National Technical
University of Athens

NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

DOCTORAL THESIS

PROCESSING AND ANALYSIS OF EEG DATA RECORDINGS WITH THE APPLICATION OF MACHINE LEARNING METHODS

(ΑΝΑΛΥΣΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ ΕΓΚΕΦΑΛΟΓΡΑΦΙΚΩΝ ΚΑΤΑΓΡΑΦΩΝ ΜΕ ΤΕΧΝΙΚΕΣ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ)

IOANNIS KAKKOS

Dissertation Number: 49988
Athens, July 2021

National Technical University of Athens, School of Electrical and Computer Engineering
Ioannis Kakkos

NATIONAL TECHNICAL UNIVERSITY OF ATHENS

SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING
DIVISION OF INFORMATION TRANSMISSION SYSTEMS AND MATERIAL TECHNOLOGY

Processing and Analysis of EEG Data Recordings with the Application of Machine Learning Methods

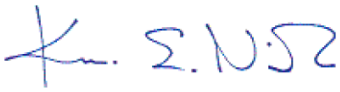
(Ανάλυση και Επεξεργασία Δεδομένων Εγκεφαλογραφικών Καταγραφών με Τεχνικές Μηχανικής Μάθησης)

IOANNIS KAKKOS

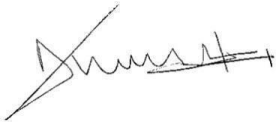
Approved by the seven-member examination board on July 21, 2021.

G. Matsopoulos
Professor N.T.U.A

E. Ventouras
Professor UNI.W.A.



K. Nikita
Professor N.T.U.A



D. Koutsouris
Professor N.T.U.A

P. Tsanakas
Professor N.T.U.A

P. Asvestas
As. Professor UNI.W.A.

Yu Sun
Professor, Zhejiang University,
Hangzhou, China

Athens, July 2021

Ιωάννης Κάκκος
Διδάκτωρ Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

This PhD thesis was co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the project "Strengthening Human Resources Research Potential via Doctorate Research" (MIS-5000432), implemented by the State Scholarships Foundation (IKY).

This work is dedicated to my son for being my inspiration to achieve greatness and keeping me sane (it did not work)

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΑΡΙΘ. ΠΡΩΤ. : 32014
» ΠΙΣΤ. :
» ΔΙΠΛ. :

ΠΙΣΤΟΠΟΙΗΤΙΚΟ

Καθώς προκύπτει από το πρακτικό εξέτασης της 7μελούς επιτροπής της Υποψήφιας Διδάκτορος ΚΑΚΚΟΣ ΙΩΑΝΝΗΣ του ΠΑΝΑΓΙΩΤΗ, Πτυχιούχου Φυσικού, Σχολής Θετικών Επιστήμων, Πανεπιστημίου Πατρών, ολοκληρώθηκε η διαδικασία του διδακτορικού της με τίτλο: "Αναλύση και Επεξεργασία Δεδομένων Εγκεφαλογραφικών Καταγραφών με Τεχνικές Μηχανικής Μάθησης".

Η Τριμελής Συμβουλευτική Επιτροπή κατέθεσε αίτημα (αριθμ.πρωτ. 26956/01.07.2021) για τον ορισμό Επταμελούς Εξεταστικής Επιτροπής το οποίο εγκρίθηκε από τη ΓΣ της Σχολής στην από 20/07/2021 συνεδρίασή της.

Η εξέταση του διδακτορικού του ενώπιον της 7μελούς επιτροπής έγινε στις 21 Ιουλίου 2021 με επιτυχία.

Για να ανακηρυχθεί διδάκτωρ του Ε.Μ.Π., της απομένει η ορκωμοσία, η οποία θα γίνει ενώπιον της Γενικής Συνέλευσης της Σχολής και προβλέπεται ότι θα οριστεί μέχρι τις 30/11/2021.

.

Μετά από αίτησή του χορηγείται το παρόν πιστοποιητικό για κάθε χρήση.

Αθήνα 26 Ιουλίου 2021
Με εντολή του Πρύτανη
Η ΑΝΑΠΛΗΡΩΤΡΙΑ ΓΡΑΜΜΑΤΕΑΣ

ΕΥΦΡΟΣΥΝΗ ΚΑΝΤΑ

Acknowledgments

The work presented in this Thesis has been carried out at the Laboratory of Biomedical Optics & Applied Biophysics and BIO-Medical Informatics Group (BIOMIG) Laboratory, Department of Electrical and Computer Engineering, National Technical University of Athens under the supervision of prof. G. Matsopoulos (who was courageous enough to accept me as a PhD student) and co-supervisor prof. E. Ventouras, to whom I am grateful for their scientific (and not only) guidance and advice. In addition, part of this thesis was carried out in collaboration with the Singapore Institute of Neurotechnology (SINAPSE), National University of Singapore under the supervision of prof. A. Bezerianos.

During this time many people have provided me their essential support, thus contributing directly or indirectly to the completion of this Thesis. As such, I would firstly like to thank A. Bezerianos for introducing me to the concept of higher-order brain function analysis and for his counseling in regard to academic life. Furthermore, I would like to express my gratitude to prof. Asvestas for his counseling and insightful research comments.

Moreover I would like to thank my good friend prof. Yu Sun for acting as a mentor on how to orchestrate research work and write scientific publications. I could not forget to also thank my colleagues and close friends G. Dimitrakopoulos, K. Gkiatis, S. Miloulis, A. Karampasi and I. Zorzos to whom I have had the pleasure to work with. They have been a source of friendship and excellent collaboration in scientific issues.

Moreover, I would like to thank my parents and sisters who supported me in all my endeavors and motivated me even during tough times in this Ph.D. pursuit.

Above all, I would like to thank my loving wife Lorena for her essential support and encouragement, especially during the final stages of this Ph.D. Above all, my special gratitude goes to our son Panos for his remarkable spirit to force me to take brakes instead of working crazy hours and his phenomenal determination in getting me up early every morning.

Table of Contents

List of Tables

List of Figures

Extended Greek Summary

Εισαγωγικό σημείωμα

Σε αυτή τη διδακτορική διατριβή, έχει εφαρμοστεί ένα ευρύ φάσμα προσεγγίσεων Μηχανικής Μάθησης (ΜΜ) σε πειράματα που σχετίζονται με τον εγκέφαλο, παρέχοντας τη βάση για αλληλεπιδράσεις Διεπαφών Ανθρώπου-Μηχανής (ΔΑΜ) και Διεπαφών Εγκεφάλου-Υπολογιστή (ΔΕΥ) ρίχνοντας φως σε υποκείμενες γνωστικές πτυχές διαφορετικών καταστάσεων όσον αφορά τη νευροεπιστήμη. Υπό αυτό το πρίσμα, πρέπει να σημειωθεί ότι η εφαρμογή ΔΕΥ απαιτεί τη συνεχή μέτρηση των εγκεφαλικών σημάτων τα οποία να μπορούν να μεταφραστούν και να εισάγουν πληροφορίες στη συνδεδεμένη συσκευή για την εκτέλεση της κάθε εργασίας. Κατ' επέκταση, μη σχετιζόμενα νευρικά ερεθίσματα ενδέχεται να συγχέουν την εσωτερική διαδικασία ταξινόμησης, επομένως είναι εξαιρετικά σημαντικό να ανιχνεύονται συγκεκριμένες νευρικές ιδιότητες με τρόπο που να είναι όσο το δυνατόν πιο διακριτές και με καθολικό τρόπο. Ο στόχος της διατριβής αυτής είναι να θέσει τα θεμέλια για την εφαρμογή ΔΕΥ αξιολογώντας τα καθολικά χαρακτηριστικά διαφορετικών γνωστικών καταστάσεων και συνθηκών στα πλαίσια ΜΜ, ενώ παράλληλα να λαμβάνει υπόψη τις εξελίξεις στις αλγοριθμικές εφαρμογές και των εφαρμογών των υπαρχόντων συστημάτων στην ανάλυση ηλεκτροεγκεφαλογραφικών (ΗΕΓ) καταγραφών.

Περίληψη

Σε αυτήν τη διδακτορική διατριβή, προτείνονται προηγμένες προσεγγίσεις Μηχανικής Μάθησης για την επεξεργασία και την ανάλυση των ηλεκτροεγκεφαλογραφικών σημάτων, αξιοποιώντας ιδιότητες σήματος στο πεδίο του χρόνου και / ή συχνότητας. Προς αυτήν την κατεύθυνση, χρησιμοποιήθηκαν τρία πειράματα ηλεκτροεγκεφαλογραφικών καταγραφών υψηλής πυκνότητας προκειμένου να μελετηθούν γνωστικές λειτουργίες υψηλότερης τάξης, αποτελώντας τη βάση για εφαρμογές πραγματικού κόσμου ειδικά στον τομέα των Διεπαφών Εγκεφάλου-Υπολογιστή.

Σε αυτήν τη διδακτορική διατριβή, προτείνονται προηγμένες προσεγγίσεις Μηχανικής Μάθησης για την επεξεργασία και την ανάλυση των ηλεκτροεγκεφαλογραφικών σημάτων, αξιοποιώντας ιδιότητες σήματος στο πεδίο του χρόνου και / ή συχνότητας. Προς αυτήν την κατεύθυνση, χρησιμοποιήθηκαν τρία πειράματα ηλεκτροεγκεφαλογραφικών καταγραφών υψηλής πυκνότητας προκειμένου να μελετηθούν γνωστικές λειτουργίες υψηλότερης τάξης, αποτελώντας τη βάση για εφαρμογές πραγματικού κόσμου ειδικά στον τομέα των Διεπαφών Εγκεφάλου-Υπολογιστή. Τα καταγεγραμμένα δεδομένα διερευνήθηκαν χρησιμοποιώντας δυναμικά που σχετίζονται με συμβάντα (Βιωματικά Δυναμικά, ΒΔ), φασματική αποσύνθεση και δίκτυα εγκεφάλου, ενώ αναπτύχθηκαν νέες μεθοδολογίες για τη μοντελοποίηση και ανάλυση των νευρολογικών δομών παρέχοντας έγκυρα αποτελέσματα. Επιπλέον, τα χαρακτηριστικά που εξήχθησαν από κάθε πείραμα συνδυάστηκαν με διαφορετικές τεχνικές ταξινόμησης και επιλογής χαρακτηριστικών (απομονώνοντας ένα μικρό υποσύνολο σημαντικών αλληλεπιδράσεων) και πετυχαίνοντας υψηλή ακρίβεια στη διάκριση μεταξύ των διαφόρων καταστάσεων. Αξίζει να σημειωθεί ότι η αποκρυπτογράφηση της ηλεκτρικής δραστηριότητας του εγκεφάλου με γενικευμένο τρόπο (άσχετο με τη μεταβλητότητα των εργασιών και των θεμάτων) είναι μια σημαντική απαίτηση για την προσαρμογή ενός κοινού πλαισίου σε πραγματικές εφαρμογές Διεπαφών Εγκεφάλου Υπολογιστή. Σε αυτό

επικεντρώνεται αυτή η διδακτορική διατριβή, μέσω της επέκτασης της τρέχουσας τεχνολογίας σε νευρογνωστικές προσαρμογές της Μηχανικής Μάθησης.

Πιο συγκεκριμένα, το πείραμα 1 (Κεφάλαιο 3) διερευνά τον αντίκτυπο διαφορετικών χρονικών παραθύρων σε ΒΔ μετά το σφάλμα σε ακουστική δοκιμασία με διαφορετικές συνθήκες πολυπλοκότητας. Ως εκ τούτου, χαρακτηριστικά σήματος (όπως πλάτος, λανθάνουσα κατάσταση κ.λπ.) που είναι αντιπροσωπευτικά των συστατικών ΒΔ που προκαλούνται από σφάλματα, χρησιμοποιήθηκαν σε ένα πλαίσιο επιλογής χαρακτηριστικών και ταξινόμησης που εξετάζει την επίδραση των διαμορφώσεων σήματος με την πάροδο του χρόνου. Το πείραμα 2 (Κεφάλαιο 4) αφορά την ανεξάρτητη (από το είδος διαδικασίας) εκτίμηση του υψηλού έναντι του χαμηλού νοητικού φόρτου εργασίας σε δύο διαφορετικά παραδείγματα μνήμης εργασίας. Στο πλαίσιο αυτό, ένα δικτύου εγκεφάλου συνδυάστηκε με φασματικά χαρακτηριστικά εγκαφαλογραφήματος για την εξαγωγή των γενικευμένων χαρακτηριστικών που είναι αντιπροσωπευτικά του νοητικού φόρτου, ενώ διαφορετικοί μέθοδοι ταξινόμησης συνδυάστηκαν με έναν αλγόριθμο επιλογής χαρακτηριστικών για την αξιολόγηση των ιδιοτήτων διακριτότητας των υποσύνολων λειτουργιών. Η περαιτέρω εξέταση των επιλεγμένων χαρακτηριστικών αποκάλυψε κοινά πρότυπα ανεξάρτητα από την εργασία σχετικά με τις φασματικές και τοπολογικές τους ιδιότητες. Τέλος, το πείραμα 3 (Κεφάλαιο 5) διερευνά τους νευρικούς μηχανισμούς του νοητικού φόρτου εργασίας και τη διαμόρφωσή τους σε συνθήκες που προσομοιώνουν σενάρια πραγματικού κόσμου. Σε αυτήν την περίπτωση, πραγματοποιήθηκε ένα καλά ελεγχόμενο πείραμα προσομοίωσης πτήσης σε περιβάλλοντα οθόνης υπολογιστή (2D) και εικονικής πραγματικότητας (3D) με πολλαπλά επίπεδα δυσκολίας. Τα προκύπτοντα δεδομένα στη συνέχεια υποβλήθηκαν σε επεξεργασία χρησιμοποιώντας εύρεση νευρολογικών πηγών και συσσωματώθηκαν σε φλοιώδεις περιοχές κατασκευάζοντας φασματικά διαστημικά δικτύων πηγής. Στη συνέχεια εφαρμόστηκαν αλγόριθμοι μηχανικής μάθησης για να συγκριθούν οι μεταβολές της συνδεσιμότητας και να εντοπιστούν τα πρότυπα αναδιοργάνωσης των εγκεφαλικών δικτύων σε σχέση με κάθε επίπεδο και κατάσταση φόρτου εργασίας.

Τα αποτελέσματα παρέχουν νέες γνώσεις στους υποκείμενους γνωστικούς μηχανισμούς, ενώ παρέχουν υποστήριξη σε ειδικούς στους σχετικούς ιατρικούς τομείς, επιτρέποντας στη συνέχεια την αποτελεσματική παρακολούθηση της εγκεφαλικής δραστηριότητας. Επιπλέον, ο συνδυασμός της ανάλυσης που παρέχεται μέσω της Μηχανικής Μάθησης με τη νευροεπιστήμη ανοίγει το δρόμο για εφαρμογές διεπαφής ανθρώπου/εγκεφάλου-μηχανής τόσο σε επιστημονικούς όσο και σε ιατρικούς τομείς. Τα πειραματικά έργα και η ανάλυση στα προαναφερθέντα πειράματα αποτελούν τη θεωρητική βάση για τέτοιες διεπαφές, των οποίων οι επιπτώσεις και η ταξινόμηση συζητούνται επίσης σε αυτό το διδακτορικό (Κεφάλαιο 6).

Κεφάλαιο 1

Στο πρώτο κεφάλαιο εισάγονται οι βασικές έννοιες και οι τεχνικές για την παρακολούθηση του εγκεφάλου, καθώς και οι ιδιότητες των νευρικών ταλαντώσεων όσον αφορά στην αναπαράσταση της φλοιώδους δραστηριότητας σε γνωστικές λειτουργίες ανώτερης τάξης. Πιο συγκεκριμένα, περιγράφονται οι τρόποι καταγραφής της δραστηριότητας του ανθρώπινου εγκεφάλου, οι τρόποι τοποθέτησης των ηλεκτροδίων, οι εγκεφαλικοί ρυθμοί, τα δυναμικά που σχετίζονται με συγκεκριμένα ερεθίσματα (Βιωματικά Δυναμικά, ΒΔ) και τα εγκεφαλικά δίκτυα (θεωρία γράφων) αναφορικά με την κατασκευή τους και τις μετρικές τους. Αναλυτικότερα, περιγράφεται το ΗΕΓ αναφορικά με τον τρόπο που λαμβάνει τις καταγραφές και η εφαρμογή του τόσο αναφορικά με τις τεχνικές όσο και με τις νευροφυσιολογικές

λεπτομέρειες. Επίσης, παρουσιάζεται η τοποθεσία των συγκεκριμένων ΗΕΓ καναλιών κατά το διεθνές σύστημα 10-20, το όποιο και χρησιμοποιήθηκε στα πειράματα που πραγματοποιήθηκαν σε αυτή τη διδακτορική διατριβή. Επιπρόσθετα, γίνεται αναφορά στα επιμέρους χαρακτηριστικά των ΗΕΓ καταγραφών με έμφαση στους διαφορετικούς εγκεφαλικούς ρυθμούς (*δ, α, θ, β, γ*) σχετικά με την συχνότητα που παρουσιάζονται, καθώς και στα ΒΔ και την ανάλυση τους σε περαιτέρω συνιστώσες (πλάτος και χρόνος εμφάνισής τους). Τέλος, αναλύονται τα εγκεφαλικά δίκτυα και πιο συγκεκριμένα η κατασκευή του δικτύου Δεικτών Καθυστέρησης Φάσης (PLI), η σημασία του και οι μετρικές που χρησιμοποιήθηκαν στην ανάλυση των δεδομένων στα πλαίσια αυτής της διδακτορικής διατριβής.

Κεφάλαιο 2

Το δεύτερο κεφάλαιο παρουσιάζει μια βιβλιογραφική επισκόπηση των υπολογιστικών εργαλείων, αλγορίθμων και των εφαρμογών στον προσδιορισμό προγνωστικών δεικτών για την ταξινόμηση διαφορετικών γνωστικών καταστάσεων και συνθηκών. Αναλυτικότερα, παρουσιάζεται λεπτομερώς η προεπεξεργασία του ΗΕΓ σήματος και οι τεχνικές αποθορυβοποίησής του με τη χρήση φίλτρων πεπερασμένης και άπειρης κρουστικής απόκρισης (FIR, IIR), Ανάλυσης Ανεξάρτητων Συστατικών (ICA) και διόρθωσης γραμμής βάσης. Επίσης αναλύεται ο τρόπος με τον οποίο γίνεται ο εντοπισμός εγκεφαλικής πηγής με την μέθοδο που χρησιμοποιείται στις αναλύσεις των παρακάτω κεφαλαίων (ηλεκτρομαγνητική τομογραφία εγκεφάλου χαμηλής ανάλυσης, LORETA) και ο τρόπος υπολογισμού των πηγών. Τέλος, παρουσιάζονται τα εργαλεία ΜΜ αναφορικά με τους αλγορίθμους ταξινόμησης και πιο συγκεκριμένα αναφέρονται οι μεθοδολογίες των ταξινομητών, οι τρόποι εξαγωγής χαρακτηριστικών, η επιλογή τους και οι μετρικές για την αξιολόγηση των αποτελεσμάτων που προκύπτουν.

Κεφάλαιο 3

Το Κεφάλαιο 3 αφορά στην ταξινόμηση του γνωσιακού εγκεφαλικού μηχανισμού σφάλματος σε σχέση με εξαγόμενα χαρακτηριστικά ΒΔ, εστιάζοντας σε γενικευμένες νευροφυσιολογικές πτυχές που δεν λαμβάνουν υπόψη την πολυπλοκότητα των εργασιών. Πιο συγκεκριμένα, βάση της υπόθεσης ότι οι εγκεφαλικοί μηχανισμοί επεξεργασίας σφαλμάτων περιέχουν υψηλή μορφολογική ευαισθησία, οι μέθοδοι ΜΜ είναι ιδανικοί για να αποκαλύψουν τα κρυμμένα χαρακτηριστικά που αντικατοπτρίζουν την επεξεργασία τους, οδηγώντας στον εντοπισμό των χαρακτηριστικών αυτών που δεν εξαρτώνται από την τη δυσκολία μιας εργασίας. Ως εκ τούτου, εφαρμόστηκε ένα πείραμα ακουστικής αναγνώρισης ΗΕΓ με δύο συνθήκες πολυπλοκότητας με τα προκύπτοντα δεδομένα να αναλύονται υπό την προϋπόθεση ότι η επεξεργασία σφαλμάτων είναι ένα δυναμικό φαινόμενο, με διακριτά μορφολογικά χαρακτηριστικά ΒΔ. Η ανάλυση χωρίστηκε σε δύο διακριτές φάσεις, συμπεριλαμβανομένων των προσεγγίσεων ΜΜ σε μορφολογικά χαρακτηριστικά με διαφορετικά ή συνδυασμένα χρονικά παράθυρα. Η πρώτη προσέγγιση έλαβε υπόψη το σύνολο των καναλιών και αφορούσε δράστες και παρατηρητές χρησιμοποιώντας μια τροποποιημένη έκδοση επιλογής χαρακτηριστικών (ΕΧ), ενώ η δεύτερη χρησιμοποίησε ένα πλαίσιο ΕΧ σε συγκεκριμένα κανάλια των δραστών για τις ταξινομήσεις των σημάτων ξεχωριστών και από κοινού καταστάσεων πολυπλοκότητας. Και οι δύο αναλύσεις ήταν σε θέση να παρέχουν υψηλή απόδοση ταξινόμησης, και παρουσιάζονται παρακάτω.

Τα δεδομένα ΗΕΓ συλλέχθηκαν από δώδεκα άτομα τα οποία συμμετείχαν σε ένα ακουστικό πείραμα αναγνώρισης με δύο επίπεδα πολυπλοκότητας. Τα δεδομένα πάρθηκαν από 32 κανάλια ΗΕΓ σύμφωνα με το διεθνές σύστημα 10-20. Οι συμμετέχοντες χωρίστηκαν σε 6 δυάδες, όπου τόσο ο δράστης όσο και ο

παρατηρητής άκουγαν τον ίδιο τόνο από μια προεπιλεγμένη ζώνη συχνοτήτων. Στη συνέχεια, ζητήθηκε από τον δράστη να τοποθετήσει έναν κέρσορα σε μια μπάρα η οποία αντιστοιχούσε στο εύρος ζώνης συχνοτήτων, έτσι ώστε η επιλεγμένη θέση να ταιριάζει με το ακουστικό ερέθισμα. Μετά την τοποθέτηση του κέρσορα παρουσιαζόταν ο τόνος που αντιστοιχεί στη θέση της μπάρας καθιστώντας προφανές εάν η επιλογή ήταν σωστή ή εσφαλμένη. Κάθε εργασία περιελάμβανε 80 δοκιμές με τους συμμετέχοντες να εναλλάσσουν τη θέση τους ως δράστες και παρατηρητές. Στην πρώτη κατάσταση πολυπλοκότητας οι συμμετέχοντες άκουγαν ήχους από το ίδιο εύρος συχνοτήτων και έτσι η μπάρα θα ήταν πανομοιότυπη και για τους δύο, ενώ στη δεύτερη κάθε συμμετέχων θα λάμβανε διαφορετικά ακουστικά ερεθίσματα εύρους συχνοτήτων ενώ ήταν δράστης. Με αυτόν τον τρόπο θα ήταν πιο δύσκολο να χαρτογραφηθεί διανοητικά η διαφορετική μπάρα στην δεύτερη κατάσταση. Τα ληφθέντα δεδομένα ΗΕΓ τμηματοποιήθηκαν σε σχέση με τον τόνο ανάδρασης και υποβλήθηκαν σε προ-επεξεργασία με φίλτρο άπειρης κρουστικής απόκρισης Chebyshev και σε διόρθωσης γραμμής βάσης 100 ms πριν από τη διέγερση. Στη συνέχεια, υπολογίστηκε ο μέσος όρος του ΗΕΓ σήματος για κάθε συμμετέχοντα και για κάθε ηλεκτρόδιο, καταχωρώντας την απάντηση ως σωστή ή εσφαλμένη με βάση το Ισοδύναμο Ορθογώνιο Εύρος Ζώνης σύμφωνα με τη θεωρία της ψυχοακουστικής.

Για να εκτελεστεί η ταξινόμηση στις δύο κλάσεις πραγματοποιήθηκε μια σειρά εξαγωγής χαρακτηριστικών για καθεμία από τις δύο προσεγγίσεις. Και στις δύο τα χαρακτηριστικά υπολογίστηκαν με βάση την οπτική παρατήρηση των μέσων καμπυλών ΒΔ των διαφορετικών ηλεκτροδίων για διαφορετικά χρονικά παράθυρα, όπου τα σχετιζόμενα με σφάλμα ΒΔ, σύμφωνα με την βιβλιογραφία, εμφανίζουν ξεχωριστή μορφολογία και σαφή διαφοροποίηση μεταξύ των 2 κλάσεων (σωστών και λανθασμένων). Ως εκ τούτου, τα σωστά και λανθασμένα ΒΔ χωρίστηκαν περαιτέρω σε 5 χρονικά παράθυρα (0-125 ms, 125-220 ms, 220-300 ms, 300-400 ms και 0-600 ms), ενώ τα χαρακτηριστικά εξήχθησαν για τα χρονικά παράθυρα και για ανά δύο συνδυασμούς τους για κάθε θέση ηλεκτροδίου ως: μέγιστη και ελάχιστη τιμή πλάτους του σήματος ΒΔ, λανθάνων χρόνος της μέγιστης και ελάχιστης τιμής και το εμβαδό της περιοχής κάτω από την καμπύλη ΒΔ.

1η Προσέγγιση: Ανάλυση σχετιζομένων με σφάλμα ΒΔ δραστών και παρατηρητών

Όπως αναφέρθηκε και προηγουμένως τα πειραματικά δεδομένα αναλύθηκαν στην πρώτη προσέγγιση χρησιμοποιώντας ένα πλήρες σύνολο καναλιών σε δράστες και παρατηρητές σε ένα πλαίσιο ΜΜ που ενσωμάτωνε την ΕΧ και τις διαδικασίες ταξινόμησης. Πιο συγκεκριμένα εφαρμόστηκε η διαδοχική πρόσθια επιλογή (SFS) για να εντοπιστούν τα πιο ενδεικτικά χαρακτηριστικά, όσον αφορά στην ακρίβεια της ταξινόμησης.

Γενικά, η SFS ξεκινά με ένα κενό σύνολο χαρακτηριστικών και προσθέτει ένα προς ένα το χαρακτηριστικό με την μεγαλύτερη σημασία (ακρίβεια ταξινόμησης) έως ότου δεν μπορεί να υπάρξει περαιτέρω βελτίωση. Ωστόσο, το γεγονός ότι σε κάθε βήμα το SFS πρέπει να εκτιμήσει την ακρίβεια ταξινόμησης όλων των χαρακτηριστικών, έχει ως αποτέλεσμα μεγάλο υπολογιστικό κόστος, ενώ η φύση του αλγορίθμου τον παγιδεύει σε τοπικά μέγιστα. Για το λόγο αυτό, ένας τροποποιημένος αλγόριθμος SFS χρησιμοποιήθηκε, όπου τα χαρακτηριστικά ταξινομούνται με βάση τη συνολική τους σημασία στην ταξινόμηση. Στη συνέχεια, μια διαδικασία βαθμολόγησης εντοπίζει το βασικό υποσύνολο (αντί για το μηδενικό) ως αυτό που παρουσιάζει την καλύτερη απόδοση μέσω της εξαντλητικής αναζήτησης των δέκα πιο σημαντικών χαρακτηριστικών. Τέλος, ο αλγόριθμος πραγματοποιεί μια επαναληπτική δομή κατά την οποία ενσωματώνει το επόμενο πιο σημαντικό χαρακτηριστικό στο υποσύνολο μόνο αν αυξηθεί η απόδοση (χωρίς να ληφθεί υπόψη το ποσό αύξησης), ενώ εάν η συνολική απόδοση επιδεινωθεί ή δεν

αλλάξει παραλείπεται για να προστεθεί το επόμενο. Για να αξιολογηθεί η ακρίβεια της ταξινόμησης σε κάθε επανάληψη, εφαρμόστηκε μια διαδικασία διασταυρούμενης επικύρωσης, που συνεπάγει τη χρήση όλων των δεδομένων με εξαίρεση ενός, τα οποία και χρησιμοποιούνται ως σύνολο εκπαίδευσης με το αποκλειόμενο να αποτελεί το σύνολο δοκιμών. Οι ταξινομητές που χρησιμοποιήθηκαν ήταν γραμμικές Μηχανές Διανυσμάτων Υποστήριξης (SVM), SVM με συνάρτηση ακτινικής βάσης και με τετραγωνικό πυρήνα, σε συνδυασμό με τις μεθόδους εκμάθησης Διαδοχικής Ελάχιστης Βελτιστοποίησης και Ελαχίστων Τετραγώνων.

Τα αποτελέσματα την γενικής μορφής του SFS και της τροποποιημένης έκδοσης παρουσιάζονται στο Κεφάλαιο 3, Πίνακες 3.1, 3.2, 3.3 και 3.4. Συνοπτικά, τα χαρακτηριστικά που περιέχονταν στο πρώτο χρονικό παράθυρο ήταν επιλεγμένα με μεγαλύτερη συχνότητα, με το ηλεκτρόδιο P8 να επιλέγεται από τους δύο SFS αλγορίθμους πιο συχνά. Η γενική μορφή του SFS έφτασε σε αποτελέσματα υψηλής ακρίβειας με μικρά βήματα, παρόλα αυτά παγιδεύτηκε σε τοπικά μέγιστα (επιβεβαιώνοντας την αρχική υπόθεση), ενώ η τροποποιημένη έκδοση βελτίωσε την ακρίβεια ταξινόμησης (με περισσότερα βήματα όμως) και παρουσίασε σημαντική μείωση του υπολογιστικού φορτίου. Τέλος, φάνηκε ότι τα χαρακτηριστικά (όπως το πλάτος και ο χρόνος εμφάνισης) εμφανίζουν μορφολογικές διαφορές σε κάθε κατάσταση πολυπλοκότητας, καλύπτοντας συγκεκριμένα συστατικά ΒΔ. Αυτό το γεγονός αποτέλεσε το βασικό κίνητρο για περαιτέρω διερεύνηση της επίδρασης της δυσκολίας της αιτούμενης από τον δράστη εργασίας, υιοθετώντας μια νέα προσέγγιση που περιγράφεται παρακάτω.

2η Προσέγγιση: Ανάλυση σχετιζομένων με σφάλμα ΒΔ διαφορετικών καταστάσεων πολυπλοκότητας

Σε αυτή την προσέγγιση, χρησιμοποιήθηκε μια στρατηγική αναζήτησης χαρακτηριστικών που ενσωμάτωσε έναν συνδυασμό διαδοχικής κυμαινόμενης επιλογής προς τα εμπρός (SFFS) και διαδοχικής επιλογής προς τα εμπρός (SFS), επιτρέποντας την ανίχνευση των μεμονωμένων χαρακτηριστικών που παρέχουν υψηλή ακρίβεια ταξινόμησης. Επίσης, η προσέγγιση διαφοροποιήθηκε σχετικά με τους συμμετέχοντες και τα χαρακτηριστικά που χρησιμοποιήθηκαν. Ως εκ τούτου, η μεθοδολογική στρατηγική περιλάμβανε μόνο ΗΕΓ δραστών στις δύο συνθήκες πολυπλοκότητας και μια υποομάδα ηλεκτροδίων (τα 7 πιο ενδεικτικά σύμφωνα με την βιβλιογραφία που βρίσκονται στην κεντρική περιοχή του τριχωτού της κεφαλής).

Η διάκριση μεταξύ των απαντήσεων ως σωστών ή εσφαλμένων πραγματοποιήθηκε με ταξινομητή SVM με Διαδοχική Ελάχιστη Βελτιστοποίηση, Ελάχιστα Τετράγωνα και Τετραγωνικό Προγραμματισμό, ενώ δοκιμάστηκαν λειτουργίες πυρήνων SVM γραμμικοί, συναρτήσεων ακτινικής βάσης, τετραγωνικοί, πολυστρωματικών perceptron και πολυωνυμικοί. Δυο διαδικασίες έλαβαν χώρα. Στην πρώτη εφαρμόστηκε μια τυπική SFFS, η οποία περιλαμβάνει πανομοιότυπη λειτουργία με την SFS (αρχίζοντας από ένα κενό σύνολο χαρακτηριστικών), αλλά μετά από κάθε βήμα προς τα εμπρός μπορεί να εκτελεί βήματα προς τα πίσω, δεδομένου της αύξησης της τιμής της αντικειμενικής συνάρτησης (αφαιρώντας χαρακτηριστικά που προστέθηκαν σε προηγούμενα βήματα). Ως αντικειμενική συνάρτηση θεωρήθηκε η ακρίβεια του ταξινομητή χρησιμοποιώντας δεδομένα από τις δυο καταστάσεις πολυπλοκότητας ταυτόχρονα. Στην δεύτερη διαδικασία η ακρίβεια λαμβάνεται χρησιμοποιώντας τη μέση τιμή της ακρίβειας της πρώτης και δεύτερης κατάστασης ταυτόχρονα, καθώς και ξεχωριστά της πρώτης και της δεύτερης. Μετά την επιλογή του βέλτιστου υποσυνόλου χαρακτηριστικών, εφαρμόστηκε η SFS ξεχωριστά στα δεδομένα κάθε κατάστασης πολυπλοκότητας ξεκινώντας από ένα αρχικό σύνολο

χαρακτηριστικών επιλεγμένο από την SFFS. Η εκπαίδευση και δοκιμή του κάθε αλγορίθμου ταξινόμησης υλοποιήθηκε με τη χρήση της ίδιας διαδικασίας διασταυρούμενης επικύρωσης σε κάθε βήμα των μεθόδων SFFS και SFS όπως στην προηγούμενη προσέγγιση. Επιπλέον, οι διαδικασίες ΕΧ και ταξινόμησης δεν εφαρμόστηκαν μόνο για τα πέντε χρονικά παράθυρα, αλλά και για συνδυασμούς τους ανά δύο.

Τα αποτελέσματα των δυο διαδικασιών παρουσιάζονται στο Κεφάλαιο 3, Πίνακες 3.5-3.10. Η σύγκριση μεταξύ των δυο διαδικασιών δείχνει ότι η δεύτερη παρείχε καλύτερα αποτελέσματα, τόσο όσον αφορά στην ακρίβεια του ταξινομητή, όσο και σχετικά με το πλήθος των χαρακτηριστικών που επελέγησαν (σημαντικά μικρότερος). Η υψηλή ακρίβεια ταξινόμησης που επιτεύχθηκε για την ταξινόμηση μεταξύ των διαφορετικών καταστάσεων πολυπλοκότητας αλλά και στην κάθε μια κατάσταση ξεχωριστά επιβεβαιώνει την αρχική μας υπόθεση ότι, οι διαδικασίες ΜΜ μπορούν να ανιχνεύσουν επιτυχώς κρυμμένα μοτίβα σε χαρακτηριστικά ΒΔ και έτσι να εντοπιστούν οι λανθασμένες αποφάσεις, ανεξάρτητα από τη δυσκολία της κάθε εργασίας. Επίσης ο συνδυασμός χρονικών παραθύρων στηρίζει την υπόθεση ότι τα διαφορετικά συστατικά των ΒΔ προσφέρουν μεγαλύτερη ευελιξία στις διαδικασίες ταξινόμησης.

Συμπερασματικά, οι διαφορετικές προσεγγίσεις έδειξαν ότι η ταξινόμηση σε διαφορετικές συνθήκες δυσκολίας μπορεί να επιτευχθεί, ωστόσο, λόγω της μεταβλητότητας των χαρακτηριστικών, η εκτίμηση των καθολικών χαρακτηριστικών αποτελεί σημαντικό εμπόδιο. Ως εκ τούτου, συνδυάζοντας χαρακτηριστικά διαφορετικών χρονικών παραθύρων μπόρεσαν να εντοπιστούν δείκτες ενός πιθανώς γενικευμένου γνωσιακού μηχανισμού σφάλματος ανεξάρτητου από την πολυπλοκότητα των εργασιών.

Κεφάλαιο 4

Στο Κεφάλαιο 4, εφαρμόστηκε ανίχνευση του νοητικού φόρτου εργασίας, ανεξάρτητη από την εργασία που πραγματοποιήθηκε, χρησιμοποιώντας την συγχώνευση διαφορετικών φασματικών χαρακτηριστικών ΗΕΓ. Η ανάλυση που προτείνεται σε αυτό το Κεφάλαιο περιλαμβάνει ένα πλαίσιο ΜΜ για την αντιμετώπιση των εμποδίων που προκύπτουν συχνά από την ταξινόμηση φόρτου εργασίας μεταξύ εργασιών, χρησιμοποιώντας φασματικά χαρακτηριστικά ΗΕΓ για να αποκαλυφθούν οι γενικευμένοι μηχανισμοί που ρυθμίζουν τον νοητικό φόρτο εργασίας. Αναλυτικότερα, υπολογίστηκαν η λειτουργική συνδεσιμότητα και τα χαρακτηριστικά ισχύος φασματικής πυκνότητας από διαφορετικές ζώνες συχνοτήτων σε ένα πειραματικό πρωτόκολλο που αποτελούσε δύο εργασίες με διαφοροποιημένα επίπεδα δυσκολίας. Τα χαρακτηριστικά στη συνέχεια τροφοδοτήθηκαν σε ένα πλαίσιο ΕΧ και ταξινόμησης για να εκτιμηθεί η ποιότητα των χαρακτηριστικών και στη συνέχεια να εκτιμηθεί η απόδοση ταξινόμησης.

Πιο συγκεκριμένα, στο πείραμα συμμετείχαν 40 άτομα από τους οποίους ζητήθηκε η εκτέλεση δυο εργασιών λειτουργικής μνήμης (n-Back και Mental Arithmetic) σε διαφορετικά επίπεδα δυσκολίας. Σχετικά με την n-Back, χρησιμοποιήθηκε η 0-back και 2-back για χαμηλά και υψηλά επίπεδα φόρτου εργασίας αντίστοιχα, με την 0-Back να ζητείται τους συμμετέχοντες να απαντήσουν εάν βλέπουν το γράμμα «Χ» και τη 2-Back το κεφαλαίο γράμμα που αντιστοιχούσε στο γράμμα που παρουσιάστηκε δύο δοκιμές νωρίτερα. Στην εργασία Mental Arithmetic, τα δύο επίπεδα φόρτου εργασίας περιλαμβάνουν προσθέσεις με μονοψήφιους και με τριψήφιους αριθμούς. Σε κάθε δοκιμή παρουσιάστηκε μια πρόσθεση και στη συνέχεια το άθροισμα ή ένας αριθμός με παρόμοια τιμή. Οι συμμετέχοντες έπρεπε να κρίνουν εάν η απάντηση που δόθηκε ήταν σωστή.

Η κάθε καταγραφή ΗΕΓ περιελάμβανε 64 ηλεκτρόδια σύμφωνα με το διεθνές σύστημα 10-20, με τα ακατέργαστα σήματα να ψηφιοποιούνται σε συχνότητα δειγματοληψίας 256 Hz, να φιλτράρονται με φίλτρο πεπερασμένης κρουστικής απόκρισης (FIR 1 - 40 Hz), να αποθορυβοποιούνται με χρήση ICA και στη συνέχεια να πραγματοποιείται διόρθωσης γραμμής βάσης, 100ms πριν από την έναρξη του ερεθίσματος. Στη συνέχεια, τα ΗΕΓ σήματα χωρίστηκαν σε επιμέρους τμήματα όπου υπολογίστηκε η σχετική φασματική πυκνότητα ισχύος (PSD) και ο δείκτης υστέρησης φάσης (PLI) για ζώνες συχνοτήτων δέλτα, θήτα, άλφα και βήτα.

Προκειμένου να εκτιμηθεί η προγνωστική ισχύς των PSD και PLI χαρακτηριστικών, η ΕΧ χρησιμοποιήθηκε σε ολόκληρο το σύνολο χαρακτηριστικών ανεξάρτητα από την κάθε εργασία. Ο αλγόριθμος ΕΧ που χρησιμοποιήθηκε ήταν μια αναδρομική μέθοδος εξάλειψης χαρακτηριστικών με τη μέθοδο μείωσης συσχέτισης μεροληψίας (RFE-CBR) που εκτιμά τη σημασία κάθε χαρακτηριστικού (και κατά συνέπεια την κατάταξή της σημαντικότητάς του) βάσει ενός εσωτερικού γραμμικού SVM. Όπως και στο προηγούμενο Κεφάλαιο χρησιμοποιήθηκε διασταυρούμενη επικύρωση χρησιμοποιώντας κάθε φορά ένα σετ εκπαίδευσης εξαιρουμένων των δεδομένων ενός συμμετέχοντα. Η ταξινόμηση εφαρμόστηκε με την ίδια διασταυρούμενη επικύρωση με έναν γραμμικό ταξινομητή SVM. Ως βέλτιστο υποσύνολο χαρακτηριστικών θεωρήθηκε αυτό με την υψηλότερη μέση ακρίβεια στην ταξινόμηση μεταξύ των επιπέδων δυσκολίας και των δύο εργασιών ταυτόχρονα, ξεκινώντας με ένα κενό σύνολο χαρακτηριστικών και προσθέτοντας κατά σειρά ένα προς ένα τα καταταγμένα χαρακτηριστικά από τη ΕΧ. Τα αποτελέσματα της προτεινόμενης προσέγγισης βρίσκονται στον Πίνακα 4.1 του Κεφαλαίου 4. Από τα χαρακτηριστικά που επελέγησαν περίπου το 10% αφορούσαν PSD με τα υπόλοιπα ήταν PLI.

Η προσέγγιση ΜΜ που προτείνεται, έδωσε 94% ακρίβεια ταξινόμησης (την μεγαλύτερη ως τώρα στην σύγχρονη βιβλιογραφία σύμφωνα με την αναζήτηση των συγγραφέων της σχετικής εργασίας) στις διακρίσεις του νοητικού φορτίου, ενώ η ανάλυση των επιλεγμένων χαρακτηριστικών έδειξε κοινές τάσεις ανεξάρτητες από την εργασία σε ιδιότητες φασματικής ισχύος και λειτουργικής συνδεσιμότητας. Συγκεκριμένα, ανιχνεύθηκε αυξημένη ισχύς σε δ και θ ρυθμό με την αύξηση του επιπέδου του φόρτου εργασίας, ενώ εντοπίστηκαν διαφοροποιήσεις της λειτουργικής συνδεσιμότητας όσον αφορά στη συχνότητα και στις θέσεις των ηλεκτροδίων στην επιφάνεια της κεφαλής. Τέλος, τα χαρακτηριστικά που εντοπίστηκαν παρείχαν δείκτες της προγνωστικής τους ποιότητας, αλλά και επέδειξαν τον γνωστικό έλεγχο των ανασταλτικών λειτουργιών ως τον κύριο παράγοντα ενός γενικευμένου (universal) μηχανισμού φόρτου εργασίας. Ως εκ τούτου, αποδεικνύεται ότι η συνεχής νοητική ενημέρωση με κάθε νέα δοκιμή απαιτεί επιπλέον πόρους χωρητικότητας μνήμης, ενώ αγνοεί άσχετα ερεθίσματα προηγούμενων δοκιμών. Η συνολική προσέγγιση απεικονίζει την αποτελεσματικότητα του συνδυασμού των χαρακτηριστικών στην προσπάθεια ανίχνευσης ενδεικτικών βιοδεικτών νοητικού φορτίου ανεξάρτητα από την εφαρμοζόμενη εργασία.

Κεφάλαιο 5

Στο Κεφάλαιο 5, λαμβάνονται υπόψη τα αποτελέσματα και τα συμπεράσματα των προηγούμενων Κεφαλαίων και εφαρμόζεται αξιολόγηση του φόρτου εργασίας σε μια προσέγγιση πραγματικού κόσμου. Πιο συγκεκριμένα, πραγματοποιήθηκε πείραμα προσομοίωσης πτήσης σε περιβάλλον 2 και 3 διαστάσεων με προσέγγιση ΜΜ πολλαπλών επιπέδων φόρτου εργασίας, χρησιμοποιώντας λειτουργικές διαφορές εγκεφαλικών δικτύων των εγκεφαλικών δομών. Αναλυτικότερα, η λειτουργική συνδεσιμότητα από τις εγκεφαλικές πηγές υπολογίστηκε και χρησιμοποιήθηκε ως χαρακτηριστικό ταξινόμησης

δείχνοντας κοινά και μεμονωμένα χαρακτηριστικά και στα δύο περιβάλλοντα σε συγκεκριμένες συχνότητες εγκεφάλου.

Το πειραματικό πρωτόκολλο περιελάμβανε 33 άτομα τα οποία πήραν μέρος σε δύο συνεδρίες προσομοίωσης πτήσης χρησιμοποιώντας δύο περιβάλλοντα: α) μια συνεδρία χρησιμοποιώντας μια οθόνη υπολογιστή και β) μια συνεδρία εικονικής πραγματικότητας. Κάθε συνεδρία προσομοίωσης πτήσης αποτελούταν από τρία στάδια με αυξανόμενη δυσκολία, σχεδιασμένα να προκαλούν διαφορετικά επίπεδα νοητικού φόρτου εργασίας. Τα τρία στάδια περιλάμβαναν: την κατάσταση αυτόματου πιλότου, στην οποία δεν απαιτούνταν λειτουργία αεροσκάφους (ελάχιστο επίπεδο φόρτου εργασίας), την κατάσταση χειροκίνητης λειτουργίας, στην οποία ζητήθηκε από τα άτομα να χειριστούν το αεροσκάφος και την χειροκίνητη λειτουργία αεροσκάφους με ταυτόχρονα κενά αέρος και προβλήματα κατά την πτήση. Τα συνεχή δεδομένα ΗΕΓ καταγράφηκαν σε υψηλή ανάλυση από 64 ηλεκτρόδια, ενώ ψηφιοποιήθηκαν με ρυθμό δειγματοληψίας 256 Hz και εφαρμόστηκε ένα φίλτρο πεπερασμένης κρουστικής απόκρισης FIR διέλευσης ζώνης 0,5 - 40 Hz. Για την αποθορυβοποίηση χρησιμοποιήθηκε ανάλυση ανεξάρτητων συστατικών και τα συστατικά που έδειξαν υψηλή συσχέτιση με τα σήματα από τα ηλεκτρόδια στα μάτια απορρίφθηκαν.

Για τον εντοπισμό των επιπτώσεων του νοητικού φόρτου εργασίας και την αξιολόγηση των αποτελεσμάτων, χρησιμοποιήθηκε ο εντοπισμός εγκεφαλικών πηγών στις χρονικές σειρές ΗΕΓ για κάθε περίοδο λειτουργίας και επίπεδο φόρτου εργασίας στις δ, θ, α, β και γ ζώνες συχνοτήτων. Οι εγκεφαλικές πηγές που αντιστοιχούν στα επιφανειακά ΗΕΓ σήματα προσεγγίστηκαν με εφαρμογή ηλεκτρομαγνητικής τομογραφίας χαμηλής ανάλυσης (eLORETA) σε πρότυπο MNI152, όπου το αποτέλεσμα τμηματοποιήθηκε σε 80 περιοχές ενδιαφέροντος και υπολογίστηκε η λειτουργική συνδεσιμότητα μεταξύ κάθε ζεύγους περιοχών με PLI. Επίσης εκτιμήθηκε η τυπολογική γενικευμένη απόδοση και η τοπική απόδοση των εγκεφαλικών δικτύων.

Για την ΕΧ χρησιμοποιήθηκε ο αλγόριθμος RFE-CBR όπως και στο προηγούμενο Κεφάλαιο, ενώ η ταξινόμηση έγινε με μέθοδο τυχαίου υποδιαστημικού συνόλου με Ανάλυση Γραμμικής Διάκρισης (LDA) ως βασικούς ταξινομητές. Η μέθοδος τυχαίου υποδιαστημικού συνόλου αποσυνθέτει τα δεδομένα εισόδου σε τυχαία υποσύνολα χώρου δυνατοτήτων για την εφαρμογή μεμονωμένων ταξινομήσεων LDA οι οποίες εκπαιδεύονται ξεχωριστά και στη συνέχεια λαμβάνεται συλλογική απόφαση με πλειοψηφία. Το υποσύνολο χαρακτηριστικών με την υψηλότερη συνολική ακρίβεια ταξινόμησης καθορίστηκε ως το βέλτιστο για την διάκριση επιπέδου φόρτου εργασίας. Τα αποτελέσματα την προτεινόμενης προσέγγισης παρουσιάζονται στους πίνακες 5.2, 5.3 και στην εικόνα 5.4 του Κεφαλαίου 5.

Η προτεινόμενη μεθοδολογία ήταν επιτυχής στον εντοπισμό των κοινών και διακριτών εγκεφαλικών χαρακτηριστικών του νοητικού φορτίου μεταξύ των δύο περιβαλλόντων προσομοίωσης, παρουσιάζοντας υψηλή ικανότητα διάκρισης στη διαδικασία ταξινόμησης, ενώ χρησιμοποίησε μόνο ένα μικρό μέρος του πλήρους συνόλου χαρακτηριστικών. Βάσει των παραγόμενων αποτελεσμάτων εντοπίστηκε εντονότερη λειτουργία στις μετωπιαίες περιοχές στον ρυθμό α και στα δυο περιβάλλοντα προσομοίωσης με κοινές συνδέσεις που υπάρχουν σε όλα τα επίπεδα πολυπλοκότητας. Αντίθετα, οι ζώνες θ και β παρουσίασαν τυπολογικές ανισότητες στην τοπική και γενικευμένη απόδοση μεταξύ των δύο περιβαλλόντων. Αυτά τα αποτελέσματα επιτρέπουν μια αποτελεσματική αξιολόγηση των γνωστικών υποστρωμάτων που διέπουν τον φόρτο εργασίας σε ρεαλιστικά σενάρια. Τέλος, το περιβάλλον εικονικής πραγματικότητας εκτιμάται ότι προκάλεσε υψηλότερο νοητικό φόρτο εργασίας στους συμμετέχοντες σε σύγκριση με το περιβάλλον δισδιάστατης απεικόνισης. Σε αυτό δεν συνέβαλαν μόνο οι πιο απαιτητικές οπτικά γνωστικές

διαδικασίες, αλλά και τα επίπεδα άγχους όταν οι συμμετέχοντες είχαν περιορισμένο έλεγχο, καθώς οι περιοχές της αμυγδαλής του εγκεφάλου ήταν περισσότερο ενεργές στη διεπαφή τεχνητής πραγματικότητας, ενισχύοντας την ψυχική πίεση που προκαλείται από μια πιο ρεαλιστική κατάσταση που προσομοιώνει καταστάσεις σε σενάρια πραγματικού κόσμου.

Κεφάλαιο 6

Λαμβάνοντας υπόψη τα παραπάνω Κεφάλαια, στο Κεφάλαιο 6 παρουσιάζεται η βάση των βιοηλεκτρικών σημάτων και οι συνδυασμοί με τους σχετικούς αλγόριθμους ΜΜ σε σχέση με τις τεχνολογικές και μεθοδολογικές εξελίξεις του ΔΑΜ/ΔΕΥ. Αναλυτικότερα περιγράφεται η αποκωδικοποίηση της ανθρώπινης πρόθεσης και η μετάφρασή της σε πραγματικό χρόνο με μια εντολή ελέγχου για την καθοδήγηση υλικού και/ή λογισμικού, η οποία με τη σειρά της ενοποιείται ως μια αλληλεπίδραση μεταξύ ενός ατόμου και μηχανής. Στη συνέχεια αναλύονται τα δεδομένα ως βιοηλεκτρικά σήματα που αντανακλούν τη βιολογική δραστηριότητα (όπως μυϊκά ή ΗΕΓ σήματα) και συστατικά της εγκεφαλικής δραστηριότητας (όπως ΒΔ, αποσυγχρονισμού ή συγχρονισμού ΗΕΓ κ.ά.) Μεταξύ άλλων αναφέρονται μη βιολογικά δεδομένα κίνησης, τα οποία μπορούν να εξαχθούν μέσω μη επεμβατικών κινητών αισθητήρων, αναλύοντας τις παραμέτρους που περιλαμβάνουν μετατόπιση / περιστροφή των άκρων καθώς και τις δυνάμεις αντίδρασης.

Στην βάση της δημιουργίας ενός συστήματος ΔΕΥ το σύνολο των διαθέσιμων μεθοδολογιών συνήθως επηρεάζει την απόδοση της ταξινόμησης, την υπολογιστική πολυπλοκότητα και τον χρόνο απόκρισης, ενώ ο ρόλος των χρησιμοποιούμενων χαρακτηριστικών είναι ζωτικής σημασίας, καθώς αντιπροσωπεύουν την κατανόηση των διαδικασιών κίνησης. Σε αυτό το πλαίσιο, παρουσιάζεται ένα πλήθος αλγορίθμων που περιλαμβάνουν ΕΧ και ταξινομητές και έχουν χρησιμοποιηθεί για να αποκαλύψουν γραμμικές και μη γραμμικές εξαρτήσεις των χαρακτηριστικών.

Τέλος προτείνονται υβριδικά πλαίσια, μέσω της ενσωμάτωσης πολλαπλών φυσιολογικών και μη φυσιολογικών δεδομένων προσφέροντας μια ολιστική ανάλυση κίνησης σε πραγματικό χρόνο και τρόποι αντιμετώπισης των μελλοντικών προκλήσεων του πεδίου εφαρμογής ΔΕΥ.

Κεφάλαιο 7

Στο Κεφάλαιο 7 συνοψίζονται τα συμπεράσματα των προηγούμενων Κεφαλαίων και προτείνονται κάποιες κατευθύνσεις για μελλοντική εργασία. Πιο συγκεκριμένα, αναφέρονται οι κυριότερες συνεισφορές των εργασιών που πραγματοποιήθηκαν και επισημαίνονται τα πιο σημαντικά ευρήματα ανά Κεφάλαιο.

