

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ανίχνευση ανθρώπων με τη χρήση Wi-Fi CSI (Channel
State Information)



Του φοιτητή
Τζανή Βασίλειου
Αριθμός Μητρώου: 2020169

Επιβλέπων
Ιορδάνης Κιοσκερίδης
Καθηγητής

Ημερομηνία 31-05-2026

Ανίχνευση ανθρώπων με τη χρήση Wi-Fi CSI (Channel State Information)

26182

Τζανής Βασίλειος

Ιορδάνης Κιοσκερίδης

13-03-2026

31-05-2026

Βεβαιώνω ότι είμαι ο συγγραφέας αυτής της εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, έχω καταγράψει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών, εικόνων και κειμένου, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επιπλέον, βεβαιώνω ότι αυτή η εργασία προετοιμάστηκε από εμένα προσωπικά, ειδικά ως διπλωματική εργασία, στο Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του ΔΙ.ΠΑ.Ε.

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή Τζανή Βασιλείου που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης, ο συγγραφέας/δημιουργός εκχωρεί στο Διεθνές Πανεπιστήμιο της Ελλάδος άδεια χρήσης του δικαιώματος αναπαραγωγής, δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσης της εργασίας διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος. Η ανοικτή πρόσβαση στο πλήρες κείμενο της εργασίας, δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού, ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, πώληση, εμπορική χρήση, διανομή, έκδοση, μεταφόρτωση (downloading), ανάρτηση (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού.

Η έγκριση της διπλωματικής εργασίας από το Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του Διεθνούς Πανεπιστημίου της Ελλάδος, δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα, εκ μέρους του Τμήματος.

«Στην Οικογένεια μου»

Πρόλογος

Η συγκεκριμένη διπλωματική επιλέχθηκε γιατί συνδυάζει το Wi-Fi CSI με την ανάλυση δεδομένων και το machine learning. Μέσα από αυτή την εργασία έγινε πιο κατανοητό πώς τα σήματα Wi-Fi μπορούν να χρησιμοποιηθούν για να βγάλουμε χρήσιμες πληροφορίες από το περιβάλλον. Η ενασχόληση με το θέμα αυτό με βοήθησε να κατανοήσω καλύτερα τη σχέση ανάμεσα στο σήμα και την επεξεργασία του, αλλά και τις δυσκολίες που υπάρχουν σε πραγματικές συνθήκες.

Περίληψη

Η παρούσα διπλωματική εργασία ασχολείται με την ανίχνευση ανθρώπινης κίνησης σε εσωτερικούς χώρους μέσω Wi-Fi CSI (Channel State Information), χωρίς κάμερες, ραντάρ ή φορητές συσκευές. Το σύστημα στηρίζεται σε δύο πλακέτες Seeed Studio XIAO ESP32-C6 συνολικού κόστους περίπου 20€, με αρχιτεκτονική SISO, με τη μία να εκπέμπει πακέτα στα 100 Hz και την άλλη να εξάγει το CSI σε πραγματικό χρόνο. Το firmware αναπτύχθηκε σε ESP-IDF με FreeRTOS και χρησιμοποιεί το πρωτόκολλο ESPNOW για ταχεία επικοινωνία μεταξύ των μικροελεγκτών. Ένα κρίσιμο ήταν το κλείδωμα του AGC gain, ώστε οι μεταβολές στο πλάτος του σήματος να αποτυπώνουν την πραγματική κίνηση. Συλλέχθηκαν 100 καταγραφές των 60 δευτερολέπτων (50 Walk, 50 Idle), δηλαδή 600.000 ακατέργαστα πλαίσια CSI. Η προεπεξεργασία σήματος περιλαμβάνει αφαίρεση των null subcarriers (από 128 σε 114), Hampel filter για ακραίες τιμές, Butterworth χαμηλοπερατό φίλτρο στα 10 Hz, temporal differencing, PCA με 10 συνιστώσες που συλλαμβάνουν το μεγαλύτερο μέρος της συνολικής διακύμανσης (ενδεικτικά 80–98%, ανάλογα με την καταγραφή), και τέλος η κανονικοποίηση των δεδομένων. Από κάθε παράθυρο 100 δειγμάτων εξάγονται 25 χαρακτηριστικά ανά συνιστώσα, δηλαδή 250 διαστάσεις συνολικά, χωρισμένα σε στατιστικά, φασματικά (FFT) και αυτοσυσχέτισης χαρακτηριστικά. Για την αντιμετώπιση του περιορισμένου dataset εφαρμόστηκε και data augmentation που πενταπλασίασε τα παράθυρα εκπαίδευσης. Δοκιμάστηκαν έξι αλγόριθμοι Classical ML με cross-validation ($k=5$), ώστε οι καταγραφές να μη διαρρέουν μεταξύ train και test, και τελική αξιολόγηση σε ανεξάρτητο σύνολο ελέγχου (20% των καταγραφών). Το Gradient Boosting έδωσε την υψηλότερη ακρίβεια (97,17%), με Random Forest και Extra Trees να ακολουθούν πολύ κοντά (96,98% και 97,12% αντίστοιχα). Τέλος, το waveform_length αναδείχθηκε ως το κυρίαρχο χαρακτηριστικό για την κλάση Walk σε όλα τα δενδρικά μοντέλα. Συνοπτικά, αποδεικνύεται ότι αξιόπιστη ανίχνευση Motion/Idle είναι εφικτή με SISO hardware των 20€ και Classical ML σε πραγματικό χρόνο, χωρίς ανάγκη για ακριβές κάρτες δικτύου ή βαριά μοντέλα Deep Learning. Διευκρινίζεται ότι ο όρος «ανίχνευση ανθρώπων» του τίτλου αναφέρεται ειδικά στη δυαδική ανίχνευση κίνησης (Motion/Idle) και όχι στην αναγνώριση επιμέρους δραστηριοτήτων ή στον εντοπισμό θέσης ή ταυτότητας.

«Human Activity Recognition Using Wi-Fi Channel State Information»

Tzanis Vasileios

Abstract

This thesis addresses indoor human motion detection through Wi-Fi CSI (Channel State Information), without the use of cameras, radar or wearable devices. The system relies on two Sseed Studio XIAO ESP32-C6 boards with a total cost of approximately 20€, in a SISO architecture, where one transmits packets at 100 Hz and the other extracts CSI in real time. The firmware was developed in ESP-IDF with FreeRTOS and uses the ESPNOW protocol for fast communication between the microcontrollers. A critical aspect of the implementation was locking the AGC gain, so that amplitude variations in the signal could actually capture real motion. A total of 100 recordings of 60 seconds each were collected (50 Walk, 50 Idle), that is, 600,000 raw CSI frames. Signal preprocessing includes null subcarrier removal (from 128 down to 114), a Hampel filter for outliers, a Butterworth low-pass filter at 10 Hz, temporal differencing, PCA with 10 components capturing most of the total variance (indicatively 80–98%, depending on the recording), and finally normalization. From each window of 100 samples, 25 features are extracted per component, that is 250 dimensions in total, divided into statistical, frequency-domain (FFT) and autocorrelation features. To address the limited dataset, data augmentation was also applied, multiplying the training windows by five. Six Classical ML algorithms were tested with cross-validation ($k=5$), to prevent leakage between train and test sets, with final evaluation on an independent hold-out test set (20% of the recordings). Gradient Boosting yielded the highest accuracy (97.17%), with Random Forest and Extra Trees following very closely (96.98% and 97.12% respectively). The `waveform_length` feature emerged as the dominant feature for the Walk class across all tree-based models. In practice, it is demonstrated that reliable Motion/Idle detection is feasible with SISO hardware costing 20€ and Classical ML in real time, without the need for expensive network cards or heavy Deep Learning models. It is clarified that the term “human detection” in the title refers specifically to binary motion detection (Motion/Idle), and not to the recognition of individual activities or to the localization of position or identity.

Ευχαριστίες

Ολοκληρώνοντας την παρούσα εργασία, ευχαριστώ θερμά τον επιβλέποντα καθηγητή μου, κ. Κιοσκερίδη. Ευχαριστώ την οικογένειά μου, που αποτέλεσε το σταθερό μου στήριγμα και μου προσέφερε κάθε δυνατή βοήθεια στα φοιτητικά μου χρόνια. Τέλος, ευχαριστώ όσους στάθηκαν δίπλα μου σε αυτό το ακαδημαϊκό ταξίδι.

Περιεχόμενα

Πρόλογος.....	v
Περίληψη.....	vi
Abstract	vii
Ευχαριστίες	viii
Περιεχόμενα	ix
Κατάλογος Σχημάτων	xii
Κατάλογος Πινάκων.....	xiv
Συντομογραφίες.....	xvi
Κεφάλαιο 1ο: Πλαίσιο και Στόχοι της Εργασίας.....	1
1.1 Εισαγωγή	1
1.2 Κίνητρο	1
1.3 Ορισμός Προβλήματος και Αρχιτεκτονική Προσέγγιση.....	1
1.4 Συνεισφορές	2
1.5 Εμβέλεια και Περιορισμοί.....	2
1.6 Δομή Εργασίας.....	3
Κεφάλαιο 2ο: Θεωρητικό Υπόβαθρο.....	4
2.1 Εισαγωγή	4
2.2 Πρότυπο IEEE 802.11 και Πολυπλεξία OFDM.....	4
2.2.1 Διαμόρφωση Υποφορέων (Modulation) και Σχήματα MCS	6
2.3 Channel State Information (CSI).....	7
2.3.1 CSI έναντι RSSI	8
2.3.2 Πρόσβαση και Χρήση CSI	9
2.3.3 Φάση CSI και Εξυγίανση Φάσης (Phase Sanitization).....	10
2.4 Φαινόμενο Πολυδιαδρομικής Διάδοσης (Multipath) και Ανθρώπινη Κίνηση.....	10
2.4.1 Μοντέλα Διαλείψεων και Συνοχή Καναλιού.....	12
2.4.2 Ο ρόλος του ανθρώπινου σώματος.....	12
2.4.3 Ζώνες Fresnel	13
2.5 Θεωρητικοί Περιορισμοί Κεραιών: SISO Έναντι MIMO και Χωρική Ανάλυση.....	14
2.6 Αλγόριθμοι Μηχανικής Μάθησης.....	15
2.6.1 Παραδοσιακοί Αλγόριθμοι Μάθησης.....	15
2.6.2 Βαθιά Μάθηση (Deep Learning)	16
2.6.3 Σύγκριση και Κριτήρια Επιλογής.....	17

