



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Ταξινόμηση Σημάτων Ηλεκτροεγκεφαλογραφήματος Νοητικής Κίνησης με Χρήση Τεχνικών Βαθιάς Μάθησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΜΟΥΡΑΤΙΔΗ ΣΤΑΥΡΟΥ

Επιβλέπων : Ανδρέας – Γεώργιος Σταφυλοπάτης

Καθηγητής Ε.Μ.Π

Αθήνα, Σεπτέμβριος 2020



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Ταξινόμηση Σημάτων Ηλεκτροεγκεφαλογραφήματος Νοητικής Κίνησης με Χρήση Τεχνικών Βαθιάς Μάθησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΜΟΥΡΑΤΙΔΗ ΣΤΑΥΡΟΥ

Επιβλέπων : Ανδρέας - Γεώργιος Σταφυλοπάτης

Καθηγητής Ε.Μ.Π

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 29^η Σεπτεμβρίου 2020

.....
Ανδρέας – Γεώργιος Σταφυλοπάτης
Καθηγητής Ε.Μ.Π

.....
Γεώργιος Στάμου
Καθηγητής Ε.Μ.Π

.....
Στέφανος Κόλλιας
Καθηγητής Ε.Μ.Π

Αθήνα, Σεπτέμβριος 2020

.....
Μουρατίδης Σταύρος

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Σταύρος Μουρατίδης, 2020.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Οι διεπαφές εγκεφάλου – υπολογιστή (Brain Computer Interfaces - BCIs) που βασίζονται σε ηλεκτροεγκεφαλογραφήματα (EEG) νοητικών κινήσεων, αποτελούν μη επεμβατικά συστήματα τα οποία μετεφράζουν την πρόθεση κίνησης μελών του ατόμου σε σήματα ελέγχου μια εξωτερικής συσκευής. Αποτελούν έναν επιστημονικό τομέα υπό εντατική έρευνα καθώς μπορούν να αξιοποιηθούν από άτομα με σοβαρές αναπηρίες, παρέχοντας τους την δυνατότητα επικοινωνίας και αλληλεπίδρασης με το περιβάλλον. Για να γίνει η μετάφραση της πρόθεσης κίνησης στο αντίστοιχο σήμα ελέγχου είναι απαραίτητη η αποκωδικοποίηση και ταξινόμηση των EEG σημάτων. Για την ταξινόμηση, τυπικά γίνεται χρήση κλασικών τεχνικών μηχανικής μάθησης, οι οποίες αν και πετυχαίνουν αρκετά καλά αποτελέσματα, βασίζονται στην χειροκίνητη εξαγωγή χαρακτηριστικών των σημάτων. Ωστόσο η εξαγωγή των χαρακτηριστικών αυτών αποτελεί μεγάλη πρόκληση καθώς τα σήματα αυτά χαρακτηρίζονται από μεγάλη μεταβλητότητα. Παράλληλα ο τομέας της βαθιάς μάθησης εξελίσσεται ραγδαία τα τελευταία χρόνια, παρέχοντας μεθόδους μάθησης που εκπαιδεύονται να εξαγάγουν αυτόματα τα πιο χρήσιμα χαρακτηριστικά των δεδομένων τους, οδηγώντας σε εντυπωσιακά αποτελέσματα. Αντικείμενο της εργασίας είναι η μελέτη της επίδοσης τεχνικών βαθιάς μάθησης, οι οποίες θα μπορούσαν να αντικαταστήσουν τις κλασικές μεθόδους και να οδηγήσουν στην ανάπτυξη των BCIs. Υλοποιήθηκαν τέσσερα νευρωνικά δίκτυα: δυο συνελκτικά (CNN), ένα συνελκτικό βασισμένο σε σπεκτρογράμματα καθώς και ένα συνελκτικό δίκτυο μακράς βραχείας μνήμης (ConvLSTM). Τα μοντέλα μας αξιολογήθηκαν σε δεδομένα από νοητικές κινήσεις 109 ατόμων, τα οποία είναι διαθέσιμα στο Physionet. Παρατήρηθηκαν καλύτερες επιδόσεις από γνωστές μεθόδους ρηχής μάθησης καθώς επίσης και από μοντέλα βαθιάς μάθησης συγγενών εργασιών στο ίδιο dataset.

Λέξεις Κλειδιά: Ηλεκτροεγκεφαλογράφημα, Διεπαφή Εγκεφάλου – Υπολογιστή, Νοητική Κίνηση, Αναγνώριση πρόθεσης κίνησης, Μηχανική Μάθηση, Εξαγωγή Χαρακτηριστικών, Βαθιά Μάθηση.

Abstract

Motor imagery electroencephalography (EEG) based Brain – Computer Interfaces (BCIs) are non - invasive systems that translate a subject’s movement intention into a control signal for an external device. BCIs are under intense research over the last two decades, as they can provide an alternative path of communication and environmental interaction for disabled people, such as patients suffering from locked-in syndrome. This translation requires the decoding and classification of the acquired EEG signal. Typically, traditional machine learning techniques are used for EEG classification. These techniques rely on hand-crafted feature extraction, which is a very difficult task because of the high-stationarity of EEG signals. Meanwhile, deep learning is a rapidly growing field with many successful applications and provides models that are able to automatically extract the desired features of raw data leading to end – to – end learning with impressive results. The subject of this diploma thesis is studying of deep learning methods in classifying EEG signals, that could possibly replace the traditional methods and benefit greatly motor imagery BCIs. Four models were developed: two convolutional (CNN) models, one spectrogram-based CNN and one convolutional Long Term Short Term Memory network (ConvLSTM). The above models were evaluated on motor imagery data from 109 subjects that are publicly available in Physionet Motor Imagery EEG dataset. Our deep learning models achieved better classification results on this dataset than state-of-the art machine learning techniques as well as compared to deep learning models from past research papers.

Key Words : Electroencephalography (EEG), Brain – Computer Interface (BCI), Motor Imagery, Movement intention decoding, Machine Learning, Feature Extraction, Deep Learning.

Ευχαριστίες

Αρχικά, θα ήθελα να ευχαριστήσω τον κ. Ανδρέα – Γεώργιο Σταφυλοπάτη, Καθηγητή Ε.Μ.Π, για την ευκαιρία που μου έδωσε να εκπονήσω την διπλωματική μου εργασία και να ασχοληθώ με ένα ιδιαίτερα ενδιαφέρον θέμα για εμένα. Ακόμα θα ήθελα να ευχαριστήσω ιδιαίτερα τον κ. Γεώργιο Σιόλα, Ερευνητή στο Εργαστήριο Ευφυών συστημάτων για την καίρια βοήθεια του κατά την εκπόνηση της εργασίας και την άψογη συνεργασία. Τέλος θα ήθελα να εκφράσω την ευγνωμοσύνη μου στην οικογένεια μου, την Ειρήνη και τους φίλους μου που ήταν πάντα δίπλα μου και με στήριζαν κατά τη διάρκεια της διπλωματικής αλλά και σε όλα τα χρόνια των σπουδών μου.

Περιεχόμενα

Περίληψη	5
Abstract	7
Ευχαριστίες	9
Κεφάλαιο 1: Εισαγωγή.....	13
1.1) Διεπαφές εγκεφάλου - υπολογιστή.....	13
1.2) Αντικείμενο της εργασίας	13
1.3) Διάρθρωση της εργασίας.....	15
Κεφάλαιο 2: Θεωρητικό Υπόβαθρο	16
2.1) Ηλεκτροεγκεφαλογραφία και διεπαφές εγκεφάλου - υπολογιστή.....	16
Ηλεκτροεγκεφαλογραφία:.....	16
Ηλεκτροεγκεφαλικά Σήματα (EEG signals):	19
Φαντασιακή Κίνηση (Motor Imagery):.....	21
Διεπαφές εγκεφάλου - υπολογιστή.....	21
BCIs με χρήση EEG.....	22
2.2) Προεπεξεργασία.....	24
Αποθορυβοποίηση και φίλτρα	24
Εξαγωγή Χαρακτηριστικών	25
2.3) Μηχανική μάθηση και νευρωνικά δίκτυα.....	28
Μηχανική Μάθηση	28
Τεχνητα Νευρωνικά Δίκτυα.....	29
Πολυστρωματικό Perceptron	31
Συνελκτικά Νευρωνικά Δίκτυα (CNN).....	32
Αναδρομικά Νευρωνικά Δίκτυα (RNN)	36
Αναδρομικό Συνελκτικό Δίκτυο (Recurrent Convolutional Network)	39
Κεφάλαιο 3: Ανασκόπηση Βιβλιογραφίας.....	41
3.1) Αποκωδικοποίηση ΜΙ με χρήση κλασσικών τεχνικών μηχανικής μάθησης.....	41
3.2) Αποκωδικοποίηση ΜΙ με χρήση τεχνικών βαθιάς μάθησης.....	43
Αναδρομικά Νευρωνικά Δίκτυα (RNN)	43
Συνελκτικά Νευρωνικά Δίκτυα (CNN).....	44
Αναδρομικά Συνελκτικά Δίκτυα (RCNN)	45
Κεφάλαιο 4: Πειραματική Διαδικασία.....	46

4.1) Περιγραφή Dataset.....	46
Μεθοδολογία:.....	47
4.2) Υλοποίηση	48
Προεπεξεργασία:.....	48
Περιγραφή των χρησιμοποιούμενων Νευρωνικών Δικτύων.....	49
Κεφάλαιο 5: Αποτελέσματα και Αξιολόγηση.....	56
EEGNet	56
ShallowConvNet	57
SpectrogramConv.....	58
ConvLSTM.....	58
Κεφάλαιο 6: Συμπεράσματα και Μελλοντική Εργασία	61
6.1) Σύνοψη και συμπεράσματα.....	61
6.2) Μελλοντικές κατευθύνσεις	62
Δεδομένα	62
Μελέτη Περισσότερων Κινήσεων και Καταστάσεων.....	62
Online Αξιολόγηση	63
Υλοποίηση τεχνητού μέλους – εξωσκελετού.....	63
Βελτίωση απόδοσης των Νευρωνικών Δικτύων	63
Βιβλιογραφία	65

Κεφάλαιο 1: Εισαγωγή

1.1) Διεπαφές εγκεφάλου - υπολογιστή

Η ανάπτυξη συστημάτων για την διασύνδεση εγκεφάλου – υπολογιστή (Brain Computer Interface – BCI) αποτελεί αντικείμενο εντατικής έρευνας τα τελευταία χρόνια. Τα BCIs έχουν ιδιαίτερα μεγάλη αξία καθώς μπορούν να αποτελέσουν ένα εναλλακτικό κανάλι επικοινωνίας για ανθρώπους που δεν μπορούν να επικοινωνήσουν ούτε μέσω φυσιολογικών οδών όπως η ομιλία, ούτε με την χρήση υπολογιστών. Ωστόσο η συνεισφορά τους δεν τελειώνει εκεί καθώς θα μπορούσαν να αποτελέσουν και συστήματα επαύξησης των δυνατοτήτων υγείων ατόμων, δίνοντάς τους την δυνατότητα για έλεγχο συσκευών με την σκέψη τους είτε για ψυχαγωγικούς είτε για άλλους σκοπούς. Μια κατηγορία των BCIs με πολύ μεγάλο ενδιαφέρον είναι τα BCIs βασισμένα σε φαντασιακές κινήσεις (Motor Imagery - MI), όπου ο έλεγχος του συστήματος γίνεται με την σκέψη της κίνησης ενός μέλους του σώματος. Η ανάπτυξη αυτών των συστημάτων μπορεί να φέρει επανάσταση στην νευροπροσθετική και είναι πολύ ελπιδοφόρα για άτομα με κινητικές αναπηρίες. Για την καταγραφή της εγκεφαλικής δραστηριότητας μπορούν να αξιοποιηθούν πολλές εναλλακτικές μέθοδοι όπως η Μαγνητοεγκεφαλογραφία (MEG), η ηλεκτροκορτικογραφία (ECoG), η λειτουργική απεικόνιση μαγνητικού συντονισμού (fMRI) και η ηλεκτροεγκεφαλογραφία (EEG). Από αυτές τις μεθόδους πιο διαδεδομένη είναι η ηλεκτροεγκεφαλογραφία καθώς είναι μη επεμβατική, φθηνότερη και πιο εύκολα φορητή από τις υπόλοιπες. Ωστόσο η αποκωδικοποίηση των EEG αποτελεί μια πρόκληση καθώς τα σήματα αυτά είναι ασθενή, θορυβώδη, έχουν μεγάλη ευαισθησία σε εξωτερικές συνθήκες και εμφανίζουν μεγάλη μεταβλητότητα στον χρόνο καθώς επίσης και από άνθρωπο σε άνθρωπο. Τυπικά για την ανάλυση τους γίνεται χρήση εξειδικευμένων τεχνικών για την εξαγωγή συχνοτικών και χωρικών χαρακτηριστικών τους, όπως ο αλγόριθμος FBCSP (Filter bank Common Spatial Filters), ενώ για την ταξινόμησή τους γίνεται χρήση κλασσικών μοντέλων ταξινόμησης όπως μηχανές διανύσματος υποστήριξης (Support Vector Machine - SVM) και ταξινομητές γραμμικής διακριτικής ανάλυσης (Linear Discriminant Analysis - LDA). Ωστόσο τα τελευταία χρόνια έχει γνωρίσει τεράστια ανάπτυξη ο κλάδος της βαθιάς μάθησης φέρνοντας state of the art αποτελέσματα σε σχεδόν κάθε τομέα που εφαρμόζεται. Αξίζει λοιπόν να μελετηθεί κατά πόσο αυτή η τεχνολογία έχει την δυνατότητα να επιτύχει καλύτερα αποτελέσματα για την αποκωδικοποίηση των EEG σημάτων σε σχέση με τις έως τώρα τεχνικές και να αποτελέσει καταλύτη στην ανάπτυξη αποτελεσματικών MI BCI.

1.2) Αντικείμενο της εργασίας

Αντικείμενο της παρούσας διπλωματικής είναι η μελέτη ως προς την δυνατότητα της βαθιάς μάθησης να αντικαταστήσει τις έως τώρα χρησιμοποιούμενες τεχνικές για την ανάλυση και κατηγοριοποίηση των EEG σημάτων. Φυσικά στόχος είναι και η ανάπτυξη ενός όσο το δυνατόν πιο αποτελεσματικού μοντέλου. Μελετήθηκαν ηλεκτροεγκεφαλικά σήματα τα οποία αφορούν φαντασιακές κινήσεις χεριών και ποδιών από 109 άτομα. Πιο συγκεκριμένα οι υπό εξέταση MI κινήσεις είναι κλείσιμο δεξιού και αριστερού χεριού, καθώς επίσης και των ποδιών. Το συγκεκριμένο dataset αποτελεί την μεγαλύτερη συλλογή EEG MI δεδομένων, ειδικά όσον αφορά το πλήθος των εξεταζόμενων υποκειμένων. Δίνει έτσι την δυνατότητα αξιολόγησης των υλοποιούμενων μοντέλων για προβλέψεις τόσο εντός του κάθε υποκειμένου ξεχωριστά όσο και καθολικά ανάμεσα στα υποκείμενα. Η δεύτερη περίπτωση είναι μεγάλης σημασίας και δεν έχει μελετηθεί εκτεταμένα καθώς δεν υπάρχουν διαθέσιμα πολλά σύνολα δεδομένων που να το επιτρέπουν. Για την ταξινόμηση των κινήσεων υλοποιήθηκαν 4 διαφορετικές αρχιτεκτονικές βαθιάς μάθησης:

- EEGNet: Μικρό συνελκτικό δίκτυο για την αποκωδικοποίηση EEG σημάτων διαφόρων προβλημάτων βασισμένο στο EEGNet δίκτυο που ανέπτυξαν οι (Lawhern et al., 2018).
- ShallowConvNet: Αποτελεί επίσης ένα ελαφρύ συνελκτικό δίκτυο με εξαιρετικά αποτελέσματα στην ανάλυση MI EEG. Βασισμένο στο sCNN των (Dose et al., 2018).
- SpectrogramConv: Μετατροπή των EEG σημάτων σε «εικόνες» μέσω του μετασχηματισμού Fourier σύντομου χρόνου (Short Time Fourier Transform - STFT) και ταξινόμηση της εικόνας με χρήση συνελκτικού νευρωνικού δικτύου
- ConvLSTM: Αναδρομικό συνελκτικό δίκτυο μακράς βραχείας μνήμης.

Τα δίκτυα βαθιάς μάθησης μπορούν να λάβουν ως είσοδο ακατέργαστα (raw) δεδομένα και να διακρίνουν μόνα τους όλα τα απαραίτητα χαρακτηριστικά για την κατηγοριοποίησή τους. Η ιδιότητα αυτή έχει μεγάλη αξία στον τομέα των BCIs καθώς η εξαγωγή των βέλτιστων χαρακτηριστικών των EEG σημάτων, η οποία είναι απαραίτητη για την αποτελεσματικότητα των κλασικών τεχνικών μηχανικής μάθησης, αποτελεί μεγάλη πρόκληση. Στα πλαίσια της εργασίας γίνεται σύγκριση της ορθότητας των παραπάνω μοντέλων, τόσο με κλασικές τεχνικές μηχανικής μάθησης με γνώμονα την σύγκριση Βαθιάς – Ρηχής μάθησης, όσο και με μοντέλα Βαθιάς μάθησης για την αξιολόγηση των υλοποιούμενων μοντέλων. Η σύγκριση γίνεται με αποτελέσματα από την διαθέσιμη βιβλιογραφία και όπως θα αναλυθεί στο 5^ο Κεφάλαιο της εργασίας το ShallowConvNet και το ConvLSTM πετυχαίνουν state of the art αποτελέσματα για το συγκεκριμένο dataset. Τέλος εξετάζεται η επίδραση που έχει η μείωση των καναλιών καταγραφής και του χρονικού παραθύρου στις επιδόσεις των μοντέλων λόγω της πρακτικής τους σημασίας για την ανάπτυξη ενός συστήματος BCI.

1.3) Διάρθρωση της εργασίας

Στο 2^ο Κεφάλαιο της παρούσας διπλωματικής εργασίας γίνεται επισκόπηση του ερευνητικού πεδίου των διεπαφών εγκεφάλου υπολογίστη, καθώς και βασικών θεωρητικών εννοιών, σχετικών με την επεξεργασία των σημάτων και με τον κλάδο της μηχανικής μάθησης, που θεωρούνται απαραίτητες για την κατανόηση της εργασίας.

Στο 3^ο Κεφάλαιο γίνεται ανασκόπηση της βιβλιογραφίας σχετικά με ερευνητικές μελέτες που έχουν προηγηθεί στον τομέα των BCIs βασισμένων σε EEG φαντασιακών κινήσεων.

Στο 4^ο Κεφάλαιο περιγράφεται η πειραματική διαδικασία που ακολουθήσαμε. Παρουσιάζεται το πειραματικό σχήμα, η προεπεξεργασία των δεδομένων καθώς και λεπτομέρειες για την υλοποίηση της κάθε αρχιτεκτονικής.

Στο 5^ο Κεφάλαιο παρατίθενται τα αποτελέσματα των προαναφερθέντων μοντέλων, εξετάζοντας την καθολική και κατά άτομο ορθότητα, καθώς επίσης και την επίδραση των καναλιών καταγραφής και του μήκους του χρονικού παραθύρου.

Τέλος στο 6^ο Κεφάλαιο παρουσιάζονται συμπεράσματα από την εργασία και δίνονται μελλοντικές κατευθύνσεις για την περαιτέρω ανάπτυξη των υπό εξέταση συστημάτων.

Κεφάλαιο 2: Θεωρητικό Υπόβαθρο

2.1) Ηλεκτροεγκεφαλογραφία και διεπαφές εγκεφάλου - υπολογιστή

Ηλεκτροεγκεφαλογραφία:

Η ηλεκτροεγκεφαλογραφία (EEG) είναι μια ηλεκτροφυσιολογική μέθοδος καταγραφής της ηλεκτρικής δραστηριότητας του εγκεφάλου, όπως αυτή σχηματίζεται από την επαλληλία των ηλεκτρικών ρευμάτων που παράγονται από τους νευρώνες του. Συνήθως είναι μη επεμβατική, τοποθετώντας ηλεκτρόδια στο ανθρώπινο κρανίο προκειμένου να μετρηθούν διακυμάνσεις της τάσης προερχόμενες από το ιοντικό ρεύμα που διαπερνά τους εγκεφαλικούς νευρώνες. Οι εφαρμογές που αξιοποιούν το EEG εστιάζουν κατά βάση είτε στα πιθανώς σχετικά με κάποιο γεγονός δυναμικά (Event-related Potentials), είτε στο φασματικό περιεχόμενο του σήματος. Στη πρώτη κατηγορία μελετώνται πιθανές διακυμάνσεις στο δυναμικό χρονικά σχετιζόμενες με κάποιο συγκεκριμένο εξωτερικό ερέθισμα, όπως ένα οπτικοακουστικό. Στη δεύτερη κατηγορία γίνεται ανάλυση των νευρικών ταλαντώσεων (ή αλλιώς εγκεφαλικών κυμάτων) οι οποίες παρατηρούνται στα σήματα EEG στο πεδίο της συχνότητας.

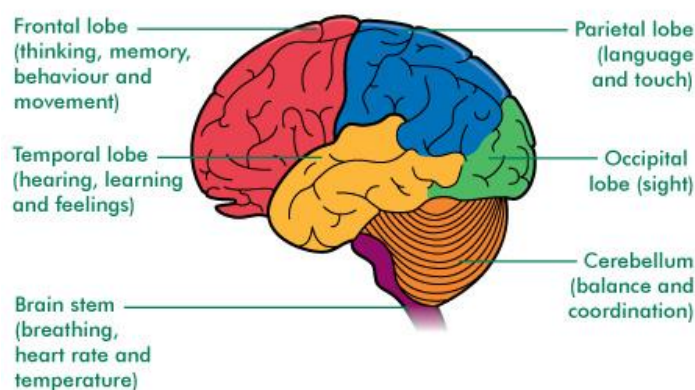
Για την καταγραφή τοποθετούνται ηλεκτρόδια εξωτερικά στο ανθρώπινο κρανίο, συνήθως με ειδικό gel για μεγαλύτερη αγωγιμότητα μεταξύ του ηλεκτροδίου και του δέρματος και τελικά, καλύτερη διάδοση του EEG σήματος. Οι θέσεις των ηλεκτροδίων, καθώς και η ονομασία τους είναι προκαθορισμένη συνήθως από το Διεθνές Σύστημα Τοποθέτησης 10-20 (International 10-20 Placement System). Το σύστημα αυτό παρέχει ένα συγκεκριμένο πρότυπο ώστε να διασφαλίζεται η κοινή ονομασία των ηλεκτροδίων και να καθίσταται ευκολότερη η δυνατότητα αναπαραγωγής των μελετών μεταξύ διαφορετικών ερευνητικών ομάδων. Σύμφωνα με αυτό το σύστημα:

- ο Το όνομα κάθε ηλεκτροδίου αποτελείται από ένα γράμμα εκ των F, T, P, O, C που αναφέρεται στην περιοχή του εγκεφάλου, ή αλλιώς λοβό, που καταγράφει: μετωπικό (Frontal), κροταφικό (Temporal), βρεγματικό (Parietal) και ινιακό (Occipital) λοβό αντίστοιχα. Στα C ηλεκτρόδια το γράμμα αναφέρεται στο central αν και δεν υπάρχει κεντρικός λοβός σαν ξεχωριστή περιοχή του εγκεφάλου. Τα ηλεκτρόδια αυτά συνήθως παρουσιάζουν δραστηριότητα που έχει να κάνει με την μετωπική, την κροταφική και ελαφρώς από την βρεγματική – ινιακή περιοχή. Η κινητική

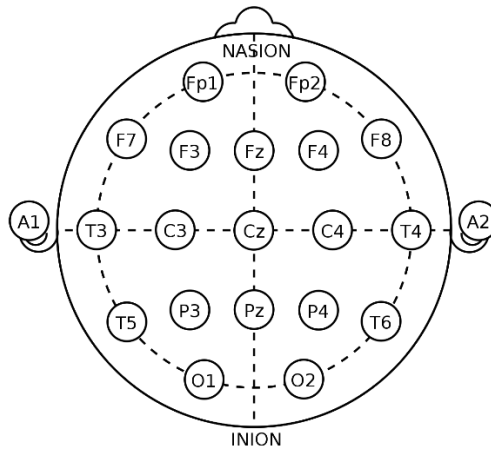
δραστηριότητα που αποτελεί σημείο ενδιαφέροντος στην εργασία εντοπίζεται κυρίως σε αυτή την περιοχή.