Abstract

Cognition refers to all conscious mental activity, involved in thinking, remembering, and reasoning. As such, the different brain functions range from volitional movement to language, imagination and planning. Conventionally, brain research focuses on the perception of the external world by examining how information is processed by analyzing biomedical records based on statistical or quantitative characteristics. However, the complex attributes and interactions in the human brain pose several challenges in order to effectively elucidate the underlying neural substrates that govern higher-order cerebral functions. To address this, recent studies have suggested the inclusion of machine learning designs as a way for an effective and accurate analysis of the complicated properties of brain signals, unveiling (hidden) cognitive associations and characteristics.

In this PhD dissertation, advanced machine learning approaches for processing and analysis of electroencephalographic (EEG) signals are proposed, exploiting the signal properties in the time and/or frequency domains. Towards this direction, three high-density EEG experiments were employed in order to study higher-order cognitive functions, forming the basis for real-world applications especially in the field of Brain-Computer Interfaces (BCI). The recorded EEG data were investigated using event-related potentials (ERPs), spectral decomposition and brain networks, while new methodologies were developed for the modeling, reconstruction and analysis of neurological structures providing valid results. Moreover, the features extracted from each experiment (representing different properties of the various brain structures) were combined with different classification and feature selection techniques, isolating a small subset of important interactions which achieved high accuracy in the discrimination between the different conditions. Of note is that the deciphering of the brain electrical activity in a global manner (irrelevant of task and subject variability) is a major requirement for the adaptation of a common framework to real world BCI applications. As such, this Doctoral Thesis focuses on addressing this through expanding the current state-of-the art in neurocognitive adaptations of Machine Learning.

Specifically, experiment 1 (Chapter 3) investigates the impact of different time windows in time-locked ERP components after an error was committed, in an auditory task with varying complexity conditions. As such, signal characteristics (like amplitude, latency etc.) that are representative of error-elicited ERP components were utilized in a feature selection and classification framework examining the effect of EEG signal modulations through time. Experiment 2 (Chapter 4) regards the task-independent assessment of high *vs* low mental workload in two different working memory paradigms. Within this scope, a brain network design was combined with EEG spectral characteristics to extract the global traits that are representative of mental load, while different classification schemes were paired with a feature selection algorithm to evaluate the distinguishability properties of the fused feature subsets. Further examination of the selected features revealed common task-independent patterns regarding their spectral and localization properties. Lastly, experiment 3 (Chapter 5) investigates the neural mechanisms of mental workload and their modulation in conditions simulating real-world scenarios. On this premise, a well-controlled EEG flight simulation experiment was conducted in computer screen (2D) and virtual reality (3D) conditions with multiple levels of difficulty. The resulting data were subsequently processed using electrical source imaging (ESI) and parcellated into cortical regions constructing spectral source space networks. Machine learning algorithms were then applied to compare the connectivity alterations and identify the brain networks reorganization patterns in regard to each workload level and condition.

The outcomes provide new insights in the underlying cognitive mechanisms, while provide support to experts in the relevant medical fields, subsequently allowing effective monitoring of brain activity. In addition, the combination of machine learning analysis and processing with neuroscience pave the way for Human-Machine Interface (HMI) and BCI applications in both scientific and medical fields. The experimental works and analysis in the aforementioned experiments form the theoretical basis for HMI/BCI, the implications and taxonomy of which is also discussed in this PhD (Chapter 6).

Aims and Objectives

This Doctoral Thesis presents new techniques for the effective analysis of higher-order cognitive functions, employing Machine Learning methods. The applications of such methods in Neurosciences provides a unified approach that incorporates biomarker extraction and data-driven approaches in the detection of (otherwise hidden) cognitive operations. To that end, this Thesis's aim was to build the groundwork for BCI frameworks by carrying out multifactorial analysis of the characteristics of the signals at multiple levels giving information on neurological structures and revealing global neural substrates utilizing Machine Learning schemes. Specifically, given the tendency to consolidate available information from different sources and modes of analysis, Machine Learning techniques offer a unique opportunity to study, predict and monitor the progression of various conditions, diseases and states providing valid answers to various scientific questions. In this regard, electroencephalography experiments were performed to study complex states of the brain, such as mental load and error-related cognition. Data were analyzed using standard approaches, such as ERP analysis, power density analysis in the field of frequency and topological analysis of functional brain networks in source and sensor space. Subsequently, the above-mentioned analyses were combined with machine learning and classification algorithms to categorize the various brain states effectively and by extension find links and differences between the employed biomarkers.

The main objectives of this Thesis include – yet are not limited to – the following:

- dynamic study of the neurophysiological mechanisms underlying higher order cognitive states,
- identification of cognitive indicators related to brain mechanisms,
- development of advanced computational tools for automated effective classification,
- applicability of system transfer to real-life conditions.

PhD Structure

In a nutshell, Chapters 1 and 2 are introductory regarding the brain recordings, structures, measured cortical responses and computational tools with emphasis on characteristic indicators of electroencephalographic recordings. Chapter 3, 4 and 5 refer to the experiments employed for the implementation of this Thesis, their analyses and the implications of the results from a neuroscience and engineering perspective. Chapter 6 takes into account the research analysis in cognitive conditions and states and implements them into a theoretical and experimental approach of Machine Learning in Brain-Computer Interface scenarios.

The first chapter introduces the basic concepts and techniques for brain monitoring as well as the properties of neural oscillations in terms of the representation of the cortical activity and especially in higher-order cognitive functions.

The second chapter presents a bibliographic overview of the computational tools and algorithms and its applications in the identification of predictive indicators for the classification of different cognitive states and conditions.

The third chapter concerns the study of error-related brain activity, relevant to different complexity conditions of an auditory identification task. As such, the cognitive properties of Event Related Potentials are analyzed and a novel method for subsequent classification is proposed based on different time-windows resulting high condition-independent performance and efficient evaluation of the employed cognitive attributes.

The fourth chapter analyzes the cross-task workload discrimination limitations and introduces an efficient framework to alleviate them. Specifically, encephalographic data in two different working memory tasks, a numeric and an n-Back task, in two difficulty levels each were employed fusing sensor derived networks and spectral power density features for the machine learning application. The aforementioned analysis succeeded in high classification accuracy results (the highest so far in the literature, to the knowledge of the authors of the related study), while minimizing the number of electrodes required.

The fifth chapter focuses on multi-level workload classification in flight simulation employed in computer screen and virtual reality, approaching real-world scenarios. For the purpose of workload assessment, brain networks were constructed in source space, while the introduced feature selection and classification framework was able to detect the common and separate brain structures governing pilot mental load.

The sixth chapter takes into account the experiments conducted as well as the current state–of –the-art in Brain-Computer Interfaces and suggests advances in algorithmic applications and system implementation, while incorporating basic analysis in real and imaginary motion recordings.

Finally, the seventh chapter summarizes the main contributions and conclusions of the dissertation and lists some ideas for future extensions.

Chapter 1

1.1 Recording the Activity of the Human Brain

The human brain is the main organ of the human nervous system consisting of several structures, each responsible for specific processing, integration and coordination of the received information. The main taxonomy of the brain include the cerebrum, the brainstem and the cerebellum, the later further divided into two hemispheres which are connected through a mass of nerve cells (neurons). There are approximately 86 billion neurons in the human brain, carrying out the vast majority of communication through electric impulses (resulting in neurotransmitters' release between cells), forming complex neural pathways and circuits [1]. These subtle electrical fields generated (called post-synaptic potentials) -when averaged through thousands of neurons- represent the synchronous activity of larger brain areas and thus result in a significant reflection of electrical activity generated by the brain. Accordingly, electrodes placed on the skin (scalp) surface are able to monitor the brain activity as propagated through the different anatomical head layers (meninges, skull, skin, hair) [2].

One of the most common monitoring methods is the electroencephalogram (EEG), a typically non-invasive recording technique that can measure the voltage fluctuations of the ion flows of oriented neuron populations near the surface of the cerebrum. EEG presents several advantages over other brain imaging methods with excellent time resolution (allowing thousands of "snapshots" of electrical activity from multiple electrodes within a single second), minimal cost and the fact that recordings are passive being the predominant ones. An EEG device contains electrodes that record the brain wave patterns and the EEG machine sends the data to a computer. There are 2 types of electrodes, wet and dry. Wet electrodes are small trays made of stainless steel, tin, gold or silver coated with silver chloride while utilizing electrolytic gel as a conductor between the scalp and electrodes [3]. Dry electrodes, on the other hand, depend on mechanical contact [4]. They are significantly smaller from wet electrodes and no electrolyte is used and no skin preparation is required. However, the number of daisy chained dry EEG devices is so far significantly limited in comparison to wet EEG systems (up to 30 channels). In this Thesis we will focus on high-density wet electrode EEG (32-64 channels), although dry EEG is proposed for real-world practical applications (e.g. wearable EEG).

1.2 EEG Electrode Placement and Montages

In order to acquire a generalized interpretation of the EEG recordings, internationally recognized methods to describe and apply the location of scalp electrodes in the context of an EEG examination have been proposed. These are the "10-10" and more often "10–20" systems which are based on the relationship between the location of the electrodes and the underlying area of the brain. As such, the "10" and "20" refer to the fact that the actual distances between adjacent electrodes are either 10% or 20% of the total front–back or right–left distance of the skull [5]. The experiments employed in this Thesis (Chapters 3, 4, 5) utilize the 10-20 system with each electrode placement denoting the lobe and area of the brain. In this regard, even-numbered electrodes indicate right side locations, while odd numbers refer to areas on the left side of the scalp (**Figure 1.1**). Moreover, the pre-frontal (Fp), frontal (F), temporal (T), parietal (P), occipital (O), and central (C) denote the main recording areas with (z) sites electrodes indicating the midline sagittal plane of the skull, (Fpz, Fz, Cz, Oz). Several combinations of the aforementioned letters

designate intermediate scalp locations like AF (between Fp and F), FC (between F and C), FT (between F and T), CP (between C and P), TP (between T and P) and PO (between P and O). Moreover the "M" electrodes usually mark mastoid areas (i.e., found just behind the outer ear) and Iz placed over the inion which are commonly recorded, although they are fiducial positions they usually don't represent higher-order cognitive processes [6].

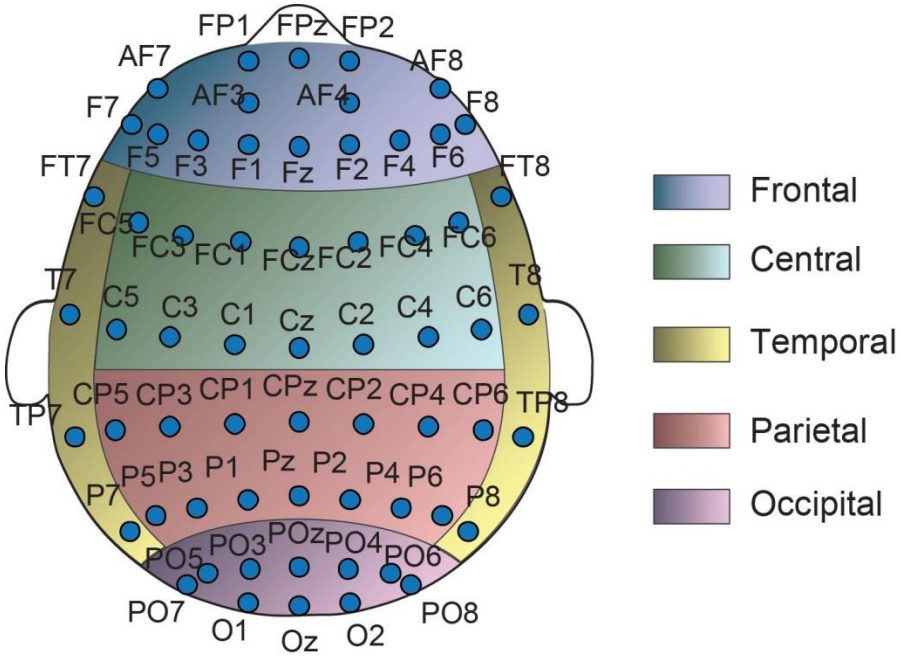


Figure 1.1 Electrode locations of 64 EEG channel according to the International 10-20 system.

1.3 EEG Rhythms and Oscillations

EEG signals can be described in terms of rhythmic activity based on signal morphologies of specific oscillations being the frequencies of the harmonics of which they are composed (spectral components). These are subdivided into bandwidths that correlate to brain functioning or condition. There are 5 main sub-bands: delta (δ, 0.5 – 3.5 Hz), theta (θ, 3.5-7 Hz), alpha (α, 7-15 Hz), beta (β, 15-30 Hz) and gamma (γ, 30-70Hz) [7] (**Figure 1.2**). Of note is that the range of the above sub-bands is not precisely defined leading to small variations between studies, while there are various scientific works that focus on lower or higher frequencies, although activity below or above these frequencies could prove to be artifactual, under standard clinical recording techniques. Other spectral components include power characteristics (i.e., the amount of energy in a frequency band, typically expressed as squared amplitude) and phase characteristics (i.e., the synchronization across several generators), while several theories have been proposed on how illness, age and external stimuli cause changes in the internal amplitude and synchronization patterns [8], [9].

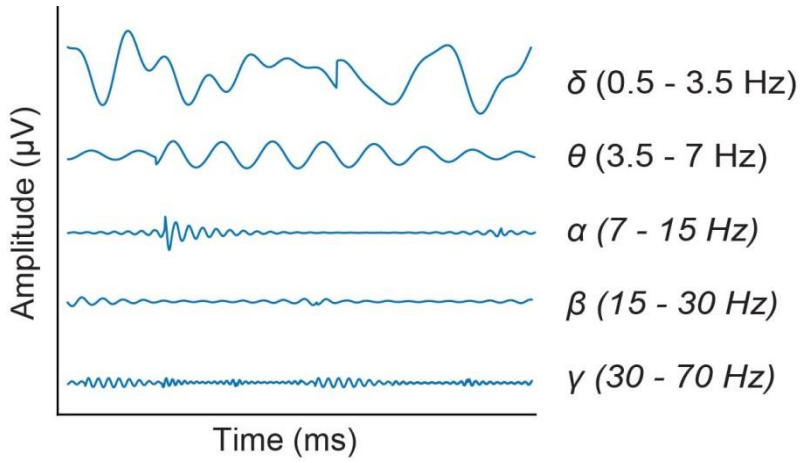


Figure 1.2. The usual frequency bands for EEG signal analysis

1.4 Event-related Potentials (ERPs)

In the EEG analysis of the diverse brain signals, a stereotyped electrophysiological response to a stimulus (whether it may result from a sensory, cognitive, or motor event) provide a robust measurement of cognitive processing between a stimulus and a response and the way brain functions might be affected by specific experimental manipulations [10], [11]. As such, these event-related potentials (ERPs) include time-locked voltage deflections of either positive or negative signal deviations, indicating so called "components" (**Figure 1.3**). Although commonly ERP components are denoted by a letter indicating polarity (N, negative / P, positive) followed by a number referring to the latency or their position in the waveform, several components (as those analyzed in this Thesis are referred to with an acronym (e.g. Error-related Negativity, ERN) [12].

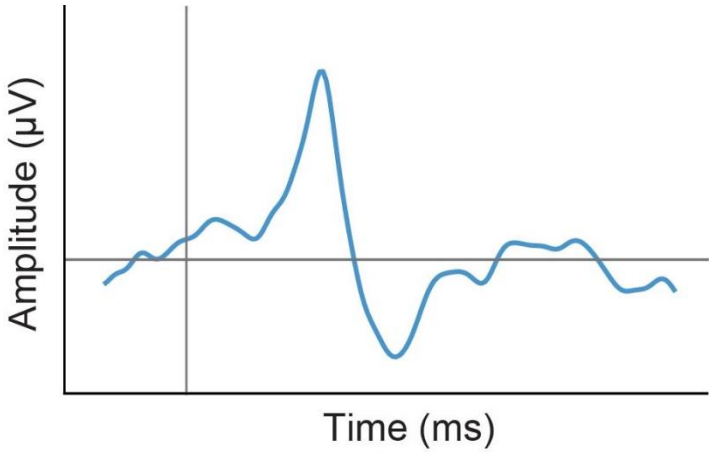


Figure 1.3. A sample ERP waveform

Due to the fact that ERPs are significantly small-sized (tens of microvolts), it usually takes a large number of trials to accurately measure it correctly. Since EEG recorded signal looks different each time (because or random variation) averaging across many trials cancels out the random variation leaving only the unchanging task-related components shown. As such, the interpretation of ERPs takes into account the assumption that the components of interest result from event-locked ERPs with invariable latency and shape and can be averaged as:

$$\bar{x}(t) = \frac{1}{N}\sum_{k=1}^{N} x_k(t) \tag{1.1}$$

where $\bar{x}$ and x_k denote the average and single (trial) ERP, N is the number of trials, k the trial number, and t is the time elapsed after the k^{th} event.

1.5 Brain Networks (Graph Theory)

1.5.1 Network Construction

Another type of brain activity related signals is related to the relationships between signals recorded by sensors in different regions (or brain areas in source space) and to construct brain networks in order to understand various cognitive mechanisms [13], [14]. In this Thesis weighted brain networks were estimated via Phase Lag Index (PLI) in sensor (Chapter 4) and source space (Chapter 5) [15]. Specifically, PLI is a two-dimensional approach to calculate the functional connectivity in terms of phase synchronization while alleviating the volume conduction limitations of other phase locking connectivity methods [16]. In detail, for a pair of channels, channel A, and channel B, where $x_A(t_n)$ and $x_B(t_n)$ denote signals in an *n* epoch that have been band-pass filtered to a predefined frequency range, let $z_A(t_n)$ and $z_B(t_n)$ be the Hilbert transform of $x_A(t_n)$ and $x_B(t)$, respectively:

$$z_A(t_n) = Z_A(t_n)e^{i\varphi_A(t_n)} \tag{1.2}$$

$$z_B(t_n) = Z_B(t_n)e^{i\varphi_B(t_n)} \tag{1.3}$$

where $Z_A(t_n)$ and $Z_B(t_n)$ indicate amplitude and $\varphi_A(t_n)$ and $\varphi_B(t_n)$ designate phases at time point t_n derived through the Hilbert transform. Then PLI for *N* epochs can be estimated as:

$$PLI_{A\to B} = \frac{1}{N}\sum_{n=1}^{N} e^{sign\left(\varphi_B(t_n)-\varphi_B(t_n)\right)} \tag{1.4}$$

PLI ranges in the interval [0, 1], with 0 indicates either no coupling or coupling with a phase difference at 0 or π and a value of 1 indicates exact phase locking with a consistent phase difference other than 0 or π. Furthermore, the PLI graphs generated are non-directional, meaning that connections between nodes are symmetrical (i.e., $PLI_{A\to B} = \ PLI_{B\to A}$).

1.5.2 Network Metrics

A network has the "small world" property if its structure is characterized by a high clustering factor and a similar characteristic path length compared to random networks [17] (**Figure 1.4**). Qualitatively, in a small world network, the majority of nodes are directly connected to most of their neighbors as well as to a few remote nodes, so that any pair of nodes is connected with a relatively short path length. There is convergent evidence that brain connectivity is characterized by a small world topology as a result of local specialization and universal integration, while disruption of its optimal topology may present signs of illness or burdened mental state [14], [18]. Therefore, in this Thesis we examined the evolution of

network topology over time through the metrics of characteristic path length (L) and its derivatives, i.e., global (E_G) and local (E_L) efficiency [19].

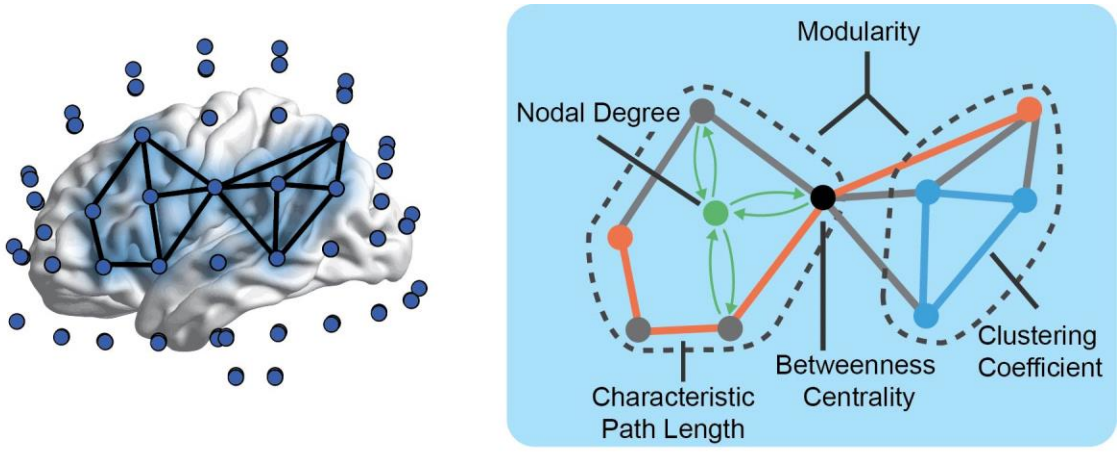


Figure 1.4. In the left corner a Schematic of a brain network modeled as a collection of nodes-representing regions of interest and interactions between brain regions. The topological properties of the different metric are presented in the right corner.

The path length of an edge in a weighted graph is defined as the reciprocal of the edge weight. The shortest path length L_{AB} between two nodes A, B is the minimum sum of the edge weights of all possible paths between the two nodes. Characteristic path length L is defined as the average of the minimum path lengths of all *N* nodes:

$$L = \frac{1}{N(N-1)} \sum_{A \neq B} L_{AB} \tag{1.5}$$

The global efficiency (E_G) of a network represents how efficiently the information is exchanged across the whole network where information is concurrently exchanged. (E_G) is the average inverse shortest path length in the network, inversely associated to the characteristic path length (L).

For a weighted brain network *G(V, E, W)* with a set of *V* nodes and *E* edges with *W* weights, where *N*=|*V*|, the global efficiency E_G of a graph is defined as:

$$E_G = \frac{1}{N(N-1)} \sum_{A,B \in V, A \neq B} \frac{1}{L_{AB}} \tag{1.6}$$

where L_{AB} denotes the shortest path length between nodes *A* and *B*.

The local efficiency (E_L) determines the network's resistance to failure on a small scale, quantifying how efficiently information is exchanged by its neighboring nodes when it is removed. The E_L is calculated on the neighborhood of each node without including it:

$$E_L = \frac{1}{N(N-1)} \sum_{k \in V} \sum_{A,B \in neibors(k), A \neq B} \frac{1}{L_{AB}} \tag{1.7}$$

Chapter 2

2.1 EEG Artifacts and Noise

The brain dynamics in the EEG recordings are highly susceptible to various forms and sources of noise presenting several challenges in the effort of analyzing and interpreting signal properties, especially when the signal is a combination of desired brain dynamics and noise. These signal contaminations (artifacts) generally originate from non-cerebral functions and may be the result of environmental factors, equipment or biological sources [20]. As such, AC power lines, lighting and a large array of electronic equipment (from computers, mobile phones etc.) usually present a systematic noise in the EEG signals, which subsequently is amplified along with cortical signals and thus is embedded in the recorded data [21]. Physiological artifacts generated from cardiac signals, muscle contraction and eye blinks are non-systematic and therefore they cannot be predicted or prevented [22].

2.2 Artifact Correction - Preprocessing

To address the artificial signal contamination issue and subsequently provide a robust EEG analysis, certain measures can be employed during recording procedures and, following the recording, as preprocessing steps. In that scope, the use of a Faraday cage, conductive housing on cables, use of fluorescent lamps and correctly grounded instruments insulate the recording room from most environmental noise, while subjects participating in EEG experiments are usually required to minimize their movements/talking and keep an optimal distance from electronic equipment while the recordings take place [23]. The foregoing actions were taken into consideration in the conduction of the experiments described in this Thesis. In addition, to effectively deal with non-systematic artifacts, a number of preprocessing strategies was utilized involving the following steps for removing irrelevant noise and facilitate subsequent analysis:

- Resampling: Although high temporal resolution is generally desirable, it has the disadvantage of delivering a large amount of data, which in turn is much slower in the subsequent processing. In this Thesis, data were downsampled by applying an anti-aliasing filter, taking into account the Nyquist rule to determine the extent to which the sampling frequency can be reduced (i.e., the sampling frequency must be at least twice the highest frequency of the analysis) [24].
- Re-referencing: To determine that electrical activity is not dependent on reference, a re-calculation of sensor values took place in the analysis of this Thesis' data, using average reference or unipolar reference based on mastoid electrodes [25]. As such, the signal of the new reference was subtracted from each EEG channel, leading the voltage at these channels to reflect the difference to the new reference, reducing systematic signal noise.
- Filtering: Each signal consists of sine waves and cosine waves with different frequencies. A digital filter is a signal processor that selectively attenuates a signal for each of its frequencies. The pass-band frequencies of the signal will pass unchanged from input to output, while the cut-off band frequencies will be completely attenuated at the filter output. In the preprocessing filtering procedures of the EEG data utilized in this Thesis, Finite Impulse Response (FIR) band-pass and Infinite Impulse Response (IRR) Chebyshev low- and high-pass filters were employed [26].

For FIR filter of *N* order, the output sequence is calculated as:

$$y[n] = \sum_{i=0}^{N} b_i x[n-i] \tag{2.1}$$

where *N* is the filter order, b_i is the impulse response of the filter, *x[n]* is the input signal, and *y[n]* is the output signal.

For IIR filter the output signal is computed as:

$$y[n] = \frac{1}{a_0}\left(\sum_{i=0}^{P} b_i x[n-i] - \sum_{j=1}^{Q} a_j y[n-j]\right) \tag{2.2}$$

where P is the feedforward filter order, Q is the feedback filter order, b_i are the feedforward filter coefficients, a_i are the feedback filter coefficients, *x[n]* is the input signal, and *y[n]* is the output signal.

The transfer function of the Chebyshev filter is defined as:

$$H(z) = \frac{b_1 + b_2 z^{-1} + \cdots + b_{N+1} z^{-N}}{a_1 + a_2 z^{-1} + \cdots + a_{N+1} z^{-N}} \tag{2.3}$$

where *z* is the z-transformation of the discrete signal and *N* is filter order,

- De-trending: Trends can cause distortion during analysis in the domains of time and frequency. As such linear trends can be calculated (e.g. via least squares method) and subtracted from the data. In this Thesis the EEG signals were de-trended by estimating the least-squares line (considered as the best fitted) and subtracting it from the data [27].
- Independent Component Analysis (ICA): ICA involves the separation of a multi-dimensional signal into components, assuming that the signals are statistically independent of each other. In this context unwanted components that correspond to artifacts based on signal characteristics can be removed (e.g. components that are highly correlated with EOG signals) [28].

  As such, ICA can separate different signal sources (such as artifacts) hypothesizing a linear mix between independent random variables. Specifically, let $S_1, S_2 \ldots S_n$ be the independent signals and noise with unknown distribution then $\mathbf{S} = (S_1, S_2 \ldots S_n)^T$ and a $n \times n$ non-singular matrix $\mathbf{W}$ (unmixing matrix, where $\mathbf{W}^{-1}$ is a mixing one), then the recorded EEG signal ($\mathbf{X}$) would be:

$$\mathbf{X} = \mathbf{W}^{-1}\mathbf{S} \tag{2.4}$$

  However, in order to apply the ICA algorithm, the signal has to be transformed in such a way that the covariance is equal to 0 and thus there is no correlation between its components (i.e., to whiten the signal). To do so, the eigenvalue decomposition ($\tilde{\mathbf{X}}$) of its covariance matrix has to be calculated as:

$$\widetilde{\mathbf{X}} = \mathbf{E}\mathbf{D}^{-\frac{1}{2}}\mathbf{E}^{T}\mathbf{X} \tag{2.5}$$

where **E** is an orthogonal matrix of eigenvectors and **D** is a diagonal matrix of eigenvalues $(\lambda_1, \lambda_2, \dots, \lambda_n)$ of the covariance matrix:

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} \tag{2.6}$$

Once the processing of signal is finished, for each component w, the values of the de-mixing matrix are updated until the algorithm has converged (i.e., product of w and its transpose is roughly equal to 1) or the maximum number of iterations has been reached. Therefore, each IC (w_k) is calculated (starting from k = 0) as:

$$repeat\ until\ w_k^T w_{k+1} \approx 1:$$

$$w_k = \frac{1}{n}\sum_{i}^{n} \mathbf{X} g(\mathbf{W}^T\mathbf{X}) - \frac{1}{n}\sum_{i}^{n} g'(\mathbf{W}^T)\mathbf{X}\mathbf{W}) \tag{2.7}$$

$$w_k = w_k - \sum_{j=1}^{k-1} (w_k^T w_j) w_j \tag{2.8}$$

$$w_k = \frac{w_k}{\|w_k\|} \tag{2.9}$$

$\mathbf{W} = [w_1, w_2, \dots]$, $g(u) = \tanh(u)$, $g'(u) = 1 - \tanh^2(u)$

In brief, ICA can be described as following:

```
Step 1: Center X by subtracting the mean
Step 2: Whiten X
Step 3: Choose a random initial value for the de-mixing matrix W
Step 4: Calculate new value for W
Step 5: Normalize W
Step 5: Check if algorithm has converged; if not go to step 4
Step 6: Take the product of W and X to get the ICs
```

Finally, the ICs that have properties related to noise are rejected (using visual inspection or more automated procedures), and the "cleared" signal is computed using the mixing matrix and the **S** matrix that has a zero column for each rejected IC.

- Baseline correction: Baseline correction is a linear operation which is employed to eliminate very low frequency voltage (quasi-DC) amplitude EEG shifts that for some reasons might occur after the stimulating event has occurred and is typically applied when the data are divided into different event-related epochs. In this context, in each EEG data segmentation the computation of the average of the points from the so-called "baseline" period (varying in each experiment in this Thesis) is subtracted from each point in the waveform [29]. Although baseline correction is not considered an explicit artifact correction procedure, it is

regarded as an essential pre-processing step to alleviate the offset influences, which are not expected to be related to the brain mechanisms under investigation, for subsequent analysis.

2.3 Source Localization

One of the major issues in neuroscience and clinical neurology is the identification of the active region sites in the human brain. However, high spatial resolution methods (such as functional Magnetic Resonance Imaging, fMRI and Positron Emission Tomography, PET) present difficulties in locating the functional activity due to the time duration and the nature of their implementation [30]. In contrast, EEG can be employed to identify the unknown distribution of the brain's electrical sources, at the time resolution of the voltage sampling process, given the potential values as measured on the surface of the head (the so-called “inverse problem”). As such, solving the inverse problem allows the direct correlation of the brain anatomy to the dynamics measured at the surface of the head, providing valuable knowledge about brain functions [31].

Solving the inverse problem is a complex process as there is no single solution, since different source distributions can cause the same surface dynamics. Therefore, the inverse problem cannot have a finite number of solutions [32]. Moreover, due to the fact that the function that describes the potential distribution is unknown and discrete values are only recorded by few electrodes (relative to the number of brain sources) the problem becomes even more ill-posed, requiring often unrealistic assumptions and constraints about the distribution of sources, their locations and their type [33]. Virtually all source localization methods require a model (volume conductor) which is the "key" to solving the direct (and inverse) problem. This model determines how the sources located in different parts of the brain cause the dynamics on the surface of the head by corresponding multiple dipoles to fixed position, variable size and orientation [34]. Different models have been proposed over the years to describe the distribution of monopolar or bipolar sources. These include head geometry to be considered as a set of spherical or elliptical cortices or based on a realistic model reconstructed from anatomical information [35]. In this Thesis a realistic model that utilizes the “low resolution brain electromagnetic tomography, LORETA” method [36] in order to simulate the electrical activities of the head to provide more accurate and reliable results.

2.3.1 Solving the Inverse Problem

To solve the inverse problem several approaches have been developed requiring no prior knowledge of the sources of electrical activity [37]. As such, the relationship between the distribution of sources in the brain and the discrete potentials presented and recorded on the scalp surface is a linear relationship expressed as:

$$\mathbf{d} = \mathbf{Ls} \tag{2.10}$$

where $\mathbf{d}$ is an one-dimensional vector array (with size $m \times 1$) of the potentials presented at each specific position of the m electrodes, $\mathbf{s}$ is an one-dimensional vector (with size $n \times 1$) of the n sources with defined direction and orientation and $\mathbf{L}$ is the dimension array ($m \times n$) known as the lead-field matrix which contains information about the geometry and conductivity of the head [38].

Each column of the lead-field matric contains information about how the potentials generated by this source are distributed to the electrodes on the surface of the head individually (forward problem) determined by the geometry of the head. The solution to the inverse problem lies in solving equation 2.10 for the unknown source distribution **s**.

Since the source distribution **s** contains more independent variables than the unknown EEG variables (i.e., $n>m$), it is not possible to exactly determine the source distribution. However, due to specific mathematical properties regarding non-zero sources and their relationship with the lead-field matrix, several solutions of the inverse problem (that satisfy the equation 2.10) can be extracted resulting in:

$$\{\mathbf{s} = \mathbf{s_1} + \mathbf{s_2} + \cdots + \mathbf{s_n}\,; \mathbf{p} = \mathbf{Ls}\} \Rightarrow \{\mathbf{Ls_1} = \mathbf{p},\ \mathbf{Ls_2} = 0, \dots,\ \mathbf{Ls_n} = 0\} \tag{2.11}$$

where **p** stands for each specific point in solution space (**d** in equation 2.10), which refers to a vector in the lead field matrix. Of note however is that sometimes the algorithm converges to single solution (with only one $\mathbf{s_n} \neq 0$), omitting other important sources.

A unique solution can be determined by combining constraints on both the solution and the data as a linear framework:

$$(\hat{\mathbf{s}} - \widehat{\mathbf{s_0}})^{\mathrm{T}}\mathbf{C_s}(\hat{\mathbf{s}} - \widehat{\mathbf{s_0}}) = min \tag{2.12}$$

and

$$(\mathbf{L}\hat{\mathbf{s}} - \mathbf{d})^{\mathrm{T}}(\mathbf{L}\hat{\mathbf{s}} - \mathbf{d}) = min \tag{2.13}$$

where $\hat{\mathbf{s}}$ the estimated solution, $\widehat{\mathbf{s_0}}$ is an a-priori approach to the solution and $\mathbf{C_s}$ is a matrix representing the metrics associated with the source space; $\mathbf{L}\hat{\mathbf{s}}$ are the predicted and d the measured data.

If $\mathbf{C_s}$ is positively defined the solution becomes:

$$\hat{\mathbf{s}} = \widehat{\mathbf{s_0}} + \mathbf{C_s^{-1}L^T}\left(\mathbf{LC_s^{-1}L^T}\right)^{-1}(\mathbf{d} - \mathbf{L}\widehat{\mathbf{s_0}}) \tag{2.14}$$

The matrix $\boldsymbol{C_s}$ can also be used to incorporate prior information about the areas of the brain where active sources are expected (i.e., where fMRI data is also available). However, if sources are expected at any location in the source space, each location is treated with the same gravity.

$$\hat{\mathbf{s}} = \mathbf{L}^{\mathrm{T}}\left(\mathbf{LC_s^{-1}L}^{\mathrm{T}}\right)^{-1}\mathbf{d} \tag{2.15}$$

In this Thesis the "exact Low Resolution Electromagnetic Tomography Activity" (eLORETA) was employed in Chapter 5, estimating the current density given by the minimum norm solution with a sophisticated regularization, which is utilizing a discrete Laplace operator that selects preferentially spread source distributions.

2.4 Machine Learning Tools

In this PhD dissertation, the focus relies on the application of advanced machine learning approaches for the processing and analysis of EEG signals, thus exploiting the signal properties in the time and/or frequency domains. In this regard, Machine Learning (ML) offers the unique advantage to produce models that can adapt to different conditions and tasks, uncovering hidden characteristics of cognitive processes, while improving the model automatically through experience and data-driven approaches [39].

Typically, ML dictates a model to be trained by the input data, while distinct separation of the non-trained data (testing data) into classes are given, based on the extracted characteristics. As a general rule ML integrates specific procedures that can be outlined as: a) feature extraction, b) feature selection and c) classification, each including multiple steps and internal algorithmic processes. Specifically:

- Feature extraction: Regarding the input characteristics (features), the data are processed in accordance to detailed rules so that representations of them can be quantitatively calculated, expressing their properties as vector values. As such, the features extracted are expected to contain relevant information in regard to the tasks/conditions applied, while illuminating the differentiation in such a manner that could lead to better human interpretations. In this context, the analysis performed in this Thesis contains features deriving from ERP signal morphological characteristics, network related analysis (in sensor and source space), and spectral density of different frequency bands. Each feature extraction procedure utilized for the experiments employed is detailed in the corresponding chapters.

- Feature selection: Feature selection is related to reducing the input data due to their large size (thus making the model susceptible to overfitting, while also augmenting computational cost), their inherited small informative nature, or if they are suspected to be redundant and therefore result in subsequent classification bias. On this premise, the full feature set can be transformed into a reduced subset [2], so that the succeeding tasks can be performed with reduced representation instead of the complete initial data. More importantly, a small feature set can facilitate in model simplification, highlighting the informative feature vectors and thus provide effective identification, generalization and interpretation of the employed attributes, relative to the task/condition applied. In this Thesis various feature selection methods were incorporated in the ML frameworks, fully described in the related chapters.

- Classification: The classification procedure refers to the training of models based on the feature set employed, so that new data can be separated into classes. Although a vast number of different classification algorithms exist, in this Thesis supervised learning algorithms were used building the corresponding models from a set of data that contains both the inputs and the desired outputs [40]. These data comprise of feature vectors (instances) and typically consist of a training set (i.e., the data utilized to build the mathematical model) and a training set (i.e., the data utilized to test the model efficiency). Supervised learning methods commonly map the training data (features) as points in a multi-dimensional space separating the different categories (labels) by a set of rules (forming a gap between them) and then predict the testing data labels on their assigned allocation. In this Thesis several classification algorithms were employed such as Support Vector Machines (SVM) with multiple kernels, k-Nearest Neighbor (k-NN), Linear Discriminant Analysis (LDA) and Random Forests (RF). In addition, ensemble classification was applied (Chapter 5), which incorporates a set of classifiers, combining their predictions for the classification of unseen instances in the form of majority voting. Of note is that although more recent approaches (such as deep learning classifiers and feature learning operators) may present higher performance than the more conventional methods applied here, these methods diffuse the information in a way that is exceedingly difficult to decipher [41]. As such, the objective of this Thesis was not only to obtain high accuracy classification, but also interpret the features and models applied as part of the underlying mechanisms that govern higher order cognitive functions and states.

2.4.1 Classification Evaluation Metrics

In order to assess the performance of the ML frameworks included in this Thesis the classification accuracy, sensitivity and sensitivity were calculated with regard to the true vs the predicted labels [42]. Accordingly:

- Classification accuracy is defined as the ratio of the number of correctly classified instances, i.e., the number of true positives plus the number of the true negatives, to the total number of instances.

$$Accuracy = \frac{\sum True\ Positives + \sum True\ Negatives}{\sum Total\ number\ of\ cases} \tag{2.16}$$

- Sensitivity is the ratio of the number of true positives, to the total number of relevant positive elements.

$$Sensitivity = \frac{\sum True\ Positives}{\sum Total\ number\ of\ Positives} \tag{2.17}$$

- Specificity is the ratio of the number of true negatives, to the total number of relevant negative elements.

$$Specificity = \frac{\sum True\ Negative}{\sum Total\ number\ of\ Negatives} \tag{2.18}$$

In addition, the area under the curve of the receiver operating characteristic (ROC) curves have been utilized to further evaluate the trade-off between classification true- and false-positive rates [43].

Chapter 3

3.1 Error-related Cognitive Monitoring

The cognitive processing of error detection is extremely important in the adaptation of the behavioral and learning processes, resulting from the cognitive evaluation of an outcome that is considered undesired or mismatches an expected response. This brain activity during correct and incorrect responses is often reflected as distinct patterns of specific ERPs that elucidate the complex cerebral responses to deviant sensory stimuli. As such, the development of accurate error detection systems can facilitate in the detection and interpretation of the cerebral responses to erroneous stimuli both regarding the complex neural mechanisms of decision making and the utilization of the resulted outputs for practical applications. In this thesis, an audio identification experiment was implemented with two levels of complexity to investigate the neurophysiological error processing mechanisms. As such, an analysis of the variations of the erroneous processing was carried out via ML procedures for each level of complexity. For a more thorough examination of the error processing on error-related components, two scientific approaches were implemented on different time windows. The first incorporating a full channel analysis in actors and observers, utilized to illuminate the different properties of the error-related ERPs, whereas the second approach, building upon the firsts results and conclusions, used specific (theoretically established by international literature) electrode positions to facilitate a cross-condition evaluation of the variations of error-related ERP components.

3.2 Introduction

Decision making is an everyday procedure, where the ability of the brain to recognize the errors that occur during the various mental operations is the key for the optimization of human behavior. Noninvasive electroencephalography (EEG) and in particular the study of the event-related potentials (ERPs) triggered when an individual performs an incorrect action or observes errors committed by others, is considered ideal to decode the complex neural mechanisms since it employs brain activity measurements with very high temporal resolution [44]. These ERPs consist of several components, elicited when correct or erroneous choices of individuals are made based on with external stimuli.

Falkestein at al. and Gehring et al. [45], [46] were the first to report a negative deflection of a response-locked ERP peaking around 100 ms after the commission of an error (error-related negativity, Ne, ERN). This has become evident as other studies [47]–[49] also presented an ERN component, peaking around 40-150 ms after the erroneous response onset. Following the ERN, a positive ERP component (error-positivity, PE), with amplitude maximum typically appearing 200-500 ms after erroneous responses, has been reported to reflect error awareness [50]–[54], while a time-locked negative ERP have been detected peaking approximately 250-300 ms when feedback is provided on erroneous actions (feedback related negativity, fERN) [55]–[60]. The components related to error monitoring are not limited only with regard to deviant outcomes. Several studies have identified a component (similar to ERN in terms of latency and morphology), elicited after correct trials (correct-related negativity, CRN), theorized as a possible cognitive mechanism to prevent errors from happening [61]–[64]. As a means to unveil the underlying processes of error cognition, ample evidence suggest that the various ERPs stated above (a term that will

subsequently be utilized as error-related ERPs, ErrPs), are similar in term of topology and present mostly central, centro-parietal and fronto-central distributions [65]–[69]. Furthermore, functional Magnetic Resonance Imaging (fMRI) and source localization studies have identified the anterior cingulate cortex as the generator of error processing, thus making the ErrPs' negative and positive deflections be more apparent in midline scalp positions [70]–[73].

The robustness of the ErrPs has been significant factor for the effective analysis of the error-related EEG signals providing the basis for the theoretical background utilized in numerous multivariate studies. Most of these studies incorporate ML methods to recognize distinguishable patterns in the recorded EEG signals, applying algorithmic models to represent signal features as points in space, mapping and classifying the error-related brain activity as correct or incorrect [74]–[77]. The ML models can thus enable the complete parameterization of the signal characteristics, allowing for future real life purposes. For instance, Ventouras et. al., [78] achieved high classification accuracy by comparing different feature selection methods, while utilizing 2nd-order-statistical features and various time windows in an SVM and a k-NN ML framework. Plewan et. al., 2016 [79] successfully classified ERPs between subjects by employing Independent Components as features and a radial basis function SVM, on a modified flanker and a mental rotation task. Nevertheless, when they applied the same method in cross-task classification, accuracy result deteriorated. Moreover, Spüler et. al. 2015 [77] performed continuous feedback EEG classification on different type of severity errors, utilizing frequency and time- locked ERP features, suggesting that difference in classification performance can be attributed to the task complexity. This suggestion is in line with research indicating that error cognition is a multifactorial neurological process which depends on a large number of conditions such as workload, psychological/ emotional status, attention, etc. As such, ErrPs can present large fluctuations in terms of amplitude and latency between different individuals, tasks and difficulty levels [80]–[83]. More importantly, the morphology and other ErrPs characteristics exhibit significant variations as a result of a wide range of conditions including intent, motivation, substance abuse and age [84]–[86], while psychological conditions and anxiety have also been reported as a major influence affecting ErrP attribute [87]–[93]. Iturrate et al. [94] investigated the task condition effect in ErrPs by classification training and testing under different task conditions, reporting reduced classification accuracy due to signification variations on the extracted EEG features. Furthermore, Van der Borght et al. [95] found significant decrements in different ErrPs (ERN, CRN and in the early PE), due to complexity increments of a 2-condition flanker task. In a similar manner, Endrass et al. [96] found decreased ERN and CRN in the highest level of difficulty of a visual size discrimination task with three difficulty conditions, applying Principal Component Analysis.

Taking the above into consideration, it can be inferred that due to high morphological sensitivity of the ErrPs, the error processing mechanisms might be masked by the complexity of the task employed. As such, ML methods can unveil the hidden attributes reflecting error processing, leading to the identification of global condition-independent ErrPs characteristics and thus allowing the detection of incorrect decisions, irrespective of the difficulty of a task. This in turn will enable to improve the classification performance for the different levels of complexity and provide indications for condition-adjusted cognitive mechanisms. To address this, an EEG auditory identification experiment with two conditions of complexity was implemented with the resulting data analyzed under the premise that error processing is a dynamic phenomenon, with distinct ErrP morphological attributes.
The analysis was divided in two discrete phases including ML approaches on time-windowed morphological features. The first approach took into account a full channel feature set in actors and

observers along with a modified Sequential Forward Selection (scoring SFS) to minimize the frequently occurring nesting effects, while the second utilized a feature selection framework (in the significant channels in terms of ErrPs) based on the combination of Sequential Forward Floating Selection (SFFS) and Sequential Forward Selection (SFS) to facilitate cross-condition and within-condition classifications in actor responses. Both analyses were able to provide high classification performance, whereas the implications of each are discussed in detail below.

3.3 Materials and Methods

3.3.1 Participants

The EEG data were collected from 14 healthy individuals (8/6 male/female) with average age 26.6 ± 2.9 years and high level of education (17.7 ± 2.3 years). All participants were right-handed and reported no history of any neurological diseases or drug intake. Furthermore, prior to the recordings the subjects' normal hearing ability was estimated by a pure-tone audiogram (thresholds <15 dB HL), while informed consent was obtained from all participants. The present study was conducted according to the declaration of Helsinki and approved by the university's ethics committee.

3.3.2 Experimental Design

The participants were requested to undertake an auditory identification experiments in two conditions of complexity as actor or observers as detailed below (**Figure 3.1**). Both sessions were performed on the same date and required the determination of the specific frequencies corresponding to the acoustic stimuli. Initially, subjects were divided into 7 dyads, sat opposite from each other and had computer screens in front of them, displaying a slider and a cursor, effectively screening them. Each condition consisted of 80 trials for each dyad, whereas dyad members alternated roles as actor or observer from one trial to the next (resulting in 40 trials for each subject when participating as an actor).

At the beginning of each trial (operating phase), both individuals were presented a stimulus tone of 1 sec duration through headphones. The tone of the stimulus was randomly selected from a fixed frequency range with a bandwidth of 400 Hz (within a block of trials) represented by the slider bar, while the position of the slider corresponded to a specific tone. Four frequency ranges were used for the auditory stimuli, specifically 200-600 Hz, 620-1020 Hz, 1040-1440 Hz and 1460-1860 Hz. Then, the actor was asked to match the frequency of the stimulus tone by positioning a cursor in a slider, appearing in both participants' computer screen, with the use of a gamepad. The slider represented the fixed frequency range within a block of trials, with the position of the cursor corresponding to a specific tone within this range. Individuals were unaware of the frequency range scale in which they had to place the cursor, while during the handling of the gamepad neither the actor nor the observer could hear the sound corresponding to the position chosen. The non-movement of the gamepad for 0.5 sec signaled the end of the operating phase. Following the operating phase, both dyad members were asked to assess the correctness of the position chosen by the actor participant utilizing a two-button controller corresponding to correct or incorrect estimation. After the first judgment, the tone corresponding to the position chosen by the actor was presented to dyad members (feedback tone, FBT), who were requested to evaluate if the FBT was the same as the primary tone. The FBT judgment was then followed by a knowledge-of-results (KOR) tone,

for each individual indicating, whether the position chosen was correct (as a 500 Hz tone) or erroneous (as a 3 kHz tone) along with the announcement of the words "correct" or "incorrect" respectively.

The difference of the complexity of the conditions was based on whether the frequency ranges of the acoustic stimuli were the same ("easy" condition, *Joint1*) or different ("difficult" condition, *Joint2*) as the participants switched roles as actors and observers between trials. Specifically, in the first condition (Joint1) acoustic stimuli of the same frequency range were presented to both dyad members, whereas in the second condition (Joint2) the stimulus presented to each participant as an actor differed from the stimulus presented to his/her partner (when the partner was the actor) in terms of the fixed frequency range. On this premise, it might be assumed that actors in *Joint1* condition could match the stimulus sound more easily than in the *Joint2* condition, since they could effectively map the frequency range of the slider bar mentally by observing their partner being actor in the previous trial. In contrast, in *Joint2* condition, participants could not use the mental map of the slider frequency range bar by observing the actor in previous trials, while this procedure could mentally disorient them when they were requested to assign the cursor to the stimulus tone as actors on the next trial. As such, the *Joint2* condition would be significantly more to challenging in the designation of the correct position (within the frequency range employed) hindering the identification process.

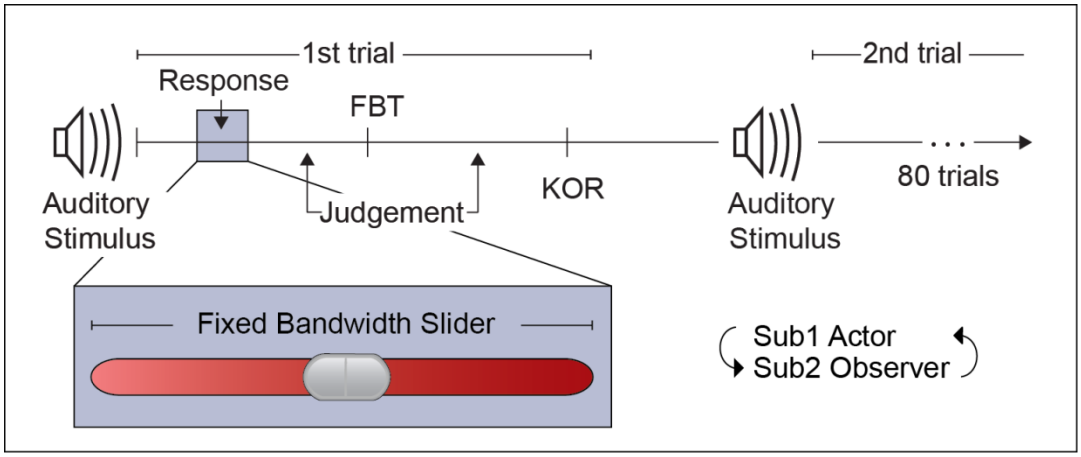


Figure 3.1. The experimental design. In each trial both subjects heard the same auditory tone. A) In *Joint1* condition the stimuli was presented from the same frequency range in all trials. B) In *Joint2* condition the stimuli presented were of different frequency ranges in successive trials.

Based on existing literature, the aforementioned experimental protocol included two feedback responses, the FBT and KOR. As such, the FBT can be regarded as a first-level feedback with indirect information being provided to the individual after each response, while the second-level feedback, i.e., the KOR, allows clear evidence concerning the correctness of the auditory identification. Taking this into account, only the FBT ERPs were investigated in the following analysis, since the first feedback should trigger a cognitive response closer in terms of temporal proximity than the second feedback.

Prior to the experiment each subject performed an acoustic pre-test to assess his/her hearing ability in the four frequency ranges used. This included identifying the higher of two tones. The frequencies of the two tones were set to the 25% and the 75% of each range of 400 Hz bandwidth employed. All participants were able to successfully discriminate between the tones presented

3.3.3 Data Acquisition and pre-processing

The experimental setup involved a Faraday room, optical receiver for trigger inputs and electrode bundled cables to eliminate potential electric and magnetic interference. The bioelectrical brain activity recording via EEG occurred simultaneously for both participants of each dyad (actor and observer), using two different recording systems, daisy-chained in a master-slave relationship. Specifically, each system included a 32-channel electrode cap (Biosemi, Activetwo System) according to the International 10–20 EEG system, with Fp1, AF3, F7, F3, FC1, FC5, T7, C3, CP1, CP5, P7, P3, Pz, PO3, O1, Oz, O2, PO4, P4, P8, CP6, CP2, C4, T8, FC6, FC2, F4, F8, AF4, Fp2, Fz and Cz (**Figure 3.2**) electrode positions. Horizontal and vertical electrooculograms were recorded from electrodes placed above and below the eyes and at the outer canthi of the eyes. EEG data were digitized at 256 Hz, re-referenced to the average of the electrode recordings and filtered off-line by applying an IIR low-pass and a high-pass Chebyshev filter with cut-off frequencies 35 Hz and 0.05 Hz respectively. Subsequently, the EEG signals were de-trended and segmented into 2.5 sec ERP epochs (0.5 sec before and 2 sec after the FBT), resulting in 40 × 14 × 2 = 1,120 trials acquired from the total of 40 trials for the 2 complexity conditions of the 14 participants and for actors and observers. After segmentation, each ERP was baseline-corrected relative to a 100 ms pre-stimulus baseline, while the trials contaminated with artificial ocular noise were manually removed. The EEG data of each trial were baseline-adjusted relative to a 100 ms pre-stimulus baseline. Due to significant artifact contamination, data from one dyad were excluded leaving 12 subjects for subsequent processing. More details for each of the pre-processing steps employed can be found in Chapter 2.