2.7	Θεωρητικό Υπόβαθρο Επεξεργασίας Σήματος	18
2.8	Θεωρητικό Υπόβαθρο Χαρακτηριστικών και Μετρικών Αξιολόγησης	18
2.9	Σύνοψη του κεφαλαίου.....	19
Κεφάλαιο 3ο: Βιβλιογραφική Ανασκόπηση		21
3.1	Εισαγωγή.....	21
3.2	Η Εξέλιξη του CSI-based HAR (2007–2026).....	21
3.3	Συστήματα Χαμηλού Κόστους και Υλικός Εξοπλισμός (Hardware).....	22
3.3.1	Συστήματα Intel 5300 NIC (MIMO).....	22
3.3.2	Nexmon CSI (Raspberry Pi).....	23
3.3.3	ESP32 Η Πλατφόρμα Νέας Γενιάς.....	23
3.4	Προσεγγίσεις Κλασικής Μηχανικής Μάθησης (Classical ML).....	24
3.5	Προσεγγίσεις Βαθιάς Μάθησης (Deep Learning).....	25
3.5.1	Convolutional Neural Networks (CNN).....	25
3.5.2	LSTM και Παραλλαγές	25
3.5.3	Vision Transformers και Foundation Models.....	25
3.6	Πλεονεκτήματα: Ιδιωτικότητα και Ασφάλεια	26
3.7	Ανοιχτές Προκλήσεις: Data Leakage και Cross-Environment.....	26
3.8	Συγκριτικός Πίνακας Βιβλιογραφίας	27
Κεφάλαιο 4ο: Σχεδιασμός και Υλοποίηση Συστήματος.....		30
4.1	Αρχιτεκτονική Συστήματος.....	30
4.1.1	AGC Gain Locking.....	33
4.1.2	Δυνατότητες Wi-Fi 6 (802.11ax), ο λόγος μη αξιοποίησής τους, και μέγεθος του Μικροελεγκτή.....	34
4.1.3	Πλακέτα Seeed Studio XIAO ESP32-C6 και Εξωτερική Κεραία.....	35
4.2	Firmware Εξαγωγής CSI.....	35
4.2.1	Ρυθμός Δειγματοληψίας και FreeRTOS Ουρά.....	35
4.2.2	Δομή και Περιεχόμενο CSI Πλαισίου	36
4.3	Πειραματικό Περιβάλλον και Χωρική Διάταξη.....	37
4.4	Πρωτόκολλο Συλλογής Δεδομένων	39
4.5	Σύνολο Δεδομένων και Επαύξηση.....	40
4.5.1	Ακατέργαστα Δεδομένα και Παράθυρα sliding window.....	40
4.5.2	Επαύξηση Δεδομένων (Data Augmentation)	40
4.6	Προεπεξεργασία Σήματος	42
4.6.1	Αφαίρεση Null Subcarriers.....	43
4.6.2	Φίλτρο Hampel — Αφαίρεση Ακραίων Τιμών	50

4.6.3	Φίλτρο Butterworth Χαμηλής Συχνότητας.....	54
4.6.4	Χρονική Διαφοροποίηση (Temporal Differencing)	56
4.6.5	Ανάλυση Κύριων Συνιστωσών (PCA)	59
4.6.6	Κανονικοποίηση Δεδομένων.....	61
4.7	Εξαγωγή Χαρακτηριστικών	62
4.7.1	Στατιστικά Χαρακτηριστικά Πεδίου Χρόνου.....	63
4.7.2	Χαρακτηριστικά Πεδίου Συχνότητας	65
4.7.3	Χαρακτηριστικά Αυτοσυσχέτισης και Μορφολογίας	67
4.8	Αλγόριθμοι Ταξινόμησης.....	70
4.9	Γραφικό Περιβάλλον Σε Πραγματικό Χρόνο	72
Κεφάλαιο 5ο:	Αποτελέσματα και Αξιολόγηση.....	79
5.1	Αξιολόγηση και Αποτελέσματα ανά Αλγόριθμο.....	79
5.2	Ανάλυση Σημαντικότητας Χαρακτηριστικών	80
5.3	Πίνακες Σύγχυσης	84
5.4	Αξιολόγηση μέσω Καμπυλών ROC και Μετρικής AUC.....	87
5.4.1	Η ROC καμπύλη.....	87
5.4.2	Το AUC	88
5.4.3	Αξιολόγηση Αποτελεσμάτων	88
5.5	Μελέτη Απομόνωσης Σταδίων (Ablation Study).....	92
5.6	Αξιολόγηση Γενίκευσης σε Νέο Περιβάλλον (Cross-Environment)	97
5.7	Υπολογιστική Απόδοση και Συμπεριφορά Πραγματικού Χρόνου	99
5.8	Συζήτηση Αποτελεσμάτων.....	102
Κεφάλαιο 6ο:	Συμπεράσματα και Μελλοντική Επέκταση	104
6.1	Σύνοψη της Εργασίας.....	104
6.2	Κύρια Συμπεράσματα.....	104
6.3	Περιορισμοί.....	104
6.4	Κατευθύνσεις Μελλοντικής Έρευνας	105
BIBΛΙΟΓΡΑΦΙΑ.....		106
ΠΑΡΑΡΤΗΜΑ Α : Κώδικας υλοποίησης Εξαγωγής CSI.....		110
ΠΑΡΑΡΤΗΜΑ Β : Κώδικας υλοποίησης Ρυθμού Δειγματοληψίας		112
ΠΑΡΑΡΤΗΜΑ C : Πρωτόκολλο Συλλογής Δεδομένων.....		112
ΠΑΡΑΡΤΗΜΑ D : Επαύξηση Δεδομένων (Data Augmentation)		114
ΠΑΡΑΡΤΗΜΑ Ε : Πλήρης Πηγαίος Κώδικας και Αποθετήριο (GitHub)		116

Κατάλογος Σχημάτων

Εικόνα 2.1. Φαινόμενο Πολυδιαδρομικής Διάδοσης (Multipath) και Ανθρώπινη Κίνηση [2].	11
Εικόνα 2.2. Ζώνη Fresnel [12].	13
Εικόνα 2.3. AoA (Angle of Arrival) [2].	14
Εικόνα 2.4. Ένα Βαθύ Νευρωνικό Δίκτυο [18].	16
Εικόνα 2.5. Φασματόγραμμα (αναπαράσταση χρόνου–συχνότητας).	17
Εικόνα 4.1. Αρχιτεκτονική της πλακέτας σε υψηλό επίπεδο.	30
Εικόνα 4.2. Ο Μικροελεγκτής και το Κύκλωμα RF.	31
Εικόνα 4.3. Τροφοδοσία και Διασύνδεση (Power & USB).	32
Εικόνα 4.4. Μικροελεγκτής ESP32-C6.	33
Εικόνα 4.5. Τοποθέτηση του Πομπού στο χώρο της δοκιμής εμβέλειας στα 14 μέτρα, απόσταση από τον Δέκτη (LoS).	37
Εικόνα 4.6. Τοποθέτηση του Δέκτη στο χώρο της δοκιμής εμβέλειας στα 14 μέτρα, απόσταση από τον Πομπό (LoS).	38
Εικόνα 4.7. Επαύξηση δεδομένων σε μη πραγματικά δεδομένα	40
Εικόνα 4.8. Επαύξηση δεδομένων σε πραγματικά δεδομένα στο αρχείο no_activity_11_room1_vasilis_1777973010.txt	40
Εικόνα 4.9. Επαύξηση δεδομένων σε πραγματικά δεδομένα στο αρχείο walk_activity_45_room1_vasilis_1777806476.txt	41
Εικόνα 4.10. Το ακατέργαστο σήμα από το αρχείο no_activity_11_room1_vasilis_1777973010.txt (Heatmap Diagram).	43
Εικόνα 4.11. Αφαίρεση μηδενικών τιμών από το αρχείο no_activity_11_room1_vasilis_1777973010.txt (Heatmap Diagram).	44
Εικόνα 4.12. Το ακατέργαστο σήμα από το αρχείο no_activity_14_room1_vasilis_1777973321.txt από 128 υποφορείς σε 114 (Plot Diagram)	44
Εικόνα 4.13. Αφαίρεση μηδενικών τιμών από το αρχείο no_activity_14_room1_vasilis_1777973321.txt από 128 υποφορείς σε 114	45
Εικόνα 4.14. Το ακατέργαστο σήμα από το αρχείο walk_activity_12_room1_vasilis_1777803073.txt (Heatmap Diagram).	46
Εικόνα 4.15. Αφαίρεση μηδενικών τιμών από το αρχείο walk_activity_12_room1_vasilis_1777803073.txt (Heatmap Diagram).	47
Εικόνα 4.16. Το ακατέργαστο σήμα από το walk_activity_12_room1_vasilis_1777803073.txt από 128 υποφορείς σε 114 (Plot Diagram)	48
Εικόνα 4.17. Αφαίρεση μηδενικών τιμών από το αρχείο walk_activity_12_room1_vasilis_1777803073.txt από 128 υποφορείς σε 114 (Plot Diagram)	48

Εικόνα 4.18. Εφαρμογή του Φίλτρου Hampel στο αρχείο no_activity_14_room1_vasilis_1777973321.txt (Plot Diagram)	50
Εικόνα 4.19. Εφαρμογή του Φίλτρου Hampel στο walk_activity_12_room1_vasilis_1777803073.txt (Plot Diagram).....	51
Εικόνα 4.20. Εφαρμογή του Φίλτρου Hampel στο walk_activity_12_room1_vasilis_1777803073.txt (Heatmap Diagram).....	52
Εικόνα 4.21. Εφαρμογή του Φίλτρου Hampel στο no_activity_14_room1_vasilis_1777973321.txt (Heatmap Diagram).....	52
Εικόνα 4.22. Εφαρμογή του Φίλτρου Butterworth στο no_activity_14_room1_vasilis_1777973321.txt (Heatmap Diagram).....	53
Εικόνα 4.23. Εφαρμογή του Φίλτρου Butterworth στο walk_activity_12_room1_vasilis_1777803073.txt (Heatmap Diagram).....	54
Εικόνα 4.24. Χρήση του Temporal Differencing στο no_activity_14_room1_vasilis_1777973321.txt (Heatmap Diagram).....	55
Εικόνα 4.25. Χρήση του Temporal Differencing στο walk_activity_12_room1_vasilis_1777803073.txt (Heatmap Diagram).....	57
Εικόνα 4.26. Χρήση του Temporal Differencing στο walk_activity_12_room1_vasilis_1777803073.txt (Plot Diagram).....	57
Εικόνα 4.27. Χρήση του Temporal Differencing στο no_activity_14_room1_vasilis_1777973321.txt	58
Εικόνα 4.28. Χρήση του PCA στο no_activity_14_room1_vasilis_1777973321.txt.....	59
Εικόνα 4.29. Χρήση του PCA στο walk_activity_12_room1_vasilis_1777803073.txt.....	59
Εικόνα 4.30. Κανονικοποίηση δεδομένων στο walk_activity_12_room1_vasilis_1777803073.txt.....	60
Εικόνα 4.31. Κανονικοποίηση δεδομένων στο no_activity_14_room1_vasilis_1777973321.txt.....	61
Εικόνα 4.32. Ανάλυση αρχείου no_activity_11_room1_vasilis_1777973010.txt Amplitude vs Time	62
Εικόνα 4.33. Ανάλυση αρχείου walk_activity_35_room1_vasilis_1777805631.txt Amplitude vs Time	63
Εικόνα 4.34. Signal Energy και Variance αρχείου walk_activity_35_room1_vasilis_1777805631.txt	63
Εικόνα 4.35. Signal Energy και Variance αρχείου no_activity_11_room1_vasilis_1777973010.txt ...	64
Εικόνα 4.36. Ανάλυση του αρχείου no_activity_11_room1_vasilis_1777973010.txt στο πεδίο της συχνότητας	65
Εικόνα 4.37. Ανάλυση του αρχείου walk_activity_35_room1_vasilis_1777805631.txt στο πεδίο της συχνότητας.	66
Εικόνα 4.38. Πρακτική εφαρμογή όλων των βημάτων για την εισαγωγή τους στα μοντέλα μηχανικής μάθησης. Αρχείο no_activity_14_room1_vasilis_1777973321.txt.....	70