- Κάθε ηλεκτρόδιο χαρακτηρίζεται ακόμα από ένα νούμερο. Τα άρτια νούμερα (2,4,6,8) αναφέρονται σε ηλεκτρόδια που έχουν τοποθετηθεί στο δεξί ημισφαίριο του κρανίου, ενώ αντίθετα οι περιττοί αριθμοί (1,3,5,7) αναφέρονται στο αριστερό. Υπάρχει ακόμα μια τρίτη κατηγορία ηλεκτροδίων που αντί για αριθμό έχουν το γράμμα «Z». Τα ηλεκτρόδια αυτά τοποθετούνται στο κέντρο του κεφαλιού, εκεί που ενώνονται τα δυο εγκεφαλικά ημισφαίρια, και χρησιμοποιούνται συνήθως ως ηλεκτρόδια αναφοράς για τις μετρήσεις.
- Τέλος τα ειδικά σύμβολα A, Pg, και Fp αναφέρονται σε ηλεκτρόδια που τοποθετούνται κοντά στους λοβούς των αυτιών, στην ρινοφαρυγγική μετωπιαία περιοχή και στην πολική μετωπιαία περιοχή αντίστοιχα.
- Για την τοποθέτηση των ηλεκτροδίων χρησιμοποιούνται ως σημεία αναφοράς το ριζορρίνιο (nasion), το οποίο είναι το σημείο πάνω από την μύτη και ανάμεσα στα μάτια, και το ίνιο που βρίσκεται στο πίσω μέρος και χαμηλά στο κρανίο. Κατά μήκος της νοητής γραμμής που ενώνει αυτά τα δύο σημεία τοποθετούνται τα Z-ηλεκτρόδια σε αποστάσεις 10%,20%,20%,20%,20%,10% της συνολικής. Στη συνέχεια τοποθετούνται τα υπόλοιπα ηλεκτρόδια κατά πλάτος των Z με διαστήματα και πάλι 10% και 20%. Για αυτό το λόγο το σύστημα αυτό ονομάστηκε 10-20 System.

Παρακάτω βλέπουμε τις περιοχές του ανθρώπινου εγκεφάλου και πως τοποθετούνται τα ηλεκτρόδια με το σύστημα που περιγράφηκε παραπάνω:

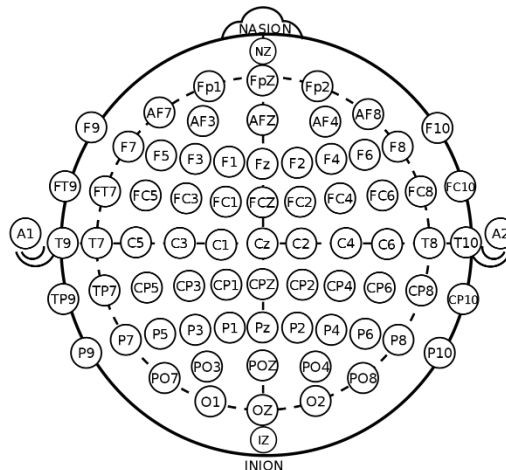


Εικόνα 2.1: Περιοχές του εγκεφάλου και αντίστοιχες λειτουργίες



Εικόνα 2.2: Τοποθέτηση ηλεκτροδίων κατά το διεθνές σύστημα 10-20

Εδώ να σημειωθεί πως με το παραπάνω σύστημα χρησιμοποιούνται συνήθως 21 ηλεκτρόδια. Σε συστήματα μεγαλύτερης ακρίβειας τοποθετούνται περισσότερα ηλεκτρόδια με την τεχνική της διαίρεσης του 10%, όπου τοποθετούνται νέα ηλεκτρόδια στον ενδιάμεσο χώρο μεταξύ ηλεκτροδίων του συστήματος International 10-20. Το νέο σύστημα που προκύπτει ονομάζεται Modified Combinatorial Nomenclature ή αλλιώς διεθνές σύστημα 10 – 10 και παρουσιάζεται στο παρακάτω σχήμα:



Εικόνα 2.3: Τοποθέτηση ηλεκτροδίων κατά το διεθνές σύστημα 10-10

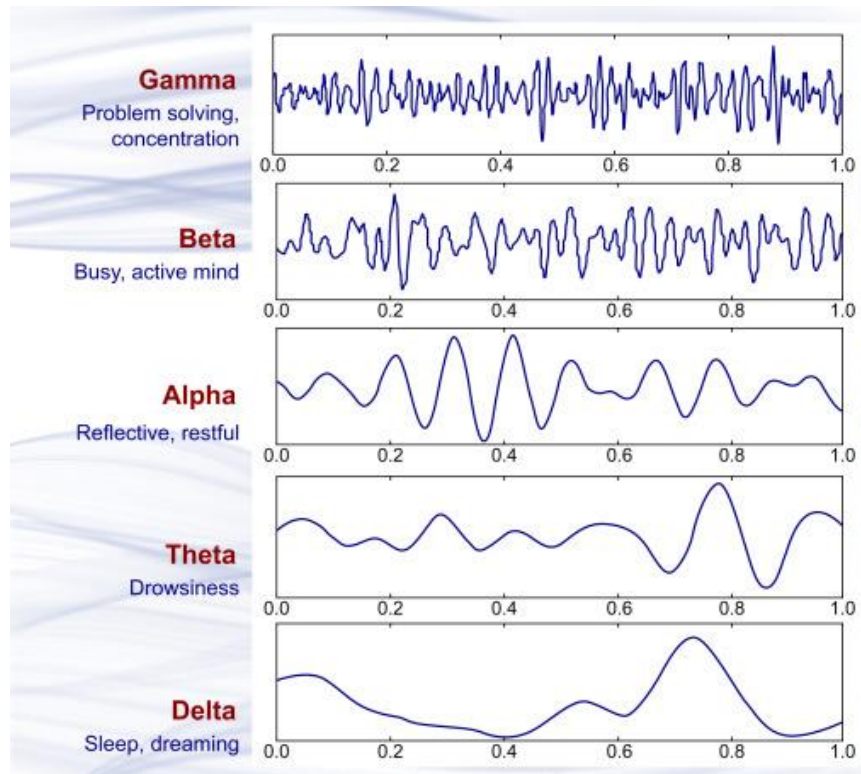
Τα επιπλέον σύμβολα που προστέθηκαν για την ονοματοδοσία των ηλεκτροδίων είναι τα εξής:

- AF: Ανάμεσα σε Fp και F
- FC: Ανάμεσα σε F και C
- FT: Ανάμεσα σε F και T
- CP: Ανάμεσα σε C και P
- TP: Ανάμεσα σε T και P
- PO: Ανάμεσα σε P και O

Ηλεκτροεγκεφαλικά Σήματα (EEG signals):

Τα ηλεκτροεγκεφαλικά σήματα χαρακτηρίζονται από μικρό σηματοθορυβικό λόγο (Signal to noise ratio - SNR), και από ασθενές πλάτος της τάξης του 1-100 μV και παρουσιάζουν ταλαντώσεις σε ένα εύρος συχνοτήτων 0-100Hz. Ωστόσο περιοχές από αυτό το φάσμα έχουν αντιστοιχηθεί, μετά από μελέτες, σε συγκεκριμένες εγκεφαλικές λειτουργίες και σε συγκεκριμένες περιοχές του εγκεφάλου. Αυτές οι ζώνες συχνοτήτων ονομάζονται και εγκεφαλικοί ρυθμοί και είναι οι παρακάτω:

- Δέλτα ($f < 4\text{Hz}$): Τα κύματα δέλτα παρουσιάζουν το μεγαλύτερο πλάτος και την μικρότερη συχνότητα. Σχετίζονται με τον βαθύ αλλά χωρίς όνειρα ύπνο και εμφανίζονται πολύ συχνά σε μωρά και μικρά παιδιά.
- Θήτα ($4\text{Hz} < f < 8\text{Hz}$): Τα κύματα θήτα σχετίζονται με την φαντασία, την περισυλλογή και την χαλάρωση.
- Άλφα ($8\text{Hz} < f < 12\text{Hz}$): Τα κύματα άλφα εντοπίζονται στο οπίσθιο μέρος της κεφαλής και συνδέονται με την πνευματική χαλάρωση αλλά και με την παθητική προσοχή σε κάτι.
- Βήτα ($12\text{Hz} < f < 30\text{Hz}$): Τα κύματα βήτα εντοπίζονται συνήθως στο μπροστινό μέρος και των δυο ημισφαιρίων του εγκεφάλου. Συνδέονται με την κινητική δραστηριότητα και την ενεργητική προσήλωση.
- Γάμμα ($f > 30\text{Hz}$): Τα κύματα γάμμα είναι τα κύματα με την μεγαλύτερη συχνότητα και θεωρείται ότι σχετίζονται με υψηλή συγκέντρωση και δραστηριότητες που απαιτούν υψηλή νοητική και κινητική λειτουργία.

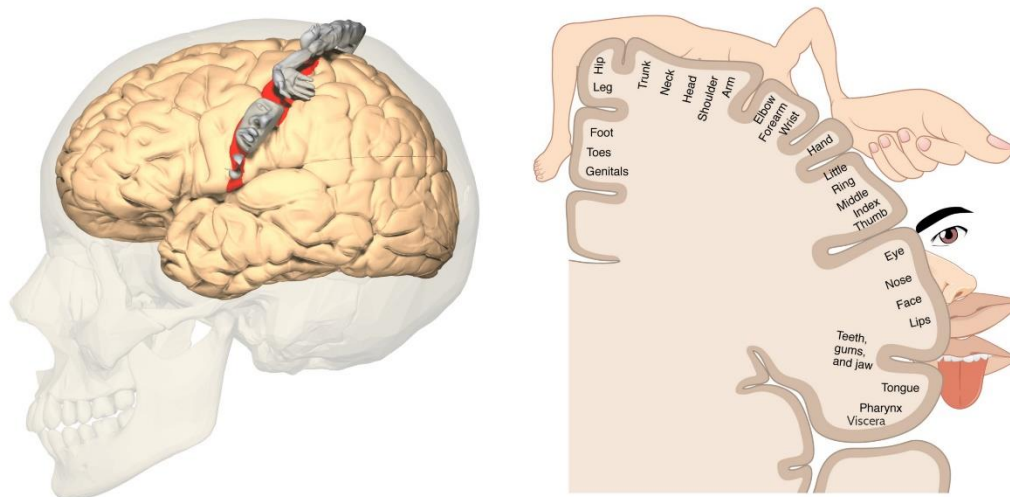


Εικόνα 2.4: Εγκεφαλικοί ρυθμοί

Πρέπει να σημειωθεί ότι τα ακριβή όρια συχνοτήτων μεταξύ των ρυθμών είναι πεδίο διαφωνιών στην επιστημονική κοινότητα. Ακόμα στη βιβλιογραφία τα κύματα της άλφα ζώνης που καταγράφονται στην αισθησιοκινητική περιοχή ονομάζονται κύματα μ (mu or sensorimotor activity). Αλλαγές στα κύματα μ και τα κύματα βήτα χρησιμοποιούνται συνήθως για την αναγνώριση του τύπου της φαντασιακής κίνησης. Η φασματική ισχύς σε μια ζώνη μπορεί να μεταβληθεί ως απόκριση σε κάποιο ερέθισμα. Για παράδειγμα ένα οπτικό ερέθισμα μειώνει την ισχύ του άλφα ρυθμού, ενώ μια κίνηση μπορεί να μειώσει την ισχύ της μ περιοχής (Pfurtscheller, 2001). Όταν η φασματική ισχύς σε μια συγκεκριμένη ζώνη ενισχύεται, ονομάζεται συγχρονισμός σχετιζόμενος με γεγονός (event-related synchronization - ERS) ενώ η μείωση της ονομάζεται αντίστοιχα αποσυγχρονισμός σχετιζόμενος με γεγονός (event-related desynchronization - ERD). Τα ERS/ERD εμφανίζονται σε μια συγκεκριμένη ζώνη ανάλογα με το ερέθισμα (band-specific), ενώ είναι κλειδωμένα ως προς το χρόνο σε σχέση με το ερέθισμα αλλά όχι ως προς την φάση.

Φαντασιακή Κίνηση (Motor Imagery):

Με τον όρο motor imagery εννοούμε την διαδικασία κατά την οποία ένας άνθρωπος φαντάζεται ότι κινεί ένα μέλος του, όπως για παράδειγμα να φαντάζεται ότι ανοιγοκλείνει την παλάμη. Μελέτες έχουν δείξει ότι η φαντασιακή κίνηση προκαλεί παρόμοια εγκεφαλική δραστηριότητα με την προετοιμασία της κανονικής κίνησης. Το motor imagery (MI) έχει εφαρμοστεί σε διάφορους τομείς, όπως στον αθλητισμό ή την μουσική ως μια μορφή νοητικής προετοιμασίας, στην αποκατάσταση νευρολογικών παθήσεων, καθώς και ως ερευνητικό εργαλείο σε μελέτες για το περιεχόμενο νοητικών διεργασιών που προηγούνται της εκτέλεσης μια κίνησης. Όταν ένα άτομο εκτελεί ή φαντάζεται την εκτέλεση μιας κίνησης, προκαλείται ERD στον mu εγκεφαλικό ρυθμό, δηλαδή μειώνεται η φασματική ισχύς της συγκεκριμένης ζώνης. Η δραστηριότητα αυτή γίνεται εμφανής στον κινητικό φλοιό του εγκεφάλου ενώ ανάλογα με το μέλος που κινείται, ενεργοποιείται διαφορετικό μέρος του φλοιού, όπως φαίνεται στην παρακάτω εικόνα:



Εικόνα 2.5: Χάρτης διέγερσης εγκεφάλου από κινητική δραστηριότητα άκρων

Τα ηλεκτρόδια που καταγράφουν καλύτερα την δραστηριότητα που αφορά την κίνηση χεριών και ποδιών είναι τα C3,C4,Cz.

Διεπαφές εγκεφάλου - υπολογιστή

Η διεπαφή εγκεφάλου – υπολογιστή (Brain Computer Interface), ή αλλιώς BMI (brain machine interface), είναι μια διεπαφή που επιτρέπει την απευθείας επικοινωνία μεταξύ του εγκεφάλου και μιας εξωτερικής μηχανής, όπως για παράδειγμα έναν υπολογιστή ή ένα προσθετικό μέλος, χωρίς την συμμετοχή περιφερειακών νεύρων και μυών. Οι εφαρμογές

τους ποικίλουν από τον έλεγχο και την χρήση φυσικών συσκευών με την σκέψη, μέχρι την αλληλεπίδραση με παιχνίδια. Μια από τις πιο σημαντικές εφαρμογές των BCIs είναι η υποστήριξη ανθρώπων με αναπηρία προκειμένου να μπορέσουν να επικοινωνήσουν και να αλληλεπιδράσουν με το περιβάλλον τους. Απαραίτητη για ένα τέτοιο σύστημα είναι η συσκευή καταγραφής της εγκεφαλικής δραστηριότητας που συνήθως αποτελείται από ένα σύνολο ηλεκτροδίων. Ανάλογα με το που τοποθετούνται τα ηλεκτρόδια, τα BCIs κατατάσσονται σε τρεις κατηγορίες:

- Παρεμβατικά: Στα παρεμβατικά BCIs τα ηλεκτρόδια εμφυτεύονται χειρουργικά στο εσωτερικό του εγκεφάλου. Το πλεονέκτημα αυτών των συστημάτων είναι το υψηλότερης ποιότητας εγκεφαλικό σήμα. Ωστόσο πάσχει από σημαντικά μειονεκτήματα. Μετά την επέμβαση υπάρχει η πιθανότητα να δημιουργηθούν ουλικοί ιστοί στον εγκέφαλο που μπορεί να επηρεάσουν την ποιότητα του σήματος, ή ακόμα το σώμα μπορεί να μην δεχθεί το εμφύτευμα και να οδηγηθεί ο ασθενής σε ιατρικές επιπλοκές (Abdulkader et al. 2015).
- Ημι – Παρεμβατικά: Σε αυτή την κατηγορία ηλεκτρόδια τοποθετούνται εσωτερικά του κρανίου αλλά εξωτερικά του εγκεφάλου. Έτσι παρέχουν καλύτερης ποιότητας σήματα από τα μη επεμβατικά BCIs όπου το οστό του κρανίου παρεμβάλλεται και παραμορφώνει το σήμα που λαμβάνουν τα ηλεκτρόδια. Αυτή η μέθοδος καταγραφής ονομάζεται ηλεκτροκορτικογραφία (ECoG) και είναι πολλά υποσχόμενη στον τομέα των brain computer interfaces.
- Μη – Παρεμβατικά : Σε αυτή την κατηγορία ανήκει η ηλεκτροεγκεφαλογραφία που περιγράφηκε παραπάνω, η ηλεκτροοφθαλμογραφία (EOG) η οποία καταγράφει το δυναμικό που δημιουργείται μεταξύ του μπροστινού και οπίσθιου μέρους του οφθαλμού κατά την κίνηση των οφθαλμών και χρησιμοποιεί αυτό το σήμα ως σήμα ελέγχου της μηχανής, καθώς και άλλες τεχνικές όπως η μαγνητοεγκεφαλογραφία (MEG) και η λειτουργική απεικόνιση μαγνητικού συντονισμού (fMRI).

BCIs με χρήση EEG

Η ηλεκτροεγκεφαλογραφία χρησιμοποιείται ευρέως για την καταγραφή των εγκεφαλικών σημάτων σε συστήματα BCI καθώς είναι μη επεμβατική, έχει υψηλή χρονική ανάλυση, το σύστημα καταγραφής είναι εύκολα μεταφέρσιμο και έχει χαμηλό κόστος συγκριτικά με άλλες μη επεμβατικές μεθόδους όπως MEG και fMRI. Ωστόσο η επεξεργασία και χρήση αυτών των σημάτων ενέχει και πολλές προκλήσεις καθώς όπως έχει προαναφερθεί είναι μη στατικά, εμφανίζουν μικρό σηματοθορυβικό δείκτη και επηρεάζονται πολλοί από τέχνηρα(artefacts) όπως παρεμβολές EOG, και δυναμικά από μύες (Stawicki et al., 2017).Ακόμα μπορούν να

επηρεαστούν από την στάση σώματος και την διάθεση του υποκειμένου (Kengic & Subasi, 2017). Για παράδειγμα η όρθια στάση τείνει να βελτιώνει την ποιότητα της καταγραφής και να έχει ισχυρότερο περιεχόμενο υψηλών συχνοτήτων σε σχέση με κάποιον που ξαπλώνει (Caldwell et al., 2003). Οι διεπαφές εγκεφάλου-υπολογιστή βασισμένα στο EEG υπάγονται σε δυο κατηγορίες: τις προκλητές και τις αυθόρμητες ή αλλιώς, εξωγενείς και ενδογενείς. Στα προκλητά BCIs δίνεται ένα εξωτερικό ερέθισμα, οπτικό, ακουστικό ή σωματοαισθητικό. Το ερέθισμα προκαλεί μια εγκεφαλική απόκριση την οποία καταγράφει και αναγνωρίζει το σύστημα ώστε να καταλάβει την πρόθεση του χρήστη. Στα αυθόρμητα BCIs δεν υπάρχει εξωτερικό ερέθισμα και το σύστημα λαμβάνει αποφάσεις ανάλογα με την εγκεφαλική δραστηριότητα που παράγεται από τις νοητικές διεργασίες του υποκειμένου. Στον πίνακα 2.1 παρατίθενται κλασσικές εφαρμογές BCI αυτών των δυο κατηγοριών:

Τύπος	Κλάση	Εφαρμογή	Τρόπος Λειτουργίας	Αριθμός υποκειμένων	Μέση Ακρίβεια	Ρυθμός Μετάδοσης
Προκλητό	VEP	SSVEP/Speller (Stawicki et al., 2017)	Look at one of 30 flickering target stimuli associated with desired character	32	90.81%	35.78 bpm
	ERP	P300/Speller (Speier et al., 2015)	Focus on the desired letter until it next flashes	15	69.28%	20.91bpm
Αυθόρμητο	N/A	Blinks/Virtual keyboard (Chambayil et al., 2010)	Choose from 29 characters using eye blinks to navigate/select	14	N/A	1 char/min
		Motor imagery / Exoskeleton control (Tang et al., 2016)	Control an exoskeleton of the upper limbs using right and left hand MI	4	84.29%	N/A

Πίνακας 2.1: Εφαρμογές BCI

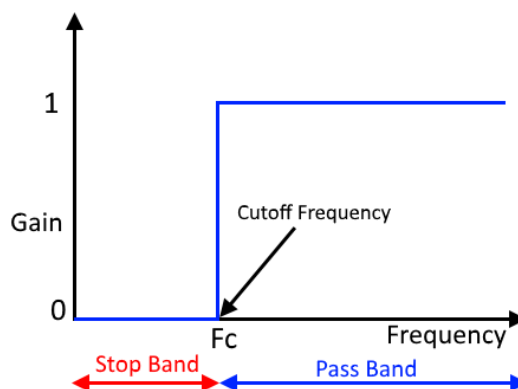
Αν και ένα BCI μπορεί να αξιοποιεί τα σήματα EEG με διάφορους τρόπους, τα motor imagery BCI, στα οποία οι χρήστες φαντάζονται κινήσεις των άκρων τους προκειμένου να ελέγξουν το σύστημα έχουν υποβληθεί σε εκτεταμένη έρευνα. Το μεγάλο ενδιαφέρον οφείλεται στην δυναμική ευρεία εφαρμογή τους σε πεδία όπως η νευροπροσθετική, η ανάρωση ασθενών με κινητικά προβλήματα αλλά και τομείς εκτός του ιατρικού όπως τα ηλεκτρονικά παιχνίδια, όπου η αποκωδικοποίηση της ανθρώπινης σκέψης μπορεί να αποδείχθει εξαιρετικά πολύτιμη (Padfield et al., 2019). Αντικείμενο της παρούσας διπλωματικής είναι η ανάπτυξη ενός MI BCI όπου το υποκείμενο καλείται να φανταστεί την κίνηση κάποιου άκρου του (χεριού ή των

δου ποδιών ποδιών), συλλέγεται το σήμα EEG και στη συνέχεια γίνεται αποκωδικοποίηση της κίνησης με χρήση τεχνικών βαθιάς μάθησης.