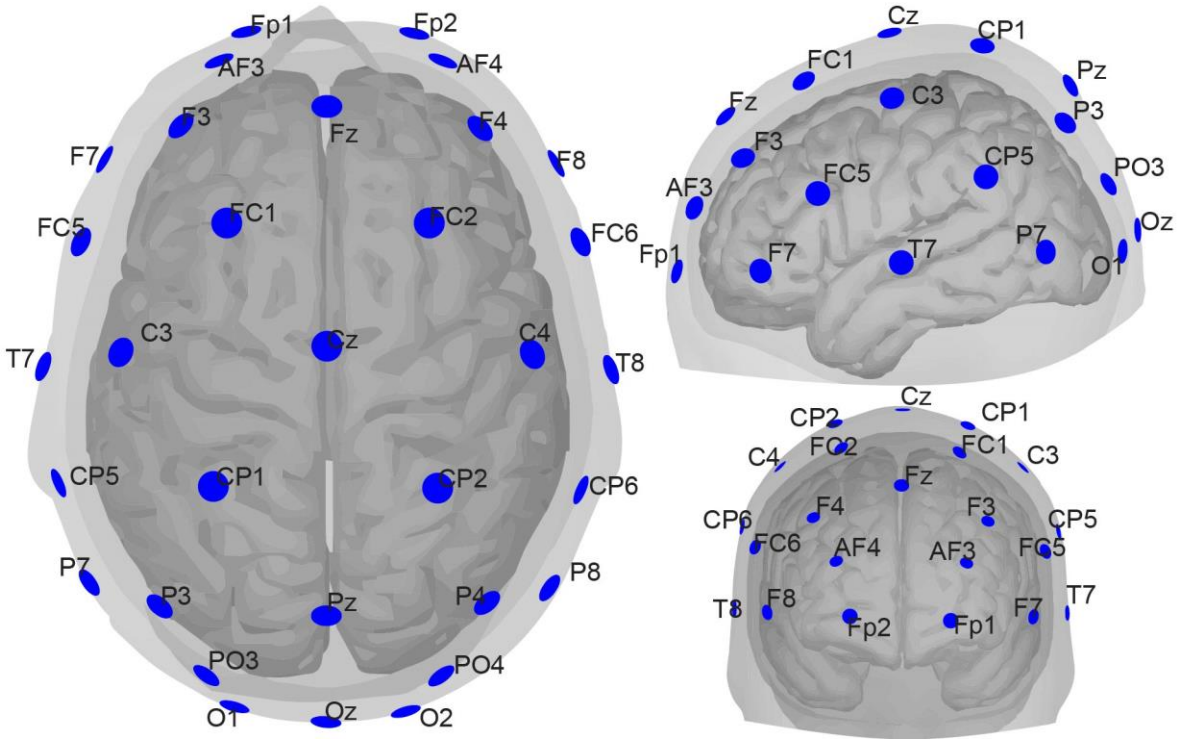


Figure 3.2. The electrode locations for the EEG acquisition according to the International 10–20 system.

3.3.4 Definition of the Correctness of Participants' Responses

To determine the ability of each participant to differentiate tones and his/her auditory frequency perception resolution, the psychoacoustic theory was taken into account. Specifically, due to the fact that two tones can be misinterpreted as the same sound if they are close in terms of frequency, similar stimuli tone and FBT may not trigger error-related cognitive reaction. As such the proximity of the FBT and stimuli tones was utilized as an indicator for correct or incorrect responses, quantified via the psychoacoustic function of Equivalent Rectangular Bandwidth (ERB) [97]. This function cites that the ability of each individual to perceive and distinguish different tones, is determined by a function of a central frequency by modeling the human hearing filters as rectangular band-pass filters, thus approximating the frequency range in which the auditory stimuli are considered the same. As such, whether individual responses were considered correct or erroneous was calculated according to the following formula:

$$B_e = 6.23\ 10^{-6} f^2 + 9.339\ 10^{-2} f + 28.52 \qquad 3.1$$

where B_e is the bandwidth of the filter in Hz and f is the central frequency (presented as the stimulus tone) of the filter in Hz.

Despite the fact that the ERB appears to be linear in low frequencies, the function is of non-linear nature (**Figure 3.3B**). As such, a pre-define criterion (e.g. such as the ratio f/Be) to denote the correct or erroneous responses could in fact render the distinction between them. In this regard, the participant's response was compared to the stimulus tone plus/minus the ERB bandwidth in each trial, estimating if the position selected in the operating phase was within this range to be considered correct or erroneous (Fig.3.3A). Although, no participant was recorded as having all-correct or all-incorrect responses, due to the sensitivity of ErrPs in terms of subject variability, as well as the dissimilarity of the number of correct and incorrect ERPs responses (per participant and condition), the resulting ERPs were averaged in each condition per subject and class (correct/incorrect). In detail, the mean ERPs were calculated for the two conditions for each of the 32 electrode positions and each participant on the basis of his/her response (with the classes being correct or incorrect). In this regard, the classes for the subsequent feature calculation and classification processes would be balanced alleviating classification bias. From the available recordings of each *Joint* condition, 12 × 32 × 2 = 768 ERP recordings corresponded to correct responses and the rest 12 × 32 × 2 = 768 recordings corresponded to incorrect responses for actors and observers.

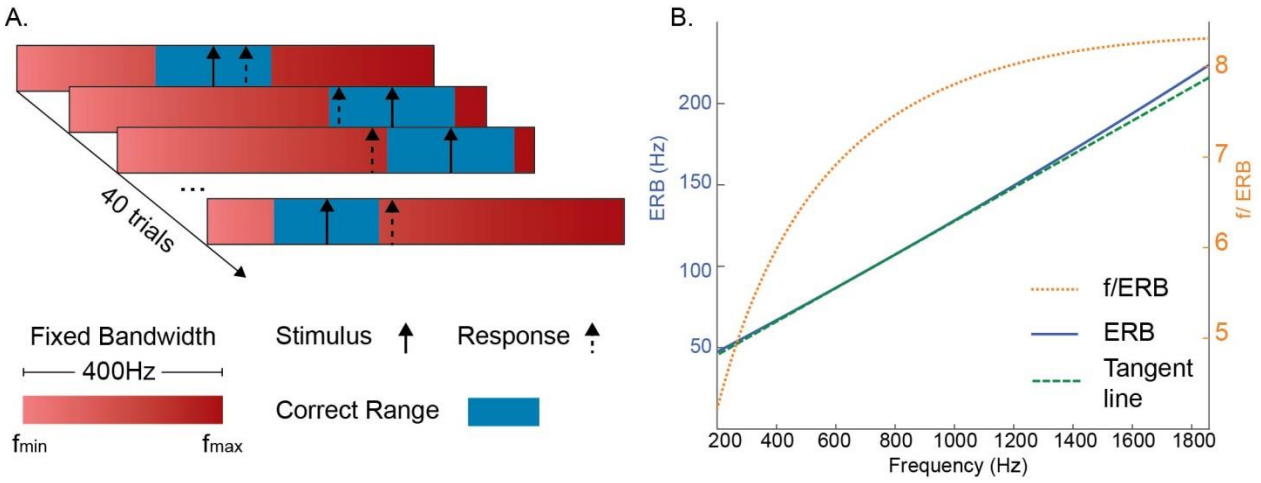


Figure 3.3. Definition of responses using the Equivalent Rectangular Bandwidth (ERB) with regard to the stimulus tone. In A) the solid line arrow represent the stimulus on each trial the stimulus randomly selected from the fixed band frequency (gradient bar). The response is indicated by a dashed line arrow, denoted as correct or erroneous based on whether or not within the ERB range (solid filled bar). In B) the ERB function as a result of stimulus frequency.

3.3.5 Feature Calculation

To perform the classification in the two conditions a series of feature extraction, combination and selection processes took place for each of the two discrete approaches. In both frameworks, features were calculated based on the visual observation of the averaged ERP curves of the different electrodes for different time windows, where the ErrP components show distinct morphology and clear differentiation between the two groups of correct and incorrect answers (as indicated by the literature presented in the introduction section).

As such, the correct and incorrect ERPs were further segmented into 5 time windows, starting after the presentation of the FBT (0ms). Specifically:

- time window 1 (tw_1) starting at 0 ms and ending at 125 ms after response
- time window 2 (tw_2) starting at 125 ms and ending at 220 ms after response
- time window 3 (tw_3) starting at 22 0ms and ending at 300 ms after response
- time window 4 (tw_4) starting at 300 ms and ending at 400 ms after response and
- time window 5 (tw_5) which included the previous as well as an additional 200 ms, starting at 0 ms and ending at 600 ms

Each time window was considered to be indicative of specific ErrP component, while the inclusion of the after-FBT stimulus ERP (tw_5) could further demonstrate useful information that might could be masked when calculating the separate (small-duration) time windows features.

The features were extracted for of the time windows and each electrode position, based on latency and shape characteristics describing ErrPs [98], [99] (**Figure 3.4**) calculated as:

- MaxA: the maximum amplitude value of the ERP signal for each time window

- MinA: the minimum amplitude value of the ERP signal for each time window
- MaxT: the latency of the maximum time windowed signal value, corresponding to the time MaxA occurred
- MinT: the latency of the minimum time windowed signal value, corresponding to the time MinA occurred
- AUC: the area under the ERP curve, representing the overall signal energy, estimated by calculating the corresponding time window ERP integral.

Of note is that no normalization took place despite the different type and value range of features. This was due to the fact that further BCI expansion could not (in effect) perform normalization online and thus the results of this study would not be applicable. Hence, from each averaged ERP 5 features were calculated for each of the 5 time windows and each of the 32 electrode positions, resulting in 32 × 5 × 5 = 800 features in total for each individual. This process was done for actors and observers, for *Joint1* and *Joint 2*.

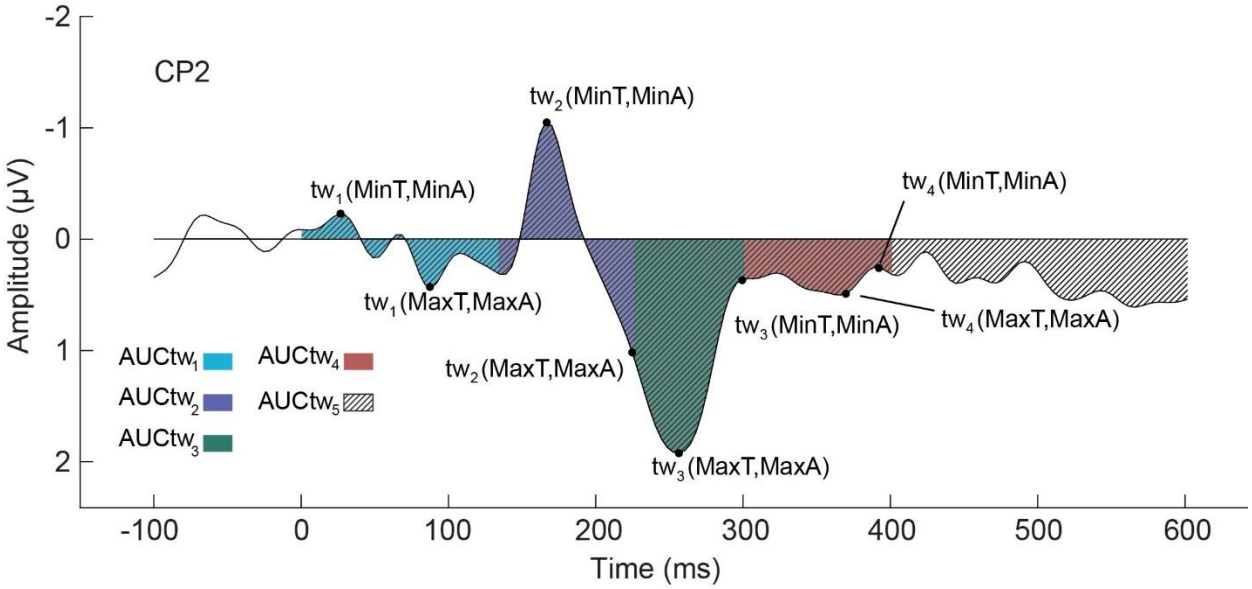


Figure 3.4. The features extracted for the different time windows of electrode CP2. For each time window the latency and amplitude of the minimum and maximum values as well as the area under the ERP curve were estimated. Time window 5 includes all time windows with global latency and amplitude while the stripped pattern denotes the area under the ERP curve.

3.4 Behavioral Results

As mentioned before, in *Joint1* condition participants received tones from the same frequency range and thus the slider bar represented an identical range in both of them. On the other hand in *Joint2* condition each participant was designated to a different frequency range and so the auditory stimuli while being actor was alternating between them. As such, the dissimilarity of the frequency tone presented in each condition (*Joint1* and *Joint2*) would increase the complexity of the task making it more challenging to identify the correct position of the slider bar. This was due to confusion created by difficulty to mentally map the different slider bars in *Joint2* condition, while confusing orientation of the individual frequency range slider when each subject would operate again as an actor. To estimate whether this was the case a

behavioral analysis took place by implementing a one-way Analysis of Variance (ANOVA) on the number of the correct responses of the participants.

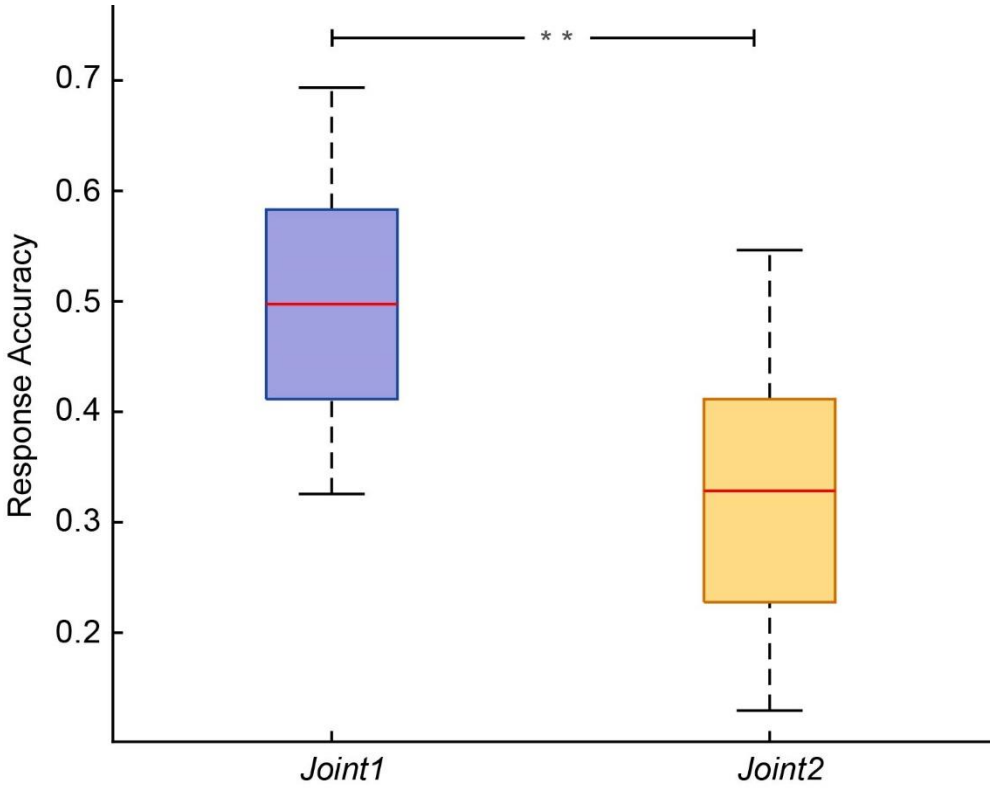


Figure 3.5. Accuracy of the actors' responses given for *Joint1* and *Joint2* condition.

The responses given in each trial was assessed by the correctness of the position of the slider bar via ERB calculation. In this regard, the KOR was deemed irrelevant and thus since the actors were solely able to manipulate the slider, observers behavioral results could not be recorded. In **Figure 3.5** the accuracy of the responses given by the actors for the two *Joint* conditions is presented. The ANOVA results indicated that individuals' performance was significantly ($p < 0.01$) affected by the complexity of the tasks. As such, the behavioral results verify that *Joint2* condition was considered more difficult than *Joint1* condition.

3.5 Full Channel ML Analysis in Actor and Observer ErrPs

As mentioned in the previous sections the experimental data were analyzed in two distinct approaches. In the first approach a full channel feature set in actors and observers was utilized in a ML framework that incorporated Feature Selection (FS) and classification procedures. The rationale behind this was the consideration that the classification performance is not only prone to the distinctness (and reliability) of the features extracted, but also affected by redundant or irrelevant features (in terms of classification) generally included in large feature sets). On this premise, the FS method applied would facilitate in problem optimization by reducing variance and thus improve the performance of the classifiers. Specifically we employed a modified Sequential Forward Selection (scoring SFS) and Support Vector Machines (SVM) classifiers with different kernels to detect the most prominent features/electrode channels with regard to the time windows and morphological features of the ErrPs.

3.5.1 Feature Selection and Classification

Since the number of features was very large ($5 \times 5 \times 32 = 800$, 5 features for each of the 5 time windows and each of the 32 electrode positions) in comparison to the instances for the subsequent classification (24 instances, 12 correct and 12 incorrect), the Sequential Forward Selection (SFS) was applied [7]. By doing so, non-useful information would be removed from the classification processes while the most indicative features, in terms of classification accuracy, would be detected.

In general, SFS starts with an empty feature set and adds one-by-one the feature with the highest significance (i.e., classification accuracy) until there can be no further improvement. In detail, the features included in the full feature set are estimated based on their individual accuracy by implementing classification on each. Next, the SFS algorithm adds another feature in the subset and estimates the classification performance. Subsequently all the features are used and the algorithm stores the subset (now including 2 features) with the highest accuracy, provided it is higher than the 1-feature subset accuracy. This step is repeated until no accuracy increment can be made. However, the fact that in each step the SFS has to estimate the classification accuracy of all features, results in large computational cost, while the greedy nature of the algorithm make it prone to nesting effects. This means that due to the algorithm including the most significant feature irrespective of other combinations, it is often trapped in local maxima. For this reason, we modified the standard SFS (indicated as scoring SFS, in the present thesis) as follows (**Figure 3.6**). Specifically, all features are sorted by their overall significance by the resulting classification accuracy. Then, our method results to a scoring process, setting the base subset instead of empty to the one that present the best performance through the exhaustive search combination of the 10 most significant features (arbitrarily selected for optimal computational cost). The second phase incorporates the next most significant feature to the subset if performance increases (disregarding the increment amount), while if the added feature deteriorates or didn't change overall performance it would be skipped and the next one would be added. When no further progress in the classification accuracy could be made the algorithm would stop.

In this approach, the FS and the classification accuracy results are entangled and thus optimal classification entails the feature subset resulted from the SFS method. To evaluate the classification accuracy in each iteration, a leave-one-out cross-validation procedure was implemented due to the limited data available, involving the employment of a single instance as the testing set, while considering the

training set to be the remaining data. The classifiers employed were SVM with linear ($K(\vec{x},\vec{z}) = (\vec{x}^T\vec{z})$), radial basis function (RBF) ($K(\vec{x},\vec{z}) = e^{-\gamma\|\vec{x}-\vec{z}\|^2},\ \gamma = 0.5$) and quadratic kernels ($K(\vec{x},\vec{z}) = (c + \vec{x}^T\vec{z})^d$, $c = 1,\ d = 2$), paired to Sequential Minimal Optimization (SMO) and Least Squares (LS) learning methods, while the soft margin regularization parameter C was set to 1 for all ML procedures [100].

Due to the limited number of instances available for classification, further cross-validation evaluation splitting of the data or bootstrapping was considered impractical and thus 1,000 runs of permutation tests on random labels were performed to ensure no selection bias or overfitting existed in the FS and classification processes. Permutation p-value was calculated as the ratio of classifiers trained on the randomized labels that outperformed the classifiers trained on the original samples to the number of total permutations and is presented in a parenthesis in the results tables. All algorithms were implemented to discriminate between correct and incorrect responses for actors and observer in MATLAB R2015b (MathWorks Inc., USA).

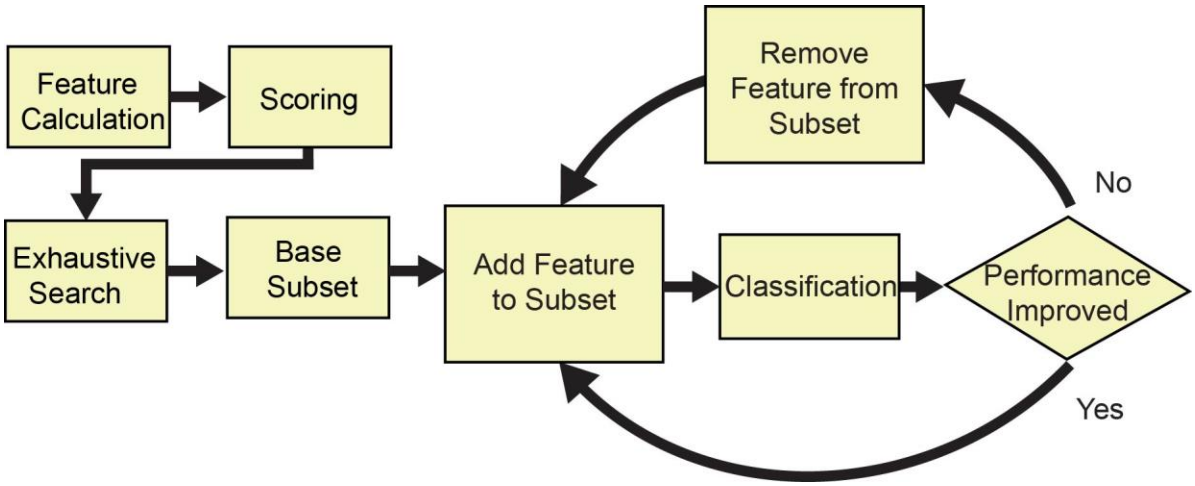


Figure 3.6. The flowchart of the scoring SFS method

3.5.2 Results and Implications

The classification results from the standard and the scoring SFS are presented in TABLES 3.1 and 3.2 accordingly.

TABLE 3.1. STANDARD SFS CLASSIFCATION RESULTS

	Kernel	SMO Accuracy	LS Accuracy
***Joint1* Actor**	linear	0.85 (p<0.01)	0.90 (p<0.001)
	RBF	0.91 (p<0.001)	0.87 (p<0.001)
	quadratic	0.83 (p<0.001)	0.92 (p<0.001)
***Joint2* Actor**	linear	0.87 (p<0.010)	0.85 (p<0.01)
	RBF	0.94 (p<0.001)	0.96 (p<0.001)
	quadratic	0.97 (p<0.001)	1 (p<0.001)
Mean Actor Accuracy		0.91	
***Joint1* Observer**	linear	0.83 (p<0.001)	0.75 (p=0.002)
	RBF	0.85 (p<0.001)	0.86 (p<0.001)
	quadratic	0.97 (p<0.001)	1 (p<0.001)
***Joint2* Observer**	linear	0.87 (p<0.001)	0.82 (p<0.001)
	RBF	0.91 (p<0.001)	0.87 (p<0.001)
	quadratic	0.95 (p<0.001)	1 (p<0.001)
Mean Observer Accuracy		0.89	

Note: Permutation p-values are presented in the parenthesis after classification accuracy.

.

TABLE 3.2. SCORING SFS CLASSIFICATION RESUTLS

	Kernel	SMO Accuracy	LS Accuracy
***Joint1* Actor**	linear	0.88 (p<0.001)	0.96 (p<0.001)
	RBF	0.88 (p<0.001)	0.88 (p<0.001)
	quadratic	0.92 (p<0.001)	1 (p<0.001)
***Joint2* Actor**	linear	0.88 (p=0.001)	0.92 (p<0.001)
	RBF	0.96 (p<0.001)	0.96 (p<0.001)
	quadratic	1 (p<0.001)	1 (p<0.001)
Mean Actor Accuracy		0.93	
***Joint1* Observer**	linear	0.83 (p<0.001)	0.79 (p <0.01)
	RBF	0.92 (p<0.001)	0.92 (p<0.001)
	quadratic	0.92 (p<0.001)	1 (p<0.001)
***Joint2* Observer**	linear	0.96 (p<0.001)	0.92 (p<0.001)
	RBF	0.92 (p<0.001)	0.92 (p<0.001)
	quadratic	1 (p<0.001)	0.88 (p<0.001)
Mean Observer Accuracy		0.91	

Note: Permutation p-values are presented in the parenthesis after classification accuracy

In general, both FS methods achieved high overall performance with the mean classification accuracy of the standard SFS being 0.91 for actors and 0.89 for observers, while the scoring SFS marginally exceeded that with mean accuracy 0.93 for actors and 0.91 for observers. In addition, the low p-values of the permutation tests employed indicate that both FS methods successfully selected discriminative features with no over-fitting. Of note is that the best overall accuracy was obtained with the quadratic SVM kernel despite the FS method employed, for actors and observers of both *Joint* conditions.

Concerning the features selected the standard and scoring SFS method feature subsets that provide the overall highest accuracy are presented below (TABLES 3.3 and 3.4).

TABLE 3.3. STANDARD SFS FEATURES SELECTED

	Kernel	SMO Accuracy	LS Accuracy
***Joint1* Actor**	linear	CP5_ MaxT_*tw2*, FC1_ MaxA_*tw5*	CP5_ MaxT_*tw2*, FC2_MaxA_*tw2*
	RBF	C3_ MaxA_*tw2*, C3_ MaxA_*tw1*, C4_MinT_*tw1*, CP1 _MinT_*tw3*	FC5_AUC_*tw5*, FC1_ MaxT_*tw5*
	quadratic	C3_MaxA_ *tw2*, Fp2 _ AUC_*tw1*	Cz_AUC_*tw3*, Fp2_MinA_*tw5*, CP2_ AUC_*tw1*
***Joint2* Actor**	linear	P8_AUC_*tw5*, Fp2_ MinT_*tw1*	T7_MaxT_*tw1*
	RBF	P3_MaxA_*tw1*, CP1 _AUC_*tw3*	P3_MaxA_*tw1*, FC1 _AUC_*tw1*, PO3_MinA_*tw1*
	quadratic	P3_MaxA_*tw1*, T7_AUC_*tw1*	P3_MaxA_*tw1*, F8_MaxA_*tw1*, O2_MinT_*tw4*
***Joint1* Observer**	linear	CP5 _MinA_*tw1*, C3_MaxT_*tw1*, CP1 _MinT_*tw5*	Fp1_MaxT_*tw1*, P8_MinT_*tw*
	RBF	FC5_MinA_*tw1*, FC5_AUC_*tw1*	P3_AUC_*tw*, T7_MaxT_*tw1*
	quadratic	FC5_MinA_*tw1*, Fp2 _AUC_*tw1*, FC1_MaxT_*tw1*	P3_MaxA_*tw4*, P4_MaxT_*tw4*, C4_MaxA_*tw1*, CP5_MinA_*tw1*
***Joint2* Observer**	linear	Cz_AUC_*tw3*, PO4_MaxT_*tw2*	PO4_MaxT_*tw1*, F3_MinA_*tw1*
	RBF	C3_MinT_*tw1*, Cz_MaxT_*tw4*	C3_MinT_*tw1*, Oz_MaxT_*tw1*
	quadratic	C3_MinT_*tw1*, O2_MaxT_*tw2*, Oz_MinA_*tw3*	C3_MinT_*tw1*, F7_MinT_*tw1*, P8_MaxA_*tw1*, C3_AUC_*tw1*

Note: Each feature is denoted including the electrode from which the feature was extracted, followed by the feature itself and the time window employed, i.e. “electrode_feature_time-window”.

TABLE 3.4. SCORING SFS FEATURES SELECTED

	Kernel	SMO Accuracy	LS Accuracy
***Joint1* Actor**	linear	Fz_MaxT_*tw2*, PO3_MaxA_*tw4*, F8_MinT_*tw4*, CP5_MaxT_*tw2*, Fp1_MinT_*tw4*	P3_MaxT_*tw2*, FC6_MinT_*tw2*, F4_MaxT_*tw2*, Fz_MaxT_*tw2* Fz_MinT_*tw3*, Fp1_MaxT_*tw1*, O1_MinA_*tw5*, P4_MaxT_*tw1*, F7_MinA_*tw2*
	RBF	F8_MaxT_*tw1*, FC1_MaxA_*tw2*, P3_MinT_*tw2*, T8_MaxT_*tw4*, CP5_MinT_*tw1*, CP2_MaxT_*tw1*	Oz_MinT_*tw4*, FC6_MinT_*tw3*, F8_AUC_*tw5*
	quadratic	F8_MaxT_*tw1*, F4_MaxT_*tw2*, Fz_MaxT_*tw2* P8_MaxA_*tw1*, FC1_MinT_*tw3*	F8_MaxT_*tw1*, CP1_MinA_*tw3*, Fp1_MaxT_*tw4*, Cz_AUC_*tw3*, CP6_MaxT_*tw4*
***Joint2* Actor**	linear	P8_MinA_*tw4*, FC6_MinT_*tw3*, P4_AUC_*tw4*	P8_AUC_*tw5*, AF4_AUC_*tw1*, F3_MinT_*tw2*, PO3_MinT_*tw3*, FC5_MinT_*tw2*, Cz_MaxA_*tw2*
	RBF	CP6_MinT_*tw4*, P8_AUC_*tw5*, T7_MaxT_*tw1*, C3_AUC_*tw5*	FC1_MinT_*tw2*, CP6_MinT_*tw4*, P8_AUC_*tw5*, FC1_MinT_*tw3*
	quadratic	P8_AUC_*tw4*, FC1_MinT_*tw2*, FC6_MinT_*tw3*, P3_MaxA_*tw1*, CP5_AUC_*tw3*	FC1_MinT_*tw2*, P8_MinA_*tw4*, CP6_MinT_*tw4*, P8_AUC_*tw5*
***Joint1* Observer**	linear	CP2_MinT_*tw1*, FC6_MaxA_*tw1*, FC2_MaxA_*tw4*,	P7_AUC_*tw3*, P8_MinT_*tw5*
	RBF	Fp1_MaxA_*tw1*, FC6_AUC_*tw1*, P8_MinT_*tw1*, P8_AUC_*tw2*	CP1_AUC_*tw1*, CP1_MinA_*tw4*, T7_MaxT_*tw1*, C3_MaxT_*tw1*, CP6_MaxT_*tw5*
	quadratic	Fp2_AUC_*tw1*, AF3_MaxA_*tw2*, C3_MinA_*tw2*, F8_AUC_*tw4*	Pz_AUC_*tw1*, CP2_MinT_*tw1*, C3_MinA_*tw2*, O2_MinT_*tw2*, FC6_AUC_*tw1*, C4_MinT_*tw3*, F4_MinT_*tw2*
***Joint2* Observer**	linear	CP6_MaxA_*tw1*, T8_MinA_*tw1*, FC5_MinA_*tw2*, O1_MinT_*tw2*, Fz_AUC_*tw1*	P8_MaxT_*tw1*, T8_MAX_*tw1*, T8_MaxT_*tw1*, F7_AUC_*tw4*
	RBF	FC1_MinT_*tw1*, P8_MaxT_*tw1*, F8_MaxT_*tw1*	P8_MaxT_*tw1*, Cz_MinT_*tw1*, C3_MinT_*tw1*, FC2_MinT_*tw1*, O2_MaxT_*tw3*_
	quadratic	F8_MaxT_*tw1*, F7_MinT_*tw4*, CP5_MaxA_*tw4*, P4_AUC_*tw4*, C3_MinT_*tw2*	FC1_MinT_*tw1*, P8_MaxT_*tw1*, Cz_MaxT_*tw4*

Note: Each feature is denoted including the electrode from which the feature was extracted, followed by the feature itself and the time window employed, i.e. “electrode_feature_time-window”.

In the standard SFS algorithm (although no clear trend can be detected from specific electrodes) positions central and frontocentral locations were the most prominent, with *tw1* being the most frequent utilized. Similarly, *tw1* was the most common in the scoring FS with P8 electrode position to be the most prominent. However, subsequent analysis in the P8 electrode position exhibited little similarity between *Joint1* and *Joint2* conditions (**Figure 3.7**).

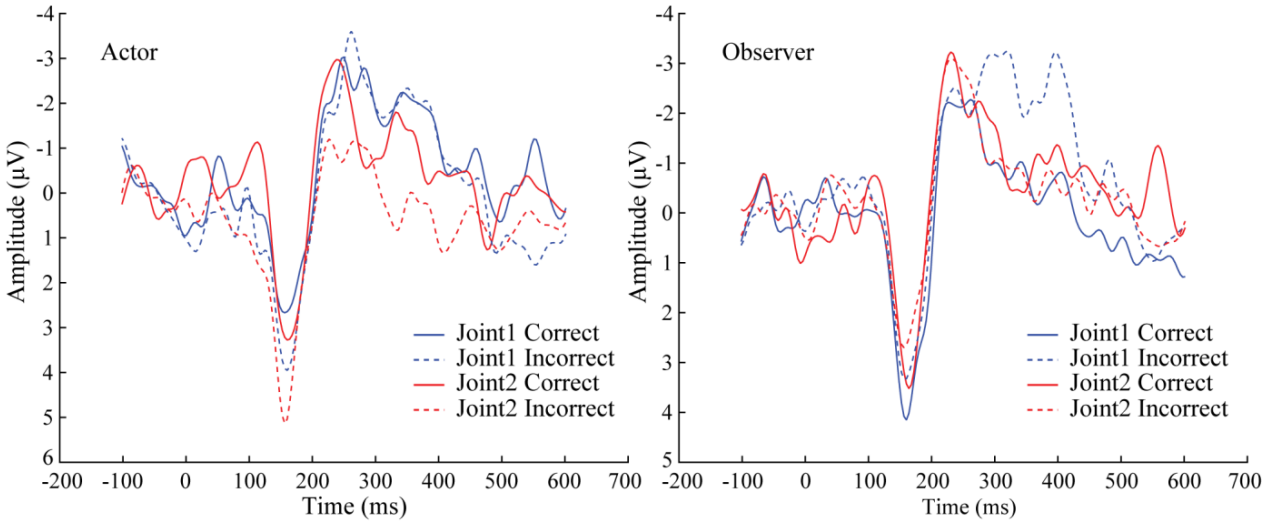


Figure 3.7. Grand averages of the P8 electrode

Next, a comparison between the complexities of the two FS methods took place estimating the overall performance in relation to the computational cost. As such, the standard SFS achieved high accuracy with small steps (including less features), although it reaches local maxima (nesting effects) and thus it cannot add more features. In comparison the scoring SFS surpassed the standard in terms of classification accuracy, while facilitating the nesting effects of the standard SFS greedy nature (**Figure 3.8**). Moreover, due to the scoring process (adding the next feature on the subset based on its significance without comparing all features) it presents an important decrement in the FS computational burden emerging from large feature sets.

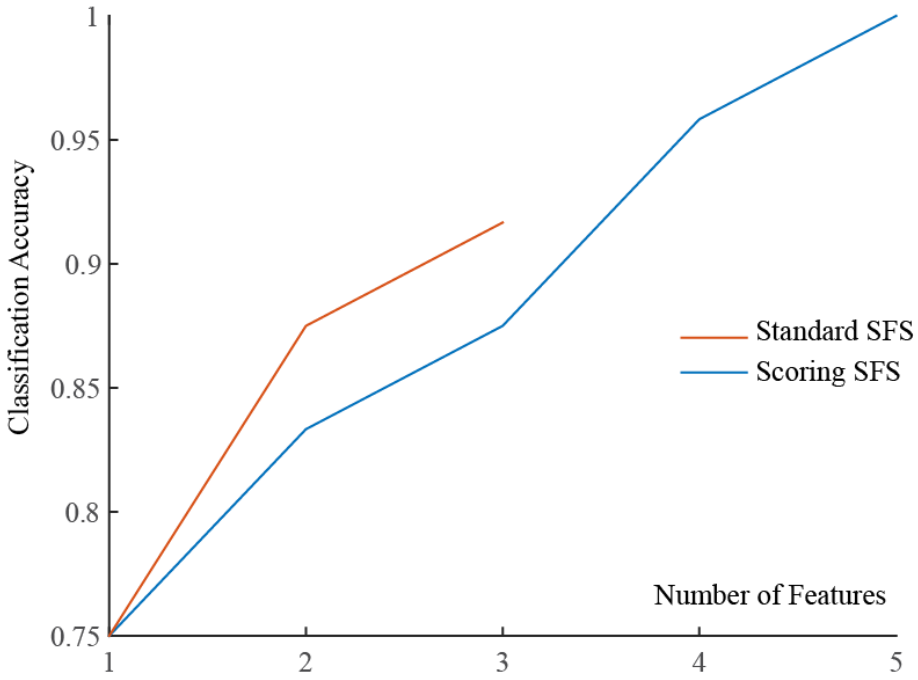


Figure 3.8. Accuracy per SFS step for standard and scoring SFS.

The results of this analysis don't only illustrate the importance of the FS methods in ML frameworks, but also highlight neural substrates that may be masked, especially in complex higher order cognitive functions. For instance, the *tw* selection of both FS methods applied, implying that ErrPs of consistently reported error related ERPs, such as the ERN [101], were detected although in different areas than expected. Furthermore, the utilization of ERP signal representation as features (such as the peak amplitude and time of appearance) in various time windows demonstrated their potential as biomarkers for ML applications. Interestingly, the small similarity between the different task-complexity ErrP signals could also infer that the difficulty in the tasks employed could exhibit morphological differences in the elicited ErrPs, masking specific components, while the use of ML could help in reveal their hidden neural substrates.

These facts taken together have been the key motivation for further investigation on the influence of different brain areas and task difficulty in the cognitive functions of error recognition. As such, a new approach was adopted described in the following section.

3.6 Cross-condition and Within-condition ML Analysis in Actor ErrPs

As previously mentioned the results and implications of the full channel analysis indicated the effectiveness of the time windows and waveform characteristics of ErrPs as features. In addition, the complexity conditions demonstrated potential modulations of the signal properties with regard to specific condition and time window. Taking the above into account, the subsequent analysis investigates the hypothesis of a small feature subset being able to provide high cross-condition classification performance, while detecting features specific for each condition to further increase within-condition accuracy.

In this context, a feature search strategy was utilized that incorporated a Sequential Forward Floating Selection (SFFS) and a Sequential Forward Selection (SFS) combination, allowing the detection of the individual features that provide high classification accuracy both in relation to task complexity and complexity-independent. The prominence of selected features where also assessed as being common to the two conditions and specific to each condition separately, successfully discriminating between correct and erroneous responses. The results of this analysis indicated the effectiveness of ML in the effort to detect distinct ErrP differentiations between correct and incorrect decisions, despite ErrP characteristics being affected due to task difficulty. Furthermore, they highlighted the linear nature of the ErrP signal attributes and the latency distortion of the typical ErrP time windows as a result of condition complexity manipulation. More importantly, they suggested a common underlying error detection cognitive mechanism and also modifications of that mechanism depending on the complexity of a task. This study was the first (to the best of our knowledge) to implement cross-condition ML approaches with regard to error-related responses.

3.6.1 Dissimilarities with the Full Channel ML Approach

Despite FS and classification differences, this approach also diversified in the participants and features employed. As such, the methodological strategy only included actors EEG recordings (as ErrPs in actors would theoretically include larger modulations due to complexity manipulation) in the two *Joint* conditions and a subgroup of electrode locations.

Specifically, based on the previous approach findings as well as on evidence from literature (please refer to “Full Channel ML Analysis in Actor and Observer ErrPs” and “Introduction” sections), central electrode locations are more indicative to error-related cognitive processes (despite temporal scalp regions being more relevant regarding auditory cognition). As such, in this approach only regions strongly related to error cognition were taken into account, with the remaining electrode position exploded from the subsequent analysis [79], [102]. To that end, the electrode selection (and by extension the features incorporated in this analysis) was decided relative to their position in the central scalp region (**Figure 3.9**). There were 7 electrodes included: Cz, Fz and Pz (as the midline electrodes adjacent to Cz) and FC1, CP1, CP2, FC2 (as the non-midline electrodes closer to Cz). Hence, from each averaged ERPs (please refer to “Feature Calculation” section), 5 features for each of the 5 time windows and each of the 7 electrode positions were incorporated in the subsequent analysis, resulting in $7 \times 5 \times 5 = 175$ features.

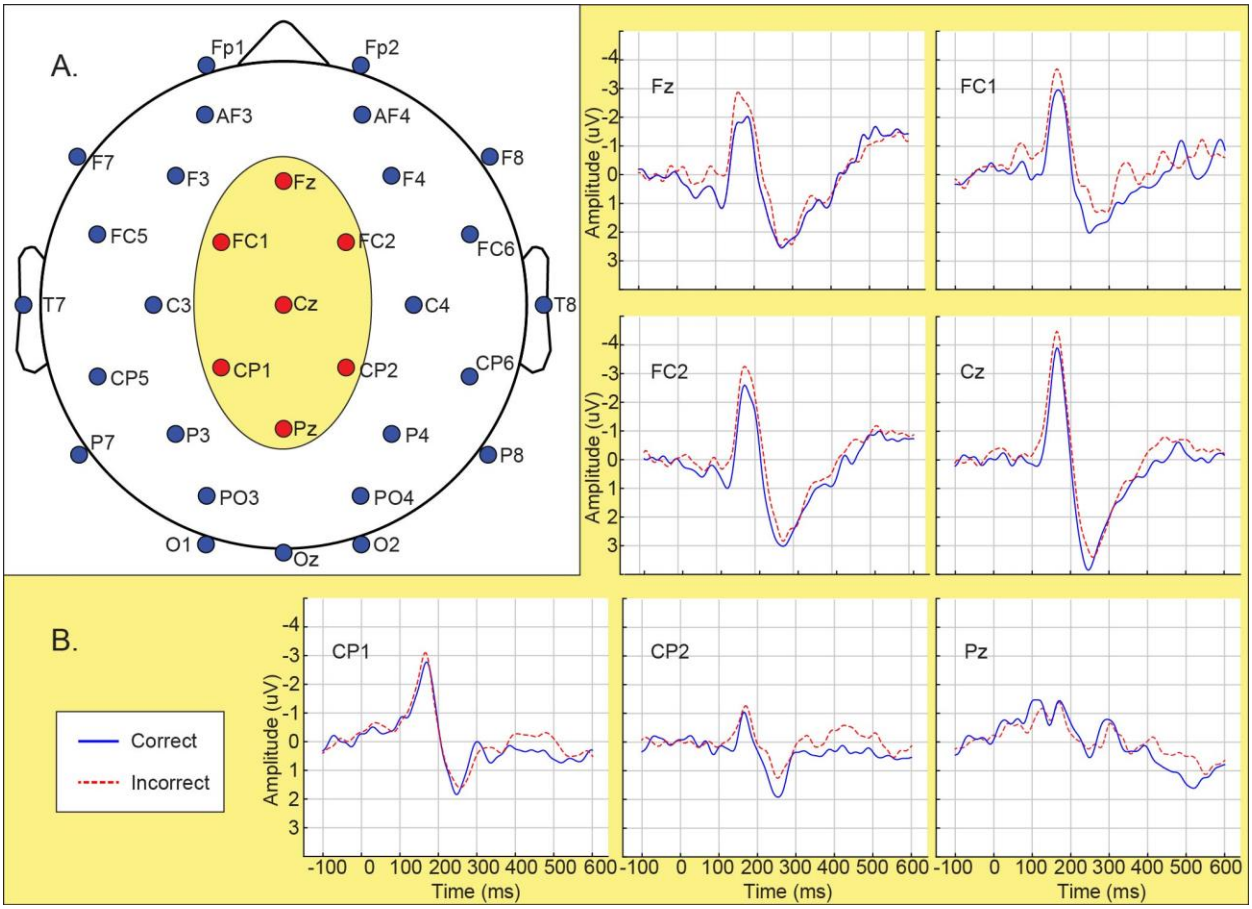


Figure 3.9. A) The electrodes position used in the EEG recordings. The ellipse includes the positions employed in the feature extraction processes. B) The average ERPs across all subjects (both conditions merged) for correct and incorrect responses

3.6.2 Feature Selection and Classification

In this approach, in order to discriminate between actors' responses as correct or erroneous, SVM classification was adopted with various learning methods and kernel functions [100], [103]. In detail, learning methods included sequential minimal optimization (SMO), least squares (LS) and quadratic programming (QP), while SVM kernels functions incorporated linear ($K(\vec{x},\vec{z}) = (\vec{x}^{T}\vec{z})$), radial basis function (RBF) ($K(\vec{x},\vec{z}) = e^{-\gamma\|\vec{x}-\vec{z}\|^2}$, $\gamma = 0.055,\ 0.08,\ 0.125,\ 0.22,\ 0.5$), quadratic ($K(\vec{x},\vec{z}) = (c + \vec{x}^{T}\vec{z})^{d}$, $c = 1,\ d = 2$), multi-layer perceptron (mlp) ($K(\vec{x},\vec{z}) = \tanh(k\vec{x}^{T}\vec{z} + d)$, $k = 1,\ d = -1$) and polynomial ($K(\vec{x},\vec{z}) = (c + \vec{x}^{T}\vec{z})^{d}, c = 1, d = 3$) kernel designs. For each learning method and kernel function combination, the overall classification accuracy, sensitivity and specificity were computed (for more details please refer to Chapter 2).

As mentioned in the previous sections, the aim of this approach was to identify condition-independent feature subsets that provide high classification performance and then detect condition-specific features that would enhance classification accuracy of the individual complexity levels. In this regard, FS and

classification was implemented concurrently for both complexity conditions and responses (i.e., classes) resulting in 12 × 2 × 2 = 48 instances (subjects × conditions × classes) to identify the optimal condition-independent feature subset and then, on top of that add condition-specific features to increase the performance on each individual condition (12 × 2 = 24 instances). Moreover, to provide a robust output on the specific features subsets that would provide the optimal classification performance and at the same time eliminate redundant and/or unnecessary features, two FS modification procedures were applied, as described in the following.

In the first procedure (FS1), a standard Sequential Forward Floating Search (SFFS) [104] was applied to the features that were computed, utilizing data from both conditions concurrently. In this manner, the final set of features (*cross-condition)* produced by the SFFS is deemed to represent the feature subset that best classifies correct or incorrect responses, irrespective of the difficulty of the task involved. SFFS alleviates the nesting problem occurring in other FS methods by including conditional repetitions containing three 3 steps: inclusion, conditional exclusion and sequential conditional exclusion. Firstly, starting from an empty set the most significant feature in terms of classification accuracy is selected thought exhaustive search. Then, a new feature is included based on its significance with respect to the existing feature subset. The next step is the evaluation of the new subset by excluding the least significant feature, provided that the resulting subset will include at least two features. If the removed feature was the one added in the subset in the proceeding step, the exclusion step is skipped and a new inclusion is made. Alternatively, a new conditional exclusion is made, provided that the new subset's accuracy is higher than the one currently found. When these conditions cease to be satisfied, the overall classification accuracy is computed, a new inclusion is conducted and the 3-step procedure is repeated. The algorithm stops when no further improvement can be made to the classification accuracy by modifications of the feature set.

Since SFFS is a sub-optimal technique, accepting the features from both conditions concurrently might (in some cases), result to over-fitting and thus make the subsequent condition-specific SFS procedure less effective (in providing features that would both improve the classification accuracy and be representative of the condition-specific mechanisms). Therefore, a second feature selection procedure (FS2) was also implemented.

In FS2, for each step of the SFFS method, the interim feature set computed by the SFFS was also used for the classification of the ERPs of actors, using data from the *Joint1* condition only and, separately, data from *Joint2* condition only. As a result, for each SFFS step and corresponding interim feature set, three overall classification accuracy values were available: the accuracy reached when both conditions were used concurrently, the accuracy when only *Joint1* condition data were used and the accuracy when only *Joint2* condition data were used. Then the average of these three accuracy values (*cross-condition*) was calculated. The average value (*task-specific average*) was calculated as the mean value of the task-specific accuracies of *Joint1* and *Joint2*. This procedure was repeated until SFFS concluded. Then the feature set that was selected for the application of the SFS method for data from each condition separately was the feature set for which cross-condition accuracy was highest. Another way of expressing the rationale for applying FS2, is that: although the features that will be finally selected might not provide the best classification for the combined data of the two conditions, stopping SFFS at an intermediate step, enables SFS to add features that are important for the classification of task difficulty-specific ErrPs (but would have been excluded by overly "fine-tuning" the classification, when both conditions were used concurrently). The workflow of the proposed FS methodology is presented in **Figure 3.10**.

Subsequent to the selection of the optimal feature subset, Sequential Forward Search (SFS) [105] was applied for the data of the two conditions separately based on the features selected by the SFFS. To that end, the SFS employed utilizes the SFFS output feature subset and iteratively includes the most significant feature (with respect to classification accuracy) through exhaustive search, until the classifier performance can no further improve. To alleviate the SFS nesting problems, the implemented algorithm would consider adding a second feature (as a two-feature addition) to the feature subset if a single feature addition did not improve classification performance as long as modified feature subset did not deteriorate in accuracy by each of the two single features added. When SFS concluded, the added features were expected to provide additional condition related information in the classification processes, enhancing the ML performance concerning individual complexity levels.

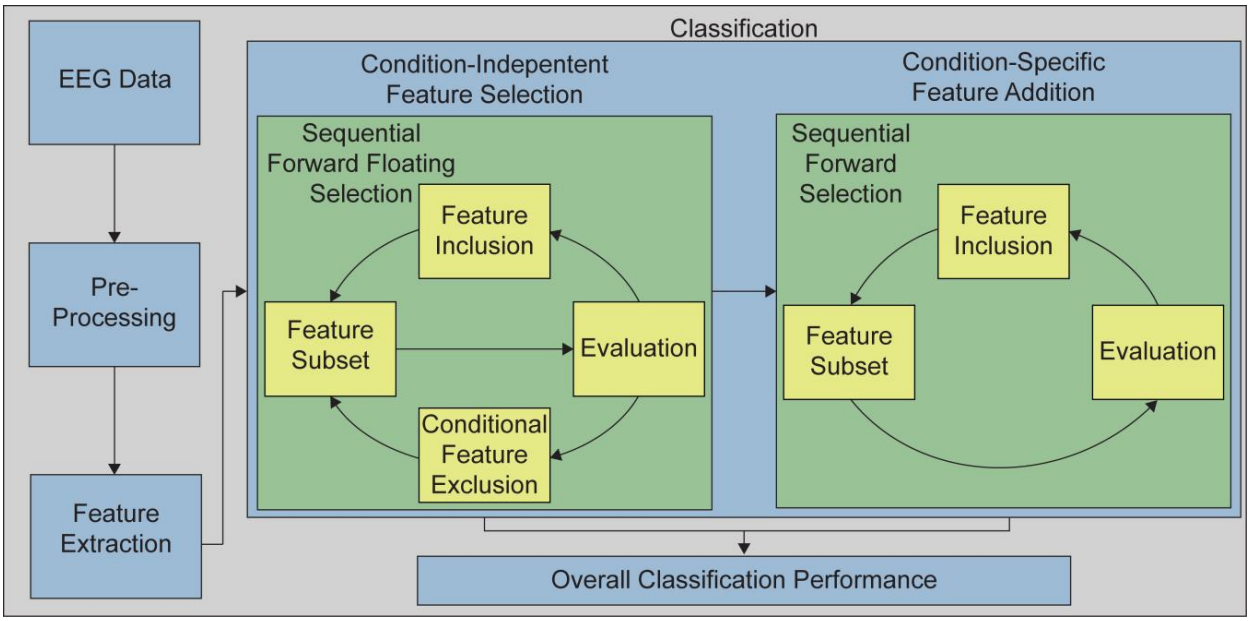


Figure 3.10. The workflow of the FS framework employed.

In this analysis, FS was done concurrently with the evaluation of the various classification algorithms, since the classification accuracies used for applying FS1 and FS2 were those provided by the respective classification algorithm, whose performance was evaluated. The training and testing of each classification algorithm was implemented using a leave-one-out cross-validation (LOOCV) procedure in every step of the SFFS and SFS methods. This procedure involves the exclusion of a single instance in the training data set, which will be utilized as the testing data set. This process is repeated for a different instance each time, until all data are employed as a testing set once. Accuracy, sensitivity and specificity are calculated as the average of each LOOCV fold. LOOCV procedures provide a reliable generalization and approximate better the actual performance of the classifiers than other cross-validation approaches, while facilitate in the prevention of over-training [106], [107].

The above procedures were repeated for each classifier configuration (i.e, learning method and kernel function) with classification accuracy being set as the objective function, allowing for FS and classification methods to be evaluated concurrently. In addition, the FS and classification processes were not only implemented for the five time windows, but also for combinations of two of them, to further investigate discriminative characteristics of the ErrPs and determine whether using features from

components belonging to adjacent time windows might improve classification. Specifically the time window combinations were: tw_1 and tw_2 ($tw_{1,2}$), tw_2 and tw_3 ($tw_{2,3}$), and tw_3 and tw_4 ($tw_{3,4}$). To avoid the inclusion of multiple components of ErrPs (which might mask the individual ErrP contribution to the classification procedures), overlapping windows were not included. This would also facilitate the effective detection of discriminative ErrP characteristics, while providing indication as to if adjacent *tw* ERP components might improve performance. Moreover, 1,000 runs of permutation tests on randomized class labels were carried out to verify that FS and classification introduced no overfitting or bias and to assess the statistical significance of the estimated performance. As in the previous approach (please refer to "full channel ML analysis in actor and observer ErrPs" section, "feature selection and classification subsection") p-values were calculated as the ratio of classifiers trained on the randomized labels that outperformed the classifiers trained on the original samples to the number of total permutations and is presented in a parenthesis in the results tables.