Εικόνα 4.39. Πρακτική εφαρμογή όλων των βημάτων για την εισαγωγή τους στα μοντέλα μηχανικής μάθησης. Αρχείο walk_activity_12_room1_vasilis_1777803073.txt	70
Εικόνα 4.40. Αρχικό dashboard της εφαρμογής.....	72
Εικόνα 4.41. Καρτέλα Signal View της εφαρμογής.....	73
Εικόνα 4.42. Καρτέλα Activity Log.....	74
Εικόνα 4.43. Καρτέλα πληροφοριών του Τρέχοντος Συστήματος.	75
Εικόνα 4.44. Ρυθμίσεις Μοντέλων Συστήματος	76
Εικόνα 5.1. Ανάλυση χαρακτηριστικών των τριών αλγορίθμων.	81
Εικόνα 5.2. Ανάλυση χαρακτηριστικών με βάση τις κλάσεις (walk, idle) των τριών αλγορίθμων.	82
Εικόνα 5.3. Πίνακες Σύγκρισης των αλγορίθμων Extra Trees και Gradient Boosting	84
Εικόνα 5.4. Πίνακες Σύγκρισης των αλγορίθμων SVM και k-NN.....	85
Εικόνα 5.5. Πίνακες Σύγκρισης των αλγορίθμων Random Forest και Logistic Regression.....	86
Εικόνα 5.6. Καμπύλες ROC (One-vs-Rest) για τους ταξινομητές Extra Trees, Random Forest και SVM στο σύνολο ελέγχου.	88
Εικόνα 5.7. Καμπύλες ROC (One-vs-Rest) για τους ταξινομητές Gradient Boosting, Logistic Regression και k-NN στο σύνολο ελέγχου.	90
Εικόνα 5.8. Συνολικός χρόνος ανά παράθυρο (p95) ανά ταξινομητή.....	101

Κατάλογος Πινάκων

Πίνακας 3.1: Συγκριτικός πίνακας ανίχνευσης ανθρώπινης δραστηριότητας, Wi-Fi CSI HAR.	28
Πίνακας 4.1: Χαρακτηριστικά ανά κύρια συνιστώσα ($25 \times 10 = 250$ διαστάσεις).	68
Πίνακας 5.1: Μετρικές απόδοσης στο αποκλεισμένο σύνολο ελέγχου (hold-out 20% των καταγραφών, 2.189 παράθυρα) για τους έξι αλγορίθμους ML.	79
Πίνακας 5.2: Πίνακες σύγκρισης ανά αλγόριθμο. (Κλάση 0 = Idle, Κλάση 1 = Walk, σύνολο ελέγχου 2.189 παράθυρα)	83
Πίνακας 5.3: Μελέτη ablation με βάση το Random Forest. Κάθε γραμμή μεταβάλλει ένα στοιχείο του pipeline κρατώντας τα υπόλοιπα σταθερά (ακρίβεια στο αποκλεισμένο σύνολο ελέγχου).	93
Πίνακας 5.4: Επίδραση της αφαίρεσης του φίλτρου Butterworth ανά μοντέλο. Χωρίς το χαμηλοπερατό φιλτράρισμα, μόνο το Random Forest παραμένει αξιόπιστο.	93
Πίνακας 5.5: Πλήρης μελέτη ablation με ακρίβεια (%) στο αποκλεισμένο σύνολο ελέγχου ανά μοντέλο, για κάθε μεταβολή ενός στοιχείου του pipeline.	95
Πίνακας 5.6: Αξιολόγηση cross-environment (LEGO). Εκπαίδευση σε έναν χώρο, έλεγχος στον άλλο.	97
Πίνακας 5.7. Υπολογιστική απόδοση των έξι ταξινομητών.....	100

Συντομογραφίες

Συντομογραφία	Επεξήγηση / Ορισμός
A-BLSTM	Attention-based Bidirectional Long Short-Term Memory
ADC	Analog-to-Digital Converter (Μετατροπέας Αναλογικού σε Ψηφιακό)
AGC	Automatic Gain Control
AoA	Angle of Arrival (Γωνία Άφιξης)
AUC	Area Under the Curve (Εμβαδόν κάτω από την καμπύλη)
BLSTM	Bidirectional Long Short-Term Memory
BPSK	Binary Phase Shift Keying
CFO	Carrier Frequency Offset
CIR	Channel Impulse Response (Κρουστική Απόκριση Καναλιού)
CNN	Convolutional Neural Networks (Συνελικτικά Νευρωνικά Δίκτυα)
CSI	Channel State Information (Πληροφορία Κατάστασης Καναλιού)
CSV	Comma-Separated Values (Αρχείο τιμών διαχωρισμένων με κόμμα)
dB / dBi	Decibel / Decibel isotropic (Μονάδες μέτρησης ισχύος / κέρδους κεραίας)
DC	Direct Current (Συνεχές Ρεύμα / Κεντρική συχνότητα στο φάσμα)
DfP	Device-free Passive
DFWS	Device-Free Wi-Fi Sensing
DL	Deep Learning (Βαθιά Μάθηση)
DNN	Deep Neural Network
DSP	Digital Signal Processing (Ψηφιακή Επεξεργασία Σήματος)
ESP-IDF	Espressif IoT Development Framework
ET	Extra Trees
FFT	Fast Fourier Transform (Μετασχηματισμός Fourier)

Συντομογραφία	Επεξήγηση / Ορισμός
FN	False Negative (Ψευδώς Αρνητικό)
FP	False Positive (Ψευδώς Θετικό)
GB	Gradient Boosting
GCoV	Grouped Coefficient of Variation
GHz / MHz / Hz	Gigahertz / Megahertz / Hertz (Μονάδες Συχνότητας)
GPIO	General-Purpose Input/Output (Ακροδέκτης γενικής χρήσης)
GRU	Gated Recurrent Unit
HAR	Human Activity Recognition (Αναγνώριση Ανθρώπινης Δραστηριότητας)
HT40	High Throughput 40 MHz
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things (Διαδίκτυο των Πραγμάτων)
ISAC	Integrated Sensing and Communication
JTAG	Joint Test Action Group (Θύρα/Πρωτόκολλο αποσφαλμάτωσης)
k-NN	k-Nearest Neighbors
LoS	Line-of-Sight (Άμεση Οπτική Επαφή)
LR	Logistic Regression (Λογιστική Παλινδρόμηση)
LEGO	Leave-Environment-Group-Out
LSTM	Long Short-Term Memory
LTF	Long Training Field (Σύμβολα Εκπαίδευσης Καναλιού)
MAC	Media Access Control (Έλεγχος Πρόσβασης Μέσου)
MAD	Median Absolute Deviation
max / min	Maximum / Minimum (Μέγιστο / Ελάχιστο)
MCS	Modulation and Coding Scheme

Συντομογραφία	Επεξήγηση / Ορισμός
MIMO	Multiple-Input Multiple-Output (Πολλαπλοί Πομποί - Πολλαπλοί Δέκτες)
ML	Machine Learning (Μηχανική Μάθηση)
mmWave	Millimeter Wave (Χιλιοστομετρικά Κύματα)
NIC	Network Interface Card (Κάρτα Δικτύου)
NLoS	Non-Line-of-Sight (Απουσία Οπτικής Επαφής / Μη άμεση οπτική επαφή)
OFDM	Orthogonal Frequency-Division Multiplexing (Ορθογωνική Πολυπλεξία Διαίρεσης Συχνότητας)
OTFS	Orthogonal Time Frequency Space (κυματομορφή/διαμόρφωση)
PBD	Packet Boundary Detection
PC	Principal Component (Κύρια Συνιστώσα της ανάλυσης, π.χ. PC1, PC2)
PCA	Principal Component Analysis (Ανάλυση Κύριων Συνιστωσών)
PSD	Power Spectral Density (Φάσμα Ισχύος)
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RBM	Restricted Boltzmann Machine
RF	Random Forest (Αλγόριθμος) καθώς και Radio Frequency (Ραδιοσυχνότητα)
RNN	Recurrent Neural Network
RSSI	Received Signal Strength Indicator (Ένδειξη Ισχύος Λαμβανόμενου Σήματος)
ROC	Receiver Operating Characteristic
Rx	Receiver (Δέκτης)
SFO	Sampling Frequency Offset
SISO	Single-Input Single-Output (Ένας Πομπός - Ένας Δέκτης)
SNR	Signal-to-Noise Ratio (Λόγος Σήματος προς Θόρυβο)

Συντομογραφία	Επεξήγηση / Ορισμός
SoC	System on a Chip
SOTA	State of the Art (Η τελευταία λέξη της τεχνολογίας / βέλτιστη επίδοση)
SPI	Serial Peripheral Interface (Δίαυλος επικοινωνίας μνήμης)
std	Standard Deviation (Τυπική Απόκλιση)
STFT	Short-Time Fourier Transform (Βραχύχρονος Μετασχηματισμός Fourier)
SVM	Support Vector Machine
THz	Terahertz
TinyML	Tiny Machine Learning
TN	True Negative (Αληθώς Αρνητικό)
TP	True Positive (Αληθώς Θετικό)
Tx	Transmitter (Πομπός)
TXT	Text File (Αρχείο Κειμένου)
UART	Universal Asynchronous Receiver-Transmitter (Σειριακή θύρα)
USB	Universal Serial Bus
ViT	Vision Transformers
Wi-Fi	Εμπορική ονομασία της Wi-Fi Alliance για ασύρματα δίκτυα IEEE 802.11
ZCR	Zero-Crossing Rate (Ποσοστό Αλλαγής Προσήμου)
Δ.Ε.	Διπλωματική Εργασία
ΔΙ.ΠΑ.Ε. / ΔΙΠΑΕ	Διεθνές Πανεπιστήμιο της Ελλάδος
Π.Ε.	Πτυχιακή Εργασία

Κεφάλαιο 1ο: Πλαίσιο και Στόχοι της Εργασίας

1.1 Εισαγωγή

Το παρόν κεφάλαιο παρουσιάζει το κίνητρο για την επιλογή του θέματος, ορίζει το τεχνικό πρόβλημα και την αρχιτεκτονική προσέγγιση, περιγράφει τις κύριες συνεισφορές και ξεκαθαρίζει τους περιορισμούς της παρούσας εργασίας.