2.2) Προεπεξεργασία

Αποθρομβοποίηση και φίλτρα

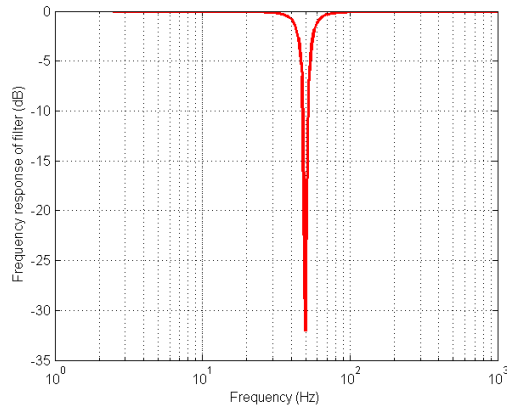
Υψιπερατό Φίλτρο(High-pass filter): Ένα υψιπερατό φίλτρο αφήνει να περάσουν οι συχνότητες του σήματος που είναι μεγαλύτερες από μια συχνότητα αποκοπής (cutoff frequency) και κόβει όσες συχνότητες είναι χαμηλότερες από αυτήν. Ένα απλουστευμένο σχήμα ενός ιδανικού υψιπερατού φίλτρου είναι το παρακάτω:



Εικόνα 2.6: Διάγραμμα υψιπερατού φίλτρου

Στα πλαίσια της εργασίας χρησιμοποιούμε ένα υψιπερατό φίλτρο με πολύ μικρή συχνότητα αποκοπής για να εξαλείψουμε τον θόρυβο μικρής συχνότητας (baseline drift).

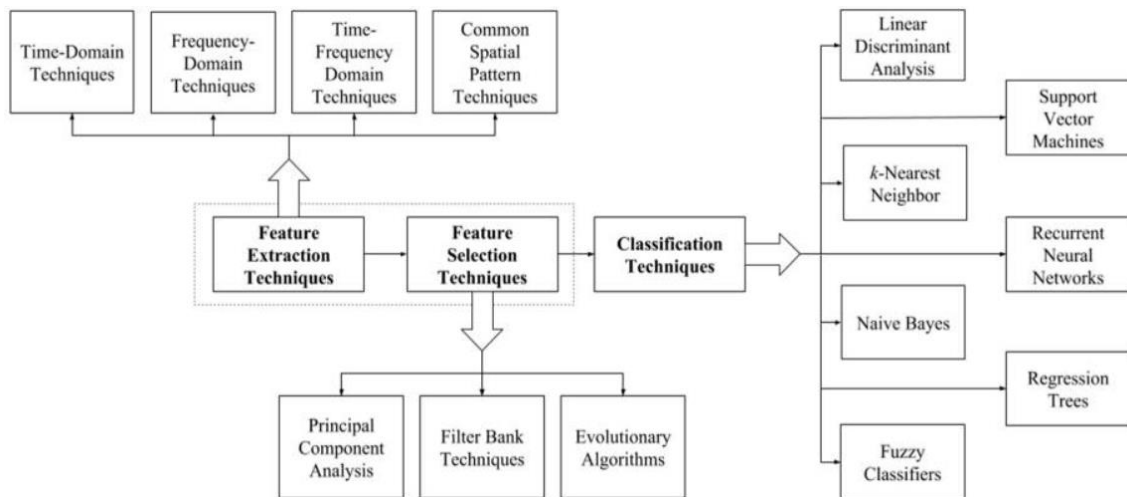
Notch-Filter: Το notch φίλτρο είναι ένα ζωνοαπερατό φίλτρο. Τα ζωνοαπερατά φίλτρα είναι φίλτρα που κόβουν τις συχνότητες που περιέχονται σε μια δοσμένη ζώνη και αφήνουν τις υπόλοιπες να περάσουν. Το χαρακτηριστικό του notch filter είναι ότι έχει πολύ μικρή ζώνη αποκοπής ιδιότητα που το καθιστά πολύ χρήσιμο για την εξάλειψη μιας πηγής θορύβου με συγκεκριμένη συχνότητα όπως για παράδειγμα από την γραμμή τροφοδοσίας (50Hz or 60Hz). Για αυτόν τον συγκεκριμένο λόγο το χρησιμοποιούμε και στην εργασία προκειμένου να εξαλείψουμε τον θόρυβο από την γραμμή τροφοδοσίας στα 50Hz. Παρακάτω φαίνεται η απόκριση αυτού του φίλτρου:



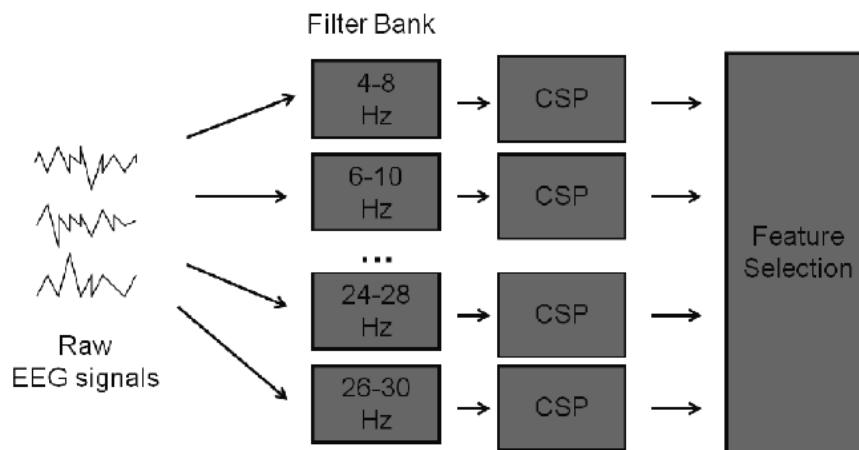
Εικόνα 2.7: Απόκριση φίλτρου Notch

Εξαγωγή Χαρακτηριστικών

Το πιο σημαντικό στάδιο σε ένα τυπικό σύστημα EEG BCI είναι η εξαγωγή των χαρακτηριστικών του σήματος και η επιλογή των καταλληλότερων. Το στάδιο αυτό αποτελεί μια μεγάλη πρόκληση λόγω της υψηλής αστάθειας των EEG σημάτων και του χαμηλού σηματοθορυβικού δείκτη τους. Συχνοτικά χαρακτηριστικά των σημάτων μπορεί να διαφέρουν πολύ από άτομο σε άτομο αλλά και να αλλάζουν εντός του ίδιου με την πάροδο του χρόνου, γεγονός που καθιστά εξαιρετικά δύσκολη την ανάλυσή τους. Οι μέθοδοι που χρησιμοποιούνται εστιάζουν στην εξαγωγή χωρικών και συχνοτικών χαρακτηριστικών. Τεχνικές που μετασχηματίζουν το σήμα από το πεδίο του χρόνου στο πεδίο της συχνότητας όπως ο Discrete Fourier Transform (DFT) και ο Hilbert Transform (HT) ή στο πεδίο χρόνου - συχνότητας, όπως ο Short Time Fourier transform (STFT) και ο Wavelet Transform (WT) χρησιμοποιούνται ευρέως για την ανάλυση του συχνοτικού περιεχομένου του σήματος. Για την χωρική ανάλυση η πιο διαδεδομένη μέθοδος είναι το Common spatial pattern (CSP). Η μέθοδος αυτή αποδίδει βάρη στα κανάλια με στόχο την μεγιστοποίηση της διαφοράς ισχύος των διαφορετικών κλάσεων. Ωστόσο δεν είναι αρκετή για συστήματα MI-BCI μεγάλης ακρίβειας καθώς δεν λαμβάνει υπόψιν τις ζώνες συχνότητων και διαφορετικά υποκείμενα παρουσιάζουν εγκεφαλική δραστηριότητα σε διαφορετικές ζώνες. Για αυτό το λόγο η πλέον δημοφιλής μέθοδος εξαγωγής χαρακτηριστικών από motor imagery EEGs είναι μια βελτίωση του CSP που ονομάζεται Filter Bank Common Spatial Patterns (FBCSP). Κατά τη μέθοδο αυτή το αρχικό σήμα περνάει από διαφορετικά ζωνοπερατά φίλτρα δημιουργώντας ένα σύνολο από διαφορετικές ζώνες του σήματος. Στη συνέχεια σε κάθε ζώνη συχνότητων εφαρμόζεται ο αλγόριθμος CSP. Στα πιο επιτυχημένα BCIs με κλασσικές τεχνικές machine learning γίνεται συνήθως χρήση του αλγορίθμου FBCSP καθώς καταφέρνει να εξάγει ταυτόχρονα συχνοτικά και χωρικά χαρακτηριστικά του σήματος με αρκετά επιτυχημένα αποτελέσματα.



Εικόνα 2.8: Τυπική διαδικασία εξαγωγής χαρακτηριστικών και ταξινόμησης EEG

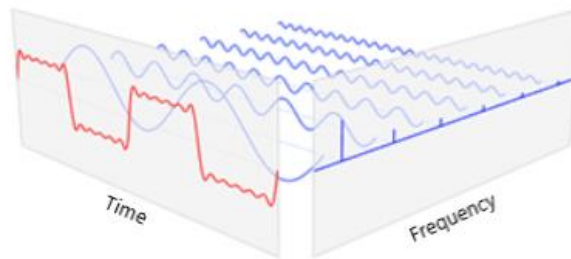


Εικόνα 2.9: Εξαγωγή χαρακτηριστικών με FBCSP

Επειδή λοιπόν όπως γίνεται κατανοητό η εξαγωγή χαρακτηριστικών από EEG αποτελεί μια διαδικασία που παρουσιάζει πολλές δυσκολίες και στην υλοποίησή τους και στην εύρεση των βέλτιστων χαρακτηριστικών, είναι σκόπιμη η ανάπτυξη ενός συστήματος που μπορεί να εξάγει αυτόματα τα πιο χρήσιμα χαρακτηριστικά των σημάτων. Έτσι στα πλαίσια της εργασίας αξιοποιήσαμε την ιδιότητα των βαθιών νευρωνικών δικτύων να εξάγουν μόνα τους τα χαρακτηριστικά των δεδομένων εισόδου και δημιουργήσαμε συστήματα end – to – end που μπορούν να ταξινομήσουν ακατέργαστα δεδομένα.

Διακριτός Μετασχηματισμός Fourier (DFT):

Ο μετασχηματισμός Fourier είναι ένας μαθηματικός μετασχηματισμός που αναλύει μια συνάρτηση (ένα σήμα στην προκειμένη περίπτωση) στις συχνότητες που το απαρτίζουν, μεταφέροντας έτσι το σήμα από το πεδίο του χρόνου στο πεδίο της συχνότητας. Η βασική ιδέα πίσω από τον μετασχηματισμό είναι ότι το κάθε σήμα μπορεί να εκφραστεί ως ένα άθροισμα ημιτόνων με διαφορετική συχνότητα το καθένα.



Εικόνα 2.10: Μετασχηματισμός Fourier ενός αθροίσματος ημιτόνων

Ο διακριτός μετασχηματισμός Fourier εφαρμόζεται στην περίπτωση σημάτων διακριτού χρόνου, όπως τα EEG σήματα που προκύπτουν με δειγματοληψία. Μαθηματικά ορίζεται ως:

$$x[k] = \sum_{n=0}^{N-1} x[n] e^{-\frac{i2\pi kn}{N}}$$

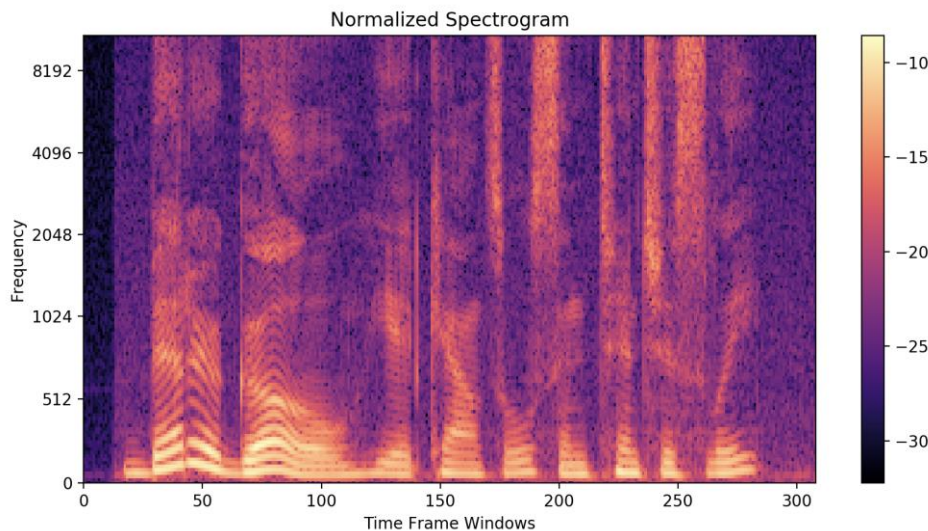
Όπου $\omega_k = \frac{2\pi}{N} k$ με $k = 0, 1, \dots, N-1$.

Στην πράξη επειδή ο παραπάνω υπολογισμός είναι πολύ χρονοβόρος υπολογιστικά ($O(N^2)$), χρησιμοποιούμε τον αλγόριθμο Γρήγορο Μετασχηματισμό Fourier (Fast Fourier Transform - FFT) για τον υπολογισμό του DFT με πολυπλοκότητα $O(N \log N)$.

Short Time Fourier Transform:

Μη στατικά σήματα παρουσιάζουν μεταβολές στο συχνοτικό τους περιεχόμενο με την πάροδο του χρόνου και η μελέτη αυτών των μεταβολών είναι ιδιαίτερα χρήσιμη. Αν σε ένα σήμα EEG εφαρμόζαμε τον DFT θα παίρναμε ένα φάσμα με τον μέσο όρο των συχνοτήτων όλης της διάρκειας του σήματος. Όμως είναι απαραίτητη η μελέτη της μεταβολής του σήματος στον χρόνο. Για αυτό το λόγο χωρίζουμε το σήμα σε μικρότερα κομμάτια (παράθυρα) και εφαρμόζουμε τον DFT σε καθένα από αυτά ξεχωριστά. Η

διαδικασία αυτή ονομάζεται Μετασχηματισμός Fourier Σύντομου Χρόνου (Short Time Fourier Transform - STFT) και μας δίνει το φασματικό περιεχόμενο για το κάθε παράθυρο ξεχωριστά. Το αποτέλεσμα του STFT είναι ένας δισδιάστατος πίνακας που κάθε στοιχείο του χαρακτηρίζεται από μια συχνότητα και μια χρονική στιγμή και η τιμή του δείχνει πόσο ισχυρή είναι αυτή η συχνότητα την συγκεκριμένη στιγμή. Για την απεικόνιση τους χρησιμοποιείται συνήθως ένα heat-map και είναι γνωστό ως σπεκτρογράφημα(spectrogram).



Εικόνα 2.11: Σπεκτρογράφημα ηχητικού σήματος

2.3) Μηχανική μάθηση και νευρωνικά δίκτυα

Μηχανική Μάθηση

Η μηχανική μάθηση ως πεδίο είναι ένας υποκλάδος της Τεχνητής Νοημοσύνης ο οποίος παρέχει σε συστήματα την δυνατότητα να μαθαίνουν και να βελτιώνονται για κάποια εργασία μέσω εμπειρίας. Πρόκειται για προγράμματα τα οποία τους παρέχονται δεδομένα και έχουν την δυνατότητα να εκπαιδευτούν σε αυτά προκειμένου να κάνουν κάποια πρόβλεψη ή να λάβουν κάποια απόφαση χωρίς να έχουν προγραμματιστεί ρητά για αυτήν. Οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται ευρέως σε πλήθος εφαρμογών όπως η όραση υπολογιστών, η αναγνώριση φυσικής γλώσσας κ.α. Η πολυπλοκότητα αυτών των προβλημάτων καθιστά πολύ δύσκολο για έναν προγραμματιστή να αναπτύξει κλασσικούς αλγόριθμους για την επίλυση τους, ωστόσο αυτή καθίσταται δυνατή παρέχοντας στον ίδιο τον υπολογιστή την δυνατότητα να εκπαιδευτεί με βάση τα δεδομένα του προβλήματος. Οι

αλγόριθμοι μηχανικής μάθησης χωρίζονται σε τρεις ευρείες κατηγορίες: την επιβλεπόμενη μάθηση, την μη επιβλεπόμενη μάθηση και την ενισχυτική μάθηση. Πιο αναλυτικά:

- **Επιβλεπόμενη Μάθηση (Supervised Learning)** είναι η διαδικασία όπου ο αλγόριθμος κατασκευάζει μια συνάρτηση που απεικονίζει δεδομένα εισόδου (σύνολο εκπαίδευσης) σε γνωστές επιθυμητές εξόδους, με απώτερο στόχο τη γενίκευση της συνάρτησης αυτής και για εισόδους με άγνωστη έξοδο. Χρησιμοποιείται σε προβλήματα Ταξινόμησης (Classification), Πρόγνωσης (Prediction) και Διερμηνείας (Interpretation)
- **Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)**, όπου ο αλγόριθμος κατασκευάζει ένα μοντέλο για κάποιο σύνολο εισόδων υπό μορφή παρατηρήσεων χωρίς να γνωρίζει τις επιθυμητές εξόδους. Χρησιμοποιείται σε προβλήματα Ανάλυσης Συσχετισμών (Association Analysis) και Ομαδοποίησης (Clustering)
- **Ενισχυτική Μάθηση (Reinforcement Learning)**, όπου ο αλγόριθμος μαθαίνει μια στρατηγική ενεργειών μέσα από άμεση αλληλεπίδραση με το περιβάλλον. Χρησιμοποιείται κυρίως σε προβλήματα Σχεδιασμού (Planning), όπως για παράδειγμα ο έλεγχος κίνησης ρομπότ και η βελτιστοποίηση εργασιών σε εργοστασιακούς χώρους, και σε video games.

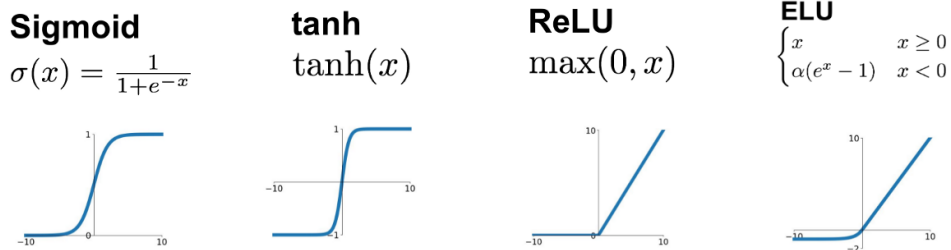
Τεχνητα Νευρωνικά Δίκτυα

Τα τεχνητά νευρωνικά δίκτυα είναι αλγόριθμοι μηχανικής μάθησης που δημιουργούν ένα σύστημα εμπνευσμένο ελαφρώς από την δομή και την λειτουργία των βιολογικών νευρωνικών δικτύων. Ένα τεχνητό νευρωνικό δίκτυο είναι ένα μοντέλο που αποτελείται από πολλούς τεχνητούς νευρώνες (Perceptrons) διασυνδεδεμένους μεταξύ τους. Ο νευρώνας είναι η θεμελιώδης υπολογιστική μονάδα του τεχνητού νευρωνικού δικτύου και μερικές φορές αναφέρεται ως κόμβος. Λαμβάνει πολλές εισόδους τις οποίες αθροίζει και προωθεί το άθροισμα τους ως έξοδο. Συνήθως κάθε είσοδος έχει ένα βάρος και το σταθμισμένο άθροισμα περνάει από μια μη γραμμική συνάρτηση, γνωστή ως συνάρτηση ενεργοποίησης. Οι πιο συχνά χρησιμοποιούμενες είναι η σιγμοειδής, η υπερβολική και η ReLU (Rectified Linear Unit). Οι μαθηματικοί τύποι που περιγράφουν το μοντέλο του τεχνητού νευρώνα είναι οι παρακάτω:

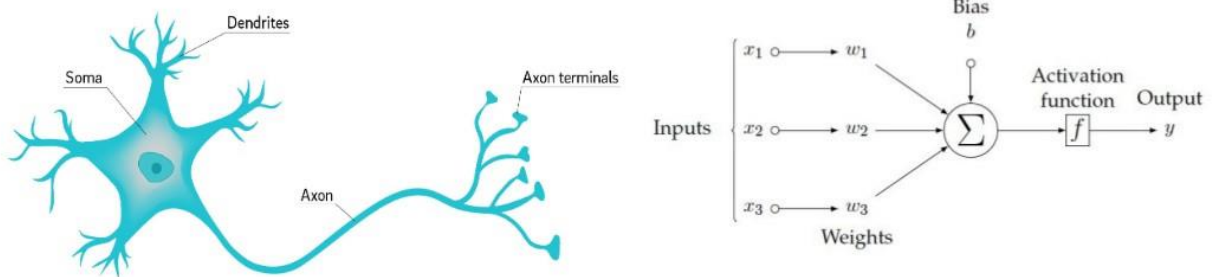
$$y = f(z) \quad (1)$$

$$z = \sum_{i=0}^N w_i x_i + b \quad (2)$$

Όπου y : Έξοδος
 X : Διάνυσμα εισόδου
 W : Διάνυσμα βαρών
 b : Σταθερός όρος (bias)
 f : συνάρτηση ενεργοποίησης



Εικόνα 2.12: Συνήθεις Συναρτήσεις Ενεργοποίησης



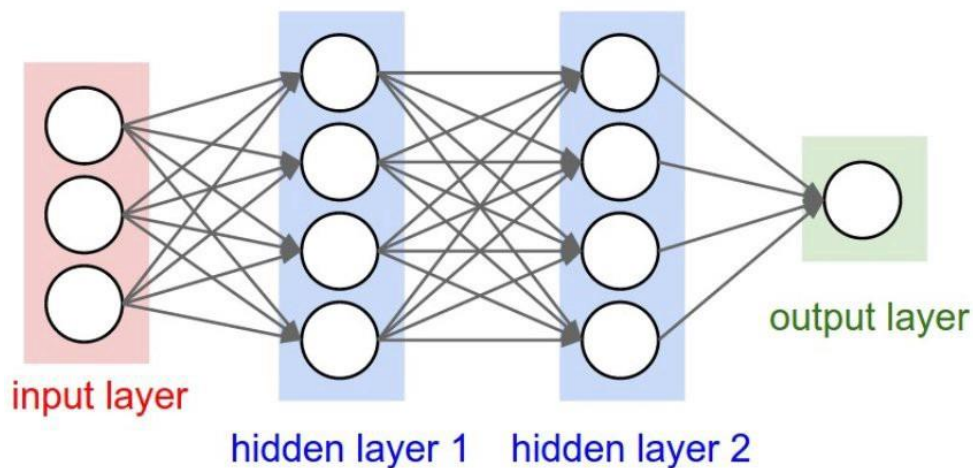
Εικόνα 2.13: Βιολογικός νευρώνας και μοντέλο Perceptron

Το μοντέλο αυτό προσομοιώνει απλοϊκά έναν πραγματικό νευρώνα. Ο βιολογικός νευρώνας αποτελείται από το κυτταρικό σώμα του, τους δενδρίτες από όπου λαμβάνει σήματα από άλλους νευρώνες και τον άξονα που συνδέεται μέσω συνάψεων με άλλους νευρώνες και μεταφέρει σήματα προς αυτούς. Όταν τα δυναμικά εισόδου των δενδριτών ξεπεράσουν αθροιστικά ένα κατώφλι, τότε ενεργοποιείται το δυναμικό δράσης που μεταδίδεται μέσω του άξονα και οδηγείται μέσω συνάψεων σε άλλους νευρώνες. Υπάρχει λοιπόν μια σαφής ομοιότητα ανάμεσα στον βιολογικό νευρώνα αντιστοιχίζοντας τα δυναμικά εισόδου των δενδριτών με τις εισόδους του μαθηματικού μοντέλου, το σώμα που επεργάζεται τις εισόδους και αποφασίζει αν θα πυροδοτήσει κάποια έξοδο με τον κόμβο και τον άξονα με την έξοδο του perceptron. Ο perceptron μπορεί να θεωρηθεί το πιο απλό μονοστρωματικό νευρωνικό δίκτυο. Η βασικότερη αδυναμία του perceptron είναι ότι μπορεί να διαχωρίσει μόνο γραμμικά

διαχωρίσιμα δεδομένα και μόνο δυο κλάσεων. Τα προβλήματα αυτά έρχεται να λύσει το πολυστρωματικό δίκτυο perceptron (MLP).