3.6.3 Results and Implications

The overall classification accuracy results applying the FS1 method are presented below. The performance on which the evaluation of the classifiers was assessed were the overall classification accuracy achieved by SFFS (*cross-condition)*, the task-specific classification accuracy where SFS applied to the data of the two conditions separately (*Joint1* and *Joint2*), and the mean value of the task-specific accuracies of *Joint1* and *Joint* as their average value *(task-specific average*). Since the different methods and kernels employed resulted in a very large number of methodological combinations, a performance evaluation criterion was arbitrarily determined, including accuracy values higher than 0.8 for cross-condition accuracy and task-specific average larger than 0.9. As such, only the corresponding results for these cases are given in TABLE 3.5, while in TABLE 3.6 the results for sensitivity and specificity are given. The features selected, presented in TABLE 3.7, are given for the cases where task-specific accuracy of *Joint1* or *Joint2* equals to 1.

TABLE 3.5. OVEARALL CLASSIFICATION ACCURACY RESULTS OF FS1 METHOD

		Overall classification accuracy				Number of features		
time window	Classifier	Cross-Condition	Joint1	Joint2	Task -Specific Average	Common	Joint1	Joint2
$tw_{1,2}$	SMO-quadratic	0.96**	0.88**	0.96**	0.92	9	9+4	9+11
$tw_{1,2}$	SMO-mlp	0.96**	0.88*	1.00**	0.94	35	35+8	35+3
$tw_{2,3}$	SMO-mlp	0.92**	0.92*	0.96**	0.94	8	8+4	8+6
$tw_{3,4}$	QP-quadratic	0.98**	0.92**	0.92**	0.92	20	20+3	20+6
$tw_{3,4}$	SMO-quadratic	0.98**	0.96**	0.92**	0.94	17	17+9	17+9

Note: The cases presented are those that exceeded the performance evaluation criterion. Asterisks mark the level of permutation p-values significance. *: $p < 0.01$; **: $p < 0.001$.

TABLE 3.6. SENSITIVITY AND SPECIFICITY RESULTS OF FS1 METHOD

		Sensitivity			Specificity		
time window	Classifier	Cross-Condition	Joint1	Joint2	Cross-Condition	Joint1	Joint2
$tw_{1,2}$	SMO-quadratic	1	0.91	0.92	0.92	0.85	1
$tw_{1,2}$	SMO-mlp	1	0.91	1	0.92	0.85	1
$tw_{2,3}$	SMO-mlp	0.88	1	0.92	0.95	0.86	1
$tw_{3,4}$	SMO-quadratic	0.96	0.92	1	1	1	0.86
$tw_{3,4}$	QP-quadratic	1	1	1	0.96	0.86	0.86

TABLE 3.7. FEATURES SELECTED IN THE FS1 METHOD

time window	Classifier	Cross-Condition	Features added for Joint1	Features added for Joint2
$tw_{1,2}$	SMO-mlp	tw_1:Fz_MinA, Fz_MinT, FC1_MinA, FC1_MaxA, FC1_AUC, FC2_MinA, FC2_MinT, FC2_MaxT, Cz_MinA, Cz_MinT, Cz_MaxA, Cz_MaxT, CP1_MinT, CP1_MaxA, CP1_AUC, CP2_MinT, CP2_MaxT, CP2_AUC, Pz_MinA, Pz_MinT, Pz_MaxT, Pz_AUC tw_2:Fz_MinA, Fz_MaxT, FC1_MinT, FC2_MinT, FC2_MaxT, FC2_MinA, Cz_MinA, Cz_MaxA, CP1_MinA, CP1_AUC, CP2_MaxA, CP2_AUC, Pz_MaxA	tw_1:Fz_MaxA, Fz_AUC, CP2_MaxA tw_2:Fz_MinT,FC1_MaxT, FC2_AUC, Cz_AUC, CP2_MinA	tw_1:FC1_MinT, FC1_MaxT, CP1_MaxT

Note: Features are given for the case in TABLE 3.5, where accuracy of Joint2 equals to 1. Each feature is denoted including the electrode from which the feature was extracted, followed by the feature itself, i.e. "electrode_feature".

The overall classification accuracy results reached following FS2 method are presented in TABLE 3.8. The overall classification accuracy values on which the classifiers were evaluated were the *cross-condition* classification accuracy, achieved by the SFFS at the step that was selected for starting SFS, the *task-specific* classification accuracy achieved by SFS applied to the data of the two conditions separately (*Joint1* and *Joint*) and the *task-specific average*. Again, the performance evaluation criterion was set as larger than 0.8 for cross-condition accuracy and higher than 0.9 for task-specific average, with the cases presented in TABLE 3.4 to be the ones that exceeded it. In **Figure 3.11** and **Figure 3.12**.the overall classification accuracy results (and performance evaluation criterion) of the FS2 method are presented. TABLE 3.9 displays the corresponding results for sensitivity and specificity, while the features selected for the cases where classification results were best (i.e., *Joint1* accuracy = 1 or *Joint2* accuracy = 1) are given in TABLE 3.10.

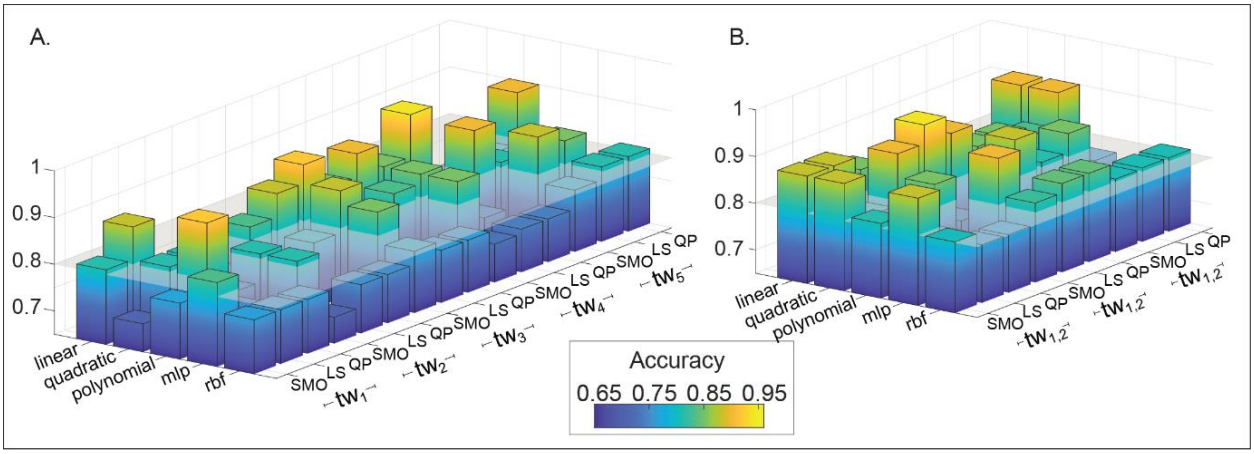


Figure 3.11. Cross-condition classification accuracy for all methods employed. In (A) for the individual time windows and in (B) for the time windows combinations. The elevated plane represents the threshold of the performance evaluation criteria of cross-condition accuracy larger than 0.8.

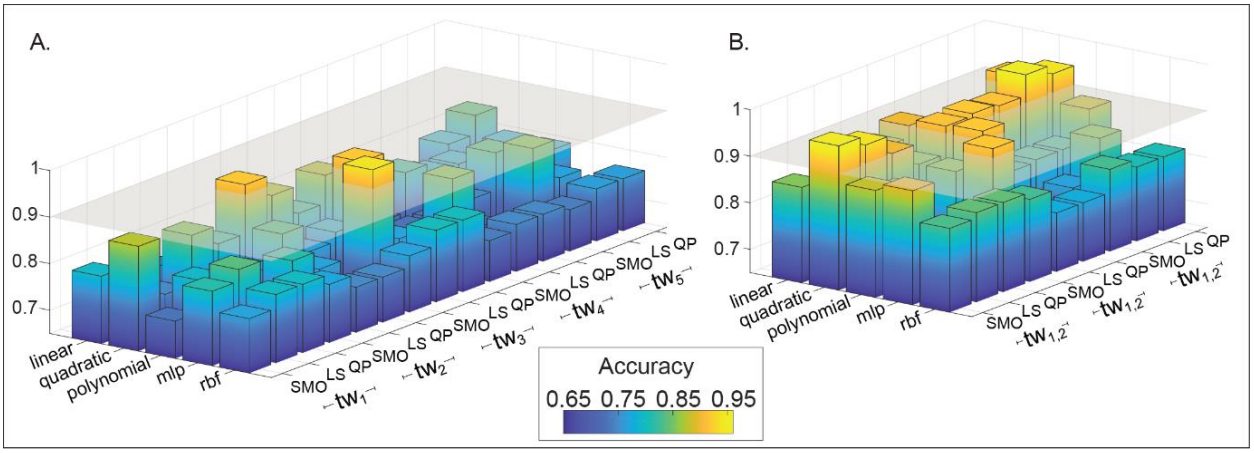


Figure 3.12. Mean classification accuracy of *Joint1* and *Joint2* for all the methods employed. In (A) for the individual time windows and in (B) for the time windows combinations. The elevated plane represents the threshold of the performance evaluation criterion of task-specific average accuracy larger than 0.9.

TABLE 3.8. OVEARALL CLASSIFICATION ACCURACY RESULTS OF FS2 METHOD

		Overall classification accuracy				Number of features		
time window	Classifier	Cross-Condition	Joint1	Joint2	Task -Specific Average	Common	Joint1	Joint2
tw_2	LS-quadratic	0.83**	0.88*	0.96**	0.92	5	*5+2*	*5+3*
tw_3	SMO-mlp	0.85**	0.96**	0.92*	0.94	5	*5+5*	*5+4*
$tw_{1,2}$	SMO-quadratic	0.88**	0.92**	1.00**	0.96	7	*7+1*	*7+5*
$tw_{2,3}$	LS-quadratic	0.90 **	1.00**	0.83*	0.92	9	*9+10*	*9+8*
$tw_{2,3}$	QP-quadratic	0.85**	0.92**	0.92**	0.92	9	*9+3*	*9+8*
$tw_{2,3}$	SMO-mlp	0.90**	0.92**	0.92*	0.92	9	*9+6*	*9+7*
$tw_{3,4}$	QP-linear	0.90**	0.92**	0.92**	0.92	12	*12+5*	*12+10*
$tw_{3,4}$	SMO-quadratic	0.85**	1.00**	0.83*	0.92	6	*6+3*	*6+1*
$tw_{3,4}$	LS-quadratic	0.83**	0.92**	1.00**	0.96	2	*2+4*	*2+3*
$tw_{3,4}$	QP-quadratic	0.90**	0.96**	0.92**	0.94	7	*7+16*	*7+2*

Note: The cases presented are those that exceeded the performance evaluation criterion. Asterisks mark the level of permutation p-values significance. *: $p < 0.01$; **: $p < 0.001$

TABLE 3.9. SENSITIVITY AND SPECIFICITY RESULTS OF FS2 METHOD

		Sensitivity			Specificity		
time window	Classifier	Cross-Condition	Joint1	Joint2	Cross-Condition	Joint1	Joint2
tw_2	LS-quadratic	0.86	0.91	1	0.81	0.85	0.92
tw_3	SMO-mlp	0.87	1	0.92	0.84	0.92	0.92
$tw_{1,2}$	SMO-quadratic	0.85	0.92	1	0.91	0.92	1
$tw_{2,3}$	LS-quadratic	0.91	1	1	0.88	1	0.75
$tw_{2,3}$	QP-quadratic	0.90	0.92	0.92	0.81	0.92	0.86
$tw_{2,3}$	SMO-mlp	0.95	0.86	0.92	0.85	1	0.91
$tw_{3,4}$	QP-linear	0.88	0.92	1	0.91	0.92	0.86
$tw_{3,4}$	SMO-quadratic	0.87	1	0.9	0.84	1	0.79
$tw_{3,4}$	LS-quadratic	0.86	0.86	1	0.81	1	1
$tw_{3,4}$	QP-quadratic	0.88	1	1	0.91	0.92	0.86

From the above TABLES it is indicated that both FS1 and FS2 did not produce results that passed the performance evaluation criterion when employing the RBF kernel. On the other hand, SVM classifier with linear, quadratic and mlp kernels presented higher results (passing the evaluation criterion) despite

different learning methods. Furthermore, in all cases for FS1 and in most cases for FS2, the performance evaluation criterion was met for features extracted from combination of 2 time windows. In contrast, the utilization of tw_5 (denoting the extended time window which included 0ms to 600ms ERP recordings) failed to meet the criterion threshold.

TABLE 3.10. FEATURES SELECTED IN THE FS2 METHOD

time window	Classifier	Cross-Condition	Features added for Joint1	Features added for Joint2
$tw_{2,3}$	LS-quadratic	tw_2: FC2_MinA, Cz_MinT, Cz_MaxA, CP1_MaxT, CP2_MinA tw_3: Fz_MinT, Cz_MinT, Cz_AUC, CP1_MinA	tw_2: FC1_MinA, FC2_MaxT, P1_AUC, CP2_MinT, CP2_MaxA, Pz_MinT, Pz_MaxA tw_3: FC2_MinA, Cz_MaxT, CP2_MaxA	tw_2: FC1_MinT, FC1_MaxA, FC2_MinT, Cz_MaxT, CP1_MinA tw_3: FC1_AUC, CP2_MinA, CP2_MaxT
$tw_{3,4}$	SMO-quadratic	tw_3: Fz_MinT, FC2_MinA, CP1_MaxA, CP1_AUC, Pz_MinA, Pz_AUC	tw_3: FC2_AUC tw_4: CP2_MaxA, Pz_MinT	tw_3: CP2_MaxT
$tw_{1,2}$	SMO-quadratic	tw_1: FC1_MaxA, FC2_MinT, CP1_MinT tw_2: Fz_MaxA, Fz_MaxT, Cz_MaxA, CP1_MaxT	tw_2: Pz_MinT	tw_1: FC2_MinA, CP1_AUC, CP2_AUC, Pz_MaxA tw_2: Pz_MinT
$tw_{3,4}$	LS-quadratic	tw_3: CP1_AUC tw_4:FC2_MaxA	tw_3: FC1_MaxT tw_4: CP2_MinA, Pz_MinA, Pz_MinT	tw_3: CP1_MinA, CP2_MinA tw_4: CP1_MaxT

Note: The first two cases correspond to the cases in TABLE 3.8, where *Joint1* classification accuracy reached 1, while the last two cases correspond to the cases in TABLE 3.8 where *Joint2* classification accuracy reached 1. Each feature is denoted including the electrode from which the feature was extracted, followed by the feature itself, i.e. "electrode_feature".

The comparison between FS1 and FS2, based on the number of cases that exceeded the criterion boundary, indicates that FS2 provided more cases. Specifically, classification accuracy equal to 1 was reached for SFS using FS1 only in one case, while in FS2 for 4 cases, using quadratic kernels. As the optimal performance attained in these cases, two of them resulted in 0.96 *task-specific average* accuracy. In addition, the permutation tests low p-values, infer that the combinations of the SVM classifiers employed were successful in detecting significant features and class labels associations, thus resulting in

high classification accuracy while avoiding overfitting. Furthermore, the high specificity and sensitivity values, as illustrated in TABLE 3.9, further corroborate the validity of the classifiers employed. The high specificity and sensitivity values in both FS methods, as illustrated in TABLES 3.6 and 3.9, further assess the effectiveness of the ML methods utilized since the small numbers of false positives and negatives indicate there was no bias included in the classification of one class over the other.

Concerning the number of features that were selected, after both SFFS and SFS were applied, it was observed that the total number of features selected from FS1, for each condition, is higher than the number of features selected from FS2, as represented by the mean of features selected by FS1 (24.2 for *Joint1* and 25.2 for *Joint2* condition) versus the mean of features selected by FS2 (12.6 and 12.8 features per condition, respectively), the mean being in each case computed from the number of cases presented in TABLES 3.5 and 3.8, respectively. A similar trend showed up in the fact that the number of features (17 and 20) selected from SFFS in FS1, in the cases that resulted in the highest cross-condition accuracy (0.98) were higher than the number of features (7, 9, 9 and 12) selected from SFFS in the FS2 cases that resulted in the highest cross-condition accuracy (0.9). Despite that however, no overall clear trend could be discerned among the selected features as for specific features of electrode positions.

Nevertheless, for the 2 cases in FS2 that t*ask -specific average* was higher than 0.9, while additionally *Joint1* accuracy reached 1, the features added by the SFS presented a central/centro-parietal majority (9 of the 13 selected features). In a similar way, for the 2 cases that the classification accuracy of *Joint2* was 1 with a task-specific average exceeding 0.9, the features selected to be added for *Joint2* demonstrated overall parietal/centro-parietal predominance (7 of the 8 selected features). Of note is that the two condition-specific subsets were different in the aforementioned cases. Specifically, inspecting the cases where *Joint1* reached 1, the features added were different from the features added for *Joint2*, starting from the same cross-condition SFFS feature set. The same held for the cases where *Joint2* reached 1, with the exception of tw_1, MinT for electrode Pz, In **Figure 3.13** the mean values and the distributions of the selected features are presented for the two cases corresponding to the optimal classification results (highest *task-specific average* and *Joint1* or *Joint2* accuracy equal to 1).

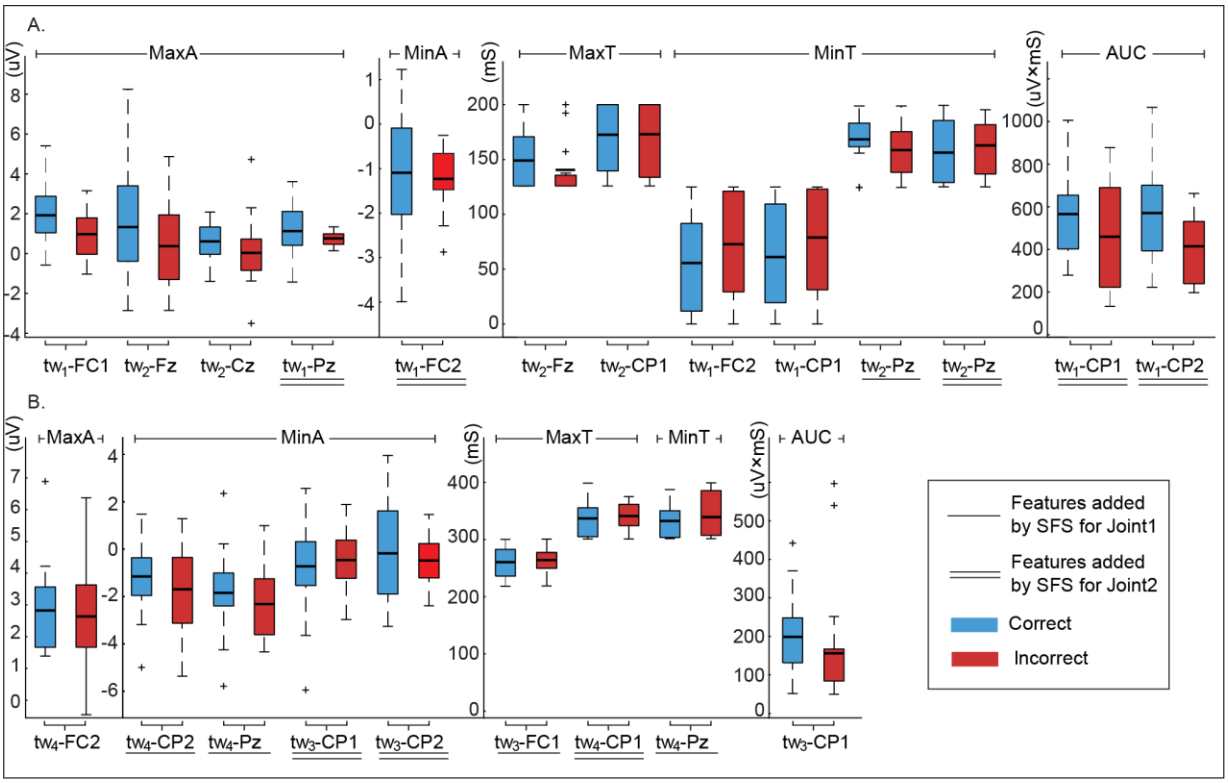


Figure 3.13. Feature values and distribution for cases (A) $tw_{1,2}$, method: SMO quadratic and (B) $tw_{3,4}$, method: LS quadratic. In each box, the mean value is indicated by the horizontal line, while the box edges denote the 25th and 75th percentile. The whiskers extend to the most extreme values and the '+' symbol marks the outlier data points.

3.6.4 Validation of Results using Additional Classifiers

To further assess the nature of the features extracted and compare the adopted framework with other ML designs, classification of different methods was also conducted (in addition to the SVM methods). As such the most commonly used methods in the classification of error-related potentials [78], [98], [108] were employed, namely k-Nearest Neighbors (k-NN) and Linear Discriminant Analysis (LDA), using the FS2 framework proposed as it was indicated to have the higher overall results.

In general, the classification performance of the classifiers tested was inferior to the SVM-based ML approach, demonstrated the effectiveness of the condition-independent and condition-specific SVM methods regarding error-related classification. The results for both methods are presented below in TABLES 3.11 and 3.12).

TABLE 3.11. OVERALL CLASSIFICATION ACCURACY RESULTS FOR THE DIFFERENT TIME WINDOWS (ADDITIONAL EVALUATION)

		Overall classification accuracy				Number of features		
time window	Classifier	Cross-Condition	Joint1	Joint2	Task –Specific Average	Common	Joint1	Joint2
tw_1	k-NN	0.69	0.71	0.75	0.73	3	*3+7*	*3+13*
tw_1	LDA	0.85	0.79	0.88	0.84	7	*7+2*	*7+3*
tw_2	k-NN	0.69	0.54	0.71	0.63	24	*24+2*	*24+5*
tw_2	LDA	0.71	0.61	0.71	0.66	3	*3+2*	*3+1*
tw_3	k-NN	0.73	0.71	0.67	0.69	22	*22+5*	*22+9*
tw_3	LDA	0.79	0.63	0.92	0.78	22	*22+1*	*22+3*
tw_4	k-NN	0.71	0.67	0.75	0.71	2	*2+11*	*2+14*
tw_4	LDA	0.71	0.91	0.87	0.89	22	*22+4*	*22+2*
tw_5	k-NN	0.87	0.59	0.70	0.65	31	*31+1*	*31+2*
tw_5	LDA	0.85	0.67	0.71	0.69	30	*30+3*	*30+2*

TABLE 3.12. OVERALL CLASSIFICATION ACCURACY RESULTS FOR THE TIME-WINDOW COMBINATIONS (ADDITIONAL EVALUATION)

		Overall classification accuracy				Number of features		
time window	Classifier	Cross-Condition	Joint1	Joint2	Task –Specific Average	Common	Joint1	Joint2
$tw_{1,2}$	k-NN	0.67	0.70	0.79	0.75	37	*37+6*	*37+6*
$tw_{1,2}$	LDA	0.88	0.92	0.95	0.94	10	*10+6*	*10+2*
$tw_{2,3}$	k-NN	0.75	0.75	0.79	0.77	3	*3+43*	*3+32*
$tw_{2,3}$	LDA	0.81	0.83	0.82	0.83	8	*8+2*	*8+1*
$tw_{3,4}$	k-NN	0.77	0.75	0.75	0.75	21	*21+26*	*21+34*
$tw_{3,4}$	LDA	0.87	0.92	0.79	0.86	12	*12+2*	*12+2*

In spite of the lower performance (relative to the SVM framework), the performance evaluation criterion of *cross-condition* accuracy larger than 0.8 and a *task-specific* average larger than 0.9 was reached in one case by an LDA classifier for $tw_{1,2}$ (**Figure 3.14**). Taking the overall results into consideration it can be inferred that linear modeling ML methods and time-window combinations are more efficient in error-related discrimination utilizing signal-based features.

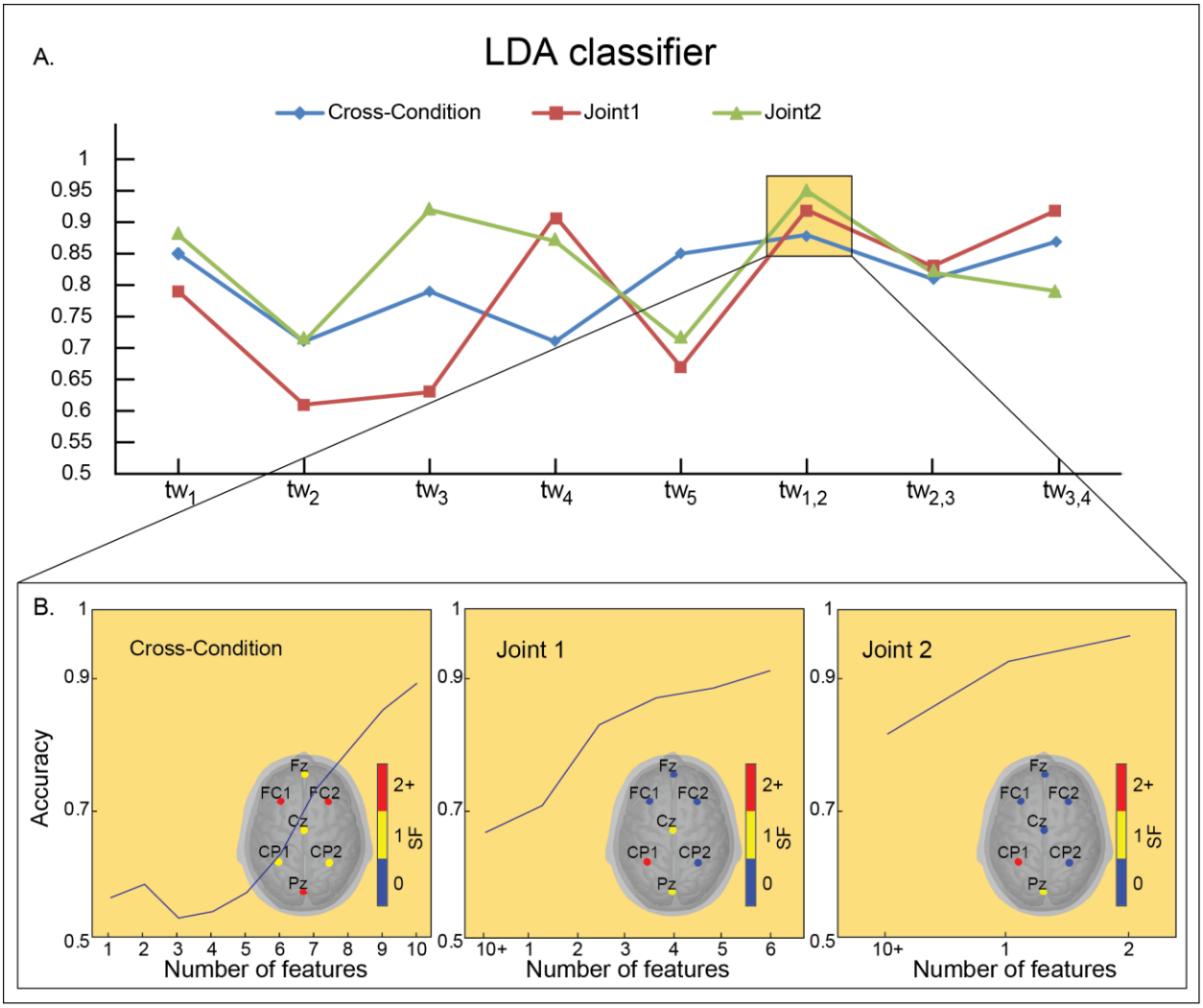


Figure 3.14. The performance of LDA classifier. In A) The classification accuracy for the different time windows and their combinations and in B) The classification performance for the case that reached the performance evaluation criterion ($\boldsymbol{tw_{1,2}}$). The line graphs demonstrate the classification accuracy as a step wise approach (inclusion of the different features), while the bottom right figures indicate the number of features for the different scalp areas.

3.6.5 Evaluation and Implications of the ML Approaches Employed with regard to ErrP Analysis

In this approach, we employed a cross-condition and within-condition ML framework in EEG recorded from an auditory task with two levels of complexity utilizing indicative ERP signal characteristics. Two feature selection methods were used to effectively identify the ErrP properties that are common to the two conditions and specific for each one resulting in effective classification between correct or incorrect ErrPs responses. The high classification accuracy reached, for cross-condition and within-condition classification for both feature selection methods, corroborates our initial hypothesis that, ML procedures can successfully detect hidden patterns in ErrP attributes and thus incorrect decisions can be detected, irrespective of the task difficulty, extracting condition specific features to improve the classification accuracy for each difficulty level.

While in practice the correct and incorrect EEG responses (when averaged across all subjects and conditions) demonstrated small distinguishable regarding error-related differentiation (indicated by previous research on the same data [109]) , the condition specific characteristics might provide indications for error processing mechanisms adjusted to task difficulty [110], [111].

Regarding the classification accuracy results, quadratic kernels demonstrated the overall best performance. Within the ML models adopted in this approach, in both FS methods RBF kernels failed to meet the performance evaluation criterion, while linear kernels presented the higher performance (despite the fact that SVM classifiers paired with the kernel trick can theoretically classify non-linear data adequately) indicated the linear nature of the features extracted [100]. This is also supported by other SVM studies that also exhibit higher classification accuracy in non-RBF kernels [112]–[114]. Additional evidence of the linearity of the features in this dataset was evident in the LDA comparison (please refer to the "validation of results using additional classifiers" section), demonstrating higher accuracy with linear modeling.

The response-related EEG signals analyzed in this chapter are those produced after the first feedback tone provided to the subjects (i.e., FBT) and as a consequence a signal similar to FRN would be expected to be generated. In this context a pre-defined time window could, in principle, be able to effectively analyze error-related components (as indicated by the majority of error-related studies) [74], [115]. However, because of the FBT nature and the variations of ErrPs amplitude and latency, due to individual subject differences and task condition manipulations [53], [102], [116] the timing and duration of the error-related ERPs could not be determined beforehand. From this standpoint, different consecutive non-overlapping time windows were investigated to refrain from excluding of latencies that provided useful information. The results suggested that relevant insight as to the time windows used, was especially extracted from combinations of adjacent instead of the shorter duration single time windows, i.e., $tw_{1,2}$ (0-220ms), $tw_{2,3}$ (125-300ms) and $tw_{3,4}$ (220-400ms). This could be attributed due to a latency extension of ErrPs present in feedback tone processing, causing error related attributes to manifest over the 200 ms time windows [117]. Task complexity could also be taken on account for distortions or masking of the multiple error-related components [118], supporting the assumption that the time windows combination could facilitate in the conditional inclusion of different ErrPs in the classification processes. In contrast, the extended tw_5 (0 ms to 600 ms) might incorporate unrelated error processing information (since large ERP peaks in auditory tasks generally appear after 400 ms [119]), thus resulting in a limited number of predictive features confounding the FS algorithms. To that end, feedback-related ErrP signals can exhibit high variability and may not be apparent to EEG recordings in view of experimental conditions and feedback relative concealment [120]–[122]. Furthermore, modulations of ErrPs attributes (even in correct trials) are consistently reported as a result of reinforced learning effects and reward expectancy [123], [124]. In the current approach the ERPs investigated and analyzed were generated from the FBT (therefore FBT originated from indirect information of the actors' responses), hence providing error-related features less clear in morphology as those that could be estimated after the presentation of a sole feedback tone (as unambiguous information on the correctness of the participants' responses).

Regarding the condition-specific features selected by the SFS when *Joint1* or *Joint2* reached accuracy equal to 1, single features added were different in the two conditions, initiating from the same SFFS set. This validates the hypothesis of the existence of ErrPs features corresponding specifically to each condition separately, eventuating divergent feature subsets for the two conditions for the cases that

provide the optimal discrimination between correct and incorrect responses. This is also supported by other recent ML studies that display discrepancies in ErrP classification performance with regards to cross-condition manipulations [94], [96], [125]. In this regard, however, conventional methods of training on one task and testing on another are likely to include condition-salient features, impairing the overall accuracy and, in some cases, making the classifier to be unreliable in cross-task classification [79], [94], [126], [127]. The adopted approach, achieves in disentangling condition-specific and condition-independent from the overall classification processes by selecting features irrespective of task complexity and complexity-related error processing variations.

Additionally, the fact that in FS2 more cases meet the performance evaluation criterion set and present task-specific accuracy equal to 1, suggest that FS2 processes provide more flexibility to the method employed, resulting in high classification performance between to correct and incorrect responses from each condition. On the other hand, the value of *cross-condition* accuracy reached by FS2 was in many cases lower than the value of *cross-condition* accuracy reached by FS1. This is to be expected, since in FS2, the feature set provided by SFFS, and used as a basis for SFS, takes also into account condition-specific classification accuracy. As such in FS1 the output SFFS set provides, by definition of the FS method, equal or greater accuracy than any other set, albeit selection bias might be included towards *Joint1* or *Joint2* classification. In addition, FS2 resulted, on average, in smaller feature subsets than FS1. This might have been anticipated considering that SFFS completes its feature inclusion cycles in FS2 taking into account the average value of three accuracies (which is lower than the optimal FS1 set) . It can also be regarded as the advantage of FS2, since in general classification systems based on fewer features are expected to perform faster in real-time systems.

Some considerations need to be taken into account in the interpretation of the results described above. Firstly, the ERP signals were average of each participant to address the unbalanced labels implications that usually reside with SVM algorithms [128], while also enhancing generalizability in the assessment of error related processes. This lead to a limited small number of instances to be classified, while since only SVM classifiers were evaluated (with the subsequent addition of k-NN and LDA for results validation), more classification methods could have led to diverse feature subsets and different performance (even surpassing the proposed ones). Moreover, despite the high overall accuracy reached by the selection of specific features for each method employed, the affinity of their correlation to the underlying error related cognitive mechanisms also raises concerns [129]. This is attributable to lack of consistency in the features subsets with regards to the different methods (TABLES 3.7 and 3.10). From this standpoint, the features that comprise the SFFS and SFS feature subsets might be unaffiliated to the error-related processes and therefore could be selected due to their ability to reduce unrelated noise. However, the potential of detecting condition-independent and condition-specific ErrP-based feature subsets for correct and incorrect classification purposes, provides indications task difficulty adjusted error-processing [110], [111].

Taking all the above into account, it can be assumed that ErrPs manifest as a global error-processing cognitive mechanism with common error-monitoring elements and salient feature alterations depending on task complexity. Towards this direction, future research could extend to explore a universal error-processing mechanism irrespective of task and/or complexity conditions, illustrating the underlying brain mechanisms that regulate the neural substrates of cognitive error responses.

Chapter 4

4.1 Task-Independent Workload Assessment

The evaluation of mental workload is a significant issue not only in the emerging field of neuroergonomics but also in real-world applications. Previous research works have attained high performance regarding efficient single-task classification, however expansion to cross-task mental workload assessment usually results in controversial results. In the work described below, an EEG experiment is adopted that incorporates two distinct tasks in two levels of task complexity. The data-driven analysis proposed in this Chapter, encompasses a ML framework to address the barriers frequently arising from cross-task workload classification, by utilizing EEG spectral attributes to uncover the global brain mechanisms regulating mental workload. In detail, functional connectivity and spectral density power features were calculated from different frequency bands in an experimental protocol that constituted two working-memory tasks of differentiated level of difficulty. The features were then fed to a feature selection and classification framework to assess the prominence of the features and subsequently estimate the task-independent classification performance. The ML approach obtained 0.94 classification accuracy in cross-task mental load discrimination, while additional analysis of the appointed features demonstrated common task-independent workload trends in spectral power and localization properties. Specifically, increased frontal, delta and theta power was detected with workload increments, while load-related differentiations of functional connectivity were also identified in terms of frequency and scalp locations. The overall approach illustrates the effectiveness of feature fusion in the endeavor of detecting indicative mental load biomarkers for the workload assessment irrelevant of the applied task.

4.2 Introduction

4.2.1 Background

Mental tasks involve various cognitive operations and processes of an individual, including storage, processing, transfer and retrieval of information, especially in strenuous task execution, deteriorating performance [130]. Nevertheless, the prolonged mental high-efficiency effort, especially in demanding tasks, necessitates more cognitive resources to be reallocated consequential leading to an increment of mental workload [131]. Recent studies report that increments in work engagement/intensity can result to mental overload, reducing operational performance and subsequently induce health conditions and burnout syndrome [132], [133]. In this regard, efficient estimation of workload-related mental states offers the opportunity for real-world assessment of cognitive burden with implication in clinical conditions as well [134], [135]. However, despite cognitive workload being consistently researched from various studies [133], the brain functions and mechanism that regulate it encounter several task-related challenges, whereas methodological implementations for real-world scenarios EEG workload detection are still nowhere near practical applications.

4.2.2 Machine Learning in EEG Workload Monitoring

As a means to evaluate the alterations of the various mental processes, electroencephalogram (EEG) has proven to be effective and practical as a means to provide low-cost, non-invasive electrophysiological

brain activity measurements with high temporal resolution [136]. Several studies have demonstrated EEG brain rhythms being associated to cognitive load, highlighting the neuronal processes, while functioning as effective indicators for workload estimation [137], [138]. In fact, load effects due to task difficulty have been observed in several brain wave aspects [139], [140]. Specifically, pronounced alterations have been evident in parieto-occipital alpha and temporal beta power with regards to complexity/workload increments [141], while significant relationships between alpha/theta band power and mental load have been consistently demonstrated in frontal and posterior areas [142], [143].

Taking into account the large volume of complex associations between the spectral oscillatory neural activities of the different brain regions, machine learning techniques allow for a robust monitoring and a comprehensive interpretation of the mechanisms that constitute mental workload. Although most of machine learning approaches focus on classification performance without considering possible physiological representation influences (i.e. data quality, unrelated noise etc.), by their very nature the distinctiveness between the different cognitive states of the features included can facilitate in the effective creation of a transparent and vigorous classification model with high prediction ability. With this in view, the utilization of EEG-based features in combination with machine learning methods can not only achieve high discrimination validity between the workload-related cognitive states, but also identify hidden patterns and unveil the underlying mental load cognitive processes, specifically in relation to the properties of brain wave frequencies and scalp locations [144]–[147]. For instance, Wang et al. [145], using a proximal support vector machine (SVM) method in a 4-level working memory (WM) task obtained high overall performance, utilizing EEG statistical, structural and power related characteristics. In their work alpha, frontal theta and posterior high beta and low gamma bands were indicated to have a significant effect in mental load. Similarly, linear discriminant analysis (LDA) exhibited a reduction in brain wave frequency power values, predominantly in centroparietal alpha and midline beta regions [144], whereas in [146] different entropy-related elements demonstrated the relevance of delta band in workload assessment particularly in frontal cortical areas.

Although the EEG ML applications have attained high discrimination between, uncovering important associations of the encompassed attributes with the cognitive states governing mental load, features generalizability in diverse tasks is still limited. On this premise, despite several studies indicating the successful detection of distinguishable patterns in single-task workload classification, the expansion to cross-task workload classification demonstrate significant obstacles that usually present deterioration in performance. A major factor for this could be the fact that most of the ML approaches include feature calculation and model generation within-task, while classification testing is performed in the other(s). As such, task-dependent characteristics that correspond to specific task, while being irrelevant to the other(s) might exclude global workload traits thus resulting in small classification efficiency. Intrinsically, workload-related cross-task efforts usually exhibit classification accuracy barely over chance level [148]–[150], with only a only few exceptions demonstrating reliable prediction results [151]. To enhance ML performance, relative studies include the combination of diverse cognitive attributes involving spatial, spectral and temporal EEG features [152] or features extracted from different types of electrophysiological signal modalities [153]. Taking the above into account it can be inferred that introduction of new features might increase cross-task classification modeling efficacy and therefore resulting accuracy [152], [154].

4.2.3 Brain Networks in Machine Learning Modeling

Most recently, accumulating studies have shown that the human brain forms a large-scale network of interconnected regions within the human connectome and such structure provides an anatomical substrate for neural communication, functional processing and information integration in the brain [155]. In consequence, functional connectivity (FC) utilization can contribute to a better perception of the complicated brain functions in diverse workload conditions, as indicative pairwise connection between the various brain areas. In that scope, alterations of FC and its corresponding network architectural aspects can expose segregated or integrated processing both locally and globally, as well as reveal the modulation of the topological properties of the brain's cognitive structures due to task load effects [155], [156].

From this standpoint, it can be presumed that FC features incorporation in ML frameworks can not only provide a higher level of interpretability regarding the brain's functional reorganization, but also contribute to the recognition of the hidden layers of constituting cognitive mental load, thus resulting in high classification performance [157]–[159]. Regardless, only a limited number of related research works integrate brain network FC in ML approaches either with regard to investigate the brain functions alterations in different tasks [160] or to illustrate brain regions communication in respect to the different frequencies and different workload levels in simulated environments [161]. In view of this, only our previous work [162] (to the best of our knowledge) employs FC as distinct workload related features to explore cross-task load discrimination. The resulted findings illustrated frontal theta and beta frequency bands power value modifications subsequently leading to 0.87 task-independent classification accuracy.

4.2.4 Current work

In light of all the above, it can be hypothesized that the already established univariate features such as spectral power with the addition of novel FC ones could improve task-independent discrimination, while providing indication of the brain related mechanisms governing mental load, promoting effective analysis. Consequently, the absence of extensive work on workload related classification in cross-task designs combining multi-variant features has been the principal motivation in this study. On this premise, in this Chapter a workload analysis framework is proposed, incorporating univariate power with multivariate FC features to improve the state-of-the arts performance regarding task-independent workload monitoring. As such, an EEG experiment incorporating two cognitive WM tasks (an n-Back and a Mental Arithmetic) in Low and High complexity conditions, assessing the different spectral attributes in terms of power and functional connectivity. Specifically, the Phase Lag Index (PLI) of pairwise electrode connections was estimated in 4 frequency bands to illuminate the FC structures and the reorganization of the mental load brain network, paired with the relative power spectral density (PSD) to identify prominent features for cross- task discrimination. This fusion feature set was then fed into a ML approach (including FS and classification), to calculate task-independent and task-dependent classification performance, while additionally providing indices in terms of their significance to the neural mental load processes.

4.3 Materials and Methods

4.3.1 Participants

In the EEG experiment, 40 individuals participated (males / females = 17 / 23) with a mean age of 21.6 ± 1.6 years. All individuals were right-hand dominant and reported normal or corrected-to-normal vision. Before the EEG experiments a phone interview screening took place, making sure that participant did not have any history of mental disease, Attention-deficit/hyperactivity disorder, sleep disorder and long-term medication intake. Individuals who were unable to obtain a full night of sleep (i.e., more than 7 hours / night) for two consequent nights prior the experiment were re-scheduled. Refraining from caffeine or alcohol consumption for at least 2 hours before recordings took place was also requested from the participants. Written informed consent was obtained from all subjects and experimental protocol was established in accordance to the Declaration of Helsinki.

4.3.2 Experimental Design

Cognitive task execution, in general, requires storage, retrieval, processing and transfer of information, with various brain function governing strategical planning, learning and WM [143]. In fact, WM is consistently reported to affect executive attention, inhibitory control, and cognitive flexibility, aspects particularly important in mental load processes synergetic with the way information is manipulated in mental workspace [163], [164]. In the experimental design presented in this Chapter, two cognitive tasks (n-Back and Mental Arithmetic***)*** related to WM were employed in low and high-workload levels (**Figure 4.1**)

The n-Back task utilizes 0-back and 2-back for low and high workload levels requiring subjects to identify target stimuli. Specifically, in 0-Back participants had to respond to the fixed target Letter '**X**'(target), whereas in the 2-Back to the uppercase letter that corresponded to the one presented two trials earlier (target). All other stimuli were considered non-target and no response was requested, while in addition in all trials no feedback was given to the subjects in regards to the correctness of their responses. Both n-Back designs comprised 150 trials of which 45 (30%) were target stimuli. Each letter was displayed for 0.5 sec following a 1.5 sec fixation cross.

In the Mental Arithmetic task, the two workload levels are handled under 1-digit numbers and two 3-digit numbers addition. As such, in each trial an addition was presented and next the sum (target) or a close number in terms of value (non-target). Participants had to judge if the answer provided was correct. In both 1-digit and 3-digit additions, there were 25 overall trials, 12 of which were target. Carrier influence was kept to a minimum with each addition concerning only one carry in both levels. Display of additions lasted 5 sec after a fixation cross for 4 sec, with answer presented for 2 sec followed by a 2 sec fixation cross.

The stimulus and interim fixation cross in each trial were presented in white fonts on a black background with each level lasting approximately 5 min. Each task was employed twice with counter-balanced order for each individual, who were requested to provide a response to both target and non-target stimuli. Preceding the EEG recordings practice n-Back and the Mental Arithmetic tasks were implemented with an arbitrary threshold of 80% for correct responses ensuring subjects understood each of the tasks

employed. Both tasks were conducted under E-Prime 2.0 software client (Psychology Software Tools, Inc.), while no feedback was provided regardless of the response accuracy.

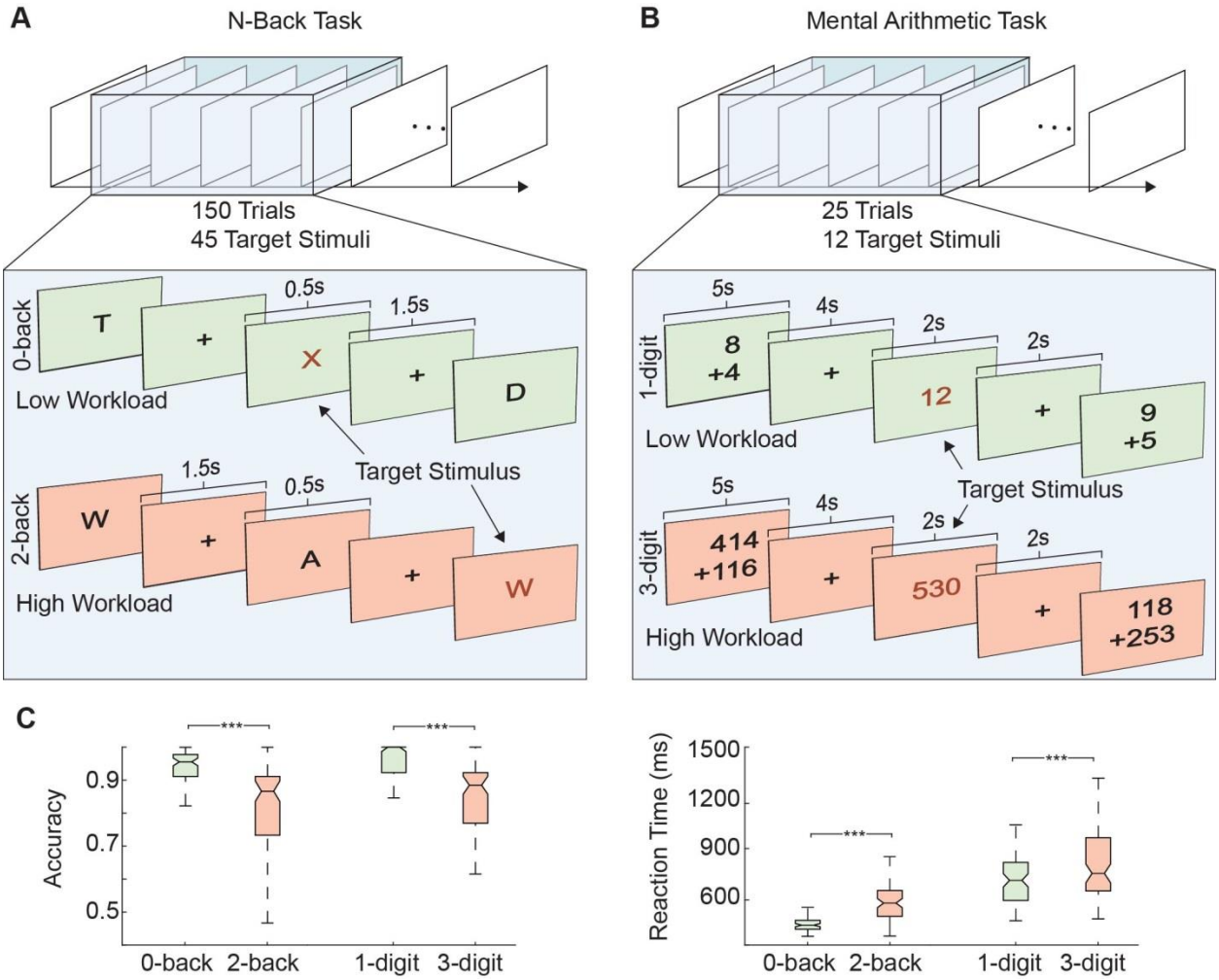


Figure 4.1. The experimental design for the high and low workload levels of the n-Back task (A) and the Mental Arithmetic task (B). Specifically, low workload level is marked as green color while high workload level is shown in light red in both tasks. Target stimuli are indicated with arrows and red color. The behavioral performance in terms of accuracy and reaction time are also presented (C). A clear workload-dependent effect on behavioral performance was revealed as reduced accuracy and increased reaction time with the increase of workload in both tasks (asterisks mark the level of significant difference between the workload levels as $p < 0.001$).

4.3.3 Data Acquisition and Preprocessing

EEG recording included 64 Ag/AgCl scalp electrodes (model: ASA Lab, ANT B. V., Netherlands), with positions corresponding to the international 10-20 system. Raw signals were digitized at a sampling frequency of 256 Hz, while electrode impedance was kept below 10 kΩ throughout the recordings, while electrooculograms were additionally measured by vertical (above and below the right eye) and horizontal electrodes (placed at the outer canthi). Preprocessing involved band-pass filtering (FIR 1 – 40 Hz), average re-referencing in relation to all electrodes, and artifact removal was performed by ICA, removing

the components highly correlated to ocular signals [165]. Subsequently, EEG data were de-trended, segmented to (stimulus-locked) epochs and baseline corrected relative to 100 ms before stimulus onset. More details of each preprocessing step can be found in Chapter 2. From the resulting data, only correct target epochs were incorporated for further analysis. Preprocessing was implemented with custom codes in the EEGLAB toolbox [166] in MATLAB R2019b (The MathWorks Inc., U.S.).

4.3.4 Feature Estimation and Fusion

In the approach presented in this Chapter, the detection of important mental load properties of EEG spectral characteristics were examined as well as their fusion in cross-task workload classification. On this premise, relative power spectral density (PSD) and phase-lag index (PLI) FC were calculated and employed as features for the subsequent classification. In detail, PSD was estimated for each channel using Welch's method with 50% overlap and 256 points on a 0.5 s window, in each trial. The resulting power values were then averaged for delta (δ, 1 – 4 Hz), theta (θ, 4 – 7 Hz), alpha (α, 8 – 12 Hz) and beta (β, 13 – 30 Hz) frequency bands. Moreover, to ensure a balanced dataset (since the number of correct target responses varied between tasks and subjects), the mean PSD was calculated for each subject, task and frequency band for all electrode locations. In a similar manner, PLI FC [15] was assessed for each of the aforementioned frequency bands, by calculating the average PLI network between all electrodes for each subject and task. Therefore, the final dataset included $40 \times 2 \times 2 \times 2 = 320$ instances (40 subjects completed two tasks twice for two workload levels) using 4 frequency bands and 62 electrodes, thus ending up with $4 \times 62 = 248$ power features and $4 \times 62 \times (62 - 1) / 2 = 7{,}564$ connectivity features, with the full fusion dataset comprising of 7,812 features (248 PSD + 7,564 PLI). The mathematical definition of PLI is provided in Chapter 1. For the purpose of the features (PSD and PLI) to display similar range in values, so that FS and classification can have a stable convergence (in terms of feature bias and weight) the relative PSD was calculated and incorporated in the final feature set. Thus, relative PSD is measured as the ratio of the band power to the total power ranging in the [0,1] interval (the same range as PLI values).

4.3.5 Feature Selection and Classification

In order to assess the predictive power of the features, and at the same time reduce variance based on the premise that high data dimensionality can hinder effective classification (since the fusion of PLI and PSD set involves a very large number of features), a feature selection (FS) framework was employed on the whole feature set. As such, the FS procedure was utilized to detect workload-related attributes incorporating them into an optimal feature subset with high predictive ability, while remove redundant or non-informational features, thus facilitation overfitting bias. In contrast to other cross-task classification studies (which generally train the classification model on one task ant test on the other), the FS and classification procedure adopted here was task-independent meaning that train and test was done while using global data from both tasks (maximizing generalizing). The algorithmic FS and classification framework is shown in **Figure 4.2**. As such, a recursive feature elimination method with correlation bias reduction (RFE-CBR) FS method was utilized that estimates the significance of each feature (consequently ranking them) on the basis of an internal linear SVM [167].