1.2 Κίνητρο

Σήμερα υπάρχουν πάνω από είκοσι ένα δισεκατομμύρια Wi-Fi συσκευές σε χρήση παγκοσμίως [32]. Αυτό σημαίνει ότι σε σχεδόν κάθε κατοικία, γραφείο ή δημόσιο χώρο υπάρχει ήδη και μια ασύρματη υποδομή. Το Wi-Fi Sensing, ή αλλιώς το Device-Free Wireless Sensing, εκμεταλλεύεται αυτή ακριβώς την υποδομή για την ανίχνευση κίνησης, χωρίς πρόσθετο εξοπλισμό και παράλληλα με την κλασική λειτουργία επικοινωνίας [1][3]. Η πρακτική αξία ενός τέτοιου συστήματος είναι μεγάλη. Οι εφαρμογές επεκτείνονται από την υγειονομική παρακολούθηση, όπου ένα τέτοιο σύστημα μπορεί να ειδοποιήσει σε περίπτωση πτώσης ηλικιωμένου ή να παρακολουθεί μια κινητική δραστηριότητα [2][13][14][33]. Σε έξυπνα σπίτια μπορεί να επιτρέψει αυτόματο φωτισμό και θέρμανση βάσει πραγματικής παρουσίας [1][5][34], ενώ στον τομέα της ασφάλειας μπορεί να λειτουργήσει ως σύστημα ανίχνευσης εισβολής χωρίς κάμερα [3][23].

Σε αντίθεση με τις κλασικές κάμερες όπου μπορεί να παραβιαστεί εύκολα η ιδιωτικότητα, το Wi-Fi CSI (Channel State Information) δεν αποκαλύπτει την ταυτότητα του ατόμου εύκολα [3][4], λειτουργεί χωρίς οπτική επαφή και είναι ανεξάρτητο από τον φωτισμό (π.χ. Ημέρα ή Νύχτα). Πρακτικά, δεν χρειάζεται ο χρήστης να φοράει καμία συσκευή [1][4], και το σύστημα μπορεί να λειτουργεί ακόμα και σε συνθήκες Non-Line-of-Sight, δηλαδή με εμπόδια μεταξύ πομπού και δέκτη [5].

Το ESP32-C6 της Espressif Systems κοστίζει μερικά ευρώ και μπορεί να εξάγει πραγματικά CSI (Channel State Information)[8][17], βέβαια, εκτός από το κόστος η απουσία πρόσθετης υποδομής και η ανεξαρτησία από φωτισμό κάνουν εφικτή την ανάπτυξη ενός τέτοιου συστήματος σε πραγματικές συνθήκες.

1.3 Ορισμός Προβλήματος και Αρχιτεκτονική Προσέγγιση

Το RSSI είναι γενικά μία συγκεντρωτική τιμή της συνολικής ισχύος του λαμβανόμενου σήματος για ένα ολόκληρο κανάλι ή πακέτο δεδομένων. Το CSI (Channel State Information), αντίθετα, μετρά την ανταπόκριση του ασύρματου καναλιού ανά υποφορέα OFDM, οπότε αντιδρά πολύ πιο ευαίσθητα στις μεταβολές που προκαλεί η ανθρώπινη κίνηση [2][3][5]. Επίσης, κατά πόσο μπορεί να αξιοποιηθεί στο έπακρο η πληροφορία αυτή εξαρτάται ωστόσο και από την αρχιτεκτονική κεραιών (SISO vs MIMO).

Η εργασία υλοποιήθηκε σε περιβάλλον με Single-Input Single-Output (SISO) και με ένα ζεύγος ESP32-C6 (Tx/Rx). Ένα SISO σύστημα δεν μπορεί να εκτιμήσει τη γωνία άφιξης (Angle of Arrival) ούτε να διαχωρίσει χωρικά σήματα από διαφορετικές κατευθύνσεις [5][8], άρα η αναγνώριση πολλών κλάσεων δραστηριότητας δεν ήταν εφικτή από την αρχή σε ένα πραγματικό περιβάλλον.

Ωστόσο, κάθε άνθρωπος όταν κινείται αφήνει μια χαρακτηριστική ενεργειακή υπογραφή στα σήματα CSI, ανεξάρτητα από τον τύπο κίνησης [8][9]. Η ανίχνευση αυτής της υπογραφής δεν απαιτεί χωρική πληροφορία και έτσι προκύπτει η δυνατότητα να πραγματοποιηθεί δυαδική ταξινόμηση μεταξύ Motion (παρουσία ατόμου σε κίνηση) και Idle (απουσία κίνησης) [4][11]. Διευκρινίζεται ότι ο όρος «ανίχνευση

ανθρώπων» του τίτλου ορίζεται στην παρούσα εργασία ως ανίχνευση ανθρώπινης κίνησης όπου δηλαδή ένας εντελώς ακίνητος άνθρωπος ταξινομείται στην κλάση Idle. Δεν πραγματοποιείται πλήρης αναγνώριση κινήσεων του ανθρώπου όπως το περπάτημα, το κάθισμα, το πέσιμο κτλ. αλλά πραγματοποιείται μια δυαδική αναγνώριση μεταξύ περπατήματος και ακινησίας στον χώρο.

Τέλος, για την ταξινόμηση επιλέχθηκαν αλγόριθμοι Classical Machine Learning αντί Deep Learning, κυρίως επειδή επιτρέπουν την χρήση σε πραγματικό χρόνο (real-time edge inference) και σε IoT πλατφόρμες με περιορισμένη υπολογιστική ισχύ [5][10][12].

1.4 Συνεισφορές

Οι κύριες συνεισφορές της εργασίας είναι τέσσερις. Πρώτον, σχεδιάστηκε και υλοποιήθηκε ένα ολοκληρωμένο σύστημα, από το firmware του ESP32-C6 έως και την πρόβλεψη σε πραγματικό χρόνο, καλύπτοντας κάθε στάδιο από τη συλλογή των δεδομένων έως την εκπαίδευση και το τελικό αποτέλεσμα (inference). Δεύτερον, η αξιολόγηση έγινε με StratifiedGroupKFold cross-validation σε επίπεδο καταγραφής, ώστε κάθε καταγραφή να εμφανίζεται αποκλειστικά στην εκπαίδευση ή στον έλεγχο (testing) ώστε έτσι να αποφευχθεί εντελώς η διαρροή δεδομένων (data leakage) [7]. Η τελική απόδοση μετρήθηκε σε αποκλεισμένο hold-out σύνολο ελέγχου (20% των καταγραφών). Τρίτον, συλλέχθηκε dataset 100 καταγραφών (50 Walk, 50 Idle, 60 δευτερόλεπτα, 100 λεπτά συνολικά). Λόγω όμως του περιορισμένου όγκου καταγραφών, εφαρμόστηκαν τεχνικές Data Augmentation που πενταπλασίασαν τα παράθυρα εκπαίδευσης χωρίς πρόσθετη συλλογή δεδομένων. Η επαύξηση στοχεύει κυρίως στην κανονικοποίηση και στην ανθεκτικότητα του μοντέλου σε θόρυβο και μετατοπίσεις, και όχι στη βελτίωση της ακρίβειας, όπως επιβεβαιώνεται και από τη μελέτη ablation του Κεφαλαίου 5. Τέλος, πραγματοποιήθηκε σύγκριση έξι αλγόριθμων Classical ML (Gradient Boosting, Random Forest, Logistic Regression, Extra Trees, k-NN, SVM) με ανάλυση των σημαντικότερων χαρακτηριστικών (Feature Importance) ανά αλγόριθμο [5]. Για την πλήρη διαφάνεια της υλοποίησης και τη διευκόλυνση όσων επιθυμούν να αναπαράγουν τα πειράματα, ολόκληρος ο πηγαίος κώδικας βρίσκεται αναρτημένος στο GitHub: <https://github.com/vasilstzanis/Wifi-CSI-Human-Detection>.

1.5 Εμβέλεια και Περιορισμοί

Η ταξινόμηση στην παρούσα εργασία είναι αποκλειστικά δυαδική (Motion/Idle) και ο στόχος είναι η ανίχνευση παρουσίας/κίνησης και όχι η αναγνώριση συγκεκριμένων δραστηριοτήτων. Τα δεδομένα συλλέχθηκαν κυρίως σε ένα δωμάτιο από ένα συγκεκριμένο άνθρωπο, με ένα μικρό υποσύνολο καταγραφών σε δεύτερο χώρο (σαλόνι). Η γενίκευση χώρου (cross-environment) είναι ένα γνωστό ανοιχτό πρόβλημα στο χώρο του CSI [4], καθώς τα σήματα CSI είναι ευαίσθητα στις πολλαπλές διαδρομές (multipath propagation). Κάθε χώρος έχει τη δική του μοναδική γεωμετρία και διάταξη αντικειμένων, με αποτέλεσμα το προφίλ των ανακλάσεων να αλλάζει ριζικά όταν το σύστημα μεταφέρεται σε ένα νέο, άγνωστο περιβάλλον. Στην συνέχεια οι Αλγόριθμοι Deep Learning (CNN, LSTM, Transformer) δεν χρησιμοποιούνται, καθώς θέτουν αυξημένες απαιτήσεις σε δεδομένα εκπαίδευσης, τα οποία δεν μπορούν να παραχθούν χρησιμοποιώντας το συγκεκριμένο κομμάτι hardware από μόνο ένα άτομο και δεν εκτελούνται εύκολα σε πραγματικό χρόνο (real-time) και ειδικά σε IoT συσκευές [10][12][26]. Παρά τους παραπάνω περιορισμούς, τα αποτελέσματα δείχνουν ότι η δυαδική ταξινόμηση με SISO hardware και Classical ML μπορεί να αποδώσει αξιόπιστα, επαναλήψιμα αποτελέσματα σε πραγματικό χρόνο.