Πολυστρωματικό Perceptron

Ένα πολυστρωματικό δίκτυο αποτελείται από πολλά στρώματα νευρώνων όπου κάθε νευρώνας του i -οστού στρώματος δέχεται ως εισόδους όλες τις εξόδους από τους νευρώνες του προηγούμενου στρώματος. Το πρώτο και το τελευταίο στρώμα ονομάζονται στρώμα εισόδου και εξόδου αντίστοιχα. Τα ενδιάμεσα στρώματα ονομάζονται κρυφά στρώματα. Ο αριθμός των στρωμάτων του δικτύου ονομάζεται βάθος. Αν έχει δυο ή παραπάνω κρυφά στρώματα θεωρείται βαθύ νευρωνικό δίκτυο (deep neural network) και η μάθηση που επιτελεί είναι η λεγόμενη βαθιά μάθηση (deep learning). Αντίστοιχα ο αριθμός των νευρώνων κάθε στρώματος αποτελεί το πλάτος του δικτύου. Η διασύνδεση των νευρώνων σε συνδυασμό με τις μη γραμμικές συναρτήσεις ενεργοποίησης δίνει στο δίκτυο την δυνατότητα να διαχωρίσει μη γραμμικώς διαχωριζόμενα δεδομένα. Επίσης, σε αντίθεση με τον perceptron μπορεί να ταξινομήσει δεδομένα πολλών κλάσεων έχοντας στο στρώμα εξόδου ισάριθμο με τις κλάσεις αριθμό νευρώνων. Γενικά μεγαλύτερα δίκτυα τόσο σε πλάτος όσο και σε βάθος έχει παρατηρηθεί ότι μπορούν να μάθουν και να διαχωρίσουν πιο περίπλοκα δεδομένα ωστόσο αν υπάρχει μικρός σχετικά όγκος δεδομένων εκπαίδευσης, είναι πιο εύκολο να υπερεκπαιδευτούν σε αυτά και να μην μπορούν να γενικεύσουν σε άγνωστα δεδομένα.



Εικόνα 2.14: Αναπαράσταση ενός πολυεπίπεδου δικτύου Perceptron

Εκπαίδευση Νευρωνικού Δικτύου:

Η αξία των τεχνητών νευρωνικών δικτύων όπως και σε όλων των αλγορίθμων μηχανικής μάθησης έγκειται στην ικανότητά τους να εκπαιδεύονται από τα δεδομένα τους. Με τον όρο εκπαίδευση αναφερόμαστε στην διαδικασία με την οποία το νευρωνικό δίκτυο

αναπροσαρμόζεται ώστε να εκτελέσει πιο επιτυχημένα την εργασία που του έχει δοθεί, με βάση τα δεδομένα εκπαίδευσής του. Η διαδικασία αυτή διεκπεραιώνεται με την αναπροσαρμογή των βαρών και των πολώσεων του δικτύου ώστε να βελτιωθεί η ακρίβειά του, ελαχιστοποιώντας την συνάρτηση σφάλματος. Το σφάλμα είναι η απόσταση (όπως αυτή ορίζεται από την συνάρτηση σφάλματος) της εξόδου του νευρωνικού δικτύου από την αντίστοιχη πραγματική έξοδο. Αφού υπολογιστεί, προωθείται στα προηγούμενα επίπεδα ανανεώνοντας τις παραμέτρους κατάλληλα ώστε να μειωθεί το σφάλμα, μέσω του κανόνα της αλυσιδωτής παραγωγίσης. Ο αλγόριθμος εκπαίδευσης αυτός ονομάζεται αλγόριθμος οπισθοδιάδοσης (Backpropagation Algorithm).

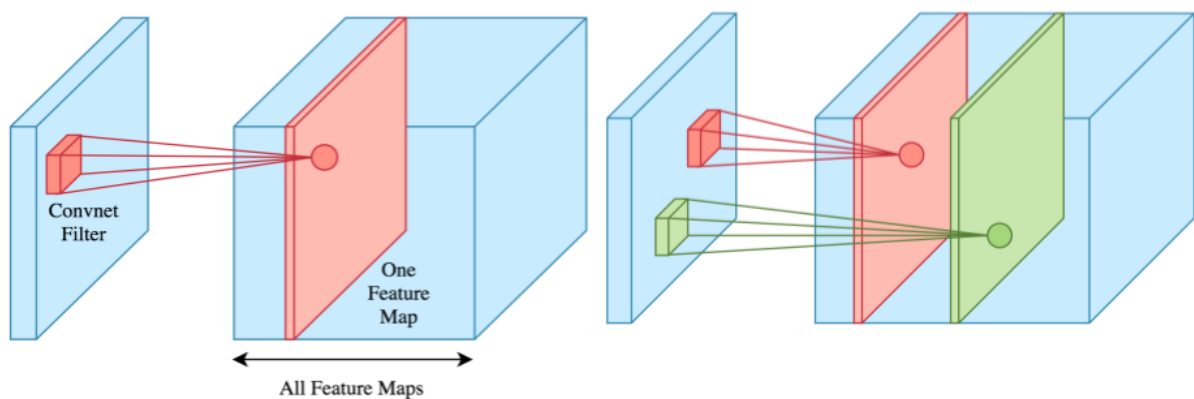
Τα βάρη και οι πολώσεις που αποτελούν τα εκπαιδευσιμα στοιχεία του δικτύου ονομάζονται και παράμετροι του. Ωστόσο η απόδοση του δικτύου κρίνεται και από άλλες παραμέτρους οι οποίες αφορούν την δομή του δικτύου (π.χ αριθμός στρωμάτων, αριθμός νευρώνων/στρώμα κ.α) ή τον αλγόριθμο εκπαίδευσης (π.χ learning rate, batch size κ.α). Αυτές οι παράμετροι ονομάζονται υπερπαραμέτροι και παίζουν καθοριστικό ρόλο στην ποιότητα και την ταχύτητα της εκπαίδευσης ωστόσο δεν ρυθμίζονται κατά την εκπαίδευση αλλά από τον ίδιο τον προγραμματιστή. Δίκτυα με μικρό αριθμό παραμέτρων τείνουν να δυσκολεύονται να διαχωρίσουν τα δεδομένα εκπαίδευσης, ένα φαινόμενο που ονομάζεται underfitting. Αντιθέτως όπως έχει προαναφερθεί μεγάλος αριθμός παραμέτρων μπορεί να οδηγήσει σε υπερεκπαίδευση (overfitting), όπου το δίκτυο δυσκολεύεται να γενικεύσει σε άγνωστα δεδομένα. Στόχος στην σχεδίαση ενός νευρωνικού δικτύου είναι η ισορροπημένη μάθηση όπου και θα έχει μάθει περίπλοκα μοτίβα των δεδομένων του αλλά και θα μπορεί να γενικεύει σε άγνωστα δεδομένα.

Συνελικτικά Νευρωνικά Δίκτυα (CNN)

Τα MLP δέχονται ως είσοδο δεδομένα σε μορφή διανύσματος και έτσι δεν μπορούν να χρησιμοποιηθούν απευθείας για την ανάλυση δεδομένων σε μορφή πλέγματος, όπως π.χ οι εικόνες οι οποίες έχουν μορφή δισδιάστατου πίνακα. Για να δοθεί μια εικόνα ως είσοδος σε ένα MLP πρέπει πρώτα τα δεδομένα να μετασχηματιστούν σε μορφή διανύσματος. Ωστόσο ακόμα και με αυτόν τον τρόπο τα MLP δεν εμφανίζουν επιτυχία σε αυτές τις περιπτώσεις καθώς χάνεται η χωρική πληροφορία των δεδομένων. Την λύση σε τέτοιες περιπτώσεις έρχονται να δώσουν τα συνελικτικά νευρωνικά δίκτυα (CNN). Τα CNN προτάθηκαν πρώτη φορά από τον Yann LeCun (Lecun et al., 1998) και αποτελούν μια κατηγορία βαθιών νευρωνικών δικτύων που συνήθως χρησιμοποιούνται για ανάλυση εικόνων. Η λειτουργία τους βασίζεται στην μαθηματική διαδικασία της συνέλιξης. Αποτελούνται από μια σειρά στρωμάτων κατά τον τρόπο που είναι δομημένα και τα MLP, ωστόσο ένα CNN αντί για

στρώματα με πλήρως συνδεδεμένους νευρώνες, έχει μια σειρά από στρώματα συνέλιξης (Convolutional Layers), στα οποία παρεμβάλλονται ενδιάμεσα συνήθως στρώματα υποδειγματοληψίας (Pooling layers).

Το στρώμα συνέλιξης είναι το βασικό συστατικό του CNN και χαρακτηρίζεται από έναν αριθμό φίλτρο ή πυρήνων (kernels) με τα οποία συνελίσσεται ο τένσορας εισόδου και δίνει ως έξοδο έναν χάρτη ενεργοποίησης (activation map). Τα φίλτρα αποτελούν τις παραμέτρους του δικτύου κατά τον ίδιο τρόπο που ισχύει για τα βάρη των πλήρως συνδεδεμένων νευρώνων στα MLP. Ένα φίλτρο έχει αρκετά μικρότερο μέγεθος. Κατά την συνέλιξη υπολογίζεται το εσωτερικό γινόμενο των στοιχείων του φίλτρου με τα αντίστοιχα στοιχεία της εισόδου και προκύπτει μια βαθμωτή τιμή ως αποτέλεσμα. Το φίλτρο σαρώνει την επιφάνεια της εισόδου τόσο κατά πλάτος όσο και καθ' ύψος και σε κάθε θέση υπολογίζεται μια τιμή ως αποτέλεσμα της συνέλιξης. Με την διαδικασία αυτή προκύπτει ο λεγόμενος χάρτης ενεργοποίησης (feature/activation map) όπου το κάθε στοιχείο του είναι το αποτέλεσμα της συνέλιξης του φίλτρου με την είσοδο στην αντίστοιχη θέση. Έτσι η χωρική πληροφορία παραμένει και το CNN εκπαιδεύεται ώστε τα φίλτρα του να ενεργοποιούνται όταν ανιχνεύεται κάποιο χρήσιμο χωρικό χαρακτηριστικό. Οι χάρτες ενεργοποίησης από όλα τα φίλτρα του στρώματος στοιβάζονται κατά βάθος σχηματίζοντας έτσι την τελική έξοδο του convolutional layer με διαστάσεις (πλάτος χάρτη, ύψος χάρτη, αριθμός φίλτρου).



Εικόνα 2.15: Επίπεδο συνέλιξης. Τα φίλτρα (κόκκινο και πράσινο) συνελίσσονται με την είσοδο και παράγουν τους αντίστοιχους χάρτες ενεργοποίησης

Κάθε στοιχείο του feature map μπορεί να θεωρηθεί ως η έξοδος ενός νευρώνα κατά τον τρόπο που τον είδαμε στα MLP. Κατά τον ίδιο τρόπο οι έξοδοι αυτοί περνούν από μια μη γραμμική συνάρτηση ενεργοποίησης (συνήθως ReLU). Ωστόσο σε αντίθεση με τους πλήρως συνδεδεμένους νευρώνες ο «νευρώνας» σε αυτή τη περίπτωση λαμβάνει ως είσοδο μια μικρή περιοχή της εισόδου και μοιράζεται τις παραμέτρους του με όλους τους υπόλοιπους «νευρώνες» του ίδιου χάρτη ενεργοποίησης. Αυτές οι δύο ιδιότητες είναι πολύ χρήσιμες για την απόδοση των CNN και για αυτό το λόγο θα αναλυθούν παρακάτω. Ο χάρτης στην

συνέχεια περνάει από ένα επίπεδο ενεργοποίησης το οποίο εφαρμόζει μια μη γραμμική συνάρτηση ενεργοποίησης (συνήθως την ReLu) κατά αντιστοιχία με την εφαρμογή της συνάρτησης ενεργοποίησης στην έξοδο του perceptron

Τοπική συνδεσιμότητα

Όπως έχει προαναφερθεί οι νευρώνες ενός επιπέδου σε ένα συνελκτικό δίκτυο συνδέονται μόνο με ένα μικρό μέρος γειτονικών νευρώνων του προηγούμενου επιπέδου. Η ιδιότητα αυτή ονομάζεται τοπική συνδεσιμότητα και είναι πολύ χρήσιμη διότι έτσι διατηρείται η χωρική συσχέτιση των δεδομένων. Αν όλοι οι νευρώνες ήταν συνδεδεμένοι με το σύνολο του προηγούμενου επιπέδου όπως στα MLP, η πληροφορία αυτή θα ήταν αδύνατο να αποκωδικοποιηθεί. Ακόμα, στην περίπτωση των πλήρως συνδεδεμένων νευρώνων, δεδομένα υψηλής διαστατικότητας θα οδηγούσαν σε τεράστιο αριθμό παραμέτρων, αυξάνοντας έτσι την υπολογιστική πολυπλοκότητα της εκπαίδευσης και τον κίνδυνο υπερπροσαρμογής. Το εύρος της συνδεσιμότητας του συνελκτικού δικτύου αποτελεί και ρυθμίσιμη υπερπαραμέτρο του.

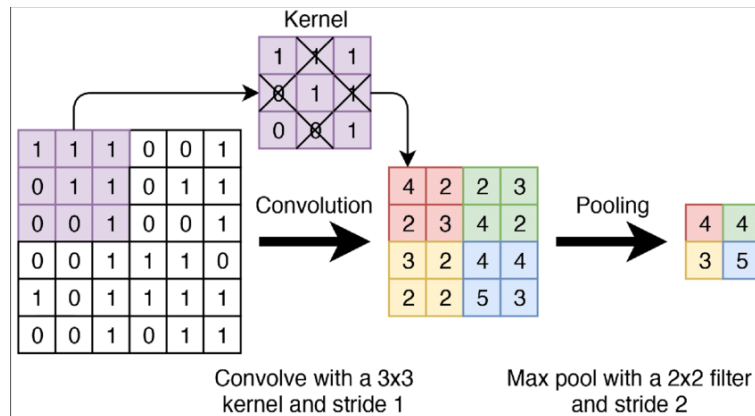
Διαμοιρασμός Παραμέτρων

Όλα τα στοιχεία ενός feature map προκύπτουν από το ίδιο φίλτρο, επομένως εξαρτώνται από τις ίδιες παραμέτρους. Η ιδέα πίσω από αυτή τη σχεδιαστική επιλογή στα CNN είναι ότι εάν ένα φίλτρο με τις τιμές του έτσι όπως έχουν οριστεί, είναι ικανό να εξάγει κάποιο χρήσιμο χαρακτηριστικό από κάποια περιοχή των δεδομένων, τότε είναι πολύ πιθανό να είναι χρήσιμο η εύρεση του ίδιου χαρακτηριστικού και σε άλλες περιοχές. Για παράδειγμα σε ένα φίλτρο μπορεί να ρυθμιστούν οι τιμές του έτσι ώστε να αναγνωρίζει τις κάθετες ακμές μιας εικόνας. Όσο χρήσιμη είναι η εξαγωγή μιας ακμής σε μια περιοχή της εικόνας, τόσο λογικά θα είναι και σε κάποια άλλη. Αυτή η ιδιότητα συμβάλει επίσης στην μείωση των διαφορετικών παραμέτρων που πρέπει να εκπαιδευτούν, καθιστώντας πολύ πιο εύκολη την εκπαίδευση του δικτύου και μειώνοντας τον κίνδυνο υπερεκπαίδευσης.

Επίπεδα υποδειγματολειψίας

Το στρώματα υποδειγματολειψίας μειώνουν τις διαστάσεις των χαρτών ενεργοποίησης, μειώνοντας έτσι τον αριθμό των παραμέτρων, άρα και της υπολογιστικής πολυπλοκότητας του δικτύου και αυξάνοντας την δυνατότητα γενίκευσης. Η είσοδος του pooling layer χωρίζεται σε ορθογώνια. Σε κάθε ορθογώνιο εφαρμόζεται μια συνάρτηση επιλογής μιας αντιπροσωπευτικής τιμής. Συνήθως για την απόφαση της τιμής εξόδου επιλέγεται ή η μέγιστη

τιμή του συμπλέγματος (Max Pooling) ή ο μέσος όρος των τιμών του (Average Pooling). Το στρώμα υποδειματολειψίας χαρακτηρίζεται από την συνάρτηση που έχει επιλεγθεί καθώς επίσης από δυο υπερπαραμέτρους: τις διαστάσεις των συμπλεγμάτων (pool size) και τον βηματισμό (stride). Η ιδέα πίσω από το pooling είναι ότι η ακριβής θέση ενός χαρακτηριστικού δεν έχει τόση σημασία σε συγκριση με την σχετική μεταξύ τους θέση. Ακόμα δίνεται η δυνατότητα συσχέτισης μακρινών χαρακτηριστικών της εισόδου από επόμενα συνελκτικά επίπεδα.

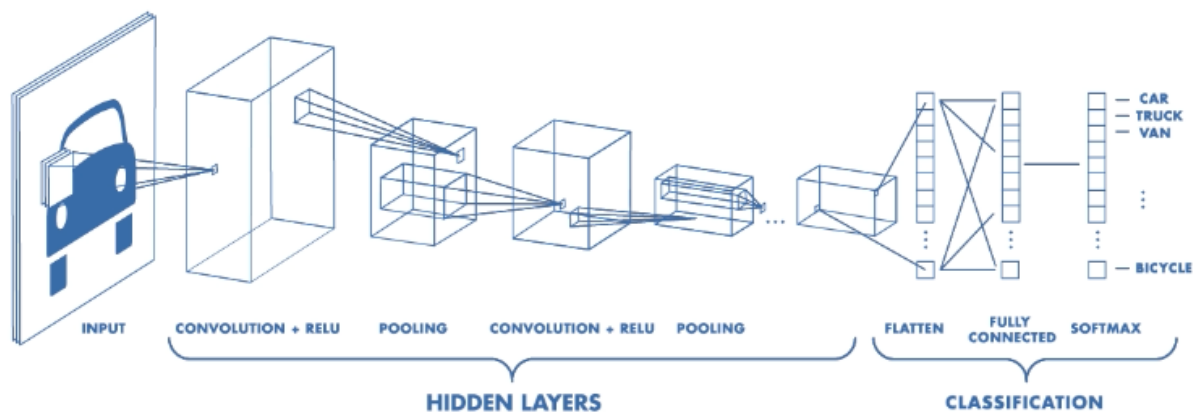


Εικόνα 2.16: Συνέλιξη ακολουθούμενη από υποδειματολειψία μεγίστους 2x2

Εξαγωγή χαρακτηριστικών

Ένα CNN που αποτελείται μόνο από διαδοχικά επίπεδα συνέλιξης και υποδειματολειψίας δεν έχει την ικανότητα ταξινόμησης των δεδομένων του. Τα επίπεδα αυτά εκπαιδεύονται ώστε να εξάγουν υψηλού επιπέδου και μικρότερων διαστάσεων χαρακτηριστικά από τα δεδομένα εκπαίδευσης. Έχει παρατηρηθεί ότι διαδοχικά στρώματα συνέλιξης και pooling μαθαίνουν ολοένα και πιο υψηλού επιπέδου χαρακτηριστικά. Για παράδειγμα με δεδομένα εισόδου εικόνες, το 1^ο convolutional layer πιθανόν θα εκπαιδευτεί να αναγνωρίζει ακμές, το 2^ο σχήματα κ.ο.κ. Ο τρόπος αυτός λειτουργίας θεωρείται παρόμοιος με τον τρόπο λειτουργίας του οπτικού φλοιού του εγκεφάλου. Ένα CNN προκειμένου να αποκτήσει την ιδιότητα του ταξινομητή θα πρέπει να καταλήγει σε έναν ταξινομητή όπως μια μηχανή υποστήριξης διανύσματος (Support Vector Machine), ή σε 1 ή παραπάνω επίπεδα πλήρως συνδεδεμένων νευρώνων. Τα fully connected layers χρησιμοποιούνται συχνά στο τέλος των CNN ωστόσο επειδή δέχονται εισόδους μόνο σε μορφή διανύσματος πρέπει να προηγηθεί ένα επίπεδο Flatten για την κατάλληλη μορφοποίηση των δεδομένων. Με την διασύνδεση αυτή γίνεται εφικτή η ανάπτυξη end – to – end συστημάτων τα οποία δέχονται ακατέργαστα δεδομένα (raw data) ως είσοδο και είναι ικανά να κάνουν αυτόματα την εξαγωγή χαρακτηριστικών και την ταξινόμησή τους με πολύ καλές επιδόσεις. Η εκπαίδευση ενός convolutional neural network γίνεται με τον αλγόριθμο οπισθοδιάδοσης κατ' αναλογία με το MLP. Οι παράμετροι

του δικτύου είναι οι τιμές των φίλτρων καθώς επίσης και τα βάρη και οι πολώσεις των FC επιπέδων του τελικού ταξινομητή. Βασικοί δομικοί υπερπαραμέτροι του δικτύου είναι ο αριθμός επιπέδων, ο αριθμός φίλτρων / επίπεδο, τα μεγέθη των φίλτρων, ο βηματισμός τους (stride) , το μέγεθος της υποδειγματολειψίας (pooling size) των αντίστοιχων επιπέδων, οι συναρτήσεις ενεργοποίησης, καθώς επίσης και ότι ισχύει από τα προηγούμενα για τα MLP.



Εικόνα 2.18: Αρχιτεκτονική συνελκτικού δικτύου με πλήρως συνδεδεμένους νευρώνες ως ταξινομητή

Αξίζει να σημειωθεί ότι τα δίκτυα CNN που περιγράφηκαν παραπάνω είναι τα 2D Convolutional Networks και αποτελούν μόνο μια κατηγορία CNN. Αποτελέσαν τα πρώτα CNN και αφορούν την ανάλυση εικόνων και άλλων 2D δεδομένων. Ωστόσο οι έννοιες που παρουσιάστηκαν παραπάνω μπορούν να μεταφερθούν τόσο σε 1D δεδομένα (π.χ. χρονοσειρές) όσο και σε 3D δεδομένα (π.χ. βίντεο όπου η τρίτη διάσταση είναι ο χρόνος). Έτσι έχουν αναπτυχθεί τα τελευταία χρόνια 1D ConvNets και 3D ConvNets με αξιοσημείωτες επιδόσεις.

Αναδρομικά Νευρωνικά Δίκτυα (RNN)

Τα νευρωνικά δίκτυα που περιγράφηκαν στις προηγούμενες υποενότητες ονομάζονται δίκτυα πρόσθιας τροφοδότησης. Ο λόγος που ονομάζονται έτσι είναι ότι όλοι οι υπολογισμοί έχουν την ίδια κατεύθυνση από την είσοδο προς την έξοδο. Δεν υπάρχει ανατροφοδότηση από έναν νευρώνα σε νευρώνες προηγούμενου ή του ίδιου επιπέδου. Υπάρχει όμως και μια δεύτερη κατηγορία νευρωνικών δικτύων τα λεγόμενα Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks). Στα δίκτυα αυτά υπάρχουν τέτοιες ανατροφοδοτήσεις προσδίδοντας στο δίκτυο μνήμη προηγούμενων χρονικών στιγμών καθώς για να υπολογιστεί η έξοδος ενός νευρώνα την χρονική στιγμή t , χρησιμοποιούνται εκτός των εισόδων της ίδιας χρονικής

στιγμής και έξοδοι από προηγούμενες στιγμές. Η ιδιότητα αυτή τα καθιστά ιδανικά σε προβλήματα είτε κατηγοριοποίησης, είτε πρόβλεψης, με ακολουθιακά δεδομένα όπως είναι η φωνή, οι λεκτικές φράσεις ή τα ηλεκτροεγκεφαλογραφήματα όπως στην περίπτωση μας.