Algorithm 1 Feature selection

Input:

$D \in \mathbb{R}^{N \times F}$: data

$subject \in \mathbb{R}^{N}$: subject labels, $subject_i \in \{1, 2, ...S\}$

$workload \in \mathbb{R}^{N}$: workload labels, $workload_i \in \{0, 1\}$

Output:

$OptimalFeatures$: set of selected features

Begin:

for i in 1:S **do** ▷ Leave-one-subject-out cross-validation

$D_{test} = D_i$

$D_{train} = D - D_i$

$rank_i = \varnothing$

for j in 1:F **do** ▷ SVM-RFE-CBR

$weights = SVM(D_{train}, workload)$

$rank_{temp} = sort(weights)$

$rank'_{temp} = CBR(rank_{temp}, D_{train})$

$f = bottom(rank'_{temp})$

$D_{train} = D_{train} - D^{f}_{train}$ ▷ remove f^{th} feature

$push(rank_i, f)$ ▷ update the rank

for j in 1:F **do**

$selected = top\text{-}j(frequency(rank^{1:j}))$

for i in 1:S **do** ▷ Leave-one-subject-out CV

$D_{test} = D_i^{selected}$

$D_{train} = D^{selected} - D_i^{selected}$

$model = SVM(D_{train}, workload)$

$prediction_i = predict(model, D_{test})$

$acc_i = evaluate(prediction_i, workload)$

$acc_j = mean(acc_i)$

$K = argmax(acc_j)$ ▷ optimal number of features

$optimalFeatures = top\text{-}K(frequency(rank^{1:K}))$

return $optimalFeatures$

Figure 4.2. The algorithm for the optimal feature subset estimation

Briefly, the RFE-CBR assesses each feature importance (as a product of the SVM weight) excluding the feature that presents the lowest significance starting with a full set. In the same manner, the whole process is repeated in succession, until no feature is left in the complete set. The RFE-CBR ranked set is then created by the opposite order of the removed features. In addition to the internal SVM evaluation, the correlation between the different features is estimated (successively removing additional features), thus providing a more stable, while facilitating possible cases where features might be wrongly estimated. In order for the FS (and the subsequent classification) procedure to be able to identify subject-invariant attributes that encode global mental load information, leave-one (subject)-out cross-validation (LOOCV) was adopted. Specifically, the each LOOCV iteration allocated whole subjects to training or test sets, excluding one participant's data (both workload levels/tasks), with the RFE-CBR FS being fed the remaining data (training dataset), thus generating 40 feature rankings. Contrary to the FS (were only

training set data were employed), in the classification processes utilized the training set to create a mathematical model to distinguish between the low and high workload levels evaluating models performance on the testing set. Classification was implemented with the same LOOCV with a Linear SVM classifier [168] determining each set accuracy as objective function for subset estimation. A schematic of the FS and classification procedures utilized in this Chapter is presented in **Figure 4.3**. The feature subset with the highest average accuracy across folds was selected as optimal, calculated by adding the most frequently shared feature (on the ranked RFE-CBR sets) of all folds in succession, starting with an empty feature set. For example, the 1st feature subset included the most common RFE-CBR feature; the 2nd, the two most common features in the ranked RFE-CBR feature and so on.

To further ensure overtraining/overfitting was avoided, in addition to the classification accuracy estimation an additional 1,000 permutation tests with class labels was employed, under the same LOOCV procedure. The permutation test were able to provide an empirical distribution of the classification accuracy calculating the p-value as the probability of a random permutation to obtain higher performance than the one obtained by the SMV classifier [169].

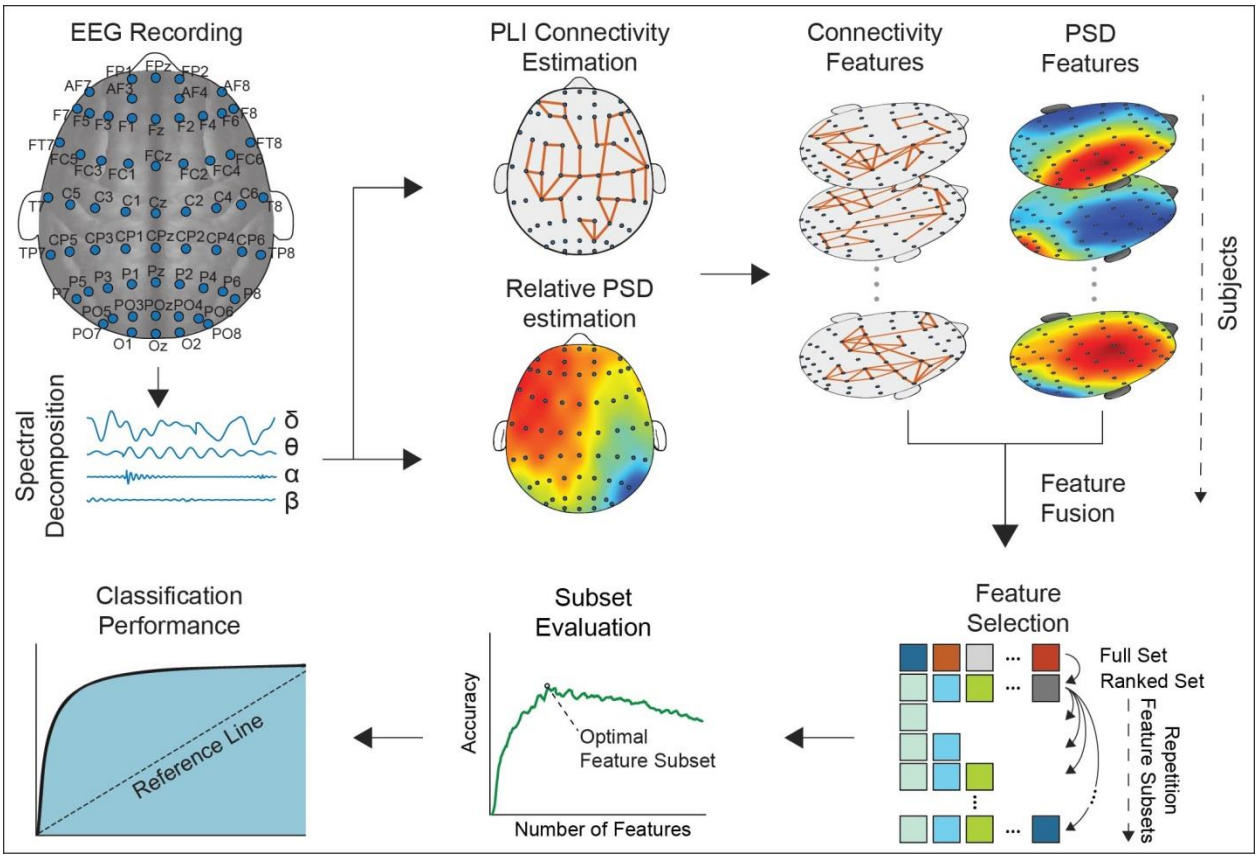


Figure 4.3. The workflow of the proposed approach

4.3.6 Validation

4.3.6.1 Comparison of Additional Classification methods

Additional validation of the selected features in terms of global predictive power (i.e. attaining of high classification accuracy utilizing the same feature subset irrelevant of the classifier), was performed by incorporating the RFE-CBR subset with different ML methods in addition to the primary approach presented above. As such, SVM with a Gaussian kernel, Linear Discriminant Analysis (LDA) and Random Forest (RF) classifiers were included, with the corresponding performance presented in the results section. To test the differentiations of the classification performance and further evaluate if the performance of the proposed approach was significantly higher from the additional classification methods used for validation, the McNemar test [170] was performed.

4.3.6.2 Leave-one-task-out Workload Classification

In general, most previous workload-related works [152], [153] implement cross-task classification by training on one task and test on the other(s). However, this approach is prone to incorporate task-specific features deterring the classifier of an overall task-independent model. In order to evaluate how the RFE-CBR ranked set affects the classification performance due to task-specific training; the selected features were further validated by including the linear SVM design in training and testing on different tasks. This approach is subsequently denoted leave-one-task-out for the remainder of this Chapter.

4.4 Results

4.4.1 Behavioral Results

As anticipated, the behavioral data (including reaction time and accuracy) presented statistically significant differences between the low and high workload levels in both tasks (**Figure 4.1C**). Specifically, the two-way ANOVA for task accuracy showed significant main task effect ($F_{1, 316} = 9.35$. $p = 0.002$) and main workload effect ($F_{1, 316} = 105.07$, $p < 0.001$), whereas effect for their interaction was not found ($F_{1, 316} = 0.005$, $p = 0.816$). In a similar way, the reaction time presented statistically significant main task effect ($F_{1, 316} = 199.01$, $p < 0.001$) and main workload effect ($F_{1, 316} = 52.49$, $p < 0.001$) with task-by-workload interaction failed to pass the significance threshold ($F_{1, 316} = 0.33$, $p = 0.569$). In particular, a pronounced decline was observed in the response accuracy, in conjunction with a reaction time increase for both tasks (absent of significant interaction effects), indicate the effectiveness of mental workload increments due to the task-related complexity manipulation.

4.4.2 Classification Performance Results

As mentioned in the previous section, cross-task classifications was implemented by applying PSD and PLI FC features from both tasks and differentiate between the two workload levels (low *vs*. high). In this regard, the overall highest accuracy was 0.94 ($p < 0.001$, 1,000 permutations) with 0.93 sensitivity and 0.94 specificity), including 166 features of which 18 (11%) were PSD and 148 (89%) were FC (TABLE 4.1). Moreover, the workload classification performance of each feature set (PLI and PDC) was evaluated separately. In detail, utilizing only the 18 PSD features the obtained performance was much lower,

namely 0.59 ($p < 0.01$, 1000 permutations) for n-Back, 0.63 ($p < 0.001$, 1000 permutations) for Mental Arithmetic and 0.61 ($p < 0.001$, 1000 permutations) for cross-task classification. On the other hand, workload classification employing the 148 PLI features resulted in high (although inferior to the PSD and PLI fusion sets) with n-Back, Mental Arithmetic and cross-task attaining 0.95 ($p < 0.001$, 1000 permutations), 0.82 ($p < 0.001$, 1000 permutations) 0.93 ($p < 0.001$, 1000 permutations) classification accuracy respectively.

TABLE 4.1. CROSS-TASK CLASSIFICATION PERFORMANCE RESULTS

Feature set	Classification Accuracy (Standard Error)			# of feature
	n-Back	Arithmetic	Task-independent	
PSD	0.59(0.02)**	0.63(0.03)***	0.61(0.02)***	18
PLI	0.95(0.02)***	0.82(0.02)***	0.90(0.01)***	123
Fusion	0.97(0.01)***	0.83(0.03)***	0.94(0.01)***	141

Note: Asterisks mark the level of significance *p < 0.05; **p < 0.01, ***p < 0.001, 1000 permutations

4.4.3 Relative PSD Features

Subsequent investigation of the PSD features revealed difference in the topographic properties and their distribution (in low and high workload) as presented in **Figure 4.4**. Interestingly, the PDS feature values in both tasks displayed comparable patterns in all frequency bands employed, further denoting the classification viability for mental load cross-task classification. Specifically, from the 18 PSD features selected, 5 were in δ, 10 in θ and 3 in α frequency bands. Notably, no PSD features were selected from β band. In detail, concerning δ features 3 were included in frontal electrode positions, with 1 in central and 1 in parietal electrode sites. In the same manner, a frontal predominance in θ frequency band was found (8 out of 10) with the rest PSD features in central areas (2 out of 10). On the contrary, a parietal-occipital majority was detected in α band, with all features being selected from parietal/parieto-occipital positions. Although no clear trend could be discerned about the δ frequency band (with 3 out of 5 PSD features decreasing indicating a complex parietal-increment-fronto-decrement pattern), θ PSD features demonstrated an overall increase from low to high workload, whereas α, an overall decrement. Interestingly, regardless of the task employed (i.e., whether it was n-Back or Mental Arithmetic), the same increasing or decreasing trends were observed in all frequency bands.

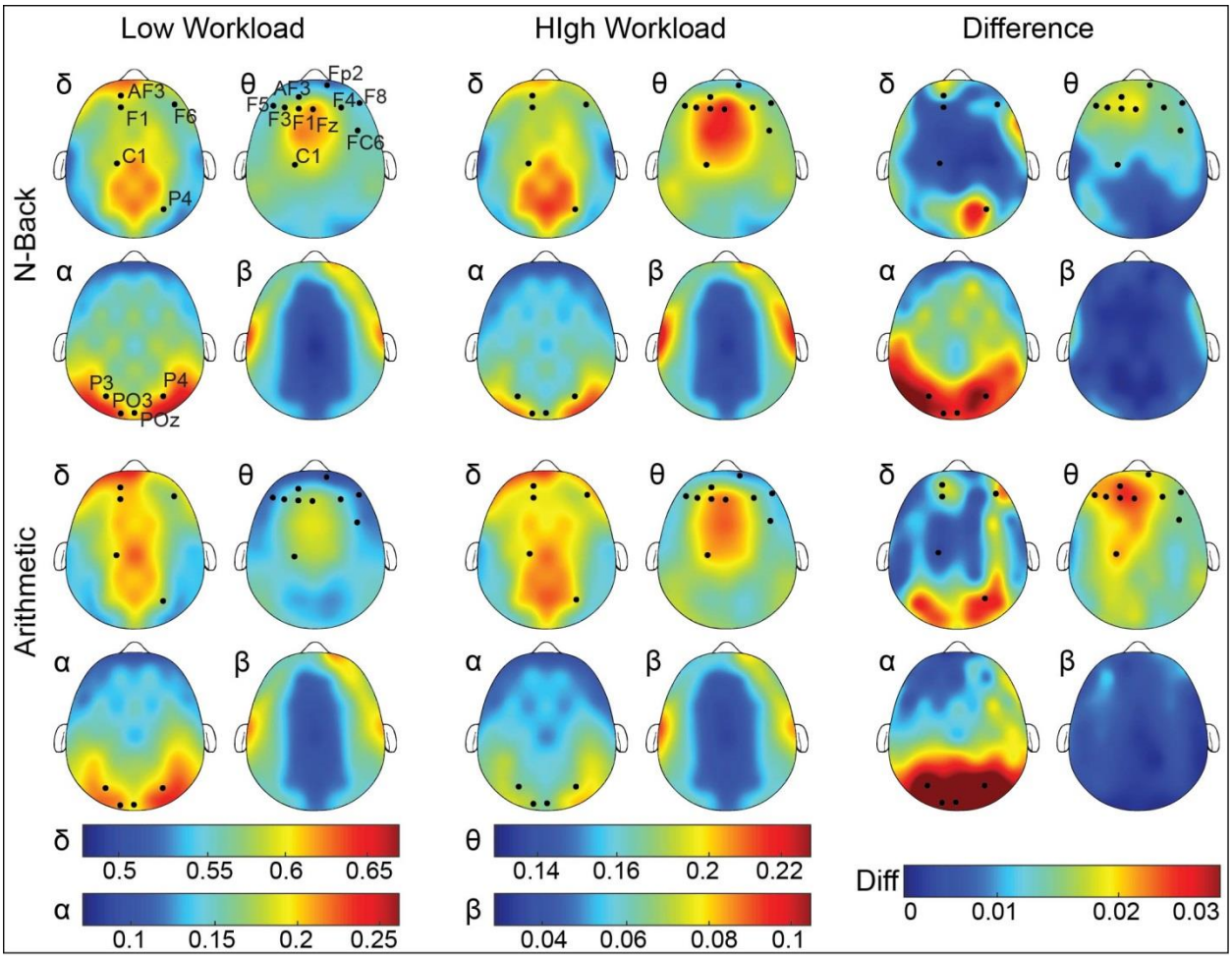


Figure 4.4. The PSD distribution in terms of topography relative to low and high workload for the n-Back (upper part) and the Mental Arithmetic (lower part) task. PSD between both tasks present similar patterns across the four frequency bands with the color bar below indicating the relative PSD mean value for each frequency band. The differences of PSD between low and high workload levels are provided in the right panel. The color of the bar indicates the absolute value of relative PSD differences.

4.4.4 Functional Connectivity Features

The FC features of the different frequency bands are presented in **Figure 4.5**. From the 148 features corresponding to PLI FC, the majority of connections included θ (n = 56) and β frequency bands (n = 54), with a smaller feature number pertaining to δ (n = 20) and α (n = 18) bands. In particular, most of the θ band connections involved frontal electrode positions; however an overall increasing or decreasing trend regarding connectivity strength could not be discerned. In contrast, 63% of the θ connections in parietal locations presented connectivity strength decrement in the high load level. At the same time, in β frequency band frontal FC strength reduction (from low to high workload) was observed in most cases (30 out of 54 features), whereas δ frequency band showed an 80% (16 out of 20) FC deterioration in all scalp locations. Quite the reverse was shown in α frequency band connections, with an overall augmentation of the FC strength (14 out of 18) in high workload in all but occipital areas where the opposite was noted. Of note is that 141 out of 148 of the PLI features presented similar trend in terms of

increases or decreases in both tasks. Nonetheless, 7 FC exhibited opposite fashion regarding connectivity strength increments in n-Back and Mental Arithmetic tasks. These included 5 θ band connections, 2 of which concerned central locations, 1 frontal/fronto-central, 1 temporal and 1 frontal/parietal locations; and two α band FC in frontal/fronto-central and parietal/occipital locations. Despite that, the discrepancy of the FC strength displayed no significant alteration in the overall position percentage.

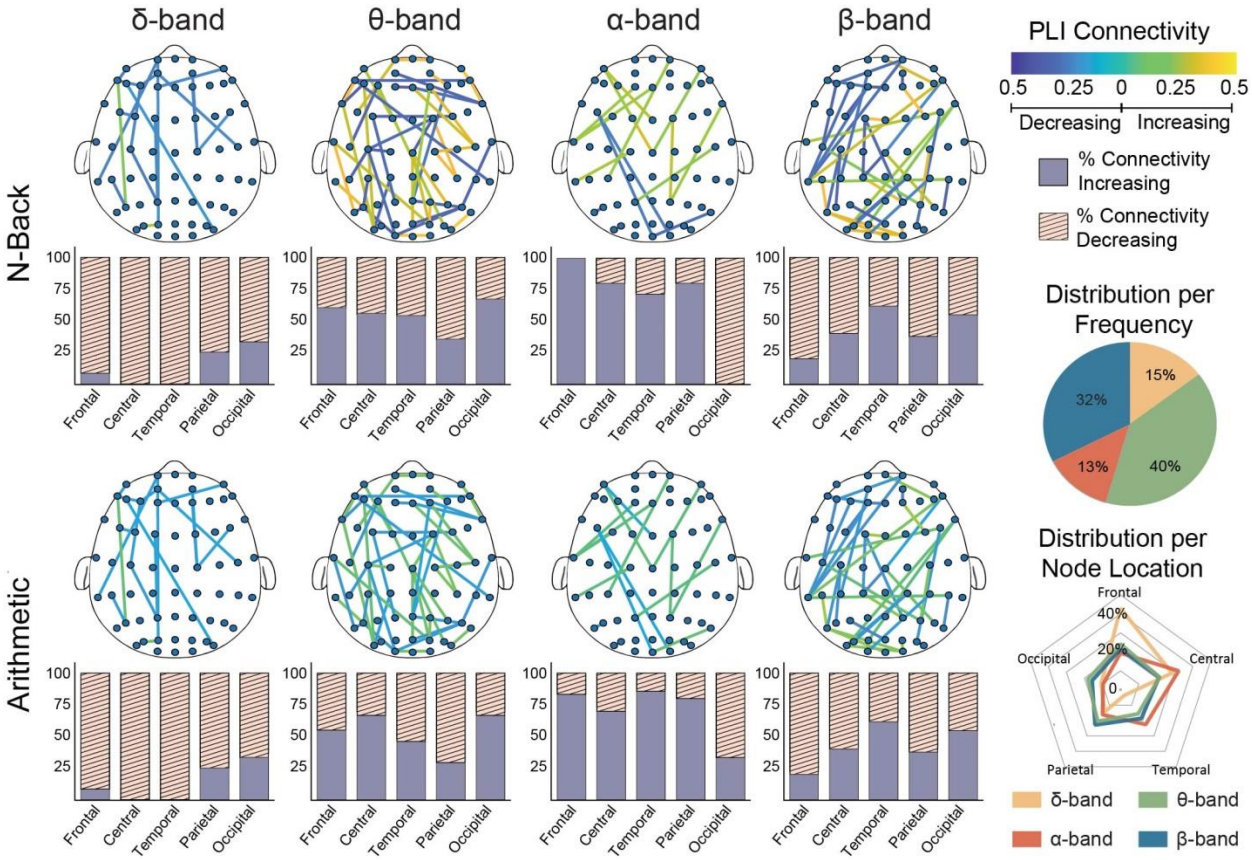


Figure 4.5. The selected PLI features for cross-task mental workload classification. To better reveal the alterations of FC with mental workload, the distribution is presented in each task (n-Back for upper panel, Mental Arithmetic for lower panel) across the four frequency bands. The color of the edges shows the mean value difference of the connectivity strength. The color bar on the top right corner represents the FC variation. The FC features are grouped according to their scalp location (frontal, central, temporal, parietal and occipital), while ration of increments and decrements in FC (relative to mental workload for each frequency band) is displayed in the bar plots below the FC features distribution. Particularly, a unilateral decreased FC pattern was revealed in δ band, while most FC features in α band exhibited an increased pattern. FC features in θ and β bands comprised over 70% of the selected FC features and showed a location-dependent complex alteration pattern. The overall distribution of FC across the four frequency bands is depicted in the pie chart, whereas the distribution per electrode position is presented in the radar plot.

4.4.5 Statistical Evaluation of Features

The variable values differences between the two workload levels were additionally explored by statistical evaluation on the selected RFE-CBR features. As such, an one-way ANOVA was implemented on each of the 166 features individually, presenting statistical significance at p-values <0.05 (95% confidence interval). Notably only 5 of the total features displayed significant p-values, none of which represented PLI features and were all related to relative PSD (**Figure 4.6**). Remarkably, all significant features presented θ band increasing values form low to high workload, with additional frontal majority (4 out of 5) and left hemispheric predominance. This is in line with several studies providing indications of frontal θ band power increments in high mental load (further details are presented in the Discussion section), nevertheless due to classification and statistical testing approaches different processing procedures, feature importance is indicated differently in each approach. For instance statistical ANOVA procedures estimate the mean value variance of low *vs* high workload by terms of variance of the confidence interval, while the FS scheme implemented determines features' prominence by the weights deriving from the internal SVM classifier mapping feature vectors in multi-dimensional space and estimating maximum hyperplane deviations. From this standpoint, statistically significant attributes might not be effective in ML methodological differentiations or in the identification of the predictive power of each feature.

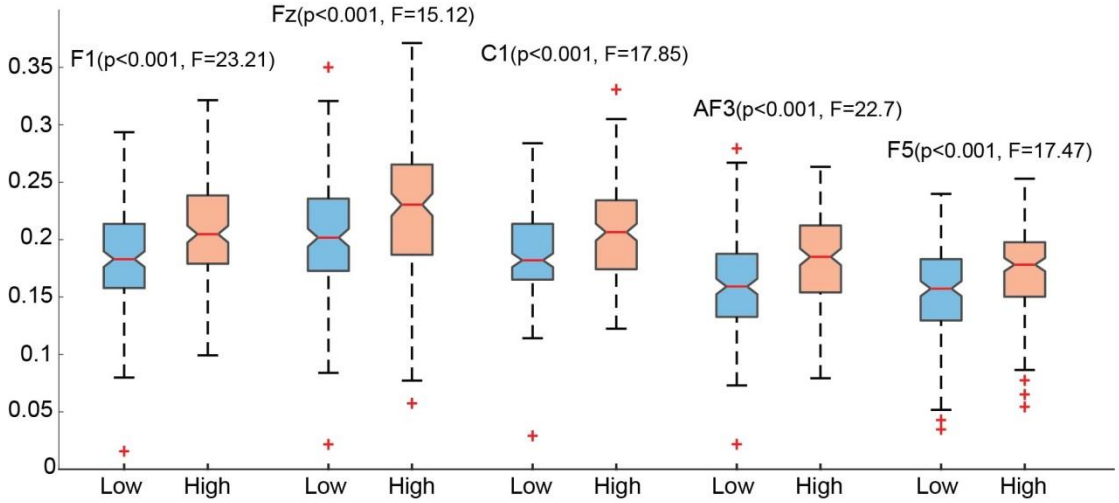


Figure 4.6. The significant feature value variations between low and high mental workload. Every box represents the feature variance (Above them is the electrode position and the ANOVA p- and F-values), with the whiskers extending to the most extreme points, the red cross ('+') displaying the data outliers and the red line denoting the median value.

4.4.6 Validation Results

The feature subset validation included four additional classifiers, namely Gaussian SVM, k-NN, LDA and RF, the performance of each is displayed in TABLE 4.2. From the table it can be indicated that all classifiers obtained high performance, in some cases comparable to the proposed approach described in this chapter (as suggested by the receiving operator characteristic, ROC provided in **Figure 4.7**), although none surpassed it in terms of classification accuracy. McNemar's test was also implemented to indicate that the classifiers where indeed different, resulting in all comparisons attaining a McNemar's p-value

<0.05. On this premise, it can be suggested that the features selected present high cross-task predictive ability in a global scale, irrespective to the classifiers employed or to algorithmic variations.

In addition to the cross-task workload classification, a leave-one-task-out scheme was employed to further estimate the universal nature of the feature. In this paradigm classification training was used on one task, while testing was done on the other providing two ML comparative results; i.e., train on n-Back / test on Mental Arithmetic, and train on Mental Arithmetic / test on n-Back (TABLE 4.3). Notably, the performance of the leave-one-task-out approach was lower than the one proposed in the cross-task approach proposed in this chapter, although it was comparable in terms of accuracy with the work in [152] (the highest so far) implementing a similar framework.

TABLE 4.2 CROSS-TASK CLASSIFICATION RESUTLS USING ADDITIONAL METHODS

Feature set	Classification accuracy (Standard Error)			
	k-NN	Gaussian SVM	LDA	Random Forest
PSD	0.57*(0.03)	0.63**(0.02)	0.59**(0.02)	0.58*(0.03)
PLI	0.71***(0.02)	0.90***(0.01)	0.86***(0.02)	0.73***(0.02)
Fusion	0.76***(0.02)	0.91***(0.02)	0.91***(0.02)	0.78***(0.02)

Note: Asterisks mark the level of significance *p < 0.05; **p < 0.01, ***p < 0.001, 1000 permutations.

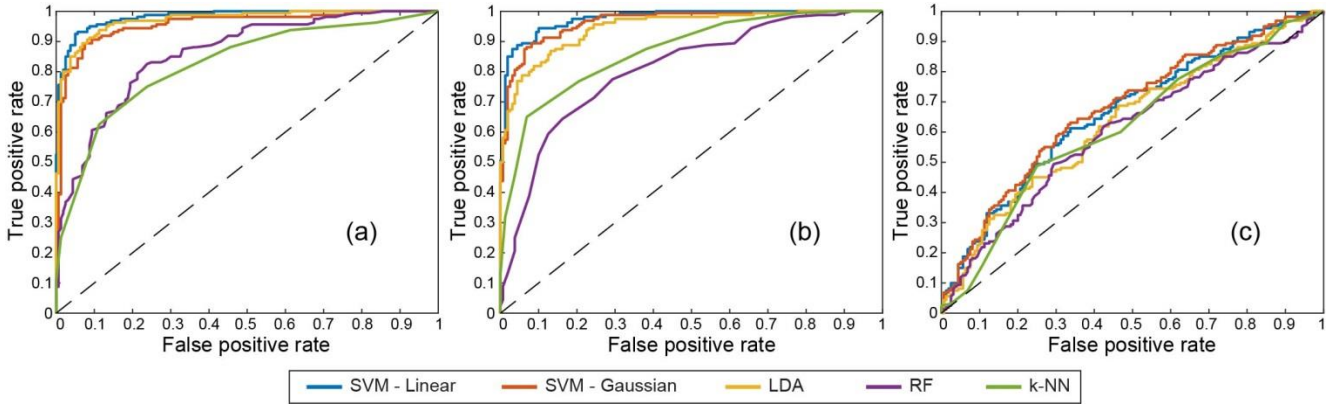


Figure 4.7. The receiving operator characteristic (ROC) curves for the methods implemented in (a) feature fusion, (b) using only PLI features and (c) using only PSD features

TABLE 4.3 LEAVE-ONE-TASK-OUT CLASSIFICATION PERFORMANCE RESULTS

Approach	Classification accuracy (Standard Error)		
	PSD	PLI	Fusion
Train on NB, Test on MA	0.61**(0.03)	0.64**(0.02)	0.90*(0.02)
Train on MA, Test on NB	0.59**(0.02)	0.78***(0.03)	0.81***(0.04)

Note: NB = N-Back, MA = Mental Arithmetic; Asterisks mark the level of significance *p < 0.05; **p < 0.01, ***p < 0.001, 1000 permutations.

4.5 Discussion

In this chapter, an analysis framework is presented regarding cross-task workload classification in two different WM mental tasks, utilizing spectral EEG characteristics. As such, a fusion of different frequency

bands relative PSD and PLI FC features was implemented, employing a ML framework incorporating FS and classification to identify the feature subset attaining the overall optimal performance. The selected features encompassing the optimal set were further analyzed in terms of frequency and scalp positions, providing indices of common (to both tasks) properties in workload underlying functions. The provided results affirm the efficiency of the approach proposed, verifying that feature fusion of power and FC attributes can obtain high classification performance, while effectively identify the hidden neural substrates governing mental workload in a global manner. Taken all of the above into account, the significance of this work does not only concern academic brain-related research, but also provide social- and health-related indices resulting from workload monitoring. For instance, besides the evident importance of mental fatigue evaluation (and by extension the decrement in individuals' performance) generated from sustaining high mental load, workload effects have been associated with various health conditions such as mental burnout, increased risk of physical, mental and emotional diseases, reduced immune system, depression etc. [171].

4.5.1 Workload-related Classification Performance

As far as the classification performance is concerned, the framework proposed in this Chapter obtained by far the highest cross-task classification accuracy, while utilizing only a small number of highly predictive features. The very low permutation tests p-value also demonstrate the fact that universal workload-related features were successfully identified, while overfitting bias was avoided. This fact, paired to McNemar's test additional evaluation, further demonstrated the efficacy of the proposed approach (with significant higher performance than the validation comparisons).

Regarding the overall performance results, the framework proposed attained cross-task accuracy of 0.94, with single task classification of 0.83 for Mental Arithmetic and 0.97 for n-Back task. The deviations in performance could be related to the fact that the selection of the features employed was based on the cross-task classification accuracy and thus the degree to which they relate to the mental load cognitive functions of individual tasks is not clear. In this regard, adding individual task-related features might enhance within-task performance [162], however we opted for the identification of workload related features that present high accuracy with global properties (i.e., independent of the task employed), thereby avoiding (task-related) selection bias. From this standpoint, cross-task classification is not a finite solvable problem with satisfactory accuracy eluding the majority of studies (TABLE 4.4). Although the tasks in the studies presented were not identical, WM manipulations as a result of mental load were reported in all of them suggesting a vague comparison. As such, the overall highest cross-task accuracy in workload classification was obtained in our previous work [162] demonstrated 0.87 accuracy, with [152] marginally improving performance (0.89). Furthermore, FC-based attributes were only employed in our previous study as distinctive features, whereas most cross-task research analyses utilize the power estimated from each frequency band. Nevertheless, in that work the brain activation in source space was evaluated by calculating the Pearson Correlation between the different brain regions, thus providing comparisons with other widely adopted fMRI studies. Alternatively, in this Chapter the FC in time-domain is estimated by implementing a more effective brain network providing insights as to relationship between the different brain locations information is transmitted. More importantly, the fusion of relative PSD and FC features not only displays the highest (so far) cross-task classification performance, but also

illustrates the capability of the proposed framework for extensive interpretation of the selected brain signal characteristics and their implications in neuroscience.

TABLE 4.4. COMPARISON WITH CROSS-TASK WORKLOAD CLASSIFICATION STUDIES

Study	Features	Method	Tasks	Accuracy
(Gevins et al., 1998) [151]	PSD	ANN	Verbal n-Back, spatial n-Back	0.83
(Baldwin and Penaranda, 2012) [148]	PSD	ANN	Reading span, spatial n-Back, Sternberg	0.45
(Walter et al., 2013) [149]	ERS/ERD	SVM	Go/no-go, verbal n-Back, reading span, algebra	0.54
(Ke et al., 2015) [150]	PSD	SVM	verbal n-Back, spatial n-Back	0.29 (†)
(Dimitrakopoulos et al., 2017) [162]	FC	SVM	Spatial-temporal n-Back, Arithmetic	0.87
(Zhao et al., 2018) [153]	Electrophysiological signals (ECG, EOG, RSP, GSR, PPG)	SVM	Anomaly detection in images	0.54 (†)
(Zhang et al., 2019) [152]	Spatial, spectral, temporal EEG	R3DCNN	Spatial n-Back, Arithmetic	0.89
The proposed framework [172]	PSD and FC	SVM	Spatial-temporal n-Back, Arithmetic	0.95

Note: (†) indicates non-binary classification (i.e. more than 2 classes).

Although, most cross-task mental load ML research works implement classification training on one task and then test on the other(s) [150], [152], [153], conversely in the proposed approach the FS and classification procedures utilized data from the two tasks concurrently. The motive behind this was to illustrate a universal (common to the two tasks) workload related brain mechanism, since task-specific attributes might be present if FS and classification were implemented separately in each task), thus providing unsatisfactory performance. This is supported by evidence that suggest high variability of the feature values when training and testing was done under different tasks/conditions [148], [149], [153], [173]. However, to assess the task-independent nature of the features subset employed, additional validation was performed by utilizing data only from one task for training and then estimate the classification accuracy by testing on the other. Notably, the obtain validation accuracy exceeded the performance of most previous studies (TABLE 4.4), being comparable to [152], although significantly reduced as against the proposed framework (0.86 *vs* 0.94).

As far as the features that comprised the optimal subset are concerned, FC features attained high classification performance excluding relative PSD ones, while when the opposite was examined (employing only PSD features) classification performance significantly declined. This manifested in both cross- and within-task classification displaying analogous (low) performance to the majority of cross-task ML studies that only employ spectral power attributes. From this standpoint, it can be inferred that by excluding PSD from the overall FS and classification a better interpretation could be achieved as only the most predictive characteristics would be considered. Nevertheless, the fusion of spectral power with FC

improved the classifier's performance indicating that feature combination acts as an efficient separator when mapping the characteristics employed in higher dimensions. Alternately, implementing FS separately in PSD and PLI and then comprise a new ranked set by their combination, could in theory improve performance. However, the RFE-CBR calculates the correlation bias as a secondary step in each repetition, thus by removing part of the full set the resulting subset might incorporate highly correlated variables rendering classification unreliable and feature interpretation questionable [174]. It is worth noting that, despite the fact that PSD attributes presented similar workload-related trends across tasks, 7 out of 148 connectivity features exhibited discrepancies from low to high workload in the two tasks (i.e., increasing in PLI value in one task, while decreasing in the other). As such, to evaluate whether these connections were related to task-specific information, the proposed framework was also utilized excluding these features. In this respect, the classification accuracy results presented an insignificant decrease (less than 1%), therefore indicating that divergence between the two tasks might be due to PLI subject variability.

4.5.2 Task-independent Spectral Features

The cognitive aspects involved in the n-Back and Mental Arithmetic tasks have been consistently reported to entail equivalent neural processes with regard to mental load [164], [175]–[177]. Specifically, WM overall capacity governs the cognitive functions of storage and retrieval of information, which in the tasks employed relate to the Mental Arithmetic additions (in respect of the carries involved) and the letters presented in the n-Back task previous trials. As such, in the high workload level the memory maintenance requirements were higher due to the sustained attention for longer periods, rendering individuals to be more focused, while ignoring previous trials cues as irrelevant (inhibitory mechanisms). In that scope, the optimal feature subset selected could identify the distinctive underlying WM cognitive processes regulated by mental workload and thus provide high discrimination between the different levels. From this standpoint, the suspected affiliation between the indicative features and mental load levels are elaborated below.

First and foremost, θ frequency frontal majority implies divergences in the WM mental load requirements related to the task difficulty. This is in line with previous research works reporting frontal θ alterations in high workload levels [138], [148], [162]. In this regard, the increment in frontal θ power is systematically described in workload-related studies, especially in WM cognitive processing in memory/arithmetic related tasks [138], [148], [177], [178], which is also present in the relative PSD features selected. On the contrary, PLI FC features did not display a clear trend with 61% of the edges involving frontal sites (14 out of 23) declining. In this context, it should be noted that PSD and PLI are not affiliated in terms of value or nature, as PLI calculated the phase synchronization between the different electrode locations which can intensify in spite of power-related decrements. On this premise, the overall results demonstrated a consistent reduction in most parietal θ band connections. These limited θ band synchronization could be suggested to be a result of task-dependent reactive control [179] and WM load manipulations due to elongated memory periods [180], [181]. In a similar fashion, α band features implicated in posterior locations display values inversely correlated to workload level. These results are corroborated by relevant studies reporting parietal and occipital α power to be reducing in high demanding tasks [138], [144], [148], [177]. This was also observed in FC features, although only in occipital locations. While α decrements have being associated to WM and visual attention elements [144], [182], [183], synchronization augmentation in α band has been related to cognitive inhibition in mental

load information handling [142], [183], [184]. This is further supported by the fact that increases in connectivity due to high workload was noted for all locations, with the exception of occipital sites. With regard to the δ band, a frontal predominance was observed in both PLI and PSD features as one might expect due to the workload-related memory and attention demands in both tasks [185], [186]. Even though, the small number of discriminative PSD features cannot definitely provide indications of an absolute decline (3 out of 5), FC values displayed a high workload decrease in PLI strength in all scalp locations. This is supported by recent studies that also report δ band decrements in memory-related and cognitive control characteristics attributed to mental load [187]–[189]. Last but not least, although β connections were the second most frequent in the optimal subset, no PSD feature was included by the RFE-CBR FS. In this regard however it should be noted that β band power values frequently illustrate little to no significant differences between different task complexity [190], [191]. On the contrary, the decreased β band FC consistent in most PLI connections has been suggested as workload indicator in relevant studies [161], [162]. Moreover, β desynchronization in frontal regions is also reported to be implicated in memory and attention demands [139], [192], [193], with β reduction related to task-related inhibition as a product of mental load [194], [195].

4.5.3 Theoretical Implications and Future Work

Even though the features selected by the proposed framework exhibit high classification performance, their interpretation should be taken with a grain of salt. The reason being that the FS process illustrates features' significance (in terms of classification accuracy) and not the degree in which they relate to the workload cognition. As such it is not clear if the features incorporated in the optimal set reflect underlying mental workload neural substrates or are included to minimize data unrelated noise and thus enhance performance [196]. From this perspective, only a small number of PSD features presented statistically significant differences (as presented in the validation results section), while FC displayed no significant value deviations with regard to the workload levels. This could be due to the variability between individuals as mental load frequency-related attributes can present a large deviation between subjects because of differences is attention, effort or task expertise [177], [186], [197].Despite that however, the optimal subset generated by the FS process illustrated high accuracy regardless of the classifier utilized (as indicated in Validation results TABLE 4.3), underlying the features quality as robust indicators of elicited high workload in a task-irrelevant manner. In addition, the combination of spectral power and FC (even though differ in both nature and behavioral motifs), present the opportunity to confirm the associations between the different cognitive functions affiliated with mental workload.

In light of all the above, the FC features included in the optimal subset present a common pattern that links cognitive control of inhibitory functions and mental load irrespective of the task employed [142], [179], [182], [183], [189], [194], [195]. In fact, the global mechanism common to both tasks indicated by the selected features of the adopted framework is mental inhibition, as individuals prolonged mental high-efficiency effort is made not to remember previous cues (i.e., such as carriers or n-Back objects), but tuning out irrelevant or distracting stimuli. In practice, the tasks implemented in this Chapter demand high WM resources, since in the high workload level involve continuous mental updating remembering previous trials stimuli and excluding others prior to them. This could also explain the universal perspective of induced mental fatigue from long periods of multi-tasking, as individuals are required to focus on each task, while at the same time disregarding all others in a constant manner.

Despite the fact that the experimental design employed here did not consider irrelevant or distraction trial effects since inhibitory control was not in particular under investigation, evidence from previous studies indicate the interference effects of mental load and its relevance to WM [198]–[201]. Consequently, it can be presumed that FC constitutes a vigorous approach for information relationship assessment between different scalp locations, while effectively providing workload-related cognitive aspects. Nevertheless, the number of relevant works that include FC analysis is very limited and thus verification of this premise is extremely difficult. Future works that incorporate features of different modalities and multivariate characteristics, as well as their fusion would provide convincing proof of global neural substrates that regulate cognitive workload.

4.6 Conclusion

In this Chapter, the fusion of EEG diverse EEG spectral characteristics has been implemented to classify between two distinct workload levels in a cross-task cognitive load classification scheme. In detail, the PLI FC and the relative PSD were calculated in four EEG frequency bands, while being incorporated a multi-dimensional feature set and then the discriminative power of each feature was assessed in regard to mental workload discrimination. To that end, a FS and classification utilized demonstrated high overall performance, being able to detect a small subset of highly discriminative attributes. Additional analysis highlighted further the robustness of this approach indicating global characteristics (common to both tasks) and deciphering the degree on how they relate to mental workload. The produced results promote the efficiency of feature fusion on cross-task workload classification, while assessing the underlying mental load neural substrates.

Chapter 5

5.1 Multi-Level Workload Classification in Real-World Scenarios

To date, mental load assessment –although well established in controlled cognitive tasks– is not considered suitable for real world implementation due to diverse neuronal information present in continuous cognitive processing. In this regard, interfaces that incorporate virtual reality provide realistic environments for lifelike interactions, offering immersive experience and perception. In this Chapter, an analysis of the modifications of pilots' brain networks has been done, with regard to three different workload levels in computer screen (2D) and virtual reality (3D) environments. In detail, the FC in EEG source space of different brain waves has been employed to assess the reorganization of the brain networks in terms of topology and efficiency. Moreover, the individual FC features has been utilized in a FS and classification scheme to detect the mental load topological differences between the two environments. The framework presented here, identified a frontal alpha band majority in both simulation interfaces with common connections existing in all complexity levels. On the contrary, theta and beta bands displayed dissimilarities in the local and global efficiency between the two environments. These results allow for an efficient evaluation of the cognitive substrates that govern workload in real-world scenarios.

5.2 Introduction

Cognitive processing refers to the individual's ability to complete mental tasks by the assimilation of information gained through experience and perception. The increasing demands of a task (which can contribute to poor performance), known in the literature as high mental workload, necessitate additional brain resources and in combination with an individual's limited mental capacity lead to deteriorated performance [202], [203]. In fact, prolonged task with a high workload has been identified as the primary cause of mental fatigue [204]. Continuous workload monitoring, for example, would greatly benefit operations such as aircraft piloting and air traffic control, with the goal of enhancing the efficiency and safety of everyone involved [204]–[206]. Monitoring mental stress with the ultimate goal of maximizing human performance and reducing human errors in real-world settings is consequently of great interest in the nascent field of neuroergonomics [207], [208]. The electroencephalogram (EEG), a non-invasive measurement of electrical activity generated by the brain, has been routinely used in mental workload research with promising results [209], [210]. As a result, convergent data demonstrate considerable changes in alpha and theta band power, which may serve as the most important indications for workload estimate [211]–[213], while numerous research indicate links between delta, beta, and gamma band oscillations and mental load [214]–[216]. In fact, Spüler et al. [217] demonstrated workload-related effects in the theta, alpha, and lower beta frequency bands at the parieto-occipital electrode locations by varying the difficulty of arithmetic addition tasks. In a relevant working memory paradigm [218], increased cognitive load was linked to lower alpha and beta band power at all midline electrodes, with extra alpha decrements in centroparietal sites. Borghini et al. [219] also developed a workload index of difficulty level in a driving exercise based on theta and alpha power spectra. In more demanding situations (equivalent to higher workload), theta band increments over prefrontal areas and alpha band decrements in parietal areas were found. Given that a wide range of brain areas have been linked to mental workload, it's plausible to assume that the neural mechanisms underlying workload involve

connection alterations, in line with the recent conceptualization of the human brain as a large-scale network of interconnected regions [220].

To explicate the complicated interactions between numerous cerebral regions involved in human cognition, as well as to understand the functional architecture and reconfiguration of the brain, an increasing number of research works are using network modeling and graph theory methodologies [14], [221], [222]. In this regard, network metrics such as global and local efficiency (measures of a network's information transfer efficiency) have been shown to be strongly affected by varying workload levels [223]–[225]. Based on phase coherence networks in an n-Back task, Kitzbichler et al. [223] showed a constant increase in global efficiency and a decrease in local efficiency in the alpha, beta, and gamma frequency regions. Increased local and global efficiency in brain networks of the delta, theta, alpha, and beta bands was found under high workload condition in another study comparing network reorganizations between low and high load conditions [224]. Furthermore, Klados et al. [225] found increased local and global efficiency during arithmetic tasks with increasing difficulty in the delta, theta, and alpha bands, with the frontal and fronto-parietal regions showing the most evident connections. Functional brain networks have also been utilized in conjunction with classification algorithms to distinguish between various cognitive states [226]–[228]. Indeed, classification may not only give more comprehensive knowledge of the global input data by finding hidden paths of network reorganizations under varied workload levels, but it can also achieve high performance in workload monitoring for possible real-world applications. In this scope, previous research proved the viability of using cross-frequency coupling between high alpha and theta bands to classify a 5-level mental arithmetic task's mental effort [221]. Moreover, extension of this work on single-task workload classification to cross-task mental workload evaluation using multi-band EEG functional connectivity achieved satisfactory classification accuracy [229]. Further examination revealed that the most discriminative connectivity patterns were found in the frontal lobe of the theta band, but beta-band connections weakened as task demands increased. Although the concept of mental workload is widely recognized in the neuroergonomics community, real-world applications are still a long way off. This is largely due to the fact that most relevant studies use well-controlled cognitive tasks, with only a small number of them using virtual reality (VR)-based simulation experiments [230], [231].

Given the documented differences in neurological foundation between 2D and 3D presentations [232], [233], it's possible that functional brain connection reorganizes differently in the two simulated settings. Using a combination of brain network analysis and connectivity-based workload classification, this chapter aims to: a) reveal the neural basis of mental workload in a flight simulation experiment; and b) delineate the convergent and divergent network changes induced by various workload levels between standard computer screen (2D) and virtual reality (3D) environments. To that end, in this chapter, the effects of the two simulated environments in multi-level workload circumstances were analyzed using cortical connections and ML. Specifically, a thorough investigation on changes in functional brain networks using graph theoretical metrics and the brain connectome was made, in a FS and classification framework, while at the same time the relationship and consequences of spatial and frequency network characteristics on pilot mental strain was assessed.

5.3 Materials & Methods

5.3.1 Subjects

In this work, 33 male students and staff members from the National University of Singapore (NUS) were recruited (mean age = 23.2 ± 2.2 years). All participants had normal or corrected-to-normal vision and no previous experience in simulated aircraft operation. Participants were excluded if they met the criteria: history of psychiatric, neurological or cognitive impairment, sleep disorder or childhood history of ADHD and long-term medication intake. Participants who did not have a full night of sleep (> 7 h) for two nights before the experiment were re-scheduled. In addition, on the day of the experiment, participants were asked to avoid strenuous exercise 2 h prior to the recordings as well as consumption of caffeine or alcohol. Signed written consent form was obtained from all participants after the description of the experimental protocol. The study was approved by the Institutional Review Board and was conducted in accordance with the Declaration of Helsinki.

5.3.2 Experimental Design

All subjects performed two flight simulation sessions using two interfaces: a) a standard session using a 24-inch LED computer screen (CS) (Model: U2412Mb, Dell Inc., Texas, USA), and b) a virtual reality session (VR) via the Oculus Rift headset (Model: Development Kit 2, Oculus VR, California, USA) (**Figure 5.1**), both depicting the view in the aircraft cockpit. The flight simulation software was the Microsoft Flight Simulator X using a joystick (Model: Extreme 3D Pro, Logitech, Switzerland). Before the EEG recordings, a training session was performed in both environments until satisfying operational control of the aircraft was accomplished. Each flight simulation session consisted of three stages with increasing task difficulty (Stage 1, Stage 2 and Stage 3, abbreviated to *S1*, *S2* and *S3*), designed to induce different levels of mental workload (**Figure 5.1 (b)**). Subjects were instructed to fly the aircraft following a predefined flight route for each stage (green dashed line in **Figure 5.1 (b)**), which lasted 4 min with an inter-stage interval of 1 min for experiment reconfiguration (**Figure 5.1 (c)**). In detail, the three stages included: *S1*, autopilot state, in which no aircraft operations was required (minimum workload level); *S2*, manual operation state, in which subjects were required to manually operate the aircraft to complete seven consecutive turns while maintaining an altitude of 3,000 feet (medium workload level); and *S3*, manual operation with aircraft failure, in which *S2* was performed while additionally the aircraft was under the influence of aileron and rudder failures, increasing control complexity (maximum workload level). The order of the sessions and stages were counterbalanced across the subjects.

To confirm that the task difficulty manipulation was effective, participates were also requested at the end of each session to rate the perceived task difficulty of each stage in a scale of 1 – 7 with higher scores indicating more difficult conditions. Moreover, to examine the users' experience differences, each subject was asked to fill a short questionnaire at the end of each task. The questionnaire comprised 10 questions about system usability, 6 questions about perceived usefulness, 6 questions about perceived user-friendliness, 7 questions about satisfaction and 6 questions about overall reaction. All answers were provided in a scale of 1 – 5, with higher score indicating better user experience at each category.

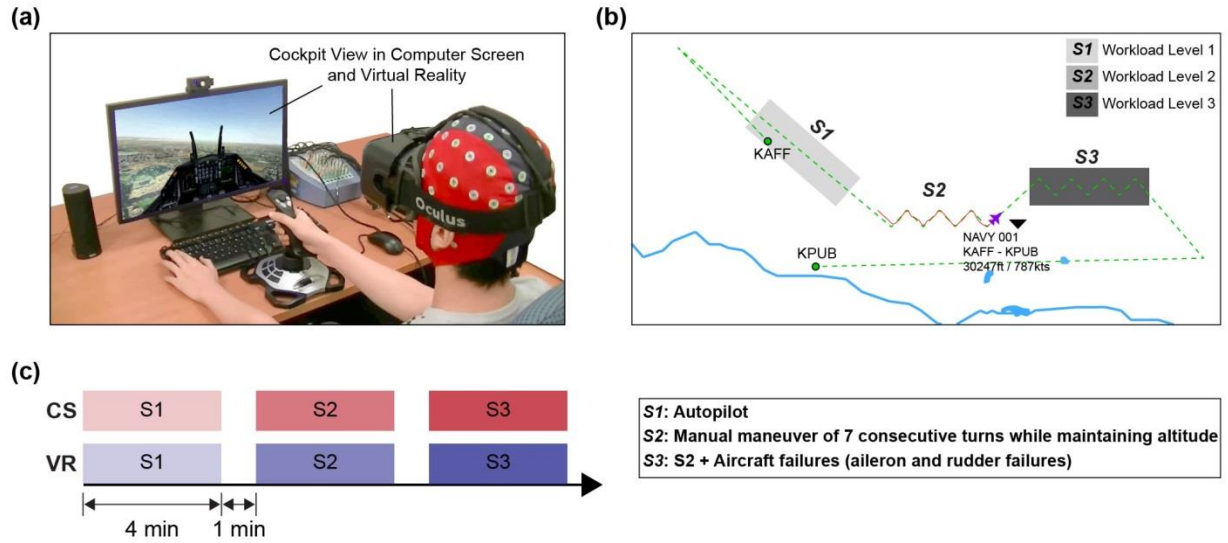


Figure 5.1. The experimental protocol employed. (a) The view of the cockpit as shown in Computer Screen (CS) and Virtual Reality (VR). (b) The flight route of three workload stages (S1, S2, and S3) in green. The predefined flight route is shown in red (c) Each participant was required to complete one session of each simulation tasks (i.e., VR and CS) for the three stages.

5.3.3 Data acquisition and preprocessing

Continuous EEG data were recorded in high-resolution from 64 Ag/AgCl scalp electrodes using the international standard 10-20 system (ASA Lab, ANT B.V., Netherlands). Bipolar electrooculogram (EOG) signals were also acquired from electrodes placed above and below the right eye (VEOG) and at the outer canthi of both eyes (HEOG). Impedance of electrodes was kept below 10 kΩ throughout the duration of the experiment. A band-pass filer (0.5 – 70 Hz) was applied for anti-aliasing and additionally a 50 Hz notch filter was used to remove main interferences. Raw EEG signals were digitized at a sampling rate of 256 Hz, filtered applying a 0.5 – 40 Hz FIR band-pass filter and referenced to the average of both mastoids (M1 and M2). Data were then de-trended and baseline adjusted. To exclude eye-blink artifacts, independent components analysis was employed and the components showing high correlation to the HEOG and VEOG were discarded [226]. More details about EEG pre-processing can be found in Chapter 2. Data from four subjects were removed due to high noise contamination or error recordings, thus data from 29 participants remained for further analysis. All preprocessing steps were implemented in EEGLAB 13.6.5 [234] in MATLAB R2017b (The MathWorks Inc, Massachusetts, USA). A flowchart of the data analysis framework is shown in **Figure 5.2**.