1.6 Δομή Εργασίας

Το Κεφάλαιο 2 καλύπτει το θεωρητικό υπόβαθρο, πώς λειτουργεί το Wi-Fi σε επίπεδο φυσικού στρώματος (IEEE 802.11, OFDM), τι είναι το CSI και πώς εξάγεται. Το Κεφάλαιο 3 ανασκοπεί τη σχετική βιβλιογραφία και τοποθετεί την παρούσα εργασία στο ερευνητικό τοπίο. Το Κεφάλαιο 4 περιγράφει τον σχεδιασμό και την υλοποίηση, από την επιλογή hardware έως τη συλλογή δεδομένων και την πραγματική εκτέλεση του προγράμματος. Το Κεφάλαιο 5 παρουσιάζει τα αποτελέσματα, αξιολογεί τα μοντέλα ταξινόμησης και αναλύει τη σημαντικότητα των χαρακτηριστικών και παρουσιάζει ablation study σχετικά με όλες τις παραμέτρους της εργασίας και μελετά τους χρόνους επίδοσης του κάθε ταξινομητή. Τέλος το Κεφάλαιο 6 συνοψίζει τα συμπεράσματα, και προτείνει κατευθύνσεις για την μελλοντική επέκταση.

Κεφάλαιο 2ο: Θεωρητικό Υπόβαθρο

2.1 Εισαγωγή

Το κεφάλαιο αυτό παρουσιάζει το θεωρητικό υπόβαθρο που είναι απαραίτητο για την κατανόηση του συστήματος. Ξεκινά από τις αρχές λειτουργίας του Wi-Fi σε επίπεδο φυσικού στρώματος, με έμφαση στο πρότυπο IEEE 802.11 και στη διαμόρφωση OFDM, καθώς αυτά καθορίζουν τον τρόπο με τον οποίο το σήμα διαδίδεται στο χώρο και αλληλεπιδρά με τα κινούμενα αντικείμενα. Στη συνέχεια εξηγείται η έννοια του CSI, πώς εξάγεται από μια συσκευή και γιατί παρέχει πλουσιότερη πληροφορία σε σχέση με το RSSI. Τέλος, αναφέρεται και στους αλγορίθμους μηχανικής μάθησης όπου χωρίζονται σε 2 κατηγορίες, παρουσιάζοντας τα πλεονεκτήματα και τα μειονεκτήματα του καθενός.

2.2 Πρότυπο IEEE 802.11 και Πολυπλεξία OFDM

Το πρότυπο IEEE 802.11 αποτελεί τη βάση των σύγχρονων ασύρματων δικτύων Wi-Fi. Λειτουργεί στις ζώνες 2,4 GHz και 5 GHz με εμβέλεια δεκάδων μέτρων σε εσωτερικούς χώρους, η οποία είναι αρκετή ώστε για ένα μόνο ζεύγος συσκευών να καλύπτει ένα τυπικό διαμέρισμα [1]. Για τον υπολογισμό του μήκους κύματος μπορεί να χρησιμοποιηθεί ο θεμελιώδης τύπος της κυματικής (2.1) που ορίζεται παρακάτω ως εξής:

$$\lambda = \frac{c}{f} \quad (2.1)$$

όπου c είναι η ταχύτητα του φωτός και ισούται περίπου με $3 \times 10^8 \text{ m/s}$ και f η συχνότητα. Έτσι με την χρήση του τύπου το μήκος κύματος για την συχνότητα 2,4 GHz υπολογίζεται περίπου στα 12,5 εκατοστά (0,125 m).

Αυτό το μήκος κύματος (0,125m) παρουσιάζει τρία πλεονεκτήματα σχετικά με την ασύρματη ανίχνευση:

1. Διαπερατότητα Υλικών (Material Penetration): Τα κύματα των 2,4 GHz υφίστανται μικρότερη εξασθένηση όταν διαπερνούν διάφορα υλικά (όπως τοίχους, ξύλινες πόρτες κτλ.) σε σχέση με τα 5 GHz. Αυτό επιτρέπει την ανάπτυξη εφαρμογών ανίχνευσης και μέσα από τοίχους [22][3].
2. Αλληλεπίδραση με το Ανθρώπινο Σώμα: Καθώς το ανθρώπινο σώμα αποτελείται σε ποσοστό περίπου 60% από νερό, λειτουργεί ως ένα εξαιρετικό απορροφητικό και ανακλαστικό μέσο για τα ραδιοκύματα στη συχνότητα 2,4 GHz. Δηλαδή, μικρές αλλαγές στη στάση του σώματος προκαλούν άμεσες και μετρήσιμες διαταραχές στο ηλεκτρομαγνητικό πεδίο [1].
3. Κυματική Περίθλαση (Diffraction): Λόγω του μήκους κύματος, το σήμα έχει την ικανότητα να καμπυλώνει γύρω από αντικείμενα παρόμοιου ή ελαφρώς μικρότερου μεγέθους, επεκτείνοντας την εμβέλεια του συστήματος ακόμη και αν δεν υπάρχει άμεση οπτική επαφή (Non-Line-of-Sight - NLoS) [22][3].

Η κεντρική τεχνολογία διαμόρφωσης είναι η OFDM (Orthogonal Frequency-Division Multiplexing), η οποία χρησιμοποιείται στο Wi-Fi ήδη από τα πρότυπα 802.11a/g και διατηρείται από τη γενιά 802.11n (Wi-Fi 4) και μετά. Η ιδέα του OFDM είναι αντί να σταλούν όλα τα δεδομένα σε μία μόνο συχνότητα, διαιρείται το εύρος ζώνης του καναλιού σε πολλούς στενούς υποφορείς (subcarriers), καθένας από τους

οποίους μεταφέρει ανεξάρτητα ένα μέρος των δεδομένων. Για παράδειγμα, ένα κανάλι στα 20 MHz στο 802.11n (Wi-Fi 4) ορίζει 56 υποφορείς (52 για δεδομένα), ενώ για ένα κανάλι 40 MHz ορίζει 114 από τους συνολικούς 128. Οι εκδόσεις 802.11ac, 802.11ax (Wi-Fi 6), 802.11be (Wi-Fi 7) επεκτείνουν περαιτέρω και τις δύο αυτές παραμέτρους [4].

Σε ένα συμβατικό σύστημα πολλαπλών φερόντων, μεταξύ των υποφορέων αφήνονται κενά φάσματος για να μην παρεμβάλλονται μεταξύ τους και έτσι δημιουργείται σπατάλη του εύρους ζώνης. Το OFDM λύνει αυτό το πρόβλημα. Αντί για κενά, οι υποφορείς σχεδιάζονται έτσι ώστε η κορυφή του ενός να συμπίπτει ακριβώς με τα μηδενικά του διπλανού. Ο δέκτης διαβάζει κάθε υποφορέα στη δική του κορυφή, εκεί που οι υπόλοιποι έχουν ήδη μηδενιστεί, οπότε δεν υπάρχει παρεμβολή, παρόλο που τα σήματα αλληλεπικαλύπτονται. Αυτή η μαθηματική ιδιότητα είναι η ορθογωνιότητα, και επιτρέπει στο OFDM να συμπίσει περισσότερους υποφορείς στο ίδιο εύρος ζώνης αξιοποιώντας όλο το εύρος ζώνης.

Πιο αναλυτικά, η ορθογωνιότητα εκφράζεται ως (2.2) [3],[22]:

$$\int_0^T e^{j2\pi f_k t} e^{-j2\pi f_m t} dt = 0, \text{ για } k \neq m \quad (2.2)$$

Στη σχέση (2.2), το T είναι η διάρκεια ενός συμβόλου OFDM, η ορθογωνιότητα ισχύει όταν η διαφορά συχνοτήτων $f_k - f_m$ είναι ακέραιο πολλαπλάσιο του $1/T$, κάτι που εξασφαλίζεται επιλέγοντας απόσταση υποφορέων ίση με $1/T$. Παράλληλα, λόγω της επιλεκτικής ως προς τη συχνότητα εξασθένησης (frequency-selective fading), κάθε υποφορέας, έχοντας διαφορετική κεντρική συχνότητα, επηρεάζεται διαφορετικά από τις ανακλάσεις και διαθλάσεις που δημιουργούνται. Αντί λοιπόν, για έναν μόνο αριθμό που δείχνει την κατάσταση του καναλιού, το σύστημα παράγει έναν πίνακα μετρήσεων, δηλαδή μία τιμή ανά υποφορέα. Αυτός ο πίνακας είναι το Channel State Information [3].

Λόγω της διαίρεσης του φάσματος, το ευρυζωνικό κανάλι (Wideband) μετατρέπεται πλέον σε δεκάδες παράλληλα υποκανάλια στενής ζώνης (Narrowband Flat-Fading Channels). Η Απόκριση Συχνότητας Καναλιού (Channel Frequency Response - CFR), δηλαδή το CSI, δεν υπολογίζεται πλέον ως ένας μέσος όρος για όλο το εύρος, αλλά αξιολογείται ανεξάρτητα για κάθε έναν υποφορέα ξεχωριστά.

Εάν το σύστημα για παράδειγμα χρησιμοποιεί N υποφορείς (π.χ. στο 802.11n/20 MHz το FFT έχει 64, εκ των οποίων N=56 ενεργοί υποφορείς), τότε για ένα δεδομένο χρονικό στιγμιότυπο t, ο δέκτης εξάγει το διάνυσμα του CSI ως:

$$H(t) = [H_1(t), H_2(t), \dots, H_k(t), \dots, H_N(t)] \quad (2.3)$$

Κάθε στοιχείο H_k είναι ένας μιγαδικός αριθμός που περιγράφει την κατάσταση του k-οστού υποφορέα και δίνεται από τον τύπο [1]:

$$H_k = |H_k| e^{j \angle H_k} \quad (2.4)$$

Όπου:

$|H_k|$ αντιπροσωπεύει το Πλάτος (Amplitude), δηλαδή τη μείωση της έντασης που υπέστη η συγκεκριμένη συχνότητα.

$\angle H_k$ αντιπροσωπεύει τη Φάση (Phase), η οποία είναι ανάλογη της χρονικής καθυστέρησης της διαδρομής αυτού του υποφορέα.

Λόγω της πολυδιαδρομικής (multipath) φύσης του καναλιού, η CFR στην πράξη εκφράζεται ως άθροισμα μιγαδικών εκθετικών πάνω από L διαδρομές διάδοσης. Η πλήρης μαθηματική σχέση αναπτύσσεται στην υποενότητα 2.4, όπου εισάγεται πρώτα η Κρουστική Απόκριση Καναλιού (CIR) και στη συνέχεια προκύπτει η CFR ως ο μετασχηματισμός Fourier της [3], [19].

Η απόσταση μεταξύ γειτονικών υποφορέων (subcarrier spacing) δίνεται από τον παρακάτω τύπο (2.5), όπου B το συνολικό εύρος ζώνης και N ο συνολικός αριθμός υποφορέων (μέγεθος FFT, π.χ. 64 στα 20 MHz)

[22], [3]

$$\Delta f = \frac{B}{N} \quad (2.5)$$

Στην πράξη, η μετάβαση από το πεδίο του χρόνου στο πεδίο της συχνότητας στο OFDM υλοποιείται ψηφιακά μέσω του Αντίστροφου Γρήγορου Μετασχηματισμού Fourier (IFFT) στον πομπό και του Γρήγορου Μετασχηματισμού Fourier (FFT) στον δέκτη [3].