Απλό RNN:

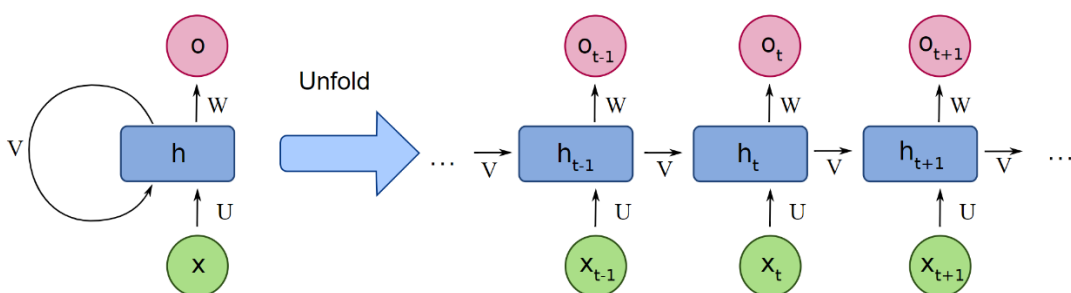
Ένα βασικό αναδρομικό νευρωνικό δίκτυο μοιάζει ως ένα σημείο με τα MLP καθώς όλοι οι νευρώνες οργανώνονται σε επίπεδα κατά τα γνωστά και κάθε κόμβος δέχεται ως εισόδους όλες τις εξόδους από το προηγούμενο επίπεδο. Επιπρόσθετα όμως ένας RNN νευρώνας διαθέτει και μια ανατροφοδότηση από την έξοδο του στην είσοδο του. Έτσι η έξοδος του την χρονική στιγμή t επηρεάζεται, εκτός από τις εισόδους του την ίδια χρονική στιγμή, από την προηγούμενη έξοδό του την χρονική στιγμή $t-1$. Φορμαλιστικά η λειτουργία του αναδρομικού νευρώνα περιγράφεται από την παρακάτω εξίσωση:

$$H_t = f (W [H_{t-1}, X_t] + b)$$

Όπου:

- H_t : έξοδος την χρονική στιγμή t
- H_{t-1} : έξοδος την χρονική στιγμή $t-1$
- f : συνάρτηση ενεργοποίησης (συνήθως \tanh)
- W : διάνυσμα βαρών
- X_t : διάνυσμα εισόδου την χρονική στιγμή t
- b : πολώσεις

Το αναδρομικό δίκτυο μπορεί να αναπαρασταθεί καλύτερα ως ένα βαθύτερο δίκτυο όπου ξεδιπλώνονται οι υπολογισμοί παλαιότερων χρονικών στιγμών και αναπαριστώνται ως βοηθητικοί νευρώνες. Οι νευρώνες αυτοί δεν είναι πραγματικοί νευρώνες του δικτύου αλλά με αυτήν την αναπαράσταση γίνεται εμφανής ο τρόπος μετάδοσης της πληροφορίας στο χρόνο.



Εικόνα 2.19: Αναδρομικός νευρώνας «ξεδιπλωμένος» στον χρόνο

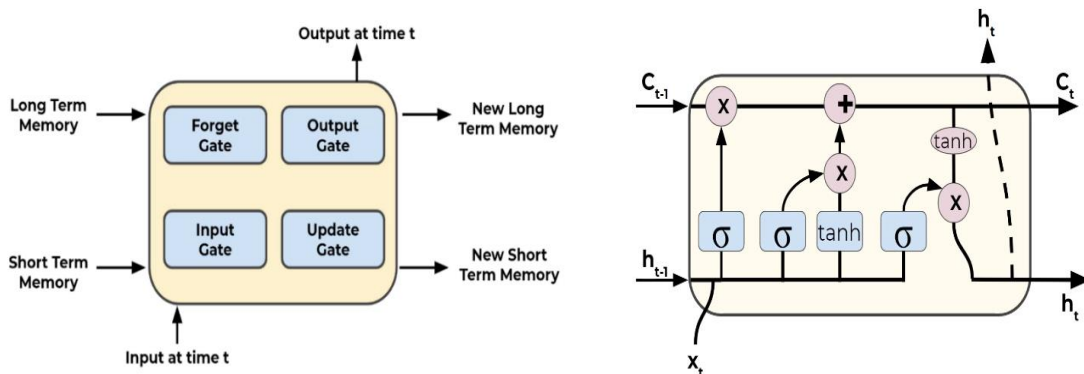
Εκπαίδευση του RNN:

Ο αλγόριθμος οπισθοδιάδοσης σφάλματος (Back propagation) με τον οποίο γίνεται η εκπαίδευση των δικτύων πρόσθιας τροφοδότησης δεν μπορεί να λειτουργήσει στην

περίπτωση των αναδρομικών δικτύων καθώς έχει τον περιορισμό ότι λειτουργεί μόνο σε ακυκλικούς γράφους. Για αυτό το λόγο αναπτύχθηκε μια παραλλαγή του back propagation που ονομάζεται αλγόριθμος οπίσθιας διάδοσης σφάλματος μέσω του χρόνου (Back Propagation through time). Στον αλγόριθμο αυτόν οι κόμβοι αρχικά ξεδιπλώνονται κατά τον τρόπο που φαίνεται και στο παραπάνω σχήμα και έτσι προκύπτει ένας ακυκλικός γράφος στον οποίο μέσω του κανόνα αλυσιδωτής παραγωγίσης (gradient descent) προσαρμόζονται τα βάρη τόσο μεταξύ των πραγματικών κόμβων όσο και των βοηθητικών, μαθαίνοντας έτσι τις σημαντικές χρονικές εξαρτήσεις. Ωστόσο η εκπαίδευση των απλών RNN δικτύων έτσι όπως έχουν περιγραφεί μέχρι τώρα εμφανίζουν δυο βασικά προβλήματα. Το ένα είναι ότι δεν μπορούν να μάθουν μακροχρόνιες εξαρτήσεις καθώς κάθε νευρώνας θυμάται μόνο την αμέσως προηγούμενη έξοδο. Το δεύτερο πρόβλημα είναι ότι η ξεδιπλωση του δικτύου στον χρόνο προκειμένου να εκπαιδευτεί οδηγεί σε ένα πολύ βαθύ νευρωνικό δίκτυο με πάρα πολλά επίπεδα. Τα πολύ βαθιά νευρωνικά πάσχουν από την εξασθένηση της κλίσης (Vanishing gradient). Κατά την οπισθοδιάδοση έχει προαναφερθεί ότι τα βάρη αλλάζουν με αλυσιδωτές παραγωγίσεις. Συνήθως διαδοχικές παραγωγίσεις οδηγούν σε ολοένα και μικρότερους παραγώγους. Έτσι όταν το σφάλμα προωθηθεί σε πολλά διαδοχικά επίπεδα καταλήγει να μικραίνει τόσο που οι παράμετροι των αρχικών στρωμάτων ουσιαστικά δεν αλλάζουν. Επομένως το δίκτυο αδυνατεί να εξάγει τα βασικά χαρακτηριστικά που διακρίνουν τα αρχικά στρώματα. Για να λυθούν αυτά τα δυο προβλήματα έχουν αναπτυχθεί μεταξύ άλλων αρχιτεκτονικών τα δίκτυα μακράς βραχείας μνήμης (Long Short Term Memory - LSTM).

LSTM

Ένας LSTM νευρώνας όπως και ένα basic RNN δέχεται ως εισόδους τις εξόδους των νευρώνων του προηγούμενου στρώματος καθώς και την έξοδο του την προηγούμενη χρονική στιγμή. Ωστόσο διαθέτει και μια επιπλέον είσοδο η οποία αφορά την μακροχρόνια μνήμη του. Η μακροχρόνια μνήμη του LSTM νευρώνα ονομάζεται και κατάσταση του. Για την επιλογή τόσο της ανανέωσης της κατάστασης από την είσοδο, όσο και της μετάβασης μιας κατάστασης στην επόμενη χρονικά κρυφή κατάσταση, γίνεται χρήση κάποιων ειδικών πυλών εντός του νευρώνα, υπεύθυνων για την απόφαση.



Εικόνα 2.20: Εσωτερική αρχιτεκτονική νευρώνα Μακράς Βραχείας Μνήμης

Οι ειδικές αυτές πύλες είναι 4 και η λειτουργία τους περιγράφεται σύντομα παρακάτω:

- Πύλη λησμονιάς (Forget Gate): Αποφασίζει ποια είσοδο να ξεχάσει από προηγούμενες στιγμές. Η εξίσωση που περιγράφει την έξοδο της πύλης είναι η $f_t = \sigma(W_f[h_{t-1}, X_t] + b_f)$.
- Πύλη εισόδου (Input Gate): Αποφασίζει ποια είσοδο αποδέχεται ο νευρώνας. Η εξίσωση που περιγράφει την έξοδο της πύλης είναι η $i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$.
- Πύλη ανανέωσης (Update Gate): Αποφασίζει πότε θα ανανεωθούν οι μνήμες. Οι εξισώσεις που την περιγράφουν είναι $\check{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c)$ και $C_t = f_t * C_{t-1} + i_t * \check{C}_t$.
- Πύλη εξόδου (Output Gate): Καθορίζει την έξοδο του νευρώνα. Οι εξισώσεις που την περιγράφουν είναι οι $O_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$ και $h_t = O_t * \tanh(C_t)$.

Αναδρομικό Συνελκτικό Δίκτυο (Recurrent Convolutional Network)

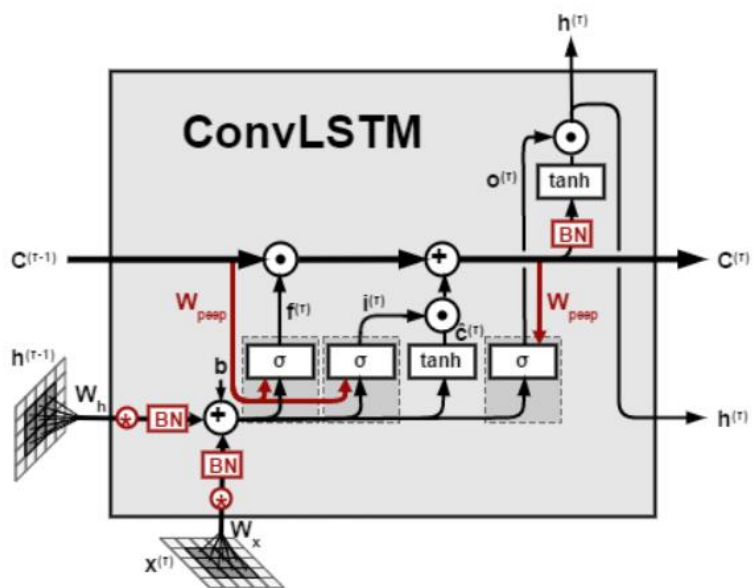
Το βασικό μειονέκτημα των πλήρως συνδεδεμένων LSTM είναι ότι δεν μπορούν να χειριστούν χωροχρονικά δεδομένα. Ο λόγος είναι ότι πριν από την επεξεργασία τους πρέπει να τα ξεδιπλώσουν σε 1D διανύσματα και ως εκ τούτου χάνεται όλη η χωρική πληροφορία τους. Για τον λόγο αυτό αναπτύχθηκε μια νέα κατηγορία νευρωνικών δικτύων, τα αναδρομικά συνελκτικά δίκτυα (Recurrent Convolutional Network – RCNN). Τα RCNN συνδυάζουν τις ιδιότητες των αναδρομικών δικτύων για την αναγνώριση χρονικών εξαρτήσεων και των συνελκτικών νευρωνικών δικτύων για την εξαγωγή χωρικών χαρακτηριστικών. Έχουν αναπτυχθεί κυρίως για την αντιμετώπιση δεδομένων σε μορφή αλληλουχιών εικόνας όπως είναι το βίντεο, καθώς μπορούν να εξάγουν και χωρικά και χρονικά χαρακτηριστικά. Στα πλαίσια της εργασίας γίνεται χρήση των Convolutional LSTM. Τα δίκτυα αυτά όπως υποδηλώνει και το όνομά τους αποτελούν ένα συγκεκριμένο είδος RCNN βασισμένο στην δομή των LSTM. Προτάθηκαν πρώτη φορά από τους Shi et al., 2015 για το πρόβλημα της πρόβλεψης της έντασης της βροχόπτωσης τοπικά σε μια περιοχή, σε ένα σχετικά σύντομο χρονικό διάστημα. Η δομή τους είναι παρόμοια με τα LSTM, όμως οι εισόδοι X_1, \dots, X_t , οι εξόδοι C_1, \dots, C_t , οι κρυφές καταστάσεις H_1, \dots, H_t , καθώς και οι πύλες i_t, f_t, o_t , έχουν αντικατασταθεί από 3D τένσορες των οποίων οι δυο διαστάσεις είναι χωρικές. Έτσι το ConvLSTM καθορίζει την μελλοντική κατάσταση ενός κελιού του τένσορα από τις εισόδους και τις παρελθοντικές καταστάσεις των γειτονικών κελιών. Αυτό επιτυγχάνεται χρησιμοποιώντας την πράξη της συνέλιξης τόσο στις μεταβάσεις εισόδου – κατάστασης όσο

και στις μεταβάσεις από μια κατάσταση σε μια άλλη. Οι βασικές πράξεις λοιπόν που πραγματοποιούνται σε ένα νευρώνα LSTM μετασχηματίζονται στις παρακάτω:

- $i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i)$
- $f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f)$
- $C_t = f_t \circ C_{t-1} + i_t \circ \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c)$
- $o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} \circ C_t + b_o)$
- $H_t = o_t \circ \tanh(C_t)$

όπου το σύμβολο * συμβολίζει την πράξη της συνέλιξης, ενώ το \circ συμβολίζει το γινόμενο Hadamard (σημείο προς σημείο).

Διατηρούνται κατά αυτό τον τρόπο οι αρχικές διαστάσεις των δεδομένων και εξάγονται από την συνέλιξη χωρικά χαρακτηριστικά, ενώ οι LSTM μνήμες μαθαίνουν να αναγνωρίζουν τις χρονικές εξαρτήσεις των δεδομένων.



Εικόνα 2.21: Εσωτερική αρχιτεκτονική νευρώνα ConvLSTM

Κεφάλαιο 3: Ανασκόπηση Βιβλιογραφίας

Οι διεπαφές εγκεφάλου υπολογιστή αποτελούν έναν τομέα υπό εντατική έρευνα τα τελευταία χρόνια με πλήθος μελετών και διαφορετικών, πολλά υποσχόμενων εφαρμογών όπως για παράδειγμα συστήματα συλλάβησης (Speier et al., 2015; Stawicki et al., 2017), συστήματα πρόβλεψης της επιληψίας (Maksimenko et al., 2017), έλεγχος αναπηρικού αμαξιδίου (Carlson & Del R. Millan, 2013) και άλλα. Εμείς στο κεφάλαιο αυτό θα αναφερθούμε σε μελέτες που έχουν προηγηθεί συγκεκριμένα στον τομέα των Motor Imagery BCIs με χρήση ηλεκτροεγκεφαλογραφημάτων. Αρχικά θα εστιάσουμε σε έρευνες που αφορούν κλασσικές μεθόδους αποκωδικοποίησης των σημάτων στις οποίες γίνεται εκτεταμένη χρήση τεχνικών εξαγωγής χαρακτηριστικών, ενώ στη συνέχεια στο δεύτερο μέρος του κεφαλαίου θα δούμε εργασίες στις οποίες γίνεται χρήση μοντέλων βαθιάς μάθησης.

3.1) Αποκωδικοποίηση MI με χρήση κλασσικών τεχνικών μηχανικής μάθησης

Διάφορες τεχνικές εξαγωγής χαρακτηριστικών χρησιμοποιούνται για την ανάλυση EEG σημάτων. Οι Rodriguez-Bermudez & García-Laencina, 2012 υπολόγισαν την πυκνότητα φάσματος ισχύος (PSD) των EEG σημάτων και με χρήση Γραμμικής Διακριτικής Ανάλυσης (LDA) για την ταξινόμηση τους, πέτυχαν ορθότητα 69.4% στο dataset 4a από τον διαγωνισμό BCI Competition III (Blankertz et al., 2006). Η χρήση ωστόσο φασματικών χαρακτηριστικών για την ταξινόμηση των EEG δεν είναι επαρκής και δεν οδηγεί σε αξιόλογα αποτελέσματα καθώς δεν παρέχει πληροφορία που να συσχετίζει το φασματικό περιεχόμενο του σήματος με το πεδίο του χρόνου. Για αυτό το λόγο χρησιμοποιείται συχνότερα ανάλυση στο πεδίο του χρόνου – συχνοτήτων. Οι Kevric and Subasi 2017 στην εργασία τους εξάγουν τέτοια χαρακτηριστικά κάνοντας χρήση του διακριτού μετασχηματισμού κυματίων (Discrete Wavelet Transform - DWT) . Τα σήματα που προκύπτουν από τον μετασχηματισμό αυτό περιέχουν φασματική πληροφορία σε συνάρτηση με τον χρόνο και, μάλιστα, με χρήση μιας επέκτασης του DWT, τον αλγόριθμο Wavelet Packet Decomposition (WPD), κατάφεραν να εξάγουν άκομα πιο λεπτομερείς πληροφορίες για το φασματικό περιεχόμενο των EEG. Στη συνέχεια έγινε επιλογή των καταλληλότερων χαρακτηριστικών με χρήση του κριτηρίου Kaiser, ενώ για την απομόνωση του θορύβου χρησιμοποιήθηκε η μέθοδος PCA. Τέλος, ως ταξινομητής χρησιμοποιήθηκε ο kNN και επιτεύχθηκε ορθότητα 92.8%.

Μια από τις πιο συχνά χρησιμοποιούμενες είναι ο αλγόριθμος Common Spatial Patterns (CSP) ο οποίος αποσκοπεί στην εύρεση χωρικά κωδικοποιημένης πληροφορίας και το

πετυχαίνει δίνοντας κατάλληλα βάρη στο κάθε κανάλι EEG. Αξιοποιεί το γεγονός ότι διαφορετικές νοητικές εργασίες ενεργοποιούν άλλες περιοχές του εγκεφάλου με αποτέλεσμα να δημιουργούν διαφορετικά δυναμικά στην επιφάνεια του κρανίου και προτιμάται σε πολλές εργασίες σχετικές με ταξινόμηση MI καθώς πετυχαίνει πολύ καλά αποτελέσματα. Οι Baig et al. 2017 φιλτράροντας τα σήματα με CSP και στη συνέχεια κάνοντας χρήση εξελκτικών αλγορίθμων για την επιλογή των βέλτιστων χαρακτηριστικών, όπως τον αλγόριθμο βελτιστοποίηση σμήνους σωματιδίων (PSO) πέτυχαν ορθότητα 90.4%. Με χρήση βελτιστοποίησης ABΓ (ABC optimization) για την επιλογή χαρακτηριστικών η ορθότητα αυξήθηκε σε 94.48% Για την ταξινόμηση έκαναν χρήση κλασσικών ταξινομητών όπως SVM, LDA και kNN. Πρέπει να σημειωθεί ότι και οι δυο προηγούμενες έρευνες πραγματοποιήθηκαν με το dataset 4a από το BCI Competition III το οποίο έχει μικρό μέγεθος με καταγραφές από μόλις 5 υποκείμενα.

Όπως έχει προαναφερθεί στο κεφάλαιο 2, ο CSP πάσχει από κάποιες αδυναμίες καθώς δεν μπορεί να διαχωρίσει την βέλτιστη ζώνη συχνοτήτων για την αναγνώριση της εγκεφαλικής δραστηριότητας στο κάθε άτομο ξεχωριστά. Επομένως σε πολλές εργασίες προτείνονται BCIs που βασίζονται σε βελτιώσεις του CSP. Οι Kim et al. 2016 προτείνουν αρχικά το διαχωρισμό του σήματος στους μ και β ητα εγκεφαλικούς ρυθμούς και στη συνέχεια την εφαρμογή του αλγορίθμου SUTCCSP (strong uncorrelating transform complex common spatial patterns) ο οποίος αποσυσχετίζει τους δυο ρυθμούς και παρέχει πληροφορία σχετικά με την διαφορά ισχύος τους ενώ παρέχει και πληροφορία σχετικά με την χωρική κατανομή της ισχύος του σήματος μέσω του CSP τμήματος του αλγορίθμου. Τέλος κάνοντας χρήση του αλγορίθμου τυχαίου δάσους (Random Forest) πέτυχε ορθότητα 80.05% στην ταξινόμηση δεξιού και αριστερού χεριού με το Motor Imagery Dataset του PhysioNet. Οι Handiru and Prasad 2016 επίσης στο ίδιο dataset προτείνουν την μέθοδο FBCSP όπως έχει περιγραφεί στο κεφάλαιο 2 για την εξαγωγή των χαρακτηριστικών των σημάτων, ενώ μετά τη σύγκριση διαφόρων μεθόδων για την βέλτιστη επιλογή καναλιών καταγραφής, όπως ανάλυση κυρίων συνιστωσών (PCA), PSO και την μέθοδο επαναληπτικής βελτιστοποίησης πολλαπλών στόχων (IMOCS) επιλέχθηκε η τελευταία. Το αποτέλεσμα τους ήταν 63.62% ορθότητα για ταξινόμηση δεξιού αριστερού χεριού ενώ όταν κράτησαν τα 35 καλύτερα υποκείμενα η ορθότητα ανέβηκε στο 79%.

<i>Ερευνητική ομάδα</i>	Εξαγωγή Χαρακτηριστικών	Επιλογή Χαρακτηριστικών	Ταξινομητής	Ορθότητα
<i>Rodríguez-Bermúdez et al. 2012</i>	PSD	Lars	LDA	69.4%
<i>Kevric and Subasi 2017</i>	DWT, WPD	Kaiser criterion	kNN	92.8%
<i>Baig et al. 2017</i>	CSP	PSO, ABC optimization	SVM, LDA, kNN	90,4% (PSO) 94.2% (ABC opt)
<i>Kim et al. 2016</i>	SUTCSSP	-	Random Forest	80.05%*
<i>Handiru and Prasad 2016</i>	FBCSP	IMOCS	SVM	63.64%

Πίνακας 3.1: Μέθοδοι κατηγοριοποίησης EEG MI σημάτων από παλαιότερες έρευνες

*Πρέπει να σημειωθεί ότι η ορθότητα που αναφέρεται στην μελέτη των Kim et al. 2016 αφορά εκπαίδευση εντός κάποιων επιλεγμένων υποκειμένων και όχι καθολική.