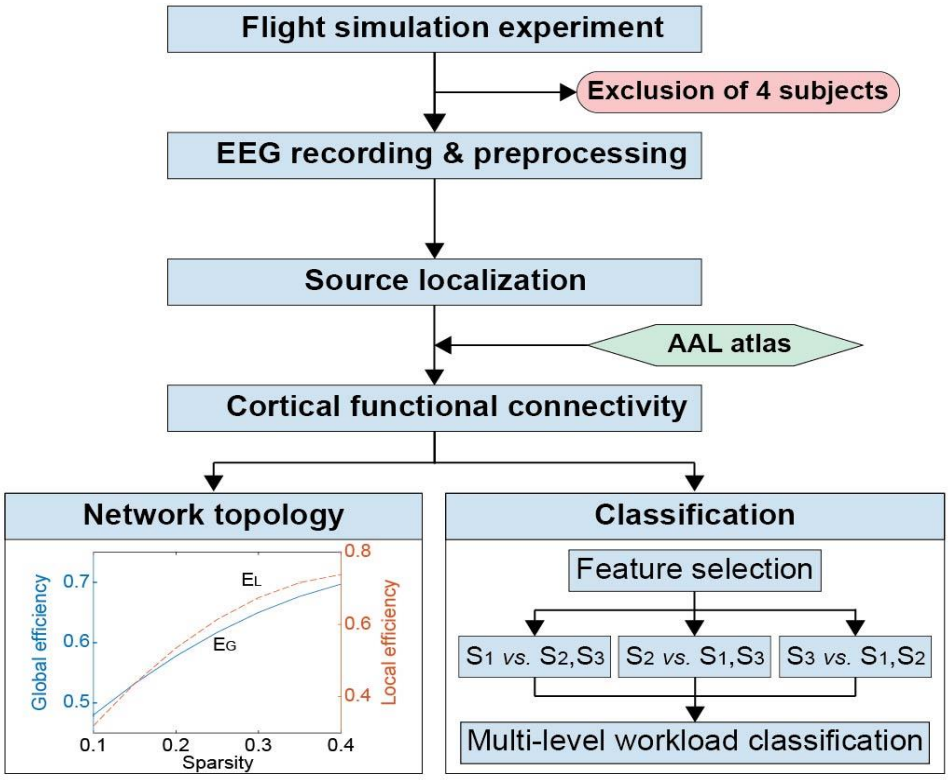


Figure 5.2. A schematic of the workflow of the proposed framework

5.3.4 Cortical connectivity

Aiming to localize mental workload effects and evaluate the reproducibility of the main results, source localization was employed to reconstruct the cortical signals. In detail, the EEG time series for each session and workload level were divided into delta (δ, 0.5–4 Hz), theta (θ, 4–7 Hz), alpha (α, 8–12 Hz), beta (β, 13–30 Hz) and gamma (γ, 30–45 Hz) frequency bands and then the intracerebral electrical sources corresponding to the recorded surface signals were approximated by applying exact low-resolution electromagnetic tomography (eLORETA) [235]. eLORETA utilizes the MNI152 template [236] to calculate the source space current density (A/m^2) in 6239 voxels at 5 mm spatial resolution with the 3-D solution space restricted to cortical gray matter (more details about source localization can be found in Chapter 2). Then, the solution space calculated was parcellated into 116 regions of interest (ROIs) according to the previously validated Automatic Anatomical Labelling atlas AAL-116 [237]. From the 116 ROIs, 36 corresponding to cerebellum and sub-cortical regions were removed to avoid possible depth bias [238], with 80 brain regions remaining for further analysis (TABLE 5.1). Source localization was carried out with the LORETA-Key software (http://www.uzh.ch/keyinst/loreta.htm). To calculate the functional connectivity between each pair of ROIs the Phase Locking Index (PLI) [239] was employed (please refer to Chapter 1).

Data were divided into epochs of 60 sec with 50% overlap and one PLI network was estimated for each epoch and frequency band (more details about PLI calculation can be found in Chapter 1).

TABLE 5.1. THE NAMES AND THEIR CORRESPONDING ABBREVIATIONS FOR THE AAL ROIS

Region name	Abbreviation	Classes
amygdala	AMYG	Paralimbic
orbitofrontal cortex (superior)	ORBsup	Paralimbic
angular gyrus	ANG	Association
orbitofrontal cortex (inferior)	ORBinf	Paralimbic
anterior cingulate gyrus	ACG	Paralimbic
orbitofrontal cortex (medial)	ORBmed	Paralimbic
calcarine fissure	CAL	Primary
orbitofrontal cortex (middle)	ORBmid	Paralimbic
cuneus	CUN	Association
lobule	PCL	Association
fusiform gyrus	FFG	Association
parahippocampal gyrus	PHG	Paralimbic
gyrus rectus	REC	Paralimbic
postcentral gyrus	PoCG	Primary
heschl gyrus	HES	Primary
posterior cingulate gyrus	PCG	Paralimbic
inferior frontal gyrus (opercular)	IFGoperc	Association
precentral gyrus	PreCG	Primary
inferior frontal gyrus (triangular)	IFGtriang	Association
precuneus	PCUN	Association
inferior occipital gyrus	IOG	Association
rolandic operculum	ROL	Association
inferior parietal lobule	IPL	Association
superior frontal gyrus (dorsal)	SFGdor	Association
inferior temporal gyrus	ITG	Association
superior frontal gyrus (medial)	SFGmed	Association
insula	INS	Paralimbic
superior occipital gyrus	SOG	Association
lingual gyrus	LING	Association
superior parietal gyrus	SPG	Association
middle cingulate gyrus	MCG	Paralimbic
superior temporal gyrus	STG	Association
middle frontal gyrus	MFG	Association
supplementary motor area	SMA	Association
middle occipital gyrus	MOG	Association
supramarginal gyrus	SMG	Association
middle temporal gyrus	MTG	Association
temporal pole (middle)	TPOmid	Paralimbic
olfactory	OLF	Paralimbic
temporal pole (superior)	TPOsup	Paralimbic

5.3.5 Network Topology Analysis

Before the network analysis, a threshold was applied to each network matrix to convert it to a fixed sparsity level to ensure the equivalent wiring cost of each subject. The sparsity value represents the ratio of the actual edges to all possible network edges. A sparsity range (10% - 40%) was applied to both keep the reachability of the network and allow prominent small-world characteristics. Aiming to offer a clear and direct physical meaning to the concept of small-word properties in terms of information flow [19], [240], global efficiency (E_G) and local efficiency (E_L) of brain networks were estimated (more details about networks metrics calculation can be found in Chapter 1).

The graph theoretical analysis was implemented with the Brain Connectivity Toolbox [14]. To avoid multiple comparisons at each individual sparsity threshold and to reduce the dependency of the significant alterations on the arbitrary selection of a single threshold, integrated network metrics were calculated over the predefined sparsity range (corresponding to the area under the curve measurement) [19]. Then, statistical analyses were performed on the integrated network metrics as described below.

5.3.6 Network-based Feature Selection and Classification

On the basis that variability in cognitive states display specific discriminative aspects in functional brain networks, the subsequent workload level classification was employed by considering the network edge weights as feature vectors. For this purpose, the PLI values were averaged for each participant and workload condition ($80 \times (80-1) / 2 = 3160$ unique values) across each of the five frequency bands epochs resulting in $5 \times 3160 = 15800$ features per workload level for each participant. Subsequently, FS was utilized to exclude redundant features and enhance the performance of the classifiers, while avoiding possible overfitting due to the large number of features compared to the number of samples. Hence, Recursive Feature Elimination with Correlation Bias Reduction (RFE-CBR) [241] based on linear support vector machines (SVM) was used to calculate the dependencies among features and recursively reject minimal evaluated features. Specifically, RFE-CBR is a backward elimination method that evaluates the features' influence and removes them in succession from the feature space using the coefficients obtained from the SVM models. Moreover, it detects the highly correlated features to mitigate correlation bias [242], subsequently creating a ranked feature space based on the features' significance (more details regarding RFE-CBR can be found in Chapter 4 where it was also implemented). Here, RFE-CBR feature selection was applied on each workload level as a binary process, resulting in three feature ranked sets.

Classification was performed by a random subspace ensemble method with Linear Discriminant Analysis (LDA) as the base classifiers [243]. The random subspace ensemble method deconstructs the input data into random subsets of feature space to implement individual LDA classifications which are trained separately and then a collective decision is made by majority vote. To identify the number of features to be included in the multi-class feature set, binary classification was performed utilizing the *one-vs-all* strategy for each workload level (FS1: S1 *vs* S2, S3; FS2: S2 *vs* S1, S3; FS3: S3 *vs* S1, S2). The feature subset with the highest overall classification accuracy was determined as the optimal for workload level discrimination, estimated by including one-by-one the ranked RFE-CBR features in succession starting with a null feature set. The multi-class classification feature set was then generated by merging the individual workload level feature subsets. In order to attain maximum performance, the number of classification learners and subspace dimensions were tested with different configurations. Final results

were estimated by conducting 100 repetitions of 10-fold cross validation in binary and multi-class classification, thus minimizing the possible effect of the training set variability.

Additionally, permutation tests was performed to examine the classification accuracy significance (as different from random level), which is appropriate for small sample data with unknown population distribution [244]. Thus, classification was executed 1000 times with random class labels permutations to calculate the distribution of classification accuracy. Probability p was estimated as the number of classifiers trained on the randomized samples that outperformed the classifiers trained on the original samples divided by the number of total permutations.

5.3.7 Statistical Analysis

To examine the workload effect on the perceived task difficulty, repeated-measures two-way ANOVA was used with #1 factor being workload (comparing the task difficulty across the three workload levels), and #2 factor being task (comparing VR and CS environments). Moreover, subjective user experiences reported in the questionnaire (referring to system usability, perceived usefulness, perceived user-friendliness, satisfaction, and overall reaction) between the two interfaces were compared using one-way ANOVA. To evaluate the mental load effects on the network properties between the two tasks, repeated-measures two-way ANOVA was applied on the integrated network metrics, with workload and task being the main factors, and workload-by-task considered as interaction. Significance value was set as $p < 0.05$. Then, post-hoc analysis was carried out for the significant ($p<0.05$) interactions using a paired t-test. All statistical analyses were implemented in SPSS 21 (IBM, New York, USA).

5.4 Results

5.4.1 Behavioral Results

A significant workload effect was observed in the perceived task difficulty in the two-way repeated measures ANOVA test ($F_{2,56} = 259.993$, $p < 0.001$), confirming the efficacy of the experimental design in inducing different levels of mental workload. However, main task effect ($F_{1, 28} = 3.045$, $p = 0.092$) and interaction ($F_{2, 56} = 0.877$, $p = 0.422$) did not show any Significant result. Moreover, regarding the user experience, the system usability measure had a significantly higher score in CS compared to VR ($F_{1,28} = 4.409$, $p = 0.045$), but no significant task effect was found in the other four measures of user experience ($p > 0.05$).

5.4.2 Network Topology

A two-way repeated measures ANOVA was additionally employed to examine the topological properties between both groups across the different workload levels (TABLE 5.2 and **Figure 5.3**). In detail, significant main workload effects were detected for E_G in δ ($F_{2,56} = 3.397$, $p = 0.041$), α ($F_{2,56} = 7.708$, $p = 0.001$), and β ($F_{2, 56} = 4.712$, $p = 0.013$) frequency bands, as well as for E_L in α ($F_{2,56} = 20.773$, $p < 0.001$) and β ($F_{2,56} = 11.175$, $p < 0.001$) frequency bands. Moreover, E_G showed significant main task effects in β band ($F_{1, 28} = 7.895$, $p = 0.009$), while E_L in θ ($F_{1,28} = 8.098$, $p = 0.008$) and β ($F_{1,28} = 6.394$, $p = 0.017$) bands. Remarkably, a significant interaction effect was found for E_G in β band ($F_{2,56} = 4.886$, $p = 0.011$). The post-hoc analysis showed that the significant interaction was based on a significant decreasing trend

of E_G ($F_{2,56}$ = 7.512, p = 0.001) with increasing workload levels in the CS environment compared to a non-significant alteration in the VR environment ($F_{2,56}$ = 1.563, p = 0.218).

TABLE 5.2. RESULTS OF THE STATISTICAL ANALYSIS OF THE NETWORK METRICS FOR EACH FREQUENCY BAND

Frequency band	Metrics	Workload $F_{2,56}$ (p-value)	Task $F_{1,28}$ (p-value)	Interaction $F_{2,56}$ (p-value)
δ	E_G	**3.397(0.041)**	0.000(0.987)	0.653(0.524)
	E_L	2.496(0.092)	0.467(0.500)	1.301(0.280)
θ	E_G	1.919(0.156)	0.385(0.540)	0.694(0.471)
	E_L	0.091(0.913)	**8.098(0.008)**	0.257(0.775)
α	E_G	**7.708(0.001)**	1.733(0.199)	0.425(0.616)
	E_L	**20.773(< 0.001)**	1.252(0.273)	0.084(0.919)
β	E_G	**4.712(0.013)**	**7.895(0.009)**	**4.886(0.011)**
	E_L	**11.175(< 0.001)**	**6.394(0.017)**	1.757(0.182)
γ	E_G	0.948(0.394)	1.606(0.216)	0.948(0.393)
	E_L	0.369(0.693)	0.418(0.523)	0.025(0.975)

Note: E_G indicates global efficiency, E_L indicates local efficiency. Bold fonts indicate significant effects.

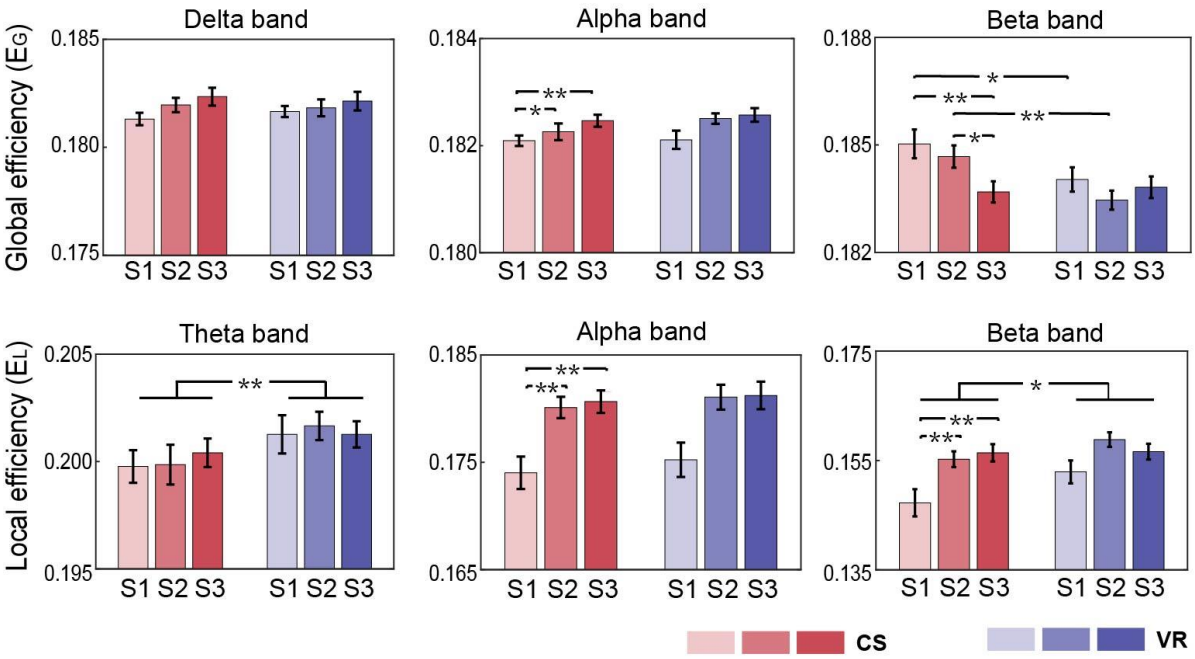


Figure 5.3. Post-hoc analysis with regard to global and local efficiency with Bonferroni correction for multiple comparisons. Each bar represents the mean± the standard error. * $p < 0.05$; ** $p < 0.01$.

5.4.3 Classification Performance

The classification on workload level was performed based on the features extracted from the PLI connectivity. The workload level feature subsets that provided the highest individual level accuracy were included in multi-class classification (TABLE 5.3). In detail, regarding the CS environment the best accuracy was 0.73, 0.87 and 0.85 using 20, 22 and 12 features for FS1, FS2 and FS3 levels respectively. Moreover, the multi-class feature set comprising their combination, incorporated 53 features (1 common) among 55 unique ROIs and achieved 0.82 ($p < 0.001$) accuracy (sensitivity = 0.82, specificity = 0.91), while 26 subspace dimensions and 30 learners were estimated as optimal parameters. Utilizing the same method in the VR environment, 0.78, 0.87 and 0.89 classification accuracy was obtained for FS1, FS2 and FS3 using 12, 21 and 19 features, whereas the multi-class feature subset comprising 50 features (2 common) among 55 unique ROIs provided 0.82 ($p < 0.001$) multi-class accuracy (sensitivity = 0.91, specificity = 0.82) using 25 subspace dimensions and 30 learners.

TABLE 5.3. CLASSIFICATION PERFORMANC RESULTS

Interface	Workload Level	Features Num.	Accuracy (p-value)	Sensitivity	Specificity
CS	FS1	20	0.73 (p = 0.004)	0.76	0.63
	FS2	22	0.87 (p < 0.001)	0.90	0.82
	FS3	12	0.85 (p < 0.001)	0.86	0.82
	All	53	0.82 (p < 0.001)	0.82	0.91
VR	FS1	12	0.78 (p < 0.001)	0.79	0.78
	FS2	21	0.87 (p < 0.001)	0.85	0.92
	FS3	19	0.89 (p < 0.001)	0.88	0.94
	All	50	0.82 (p < 0.001)	0.91	0.82

5.4.4 Functional Connectivity Characteristics

The selected functional connectivity features on both CS and VR interfaces are presented in **Figure 5.4**, with each feature's color corresponding to a specific frequency band. The majority of the network edges in both environments belong to α frequency band and include frontal regions. In detail, in the CS interface 59% (31 out of 53) connections are detected in α, 19% (10 out of 53) in β, 9% (5 out of 53) in δ and θ and 4% (2 out of 53) in γ bands. Moreover, 53% (28 out of 53) of the connections involve frontal areas ROIs, while 42% of them include ROIs in temporal areas (22 of 53). Moreover, 20 out of 53 features are intra-hemisphere connections and the rest 33 out of 53 are inter-hemispheric.

Likewise, in the VR environment, 62% (31 out of 50) of the connectivity edges are observed in α band, while 16% (8 out of 50), 8% (4 out of 50) and 14% (7 out of 50) of the connections are detected in θ, β and γ bands, respectively. Remarkably, no connections from δ frequency band were chosen by the FS. Most of these edges (52%, 26 out of 50) connect ROIs in frontal areas and 46% (23 out of 50) in temporal ones. In contrast to the CS environment where a predominant inter-hemispheric pattern was distinguished, only 16 out of 50 edges include inter-hemispheric areas, while most connections are linking areas in the same hemisphere (34 out of 50).

The multi-class feature sets for the classification between the different workload levels contained one common feature in the CS and two in the VR environments. In detail, the SMA.R to STG.R connection in α band was considered important by the RFE-CBR FS procedure both in the FS1 and FS2 in the CS environment. Moreover, the bilateral connections of PoCG.L and PoCG.R to PCUN.L in α band were also selected for both FS1 and FS2 in the VR environment. Notably, SMA.R to STG.R and PoCG.L to PCUN.L were selected in both environments.

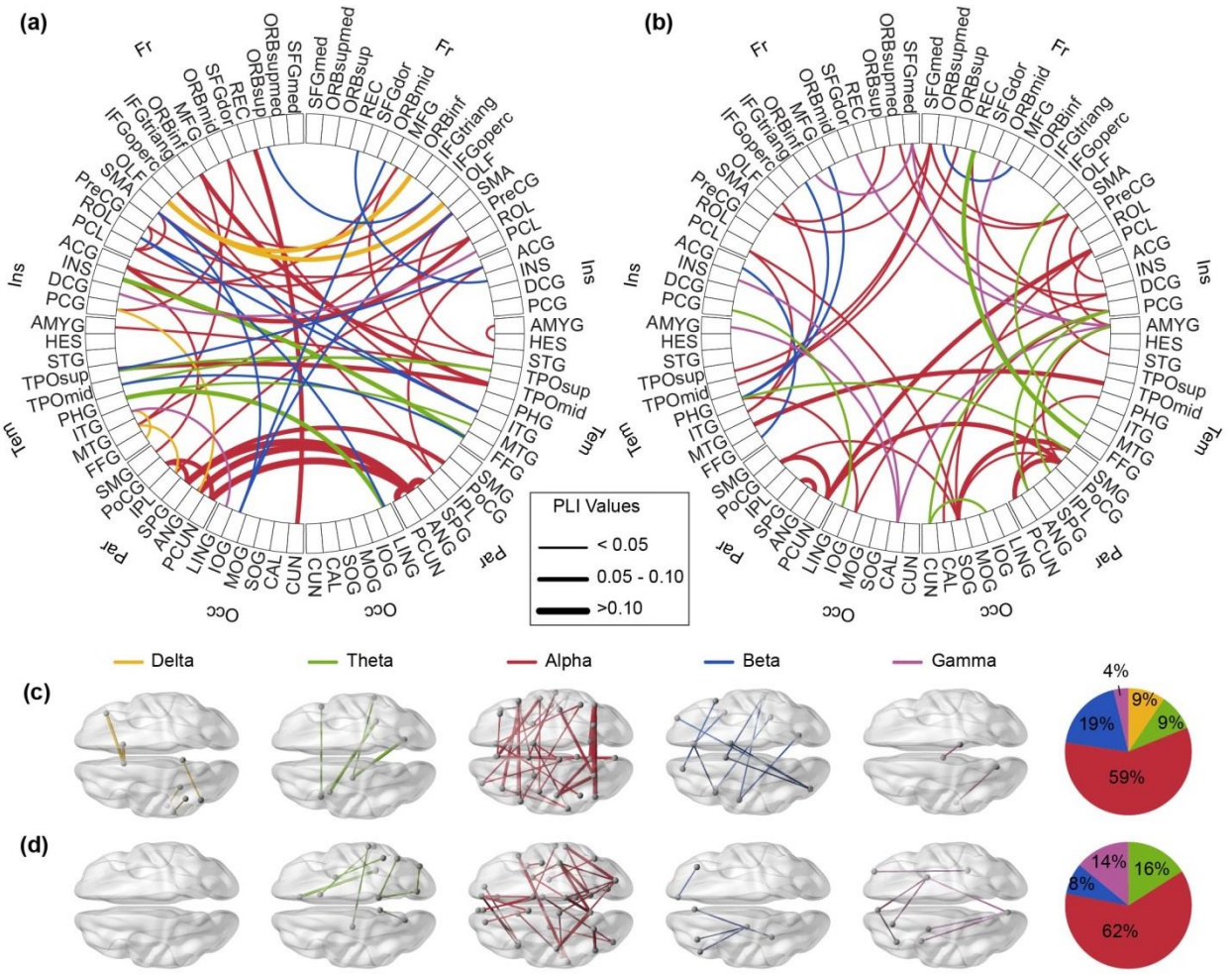


Figure 5.4. The features selected are presented in circular diagrams in the upper part of the figure for (a) CS and (b) VR environments; corresponding brain areas are displayed in the bottom part of the figure for (c) CS and (d) VR per frequency. The frequency band distribution is shown in the pie charts. Each interaction is color coded for each frequency band, with the width corresponding to the PLI connectivity strength.

To examine in depth the workload effect on individual features, one-way ANOVA was employed to obtain indications of the significance of the connectivity changes in the multi-task classification. Thus, 34 out of 53 (among 38 unique ROIS) and 29 out of 50 (among 37 unique ROIS) connections displayed a significant effect ($p < 0.05$) for CS and VR environments respectively (**Figure 5.5**). Specifically, the significant connections in the CS environment were found mainly in α and β frequency bands (23 α, 10 β and 1 δ), contrary to the VR environment, where the significant features contained θ (5), α (17), β (2) and γ (5) band connections. Additionally, mainly frontal (19 in CS, 13 in VR) and parietal (11 in CS, 13 in VR) ROIs were detected for both environments. Despite that for α band edges an overall clear trend could not be determined, all β band connections showed reduced strength with higher workload in both environments as well as frontal areas involvement, while θ band VR edges presented overall increments. Remarkably, γ band connections were only selected in the VR environment, while 4 out of 5 of them included the amygdalae nuclei (AMYG).

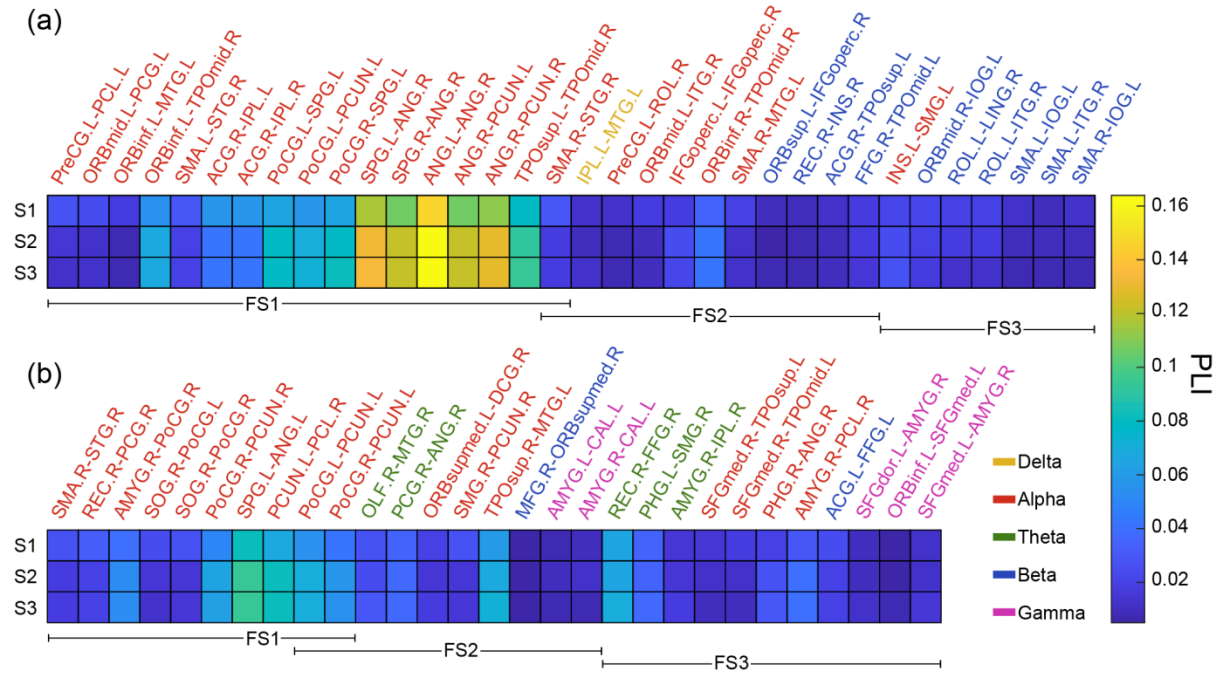


Figure 5.5. The PLI connectivity strength for the significant features selected for (a) CS and (b) VR environment for each workload level. The interaction name color indicates each frequency band. Common FC of each condition are presented via the feature overlapping.

5.5 Discussion

In this Chapter cortical connectivity and network characteristics were analyzed to examine the neural basis of pilot mental workload in 2D and 3D simulated environments. It was found that the different environments led to distinct network reorganizations in the various workload levels, attributed to the variations of the information flow between less-realistic (CS) and close-to-real-world interfaces (VR). Moreover, the connectivity features were utilized in a FS and classification scheme achieving effective discrimination among the three workload levels making possible a consistent analysis of the results to understand their role from Neuroscience point of view. Based on the significant workload effect in the behavioral results regarding the perceived task difficulty, it was confirmed that each session could induce a different mental load level, i.e., the increasing effort needed to fly the aircraft would result in higher workload, since higher attentional and cognitive control were required. Thus, the behavioral results confirm the effectiveness of the experimental design with regard to workload induction, thus providing the basis for the following steps of the methodology proposed to analyze the spatial and frequency characteristics of the functional connections.

With regard to classification, high accuracy was achieved in both environments, which was validated by the very low p-values of the permutation test. Moreover, effective classification was achieved utilizing only a very small number of the available features. These facts combined with the training of classifiers across subjects, the high performance shows the generalization ability of the employed framework, indicating that no overfitting had occurred. Remarkably, the FS and classification procedures use different machine learning algorithms. Taking into account that the aim of this Chapter was to examine the

functional connections and considering the binary nature of SVM, an independent classifier for an unbiased validation of the FS was employed.

5.5.1 Network Topology

As indicated by the network metrics examination, significant differences were found in global and local efficiency. As efficiency is inversely related to path length, the detected increments show an enhanced connectivity between nodes that is shorter path lengths, hence showing a larger degree of information transfer between brain areas. Therefore, increments of global efficiency in δ and α bands along with increment of local efficiency in α and β bands support that brain regions are interacting and cooperating to a larger extent with higher mental workload. The increments of local or global efficiency in various frequency bands are consistent with workload-related literature [223]–[225], as aforementioned in the introduction section, supporting that the brain reorganizes and employs more resources to effectively respond to higher task difficulty [202], [245]. Additionally, differences between the two environments were detected, with CS displaying significantly lower local efficiency in θ and β bands. Moreover, contrary to other frequency bands, global efficiency in β band had a significantly decreasing trend in CS, as well as larger values were observed for CS compared to VR. Thus, the majority of these findings indicate that the VR environment seems to evoke higher mental workload to the participants compared to the CS environment, probably due to the more visually demanding cognitive processes [246], [247].

5.5.2 Workload Dependent Functional Characteristics

The selected features in both CS and VR environments contain functional connections from different frequency bands. Specifically, most of the features correspond to α frequency band and are mainly located in frontal and parietal areas. The ANOVA test further confirmed the prevalence of α band among the significant workload-related features. These observations are to be expected as α band variations in frontal and parietal regions has been associated with workload by numerous related studies [205], [248], [249], although no consistent trend regarding the connectivity PLI values could be detected between the different workload levels. Furthermore, β features displayed reduced functional connectivity strength with higher workload involving frontal areas, in accordance with previous workload-related studies [172], [229]. Moreover, the β band suppression can be attributed to the increment of cognitive control and attention demands imposed by increasing task demands [250], [251].

The brain regions corresponding to the common selected connections between the two environments (SMA.R - STG.R, PoCG.L - PCUN.L) have been related with the mental workload. Specifically, higher task demands have displayed SMA activation [252], [253], indicating visuomotor leaning [254] and attention mediating time-on-task processing [255]. Taking into account that SMA was selected for FS1 in both interfaces also highlights its role in the discrimination between the different cognitive workload levels. Moreover, the right STG activation has been associated with higher workload levels, regulating attention, spatial perception and visual recognition [256], [257]. Similar findings have been reported for PoCG engagement as a result of visual and sustained attention mental load [258], [259]. Additionally, cognitive workload in visuospatial mental operations and task-related attentional demands has been consistently associated with PCUN activation alterations [260], [261].

5.5.3 Environment Dependent Functional Connectivity Reorganization

Despite the similarities regarding the mental workload characteristics in the two environments, clear distinctions in both frequency and spatial aspects were observed. For example, the significant θ band connections showed higher connectivity values with increasing cognitive load in the VR interface, which was frequently displayed in various workload-related works [205], [229], [262]. Still, θ band edges displayed no significant variations between the different workload levels in the CS. This observation could be explained by the higher sense of the subjects' presence and a more active task-related processing in the VR compared to the CS environment that has been shown to significantly affect θ band activity [247], [263]. Moreover, the ANOVA test displayed significant γ edges in the VR interface, mainly containing the AMYG. Despite that this accounts for the spatial navigation and visual representation differences between the two environments [264], the additional stress evoked by the higher workload sessions when subjects were attempting to control the aircraft should also be taken into account. It has been observed that the difficulty in the aircraft maneuvering increases the anxiety and stress levels particularly for novice pilots [245], [265]. This could specifically apply in the VR environment as a more realistic condition and especially during the periods when participants had limited control (S3). In addition, since no subject had previous experience in the simulated flying operation, the VR stronger presence environment could enhance the mental pressure evoked by the ever-increasing cognitive workload [266], [267], resulting in the elevated connectivity between these areas.

5.5.4 Limitations and Future Considerations

Some points of this work should be considered when interpreting the findings. The eLORETA source localization procedure uses a standard MRI image, hence not taking into consideration the individual brain anatomy of the subjects. Also, we employed large areas for the cortical parcellation in order to minimize the estimation error and avoid possible mislocalization. Moreover, only cortical brain areas were utilized in this Chapter and excluded sub-cortical and cerebellum brain regions, due to the frequent depth bias of source localization approaches. Hence, brain regions with known role in spatial memory and navigation, e.g. the hippocampus [268], were excluded from the functional connectivity calculation. Future studies could include individual fMRI images aiming to improve source localization calculation in order to confirm our observations and clarify subject specific workload influences. In addition, participants with no previous experience in flight simulation were recruited, thus further investigations could illuminate potential differences and training effects concerning not only novice but also experienced pilots.

5.6 Conclusion

In this Chapter, an EEG experiment of 2D and 3D flight simulation was utilized to detect the of functional brain networks differences in source space with regard to three distinct levels of complexity. In this regard, the FC of cortical regions was utilized as a means to provide discriminative features for mental load classification and thus compare the various workload stages. The proposed methodology was successful in identifying the common and distinct brain characteristics of mental load between the two simulation environments, thus illustrating high discriminative ability in classification procedure while utilizing only a small portion of the full feature set. The produced results indicate the robustness of the

proposed framework unveiling the underlying workload functions and their implications to real-world scenarios.

Chapter 6

6.1 Human-Machine Interfaces for Motor Rehabilitation

Motor-related conditions may involve significant implications for the afflicted individuals, concerning various aspects of mental and physical processes. On top of the prescribed medication, the use of non-pharmacological interventions has been added to the available arsenal towards the facilitation of everyday activities as well as the potential rehabilitation of patients. In this regard, researchers have pursued the path of establishing an alternative connection between human intent - as expressed through cognitive processing and overt residual motor activity - and software/hardware receptors that decode the desired user actions. This notion represents a Human-Machine Interface (HMI), where this Chapter focuses in analyzing the theoretical basis and implementation concerns of non-invasive applications that exploit multimodal real-time data recording. As such, using the groundwork from the experimental designs and analysis presented in previous Chapters and the recent state of the art literature, the available modalities are presented on the grounds of bioelectrical and movement phenotype signals and in conjunction with the related algorithms and clinical practice demands.

6.2 Background

Reports from WHO have demonstrated the wide impact of neurological conditions such as strokes and spinal cord injuries to the global population [269]. The high annual incidence rates are associated with a conjointly high prevalence of lingering health issues that include movement disturbance by means of interfering with the communication channel responsible for transmitting the motor command among the various nodes including the encephalon (source node), the spinal cord (intermediate node) and the muscular system (end node). If any of the aforementioned nodes is affected, the result may translate to motor disability, even if all the remaining structures remain fully unharmed. As such, the conditions mentioned above may induce implications regarding the source and intermediate nodes respectively, without diminishing movement capacity [270], [271]. In this regard, the limits imposed to the expressed movement are attributed to the deficient information transmission to the end node, itself retaining its full potential. However, the incomplete activation induced might eventually lead to perpetual motor problems such as atrophy, general ADL (activities of daily living) impotence and frailty [272], [273].

Based on the assumption of the unaffected muscular potential and the remaining brain/spinal cord functionality, specialists seek to promptly conduct restoration protocols in order to prevent or alleviate atrophy, permanent motor impairments and negative neuroplasticity, by promoting efficient neural processing during cognitive compensation [274], [275]. Significant rehabilitation attainment is not implausible, albeit the path is particularly challenging for the individuals with respect to the time, cost and commitment required [276]. Combined with slow progress, potential setbacks, limited financial resources and the often inadequate adjustment of global protocol parameters in a subject-specific manner [277], [278], adherence concerns have arisen [279]. Under these circumstances, the need for personalized frameworks targeting function execution assistance and restoration has yet to be fulfilled, although the existing groundwork on software and hardware tools has been laid towards affordable rehabilitation systems and functional modeling on subject level.

6.3 Human-Machine Motor Rehabilitation Interfaces

The decoding of human intent and its real-time translation to a control command for guiding hardware and/or software, which in turn orchestrates an interplay between an individual and their external environment, corresponds to the implementation of a Human Machine Interface (HMI) [280]. The basic concepts and components of a HMI system implementation are shown in **Figure 6.1**. A subclass of HMI is the Brain Computer Interface (BCI) which as the name suggests only utilizes brain signals for the individual-to-device path (computer module). These traits combined with the adjustability offered by such systems allows HMI application to motor-related disorders, where the decoded user intent corresponds to a desired movement which the receiver module aims to augment or even fully execute [280]. Interpretation of a planned movement may utilize locomotion phenotype data such as limb displacement or ground forces, as well as bioelectrical signals reflecting biological activity such as electroencephalography (EEG) or electromyography (EMG) recordings [281], [282]. The latter are commonly labeled as physiological data, which can be captured via non-invasive sensors and conjointly analyzed for correlating cortical activity with its muscle counterpart [282]. However, such modeling requires the accurate extraction of movement biosignals while dismissing uncorrelated activity. This procedure is achieved through data recording during rigorously designed protocols, where execution of real or imagined movements is guided by sensory cues (usually visual or acoustic) [283].

Data analysis may be based either on a single modality (e.g. EEG, EMG, locomotion phenotype) or on a combination of the available modalities, feeding the descriptive motor-related features into a machine learning scheme which in turn generates a command that echoes the initial intent onto the receiver module (e.g. a computer or an actuator) [280] that can support a function and bolster rehabilitation mechanisms. Moreover, the real-time nature of this functionality necessitates the concurrent analysis of the HMI system response and its kinematic reflection for evaluating the HMI's performance and adjusting output signals accordingly [284], [285]. In this scope, interfaces are implemented within a closed-loop design that utilizes real-time feedback signals that mirror the system's output and are used for comparing the intended action (as interpreted by the classifier) with the eventual movement. The corresponding goal is to match the two states (HMI-interpreted *vs* HMI-induced) for achieving a high degree of subject-level adaptation [284].

In the field of clinically applied rehabilitation for motor-impaired individuals, the receiver modules are classified to specific movement actuator classes [286]–[288]. Orthotic and exoskeleton devices represent the first two classes [289], bearing a resistive, assistive or passive role during movement, in reference to whether the desired function is partially inhibited (for muscle strengthening), amplified or fully performed by the hardware [288], [290]. The third receiver class refers to stimulation devices that implement Functional Electrical Stimulation (FES) [291] by administering electrical pulses that induce muscle contraction. This process basically aims to stand in for the impaired biological processes (implicating the brain and the spinal cord) that fail to provoke a full natural contraction. On this premise, FES application may concurrently enhance a specific movement, as well as trigger neurorehabilitation operations within the cortical activity [292]. The fundamental prerequisite for the related treatment is the existence of residual motor capabilities [287], that the stimulation taps into for augmenting muscle activation and aiding patients to reprogram their cortical kinetic patterns. However, researchers should consider the key differences distinguishing the stimulation-triggered muscle activation from a natural contraction, mainly its irregular structure that resembles a step function following the "all-or-none" law [293], as well as the

external parameter dependability. Namely, application parameters such as electrode placement and stimulation waveform properties introduce a substantial variability [293], [294] that requires rigorous testing for maximum adaptability on subject-level [294].

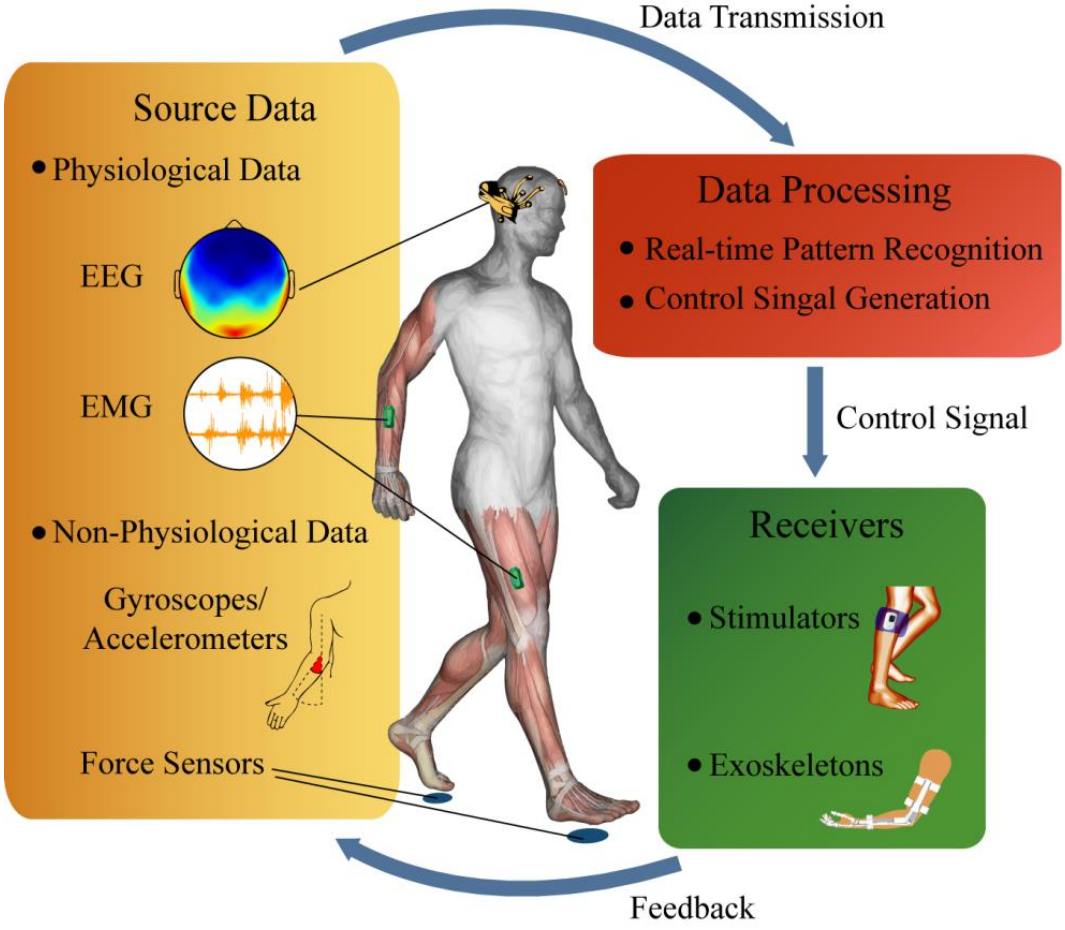


Figure 6.1. A schematic design for HMI rehabilitation

6.4 Computational Intelligence Tools

The aforementioned interpretation of a user's intended action through the explanatory features within the input motion data implies the definition of distinct output classes, denoting the different motion states that can be inferred by a pattern recognition paradigm [280], [295]. Such paradigms aim to optimally discern the HMI response motion states based on the annotated data, recorded during experimental protocols implicating real or imaginary motion execution. In this context, a multitude of algorithms including k-NN, SVM, LDA, soft labeling techniques [296]–[298], and network-based implementations [295], [299] have been utilized to unveil linear and non-linear dependencies, either individually or within ensemble classification schemes [300].

In this perspective, choosing an HMI scheme from the pool of available methodologies is not a straightforward task, usually bearing a tradeoff among classification performance, computational complexity and response time [301]. Regarding performance, the role of the employed features is vital, since they effectively represent our comprehension concerning motion processes. For this reason, motion

conditions incorporated within experimental designs must be as realistic as possible (accounting for environmental constraints [302] and participant safety), while considering all parameters involved, for successfully determining explanatory motor features [303]. Each parameter contribution can be studied through multi-level modeling entailing biological and locomotion processes in conjunction with data recording. The latter refers to the interactions among the studied functions and the measuring equipment, as well as the properties of the recorded signals. For example, the phenomena taking place regarding bioelectrical/kinematic sensor placement and continuous function such as skin contact [304], muscle contraction and joint movement [305] can be modeled towards a holistic analysis. Indicatively, this data should be utilized to fully represent all aspects of a gait cycle [306] or grasping actions [281], allowing the testing of multiple scenarios [307] (e.g. walking, stair climbing, falling, finger movements). The study of kinematic patterns involves the application of differential equations under simulated locomotion task conditions, implicating antagonistic muscle pairs [305], [308].

Within this abundance of available variables and under the requirement for real-time motion interpretation, computational speed becomes a priority that limits the "allowed" computational cost [309]. In this regard, the vast feature pool must be reduced to a small number of discriminative features that offer high classification performance within a "reasonable" time frame for real-time HMI response (in the order of ms) [310], [311]. This feature selection yields simpler and more reliable models that depend on critical motor-related markers rather than a huge number of non-interpretable feature combinations that mostly correspond to overtraining bias [229], [310]. The above background is summarized in **Figure 6.2**, illustrating a typical workflow towards the design and implementation of an HMI:

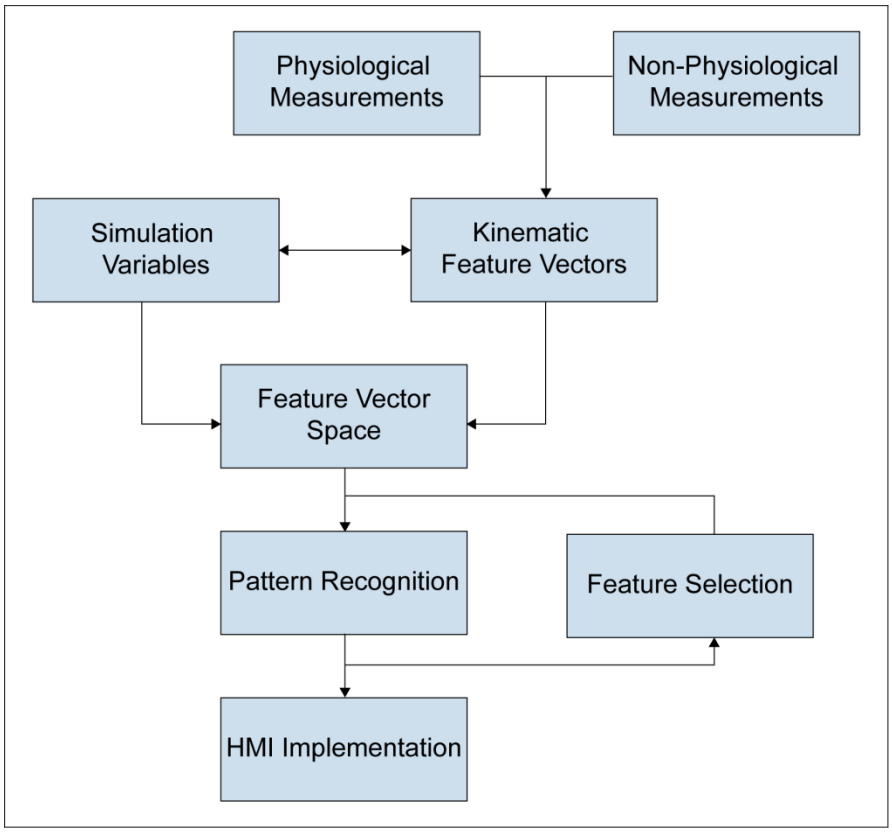


Figure 6.2. Optimal HMI implementation model design

6.5 Movement Phenotype Modalities

As presented above, data exploitation for motor-related HMIs may incorporate biosignals and/or non-biological locomotion data, extractable via non-invasive mobile sensors. The latter modality serves the evaluation of locomotion mechanics concerning parameters that include displacement/rotation of limbs and joints, as well as reaction forces (e.g. ground). In relation to upper-limb support and rehabilitation, typical measurements include elbow, forearm and wrist angles [312], while setups for lower limbs analyze knee/hip joint angles [313] for computing gait activity variables [299], [306]. In both cases flexion data and resistance forces provide additional information on motion range and extension [314], [315], contributing to a holistic rehabilitation progress evaluation. Overall, force sensors, inertia measurement units (IMUs) [299], e-textiles [316], bio-impedance sensors [317], and visual motion capture technology [318] constitute eligible hardware for unobtrusive data collection, either in laboratory or everyday settings. Moreover, part of the equipment is often incorporated in specialized wearable modules such as smart clothing [319].

By implementing dedicated classification algorithms that utilize multiple trials of fundamental movement patterns [320], sensor data are used to guide orthotic/prosthetic/exoskeleton [285], [299] equipment or even FES devices. Notably, sensor-guided stimulation constitutes a viable and more adaptable alternative to traditional paradigms that require continuous manual adjustments by supervising experts [303]. In this fashion, adaptive user training based on objective quantitative metrics has the potential to limit rehabilitation setbacks [321] and enhance protocol adherence and self-perceived progress. Furthermore, the availability of such recordings provides networking capabilities through bidirectional data exchange with cloud infrastructure and mobile applications [322], enabling remote rehabilitation monitoring.

6.6 Bioelectrical Modalities

In unveiling the complex biological processes governing motor functions, research focuses on neurological cognitive functions and electromyographic activity, both separately and in conjunction with each other. Notably, the relative interactions during motor planning, preparation and execution have been studied [323] in pursuance of objective metrics describing the neuromuscular functional associations. Such metrics bear the added value of providing evaluation capabilities regarding neurorehabilitation protocols, where the localization and quantitative estimation of electrophysiological activity are critical for assessing patient progress. In the context of evaluating interactions engaging brain and muscle activity, corticomuscular coherence is an established metric that seeks associations in the frequency domain [324], showcasing specific bands where cortical activity is related to muscle contraction [325].

Another technique with very significant contribution to cognitive function study through electrical activity is event-related analysis that has revealed a number of critical components of brain activity during the various motion stages [326]. Especially concerning the time window corresponding to motion planning and preparation (about 0.5-2 sec prior to execution) [326], [327], the MRCP (Movement-Related Cortical Potential) comprises a low-frequency (<5Hz) EEG component that has been verified under imaginary motion experiments, thus being suitable for support and rehabilitation applications where a patient exhibits limited motion [326]. Additional components occurring over the timespan preceding kinematic manifestation (mostly 1.0-1.5 sec before onset) comprise the Contingent Negative Variation

(CNV) [328] and the Bereitschaftspotential (BP) [329], appearing mainly over the premotor cortex and the primary motor cortex respectively.

On top of the individual components, event-related analysis also involves the study of synchronization / desynchronization (ERS/ERD) within EEG activity. More specifically, alpha (8-12 Hz) and central beta frequency bands (16-24 Hz) display evident ERD during both real and imaginary movement preparation [326], [327]. Namely, motor imagery and real motion employ common neural pathways, thus the former is considered a sufficient substitute for experimental designs where participants display severely limited motor abilities [283].

Generally, the advances in EEG analysis and its ascertained contribution to the study of motor functions have rendered electroencephalography as a key modality in HMI rehabilitation designs [301], [330]. The main idea in such implementations is the establishment of a real-time information bus engaging brain activity and a receiver module, trained via motor imagery data. Both exoskeletons [331]–[333] and functional electrical stimulation devices [67,68] have served as the receiver device responsible for motion execution, implicating upper and lower limbs in applications such as generalized movement support and foot drop treatment. Stroke, spinal cord injury and spastic cerebral palsy [336] are a few indicative examples where the EEG-FES combination has been successfully employed, having shown greater and more persistent benefits compared to traditional FES treatment [337], [338]. In an analogous manner, the control of these receiver modules has also exploited non-invasive electromyography signals reflecting contraction intensity, related both to extended muscle groups (e.g. leg extension) and to more complex detailed movements (e.g. fingers). Such data bear the potential to accurately characterize specific movement patterns that may not be discerned via EEG [295], [339], [340]. Particularly, EMG produced by residual muscular activity can be used as a regulator for FES treatment and exoskeletons, aiming to support motion and trigger existing kinematic capacity, thus assisting in repossession of sufficient muscular activation [341], [342]. Over the course of rehabilitation, training protocols often exploit recordings derived from a non-afflicted limb for target feedback, contralateral control and progress evaluation [343], [344].

By combining the different modalities into hybrid frameworks, a wide variety of additional support and therapeutic options arises through the incorporation of the assets of each modality. The use of multiple physiological modalities or mixtures of physiological and non-physiological data offers enhanced computational validity since different information sources [345]–[348] (angle, force, biological activity, etc.) can be exploited concurrently for motion decoding and support [286], [324] using the aforementioned receiver modules. Such holistic real-time movement analysis constitutes substantial groundwork for strengthening the functional connections between cognitive processing and kinematic phenotype, seeking to promote convergence of muscular and cortical electrical activity in the context of rehabilitation [349]. The former (muscular activity) represents the motion execution aspect while the latter (cortical activity) also contains the motion intention aspect, thus hybrid protocols utilizing both EEG and EMG aim at restoring – at the highest degree possible – their natural correlation, accounting for side events including non-voluntary motions such as spasms and tremors [350].

6.7 Future Challenges

Although technological and methodological advances have given a great boost to HMI research, accessible use in clinical and everyday practice is still not a reality. The lack of widely available

commercial end-products that can provide adaptable real-time support and rehabilitation services at an affordable price is yet to be fulfilled by the large-scale research frameworks currently studied [288], [330]. Implementation challenges include reliable EMG-based motion analysis, specifically with regard to joint displacement and electrode alterations as well as algorithm personalization and accounting for workload and floor/ceiling effects [351], [352].

An additional research question yet to be addressed is related to neurorehabilitation mechanisms triggered during training [353]. The human brain bears the ability to constantly adapt to new conditions via plasticity, with the use of HMIs being no exception. Type of impairment along with general patient status and specific protocol parameters may dictate the course of neuroplasticity, while it has been suggested that not all plasticity is beneficial [354]. Compensatory mechanisms are not always optimal, a critical factor being the temporal window of their development [276], also related to the timing of the rehabilitation training.

From a data analysis and interpretation standpoint, novel methods such as brain connectivity are progressively employed for studying corticomuscular interactions [220], [226], albeit there is still a lot to be done on deciphering brain electrical activity during motor-related processes. Moreover, the highly subject-specific nature of HMI training poses additional challenges regarding the adaptation of a common framework to every individual user through parameter adjustments. Customization is further complicated by the variable phenotype of motor impairments for each subject, necessitating comprehensive and user-tailored training [278]. Finally, in pursuance of optimal standards and methodological potency, tools and models should be validated using third-party (including open access) data [355], so that the scientific community is able to refine the results and filter best practices on rehabilitation and interpretation of biological and non-biological activity describing human motion.

Chapter 7

Main Contributions and Future Work

This Chapter summarizes the conclusions of previous Chapters and suggests guidelines for future work. In this Doctoral Thesis, a wide array of ML approaches have been implemented in different brain related experiments providing the basis for applicable HMI/BCI interactions and shedding light to underlying cognitive aspects of different states in terms of neuroscience. In this regard, it should be noted that BCI implementation requires the continuous measurement of the brain signals translating them to input information for the connected device to execute each task at hand. By extension, irrelevant neural stimuli might confound the internal classification processing, thus it is of outmost importance to detect specific neural attributes in a way that they are as discriminative as possible in the most universal way. This Thesis' aim was to establish the groundwork for BCI application by assessing universal characteristics of different cognitive states and conditions in ML frameworks, putting forward and exploring advances in algorithmic applications and system implementation, while incorporating basic analysis in real and imaginary motion recordings.