2.2.1 Διαμόρφωση Υποφορέων (Modulation) και Σχήματα MCS

Αφού το OFDM χωρίσει το κανάλι σε υποφορείς, κάθε υποφορέας μεταφέρει bits μέσω ενός σχήματος ψηφιακής διαμόρφωσης. Ανάλογα με την ποιότητα του καναλιού χρησιμοποιείται BPSK, QPSK, 16-QAM, 64-QAM και στα νεότερα πρότυπα έως και 256-QAM. Κάθε σχήμα αντιστοιχίζει τα bits σε ένα μιγαδικό σημείο αστερισμού (constellation point) [11]. Ένα σχήμα M-QAM μπορεί να μεταφέρει $\log_2(M)$ bits ανά σύμβολο, οπότε όσο μεγαλώνει το M αυξάνεται ο ρυθμός αλλά απαιτείται και μεγαλύτερο SNR.

Το πρότυπο προσαρμόζει δυναμικά τον συνδυασμό διαμόρφωσης και κωδικοποίησης μέσω του MCS (Modulation and Coding Scheme), ανάλογα με τις συνθήκες του καναλιού [22][3]. Για το Wi-Fi sensing όμως, αυτό που μετράται δεν είναι η διαμόρφωση των δεδομένων αλλά η απόκριση του καναλιού H , η οποία εκτιμάται από τα γνωστά σύμβολα εκπαίδευσης (LTF). Δηλαδή η διαμόρφωση των ωφέλιμων δεδομένων δεν επηρεάζει την εκτίμηση του CSI, αφού το H προκύπτει από τα γνωστά pilots.

Γι' αυτόν τον λόγο πολλά εργαλεία εξαγωγής CSI σταθεροποιούν ένα απλό, χαμηλής τάξης σχήμα. Το Wi-ESP, για παράδειγμα, εκπέμπει BPSK στα 6 Mbps χωρίς channel coding και scrambling, ακριβώς για να παίρνει καθαρό και επαναλήψιμο CSI [17]. Η ίδια λογική ακολουθήθηκε στην παρούσα εργασία, με τη διαφορά ότι το σχήμα διαμόρφωσης δεν μένει στην τύχη του αλλά ορίζεται ρητά μέσα από τον κώδικα σε MCS0, δηλαδή με BPSK και με ρυθμό κωδικοποίησης $1/2$. Είναι το πιο απλό και ανθεκτικό σχήμα, οπότε κάθε πακέτο ESPNOW φεύγει με σταθερή, χαμηλής τάξης διαμόρφωση και το CSI που διαβάζεται από τα training fields βγαίνει καθαρό και επαναλήψιμο. Παράλληλα κλειδώνεται και το AGC, όπως αναλύεται στο Κεφάλαιο 4. Σε πιο πρόσφατη κατεύθυνση, νέες κυματομορφές όπως το

OTFS προτείνονται για ISAC και 6G, όπου το ίδιο σήμα εξυπηρετεί ταυτόχρονα επικοινωνία και sensing [22], κάτι που συζητείται στο Κεφάλαιο 6.

2.3 Channel State Information (CSI)

Το CSI καταγράφει τη συνολική επίδραση όλων των φαινομένων διάδοσης όπως είναι η σκέδαση, εξασθένηση, απώλεια ισχύος λόγω απόστασης κατά τη μετάδοση μεταξύ πομπού και δέκτη [1].

Για κάθε υποφορέα i , το CSI εκφράζεται όπως ορίστηκε ήδη στις σχέσεις (2.3) και (2.4) ως μιγαδικός αριθμός H_i :

$$H_i = |H_i| \cdot e^{j\varphi_i} \quad (2.6)$$

όπου $|H_i|$ είναι το πλάτος (amplitude), δηλαδή πόσο εξασθένησε το σήμα σε αυτή τη συχνότητα και φ_i είναι η φάση (phase), η χρονική καθυστέρηση που εισάγει το κανάλι [3]. Ο συνολικός πίνακας για N υποφορείς γράφεται ως:

$$\mathbf{H} = [H_1, H_2, \dots, H_N]^T \quad (2.7)$$

Πιο συγκεκριμένα κάθε πακέτο Wi-Fi που στέλνει μια συσκευή περιέχει στην αρχή του τα γνωστά σύμβολα εκπαίδευσης (Long Training Symbols), τα οποία ο δέκτης αναμένει να λάβει με συγκεκριμένη μορφή. Στην πράξη όμως το σήμα φτάνει αλλοιωμένο, αφού έχει ταξιδέψει από τον πομπό στον δέκτη μέσω πολλαπλών διαδρομών (Multipath), μέσω ανακλάσεων από τοίχους, έπιπλα ή κινούμενα άτομα είτε απευθείας. Ο δέκτης συγκρίνει αυτό που έλαβε με αυτό που περίμενε και από τη διαφορά εκτιμά τον πίνακα $\mathbf{H}$, δηλαδή το CSI. Για κάθε υποφορέα ξεχωριστά, ο πίνακας $\mathbf{H}$ περιγράφει πόσο αλλοιώθηκε το σήμα σε πλάτος και σε φάση κατά τη διάδοση. Αν κάποιος κινηθεί στον χώρο, οι διαδρομές αλλάζουν, οι ανακλάσεις αλλάζουν και συνεπώς αλλάζει και το $\mathbf{H}$ (2.3). Αυτή ακριβώς την μεταβολή είναι που εκμεταλλεύεται το σύστημα για την ανίχνευση κίνησης [1][3][17].

Η σχέση εισόδου-εξόδου του ασύρματου καναλιού εκφράζεται απλά ως:

$$y = \mathbf{H} \cdot x + \eta \quad (2.8)$$

όπου x είναι το εκπεμπόμενο σήμα (γνωστό pilot), y το λαμβανόμενο σήμα, $\mathbf{H}$ ο πίνακας CSI, και η το θόρυβος. Η εξίσωση αυτή δείχνει ότι το CSI δεν μετρείται άμεσα, αλλά εκτιμάται μέσα από τη σύγκριση του αναμενόμενου και του πραγματικού σήματος [3][9][12][17] [21].

Στην ακατέργαστη μορφή του, το CSI εξάγεται από τον δέκτη ως ένα σύνολο μιγαδικών αριθμών, όπου ο κάθε υποφορέας περιγράφεται από ένα πραγματικό (real) και ένα φανταστικό (imaginary) μέρος. Βάσει αυτών των τιμών, το πλάτος (amplitude) του σήματος υπολογίζεται ως το μέτρο του μιγαδικού αριθμού, ενώ η φάση προκύπτει από το τόξο εφαπτομένης (μέσω της atan2, για πλήρη προσδιορισμό στο διάστημα $(-\pi, \pi]$). Στην παρούσα εργασία επιλέχθηκε να αξιοποιηθεί η πληροφορία του πλάτους και να αγνοηθεί η φάση (phase). Ο βασικός λόγος είναι ότι, στον εμπορικό εξοπλισμό Wi-Fi, η μέτρηση

της φάσης είναι αρκετά ευάλωτη λόγω του hardware. Πιο συγκεκριμένα, η έλλειψη απόλυτου συγχρονισμού ανάμεσα στα εσωτερικά ρολόγια (oscillators) του πομπού και του δέκτη προκαλεί συνεχείς και τυχαίες γραμμικές μετατοπίσεις, οι οποίες είναι γνωστές ως CFO (Carrier Frequency Offset) και SFO (Sampling Frequency Offset). Τα φαινόμενα αυτά αλλοιώνουν σε τέτοιο βαθμό τα δεδομένα της φάσης, που η εξαγωγή χρήσιμης πληροφορίας θα απαιτούσε μια απαιτητική διαδικασία καθαρισμού και βαθμονόμησης (phase sanitization). Αντίθετα, το πλάτος παρουσιάζει πολύ μεγαλύτερη ανθεκτικότητα απέναντι σε αυτά τα σφάλματα και μπορεί και αποτυπώνει με αξιοπιστία τις ενεργειακές διακυμάνσεις που προκαλεί η ανθρώπινη κίνηση στο χώρο. Στα συστήματα MIMO θα μπορούσε να χρησιμοποιηθεί η φάση μέσω του AoA (Angle of Arrival) και να συλλεχθεί αξιοποιήσιμη πληροφορία από την φάση η οποία είναι και ακόμα πιο ευαίσθητη από το πλάτος στην ανθρώπινη κίνηση. [9][19]

Τυπικά, για κάθε υποφορέα k και χρονική στιγμή t , η μετρούμενη φάση CSI μπορεί να αναλυθεί ως [3]:

$$\Phi_{meas}(k, t) = \Phi_{true}(k, t) + 2\pi \cdot \Delta f \cdot t + 2\pi \cdot k \cdot \eta + \theta_0 \quad (2.9)$$

όπου $\Phi_{true}(k, t)$ η πραγματική φάση από τα multipath φαινόμενα, Δf το Carrier Frequency Offset (CFO) λόγω ασυγχρόνιστων ρολογιών πομπού-δέκτη, ο όρος η που οφείλεται στο Sampling Frequency Offset (SFO) και μεταβάλλεται γραμμικά με τον δείκτη υποφορέα k , και θ_0 ο σταθερός όρος που περιλαμβάνει το packet detection delay και άλλες σταθερές μετατοπίσεις φάσης [1][3]. Στο σύστημα, εφόσον χρησιμοποιείται αποκλειστικά το πλάτος, οι συνιστώσες CFO/SFO δεν χρειάζεται να διορθωθούν. Στη σχέση (2.9) το Δf δηλώνει το CFO και το η τον όρο SFO. Τα σύμβολα αυτά αφορούν αποκλειστικά την πληροφορία της φάσης και, όπως αναφέρθηκε, δεν επηρεάζουν το πλάτος που αξιοποιείται στην παρούσα εργασία.

2.3.1 CSI έναντι RSSI

Ο RSSI (Received Signal Strength Indicator) συνοψίζει σε έναν μόνο αριθμό ανά πακέτο τη μέση ισχύ του εισερχόμενου σήματος. Είναι δηλαδή μια μέτρηση που δεν μπορεί να συλλάβει τις λεπτές αλλαγές από κίνηση ανθρώπινου σώματος, κυρίως όταν το άτομο δεν βρίσκεται μεταξύ πομπού και δέκτη [1]. Επιπλέον, ακόμη και σε στατικό περιβάλλον, εσωτερικές μεταβολές κατάστασης (ρυθμίσεις μετάδοσης, επίπεδα αναφοράς) αλλάζουν την τιμή RSSI, που την κάνει αναξιόπιστη [2].