3.2) Αποκωδικοποίηση MI με χρήση τεχνικών βαθιάς μάθησης

Αναδρομικά Νευρωνικά Δίκτυα (RNN)

Από τις παραπάνω μελέτες και από πολλές άλλες ακόμα, γίνεται κατανοητό ότι η ανάλυση των EEG σημάτων απαιτεί προηγμένες τεχνικές προεπεξεργασίας. Πολλές διαφορετικές προσεγγίσεις και βελτιστοποιήσεις έχουν αναπτυχθεί ωστόσο τα αποτελέσματα τους έχουν αρκετό χώρο για βελτίωση. Η βαθιά μάθηση, η οποία τα τελευταία χρόνια έχει πετύχει σημαντικά αποτελέσματα σε διάφορες εφαρμογές αποτελεί μια πολλά υποσχόμενη τεχνολογία για την αποτελεσματικότερη αποκωδικοποίηση των EEG και την περαιτέρω ανάπτυξη των BCIs. Επομένως όπως είναι φυσικό πλήθος ερευνών διεξάγεται για αυτό το σκοπό. Ο Greaves 2011 έκανε χρήση αναδρομικών δικτύων για την αναγνώριση τρισδιάστατης αντίληψης των υποκειμένων μέσω EEG. Συγκεκριμένα χρησιμοποίησε δίκτυα Elman (Elman, 1990) για την ταξινόμηση ηλεκτροεγκεφαλογραφημάτων κατά τα οποία το

υποκείμενο κοιτούσε είτε ένα δισδιάστατο είτε ένα τρισδιάστατο αντικείμενο. Με την χρήση αναδρομικών δικτύων πέτυχε ορθότητα μόλις 58% ενώ με την χρήση απλών πολυεπίπεδων perceptron το ποσοστό έφτασε στο 72%. Έτσι συμπαίρανε πως τα αναδρομικά νευρωνικά δίκτυα δεν είναι κατάλληλα για την ανάλυση EEG. Ωστόσο, οι Hema et al. 2008 χρησιμοποιώντας δίκτυα Elman και αξιοποιώντας μόνο δυο κανάλια καταγραφής, C3 και C4, πέτυχαν ορθότητα 93% σε ένα dataset που περιείχε τέσσερις MI κινήσεις από δυο άτομα. Οι Zhang et al. 2017 ανέπτυξαν ένα βαθύ αναδρομικό δίκτυο με 7 επίπεδα RNN και 2 επίπεδα LSTM και κατάφεραν να πετύχουν 95.53% για 5 MI κινήσεις (δεξί χέρι, αριστερό χέρι, δυο χέρια, δυο πόδια, ηρεμία) ανάμεσα σε 109 άτομα από το PhysioNet eegmi dataset.

Συνελικτικά Νευρωνικά Δίκτυα (CNN)

Η εξαγωγή χαρακτηριστικών των EEG σημάτων αποτελεί μεγάλη πρόκληση όπως έχει προαναφερθεί. Ως λύση σε αυτό το πρόβλημα έχουν αξιοποιηθεί σε διάφορες μελέτες τα CNN, τα οποία έχουν την ιδιότητα να εξάγουν αυτόματα χρήσιμα χαρακτηριστικά των δεδομένων. Πρωτοπόροι στην ανάπτυξη συνελικτικών δικτύων για την αποκωδικοποίηση πληροφορίας σχετικής με κίνηση από ακατέργαστα EEG δεδομένα υπήρξαν οι Schirrmeyer et al. 2017. Στην εργασία τους ανέπτυξαν ένα βαθύ και ένα ρηχό CNN (dCNN & sCNN). Και στις δυο περιπτώσεις το πρώτο συνελικτικό στρώμα εφάρμοζε συνέλιξη ως προς το χρόνο και εξήγαγε χρονικά χαρακτηριστικά ενώ το δεύτερο συνελικτικό στρώμα εφάρμοζε χωρική συνέλιξη ως προς τα κανάλια καταγραφής εξάγοντας έτσι και χωρική πληροφορία από το σήμα. Το υπόλοιπο δίκτυο και στο dCNN και στο sCNN αποτελούνταν από τυπικά μπλοκ συνέλιξης – υποδειγματολειψίας μεγίστου, ενώ για την ταξινόμηση υπήρχε ένα πλήρως συνδεδεμένο επίπεδο στο τέλος. Επιπρόσθετα τα αποτελέσματα αξιολογήθηκαν σε διαφορετικές ζώνες συχνοτήτων για την εύρεση της βέλτιστης. Η ορθότητα των μοντέλων αυτών ήταν 71.9% και 70.10% αντίστοιχα στο dataset 4a BCI competition III (Blankertz et al., 2006). Οι Lawhern et al., 2018 ανέπτυξαν το EEGNet, ένα ελαφρύ συνελικτικό δίκτυο το οποίο είχε την ικανότητα να επιλύει πολλά διαφορετικά BCI προβλήματα. Οι συγγραφείς της εργασίας είχαν ως σκοπό την δημιουργία ενός δικτύου γενικού σκοπού, το οποίο δεν ήταν εξειδικευμένο μόνο για ένα συγκεκριμένο πρόβλημα. Οι αρχές λειτουργίας του είναι εμπνευσμένες από τον αλγόριθμο FBCSP και εκτελεί αρχικά μια χρονική και μια χωρική συνέλιξη, ενώ στη συνέχεια κάνει χρήση διαχωρίσιμης συνέλιξης για την βέλτιστη εξαγωγή χωροχρονικής πληροφορίας. Δοκιμάστηκε με επιτυχία σε διαφορετικών ειδών EEG datasets μεταξύ των οποίων και το BCI competition IV 2a (Tangermann et al., 2012), ένα MI dataset με 4 φανταστικές κινήσεις από 9 άτομα, συγκεκριμένα δεξί χέρι, αριστερό χέρι, πόδια, γλώσσα, πετυχαίνοντας ορθότητα 70%.

Αναδρομικά Συνελικτικά Δίκτυα (RCNN)

Ενδιαφέρουσα προσέγγιση είχαν οι (Popov & Fomenkov, 2016), οι οποίοι χρησιμοποίησαν αναδρομικά συνελικτικά δίκτυα (RCNN) για την ταξινόμηση 6 διαφορετικών κινήσεων των χεριών πετυχαίνοντας ορθότητα 94.8%. Έδειξαν ακόμη ότι η RCNN αρχιτεκτονική τους, η οποία αναλύει τα eeg ως μια ακολουθία «εικόνων», επιτυγχάνει καλύτερα αποτελέσματα σε σχέση με τα κλασικά CNN. Ωστόσο πρέπει να σημειωθεί ότι το μοντέλο τους αξιολογήθηκε μόνο σε πραγματικές κινήσεις και όχι φαντασικές. RCNN μοντέλο για την ταξινόμηση MI-EEG δεδομένων που σύλλεξαν οι ίδιοι από 20 άτομα και συγκεκριμένα για την ταξινόμηση δεξιού – αριστερού χεριού, χρησιμοποίησαν μεταξύ άλλων αρχιτεκτονικών και οι (Tayeb et al., 2019). Η μέση ορθότητα που πέτυχε το συγκεκριμένο μοντέλο ήταν 77.72%, αρκετά υψηλότερη από άλλες μεθόδους που αξιολογήθηκαν στην εργασία. Παρόλα αυτά δεν ήταν το πιο αποδοτικό μοντέλο. Η προτεινόμενη μέθοδος από τους συγγραφείς είχε ως αρχικό στάδιο την μετατροπή των EEG σημάτων σε μορφή εικόνας μέσω της μετατροπής σε σπεκτρογράμμα με χρήση του μετασχηματισμού STFT. Στην συνέχεια ακολουθούν τρια μπλοκ συνέλιξης – δειγματοληψίας μεγίστου με χρήση Batch Normalization και ReLU συνάρτηση ενεργοποίησης. Τέλος ένα πλήρως συνδεδεμένο επίπεδο είναι υπεύθυνο για την πρόβλεψη. Με την αρχιτεκτονική αυτή οι συγγραφείς πέτυχαν 84.24% μέση ορθότητα ενώ δοκιμάστηκε και στο δημοσίως διαθέσιμο EEG Graz dataset πετυχαίνοντας ανάλογα αποτελέσματα. Για την αποφυγή της υπερεκπαίδευσης με το μοντέλο αυτό έκαναν χρήση Dropout πριν το πλήρως συνδεδεμένο επίπεδο καθώς επίσης και επαύξηση των δεδομένων εκπαίδευσης χωρίζοντας την κάθε δοκιμή MI σε επικαλυπτόμενα διαστήματα.

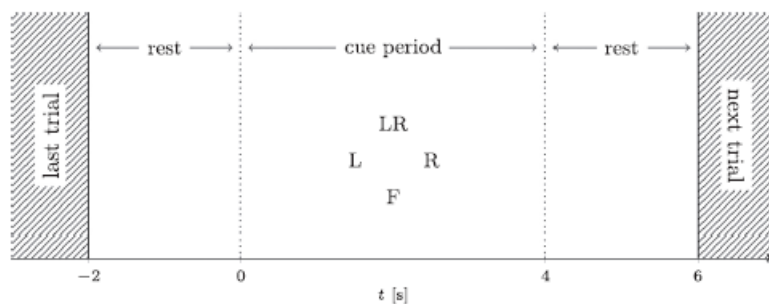
Ερευνητική ομάδα	Αρχιτεκτονική	BCI πρόβλημα	Ορθότητα
Greaves 2011	Elman RNN	2D vs 3D αντίληψη	58%
Hema et al. 2008	Elman RNN	4 MI κινήσεις 2 ατόμων	93%
Zhang et al. 2017	Deep LSTM	5 MI Physionet dataset	95.53%
Schirrmester et al. 2017	dCNN – sCNN	4a BCI Competition III	71.9% - 70.10%
Lawhern et al. 2018	EEGNet	2a BCI Competition IV	70%
Popov and Fomenkov 2016	RCNN	6 πραγματικές κινήσεις χεριού	94.8%
Tayeb et al. 2019	RCNN, STFT-CNN	MI δεξιού & αριστερού χεριού από 20 άτομα	77.72% - 84.24%

Πίνακας 3.2: Αρχιτεκτονικές βαθιάς μάθησης EEG MI σημάτων από παλαιότερες έρευνες

Κεφάλαιο 4: Πειραματική Διαδικασία

4.1) Περιγραφή Dataset

Για την διεξαγωγή των πειραμάτων μας χρησιμοποιούμε το Physionet EEG Motor Movement/Imagery Dataset (A. L. Goldberger et al.) το οποίο είναι διαθέσιμο δημόσια και περιέχει καταγραφές EEG από 109 υποκείμενα. Τα σήματα καταγράφηκαν με το σύστημα καταγραφής BCI2000 (Schalk et al., 2004) το οποίο αποτελεί μια πλατφόρμα με εργαλεία για διεξαγωγή πειραμάτων BCI. Η καταγραφή έγινε με χρήση 64 ηλεκτροδίων τοποθετημένων κατά το διεθνές πρότυπο 10-10, ενώ η συχνότητα δειγματοληψίας ήταν στα 160Hz. Κάθε υποκείμενο κλήθηκε να συμμετάσχει σε 14 συνεδρίες. Στις πρώτες δυο οι καταγραφές ήταν διάρκειας 1 λεπτού η καθεμιά και σε κατάσταση ηρεμίας (baseline run). Η μία με ανοιχτά μάτια και η άλλη με κλειστά. Στις υπόλοιπες συνεδρίες το κάθε υποκείμενο κλήθηκε είτε να εκτελέσει κάποια κίνηση, είτε να την φανταστεί. Συγκεκριμένα σε 3 συνεδρίες έσφιγγε είτε την αριστερή είτε την δεξιά γροθιά, σε 3 έσφιγγε είτε και τις δυο γροθιές του είτε τα πόδια του, σε 3 φανταζόταν ότι έσφιγγε είτε την δεξιά είτε την αριστερή γροθιά (motor imagery) και στις τελευταίες 3 φανταζόταν ότι έσφιγγε είτε και τα δυο του χέρια, είτε και τα δυο του πόδια. Κάθε συνεδρία διαρκεί 2 λεπτά και περιλαμβάνει 14 MI δοκιμές διάρκειας 4 δευτερολέπτων. Τη χρονική στιγμή $t=0$ εμφανίζεται ένα σημάδι στην οθόνη που ανάλογα το σημείο εμφάνισής του, υποδεικνύει την κίνηση (νοητική ή μη) που έχει να εκτελέσει το υποκείμενο. Αν εμφανιστεί δεξιά ή αριστερά σημαίνει σύσφιξη του αντίστοιχου χεριού ενώ πάνω ή κάτω σημαίνει σύσφιξη χεριών ή ποδιών αντίστοιχα. Ο στόχος παρέμενε στην οθόνη για τα 4 δευτερόλεπτα και όταν εξαφανιζόταν δίνονταν 4 δευτερόλεπτα ξεκούρασης. Έτσι για κάθε υποκείμενο έχουμε 21 δείγματα ανά κλάση. Στην μελέτη μας δεν έχουμε λάβει υπόψιν τα πειράματα με τις πραγματικές κινήσεις. Ο λόγος είναι ότι η αποκωδικοποίηση της πραγματικής κίνησης δεν έχει τόσο μεγάλη αξία όσο της φαντασιακής, καθώς η δεύτερη ανοίγει τον δρόμο για την ανάπτυξη BCIs για άτομα που έχουν χάσει πλήρως την κινητικότητα των άκρων τους και την δραστική βελτίωση της ζωής τους.



Εικόνα 4.1: Πειραματική διαδικασία κατά την καταγραφή του Physionet EEG dataset. Μετά από μια περίοδο ξεκούρασης το υποκείμενο καλείται να φανταστεί μια κίνηση για 4s. Πηγή εικόνας: (Dose et al., 2018)

Μεθοδολογία:

Από τα 109 υποκείμενα έχουν αποριφθεί οι S88, S92, S100 και S106 λόγω ασυνέπειας των καταγραφών τους με των υπολοίπων. Έτσι στο dataset έμειναν 105 υποκείμενα. Από το baseline run έχουμε εξάγει παράθυρα των 3 δευτερολέπτων, ενώ από κάθε MI trial έχουμε κρατήσει τα 3 πρώτα δευτερόλεπτα. Με συχνότητα δειγματοληψίας 160Hz και 64 κανάλια καταγραφής το μέγεθος του κάθε δείγματος προκύπτει 480x64. Για την ταξινόμησή τους έχουν μελετηθεί τρεις περιπτώσεις ως πριν τον αριθμό των κλάσεων :

- 2-class : Δεξί χέρι / Αριστερό χέρι
- 3-class : Δεξί χέρι / Αριστερό χέρι / Ηρεμία
- 4-class : Δεξί χέρι / Αριστερό χέρι / Ηρεμία / Πόδια.

Ακόμα μελετήθηκε η επίδραση που έχει στα αποτελέσματα των μοντέλων η μείωση του παραθύρου από 3 δευτερόλεπτα σε 2 και 1, καθώς επίσης και η χρήση λιγότερων καναλιών από 64. Συγκεκριμένα τα διαφορετικά σχήματα στον αριθμό των καναλιών περιλαμβάνουν 5 περιπτώσεις:

- 64 κανάλια. (Η αρχική διαρρύθμιση κατά το πρότυπο 10-10)
- 38 κανάλια. (Ενδιάμεση διαρρύθμιση μεταξύ του προτύπου 10-10 και 10-20 με κάλυψη όλου του εγκεφάλου με ίσες αποστάσεις)
- 19 κανάλια. (Διαρρύθμιση κατά το πρότυπο 10-20 χωρίς τα κανάλια A1 και A2)
- 8 κανάλια. (Διαρρύθμιση βασισμένη στο εμπορικό headset της BitBrain)
- 3 κανάλια. (Κανάλια C3, Cz, C4, τα οποία πρέπει να εμφανίζουν την μεγαλύτερη δραστηριότητα κατά την φαντασική κίνηση των χεριών και των ποδιών.)

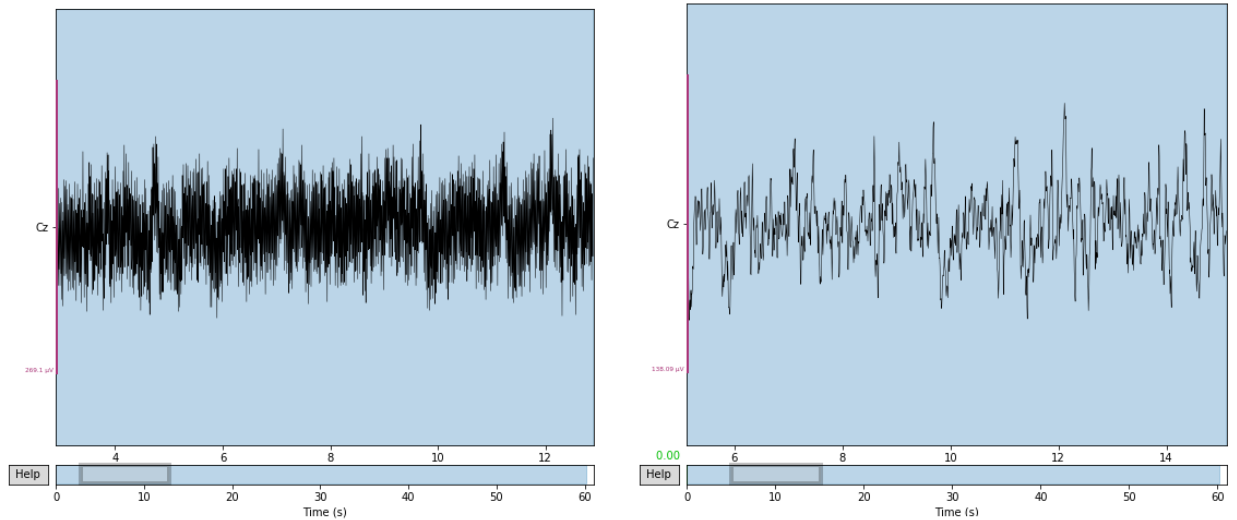
Για την αξιολόγηση των αποτελεσμάτων έχει γίνει διαχωρισμός της καθολικής ορθότητας μεταξύ όλων των υποκειμένων και της ορθότητας για συγκεκριμένο άτομο. Η καθολική ορθότητα υπολογίζεται με 5fold cross validation(5-fold CV), δηλαδή τα μοντέλα εκπαιδεύονται σε 84 άτομα και αξιολογούνται στα υπόλοιπα 21. Προκύπτει ως ο μέσος όρος των 5 αξιολογήσεων του CV. Το κάθε μοντέλο στη συνέχεια προσαρμόζεται στο κάθε υποκείμενο ξεχωριστά συνεχίζοντας την εκπαίδευση με χρήση 4-fold CV στα δεδομένα του συγκεκριμένου ατόμου για 10 ακόμα epochs και υπολογίζεται ο μέσος όρος της επίδοσης στο κάθε άτομο ξεχωριστά (Subject Specific Transfer Learning SS-TL).

Στο σημείο αυτό πρέπει να σημειωθεί ότι ο σχεδιασμός της πειραματικής διαδικασίας που περιγράφηκε παραπάνω ακολουθεί την δομή του πειράματος των Wang et al., 2020. Οι λόγοι που επιλέχθηκε αυτό το πειραματικό σχήμα είναι δυο. Πρώτον οι παράμετροι που μελετήθηκαν στην εργασία των Wang et al. κρίνονται πολύ σημαντικοί για το ενδεχόμενο πραγματικής ανάπτυξης ενός BCI βασισμένο σε motor imagery EEG καθώς ο αριθμός των καναλιών κρίνει την απαιτούμενη υπολογιστική ισχύ του συστήματος, ενώ όσο μικρότερο είναι ένα χρονικό παράθυρο τόσο πιο μικρή είναι η καθυστέρηση του. Δεύτερον με αυτόν τον τρόπο μπορούμε να επιβεβαιώσουμε τα αποτελέσματα της προηγούμενης επιστημονικής μελέτης αλλά και να έχουμε ένα αξιόπιστο μέτρο σύγκρισης για τα δικά μας.

4.2) Υλοποίηση

Προεπεξεργασία:

Όπως έχει προαναφερθεί σκοπός της εργασίας είναι η ανάπτυξη ενός end – to – end συστήματος που μπορεί να ταξινομήσει με ακρίβεια raw eeg data χωρίς να έχει προηγηθεί εκτεταμένη εξαγωγή χαρακτηριστικών όπως απαιτούνται από τις κλασσικές μεθόδους machine learning. Έτσι η προεπεξεργασία των δεδομένων που προηγήθηκε της εκπαίδευσης ήταν ελάχιστη. Συγκεκριμένα εφαρμόστηκε στα σήματα ένα Notch Filter στα 60Hz για την απομόνωση του θορύβου από την γραμμική τροφοδοσία. Στην συνέχεια τα σήματα περάστηκαν από ένα High-Pass φίλτρο με πολύ μικρή συχνότητα αποκοπής στα 0.5Hz για την απομόνωση του baseline drift. Ακόμα έγινε κανονικοποίηση των δεδομένων αφαιρώντας τη μέση τιμή τους και διαιρώντας με την τυπική απόκλιση. Για την επεξεργασία των δεδομένων έχει γίνει χρήση της βιβλιοθήκης MNE της python (Gramfort et al., 2013) η οποία περιέχει πολλά εξειδικευμένα εργαλεία για την ανάλυση EEG, MEG και EMG.



Εικόνα 4.2: EEG σήμα από το κανάλι Cz πριν και μετά την αποθρονοβοποίηση γραμμής με Notch Filter στα 60Hz

Τέλος να σημειωθεί ότι και στα 4 δίκτυα που ακολουθούν δοκιμάστηκε και η τεχνητή επαύξηση των δεδομένων για την μείωση της υπερεκπαίδευσης. Για την επαύξηση κάθε MI δοκιμή των δεδομένων εκπαίδευσης χωρίστηκε σε περισσότερα του ενός επικαλυπτόμενα παράθυρα των 3 δευτερολέπτων. Ωστόσο δεν καταγράφηκε κάποια βελτίωση στα αποτελέσματα και για αυτό τον λόγο δεν υιοθετήθηκε τελικώς αυτή η μέθοδος.

Περιγραφή των χρησιμοποιούμενων Νευρωνικών Δικτύων

Για την διεξαγωγή των πειραμάτων μας αναπτύξαμε τέσσερα διαφορετικά νευρωνικά δίκτυα. Το πρώτο δίκτυο είναι βασισμένο στο sConv που έχουν προτείνει οι R. Schirmer et al 2017 καθώς και σε μια παραλλαγή του από τους H. Dose et al./2018, και αποτελεί ένα σχετικά απλό συνελκτικό δίκτυο με λίγες παραμέτρους το οποίο έχει φέρει καλά αποτελέσματα στην ταξινόμηση των eeg σημάτων. Το δεύτερο βασίζεται στο EEGNet που ανέπτυξαν οι Vernon J. Lawhern et al. 2018 το οποίο αποτελεί επίσης ένα σχετικά μικρό δίκτυο γενικής χρήσης για eeg προβλήματα με εξαιρετικά αποτελέσματα σε μια σειρά από eeg datasets διαφορετικού περιεχομένου και αποτελεί μια εξέλιξη του sConv. Η Τρίτη προσέγγιση μας περιλαμβάνει την μετατροπή των σημάτων σε εικόνες κατασκευάζοντας το σπεκτρογράμμά τους και χρησιμοποιώντας ένα CNN δίκτυο για την ταξινόμηση των εικόνων. Τέλος δοκιμάστηκαν η επίδοση ενός ConvLSTM δικτύου στο dataset μας. Στη συνέχεια ακολουθούν πληροφορίες σχετικά με την ανάπτυξη του κάθε δικτύου:

ShallowConvNet:

Το ShallowConvNet αποτελείται από ένα συνελκτικό στρώμα με 40 φίλτρα μεγέθους (30,1) το οποίο εκτελεί συνέλιξη στο χρονικό πεδίο, ενώ ακολουθείται από ένα δεύτερο επίπεδο συνέλιξης μεγέθους (1,Nchannels) το οποίο εκτελεί χωρική συνέλιξη. Αυτά τα δύο στρώματα λειτουργούν παρόμοια με τα πρώτα βήματα του FBCSP καθώς η χρονική συνέλιξη εξάγει χαρακτηριστικά ανάλογα των ζωνοπερατών φίλτρων του πρώτου βήματος του FBCSP ενώ το δεύτερο συνελκτικό στρώμα εξάγει χωρικά χαρακτηριστικά ανάλογα με το βήμα CSP. Κάθε συνελκτικό στρώμα ακολουθείται από BatchNormalization το οποίο κανονικοποιεί την έξοδο κάθε επιπέδου σε μηδενική μέση τιμή και μοναδιαία απόκλιση βοηθώντας στην ταχύτητα και την σταθερότητα της εκπαίδευσης. Η μη γραμμική συνάρτηση που εφαρμόζεται μετά από κάθε επίπεδο είναι η ReLU. Τέλος ακολουθεί ένα επίπεδο υποδειγματολειψίας μέσης τιμής μεγέθους (15,1) το οποίο μειώνει δραστικά την διαστατικότητα των δεδομένων μας οδηγώντας σε μεγαλύτερη ικανότητα γενίκευσης του δικτύου. Σε αντίθεση με το πρωτότυπο Shallow Conv Net των R. Schirrmeyer et al. στο οποίο μετά το average pooling ακολουθεί λογαριθμική ενεργοποίηση και με την παραλλαγή των H. Dose et al. στην οποία αντί στη θέση της λογαριθμικής συνάρτησης ενεργοποίησης υπάρχει ένα πλήρως συνδεδεμένο επίπεδο 80 νευρώνων, το δίκτυο μας καταλήγει κατευθείαν σε ένα πλήρως συνδεδεμένο επίπεδο με αριθμό νευρώνων ίσο με τον αριθμό των κλάσεων και softmax συνάρτηση ενεργοποίησης για την κατηγοριοποίηση των δειγμάτων μας. Η επιλογή αυτή σε συνδυασμό με την χρήση Dropout οδήγησε το δίκτυο μας σε μικρότερα επίπεδα υπερεκπαίδευσης και καλύτερα αποτελέσματα από τους προαναφερθέντες.