In detail, Chapters 3,4,5 refer to ML approaches in cognitive EEG experiments, each one addressing a different aspect of HMI implementation. First and foremost, Chapter 3 regards the classification of error cognition in regard to ERP-related features, focusing on global mental aspects irrespective of task complexity. The results produced indicated that classification across different difficulty conditions can be achieved, however due to subject variability estimating universal condition-specific characteristics could be problematic. Advancing contemporary research a thorough analysis has been made, not only including the typical ERN, but the entirety of ErrPs reported in previous studies by concatenating time-window features, each corresponding to specific ERPs. This allowed for efficient ML procedures in full and reduced channel analysis, providing indices of a global mechanism and efficient cross-condition evaluation.

Chapters 4 and 5 focused on the classification of mental load using frequency related attributes, such as PSD and FC networks as feature sets. In both cases satisfactory classification accuracy was achieved using FS and classification methods, while identifying a small set of relationships between different brain regions in different frequency bands, which might contribute to the understanding of the various related neural mechanisms. Task engagement is an important aspect when applying brain signals to a HMI, however prolonged task execution (especially in mentally demanding tasks) results into deteriorated performance and masks the useful BCI-relevant characteristics due to workload. Accordingly, in Chapter 4, a task-independent mental workload ML approach was implemented, taking into account a fusion of different EEG spectral characteristics (i.e., the PSD and FC) resulting in the highest (so far) cross-task workload classification. More importantly, the task-independent features identified did not only provide rigorous indicators of their inherent predictive quality, but also demonstrated cognitive control of inhibitory functions as the main factor of a global workload mechanism. This outcome can be explained as "forgetting" is more mentally exhausting than "remembering", especially in prolonged WM tasks, since continuous cognitive updating with each new trial necessitates additional memory capacity resources, while disregarding irrelevant stimuli of previous trials. Within this context, these results provide some of the first quantitative confirmation of the prevalent workload-related neural mechanisms aiming to provide significant improvements in cognitive load assessment.

In Chapter 5, building upon the previous experiments results and conclusions, workload assessment has been employed in an approximation of real-world applications. Specifically, flight simulation was utilized in a 2D and 3D environment in a multi-workload levels ML approach utilizing functional brain networks differences in source space. High classification performance was attained detecting specific ROIs in specific brain frequencies, indicating common and unique characteristics in both environments. More importantly, each environment indicated differentiation with regards to network metrics (local and global efficiency) and FC. As such, VR environment was estimated to evoke higher mental workload to the participants compared to the CS environment. Interestingly, this was not only contributed due to the more visually demanding cognitive processes, but also in relation to the anxiety and stress levels when participants had limited control, since the amygdala brain areas were only indicated as significant in the VR interface. Thus the VR stronger presence environment could enhance the mental pressure evoked by a more realistic condition simulating real-world situations.

Taking all the above into account, Chapter 6 presents the basis of bioelectrical phenotype signals and in conjunction with the related ML algorithms and clinical practice demands. As such, the technological and methodological advances of BCI/HMI are discussed, as well as the implementation challenges and the need for adaptable, real-time and personalized support accounting for workload and floor/ceiling effects. Finally, the importance of deciphering brain electrical activity in a global manner (irrelevant of task and subject variability) is presented, which is a major requirement for the adaptation of a common framework to every individual user through parameter adjustments and user-tailored training.

In future works the proposed methods can be applied to more data to verify their effectiveness and especially to extract new knowledge. The new relationships that emerge from computational analysis shed light on the functioning of the brain mechanisms involved and may suggest directions for future cognitive experiments to validate them. In addition, the ML methods can be further tested to confirm their performance, as well as to extend the classification to multiple and much more complex levels. In addition, deep learning neural networks can be tested, which have recently shown excellent performance in image classification, however, appropriate data transformation is required to be able to utilize them. Especially in the case of BCI, a progressive step is the employment of portable dry-electrode EEG devices to make predictions in real time. For this to be possible, a small number of sensors are required, which is supported by the results of the FS performed in the experimental Chapter in this Doctoral Thesis.

List of Publications

Published work included in this Doctoral Thesis dissertation:

Kakkos, I., K. Gkiatis, K. Bromis, P. A. Asvestas, I. S. Karanasiou, E. M. Ventouras, and G. K. Matsopoulos. "Classification of Error Related Brain Activity in an Auditory Identification Task with Conditions of Varying Complexity." Journal of Physics: Conference Series 931, no. 1 (2017): 012017. https://doi.org/10.1088/1742-6596/931/1/012017.

Kakkos, I., E. M. Ventouras, P. A. Asvestas, I. S. Karanasiou, and G. K. Matsopoulos. "A Condition-Independent Framework for the Classification of Error-Related Brain Activity." Medical & Biological Engineering & Computing 58, no. 3 (March 2020): 573–87. https://doi.org/10.1007/s11517-019-02116-5.

Kakkos, I., G. N. Dimitrakopoulos, L. Gao, Y. Zhang, P. Qi, G. K. Matsopoulos, N. Thakor, A. Bezerianos, and Y. Sun. "Mental Workload Drives Different Reorganizations of Functional Cortical Connectivity Between 2D and 3D Simulated Flight Experiments." IEEE Transactions on Neural Systems and Rehabilitation Engineering 27, no. 9 (September 2019): 1704–13. https://doi.org/10.1109/TNSRE.2019.2930082.

Kakkos I., G. N. Dimitrakopoulos, Yi Sun, G. K. Matsopoulos, A. Bezerianos, and Yu Sun, "EEG Fingerprints of Task-independent Mental Workload Discrimination IEEE J Biomed Health Inform, vol. PP, Jun. 2021, https://doi.org/10.1109/JBHI.2021.3085131.

Kakkos, I, S. T. Miloulis, K. Gkiatis, G. N. Dimitrakopoulos, and G. K. Matsopoulos. "Human–Machine Interfaces for Motor Rehabilitation." In Advanced Computational Intelligence in Healthcare-7: Biomedical Informatics, edited by Ilias Maglogiannis, Sheryl Brahnam, and Lakhmi C. Jain, 1–16. Studies in Computational Intelligence. Berlin, Heidelberg: Springer, 2020. https://doi.org/10.1007/978-3-662-61114-2_1.

Published work carried out alongside (but not included) in this Doctoral Thesis:

K. Gkiatis, A. Karampasi, I. Kakkos, S.T. Miloulis, G.K. Matsopoulos, Ch. Benjamin, and K. Garganis, "Presurgical mapping of language network with fMRI: Greek evaluation", Poster in 13th Panhellenic Epilepsy Conference, 7-9 June 2019, Athens, 2nd Prize Award for best poster presented.

A. Karampasi, I. Kakkos, S.-T. Miloulis, I. Zorzos, G.N. Dimitrakopoulos, K. Gkiatis, P. Asvestas, G. Matsopoulos, A Machine Learning fMRI Approach in the Diagnosis of Autism, in: 2020 IEEE Int. Conf. Big Data Big Data, 2020: pp. 3628–3631. https://doi.org/10.1109/BigData50022.2020.9378453.

Bromis, K., I. Kakkos, K. Gkiatis, I. S. Karanasiou, and G. K. Matsopoulos. "Brain Functional Connectivity in Small Cell Lung Cancer Population after Chemotherapy Treatment: An ICA FMRI Study." Journal of Physics: Conference Series 931, no. 1 (2017): 012041. https://doi.org/10.1088/1742-6596/931/1/012041.

Gkiatis, K., K. Bromis, I. Kakkos, I. S. Karanasiou, G. K. Matsopoulos, and K. Garganis. "Effects of Inaccurate Identification of Interictal Epileptiform Discharges in Concurrent EEG-FMRI." Journal of Physics: Conference Series 931, no. 1 (2017): 012042. https://doi.org/10.1088/1742-6596/931/1/012042.

Miloulis, S. T., I. Kakkos, G. N. Dimitrakopoulos, Y. Sun, I. S. Karanasiou, P. A. Asvestas, E. M. Ventouras, and G. K. Matsopoulos. “Evaluating Memory and Cognition via a Wearable EEG System: A Preliminary Study.” In Wireless Mobile Communication and Healthcare, edited by Juan Ye, Michael J. O’Grady, Gabriele Civitarese, and Kristina Yordanova, 52–66. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Cham: Springer International Publishing, 2021. https://doi.org/10.1007/978-3-030-70569-5_4.

Zorzos I., I. Kakkos, E. M. Ventouras, and G. K. Matsopoulos. “Advances in Electrical Source Imaging: Current Approaches, Applications and Challenges”, Signals 2, no. 3: 378-391. https://doi.org/10.3390/signals2030024.

Karampasi A. S., A. D. Savva, V. C. Korfiatis, I. Kakkos, and G. K. Matsopoulos. “Informative Biomarkers for Autism Spectrum Disorder Diagnosis in functional Magnetic Resonance Imaging Data on the Default Mode Network”, Applied Sciences 11, no. 13: 6216. https://doi.org/10.3390/app11136216.

Appendix A

Glossary of Terms Utilized in the Extended Greek Summary

ΒΔ	βιωματικά δυναμικά
ΔΑΜ	διεπαφή ανθρώπου-μηχανής
ΔΕΥ	διεπαφή εγκεφάλου-υπολογιστή
ΕΧ	επιλογή χαρακτηριστικών
ΗΕΓ	ηλεκτροεγκεφαλογράφημα
ΜΜ	μηχανική μάθηση
FIR	πεπερασμένη κρουστική απόκριση
IIR	άπειρη κρουστική απόκριση
ICA	ανάλυση ανεξάρτητων συστατικών
LDA	ανάλυση γραμμικής διάκρισης
LORETA	ηλεκτρομαγνητική τομογραφία εγκεφάλου χαμηλής ανάλυσης
PLI	δείκτης καθυστέρησης φάσης
PSD	φασματική πυκνότητα ισχύος
RFE-CBR	αναδρομική μέθοδος εξάλειψης χαρακτηριστικών με τη μέθοδο μείωσης συσχέτισης μεροληψίας
SFFS	διαδοχική κυμαινόμενη πρόσθια επιλογή
SFS	διαδοχική πρόσθια επιλογή
SVM	μηχανές διανυσμάτων υποστήριξης

Appendix B

Glossary of Terms Utilized in this Doctoral Thesis

AAL	automatic anatomical labelling
ADL	activities of daily living
ANOVA	analysis of variance
BCI	brain computer interface
CNV	contingent negative variation
CRN	correct-related negativity
CS	computer screen
EEG	electroengephalography
EG	global efficiency
EL	local efficiency
eLORETA	exact low resolution electromagnetic tomography
EMG	electromyography
EOG	electrooculogram
ERB	equivalent rectangular bandwidth
ERD/ERS	event-related (de)synchronization
ERN	error-related negativity
ERP	event-related Potential
ErrPs	error-related ERPs
ESI	electrical source imaging
FBT	feedback tone
FC	functional connectivity
fERN	feedback related negativity
FES	functional electrical stimulation
FIR	finite impulse response
fMRI	functional magnetic resonance imaging
FS	feature selection
HEOG	horizontal electrooculogram
HMI	human-machine interface
ICA	independent component analysis
IRR	infinite impulse response
k-NN	k-nearest neighbor
KOR	knowledge-of-results
LDA	linear discriminant analysis
LOOCV	leave-one-out cross-validation
LORETA	low resolution brain electromagnetic tomography
LS	least squares

ML	machine learning
MRCP	movement-related cortical potential
PE	error-positivity
PET	positron emission tomography
PLI	phase lag index
PSD	power spectral density
QP	quadratic programming
RBF	radial basis function
RF	random forests
RFE	recursive feature elimination
RFE-CBR	recursive feature elimination method with correlation bias reduction
ROC	receiver operating characteristic
ROI	regions of interest
S1,..2,..3	stage 1,..2,..3
SFFS	sequential forward selection
SFS	sequential forward floating selection
SMO	sequential minimal optimization
SVM	support vector machines
tw	time window
VEOG	vertical electrooculogram
VR	virtual reality
WHO	world health organization
WM	working memory

References

[1] J. K. Mai and G. Paxinos, *The Human Nervous System*. Academic Press, 2011.

[2] *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Oxford University Press.

[3] X. Xing, W. Pei, Y. Wang, Z. Liu, and H. Chen, "Design of High-Density Electrodes For EEG Acquisition," *Annu Int Conf IEEE Eng Med Biol Soc*, vol. 2018, pp. 1295–1298, Jul. 2018, doi: 10.1109/EMBC.2018.8512577.

[4] G. A. Slipher, W. D. Hairston, J. C. Bradford, E. D. Bain, and R. A. Mrozek, "Carbon nanofiber-filled conductive silicone elastomers as soft, dry bioelectronic interfaces," *PLoS One*, vol. 13, no. 2, Feb. 2018, doi: 10.1371/journal.pone.0189415.

[5] E. Niedermeyer and F. H. L. da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, 2005.

[6] P. Giacometti, K. L. Perdue, and S. G. Diamond, "Algorithm to find high density EEG scalp coordinates and analysis of their correspondence to structural and functional regions of the brain," *J Neurosci Methods*, vol. 229, pp. 84–96, May 2014, doi: 10.1016/j.jneumeth.2014.04.020.

[7] E. Başar and A. Düzgün, "The CLAIR model: Extension of Brodmann areas based on brain oscillations and connectivity," *International Journal of Psychophysiology*, vol. 103, pp. 185–198, May 2016, doi: 10.1016/j.ijpsycho.2015.02.018.

[8] C. S. Nayak and A. C. Anilkumar, "EEG Normal Waveforms," in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2021.

[9] J. J. Newson and T. C. Thiagarajan, "EEG Frequency Bands in Psychiatric Disorders: A Review of Resting State Studies," *Front. Hum. Neurosci.*, vol. 12, 2019, doi: 10.3389/fnhum.2018.00521.

[10] J. Jang, M. Jones, E. Milne, D. Wilson, and K.-H. Lee, "Contingent negative variation (CNV) associated with sensorimotor timing error correction," *NeuroImage*, vol. 127, pp. 58–66, Feb. 2016, doi: 10.1016/j.neuroimage.2015.11.071.

[11] P. E. Clayson, J. K. Wynn, Z. P. Infantolino, G. Hajcak, M. F. Green, and W. P. Horan, "Reward processing in certain versus uncertain contexts in schizophrenia: An event-related potential (ERP) study," *J Abnorm Psychol*, vol. 128, no. 8, pp. 867–880, Nov. 2019, doi: 10.1037/abn0000469.

[12] S. J. Luck and E. S. Kappenman, *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press, 2011.

[13] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nat Rev Neurosci*, vol. 10, no. 3, pp. 186–198, Mar. 2009, doi: 10.1038/nrn2575.

[14] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, Sep. 2010, doi: 10.1016/j.neuroimage.2009.10.003.

[15] C. J. Stam, G. Nolte, and A. Daffertshofer, "Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources," *Hum Brain Mapp*, vol. 28, no. 11, pp. 1178–1193, Nov. 2007, doi: 10.1002/hbm.20346.

[16] S. Aydore, D. Pantazis, and R. M. Leahy, "A Note on the Phase Locking Value and its Properties," *Neuroimage*, vol. 74, pp. 231–244, Jul. 2013, doi: 10.1016/j.neuroimage.2013.02.008.
[17] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, Jun. 1998, doi: 10.1038/30918.
[18] D. S. Bassett and E. T. Bullmore, "Small-World Brain Networks Revisited," *Neuroscientist*, vol. 23, no. 5, pp. 499–516, Oct. 2017, doi: 10.1177/1073858416667720.
[19] S. Achard and E. Bullmore, "Efficiency and cost of economical brain functional networks," *PLoS Comput. Biol.*, vol. 3, no. 2, p. e17, Feb. 2007, doi: 10.1371/journal.pcbi.0030017.
[20] M. Sazgar and M. G. Young, "EEG Artifacts," in *Absolute Epilepsy and EEG Rotation Review: Essentials for Trainees*, M. Sazgar and M. G. Young, Eds. Cham: Springer International Publishing, 2019, pp. 149–162. doi: 10.1007/978-3-030-03511-2_8.
[21] E.-R. Symeonidou, A. D. Nordin, W. D. Hairston, and D. P. Ferris, "Effects of Cable Sway, Electrode Surface Area, and Electrode Mass on Electroencephalography Signal Quality during Motion," *Sensors (Basel)*, vol. 18, no. 4, Apr. 2018, doi: 10.3390/s18041073.
[22] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of Artifacts from EEG Signals: A Review," *Sensors (Basel)*, vol. 19, no. 5, Feb. 2019, doi: 10.3390/s19050987.
[23] A. B. Usakli, "Improvement of EEG Signal Acquisition: An Electrical Aspect for State of the Art of Front End," *Comput Intell Neurosci*, vol. 2010, 2010, doi: 10.1155/2010/630649.
[24] J. W. Leis, *Digital Signal Processing Using MATLAB for Students and Researchers*. John Wiley & Sons, 2011.
[25] Y. Li, Y. Wang, B. Zhang, Y. Wang, and X. Zhou, "Electrophysiological Responses to Expectancy Violations in Semantic and Gambling Tasks: A Comparison of Different EEG Reference Approaches," *Front. Neurosci.*, vol. 12, 2018, doi: 10.3389/fnins.2018.00169.
[26] M. D. Lutovac, D. V. Tošić, and B. L. Evans, *Filter Design for Signal Processing Using MATLAB and Mathematica*. Miroslav Lutovac, 2001.
[27] M. Bianchi, M. Boyle, and D. Hollingsworth, "A comparison of methods for trend estimation," *Applied Economics Letters*, vol. 6, no. 2, pp. 103–109, Feb. 1999, doi: 10.1080/135048599353726.
[28] T. P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects," *Clin Neurophysiol*, vol. 111, no. 10, pp. 1745–1758, Oct. 2000, doi: 10.1016/s1388-2457(00)00386-2.
[29] P. M. Alday, "How much baseline correction do we need in ERP research? Extended GLM model can replace baseline correction while lifting its limits," *Psychophysiology*, vol. 56, no. 12, Dec. 2019, doi: 10.1111/psyp.13451.
[30] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *IEEE Signal Processing Magazine*, vol. 18, no. 6, pp. 14–30, Nov. 2001, doi: 10.1109/79.962275.
[31] C. M. Michel and D. Brunet, "EEG Source Imaging: A Practical Review of the Analysis Steps," *Front. Neurol.*, vol. 10, 2019, doi: 10.3389/fneur.2019.00325.
[32] L. Garnero, S. Baillet, G. Marin, B. Renault, C. Guérin, and G. Meunier, "Introducing priors in the EEG/MEG inverse problem," *Electroencephalogr Clin Neurophysiol Suppl*, vol. 50, pp. 183–189, 1999.
[33] Z. Liu, L. Ding, and B. He, "Integration of EEG/MEG with MRI and fMRI in Functional Neuroimaging," *IEEE Eng Med Biol Mag*, vol. 25, no. 4, pp. 46–53, 2006.

[34] R. J. Ilmoniemi, "Models of source currents in the brain," *Brain Topogr*, vol. 5, no. 4, pp. 331–336, Jun. 1993, doi: 10.1007/BF01128686.
[35] J. J. Ermer, J. C. Mosher, S. Baillet, and R. M. Leah, "Rapidly recomputable EEG forward models for realistic head shapes," *Phys Med Biol*, vol. 46, no. 4, pp. 1265–1281, Apr. 2001, doi: 10.1088/0031-9155/46/4/324.
[36] R. D. Pascual-Marqui, "Discrete, 3D distributed, linear imaging methods of electric neuronal activity. Part 1: exact, zero error localization," *arXiv:0710.3341 [math-ph, physics:physics, q-bio]*, Oct. 2007, [Online]. Available: http://arxiv.org/abs/0710.3341
[37] R. Grech *et al.*, "Review on solving the inverse problem in EEG source analysis," *Journal of NeuroEngineering and Rehabilitation*, vol. 5, no. 1, p. 25, Nov. 2008, doi: 10.1186/1743-0003-5-25.
[38] P. J. Durka, A. Matysiak, E. M. Montes, P. V. Sosa, and K. J. Blinowska, "Multichannel matching pursuit and EEG inverse solutions," *J Neurosci Methods*, vol. 148, no. 1, pp. 49–59, Oct. 2005, doi: 10.1016/j.jneumeth.2005.04.001.
[39] T. M. Mitchell, *Machine Learning*, 1st edition. New York: McGraw-Hill Education, 1997.
[40] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd edition. Upper Saddle River: Pearson, 2009.
[41] Q. Zhang and S. Zhu, "Visual interpretability for deep learning: a survey," *Frontiers Inf Technol Electronic Eng*, vol. 19, no. 1, pp. 27–39, Jan. 2018, doi: 10.1631/FITEE.1700808.
[42] D. L. Olson and D. Delen, "Performance Evaluation for Predictive Modeling," in *Advanced Data Mining Techniques*, D. L. Olson and D. Delen, Eds. Berlin, Heidelberg: Springer, 2008, pp. 137–147. doi: 10.1007/978-3-540-76917-0_9.
[43] T. Saito and M. Rehmsmeier, "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets," *PLoS One*, vol. 10, no. 3, Mar. 2015, doi: 10.1371/journal.pone.0118432.
[44] M. K. Scheffers and M. G. Coles, "Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors," *J Exp Psychol Hum Percept Perform*, vol. 26, no. 1, pp. 141–151, Feb. 2000.
[45] M. Falkenstein, J. Hohnsbein, J. Hoormann, and L. Blanke, "Effects of errors in choice reaction tasks on the ERP under focused and divided attention.," *Psychophysiological brain research*, vol. 1, pp. 192–195, 1990.
[46] W. J. Gehring, M. Coles, D. Meyer, and E. Donchin, "The error-related negativity: an event-related brain potential accompanying errors," *Psychophysiology*, vol. 27, p. 34, 1990.
[47] M. Falkenstein, J. Hohnsbein, and J. Hoormann, "Event-related potential correlates of errors in reaction tasks," *Electroencephalogr Clin Neurophysiol Suppl*, vol. 44, pp. 287–296, 1995.
[48] S. J. Luck and E. S. Kappenman, *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press, 2013.
[49] V. van Veen and C. S. Carter, "Error Detection, Correction, and Prevention in the Brain: A Brief Review of Data and Theories," *Clinical EEG and Neuroscience*, vol. 37, no. 4, pp. 330–335, Oct. 2006, doi: 10.1177/155005940603700411.
[50] P. L. Davies, S. J. Segalowitz, and W. J. Gavin, "Development of response-monitoring ERPs in 7- to 25-year-olds," *Dev Neuropsychol*, vol. 25, no. 3, pp. 355–376, 2004, doi: 10.1207/s15326942dn2503_6.

[51] T. Endrass, C. Franke, and N. Kathmann, "Error awareness in a saccade countermanding task," *Journal of Psychophysiology*, vol. 19, no. 4, pp. 275–280, 2005, doi: 10.1027/0269-8803.19.4.275.
[52] M. Falkenstein, J. Hoormann, S. Christ, and J. Hohnsbein, "ERP components on reaction errors and their functional significance: a tutorial," *Biological Psychology*, vol. 51, no. 2–3, pp. 87–107, Jan. 2000, doi: 10.1016/S0301-0511(99)00031-9.
[53] G. Hughes and N. Yeung, "Dissociable correlates of response conflict and error awareness in error-related brain activity," *Neuropsychologia*, vol. 49, no. 3, pp. 405–415, Feb. 2011, doi: 10.1016/j.neuropsychologia.2010.11.036.
[54] S. Nieuwenhuis, K. R. Ridderinkhof, J. Blom, G. P. H. Band, and A. Kok, "Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task," *Psychophysiology*, vol. 38, no. 5, pp. 752–760, Sep. 2001, doi: 10.1111/1469-8986.3850752.
[55] M. X. Cohen, C. E. Elger, and C. Ranganath, "Reward expectation modulates feedback-related negativity and EEG spectra," *NeuroImage*, vol. 35, no. 2, pp. 968–978, Apr. 2007, doi: 10.1016/j.neuroimage.2006.11.056.
[56] G. Hajcak, J. S. Moser, C. B. Holroyd, and R. F. Simons, "The feedback-related negativity reflects the binary evaluation of good versus bad outcomes," *Biological Psychology*, vol. 71, no. 2, pp. 148–154, Feb. 2006, doi: 10.1016/j.biopsycho.2005.04.001.
[57] G. Hajcak, C. B. Holroyd, J. S. Moser, and R. F. Simons, "Brain potentials associated with expected and unexpected good and bad outcomes," *Psychophysiology*, vol. 42, no. 2, pp. 161–170, Mar. 2005, doi: 10.1111/j.1469-8986.2005.00278.x.
[58] J. S. Moser and R. F. Simons, "The neural consequences of flip-flopping: The feedback-related negativity and salience of reward prediction," *Psychophysiology*, vol. 46, no. 2, pp. 313–320, Mar. 2009, doi: 10.1111/j.1469-8986.2008.00760.x.
[59] S. Nieuwenhuis, C. B. Holroyd, N. Mol, and M. G. H. Coles, "Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance," *Neuroscience & Biobehavioral Reviews*, vol. 28, no. 4, pp. 441–448, Jul. 2004, doi: 10.1016/j.neubiorev.2004.05.003.
[60] N. Yeung, C. B. Holroyd, and J. D. Cohen, "ERP correlates of feedback and reward processing in the presence and absence of response choice," *Cereb. Cortex*, vol. 15, no. 5, pp. 535–544, May 2005, doi: 10.1093/cercor/bhh153.
[61] S. Allain, L. Carbonnell, M. Falkenstein, B. Burle, and F. Vidal, "The modulation of the Ne-like wave on correct responses foreshadows errors," *Neuroscience Letters*, vol. 372, no. 1–2, pp. 161–166, Nov. 2004, doi: 10.1016/j.neulet.2004.09.036.
[62] R. F. Simons, "The way of our errors: theme and variations," *Psychophysiology*, vol. 47, no. 1, pp. 1–14, Jan. 2010, doi: 10.1111/j.1469-8986.2009.00929.x.
[63] D. C. Torpey, G. Hajcak, J. Kim, A. Kujawa, and D. N. Klein, "Electrocortical and Behavioral Measures of Response Monitoring in Young Children During a Go/No-Go Task," *Dev Psychobiol*, vol. 54, no. 2, pp. 139–150, Mar. 2012, doi: 10.1002/dev.20590.
[64] F. Vidal, B. Burle, M. Bonnet, J. Grapperon, and T. Hasbroucq, "Error negativity on correct trials: a reexamination of available data," *Biological Psychology*, vol. 64, no. 3, pp. 265–282, Nov. 2003, doi: 10.1016/S0301-0511(03)00097-8.
[65] Y. Arbel and E. Donchin, "Parsing the componential structure of post-error ERPs: a principal component analysis of ERPs following errors," *Psychophysiology*, vol. 46, no. 6, pp. 1179–1189, Nov. 2009, doi: 10.1111/j.1469-8986.2009.00857.x.

[66] R. G. O'Connell *et al.*, "The role of cingulate cortex in the detection of errors with and without awareness: a high-density electrical mapping study," *Eur. J. Neurosci.*, vol. 25, no. 8, pp. 2571–2579, Apr. 2007, doi: 10.1111/j.1460-9568.2007.05477.x.
[67] C. Meckler, S. Allain, L. Carbonnell, T. Hasbroucq, B. Burle, and F. Vidal, "Executive control and response expectancy: A Laplacian ERP study," *Psychophysiology*, vol. 48, no. 3, pp. 303–311, Mar. 2011, doi: 10.1111/j.1469-8986.2010.01077.x.
[68] S. Nieuwenhuis, D. J. Heslenfeld, N. J. A. von Geusau, R. B. Mars, C. B. Holroyd, and N. Yeung, "Activity in human reward-sensitive brain areas is strongly context dependent," *Neuroimage*, vol. 25, no. 4, pp. 1302–1309, May 2005.
[69] N. Yeung and A. G. Sanfey, "Independent coding of reward magnitude and valence in the human brain," *J. Neurosci.*, vol. 24, no. 28, pp. 6258–6264, Jul. 2004, doi: 10.1523/JNEUROSCI.4537-03.2004.
[70] M. A. S. Boksem, M. Tops, A. E. Wester, T. F. Meijman, and M. M. Lorist, "Error-related ERP components and individual differences in punishment and reward sensitivity," *Brain Res.*, vol. 1101, no. 1, pp. 92–101, Jul. 2006, doi: 10.1016/j.brainres.2006.05.004.
[71] M. Falkenstein, R. Willemssen, J. Hohnsbein, and H. Hielscher, "Error Processing in Parkinson's Disease," *Journal of Psychophysiology*, vol. 19, no. 4, pp. 305–310, Jan. 2005, doi: 10.1027/0269-8803.19.4.305.
[72] M. Ullsperger and D. Y. von Cramon, "The role of intact frontostriatal circuits in error processing," *J Cogn Neurosci*, vol. 18, no. 4, pp. 651–664, Apr. 2006, doi: 10.1162/jocn.2006.18.4.651.
[73] M. Ullsperger, H. A. Harsay, J. R. Wessel, and K. R. Ridderinkhof, "Conscious perception of errors and its relation to the anterior insula," *Brain Struct Funct*, vol. 214, no. 5–6, pp. 629–643, Jun. 2010, doi: 10.1007/s00429-010-0261-1.
[74] R. Chavarriaga, A. Sobolewski, and J. D. R. Millán, "Errare machinale est: the use of error-related potentials in brain-machine interfaces," *Front Neurosci*, vol. 8, p. 208, 2014, doi: 10.3389/fnins.2014.00208.
[75] E. Lopez-Larraz, I. Iturrate, L. Montesano, and J. Minguez, "Real-time recognition of feedback error-related potentials during a time-estimation task," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, Aug. 2010, pp. 2670–2673. doi: 10.1109/IEMBS.2010.5626623.
[76] J. Omedes, I. Iturrate, L. Montesano, and J. Minguez, "Using frequency-domain features for the generalization of EEG error-related potentials among different tasks," *Conf Proc IEEE Eng Med Biol Soc*, vol. 2013, pp. 5263–5266, 2013, doi: 10.1109/EMBC.2013.6610736.
[77] M. Spüler and C. Niethammer, "Error-related potentials during continuous feedback: using EEG to detect errors of different type and severity," *Front Hum Neurosci*, vol. 9, Mar. 2015, doi: 10.3389/fnhum.2015.00155.
[78] E. M. Ventouras, P. Asvestas, I. Karanasiou, and G. K. Matsopoulos, "Classification of Error-Related Negativity (ERN) and Positivity (Pe) potentials using kNN and Support Vector Machines," *Comput. Biol. Med.*, vol. 41, no. 2, pp. 98–109, Feb. 2011, doi: 10.1016/j.compbiomed.2010.12.004.
[79] T. Plewan, E. Wascher, M. Falkenstein, and S. Hoffmann, "Classifying Response Correctness across Different Task Sets: A Machine Learning Approach," *PLOS ONE*, vol. 11, no. 3, p. e0152864, Mar. 2016, doi: 10.1371/journal.pone.0152864.
[80] T. O. J. Gründler, J. F. Cavanagh, C. M. Figueroa, M. J. Frank, and J. J. B. Allen, "Task-related dissociation in ERN amplitude as a function of obsessive-compulsive symptoms,"

Neuropsychologia, vol. 47, no. 8–9, pp. 1978–1987, Jul. 2009, doi: 10.1016/j.neuropsychologia.2009.03.010.

[81] S. Hoffmann and M. Falkenstein, "Independent component analysis of erroneous and correct responses suggests online response control," *Hum Brain Mapp*, vol. 31, no. 9, pp. 1305–1315, Sep. 2010, doi: 10.1002/hbm.20937.

[82] A. N. Kaczkurkin, "The effect of manipulating task difficulty on error-related negativity in individuals with obsessive-compulsive symptoms," *Biol Psychol*, vol. 93, no. 1, pp. 122–131, Apr. 2013, doi: 10.1016/j.biopsycho.2013.01.001.

[83] C. A. Mathews, V. B. Perez, K. L. Delucchi, and D. H. Mathalon, "Error-related negativity in individuals with obsessive–compulsive symptoms: Toward an understanding of hoarding behaviors," *Biol Psychol*, vol. 89, no. 2, pp. 487–494, Feb. 2012, doi: 10.1016/j.biopsycho.2011.12.018.

[84] M. Balconi and D. Crivelli, "FRN and P300 ERP effect modulation in response to feedback sensitivity: the contribution of punishment-reward system (BIS/BAS) and behaviour identification of action," *Neurosci. Res.*, vol. 66, no. 2, pp. 162–172, Feb. 2010, doi: 10.1016/j.neures.2009.10.011.

[85] I. Van den Berg, I. H. A. Franken, and P. Muris, "Individual Differences in Sensitivity to Reward," *Journal of Psychophysiology*, vol. 25, no. 2, pp. 81–86, Jan. 2011, doi: 10.1027/0269-8803/a000032.

[86] R. Hester, L. Nestor, and H. Garavan, "Impaired error awareness and anterior cingulate cortex hypoactivity in chronic cannabis users," *Neuropsychopharmacology*, vol. 34, no. 11, pp. 2450–2458, Oct. 2009, doi: 10.1038/npp.2009.67.

[87] J. Bakic, G. Pourtois, M. Jepma, R. Duprat, R. De Raedt, and C. Baeken, "Spared internal but impaired external reward prediction error signals in major depressive disorder during reinforcement learning," *Depress Anxiety*, vol. 34, no. 1, pp. 89–96, Jan. 2017, doi: 10.1002/da.22576.

[88] M. Carrasco, S. M. Harbin, J. K. Nienhuis, K. D. Fitzgerald, W. J. Gehring, and G. L. Hanna, "Increased Error-Related Brain Activity in Youth with Obsessive-Compulsive Disorder and Unaffected Siblings," *Depress Anxiety*, vol. 30, no. 1, pp. 39–46, Jan. 2013, doi: 10.1002/da.22035.

[89] T. Endrass and M. Ullsperger, "Specificity of performance monitoring changes in obsessive-compulsive disorder," *Neuroscience & Biobehavioral Reviews*, vol. 46, Part 1, pp. 124–138, Oct. 2014, doi: 10.1016/j.neubiorev.2014.03.024.

[90] J. Klawohn, A. Riesel, R. Grützmann, N. Kathmann, and T. Endrass, "Performance monitoring in obsessive-compulsive disorder: a temporo-spatial principal component analysis," *Cogn Affect Behav Neurosci*, vol. 14, no. 3, pp. 983–995, Sep. 2014, doi: 10.3758/s13415-014-0248-0.

[91] J. S. Moser, T. P. Moran, H. S. Schroder, M. B. Donnellan, and N. Yeung, "On the relationship between anxiety and error monitoring: a meta-analysis and conceptual framework," *Front Hum Neurosci*, vol. 7, Aug. 2013, doi: 10.3389/fnhum.2013.00466.

[92] A. Weinberg, A. Riesel, and G. Hajcak, "Integrating multiple perspectives on error-related brain activity: The ERN as a neural indicator of trait defensive reactivity," *Motiv Emot*, vol. 36, no. 1, pp. 84–100, Mar. 2012, doi: 10.1007/s11031-011-9269-y.

[93] A. Weinberg, R. Dieterich, and A. Riesel, "Error-related brain activity in the age of RDoC: A review of the literature," *International Journal of Psychophysiology*, vol. 98, no. 2, Part 2, pp. 276–299, Nov. 2015, doi: 10.1016/j.ijpsycho.2015.02.029.

[94] I. Iturrate, L. Montesano, and J. Minguez, "Task-dependent signal variations in EEG error-related potentials for brain–computer interfaces," *J. Neural Eng.*, vol. 10, no. 2, p. 026024, 2013, doi: 10.1088/1741-2560/10/2/026024.
[95] L. Van der Borght, F. Houtman, B. Burle, and W. Notebaert, "Distinguishing the influence of task difficulty on error-related ERPs using surface Laplacian transformation," *Biological Psychology*, vol. 115, Jan. 2016, doi: 10.1016/j.biopsycho.2016.01.013.
[96] T. Endrass, J. Klawohn, R. Gruetzmann, M. Ischebeck, and N. Kathmann, "Response-related negativities following correct and incorrect responses: evidence from a temporospatial principal component analysis," *Psychophysiology*, vol. 49, no. 6, pp. 733–743, Jun. 2012, doi: 10.1111/j.1469-8986.2012.01365.x.
[97] B. C. Moore and B. R. Glasberg, "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.*, vol. 74, no. 3, pp. 750–753, Sep. 1983.
[98] R. Chavarriaga and J. d R. Millan, "Learning From EEG Error-Related Potentials in Noninvasive Brain-Computer Interfaces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 4, pp. 381–388, Aug. 2010, doi: 10.1109/TNSRE.2010.2053387.
[99] P. W. Ferrez and J. del R Millan, "Error-related EEG potentials generated during simulated brain-computer interaction," *IEEE Trans Biomed Eng*, vol. 55, no. 3, pp. 923–929, Mar. 2008, doi: 10.1109/TBME.2007.908083.
[100] S. Theodoridis and K. Koutroumbas, *Pattern Recognition, Fourth Edition*, 4th ed. Academic Press, 2008.
[101] M. Inzlicht and T. Al-Khindi, "ERN and the placebo: A misattribution approach to studying the arousal properties of the error-related negativity," *Journal of Experimental Psychology: General*, vol. 141, no. 4, pp. 799–807, 2012, doi: 10.1037/a0027586.
[102] I. Iturrate, R. Chavarriaga, L. Montesano, J. Minguez, and J. Millán, "Latency correction of event-related potentials between different experimental protocols," *J Neural Eng*, vol. 11, no. 3, p. 036005, Jun. 2014, doi: 10.1088/1741-2560/11/3/036005.
[103] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1023/A:1022627411411.
[104] P. Pudil, J. Novovičová, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognition Letters*, vol. 15, no. 11, pp. 1119–1125, Nov. 1994, doi: 10.1016/0167-8655(94)90127-9.
[105] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, Jan. 2000, doi: 10.1109/34.824819.
[106] T.-T. Wong, "Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation," *Pattern Recognition*, vol. 48, no. 9, pp. 2839–2846, Sep. 2015, doi: 10.1016/j.patcog.2015.03.009.
[107] N. Y. Hammerla and T. Plötz, "Let's (Not) Stick Together: Pairwise Similarity Biases Cross-validation in Activity Recognition," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, New York, NY, USA, 2015, pp. 1041–1051. doi: 10.1145/2750858.2807551.
[108] A. Kumar, E. Pirogova, and J. Q. Fang, "Classification of Error-Related Potentials using Linear Discriminant Analysis," in *2018 IEEE-EMBS Conference on Biomedical*

Engineering and Sciences (IECBES), Dec. 2018, pp. 18–21. doi: 10.1109/IECBES.2018.8626709.

[109] I. S. Karanasiou, C. Papageorgiou, E. I. Tsianaka, G. K. Matsopoulos, E. M. Ventouras, and N. K. Uzunoglu, "Behavioral and brain pattern differences between acting and observing in an auditory task," *Behav Brain Funct*, vol. 5, p. 5, Jan. 2009, doi: 10.1186/1744-9081-5-5.

[110] J. van Driel, K. R. Ridderinkhof, and M. X. Cohen, "Not all errors are alike: theta and alpha EEG dynamics relate to differences in error-processing dynamics," *J. Neurosci.*, vol. 32, no. 47, pp. 16795–16806, Nov. 2012, doi: 10.1523/JNEUROSCI.0802-12.2012.

[111] A. Gentsch, P. Ullsperger, and M. Ullsperger, "Dissociable medial frontal negativities from a common monitoring system for self- and externally caused failure of goal achievement," *Neuroimage*, vol. 47, no. 4, pp. 2023–2030, Oct. 2009, doi: 10.1016/j.neuroimage.2009.05.064.

[112] R. Singla, B. Chambayil, A. Khosla, and J. Santosh, "Comparison of SVM and ANN for classification of eye events in EEG," *Journal of Biomedical Science and Engineering*, vol. 04, no. 01, pp. 62–69, 2011, doi: 10.4236/jbise.2011.41008.

[113] H. Parvar, L. Sculthorpe-Petley, J. Satel, R. Boshra, R. C. N. D'Arcy, and T. P. Trappenberg, "Detection of event-related potentials in individual subjects using support vector machines," *Brain Inform*, vol. 2, no. 1, pp. 1–12, Nov. 2014, doi: 10.1007/s40708-014-0006-7.

[114] N. Nicolaou and J. Georgiou, "Detection of epileptic electroencephalogram based on Permutation Entropy and Support Vector Machines," *Expert Systems with Applications*, vol. 39, no. 1, pp. 202–209, Jan. 2012, doi: 10.1016/j.eswa.2011.07.008.

[115] T. U. Hauser *et al.*, "The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization," *NeuroImage*, vol. 84, no. Supplement C, pp. 159–168, Jan. 2014, doi: 10.1016/j.neuroimage.2013.08.028.

[116] R. Grützmann, T. Endrass, J. Klawohn, and N. Kathmann, "Response accuracy rating modulates ERN and Pe amplitudes," *Biological Psychology*, vol. 96, no. Supplement C, pp. 1–7, Feb. 2014, doi: 10.1016/j.biopsycho.2013.10.007.

[117] T. E. Baker and C. B. Holroyd, "Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200," *Biol Psychol*, vol. 87, no. 1, pp. 25–34, Apr. 2011, doi: 10.1016/j.biopsycho.2011.01.010.

[118] M. Gawlowska, A. Domagalik, E. Beldzik, T. Marek, and J. Mojsa-Kaja, "Dynamics of error-related activity in deterministic learning - an EEG and fMRI study," *Sci Rep*, vol. 8, no. 1, p. 14617, Dec. 2018, doi: 10.1038/s41598-018-32995-x.

[119] N. A. Choudhury, J. A. Parascando, and A. A. Benasich, "Effects of Presentation Rate and Attention on Auditory Discrimination: A Comparison of Long-Latency Auditory Evoked Potentials in School-Aged Children and Adults," *PLOS ONE*, vol. 10, no. 9, p. e0138160, Sep. 2015, doi: 10.1371/journal.pone.0138160.

[120] N. K. Ferdinand, A. Mecklinger, J. Kray, and W. J. Gehring, "The processing of unexpected positive response outcomes in the mediofrontal cortex," *J. Neurosci.*, vol. 32, no. 35, pp. 12087–12092, Aug. 2012, doi: 10.1523/JNEUROSCI.1410-12.2012.

[121] L. Kreussel, J. Hewig, N. Kretschmer, H. Hecht, M. G. H. Coles, and W. H. R. Miltner, "The influence of the magnitude, probability, and valence of potential wins and losses on

the amplitude of the feedback negativity," *Psychophysiology*, vol. 49, no. 2, pp. 207–219, Feb. 2012, doi: 10.1111/j.1469-8986.2011.01291.x.

[122] B. Opitz, N. K. Ferdinand, and A. Mecklinger, "Timing Matters: The Impact of Immediate and Delayed Feedback on Artificial Language Learning," *Front Hum Neurosci*, vol. 5, Feb. 2011, doi: 10.3389/fnhum.2011.00008.

[123] O. E. Krigolson, C. D. Hassall, and T. C. Handy, "How we learn to make decisions: rapid propagation of reinforcement learning prediction errors in humans," *J Cogn Neurosci*, vol. 26, no. 3, pp. 635–644, Mar. 2014, doi: 10.1162/jocn_a_00509.

[124] H. W. Chase, R. Swainson, L. Durham, L. Benham, and R. Cools, "Feedback-related negativity codes prediction error but not behavioral adjustment during probabilistic reversal learning," *J Cogn Neurosci*, vol. 23, no. 4, pp. 936–946, Apr. 2011, doi: 10.1162/jocn.2010.21456.

[125] M. Spüler and C. Niethammer, "Error-related potentials during continuous feedback: using EEG to detect errors of different type and severity," *Front. Hum. Neurosci.*, vol. 9, 2015, doi: 10.3389/fnhum.2015.00155.

[126] C. L. Baldwin and B. N. Penaranda, "Adaptive training using an artificial neural network and EEG metrics for within- and cross-task workload classification," *NeuroImage*, vol. 59, no. 1, pp. 48–56, Jan. 2012, doi: 10.1016/j.neuroimage.2011.07.047.

[127] E. López-Larraz, M. Creatura, I. Iturrate, L. Montesano, and J. Minguez, "EEG single-trial classification of visual, auditive and vibratory feedback potentials in Brain-Computer Interfaces," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2011, pp. 4231–4234. doi: 10.1109/IEMBS.2011.6091050.

[128] Y. Sun, A. K. C. Wong, and M. S. Kamel, "Classification of imbalanced data: a review," *Int. J. Patt. Recogn. Artif. Intell.*, vol. 23, no. 04, pp. 687–719, Jun. 2009, doi: 10.1142/S0218001409007326.

[129] S. Haufe *et al.*, "On the interpretation of weight vectors of linear models in multivariate neuroimaging," *Neuroimage*, vol. 87, pp. 96–110, Feb. 2014, doi: 10.1016/j.neuroimage.2013.10.067.

[130] L. Kester and P. A. Kirschner, "Cognitive Tasks and Learning," in *Encyclopedia of the Sciences of Learning*, N. M. Seel, Ed. Boston, MA: Springer US, 2012, pp. 619–622. doi: 10.1007/978-1-4419-1428-6_225.

[131] A. Skulmowski and G. D. Rey, "Measuring Cognitive Load in Embodied Learning Settings," *Front Psychol*, vol. 8, Aug. 2017, doi: 10.3389/fpsyg.2017.01191.

[132] A. Seidler *et al.*, "The role of psychosocial working conditions on burnout and its core component emotional exhaustion – a systematic review," *J Occup Med Toxicol*, vol. 9, no. 1, p. 10, Mar. 2014, doi: 10.1186/1745-6673-9-10.

[133] M. S. Young, K. A. Brookhuis, C. D. Wickens, and P. A. Hancock, "State of science: mental workload in ergonomics," *Ergonomics*, vol. 58, no. 1, pp. 1–17, 2015, doi: 10.1080/00140139.2014.956151.

[134] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness," *Neuroscience & Biobehavioral Reviews*, vol. 44, pp. 58–75, Jul. 2014, doi: 10.1016/j.neubiorev.2012.10.003.

[135] Y. Lean and F. Shan, "Brief review on physiological and biochemical evaluations of human mental workload," *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 22, no. 3, pp. 177–187, 2012, doi: 10.1002/hfm.20269.

[136] D. L. Schomer and F. H. L. da Silva, *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Oxford University Press, 2018.
[137] C. Mühl, C. Jeunet, and F. Lotte, "EEG-based workload estimation across affective contexts," *Front. Neurosci.*, vol. 8, 2014, doi: 10.3389/fnins.2014.00114.
[138] P. Antonenko, F. Paas, R. Grabner, and T. van Gog, "Using Electroencephalography to Measure Cognitive Load," *Educ Psychol Rev*, vol. 22, no. 4, pp. 425–438, Dec. 2010, doi: 10.1007/s10648-010-9130-y.
[139] C. M. Krause, M. Pesonen, and H. Hämäläinen, "Brain oscillatory 4-30 Hz electroencephalogram responses in adolescents during a visual memory task," *Neuroreport*, vol. 21, no. 11, pp. 767–771, Aug. 2010, doi: 10.1097/WNR.0b013e32833bfcbb.
[140] A. F. Rabbi, A. Zony, P. de Leon, and R. Fazel-Rezai, "Mental workload and task engagement evaluation based on changes in electroencephalogram," *Biomed. Eng. Lett.*, vol. 2, no. 3, pp. 139–146, Sep. 2012, doi: 10.1007/s13534-012-0065-8.
[141] Y. Chen and X. Huang, "Modulation of Alpha and Beta Oscillations during an n-back Task with Varying Temporal Memory Load," *Front. Psychol.*, vol. 6, 2016, doi: 10.3389/fpsyg.2015.02031.
[142] S. Puma, N. Matton, P.-V. Paubel, É. Raufaste, and R. El-Yagoubi, "Using theta and alpha band power to assess cognitive workload in multitasking environments," *Int J Psychophysiol*, vol. 123, pp. 111–120, 2018, doi: 10.1016/j.ijpsycho.2017.10.004.
[143] T. Radüntz, "The Effect of Planning, Strategy Learning, and Working Memory Capacity on Mental Workload," *Scientific Reports*, vol. 10, no. 1, p. 7096, Apr. 2020, doi: 10.1038/s41598-020-63897-6.
[144] R. N. Roy, S. Charbonnier, A. Campagne, and S. Bonnet, "Efficient mental workload estimation using task-independent EEG features," *J Neural Eng*, vol. 13, no. 2, p. 026019, Apr. 2016, doi: 10.1088/1741-2560/13/2/026019.
[145] S. Wang, J. Gwizdka, and W. A. Chaovalitwongse, "Using Wireless EEG Signals to Assess Memory Workload in the n-Back Task," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 3, pp. 424–435, Jun. 2016, doi: 10.1109/THMS.2015.2476818.
[146] P. Zarjam, J. Epps, F. Chen, and N. H. Lovell, "Estimating cognitive workload using wavelet entropy-based features during an arithmetic task," *Computers in Biology and Medicine*, vol. 43, no. 12, pp. 2186–2195, Dec. 2013, doi: 10.1016/j.compbiomed.2013.08.021.
[147] M. Spüler, C. Walter, W. Rosenstiel, P. Gerjets, K. Moeller, and E. Klein, "EEG-based prediction of cognitive workload induced by arithmetic: a step towards online adaptation in numerical learning," *ZDM Mathematics Education*, vol. 48, no. 3, pp. 267–278, Jun. 2016, doi: 10.1007/s11858-015-0754-8.
[148] C. L. Baldwin and B. N. Penaranda, "Adaptive training using an artificial neural network and EEG metrics for within- and cross-task workload classification," *Neuroimage*, vol. 59, no. 1, pp. 48–56, Jan. 2012, doi: 10.1016/j.neuroimage.2011.07.047.
[149] C. Walter, S. Schmidt, W. Rosenstiel, P. Gerjets, and M. Bogdan, "Using Cross-Task Classification for Classifying Workload Levels in Complex Learning Tasks," in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, Sep. 2013, pp. 876–881. doi: 10.1109/ACII.2013.164.
[150] Y. Ke *et al.*, "Towards an effective cross-task mental workload recognition model using electroencephalography based on feature selection and support vector machine regression,"

International Journal of Psychophysiology, vol. 98, no. 2, Part 1, pp. 157–166, Nov. 2015, doi: 10.1016/j.ijpsycho.2015.10.004.

[151] A. Gevins *et al.*, "Monitoring working memory load during computer-based tasks with EEG pattern recognition methods," *Hum Factors*, vol. 40, no. 1, pp. 79–91, Mar. 1998, doi: 10.1518/001872098779480578.

[152] P. Zhang, X. Wang, W. Zhang, and J. Chen, "Learning Spatial-Spectral-Temporal EEG Features With Recurrent 3D Convolutional Neural Networks for Cross-Task Mental Workload Assessment," *IEEE Trans Neural Syst Rehabil Eng*, vol. 27, no. 1, pp. 31–42, 2019, doi: 10.1109/TNSRE.2018.2884641.

[153] G. Zhao, Y.-J. Liu, and Y. Shi, "Real-Time Assessment of the Cross-Task Mental Workload Using Physiological Measures During Anomaly Detection," *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 2, pp. 149–160, Apr. 2018, doi: 10.1109/THMS.2018.2803025.

[154] E. Debie *et al.*, "Multimodal Fusion for Objective Assessment of Cognitive Workload: A Review," *IEEE Trans Cybern*, Sep. 2019, doi: 10.1109/TCYB.2019.2939399.

[155] D. S. Bassett and O. Sporns, "Network neuroscience," *Nat Neurosci*, vol. 20, no. 3, pp. 353–364, Feb. 2017, doi: 10.1038/nn.4502.

[156] Z. Dai *et al.*, "EEG Cortical Connectivity Analysis of Working Memory Reveals Topological Reorganization in Theta and Alpha Bands," *Front Hum Neurosci*, vol. 11, p. 237, 2017, doi: 10.3389/fnhum.2017.00237.

[157] M. G. Kitzbichler, R. N. A. Henson, M. L. Smith, P. J. Nathan, and E. T. Bullmore, "Cognitive Effort Drives Workspace Configuration of Human Brain Functional Networks," *J. Neurosci.*, vol. 31, no. 22, pp. 8259–8270, Jun. 2011, doi: 10.1523/JNEUROSCI.0440-11.2011.

[158] M. A. Klados *et al.*, "A Graph Theoretical Approach to Study the Organization of the Cortical Networks during Different Mathematical Tasks," *PLoS One*, vol. 8, no. 8, Aug. 2013, doi: 10.1371/journal.pone.0071800.

[159] G. N. Dimitrakopoulos *et al.*, "Functional Connectivity Analysis of Mental Fatigue Reveals Different Network Topological Alterations Between Driving and Vigilance Tasks," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 4, pp. 740–749, Apr. 2018, doi: 10.1109/TNSRE.2018.2791936.

[160] S. I. Dimitriadis, Y. Sun, K. Kwok, N. A. Laskaris, N. Thakor, and A. Bezerianos, "Cognitive Workload Assessment Based on the Tensorial Treatment of EEG Estimates of Cross-Frequency Phase Interactions," *Ann Biomed Eng*, vol. 43, no. 4, pp. 977–989, Apr. 2015, doi: 10.1007/s10439-014-1143-0.

[161] I. Kakkos *et al.*, "Mental Workload Drives Different Reorganizations of Functional Cortical Connectivity Between 2D and 3D Simulated Flight Experiments," *IEEE Trans Neural Syst Rehabil Eng*, vol. 27, no. 9, pp. 1704–1713, Sep. 2019, doi: 10.1109/TNSRE.2019.2930082.