Η σχέση του RSSI σε λογαριθμική κλίμακα είναι (2.10):

$$P(d) = P(d_0) - 10\gamma \log\left(\frac{d}{d_0}\right) + X_\delta \quad (2.10)$$

όπου $P(d)$ η λαμβανόμενη ισχύς σε απόσταση d (σε dB), $P(d_0)$ η λαμβανόμενη ισχύς στο σημείο αναφοράς d_0 , γ ο εκθέτης απώλειας διαδρομής, και X_δ μια τυχαία μεταβλητή μηδενικής μέσης τιμής που αντιπροσωπεύει τη σκίαση (log-normal shadowing) [1].

Ο RSSI λοιπόν παρουσιάζει τέσσερις αδυναμίες που τον κάνουν ακατάλληλο για HAR (Human Activity Recognition) :

- (α) αδυναμία διάκρισης στατικής παρουσίας από ενεργή κίνηση,
- (β) εξαιρετική ευαισθησία σε παρεμβολές και θόρυβο,
- (γ) αδυναμία ανίχνευσης λεπτών κινήσεων,
- (δ) μη αναπαραγώγιμα αποτελέσματα σε διαφορετικά περιβάλλοντα.

Το CSI, από την άλλη, παρέχει πλούσια πληροφορία, κάθε υποφορέας αισθάνεται διαφορετικά τα φαινόμενα πολυδιαδρομικής διάδοσης (Multipath), άλλοι ενισχύονται (επικοδομητική υπέρθεση) ενώ άλλοι εξασθενούν (καταστροφική υπέρθεση). Η κίνηση στον χώρο μεταβάλλει τη γεωμετρία των ανακλάσεων, αφήνοντας χαρακτηριστικά μοτίβα στον πίνακα CSI [1]. Αυτή είναι και η αρχή λειτουργίας του Wi-Fi sensing.

Ο πλήρης μαθηματικός τύπος για το CSI, που διαχωρίζει τη στατική από τη δυναμική συνιστώσα είναι [12]:

$$H(f, t) = H_S(f) + H_D(f, t) + n(f, t) \quad (2.11)$$

όπου $H_S(f)$ η στατική συνιστώσα (ανακλάσεις από ακίνητα αντικείμενα όπως τοίχοι και έπιπλα), $H_D(f, t)$ η δυναμική συνιστώσα (ανακλάσεις από κινούμενα σώματα), και $n(f, t)$ ο θόρυβος μέτρησης [12].

2.3.2 Πρόσβαση και Χρήση CSI

Σε εμπορικές συσκευές, η πρόσβαση στο CSI απαιτεί τροποποιήσεις υλικολογισμικού για την απόκτηση του CSI. Για το Intel 5300 NIC υπάρχει το Linux 802.11n CSI Tool, για τα Atheros chipsets το αντίστοιχο Atheros CSI Tool [2]. Για τον ESP32, εργαλεία όπως το Wi-ESP και το ESP32-CSI-Tool όπου επιτρέπουν την εξαγωγή του CSI σε πραγματικό χρόνο με κόστος μερικών ευρώ [9][10][17].

Intel 5300 αποδίδει CSI για 30 μόνο (ομαδοποιημένους) από τους 52 υποφορείς δεδομένων των 20 MHz, ενώ ο ESP32 μέσω Wi-ESP αποδίδει 128 υποφορείς (με 114 ενεργούς) [17][30][31]. Η φθηνότερη συσκευή παρέχει, δηλαδή, πλουσιότερη πληροφορία ανά πακέτο.

Ο υπολογισμός πλάτους και φάσης από τα μιγαδικά δεδομένα είναι ο εξής :

$$A(i) = \sqrt{H_r(i)^2 + H_{im}(i)^2}, \quad \varphi(i) = \text{atan2}(H_{im}(i), H_r(i)) \quad (2.12)$$

όπου $H_r(i)$ και $H_{im}(i)$ είναι το πραγματικό και φανταστικό μέρος αντίστοιχα [13], [19].

2.3.3 Φάση CSI και Εξυγίανση Φάσης (Phase Sanitization)

Όπως αναφέρθηκε, στην παρούσα εργασία χρησιμοποιείται μόνο το πλάτος και αγνοείται η φάση. Ο λόγος δεν είναι ότι η φάση δεν κρύβει πληροφορία, αλλά ότι στον εμπορικό εξοπλισμό η ακατέργαστη φάση είναι μη αξιοποιήσιμη. Το πρόβλημα προέρχεται από δύο πηγές, το Carrier Frequency Offset (CFO) και το Sampling Frequency Offset (SFO) [21][3].

Το CFO εμφανίζεται επειδή ο πομπός και ο δέκτης δεν συγχρονίζουν απόλυτα τα ρολόγια και τη φάση τους πριν από κάθε πακέτο. Στη ζώνη των 5 GHz, η διαφορά των ρολογιών μπορεί να μετατοπίσει τη φάση κατά αρκετά π, ενώ η μεταβολή που προκαλεί η κίνηση του ανθρώπου είναι συνήθως κάτω από 0,5π. Ουσιαστικά η πραγματική κίνηση καλύπτεται κάτω από το offset του ίδιου του υλικού, οπότε δεν φαίνεται καθόλου στην ακατέργαστη φάση [21].

Το SFO από την άλλη, προέρχεται από τον μετατροπέα αναλογικού σε ψηφιακό (ADC) και, σε αντίθεση με το CFO, διαφέρει ανά υποφορέα. Δηλαδή προσθέτει έναν όρο που μεταβάλλεται γραμμικά με τον δείκτη υποφορέα k , όπως φάνηκε και στη σχέση (2.9). Εφόσον ούτε το CFO ούτε το SFO είναι γνωστά εκ των προτέρων, η φάση όπως βγαίνει από τον δέκτη δεν προσφέρει αξιοποιήσιμη πληροφορία [21][3].

Υπάρχουν βέβαια τεχνικές εξυγίανσης φάσης (phase sanitization). Η πιο διαδεδομένη είναι ένας γραμμικός μετασχηματισμός κατά το μήκος των υποφορέων, που αφαιρεί τη γραμμική ως προς το k συνιστώσα και αφήνει ένα καθαρότερο σήμα, όπου η κίνηση γίνεται ξανά ορατή [3][21]. Σε διατάξεις MIMO, η εξυγιασμένη φάση μπορεί επιπλέον να τροφοδοτήσει εκτίμηση γωνίας άφιξης (AoA). Στην παρούσα SISO υλοποίηση επιλέχθηκε να αξιοποιηθεί μόνο το πλάτος, αφού το CFO/SFO δεν χρειάζεται να διορθωθεί και το πλάτος αποτυπώνει ήδη με ανθεκτικότητα τις ενεργειακές διακυμάνσεις της κίνησης. Έτσι, κάποια πληροφορία της φάσης παραμένει αναξιόποιτη, κάτι που σε ένα MIMO σύστημα θα άξιζε να ανακτηθεί.

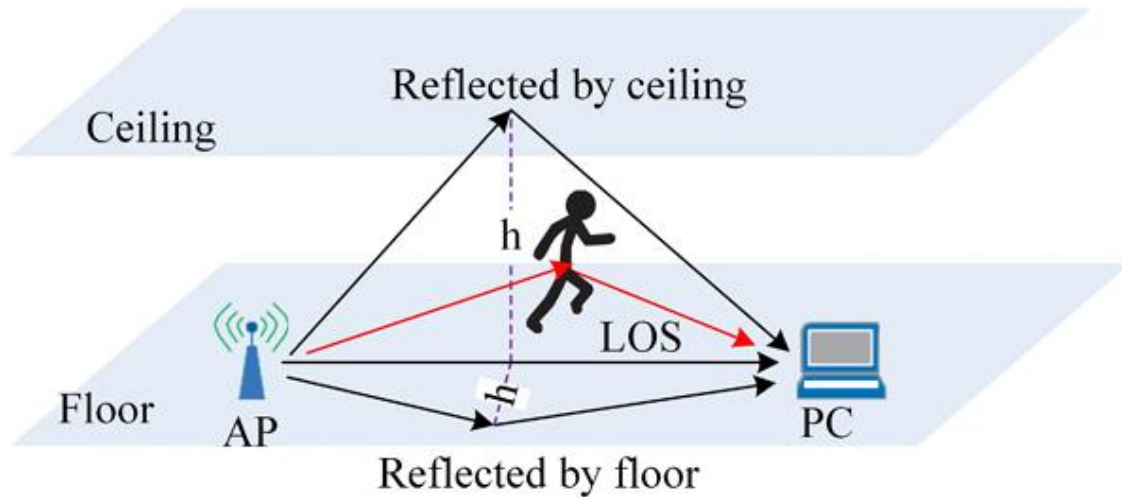
2.4 Φαινόμενο Πολυδιαδρομικής Διάδοσης (Multipath) και Ανθρώπινη Κίνηση

Σε οποιονδήποτε εσωτερικό χώρο, το σήμα Wi-Fi φτάνει στον δέκτη μέσα από πλήθος μονοπατιών είτε απευθείας (Line-of-Sight, LoS) και είτε έμμεσα (Non-Line-of-Sight, NLoS) μέσω ανακλάσεων σε τοίχους, οροφές και έπιπλα. Αυτό το φαινόμενο ονομάζεται πολυδιαδρομική διάδοση (multipath propagation) [1][2]. Παρακάτω παρουσιάζεται η έννοια του Multipath στην εικόνα 2.1.

Κατά τη μετάδοση του σήματος, τα ηλεκτρομαγνητικά κύματα έχουν 3 κύριες ιδιότητες [3][22][25]:

- Η ανάκλαση συμβαίνει όταν το κύμα συναντά επιφάνειες μεγάλες σε σχέση με το μήκος κύματός του, όπως τοίχοι και οροφή.
- Η περίθλαση εμφανίζεται στις ακμές αντικειμένων και επιτρέπει στο σήμα να φτάσει σε περιοχές που δεν έχουν άμεση οπτική επαφή με τον πομπό.

- Η σκέδαση προκαλείται από αντικείμενα με διαστάσεις συγκρίσιμες με το μήκος κύματος, όπως έπιπλα και ανθρώπινα σώματα, και μεταδίδει την ενέργεια προς τυχαίες κατευθύνσεις [22][2].



Εικόνα 2.1. Φαινόμενο Πολυδιαδρομικής Διάδοσης (Multipath) και Ανθρώπινη Κίνηση [2].

Το μοντέλο εξασθένησης εσωτερικού χώρου (indoor path loss model) δίνεται παρακάτω ως και πρόκειται για την ίδια σχέση με τη (2.10), γραμμένη με τον συμβολισμό της πηγής [1] :

$$PL(d) = PL(d_0) + 10n \cdot \log_{10} \left(\frac{d}{d_0} \right) + X_\sigma \quad (2.13)$$

όπου $PL(d)$ η εξασθένηση (dB) στην απόσταση d , $PL(d_0)$ η αναφορά στην $d_0 = 1 \text{ m}$, n ο εκθέτης εξασθένησης, και X_σ ο λογαριθμικο-κανονικός τυχαίος θόρυβος (shadowing) [1].