Επίπεδο	Μέγεθος	Σχήμα εξόδου
Conv2D	40 Φίλτρα (30x1)	(480,64,40)
BatchNormalization	-	(480,64,40)
ReLU	-	(480,64,40)
Conv2D	40 Φίλτρα (1x64)	(480,1,40)
BatchNormalization	-	(480,1,40)
ReLU	-	(480,1,40)
AvgPool2D	Pool_size (15x1), stride (8x1)	(59,1,40)
Dropout	0.5	(59,1,40)
Flatten	-	2360
Dense	4 Νευρώνες	4

Πίνακας 4.1: Αρχιτεκτονική δικτύου ShallowConvNet

EEGNet:

Το δίκτυο που φτιάξαμε βασισμένο στο EEGNet ακολουθεί και αυτό τη δομή μιας παραλλαγής του πρωτότυπου που προτάθηκε από τους Wang et al., 2020. Αποτελείται από ένα συνελκτικό επίπεδο που εκτελεί χρονική συνέλιξη με 8 φίλτρα και μέγεθος πυρήνα (128,1), από ένα επίπεδο χωρικής συνέλιξης με 16 φίλτρα και μέγεθος πυρήνα (1,Nchannels) και από ένα επίπεδο διαχωρίσιμης συνέλιξης (separable convolutional layer) 16 φίλτρων μεγέθους (16,1). Κάθε επίπεδο συνέλιξης ακολουθείται από επίπεδο BatchNormalization και εκθετική ενεργοποίηση ELU. Το στάδιο της διαχωρίσιμης συνέλιξης ακολουθείται από επίπεδο υποδειματολειςίας μέσης τιμής (8,1) και από ένα softmax πλήρως συνδεδεμένο επίπεδο για την κατηγοριοποίηση των δειγμάτων. Τα δυο πρώτα στάδια είναι παρόμοια με του Shallow ConvNet και επιτελούν την ίδια λειτουργία. Εμπνευσμένα από τον αλγόριθμο FBCSP η χρονική συνέλιξη εξάγει χάρτες χαρακτηριστικών του σήματος σε διαφορετικές ζωνοπερατές συχνότητες ενώ το δεύτερο επίπεδο εξάγει χωρικά χαρακτηριστικά του σήματος από τα διαφορετικά κανάλια. Το στοιχείο που έρχεται να προσθέσει αυτό το δίκτυο είναι το στάδιο της διαχωρίσιμης συνέλιξης. Το στρώμα αυτό δέχεται ως είσοδο 16 χάρτες ενεργοποίησης μεγέθους (60,1) οι οποίοι έχουν προέλθει από χωρική συνέλιξη και ως εκ τούτου μεταφέρουν χωρική πληροφορία για το σήμα. Η λειτουργία του χωρίζεται σε δυο διαδοχικά βήματα. Αρχικά διαχωρίζει τα εισαγόμενα feature maps ως προς το βάθος τους και εκτελεί στο καθένα συνέλιξη με μέγεθος πυρήνα (16,1) που αναπαριστά 800ms εγκεφαλικής δραστηριότητας στα 20Hz που έχει καταλήξει το σήμα από την υποδειματολειςία. Στη συνέχεια εκτελεί 16 (#πυρήνων) (1,1) σημειακές συνελίξεις συνδυάζοντας τις εξόδους. Επιτυγχάνεται λοιπόν αυτό το επίπεδο να εκπαιδευτεί στο να συνοψίσει τους χάρτες ενεργοποίησης στον χρόνο κατά το πρώτο βήμα και να τις συνδυάσει βέλτιστα κατά το δεύτερο εξάγοντας έτσι χωροχρονικά χαρακτηριστικά. Επιπρόσθετα η επιλογή αυτού του επιπέδου οδηγεί σε μείωση των παραμέτρων σε σχέση με ένα κανονικό συνελκτικό επίπεδο επιτρέποντας στο δίκτυο να γενικεύει καλύτερα. Στη συνέχεια ακολουθεί ένα πλήρως συνδεδεμένο επίπεδο για την ταξινόμηση των δειγμάτων. Για την αποφυγή overfitting έχει γίνει χρήση Dropout με ρυθμό 0.5 πριν και μετά το separable convolution.

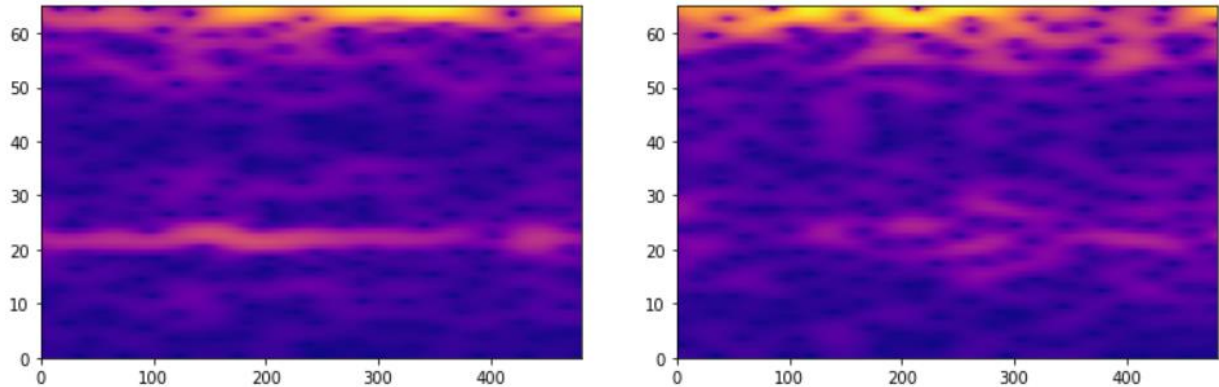
Επίπεδο	Μέγεθος	Σχήμα εξόδου
Conv2D	8 Φίλτρα (128x1)	(480,64,8)
BatchNormalization	-	(480,64,8)
Conv2D	16 Φίλτρα (1x64)	(480,1,16)
BatchNormalization	-	(480,1,16)
ELU	-	(480,1,16)
AvgPool2D	Pool_size (8x1), stride (8x1)	(60,1,16)
Dropout	0.5	(60,1,16)
SeparableConv2D	16 Φίλτρα (16x1)	(60,1,16)
BatchNormalization	-	(60,1,16)
ELU	-	(60,1,16)
AvgPool2D	Pool_size (8x1), stride (1x8)	(7,1,16)
Dropout	0.5	(7,1,16)
Flatten	-	120
Dense	4 Νευρώνες	4

Πίνακας 4.2: Αρχιτεκτονική δικτύου EEGNet

Spectrogram Conv

Τα συνελκτικά δίκτυα αν και έχουν φέρει πολύ καλά αποτελέσματα σε διάφορα είδη προβλημάτων, γνώρισαν μεγάλη ανάπτυξη και δημοφιλία τα τελευταία χρόνια λόγω των πολλά υποσχόμενων αποτελεσμάτων τους στην όραση υπολογιστών. Η επιτυχία τους αυτή στην επεξεργασία εικόνων οδήγησε σε μια ενδιαφέρουσα προσέγγιση σε προβλήματα αναγνώρισης φωνής κατά την οποία το ηχητικό σήμα μετατρέπεται πρώτα σε μορφή εικόνας μέσω του μετασχηματισμού short time fourier transform (STFT) και στη συνέχεια χρησιμοποιούνται συνελκτικά νευρωνικά δίκτυα για την κατηγοριοποίηση των εικόνων. Ωστόσο αυτή η προσέγγιση μπορεί να γενικευθεί και σε άλλου είδους δεδομένων χρονοσειρών. Ιδιαίτερα στην περίπτωση των EEG σημάτων όπου η μεταβολή της ενέργειας

διαφορετικών συχνοτήτων στην πάροδο του χρόνου αποτελεί εξαιρετικά σημαντική πληροφορία τα σπεκτρογράμματα μπορούν να δώσουν πολύτιμα time-frequency χαρακτηριστικά. Μάλιστα οι Z. Tayeb et al 2019 προτείνουν αυτή την μέθοδο ως την πλέον αποδοτική για MI classification.



Εικόνα 4.3: Σπεκτρογράμματα νοητικής κίνησης αριστερής και δεξιάς γροθιάς

Για τον υπολογισμό του σπεκτρογράμματος έγινε χρήση της βιβλιοθήκης Kapre (enter reference) η οποία παρέχει keras layers για τον υπολογισμό STFT καθώς και διάφορους χρήσιμους μετασχηματισμούς. Το δίκτυο αποτελείται από τα επίπεδα STFT, Magnitude και MagnitudeToDecibel για την μετατροπή των σημάτων σε spectrograms και ακολουθούν 4 συνελκτικά επίπεδα με μεγέθη όπως φαίνονται στο παρακάτω διάγραμμα. Κάθε επίπεδο ακολουθείται από BatchNormalization και Relu συνάρτηση ενεργοποίησης, ενώ ανά δυο επίπεδα συνέλιξης γίνεται υποδειγματολειψία μεγίστου για μείωση διαστατικότητας. Για την αποφυγή υπερεκπαίδευσης έχει γίνει και σε αυτή τη περίπτωση χρήση Dropout με ρυθμό 0.4 έπειτα από κάθε συνέλιξη.

Επίπεδο	Μέγεθος	Σχήμα εξόδου
STFT	n_fft=64, win=64, hop=2	(209, 33, 64)
Magnitude	-	(209, 33, 64)
Magnitude_to_decibel	-	(209, 33, 64)
Conv2D	32 Φίλτρα (3x3)	(209, 33, 32)
BatchNormalization	-	(209, 33, 32)
ReLU	-	(209, 33, 32)
Dropout	0.1	(209, 33, 32)
MaxPool2D	(2x2)	(104, 16, 32)
Conv2D	32 Φίλτρα (3x3)	(104, 16, 32)
BatchNormalization	-	(104, 16, 32)
ReLU	-	(104, 16, 32)
MaxPool2D	(2x2)	(52,8,32)
Dropout	0.1	(52,8,32)
Conv2D	32 Φίλτρα (3x3)	(52,8,32)
BatchNormalization	-	(52,8,32)
ReLU	-	(52,8,32)
Dropout	0.1	(52,8,32)
MaxPool2D	(2x2)	(26, 4, 32)
Conv2D	32 Φίλτρα (3x3)	(26, 4, 32)
BatchNormalization	-	(26, 4, 32)
ReLU	-	(26, 4, 32)
GlobalMaxPooling2D		32
Dense	4	4

Πίνακας 4.3: Αρχιτεκτονική δικτύου SpectrogramConv

ConvLSTM Network:

Τα convolutional LSTM δίκτυα όπως έχει προαναφερθεί είναι εξειδικευμένα για την ανάλυση αλληλουχιών εικόνων, όπως το βίντεο. Η ιδέα πίσω από την χρήση convolutional LSTM για την ταξινόμηση των EEG είναι ότι θα μπορούσαμε να δούμε τα σήματα που καταγράφονται από τα διάφορα ηλεκτρόδια ως ένα «βίντεο» όπου κάθε καρέ είναι οι τιμές των καναλιών σε κάθε χρονική στιγμή. Για το μοντέλο αυτό τα δείγματα από μέγεθος (480,64) μετατρέπονται σε (480,64,1,1), όπου τα μεγέθη αυτά αντιστοιχούν στα φυσικά μεγέθη (αριθμός πλαισίων,

πλάτος, ύψος, κανάλια). Γίνεται εμφανές ότι τα δεδομένα μας έχουν πάρει την μορφή μιας ακολουθίας 480 εικόνων 64x1x1 pixel η καθεμία. Το δίκτυο απαρτίζεται από 3 επίπεδα ConvLSTM με 10 φίλτρα το καθένα μεγέθους (9,1) ακολουθούμενα από BatchNormalization, επίπεδο υποδειγματοληψίας μεγίστου μεγέθους (2,1,1) και Dropout με ρυθμό 0.5 για αποφυγή υπερεκπαίδευσης. Το επίπεδο υποδειγματοληψίας χρησιμοποιείται για να μειώσει τις διαστάσεις στο πεδίο του χρόνου καθώς επίσης να κρατήσει πιο διακριτά χρονικά χαρακτηριστικά για τα επόμενα στρώματα.

Επίπεδο	Μέγεθος	Σχήμα εξόδου
ConvLSTM2D	10 Φίλτρα (9x1)	(480, 64, 1, 10)
BatchNormalization	-	(480, 64, 1, 10)
MaxPooling3D	(2,1,1)	(240, 64, 1, 10)
Dropout	0.5	(240, 64, 1, 10)
ConvLSTM2D	10 Φίλτρα (9x1)	(240, 64, 1, 10)
BatchNormalization	-	(240, 64, 1, 10)
MaxPooling3D	(2,1,1)	(120, 64, 1, 10)
Dropout	0.5	(120, 64, 1, 10)
ConvLSTM2D	10 Φίλτρα (9x1)	(120, 64, 1, 10)
BatchNormalization	-	(120, 64, 1, 10)
MaxPooling3D	(2,1,1)	(60, 64, 1, 10)
Dropout	0.5	(60, 64, 1, 10)
Flatten	-	38400
Dense	4 Νευρώνες	4

Πίνακας 4.4: Αρχιτεκτονική δικτύου ConvLSTM

Κεφάλαιο 5: Αποτελέσματα και Αξιολόγηση

Σε αυτό το τμήμα της εργασίας παρατίθενται τα αποτελέσματα της αξιολόγησης των παραπάνω μοντέλων. Τα αποτελέσματα που αναφέρονται ως standard αναφέρονται σε μετρήσεις με μήκος παραθύρου 3s και αριθμό καναλιών 64, δηλαδή το σύνολο της πληροφορίας. Το ακρωνύμιο SS-TL (Subject Specific Transfer Learning) αναφέρεται στη διαδικασία περαιτέρω εκπαίδευσης του δικτύου σε κάθε υποκείμενο που ανήκει στο test set ξεχωριστά ώστε να μελετηθεί κατά πόσο μπορούμε να επιτύχουμε σημαντικά βελτιωμένα αποτελέσματα αν μετά την γενική εκπαίδευση του δικτύου ακολουθήσει μια εξατομικευμένη εκπαίδευση. Οι μετρήσεις αυτές έγιναν επίσης με το σύνολο της πληροφορίας. Η μελέτη αυτή έχει πρακτική σημασία καθώς υψηλά αποτελέσματα στο SS – TL υποδεικνύουν την δυνατότητα ανάπτυξης ενός συστήματος το οποίο μπορεί πιο εύκολα και σύντομα να προσαρμοστεί στο κάθε άτομο με δεδομένο ότι έχει προηγηθεί γενική εκπαίδευση σε πλήθος καταγραφών.

EEGNet

Κλάσεις	Κανάλια						Παράθυρο	
	Standard	SS-TL	38ch	19ch	8ch	3ch	2s	1s
2 κλάσεις	80.2%	83.2%	80.1%	79.3%	75.7%	71.9%	80.4%	79%
3 κλάσεις	80.5%	84.6%	82.1%	81%	77.1%	71.9%	81.4%	75.8%
4 κλάσεις	68.5%	71.5%	69.9%	70.1%	63%	58%	68.5%	72.2%

Πίνακας 5.1: Πίνακας αποτελεσμάτων EEGNet

Το πρώτο μοντέλο που δοκιμάστηκε είναι το EEGNet το οποίο όπως είναι προφανές είναι αρκετά αποτελεσματικό στην αποκωδικοποίηση των MI κινήσεων. Συγκρίνοντάς το με τα αποτελέσματα των (Wang et al., 2020) οι οποίοι πέτυχαν 82.43%, 75.07% και 65.07% για 2,3 και 4 κλάσεις αντίστοιχα το EEGNet κατέγραψε στην παρούσα εργασία ελαφρώς καλύτερα αποτελέσματα. Αν και στο πρόβλημα των 2 κλάσεων η ορθότητα ήταν κατά 2 ποσοστιαίες μονάδες μικρότερη, στις 3 και στις 4 κλάσεις παρατηρείται +5.5% και +3.5% ορθότητα. Ιδίως στην μετάβαση από 2 κλάσεις σε 3 αντί για μείωση όπως θα ανέμενε κανείς είχαμε μια μικρή αύξηση στην ορθότητα των αποτελεσμάτων. Ωστόσο αυτό εξηγείται από την ικανότητα του δικτύου να ξεχωρίζει πολύ εύκολα την κατάσταση ηρεμίας από τις

υπόλοιπες κλάσεις με ποσοστό άνω του 90%. Ενδιαφέρον έχει το γεγονός ότι η επίδοση του δικτύου πέφτει ελάχιστα μειώνοντας κανάλια μέχρι και σε 19. Αυτό πρακτικά σημαίνει ότι όλη σχεδόν η απαραίτητη πληροφορία για την διάκριση των κινήσεων υπάρχει ενδεχομένως στο σύνολο των 19 ηλεκτροδίων και τα υπόλοιπα παρέχουν κατά βάση περιττή πληροφορία. Ωστόσο στα 8 και στα 3 κανάλια παρατηρείται αισθητή μείωση της επίδοσης. Ακόμα όσον αφορά το μέγεθος του χρονικού παραθύρου η ορθότητα του μοντέλου παραμένει σταθερή στα 2 δευτερόλεπτα ωστόσο πέφτει αισθητά όταν γίνει περαιτέρω μείωση σε 1 δευτερόλεπτο. Με την μεταφορά εκπαίδευσης σε κάθε άτομο ξεχωριστά παρατηρείται μια αύξηση της καταγραφόμενης ορθότητας. Η αύξηση αυτή ήταν της τάξης του 3-4% πολύ κοντά στο αποτέλεσμα των (Wang et al., 2020)

ShallowConvNet

Κλάσεις	Κανάλια						Παράθυρο	
	Standard	SS-TL	38ch	19ch	8ch	3ch	2s	1s
2 κλάσεις	80 %	87.5 %	79.7%	77.4%	74.9%	70.3%	78.9%	77.6%
3 κλάσεις	79.6%	85.8%	80.2%	78.8%	75.1%	69.1%	81.1%	78.9%
4 κλάσεις	70%	82.3%	70.5%	69.4%	64.6%	54.6%	71%	68%

Πίνακας 5.2: Πίνακας αποτελεσμάτων ShallowConvNet

Όπως έχει προαναφερθεί το ShallowConvNet που έχει υλοποιηθεί στα πλαίσια της εργασίας έχει βασιστεί στο μοντέλο CNN των (Dose et al., 2018). Αν και στην έρευνά τους αναφέρεται ότι το δίκτυο δεν πάσχει από overfitting, στην παρούσα παραλλαγή, αν και έχει αφαιρεθεί το πλήρως συνδεδεμένο μέρος του δικτύου με 80 νευρώνες και έχει προστεθεί επίπεδο Dropout, εμφανίζει το πρόβλημα της υπερεκπαίδευσης. Παρ'όλα αυτά τα αποτελέσματα που καταγράφηκαν ήταν σαφώς βελτιωμένα των Dose et al. 2018 (80.38%, 76.61% και 65.73%), και πιθανότατα η διαφορά αυτή οφείλεται στις παραπάνω αλλαγές. Συγκρίνοντας το με το EEGNet, αν και εμφανίζουν παρόμοια απόδοση στις standard μετρήσεις καθώς και στην μεταβολή των παραθύρων, το ShallowConvNet φαίνεται να είναι πιο ευαίσθητο στην μείωση του αριθμού των καναλιών. Ωστόσο εντυπωσιακά είναι τα αποτελέσματα του στην προσαρμογή σε κάθε υποκείμενο. Για το SS-TL τόσο συγκριτικά με το μοντέλο των Dose et al. 2018 (86.49%, 79.25%, 68.51%) όσο και με την δική μας υλοποίηση του EEGNet, το ShallowConvNet εμφανίζει πολύ μεγαλύτερη ορθότητα στις προβλέψεις του, ιδίως για την περίπτωση των 4 κλάσεων η οποία αποτελεί την μεγαλύτερη πρόκληση και στην οποία πετυχαίνει 82.3% ορθότητα.

SpectrogramConv

Κλάσεις			Κανάλια				Παράθυρο	
	Standard	SS-TL	38ch	19ch	8ch	3ch	2s	1s
2 κλάσεις	70 %	79.6 %	69.1%	69.2%	62.3%	57%	68.9%	64.7%
3 κλάσεις	65.8%	72.7%	64.7%	65.1%	60.4%	55.7%	67.9%	62.9%
4 κλάσεις	50.4%	52.2%	52.8%	49.1%	45.3%	41.5%	52.3%	49.1%

Πίνακας 5.3: Πίνακας αποτελεσμάτων SpectrogramConv

Είναι εμφανές ότι τα αποτελέσματα του SpectrogramConv είναι απογοητευτικά. Παρότι οι (Tayeb et al., 2019) στην εργασία τους παρουσίασαν την μέθοδο αυτή ως την προτιμότερη για προβλήματα MI, στην περίπτωσή μας υστερεί πολύ σε σχέση με τα υπόλοιπα μοντέλα. Ωστόσο πρέπει να σημειωθεί ότι ο αλγόριθμος STFT επιδέχεται πολλών διαφορετικών ρυθμίσεων στις παραμέτρους του και επομένως μπορεί να υπάρχει ένας βέλτιστος συνδυασμός που οδηγεί σε καλύτερα αποτελέσματα και στην περίπτωση του παρόντος dataset.