[162] G. N. Dimitrakopoulos *et al.*, "Task-Independent Mental Workload Classification Based Upon Common Multiband EEG Cortical Connectivity," *IEEE Trans Neural Syst Rehabil Eng*, vol. 25, no. 11, pp. 1940–1949, 2017, doi: 10.1109/TNSRE.2017.2701002.

[163] A. Baddeley, "Working Memory: Theories, Models, and Controversies," *Annual Review of Psychology*, vol. 63, no. 1, pp. 1–29, 2012, doi: 10.1146/annurev-psych-120710-100422.

[164] A. Diamond, "Executive Functions," *Annu Rev Psychol*, vol. 64, pp. 135–168, 2013, doi: 10.1146/annurev-psych-113011-143750.

[165] T. P. Jung *et al.*, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 2, pp. 163–178, Mar. 2000.
[166] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004, doi: 10.1016/j.jneumeth.2003.10.009.
[167] K. Yan and D. Zhang, "Feature selection and analysis on correlated gas sensor data with recursive feature elimination," *Sensors and Actuators B: Chemical*, vol. 212, pp. 353–363, Jun. 2015, doi: 10.1016/j.snb.2015.02.025.
[168] S. Abe, *Support Vector Machines for Pattern Classification*. London ; New York: Springer, 2010.
[169] P. Golland, F. Liang, S. Mukherjee, and D. Panchenko, "Permutation Tests for Classification," in *Learning Theory*, Berlin, Heidelberg, 2005, pp. 501–515. doi: 10.1007/11503415_34.
[170] T. G. Dietterich, "Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms," *Neural Computation*, vol. 10, no. 7, pp. 1895–1923, Oct. 1998, doi: 10.1162/089976698300017197.
[171] G. Morse, M. P. Salyers, A. L. Rollins, M. Monroe-DeVita, and C. Pfahler, "Burnout in mental health services: a review of the problem and its remediation," *Adm Policy Ment Health*, vol. 39, no. 5, pp. 341–352, Sep. 2012, doi: 10.1007/s10488-011-0352-1.
[172] I. Kakkos *et al.*, "EEG fingerprints of task-independent mental workload discrimination," *IEEE J Biomed Health Inform*, vol. PP, Jun. 2021, doi: 10.1109/JBHI.2021.3085131.
[173] I. Iturrate, L. Montesano, and J. Minguez, "Task-dependent signal variations in EEG error-related potentials for brain-computer interfaces," *J Neural Eng*, vol. 10, no. 2, p. 026024, Apr. 2013, doi: 10.1088/1741-2560/10/2/026024.
[174] L. Tolosi and T. Lengauer, "Classification with correlated features: unreliability of feature ranking and solutions," *Bioinformatics*, vol. 27, no. 14, pp. 1986–1994, Jul. 2011, doi: 10.1093/bioinformatics/btr300.
[175] I. Imbo, A. Vandierendonck, and S. De Rammelaere, "The role of working memory in the carry operation of mental arithmetic: number and value of the carry," *Q J Exp Psychol (Hove)*, vol. 60, no. 5, pp. 708–731, May 2007, doi: 10.1080/17470210600762447.
[176] D. DeStefano and J.-A. LeFevre, "The role of working memory in mental arithmetic," *European Journal of Cognitive Psychology*, vol. 16, no. 3, pp. 353–386, May 2004, doi: 10.1080/09541440244000328.
[177] A.-M. Brouwer, M. A. Hogervorst, J. B. F. van Erp, T. Heffelaar, P. H. Zimmerman, and R. Oostenveld, "Estimating workload using EEG spectral power and ERPs in the n-back task," *J Neural Eng*, vol. 9, no. 4, p. 045008, Aug. 2012, doi: 10.1088/1741-2560/9/4/045008.
[178] I. Seleznov *et al.*, "Detrended Fluctuation, Coherence, and Spectral Power Analysis of Activation Rearrangement in EEG Dynamics During Cognitive Workload," *Front. Hum. Neurosci.*, vol. 13, 2019, doi: 10.3389/fnhum.2019.00270.
[179] P. S. Cooper *et al.*, "Theta frontoparietal connectivity associated with proactive and reactive cognitive control processes," *Neuroimage*, vol. 108, pp. 354–363, Mar. 2015, doi: 10.1016/j.neuroimage.2014.12.028.
[180] C. Babiloni *et al.*, "Human cortical responses during one-bit short-term memory. A high-resolution EEG study on delayed choice reaction time tasks," *Clinical Neurophysiology*, vol. 115, no. 1, pp. 161–170, Jan. 2004, doi: 10.1016/S1388-2457(03)00286-4.

[181] A. Ardestani, W. Shen, F. Darvas, A. W. Toga, and J. M. Fuster, "Modulation of Frontoparietal Neurovascular Dynamics in Working Memory," *Journal of Cognitive Neuroscience*, vol. 28, no. 3, pp. 379–401, Dec. 2015, doi: 10.1162/jocn_a_00903.
[182] K. Fukuda, I. Mance, and E. K. Vogel, "α Power Modulation and Event-Related Slow Wave Provide Dissociable Correlates of Visual Working Memory," *J. Neurosci.*, vol. 35, no. 41, pp. 14009–14016, Oct. 2015, doi: 10.1523/JNEUROSCI.5003-14.2015.
[183] M. Crespo-Garcia, D. Pinal, J. L. Cantero, F. Díaz, M. Zurrón, and M. Atienza, "Working memory processes are mediated by local and long-range synchronization of alpha oscillations," *J Cogn Neurosci*, vol. 25, no. 8, pp. 1343–1357, Aug. 2013, doi: 10.1162/jocn_a_00379.
[184] A. Capilla, J.-M. Schoffelen, G. Paterson, G. Thut, and J. Gross, "Dissociated α-Band Modulations in the Dorsal and Ventral Visual Pathways in Visuospatial Attention and Perception," *Cereb Cortex*, vol. 24, no. 2, pp. 550–561, Feb. 2014, doi: 10.1093/cercor/bhs343.
[185] H. Kondo, M. Morishita, N. Osaka, M. Osaka, H. Fukuyama, and H. Shibasaki, "Functional roles of the cingulo-frontal network in performance on working memory," *NeuroImage*, vol. 21, no. 1, pp. 2–14, Jan. 2004, doi: 10.1016/j.neuroimage.2003.09.046.
[186] P. Zarjam, J. Epps, and N. H. Lovell, "Beyond Subjective Self-Rating: EEG Signal Classification of Cognitive Workload," *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 4, pp. 301–310, Dec. 2015, doi: 10.1109/TAMD.2015.2441960.
[187] S. Hanouneh, H. U. Amin, N. M. Saad, and A. S. Malik, "EEG Power and Functional Connectivity Correlates with Semantic Long-Term Memory Retrieval," *IEEE Access*, vol. 6, pp. 8695–8703, 2018, doi: 10.1109/ACCESS.2017.2788859.
[188] Y. Liu, H. Ayaz, and P. A. Shewokis, "Multisubject 'Learning' for Mental Workload Classification Using Concurrent EEG, fNIRS, and Physiological Measures," *Front. Hum. Neurosci.*, vol. 11, 2017, doi: 10.3389/fnhum.2017.00389.
[189] R. Gulbinaite, H. van Rijn, and M. X. Cohen, "Fronto-parietal network oscillations reveal relationship between working memory capacity and cognitive control," *Front. Hum. Neurosci.*, vol. 8, 2014, doi: 10.3389/fnhum.2014.00761.
[190] B.-W. Hsu, M.-J. J. Wang, C.-Y. Chen, and F. Chen, "Effective Indices for Monitoring Mental Workload While Performing Multiple Tasks," *Percept Mot Skills*, vol. 121, no. 1, pp. 94–117, Aug. 2015, doi: 10.2466/22.PMS.121c12x5.
[191] C. Walter, W. Rosenstiel, M. Bogdan, P. Gerjets, and M. Spüler, "Online EEG-Based Workload Adaptation of an Arithmetic Learning Environment," *Front Hum Neurosci*, vol. 11, May 2017, doi: 10.3389/fnhum.2017.00286.
[192] M. Bocková, J. Chládek, P. Jurák, J. Halámek, and I. Rektor, "Executive functions processed in the frontal and lateral temporal cortices: intracerebral study," *Clin Neurophysiol*, vol. 118, no. 12, pp. 2625–2636, Dec. 2007, doi: 10.1016/j.clinph.2007.07.025.
[193] P. Putman, J. van Peer, I. Maimari, and S. van der Werff, "EEG theta/beta ratio in relation to fear-modulated response-inhibition, attentional control, and affective traits," *Biol Psychol*, vol. 83, no. 2, pp. 73–78, Feb. 2010, doi: 10.1016/j.biopsycho.2009.10.008.
[194] T. P. Zanto and A. Gazzaley, "Neural Suppression of Irrelevant Information Underlies Optimal Working Memory Performance," *J Neurosci*, vol. 29, no. 10, pp. 3059–3066, Mar. 2009, doi: 10.1523/JNEUROSCI.4621-08.2009.

[195] G. T. Waldhauser, M. Johansson, and S. Hanslmayr, "Alpha/Beta Oscillations Indicate Inhibition of Interfering Visual Memories," *J. Neurosci.*, vol. 32, no. 6, pp. 1953–1961, Feb. 2012, doi: 10.1523/JNEUROSCI.4201-11.2012.

[196] S. Haufe *et al.*, "On the interpretation of weight vectors of linear models in multivariate neuroimaging," *NeuroImage*, vol. 87, pp. 96–110, Feb. 2014, doi: 10.1016/j.neuroimage.2013.10.067.

[197] F. Chen *et al.*, "Multimodal behavior and interaction as indicators of cognitive load," *ACM Trans. Interact. Intell. Syst.*, vol. 2, no. 4, pp. 22:1–22:36, Jan. 2013, doi: 10.1145/2395123.2395127.

[198] A. Heathcote, J. R. Coleman, A. Eidels, J. M. Watson, J. Houpt, and D. L. Strayer, "Working memory's workload capacity," *Mem Cogn*, vol. 43, no. 7, pp. 973–989, Oct. 2015, doi: 10.3758/s13421-015-0526-2.

[199] C. Scharinger, A. Soutschek, T. Schubert, and P. Gerjets, "When flanker meets the n-back: What EEG and pupil dilation data reveal about the interplay between the two central-executive working memory functions inhibition and updating," *Psychophysiology*, vol. 52, no. 10, pp. 1293–1304, Oct. 2015, doi: 10.1111/psyp.12500.

[200] C. Gilmore, S. Keeble, S. Richardson, and L. Cragg, "The role of cognitive inhibition in different components of arithmetic," *ZDM Mathematics Education*, vol. 47, no. 5, pp. 771–782, Sep. 2015, doi: 10.1007/s11858-014-0659-y.

[201] S. Palva and J. M. Palva, "Functional roles of alpha-band phase synchronization in local and large-scale cortical networks," *Front Psychol*, vol. 2, p. 204, 2011, doi: 10.3389/fpsyg.2011.00204.

[202] D. R. Thomson, D. Besner, and D. Smilek, "A resource-control account of sustained attention: evidence from mind-wandering and vigilance paradigms," *Perspect Psychol Sci*, vol. 10, no. 1, pp. 82–96, Jan. 2015, doi: 10.1177/1745691614556681.

[203] N. Meshkati and P. A. Hancock, *Human Mental Workload*. Elsevier, 2011.

[204] H. Ayaz, P. A. Shewokis, S. Bunce, K. Izzetoglu, B. Willems, and B. Onaral, "Optical brain monitoring for operator training and mental workload assessment," *Neuroimage*, vol. 59, no. 1, pp. 36–47, Jan. 2012, doi: 10.1016/j.neuroimage.2011.06.023.

[205] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness," *Neuroscience & Biobehavioral Reviews*, vol. 44, pp. 58–75, Jul. 2014, doi: 10.1016/j.neubiorev.2012.10.003.

[206] T. Gateau, H. Ayaz, and F. Dehais, "In silico vs. Over the Clouds: On-the-Fly Mental State Estimation of Aircraft Pilots, Using a Functional Near Infrared Spectroscopy Based Passive-BCI," *Front. Hum. Neurosci.*, vol. 12, 2018, doi: 10.3389/fnhum.2018.00187.

[207] E. Galy, M. Cariou, and C. Mélan, "What is the relationship between mental workload factors and cognitive load types?," *International Journal of Psychophysiology*, vol. 83, no. 3, pp. 269–275, Mar. 2012, doi: 10.1016/j.ijpsycho.2011.09.023.

[208] M. S. Young, K. A. Brookhuis, C. D. Wickens, and P. A. Hancock, "State of science: mental workload in ergonomics," *Ergonomics*, vol. 58, no. 1, pp. 1–17, Jan. 2015, doi: 10.1080/00140139.2014.956151.

[209] S. Lei and M. Roetting, "Influence of task combination on EEG spectrum modulation for driver workload estimation," *Hum Factors*, vol. 53, no. 2, pp. 168–179, Apr. 2011, doi: 10.1177/0018720811400601.

[210] P. Antonenko, F. Paas, R. Grabner, and T. van Gog, "Using Electroencephalography to Measure Cognitive Load," *Educ Psychol Rev*, vol. 22, no. 4, pp. 425–438, Dec. 2010, doi: 10.1007/s10648-010-9130-y.
[211] M. A. Hogervorst, A.-M. Brouwer, and J. B. F. van Erp, "Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload," *Front. Neurosci.*, vol. 8, 2014, doi: 10.3389/fnins.2014.00322.
[212] C. Mühl, C. Jeunet, and F. Lotte, "EEG-based workload estimation across affective contexts," *Front. Neurosci.*, vol. 8, 2014, doi: 10.3389/fnins.2014.00114.
[213] A.-M. Brouwer, M. A. Hogervorst, J. B. F. van Erp, T. Heffelaar, P. H. Zimmerman, and R. Oostenveld, "Estimating workload using EEG spectral power and ERPs in the n-back task," *J Neural Eng*, vol. 9, no. 4, p. 045008, Aug. 2012, doi: 10.1088/1741-2560/9/4/045008.
[214] P. Zarjam, J. Epps, and F. Chen, "Characterizing working memory load using EEG delta activity," in *2011 19th European Signal Processing Conference*, Aug. 2011, pp. 1554–1558.
[215] C. Dijksterhuis, D. de Waard, K. Brookhuis, B. Mulder, and R. de Jong, "Classifying visuomotor workload in a driving simulator using subject specific spatial brain patterns," *Front. Neurosci.*, vol. 7, 2013, doi: 10.3389/fnins.2013.00149.
[216] P. Zarjam, J. Epps, F. Chen, and N. H. Lovell, "Estimating cognitive workload using wavelet entropy-based features during an arithmetic task," *Computers in Biology and Medicine*, vol. 43, no. 12, pp. 2186–2195, Dec. 2013, doi: 10.1016/j.compbiomed.2013.08.021.
[217] M. Spüler, C. Walter, W. Rosenstiel, P. Gerjets, K. Moeller, and E. Klein, "EEG-based prediction of cognitive workload induced by arithmetic: a step towards online adaptation in numerical learning," *ZDM Mathematics Education*, vol. 48, no. 3, pp. 267–278, Jun. 2016, doi: 10.1007/s11858-015-0754-8.
[218] R. N. Roy, S. Bonnet, S. Charbonnier, and A. Campagne, "Mental fatigue and working memory load estimation: Interaction and implications for EEG-based passive BCI," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jul. 2013, pp. 6607–6610. doi: 10.1109/EMBC.2013.6611070.
[219] G. Borghini *et al.*, "Assessment of mental fatigue during car driving by using high resolution EEG activity and neurophysiologic indices," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2012, pp. 6442–6445. doi: 10.1109/EMBC.2012.6347469.
[220] O. Sporns, "The human connectome: a complex network," *Ann. N. Y. Acad. Sci.*, vol. 1224, pp. 109–125, Apr. 2011, doi: 10.1111/j.1749-6632.2010.05888.x.
[221] S. I. Dimitriadis, Y. Sun, K. Kwok, N. A. Laskaris, N. Thakor, and A. Bezerianos, "Cognitive Workload Assessment Based on the Tensorial Treatment of EEG Estimates of Cross-Frequency Phase Interactions," *Ann Biomed Eng*, vol. 43, no. 4, pp. 977–989, Apr. 2015, doi: 10.1007/s10439-014-1143-0.
[222] R. Vijayalakshmi, D. Nandagopal, N. Dasari, B. Cocks, N. Dahal, and M. Thilaga, "Minimum connected component – A novel approach to detection of cognitive load induced changes in functional brain networks," *Neurocomputing*, vol. 170, pp. 15–31, Dec. 2015, doi: 10.1016/j.neucom.2015.03.092.
[223] M. G. Kitzbichler, R. N. A. Henson, M. L. Smith, P. J. Nathan, and E. T. Bullmore, "Cognitive Effort Drives Workspace Configuration of Human Brain Functional Networks,"

J. Neurosci., vol. 31, no. 22, pp. 8259–8270, Jun. 2011, doi: 10.1523/JNEUROSCI.0440-11.2011.

[224] D. Huang *et al.*, "Combining Partial Directed Coherence and Graph Theory to Analyse Effective Brain Networks of Different Mental Tasks," *Front. Hum. Neurosci.*, vol. 10, 2016, doi: 10.3389/fnhum.2016.00235.

[225] M. A. Klados *et al.*, "A Graph Theoretical Approach to Study the Organization of the Cortical Networks during Different Mathematical Tasks," *PLOS ONE*, vol. 8, no. 8, p. e71800, Aug. 2013, doi: 10.1371/journal.pone.0071800.

[226] G. N. Dimitrakopoulos *et al.*, "Functional connectivity analysis of mental fatigue reveals different network topological alterations between driving and vigilance tasks," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. PP, no. 99, pp. 1–1, 2018, doi: 10.1109/TNSRE.2018.2791936.

[227] Y. Sun, J. Lim, J. Meng, K. Kwok, N. Thakor, and A. Bezerianos, "Discriminative analysis of brain functional connectivity patterns for mental fatigue classification," *Ann Biomed Eng*, vol. 42, no. 10, pp. 2084–2094, Oct. 2014, doi: 10.1007/s10439-014-1059-8.

[228] W. R. Shirer, S. Ryali, E. Rykhlevskaia, V. Menon, and M. D. Greicius, "Decoding Subject-Driven Cognitive States with Whole-Brain Connectivity Patterns," *Cereb Cortex*, vol. 22, no. 1, pp. 158–165, Jan. 2012, doi: 10.1093/cercor/bhr099.

[229] G. N. Dimitrakopoulos *et al.*, "Task-Independent Mental Workload Classification Based Upon Common Multiband EEG Cortical Connectivity," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 11, pp. 1940–1949, Nov. 2017, doi: 10.1109/TNSRE.2017.2701002.

[230] J. Fan, J. W. Wade, A. P. Key, Z. E. Warren, and N. Sarkar, "EEG-Based Affect and Workload Recognition in a Virtual Driving Environment for ASD Intervention," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 1, pp. 43–51, Jan. 2018, doi: 10.1109/TBME.2017.2693157.

[231] A. Lécuyer, L. George, and M. Marchal, "Toward Adaptive VR Simulators Combining Visual, Haptic, and Brain-Computer Interfaces," *IEEE Computer Graphics and Applications*, vol. 33, no. 5, pp. 18–23, Sep. 2013, doi: 10.1109/MCG.2013.80.

[232] W. H. Song, D. H. Han, and H. J. Shim, "Comparison of Brain Activation in Response to Two Dimensional and Three Dimensional On-Line Games," *Psychiatry Investig*, vol. 10, no. 2, pp. 115–120, Jun. 2013, doi: 10.4306/pi.2013.10.2.115.

[233] M. Palaus, E. M. Marron, R. Viejo-Sobera, and D. Redolar-Ripoll, "Neural Basis of Video Gaming: A Systematic Review," *Front Hum Neurosci*, vol. 11, p. 248, 2017, doi: 10.3389/fnhum.2017.00248.

[234] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004, doi: 10.1016/j.jneumeth.2003.10.009.

[235] R. D. Pascual-Marqui *et al.*, "Assessing interactions in the brain with exact low-resolution electromagnetic tomography," *Philos Trans A Math Phys Eng Sci*, vol. 369, no. 1952, pp. 3768–3784, Oct. 2011, doi: 10.1098/rsta.2011.0081.

[236] J. Mazziotta *et al.*, "A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM).," *Philos Trans R Soc Lond B Biol Sci*, vol. 356, no. 1412, pp. 1293–1322, Aug. 2001, doi: 10.1098/rstb.2001.0915.

[237] N. Tzourio-Mazoyer *et al.*, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *Neuroimage*, vol. 15, no. 1, pp. 273–289, Jan. 2002, doi: 10.1006/nimg.2001.0978.
[238] Y. Attal and D. Schwartz, "Assessment of Subcortical Source Localization Using Deep Brain Activity Imaging Model with Minimum Norm Operators: A MEG Study," *PLOS ONE*, vol. 8, no. 3, p. e59856, Mar. 2013, doi: 10.1371/journal.pone.0059856.
[239] C. J. Stam, G. Nolte, and A. Daffertshofer, "Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources," *Hum Brain Mapp*, vol. 28, no. 11, pp. 1178–1193, Nov. 2007, doi: 10.1002/hbm.20346.
[240] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Phys. Rev. Lett.*, vol. 87, no. 19, p. 198701, Nov. 2001, doi: 10.1103/PhysRevLett.87.198701.
[241] K. Yan and D. Zhang, "Feature selection and analysis on correlated gas sensor data with recursive feature elimination," *Sensors and Actuators B: Chemical*, vol. 212, pp. 353–363, Jun. 2015, doi: 10.1016/j.snb.2015.02.025.
[242] L. Toloşi and T. Lengauer, "Classification with correlated features: unreliability of feature ranking and solutions," *Bioinformatics*, vol. 27, no. 14, pp. 1986–1994, Jul. 2011, doi: 10.1093/bioinformatics/btr300.
[243] H. Kong, X. Li, J.-G. Wang, and C. Kambhamettu, "Ensemble LDA for Face Recognition," in *Advances in Biometrics*, 2005, pp. 166–172.
[244] P. Golland, F. Liang, S. Mukherjee, and D. Panchenko, "Permutation Tests for Classification," in *Learning Theory*, 2005, pp. 501–515.
[245] W.-C. Li, F.-C. Chiu, Y. Kuo, and K.-J. Wu, "The Investigation of Visual Attention and Workload by Experts and Novices in the Cockpit," in *Engineering Psychology and Cognitive Ergonomics. Applications and Services*, 2013, pp. 167–176.
[246] S. Zhang, Y. Zhang, Y. Sun, N. Thakor, and A. Bezerianos, "Graph theoretical analysis of EEG functional network during multi-workload flight simulation experiment in virtual reality environment," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jul. 2017, pp. 3957–3960. doi: 10.1109/EMBC.2017.8037722.
[247] S. M. Slobounov, W. Ray, B. Johnson, E. Slobounov, and K. M. Newell, "Modulation of cortical activity in 2D versus 3D virtual reality environments: an EEG study," *Int J Psychophysiol*, vol. 95, no. 3, pp. 254–260, Mar. 2015, doi: 10.1016/j.ijpsycho.2014.11.003.
[248] W. K. Y. So, S. W. H. Wong, J. N. Mak, and R. H. M. Chan, "An evaluation of mental workload with frontal EEG," *PLOS ONE*, vol. 12, no. 4, p. e0174949, Apr. 2017, doi: 10.1371/journal.pone.0174949.
[249] A. T. Kamzanova, A. M. Kustubayeva, and G. Matthews, "Use of EEG workload indices for diagnostic monitoring of vigilance decrement," *Hum Factors*, vol. 56, no. 6, pp. 1136–1149, Sep. 2014, doi: 10.1177/0018720814526617.
[250] M. V. Kosti, K. Georgiadis, D. A. Adamos, N. Laskaris, D. Spinellis, and L. Angelis, "Towards an affordable brain computer interface for the assessment of programmers' mental workload," *International Journal of Human-Computer Studies*, vol. 115, pp. 52–66, Jul. 2018, doi: 10.1016/j.ijhcs.2018.03.002.

[251] S. Enriquez-Geppert and F. Barceló, "Multisubject Decomposition of Event-related Positivities in Cognitive Control: Tackling Age-related Changes in Reactive Control," *Brain Topogr*, vol. 31, no. 1, pp. 17–34, Jan. 2018, doi: 10.1007/s10548-016-0512-4.
[252] Y. Yamauchi *et al.*, "Relation between parametric change of the workload and prefrontal cortex activity during a modified version of the 'rock, paper, scissors' task," *Neuropsychobiology*, vol. 68, no. 1, pp. 24–33, 2013, doi: 10.1159/000350948.
[253] T. P. K. Breckel, C. Giessing, and C. M. Thiel, "Impact of brain networks involved in vigilance on processing irrelevant visual motion," *NeuroImage*, vol. 55, no. 4, pp. 1754–1762, Apr. 2011, doi: 10.1016/j.neuroimage.2011.01.025.
[254] S. Tanaka, M. Honda, T. Hanakawa, and L. G. Cohen, "Differential contribution of the supplementary motor area to stabilization of a procedural motor skill acquired through different practice schedules," *Cereb. Cortex*, vol. 20, no. 9, pp. 2114–2121, Sep. 2010, doi: 10.1093/cercor/bhp276.
[255] J. Lim, W. Wu, J. Wang, J. A. Detre, D. F. Dinges, and H. Rao, "Imaging brain fatigue from sustained mental workload: An ASL perfusion study of the time-on-task effect," *NeuroImage*, vol. 49, no. 4, pp. 3426–3435, Feb. 2010, doi: 10.1016/j.neuroimage.2009.11.020.
[256] M. Stylianou-Korsnes, M. Reiner, S. J. Magnussen, and M. W. Feldman, "Visual recognition of shapes and textures: an fMRi study," *Brain Struct Funct*, vol. 214, no. 4, pp. 355–359, May 2010, doi: 10.1007/s00429-010-0241-5.
[257] J. Navarro, E. Reynaud, and F. Osiurak, "Neuroergonomics of car driving: A critical meta-analysis of neuroimaging data on the human brain behind the wheel," *Neuroscience & Biobehavioral Reviews*, vol. 95, pp. 464–479, Dec. 2018, doi: 10.1016/j.neubiorev.2018.10.016.
[258] K. Paschke, K. Jordan, T. Wüstenberg, J. Baudewig, and J. Leo Müller, "Mirrored or identical — Is the role of visual perception underestimated in the mental rotation process of 3D-objects?: A combined fMRI-eye tracking-study," *Neuropsychologia*, vol. 50, no. 8, pp. 1844–1851, Jul. 2012, doi: 10.1016/j.neuropsychologia.2012.04.010.
[259] E. Demeter, L. Hernandez-Garcia, M. Sarter, and C. Lustig, "Challenges to attention: A continuous arterial spin labeling (ASL) study of the effects of distraction on sustained attention," *NeuroImage*, vol. 54, no. 2, pp. 1518–1529, Jan. 2011, doi: 10.1016/j.neuroimage.2010.09.026.
[260] M. S. Korsnes, A. A. Wright, and J. D. E. Gabrieli, "An fMRI analysis of object priming and workload in the precuneus complex," *Neuropsychologia*, vol. 46, no. 5, pp. 1454–1462, Apr. 2008, doi: 10.1016/j.neuropsychologia.2007.12.028.
[261] R. Sladky *et al.*, "Neurobiological differences in mental rotation and instrument interpretation in airline pilots," *Scientific Reports*, vol. 6, p. 28104, Jun. 2016, doi: 10.1038/srep28104.
[262] Y. Lean and F. Shan, "Brief review on physiological and biochemical evaluations of human mental workload," *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 22, no. 3, pp. 177–187, 2012, doi: 10.1002/hfm.20269.
[263] J. F. Cavanagh and M. J. Frank, "Frontal theta as a mechanism for cognitive control," *Trends Cogn. Sci. (Regul. Ed.)*, vol. 18, no. 8, pp. 414–421, Aug. 2014, doi: 10.1016/j.tics.2014.04.012.
[264] D. J. White, M. Congedo, J. Ciorciari, and R. B. Silberstein, "Brain oscillatory activity during spatial navigation: theta and gamma activity link medial temporal and parietal

regions," *J Cogn Neurosci*, vol. 24, no. 3, pp. 686–697, Mar. 2012, doi: 10.1162/jocn_a_00098.
[265] Y. Zhang, H. Zheng, Y. Duan, L. Meng, and L. Zhang, "An integrated approach to subjective measuring commercial aviation pilot workload," in *2015 IEEE 10th Conference on Industrial Electronics and Applications (ICIEA)*, Jun. 2015, pp. 1093–1098. doi: 10.1109/ICIEA.2015.7334270.
[266] A. Ogawa, C. Bordier, and E. Macaluso, "Amygdala Activation Is Associated with Sense of Presence during Viewing 3D-surround Cinematography," in *Neural Information Processing*, 2013, pp. 153–160.
[267] A. R. Dores *et al.*, "Amygdala activation in response to 2D and 3D emotion-inducing stimuli," *PsychNology Journal*, vol. 12, no. 1–2, pp. 29–43, 2014.
[268] K. R. Sherrill, U. M. Erdem, R. S. Ross, T. I. Brown, M. E. Hasselmo, and C. E. Stern, "Hippocampus and Retrosplenial Cortex Combine Path Integration Signals for Successful Navigation," *J. Neurosci.*, vol. 33, no. 49, pp. 19304–19313, Dec. 2013, doi: 10.1523/JNEUROSCI.1825-13.2013.
[269] "Health topics." https://www.who.int/health-topics.
[270] *Restorative Neurology of Spinal Cord Injury*. Oxford, New York: Oxford University Press, 2011.
[271] S. M. Hatem *et al.*, "Rehabilitation of Motor Function after Stroke: A Multiple Systematic Review Focused on Techniques to Stimulate Upper Extremity Recovery," *Front Hum Neurosci*, vol. 10, p. 442, 2016, doi: 10.3389/fnhum.2016.00442.
[272] A. Thibaut, C. Chatelle, E. Ziegler, M.-A. Bruno, S. Laureys, and O. Gosseries, "Spasticity after stroke: Physiology, assessment and treatment," *Brain Injury*, vol. 27, no. 10, pp. 1093–1105, Sep. 2013, doi: 10.3109/02699052.2013.804202.
[273] M. C. Emos and S. Agarwal, "Neuroanatomy, Upper Motor Neuron Lesion," in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2019.
[274] K. Dworzynski, G. Ritchie, E. Fenu, K. MacDermott, and E. D. Playford, "Rehabilitation after stroke: summary of NICE guidance," *BMJ*, vol. 346, p. f3615, Jun. 2013, doi: 10.1136/bmj.f3615.
[275] K. Nas, L. Yazmalar, V. Şah, A. Aydın, and K. Öneş, "Rehabilitation of spinal cord injuries," *World J Orthop*, vol. 6, no. 1, pp. 8–16, Jan. 2015, doi: 10.5312/wjo.v6.i1.8.
[276] Stinear Cathy, Ackerley Suzanne, and Byblow Winston, "Rehabilitation is Initiated Early After Stroke, but Most Motor Rehabilitation Trials Are Not," *Stroke*, vol. 44, no. 7, pp. 2039–2045, Jul. 2013, doi: 10.1161/STROKEAHA.113.000968.
[277] V. L. Feigin *et al.*, "Global and regional burden of stroke during 1990-2010: findings from the Global Burden of Disease Study 2010," *Lancet*, vol. 383, no. 9913, pp. 245–254, Jan. 2014.
[278] V. Carbone, M. M. van der Krogt, H. F. J. M. Koopman, and N. Verdonschot, "Sensitivity of subject-specific models to Hill muscle–tendon model parameters in simulations of gait," *Journal of Biomechanics*, vol. 49, no. 9, pp. 1953–1960, Jun. 2016, doi: 10.1016/j.jbiomech.2016.04.008.
[279] K. B. Kortte, L. D. Falk, R. C. Castillo, D. Johnson-Greene, and S. T. Wegener, "The Hopkins Rehabilitation Engagement Rating Scale: Development and Psychometric Properties," *Archives of Physical Medicine and Rehabilitation*, vol. 88, no. 7, pp. 877–884, Jul. 2007, doi: 10.1016/j.apmr.2007.03.030.

[280] G. Andreoni, S. Parini, L. Maggi, L. Piccini, G. Panfili, and A. Torricelli, "Human Machine Interface for Healthcare and Rehabilitation," in *Advanced Computational Intelligence Paradigms in Healthcare-2*, S. Vaidya, L. C. Jain, and H. Yoshida, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 131–150. doi: 10.1007/978-3-540-72375-2_7.

[281] J. Rudisch *et al.*, "Kinematic parameters of hand movement during a disparate bimanual movement task in children with unilateral Cerebral Palsy," *Hum Mov Sci*, vol. 46, pp. 239–250, Apr. 2016, doi: 10.1016/j.humov.2016.01.010.

[282] P. Reis, F. Hebenstreit, F. Gabsteiger, V. von Tscharner, and M. Lochmann, "Methodological aspects of EEG and body dynamics measurements during motion," *Front. Hum. Neurosci.*, vol. 8, 2014, doi: 10.3389/fnhum.2014.00156.

[283] J. Munzert, B. Lorey, and K. Zentgraf, "Cognitive motor processes: The role of motor imagery in the study of motor representations," *Brain Research Reviews*, vol. 60, no. 2, pp. 306–326, May 2009, doi: 10.1016/j.brainresrev.2008.12.024.

[284] Y. H. Yin, Y. J. Fan, and L. D. Xu, "EMG and EPP-Integrated Human–Machine Interface Between the Paralyzed and Rehabilitation Exoskeleton," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 4, pp. 542–549, Jul. 2012, doi: 10.1109/TITB.2011.2178034.

[285] A. C. Villa-Parra, D. Delisle-Rodriguez, J. Souza Lima, A. Frizera-Neto, and T. Bastos, "Knee Impedance Modulation to Control an Active Orthosis Using Insole Sensors," *Sensors (Basel)*, vol. 17, no. 12, Nov. 2017, doi: 10.3390/s17122751.

[286] J. L. Pons, "Rehabilitation Exoskeletal Robotics," *IEEE Engineering in Medicine and Biology Magazine*, vol. 29, no. 3, pp. 57–63, May 2010, doi: 10.1109/MEMB.2010.936548.

[287] D. Bhatia, G. Bansal, R. p. Tewari, and K. k. Shukla, "State of art: Functional Electrical Stimulation (FES)," *International Journal of Biomedical Engineering and Technology*, vol. 5, no. 1, pp. 77–99, Jan. 2011, doi: 10.1504/IJBET.2011.038474.

[288] M. A. M. Dzahir and S. Yamamoto, "Recent Trends in Lower-Limb Robotic Rehabilitation Orthosis: Control Scheme and Strategy for Pneumatic Muscle Actuated Gait Trainers," *Robotics*, vol. 3, pp. 120–148, 2014, doi: 10.3390/robotics3020120.

[289] H. Herr, "Exoskeletons and orthoses: classification, design challenges and future directions," *J Neuroeng Rehabil*, vol. 6, p. 21, Jun. 2009, doi: 10.1186/1743-0003-6-21.

[290] Z. Yue, X. Zhang, and J. Wang, "Hand Rehabilitation Robotics on Poststroke Motor Recovery," *Behav Neurol*, vol. 2017, 2017, doi: 10.1155/2017/3908135.

[291] P. Ghosh, A. Mazumder, A. Banerjee, and D. N. Tibarewala, "Electromyography-Based Functional Electrical Stimulation (FES) in Rehabilitation," *Biomedical Image Analysis and Mining Techniques for Improved Health Outcomes*, pp. 337–355, 2016, doi: 10.4018/978-1-4666-8811-7.ch016.

[292] S. R. Soekadar, N. Birbaumer, M. W. Slutzky, and L. G. Cohen, "Brain-machine interfaces in neurorehabilitation of stroke," *Neurobiol. Dis.*, vol. 83, pp. 172–179, Nov. 2015, doi: 10.1016/j.nbd.2014.11.025.

[293] R. Martin, C. Sadowsky, K. Obst, B. Meyer, and J. McDonald, "Functional electrical stimulation in spinal cord injury:: from theory to practice," *Top Spinal Cord Inj Rehabil*, vol. 18, no. 1, pp. 28–33, 2012, doi: 10.1310/sci1801-28.

[294] K. Takeda, G. Tanino, and H. Miyasaka, "Review of devices used in neuromuscular electrical stimulation for stroke rehabilitation," *Med Devices (Auckl)*, vol. 10, pp. 207–213, 2017, doi: 10.2147/MDER.S123464.

[295] M. Atzori, M. Cognolato, and H. Müller, "Deep Learning with Convolutional Neural Networks Applied to Electromyography Data: A Resource for the Classification of Movements for Prosthetic Hands," *Front. Neurorobot.*, vol. 10, 2016, doi: 10.3389/fnbot.2016.00009.

[296] Y. Zhang *et al.*, "Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces," *Expert Systems with Applications*, vol. 96, pp. 302–310, Apr. 2018, doi: 10.1016/j.eswa.2017.12.015.

[297] M. Y. Latif *et al.*, "Brain computer interface based robotic arm control," in *2017 International Smart Cities Conference (ISC2)*, Sep. 2017, pp. 1–5. doi: 10.1109/ISC2.2017.8090870.

[298] W.-C. Hsu, L.-F. Lin, C.-W. Chou, Y.-T. Hsiao, and Y.-H. Liu, "EEG Classification of Imaginary Lower Limb Stepping Movements Based on Fuzzy Support Vector Machine with Kernel-Induced Membership Function," *Int. J. Fuzzy Syst.*, vol. 19, no. 2, pp. 566–579, Apr. 2017, doi: 10.1007/s40815-016-0259-9.

[299] J.-Y. Jung, W. Heo, H. Yang, and H. Park, "A Neural Network-Based Gait Phase Classification Method Using Sensors Equipped on Lower Limb Exoskeleton Robots," *Sensors*, vol. 15, no. 11, pp. 27738–27759, Nov. 2015, doi: 10.3390/s151127738.

[300] A. Ahangi, M. Karamnejad, N. Mohammadi, R. Ebrahimpour, and N. Bagheri, "Multiple classifier system for EEG signal classification with application to brain–computer interfaces," *Neural Comput & Applic*, vol. 23, no. 5, pp. 1319–1327, Oct. 2013, doi: 10.1007/s00521-012-1074-3.

[301] J. Wolpaw and E. W. Wolpaw, Eds., *Brain-Computer Interfaces: Principles and Practice*, 1 edition. Oxford ; New York: Oxford University Press, 2012.

[302] D. G. Dotov, B. G. Bardy, and S. Dalla Bella, "The role of environmental constraints in walking: Effects of steering and sharp turns on gait dynamics," *Sci Rep*, vol. 6, Jun. 2016, doi: 10.1038/srep28374.

[303] N. Kapadia *et al.*, "A randomized trial of functional electrical stimulation for walking in incomplete spinal cord injury: Effects on walking competency," *J Spinal Cord Med*, vol. 37, no. 5, pp. 511–524, Sep. 2014, doi: 10.1179/2045772314Y.0000000263.

[304] L. Davies, P. Chappell, and T. Melvin, "Modelling the effect of hydration on skin conductivity," *Skin Res Technol*, vol. 23, no. 3, pp. 363–368, Aug. 2017, doi: 10.1111/srt.12344.

[305] C. L. Lynch, G. M. Graham, and M. R. Popovic, "A generic model of real-world non-ideal behaviour of FES-induced muscle contractions: simulation tool," *J Neural Eng*, vol. 8, no. 4, p. 046034, Aug. 2011, doi: 10.1088/1741-2560/8/4/046034.

[306] A. Muro-de-la-Herran, B. Garcia-Zapirain, and A. Mendez-Zorrilla, "Gait Analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications," *Sensors (Basel)*, vol. 14, no. 2, pp. 3362–3394, Feb. 2014, doi: 10.3390/s140203362.

[307] J. Klein, S. J. Spencer, and D. J. Reinkensmeyer, "Breaking It Down Is Better: Haptic Decomposition of Complex Movements Aids in Robot-Assisted Motor Learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 3, pp. 268–275, May 2012, doi: 10.1109/TNSRE.2012.2195202.

[308] S. L. Delp *et al.*, "OpenSim: open-source software to create and analyze dynamic simulations of movement," *IEEE Trans Biomed Eng*, vol. 54, no. 11, pp. 1940–1950, Nov. 2007, doi: 10.1109/TBME.2007.901024.

[309] M. Verleysen and D. François, "The Curse of Dimensionality in Data Mining and Time Series Prediction," in *Computational Intelligence and Bioinspired Systems*, 2005, pp. 758–770.
[310] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Computers & Electrical Engineering*, vol. 40, no. 1, pp. 16–28, Jan. 2014, doi: 10.1016/j.compeleceng.2013.11.024.
[311] G. N. Dimitrakopoulos, I. Kakkos, N. V. Thakor, A. Bezerianos, and null Yu Sun, "A mental fatigue index based on regression using mulitband EEG features with application in simulated driving," *Conf Proc IEEE Eng Med Biol Soc*, vol. 2017, pp. 3220–3223, 2017, doi: 10.1109/EMBC.2017.8037542.
[312] Z. Song and S. Guo, "Design Process of Exoskeleton Rehabilitation Device and Implementation of Bilateral Upper Limb Motor Movement," *J. Med. Biol. Eng*, vol. 32, no. 5, pp. 323–330, 2011, doi: 10.5405/jmbe.987.
[313] X. Zhang, Z. Xiang, Q. Lin, and Q. Zhou, "The design and development of a lower limbs rehabilitation exoskeleton suit," in *2013 ICME International Conference on Complex Medical Engineering*, May 2013, pp. 307–312. doi: 10.1109/ICCME.2013.6548260.
[314] P. Polygerinos, Z. Wang, K. C. Galloway, R. J. Wood, and C. J. Walsh, "Soft robotic glove for combined assistance and at-home rehabilitation," *Robotics and Autonomous Systems*, vol. 73, pp. 135–143, Nov. 2015, doi: 10.1016/j.robot.2014.08.014.
[315] Z. Song *et al.*, "Implementation of Resistance Training Using an Upper-Limb Exoskeleton Rehabilitation Device for Elbow Joint," *J. Med. Biol. Eng*, vol. 34, no. 2, pp. 188–196, 2014, doi: 10.5405/jmbe.1337.
[316] A. Fleury, M. Sugar, and T. Chau, "E-textiles in Clinical Rehabilitation: A Scoping Review," *Electronics*, vol. 4, no. 1, pp. 173–203, Mar. 2015, doi: 10.3390/electronics4010173.
[317] B. W. Lee, C. Lee, J. Kim, and M. Lee, "Optimum conductive fabric sensor sites for evaluating the status of knee joint movements using bio-impedance," *Biomed Eng Online*, vol. 10, p. 48, Jun. 2011, doi: 10.1186/1475-925X-10-48.
[318] L. D. Vito, O. Postolache, and S. Rapuano, "Measurements and sensors for motion tracking in motor rehabilitation," *IEEE Instrumentation Measurement Magazine*, vol. 17, no. 3, pp. 30–38, Jun. 2014, doi: 10.1109/MIM.2014.6825386.
[319] N. Carbonaro, G. D. Mura, F. Lorussi, R. Paradiso, D. D. Rossi, and A. Tognetti, "Exploiting Wearable Goniometer Technology for Motion Sensing Gloves," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 6, pp. 1788–1795, Nov. 2014, doi: 10.1109/JBHI.2014.2324293.
[320] N. Miura, T. Watanabe, S. Sugimoto, K. Seki, and H. Kanai, "Fuzzy FES controller using cycle-to-cycle control for repetitive movement training in motor rehabilitation. Experimental tests with wireless system," *J Med Eng Technol*, vol. 35, no. 6–7, pp. 314–321, Oct. 2011, doi: 10.3109/03091902.2011.591480.
[321] V. Rajanna *et al.*, "KinoHaptics: An Automated, Wearable, Haptic Assisted, Physio-therapeutic System for Post-surgery Rehabilitation and Self-care," *J Med Syst*, vol. 40, no. 3, p. 60, Dec. 2015, doi: 10.1007/s10916-015-0391-3.
[322] O. Postolache, "Physical rehabilitation assessment based on smart training equipment and mobile APPs," in *2015 E-Health and Bioengineering Conference (EHB)*, Nov. 2015, pp. 1–6. doi: 10.1109/EHB.2015.7391530.

[323] B. Kim, L. Kim, Y.-H. Kim, and S. K. Yoo, "Cross-association analysis of EEG and EMG signals according to movement intention state," *Cognitive Systems Research*, vol. 44, pp. 1–9, Aug. 2017, doi: 10.1016/j.cogsys.2017.02.001.
[324] H. Enders and B. M. Nigg, "Measuring human locomotor control using EMG and EEG: Current knowledge, limitations and future considerations," *Eur J Sport Sci*, vol. 16, no. 4, pp. 416–426, 2016, doi: 10.1080/17461391.2015.1068869.
[325] J. Chai *et al.*, "Identification of gait-related brain activity using electroencephalographic signals," in *2017 8th International IEEE/EMBS Conference on Neural Engineering (NER)*, May 2017, pp. 548–551. doi: 10.1109/NER.2017.8008410.
[326] A. Shakeel, M. S. Navid, M. N. Anwar, S. Mazhar, M. Jochumsen, and I. K. Niazi, "A Review of Techniques for Detection of Movement Intention Using Movement-Related Cortical Potentials," *Computational and Mathematical Methods in Medicine*, 2015.
[327] P. Ahmadian, S. Cagnoni, and L. Ascari, "How capable is non-invasive EEG data of predicting the next movement? A mini review," *Front Hum Neurosci*, vol. 7, p. 124, 2013, doi: 10.3389/fnhum.2013.00124.
[328] Y. X. Kato, T. Yonemura, K. Samejima, T. Maeda, and H. Ando, "Development of a BCI master switch based on single-trial detection of contingent negative variation related potentials," *Conf Proc IEEE Eng Med Biol Soc*, vol. 2011, pp. 4629–4632, 2011, doi: 10.1109/IEMBS.2011.6091146.
[329] M. Schultze-Kraft *et al.*, "The point of no return in vetoing self-initiated movements," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 113, no. 4, pp. 1080–1085, Jan. 2016, doi: 10.1073/pnas.1513569112.
[330] A. Remsik *et al.*, "A review of the progression and future implications of brain-computer interface therapies for restoration of distal upper extremity motor function after stroke," *Expert Rev Med Devices*, vol. 13, no. 5, pp. 445–454, May 2016, doi: 10.1080/17434440.2016.1174572.
[331] K. K. Ang *et al.*, "A Randomized Controlled Trial of EEG-Based Motor Imagery Brain-Computer Interface Robotic Rehabilitation for Stroke," *Clin EEG Neurosci*, vol. 46, no. 4, pp. 310–320, Oct. 2015, doi: 10.1177/1550059414522229.
[332] E. López-Larraz *et al.*, "Control of an Ambulatory Exoskeleton with a Brain-Machine Interface for Spinal Cord Injury Gait Rehabilitation," *Front Neurosci*, vol. 10, p. 359, 2016, doi: 10.3389/fnins.2016.00359.
[333] N. A. Bhagat *et al.*, "Design and Optimization of an EEG-Based Brain Machine Interface (BMI) to an Upper-Limb Exoskeleton for Stroke Survivors," *Front Neurosci*, vol. 10, p. 122, 2016, doi: 10.3389/fnins.2016.00122.
[334] F. A. Jure, L. C. Carrere, G. G. Gentiletti, and C. B. Tabernig, "BCI-FES system for neuro-rehabilitation of stroke patients," *J. Phys.: Conf. Ser.*, vol. 705, no. 1, p. 012058, 2016, doi: 10.1088/1742-6596/705/1/012058.
[335] A. H. Do, P. T. Wang, C. E. King, A. Abiri, and Z. Nenadic, "Brain-Computer Interface Controlled Functional Electrical Stimulation System for Ankle Movement," *Journal of NeuroEngineering and Rehabilitation*, vol. 8, p. 49, Aug. 2011, doi: 10.1186/1743-0003-8-49.
[336] T.-W. Kim and B.-H. Lee, "Clinical usefulness of brain-computer interface-controlled functional electrical stimulation for improving brain activity in children with spastic cerebral palsy: a pilot randomized controlled trial," *J Phys Ther Sci*, vol. 28, no. 9, pp. 2491–2494, Sep. 2016, doi: 10.1589/jpts.28.2491.

[337] E. Chung, J.-H. Kim, D.-S. Park, and B.-H. Lee, "Effects of brain-computer interface-based functional electrical stimulation on brain activation in stroke patients: a pilot randomized controlled trial," *J Phys Ther Sci*, vol. 27, no. 3, pp. 559–562, Mar. 2015, doi: 10.1589/jpts.27.559.

[338] A. Biasiucci *et al.*, "Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke," *Nature Communications*, vol. 9, no. 1, p. 2421, Jun. 2018, doi: 10.1038/s41467-018-04673-z.

[339] M. Aach *et al.*, "Voluntary driven exoskeleton as a new tool for rehabilitation in chronic spinal cord injury: a pilot study," *Spine J*, vol. 14, no. 12, pp. 2847–2853, Dec. 2014, doi: 10.1016/j.spinee.2014.03.042.

[340] C. Ockenfeld, R. K. Y. Tong, E. A. Susanto, S.-K. Ho, and X. Hu, "Fine finger motor skill training with exoskeleton robotic hand in chronic stroke: stroke rehabilitation," *IEEE Int Conf Rehabil Robot*, vol. 2013, p. 6650392, Jun. 2013, doi: 10.1109/ICORR.2013.6650392.

[341] M. Chen *et al.*, "A self-adaptive foot-drop corrector using functional electrical stimulation (FES) modulated by tibialis anterior electromyography (EMG) dataset," *Med Eng Phys*, vol. 35, no. 2, pp. 195–204, Feb. 2013, doi: 10.1016/j.medengphy.2012.04.016.

[342] K. Shindo *et al.*, "Effectiveness of hybrid assistive neuromuscular dynamic stimulation therapy in patients with subacute stroke: a randomized controlled pilot trial," *Neurorehabil Neural Repair*, vol. 25, no. 9, pp. 830–837, Dec. 2011, doi: 10.1177/1545968311408917.

[343] J. S. Knutson, M. Y. Harley, T. Z. Hisel, S. D. Hogan, M. M. Maloney, and J. Chae, "Contralaterally controlled functional electrical stimulation for upper extremity hemiplegia: an early-phase randomized clinical trial in subacute stroke patients," *Neurorehabil Neural Repair*, vol. 26, no. 3, pp. 239–246, Apr. 2012, doi: 10.1177/1545968311419301.

[344] C. Loconsole *et al.*, "An emg-based robotic hand exoskeleton for bilateral training of grasp," in *2013 World Haptics Conference (WHC)*, Apr. 2013, pp. 537–542. doi: 10.1109/WHC.2013.6548465.

[345] L. Popović Maneski, I. Topalović, N. Jovičić, S. Dedijer, L. Konstantinović, and D. B. Popović, "Stimulation map for control of functional grasp based on multi-channel EMG recordings," *Medical Engineering & Physics*, vol. 38, no. 11, pp. 1251–1259, Nov. 2016, doi: 10.1016/j.medengphy.2016.06.004.

[346] L. Liu, X. Chen, Z. Lu, S. Cao, D. Wu, and X. Zhang, "Development of an EMG-ACC-Based Upper Limb Rehabilitation Training System," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 3, pp. 244–253, Mar. 2017, doi: 10.1109/TNSRE.2016.2560906.

[347] A. L. Rincon, H. Yamasaki, and S. Shimoda, "Design of a video game for rehabilitation using motion capture, EMG analysis and virtual reality," in *2016 International Conference on Electronics, Communications and Computers (CONIELECOMP)*, Feb. 2016, pp. 198–204. doi: 10.1109/CONIELECOMP.2016.7438575.

[348] A. Sarasola-Sanz *et al.*, "A hybrid brain-machine interface based on EEG and EMG activity for the motor rehabilitation of stroke patients," *IEEE Int Conf Rehabil Robot*, vol. 2017, pp. 895–900, 2017, doi: 10.1109/ICORR.2017.8009362.

[349] S. M. Rayegani *et al.*, "Effect of neurofeedback and electromyographic-biofeedback therapy on improving hand function in stroke patients," *Top Stroke Rehabil*, vol. 21, no. 2, pp. 137–151, Apr. 2014, doi: 10.1310/tsr2102-137.

[350] G. Grimaldi, M. Manto, and Y. Jdaoudi, "Quality parameters for a multimodal EEG/EMG/kinematic brain-computer interface (BCI) aiming to suppress neurological tremor in upper limbs," *F1000Res*, vol. 2, p. 282, 2013, doi: 10.12688/f1000research.2-282.v2.

[351] R. N. Khushaba, M. Takruri, J. V. Miro, and S. Kodagoda, "Towards limb position invariant myoelectric pattern recognition using time-dependent spectral features," *Neural Networks*, vol. 55, pp. 42–58, Jul. 2014, doi: 10.1016/j.neunet.2014.03.010.

[352] I. Kyranou, S. Vijayakumar, and M. S. Erden, "Causes of Performance Degradation in Non-invasive Electromyographic Pattern Recognition in Upper Limb Prostheses," *Front. Neurorobot.*, vol. 12, 2018, doi: 10.3389/fnbot.2018.00058.

[353] M. A. Dimyan and L. G. Cohen, "Neuroplasticity in the context of motor rehabilitation after stroke," *Nat Rev Neurol*, vol. 7, no. 2, pp. 76–85, Feb. 2011, doi: 10.1038/nrneurol.2010.200.

[354] M. F. Levin, J. A. Kleim, and S. L. Wolf, "What do motor 'recovery' and 'compensation' mean in patients following stroke?," *Neurorehabil Neural Repair*, vol. 23, no. 4, pp. 313–319, May 2009, doi: 10.1177/1545968308328727.

[355] "EEG Motor Movement/Imagery Dataset." https://www.physionet.org/pn4/eegmmidb/