memories"), task (*i*="implicit", *e*="explicit", n.a.="not available"), #channels, #participants, #classes, classifiers, and accuracies (n.a.="not available"). 23
2.2 State of art of stress classification 29

3.1 Classifier optimized hyperparameters and variation range 41
3.2 Accuracy (mean and standard deviation) considering a priori knowledge i.e. Asymmetry - Within-subject (Within) & Cross-subject (Cross) 42
3.3 Accuracy (mean and standard deviation) without considering a priori knowledge i.e. Asymmetry - Within-subject 42
3.4 Accuracy (mean and standard deviation) without considering a priori knowledge i.e. Asymmetry - Cross-subject 43
3.5 Accuracies obtained for each subject in the within-subject experiments when a FC-CSP Pipeline is adopted 44
3.6 Accuracy performances of the best processing solutions for both within- and cross-subject approaches at varying the number of input features selected through the Mutual Information strategy. . . . 44
3.7 Studies on emotion recognition classified according to metrological approach, number of channels and accuracy (n.a. = "not available", ✔ = "the property is verified". Only for the first line, ✔ = "Measurement") 46
3.8 Data-set composition 52
3.9 Classifier optimized Hyperparameters and variation range 54
3.10 Within-subject accuracy (mean and standard deviation percentage of the 17 subject accuracy) at varying feature and classifier 54
3.11 Within-subject accuracy of the proposed solution based on the 12 bandpass Filter Bank and Common Spatial Pattern at varying the classifier. 55

4.1 Stress index distribution (descending sort) 62
4.2 Classifier optimized iperparameters and range of variation 63
4.3 Classifiers accuracy (mean and uncertainty percentage) in Original Data (O.D.) and Principal Components Hyperplanes 64
4.4 F-measure test results for SVM (mean and uncertainty percentage) 64
4.5 Accuracy (mean and uncertainty) in Original Data (O.D.) and Principal Components Hyperplanes at varying amplitude of random gaussian noise 66

4.6 Overall mean of the intra-individual performances on cognitive engagement using three different classifiers: the balanced accuracy (BA) and the Matthews correlation coefficient (MCC) at varying the oversampling methods. 70
4.7 Overall mean of the intra-individual performances on emotional engagement using three different classifiers: the balanced accuracy (BA) and the Matthews correlation coefficient (MCC) at varying the oversampling methods. 71
4.8 Within-subject experimental results. Classification accuracies using the *Engagement Index* [152] for cognitive engagement classifications are reported. 79
4.9 Within-subject experimental results. Accuracies are reported on data preprocessed using Filter Bank and CSP for cognitive engagement and emotional engagement classifications. The best performance average values are highlighted in bold. 80
4.10 Cross-subject experimental results using FBCSP followed by TCA. Accuracies are reported with and without for-subject average removal for cognitive and emotional engagement detection. The best performance values are highlighted in bold. 80