ConvLSTM

Κλάσεις			Κανάλια				Παράθυρο	
	Standard	SS-TL	38ch	19ch	8ch	3ch	2s	1s
2 κλάσεις	79 %	90.4 %	78.7%	78%	75.8%	71.9	77.5%	77.8%
3 κλάσεις	83.4%	91.1%	83.5%	82.5%	79.7%	75.7%	81.7%	82.8%
4 κλάσεις	70.1%	88.3%	69.8%	67.2%	65.8%	59%	69.4%	68.7%

Πίνακας 5.4: Πίνακας αποτελεσμάτων ConvLSTM

Το ConvLSTM μοντέλο έχει πετύχει καθολική ορθότητα αντίστοιχη των EEGNet και ShallowConvNet στις 2 και στις 4 κλάσεις, ενώ φαίνεται να έχει καλύτερα αποτελέσματα στην περίπτωση των 3 κλάσεων, γεγονός που οφείλεται στο ότι μπορεί να διαχωρίσει καλύτερα την κατάσταση ηρεμίας. Το συγκεκριμένο δίκτυο αντιλαμβάνεται τα δεδομένα ως μια αλληλουχία «εικόνων» όπου τα κανάλια αποτελούν τα εικονοστοιχεία. Επομένως θα περιμέναμε η επιλογή λιγότερων καναλιών να επιφέρει μεγαλύτερη μείωση στην απόδοση του δικτύου. Ωστόσο τα αποτελέσματα δείχνουν ότι και με λιγότερα κανάλια η απόδοση παραμένει αντίστοιχη των προηγούμενων δικτύων. Δεν ισχύει όμως το ίδιο και στην περίπτωση της SS-TL ορθότητας. Το ConvLSTM έχει πετύχει πολύ υψηλότερα επίπεδα ορθότητας από κάθε άλλο μοντέλο που είτε έχει υλοποιηθεί στα πλαίσια της εργασίας είτε έχουμε διαβάσει για αυτό κατά την βιβλιογραφική έρευνα. Η ικανότητα του να προσαρμόζεται σε κάθε υποκείμενο του συνόλου αξιολόγησης ξεχωριστά, μετά από την γενική του εκπαίδευση στα δεδομένα εκπαίδευσης, είναι εξαιρετική αγγίζοντας τα ποσοστά ορθότητας 90.4% και 91.1% για τις 2 και 3 κλάσεις αντίστοιχα, καθώς και 88.3% για την περίπτωση των τεσσάρων κλάσεων. Πρέπει να σημειωθεί ότι για την μέτρηση με 19 και 8 κανάλια το μέγεθος των φίλτρων μειώθηκε σε 6x1 ενώ για την περίπτωση των 3 καναλιών μειώθηκε σε 3x1.

Στον παρακάτω πίνακα γίνεται η σύγκριση μεταξύ των αποτελεσμάτων των μοντέλων μας καθώς και των αποτελεσμάτων της βιβλιογραφίας που αφορούν το ίδιο dataset:

	2 κλάσεις		3 κλάσεις		4 κλάσεις	
	Standard	SS-TL	Standard	SS-TL	Standard	SS-TL
EEGNet	80.2%	83.2%	80.5%	84.6%	68.5%	71.5%
ShallowConvNet	80 %	87.5 %	79.6%	85.8%	70%	82.3%
SpectrogramConv	70 %	79.6 %	65.8%	72.7%	50.4%	52.2%
ConvLSTM	79 %	90.4 %	83.4%	91.1%	70.1%	88.3%
SUTCSP-Random Forest (Kim et al., 2016)	-	80.05%	-	-	-	-
FBCSP-SVM (Handiru & Prasad, 2016)	63.62%	-	-	-	-	-
EEGNet (Wang et al., 2020)	82.43%	84.32%	75.07%	80.07%	65.07%	70.83%
sCNN (Dose et al., 2018)	80.38%	86.49%	69.82%	79.25%	58.59%	68.51%

Πίνακας 5.5: Συγκριτικά αποτελέσματα γνωστών μοντέλων για το dataset Physionet EEGMI.

Παρατηρούμε ότι τα βαθιά νευρωνικά δίκτυα έχουν πετύχει σαφώς καλύτερα αποτελέσματα σε σχέση με τις κλασσικές τεχνικές μηχανικής μάθησης που χρησιμοποιούνταν και χρησιμοποιούνται ακόμα και σήμερα για την αποκωδικοποίηση των EEG σημάτων. Επιπρόσθετα από την σύγκριση προκύπτει η υπεροχή των ShallowConvNet και ConvLSTM έναντι των υπολοίπων, ιδίως αν λάβουμε υπόψιν την προσαρμογή κατ'άτομο.

Κεφάλαιο 6: Συμπεράσματα και Μελλοντική Εργασία

6.1) Σύνοψη και συμπεράσματα

Αντικείμενο της παρούσας είναι η ταξινόμηση σημάτων ηλεκτροεγκεφαλογραφημάτων από φαντασιακές κινήσεις, κάνοντας χρήση βαθιών νευρωνικών δικτύων και μελετώντας την απόδοσή τους σε σχέση με τις κλασικές μεθόδους ρηχής μάθησης. Απώτερο σκοπό αποτελεί η ανάπτυξη μιας διεπαφής Εγκεφάλου - Υπολογιστή που θα αποκωδικοποιεί την εγκεφαλική δραστηριότητα δίνοντας έτσι την δυνατότητα κυρίως σε άτομα με αναπηρία, να αλληλεπιδρούν με τον κόσμο μέσω εναλλακτικών οδών. Υλοποιήθηκαν 4 μοντέλα βαθιάς μάθησης εκ των οποίων τα τρία από αυτά εμφανίζουν σαφή υπεροχή σε σχέση με τις state of the art τεχνικές αποκωδικοποίησης των EEG με χειροκίνητη εξαγωγή χαρακτηριστικών. Απαντήθηκε έτσι το βασικό ερώτημα της εργασίας, δηλαδή αν η βαθιά μάθηση μπορεί να συνεισφέρει στον τομέα των EEG BCIs, αντικαθιστώντας τις συμβατικές μεθόδους.

Επιπρόσθετα τα μοντέλα δοκιμάστηκαν σε διάφορες διαρρυθμίσεις όσον αφορά τον αριθμό των κλάσεων, τον αριθμό των χρησιμοποιούμενων καναλιών καταγραφής και το μέγεθος του χρονικού παραθύρου. Από τα αποτελέσματα μπορούμε να συμπαιράνουμε ότι ο αριθμός των καναλιών και το χρονικό παράθυρο μπορούν να μειωθούν μέχρι και σε 19ch και 2s αντίστοιχα, χωρίς κάποια σημαντική μείωση στην ποιότητα των αποτελεσμάτων. Αυτή η παρατήρηση έχει μεγάλη αξία και είναι πολύ ενθαρρυντική καθώς για την υλοποίηση ενός πραγματικού BCI, πέρα από την ακρίβεια του συστήματος, έχουν σημασία και τεχνικοί παράμετροι όπως η υπολογιστική πολυπλοκότητα που καθορίζει την κατανάλωση ενέργειας όπως επίσης και το χρόνο απόκρισης.

Τέλος εξετάστηκε η ικανότητα των δικτύων να προσαρμοστούν σε μεμονωμένα άτομα. Όλα τα δίκτυα πέρα του SpectrogramConv ήταν ικανά να προβλέψουν τα δείγματα των υποκειμένων αξιολόγησης έχοντας εκπαιδευτεί σε άλλα άτομα, με μεγάλη ορθότητα, αντίστοιχη των πιο υψηλών καταγεγραμμένων αποτελεσμάτων από συγγενείς εργασίες. Παρά ταύτα όσον αφορά την μεταφορά της εκπαίδευσης σε μεμονωμένα άτομα το ShallowConvNet και το ConvLSTM κατάφεραν να ξεπεράσουν κατά πολύ τις επιδόσεις των καλύτερων μοντέλων της βιβλιογραφίας για αυτό το dataset, ιδίως με το τελευταίο να πετυχαίνει 91.4% και 88.3% για την περίπτωση των 3 και 4 κλάσεων αντίστοιχα.

Αν και το ConvLSTM υπερέρχει το ShallowConvNet ως προς την ορθότητα της ταξινόμησης, το δεύτερο υπερτερεί ως προς την ταχύτητα πρόβλεψης ενός δείγματος. Συγκεκριμένα το ShallowConvNet απαιτεί χρόνο 5ms για την ταξινόμηση ενός δείγματος 4 κλάσεων, ενώ αντίστοιχα το ConvLSTM απαιτεί 7ms. Η διαφορά μοιάζει μικρή σε απόλυτη τιμή ωστόσο η

σχετική αύξηση είναι της τάξεως του 40%. Η αύξηση αυτή του χρόνου ταξινόμησης των δεδομένων ενδεχομένως μπορεί να καταστήσει το ConvLSTM λιγότερο χρήσιμο σε ένα σύγχρονο (online) σύστημα, το οποίο κάνει διαρκή ταξινόμηση των σημάτων του φορέα, καθώς θα εισάγει μεγαλύτερη καθυστέρηση στην απόκριση του συστήματος. Επομένως πριν βγει συμπέρασμα για το ποιο από τα δύο δίκτυα είναι προτιμότερο για μια πραγματική εφαρμογή, θα πρέπει να γίνει μελέτη για τις επιδόσεις τους σε online περιβάλλον.

6.2) Μελλοντικές κατευθύνσεις

Παρότι στα πλαίσια της εργασίας απαντήθηκε το ερώτημα για την αποδοτικότητα των βαθιών νευρωνικών δικτύων σε σχέση με τις τυπικές μεθόδους ανάλυσης των ηλεκτροεγκεφαλογραφημάτων φαντασιακών κινήσεων δείχνοντας την σαφή υπεροχή τους, μια μελέτη στον τομέα των BCI αποσκοπεί πάντα στην μελλοντική πρακτική υλοποίηση της εφαρμογής. Για τον σκοπό αυτό, υπάρχει μεγάλος χώρος για περαιτέρω έρευνες και επέκταση των παραπάνω αποτελεσμάτων. Οι έρευνες αυτές θα μπορούσαν να επικεντρωθούν σε διάφορους τομείς ωστόσο παρακάτω θα αναφέρουμε τους πιο σημαντικούς:

Δεδομένα

Ένα μεγάλο πρόβλημα που αποτελεί τροχοπέδη στην ανάπτυξη των BCI είναι η δυσκολία στην συλλογή αρκετών δεδομένων και η έλλειψη διαθεσιμότητας αρκετά μεγάλων dataset. Βασική μελλοντική εργασία για ερευνητικές ομάδες θα πρέπει να είναι είτε η συλλογή επιπρόσθετων EEG δεδομένων. Εκτός όμως από την ποσότητα των δεδομένων καθοριστικό ρόλο παίζει και η ποιότητα. Έτσι λοιπόν είναι σκόπιμο να γίνουν μελέτες με πιο εξελιγμένα συστήματα καταγραφής EEG, ή ακόμα και με ενδοκρανιακά ηλεκτρόδια (ECoG) τα οποία παρέχουν πολύ υψηλότερης ποιότητας ηλεκτροεγκεφαλικά σήματα σε σχέση με τις καταγραφές του BCI2000 (Schalk et al., 2004) που είχαμε στην διάθεσή μας κατά την διεξαγωγή της εργασίας.

Μελέτη Περισσότερων Κινήσεων και Καταστάσεων

Με τα δεδομένα που είχαμε διαθέσιμα, υλοποιήσαμε μοντέλα ταξινόμησης 3 διαφορετικών νοερών κινήσεων (δεξί / αριστερό χέρι, πόδια) καθώς και της κατάστασης ηρεμίας. Ωστόσο σε ένα πραγματικό σύστημα για να είναι χρηστικό θα πρέπει να γίνει μελέτη για την αναγνώριση περισσότερων κινήσεων καθώς και για την συνεχή κίνηση των άκρων στον τρισδιάστατο χώρο. Ακόμα χρειάζεται να γίνει περισσότερη έρευνα σχετικά με την ευαισθησία αυτών των EEG σημάτων στην κατάσταση του ατόμου(ψυχολογική, στάση σώματος κ.α) καθώς και στο εξωτερικό περιβάλλον (π.χ επίδραση του φωτός), με σκοπό την

κατά το δυνατόν απομόνωση του θορύβου που δημιουργούν. Φυσικά για το σκοπό αυτό είναι απαραίτητη προϋπόθεση η συλλογή επαρκών δεδομένων όπως αναφέρθηκε παραπάνω καθώς τα μέχρι τώρα πειράματα έχουν γίνει σε εξαιρετικά ελεγχόμενες συνθήκες και μικρή ποικιλία διαφορετικών κινήσεων. Άλλωστε δεν υπάρχει λόγος να αμφισβητηθεί η δυνατότητα των βαθιών νευρωνικών δικτύων να μάθουν απομονώνουν μόνα τους τους θορύβους αυτούς, αρκεί να υπάρχει ο απαραίτητος όγκος δεδομένων.

Online Αξιολόγηση

Ένα από τα επόμενα βήματα για την επέκταση των μοντέλων που αναφέρθηκαν παραπάνω είναι η αξιολόγηση τους σε online συνθήκες. Κατά την διαρκή ταξινόμηση των καταγραφόμενων σημάτων EEG οι επιδόσεις των παραπάνω μοντέλων αναμένεται να είναι διαφορετικές. Στα πλαίσια της εργασίας έχει γίνει μέτρηση της επίδρασης διαφορετικών χρονικών παραθύρων. Τα αποτελέσματα έδειξαν όπως ήταν αναμενόμενο ότι όσο μικραίνει το παράθυρο από ένα σημείο και μετά μειώνεται και η επίδοση του ταξινομητή. Ωστόσο όσο μικρότερο το παράθυρο τόσο προτιμότερο για μια εφαρμογή πραγματικού χρόνου. Έτσι λοιπόν δημιουργείται η ανάγκη για εύρεση της βέλτιστης ισορροπίας σε αυτό το tradeoff. Ακόμα η ταξινόμηση που έγινε στην εργασία αφορούσε χρονικά παράθυρα που ξεκινούσαν πάντα από το σημείο έναρξης της MI δοκιμής $t=0$. Φυσικά σε μια συνεχή καταγραφή αυτό δεν συμβαίνει, γεγονός που καθιστά αρκετά πιθανή την μείωση της επίδοσης. Τέλος για μια εμπειρισταωμένη αξιολόγηση των μοντέλων σε online συνθήκες θα πρέπει να γίνουν μετρήσεις και για τον χρόνο απόκρισης τους, όπου θα παίζει σημαντικό ρόλο η υπολογιστική πολυπλοκότητα τους.

Υλοποίηση τεχνητού μέλους – εξωσκελετού

Τελικός στόχος μιας εφαρμογής BCI βασισμένη σε EEG MI αποτελεί ο έλεγχος ενός εξωτερικού συστήματος με την σκέψη. Εφόσον πραγματευόμαστε δεδομένα από την νοητική κίνηση των άκρων, το εξωτερικό σύστημα που έχει το περισσότερο νόημα είναι ένα προσθετικό μέλος ή ένας εξωσκελετός που θα μπορεί να λάβει ως σήμα την πρόβλεψη του συστήματος απόφασης και να εκτελέσει την αντίστοιχη κίνηση, χαρίζοντας στον χρήστη του την δυνατότητα να υποκαταστήσει ως έναν βαθμό την ενδεχόμενη χαμένη κίνηση ή και ύπαρξη των βιολογικών του άκρων.

Βελτίωση απόδοσης των Νευρωνικών Δικτύων

Τα δίκτυα που υλοποιήθηκαν στην παρούσα εργασία αν και φέρουν πολύ καλά αποτελέσματα σε σχέση με συγγενείς εργασίες, έχουν φυσικά μεγάλο περιθώριο βελτίωσης. Ακόμα και στην περίπτωση του subject specific transfer learning όπου επιτύχαμε 88.1% ορθότητα στις 4 κλάσεις, πρακτικά παραπάνω από μια στις 10 νοερές κινήσεις που θα σκεφτεί ένα άτομο, θα ταξινομηθεί εσφαλμένα και το εν δυνάμει μηχανικό σύστημα θα εκτελέσει μια λανθασμένη κίνηση. Είναι προφανές ότι μια τέτοια συμπεριφορά σε ένα πραγματικό σύστημα δεν μπορεί να είναι αποδεκτή, επομένως είναι αναγκαία η βελτίωση της απόδοσης των μοντέλων ταξινόμησης. Βάσει των αποτελεσμάτων μας αξίζει να διερευνηθεί περαιτέρω η χρήση αναδρομικών συνελκτικών δικτύων για το πρόβλημα της ταξινόμησης EEG σημάτων. Μέχρι σήμερα δεν έχει γίνει κάποια εκτεταμένη μελέτη για αυτόν τον τύπο νευρωνικών δικτύων ενώ θα μπορούσε να οδηγήσει στην βελτίωση των συστημάτων απόφασης.

Βιβλιογραφία

- Goldberger, A., L. Amaral, L. Glass, J. Hausdorff, P. C. Ivanov, R. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley. "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* [Online]. 101 (23), pp. e215–e220." (2000).
- Baig, M. Z., Aslam, N., Shum, H. P. H., & Zhang, L. (2017). Differential evolution algorithm as a tool for optimal feature subset selection in motor imagery EEG. *Expert Systems with Applications*, 90, 184–195. <https://doi.org/10.1016/j.eswa.2017.07.033>
- Blankertz, B., Muller, K.-., Krusienski, D. J., Schalk, G., Wolpaw, J. R., Schlogl, A., Pfurtscheller, G., Millan, J. R., Schroder, M., & Birbaumer, N. (2006). The BCI competition III: validating alternative approaches to actual BCI problems. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2), 153–159. <https://doi.org/10.1109/TNSRE.2006.875642>
- Caldwell, J. A., Prazinko, B., & Caldwell, J. L. (2003). Body posture affects electroencephalographic activity and psychomotor vigilance task performance in sleep-deprived subjects. *Clinical Neurophysiology*, 114(1), 23–31. [https://doi.org/10.1016/S1388-2457\(02\)00283-3](https://doi.org/10.1016/S1388-2457(02)00283-3)
- Carlson, T., & Del R. Millan, J. (2013). Brain-controlled wheelchairs: A robotic architecture. *IEEE Robotics and Automation Magazine*, 20(1), 65–73. <https://doi.org/10.1109/MRA.2012.2229936>
- Chambayil, B., Singla, R., & Jha, R. (2010). Virtual keyboard BCI using Eye blinks in EEG. *2010 IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications*, 466–470. <https://doi.org/10.1109/WIMOB.2010.5645025>
- Dose, H., Møller, J. S., Iversen, H. K., & Puthusserypady, S. (2018). An end-to-end deep learning approach to MI-EEG signal classification for BCIs. *Expert Systems with Applications*, 114, 532–542. <https://doi.org/10.1016/j.eswa.2018.08.031>
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179–211. [https://doi.org/10.1016/0364-0213\(90\)90002-E](https://doi.org/10.1016/0364-0213(90)90002-E)
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7 DEC. <https://doi.org/10.3389/fnins.2013.00267>
- Greaves, A. S. (2011). *Classification of EEG with Recurrent Neural Networks*. 1–5. <internal-pdf://179.90.41.114/GreavesAlex.pdf>
- Handiru, V. S., & Prasad, V. A. (2016). Optimized Bi-Objective EEG Channel Selection and Cross-Subject Generalization with Brain-Computer Interfaces. *IEEE Transactions on Human-Machine Systems*, 46(6), 777–786. <https://doi.org/10.1109/THMS.2016.2573827>
- Hema, C. R., Paulraj, M. P., Yaacob, S., Adom, A. H., & Nagarajan, R. (2008). Recognition of motor imagery of hand movements for a BMI using PCA features. *2008 International*

- Conference on Electronic Design, ICED 2008, February 2014*, 1–5. <https://doi.org/10.1109/ICED.2008.4786683>
- Kevric, J., & Subasi, A. (2017). Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system. *Biomedical Signal Processing and Control*, *31*, 398–406. <https://doi.org/10.1016/j.bspc.2016.09.007>
- Kim, Y., Ryu, J., Kim, K. K., Took, C. C., Mandic, D. P., & Park, C. (2016). Motor Imagery Classification Using Mu and Beta Rhythms of EEG with Strong Uncorrelating Transform Based Complex Common Spatial Patterns. *Computational Intelligence and Neuroscience*, *2016*. <https://doi.org/10.1155/2016/1489692>
- Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., & Lance, B. J. (2018). EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, *15*(5), 1–30. <https://doi.org/10.1088/1741-2552/aace8c>
- Lecun, Y., Bottou, L., Bengio, Y., & Ha, P. (1998). LeNet. *Proceedings of the IEEE, November*, 1–46. <https://doi.org/10.1109/5.726791>
- Maksimenko, V. A., Van Heukelum, S., Makarov, V. V., Kelderhuis, J., Lüttjohann, A., Koronovskii, A. A., Hramov, A. E., & Van Luijtelaaar, G. (2017). Absence Seizure Control by a Brain Computer Interface. *Scientific Reports*, *7*(1), 1–8. <https://doi.org/10.1038/s41598-017-02626-y>
- Padfield, N., Zabalza, J., Zhao, H., Masero, V., & Ren, J. (2019). EEG-based brain-computer interfaces using motor-imagery: Techniques and challenges. *Sensors (Switzerland)*, *19*(6), 1–34. <https://doi.org/10.3390/s19061423>
- Pfurtscheller, G. (2001). Functional brain imaging based on ERD/ERS. *Vision Research*, *41*(10–11), 1257–1260. [https://doi.org/10.1016/S0042-6989\(00\)00235-2](https://doi.org/10.1016/S0042-6989(00)00235-2)
- Popov, E., & Fomenkov, S. (2016). Classification of hand motions in EEG signals using recurrent neural networks. *2016 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, 1–4. <https://doi.org/10.1109/ICIEAM.2016.7911620>
- Rodríguez-Bérmudez, G., & García-Laencina, P. J. (2012). Automatic and adaptive classification of electroencephalographic signals for brain computer interfaces. *Journal of Medical Systems*, *36*(SUPPL.1). <https://doi.org/10.1007/s10916-012-9893-4>
- Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: A general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, *51*(6), 1034–1043. <https://doi.org/10.1109/TBME.2004.827072>
- Schirrmester, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggenberger, K., Tangermann, M., Hutter, F., Burgard, W., & Ball, T. (2017). Deep learning with convolutional neural networks for EEG decoding and visualization. *Human Brain Mapping*, *38*(11), 5391–5420. <https://doi.org/10.1002/hbm.23730>
- Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in Neural Information Processing Systems, 2015-January*, 802–810.

- Speier, W., Deshpande, A., & Pouratian, N. (2015). A method for optimizing EEG electrode number and configuration for signal acquisition in P300 speller systems. *Clinical Neurophysiology*, *126*(6), 1171–1177. <https://doi.org/10.1016/j.clinph.2014.09.021>
- Stawicki, P., Gemblar, F., Rezeika, A., & Volosyak, I. (2017). A novel hybrid mental spelling application based on eye tracking and SSVEP-based BCI. *Brain Sciences*, *7*(4). <https://doi.org/10.3390/brainsci7040035>
- Tang, Z., Sun, S., Zhang, S., Chen, Y., Li, C., & Chen, S. (2016). A brain-machine interface based on ERD/ERS for an upper-limb exoskeleton control. *Sensors (Switzerland)*, *16*(12), 1–14. <https://doi.org/10.3390/s16122050>
- Tangemann, M., Müller, K.-R., Aertsen, A., Birbaumer, N., Braun, C., Brunner, C., Leeb, R., Mehring, C., Miller, K., Mueller-Putz, G., Nolte, G., Pfurtscheller, G., Preissl, H., Schalk, G., Schlögl, A., Vidaurre, C., Waldert, S., & Blankertz, B. (2012). Review of the BCI Competition IV. *Frontiers in Neuroscience*, *6*, 55. <https://doi.org/10.3389/fnins.2012.00055>
- Tayeb, Z., Fedjaev, J., Ghaboosi, N., Richter, C., Everding, L., Qu, X., Wu, Y., Cheng, G., & Conradt, J. (2019). Validating deep neural networks for online decoding of motor imagery movements from eeg signals. *Sensors (Switzerland)*, *19*(1). <https://doi.org/10.3390/s19010210>
- Wang, X., Hersche, M., Tomekce, B., Kaya, B., Magno, M., & Benini, L. (2020). An Accurate EEGNet-based Motor-Imagery Brain-Computer Interface for Low-Power Edge Computing. *IEEE Medical Measurements and Applications, MeMeA 2020 - Conference Proceedings, March*. <https://doi.org/10.1109/MeMeA49120.2020.9137134>
- Zhang, X., Yao, L., Huang, C., Sheng, Q. Z., & Wang, X. (2017). Intent Recognition in Smart Living Through Deep Recurrent Neural Networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10635 LNCS*, 748–758. https://doi.org/10.1007/978-3-319-70096-0_76