Η πλήρης μαθηματική περιγραφή της πολυδιαδρομικής διάδοσης στο πεδίο χρόνου-καθυστερήσης (time-delay domain) είναι η Κρουστική Απόκριση Καναλιού (Channel Impulse Response, CIR) [1], [3]:

$$h(t, \tau) = \sum_i a_i(t) e^{-j\theta_i(t)} \delta(\tau - \tau_i(t)) \quad (2.14)$$

όπου $a_i(t)$ η εξασθένηση της i -οστής διαδρομής, $\theta_i(t)$ η μετατόπιση φάσης, $\tau_i(t)$ η χρονική καθυστέρηση άφιξης, και δ η συνάρτηση δέλτα (Dirac) [3], [19].

Ο δέκτης τελικά καταγράφει το άθροισμα όλων των διαδρομών, οι οποίες ανάλογα με τη διαφορά φάσης τους είτε ενισχύονται είτε αλληλοεξασθενούν. Μαθηματικά, η CFR γράφεται ως [1][3]:

$$H(f, t) = \sum_{l=1}^L \alpha_l(t) \cdot e^{-j2\pi f \cdot d_l(t)/c} \quad (2.15)$$

όπου L το πλήθος μονοπατιών, $d_l(t)$ το μήκος του l -οστού μονοπατιού, και $\alpha_l(t)$ ένας μιγαδικός συντελεστής που αποτυπώνει το πλάτος και τη φάση κάθε μονοπατιού [1][3].

2.4.1 Μοντέλα Διαλείψεων και Συνοχή Καναλιού

Η πολυδιαδρομική διάδοση δεν περιγράφεται μόνο γεωμετρικά αλλά και στατιστικά, μέσω μοντέλων διαλείψεων (fading). Όταν δεν υπάρχει κυρίαρχη διαδρομή άμεσης οπτικής επαφής, το πλάτος ακολουθεί την κατανομή Rayleigh. Όταν υπάρχει μια ισχυρή απευθείας συνιστώσα μαζί με τις σκεδαζόμενες, το κατάλληλο μοντέλο είναι το Rician, που χρησιμοποιείται από αρκετές εργασίες Wi-Fi sensing για τον χαρακτηρισμό του καναλιού [3]. Στη δική μας διάταξη, που είναι κυρίως LoS, το Rician περιγράφει καλύτερα την πραγματικότητα, αφού συνυπάρχουν η ισχυρή ευθεία διαδρομή και οι ανακλάσεις από τοίχους και έπιπλα.

Δύο είναι τα μεγέθη που συνδέουν αυτά τα μοντέλα με την κίνηση. Ο χρόνος συνοχής (coherence time) είναι το διάστημα στο οποίο η κρουστική απόκριση του καναλιού παραμένει περίπου σταθερή και είναι αντιστρόφως ανάλογος της διασποράς Doppler, ενώ το εύρος συνοχής (coherence bandwidth) είναι αντιστρόφως ανάλογο της διασποράς καθυστέρησης, όπως συνοψίζεται στη σχέση (2.16). Όταν κάποιος κινείται, η διασπορά Doppler αυξάνεται, ο χρόνος συνοχής μικραίνει και το κανάλι αποσυσχετίζεται γρηγορότερα. Αυτή ακριβώς η ταχύτερη αποσυσχέτιση είναι το σήμα που αξιοποιείται [3].

$$T_c \approx \frac{1}{2f_D}, \quad B_c \approx \frac{1}{5\sigma_\tau} \quad (2.16)$$

Εφόσον το εύρος συνοχής είναι μικρότερο από το εύρος ζώνης του καναλιού, προκύπτει επιλεκτική ως προς τη συχνότητα, δηλαδή διαφορετικοί υποφορείς εξασθενούν διαφορετικά. Αυτός είναι, ο λόγος που το CSI ανά υποφορέα είναι πλουσιότερο από το RSSI, κάτι που αναπτύχθηκε ήδη στην ενότητα 2.2. Πρακτικά, αυτό εξηγεί γιατί ένα παράθυρο 1 δευτερολέπτου στα 100 Hz αρκεί για να αποτυπώσει τις διακυμάνσεις που προκαλεί το βάδισμα στο σήμα. Αντίθετα, τα ακίνητα αντικείμενα διατηρούν το σήμα σταθερό, οπότε μπορούν εύκολα να φιλτραριστούν ως στατικό multipath.

2.4.2 Ο ρόλος του ανθρώπινου σώματος

Όπως αναφέρθηκε στην ενότητα 2.2, το ανθρώπινο σώμα αποτελείται περίπου κατά 60% από νερό, άρα είναι και ηλεκτρικά αγώγιμο στις συχνότητες Wi-Fi και λειτουργεί ως ισχυρός σκεδαστής ραδιοκυμάτων [1]. Όταν ένα άτομο κινείται μέσα στην εμβέλεια ενός ζεύγους Wi-Fi, μεταβάλλει την απόσταση ορισμένων μονοπατιών, με αποτέλεσμα να υπάρχει ανακατανομή ενέργειας μεταξύ υποφορέων και μια μικρή μετατόπιση συχνότητας λόγω του φαινομένου Doppler [1][35].

Η μετατόπιση Doppler για κινούμενο σώμα [1][3] δίνεται παρακάτω ως:

$$\Delta f = \frac{2v \cos \theta}{c} \cdot f_c \quad (2.17)$$

όπου v η ταχύτητα του κινούμενου σώματος, θ η γωνία της κίνησης ως προς τη διαδρομή του σήματος, c η ταχύτητα φωτός και f_c η κεντρική συχνότητα του φέροντος. Επιπλέον, για το βάδισμα η μετατόπιση αυτή είναι μικρή, αλλά ανιχνεύσιμη με FFT [2][3].

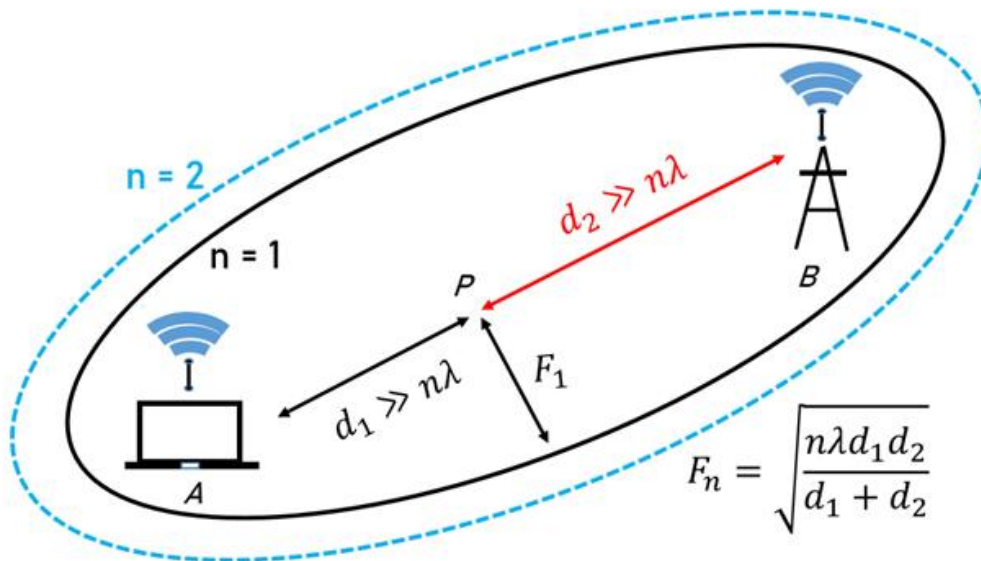
2.4.3 Ζώνες Fresnel

Η ένταση των μεταβολών CSI εξαρτάται άμεσα από το πού βρίσκεται το άτομο σε σχέση με τον πομπό και τον δέκτη. Γύρω από τη γραμμή ευθείας μεταξύ τους υπάρχει μια ελλειψοειδής περιοχή, η πρώτη ζώνη Fresnel, εντός της οποίας κάθε αντικείμενο επηρεάζει το σήμα έντονα, ενώ έξω από αυτή η επίδραση μειώνεται γρήγορα [3]. Αυτό σημαίνει ότι ένα άτομο που κινείται κοντά στον άξονα Tx–Rx διαταράσσει περισσότερες ανακλαστικές διαδρομές από κάποιον που βρίσκεται στη γωνία του δωματίου, οπότε το CSI αντιδρά πιο έντονα.

Η ακτίνα της n -οστής ζώνης Fresnel σε απόσταση d από τον πομπό [3]:

$$r_n(d) = \sqrt{\frac{n \cdot \lambda \cdot d \cdot (D - d)}{D}} \quad (2.18)$$

όπου λ το μήκος κύματος, D η συνολική απόσταση πομπού–δέκτη, d η απόσταση του ατόμου από τον πομπό, και n ο αριθμός ζώνης [3][12]. Στην εικόνα 2.2 παρατηρείται η ζώνη Fresnel.



Εικόνα 2.2. Ζώνη Fresnel [12].

2.5 Θεωρητικοί Περιορισμοί Κεραιών: SISO Έναντι MIMO και Χωρική Ανάλυση

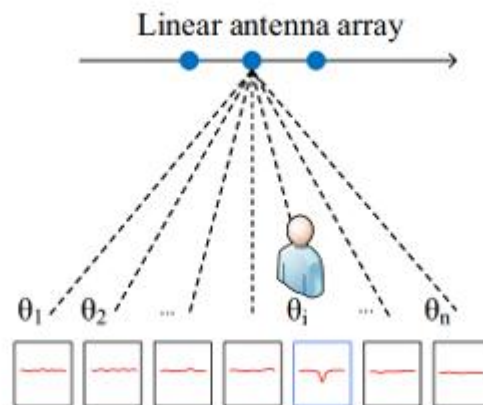
Σε ένα SISO σύστημα όλα τα μονοπάτια διάδοσης αθροίζονται σε έναν μόνο πίνακα CSI, οπότε ο δέκτης βλέπει τη συνολική μεταβολή του καναλιού χωρίς να μπορεί να πει από πού προήλθε. Δεν υπάρχει χωρική πληροφορία (spatial diversity), δεν υπάρχει Angle of Arrival (AoA), οπότε το σύστημα απαντά στο ερώτημα εάν υπάρχει κίνηση και σταματά εκεί [4]. Αυτός είναι και ο λόγος που συνήθως η αρχιτεκτονική κεραιών καθορίζει τι μπορεί να κάνει ένα Wi-Fi sensing σύστημα.

Για μια λεπτομερή αναγνώριση κίνησης του ανθρώπου (fine-grained HAR), όπως η διάκριση βηματισμού, ανύψωση χεριού ή αναγνώριση δέκα δραστηριοτήτων, χρειάζεται MIMO. Ένα 3×3 σύστημα με 56 υποφορείς δίνει θεωρητικά έως 504 ανεξάρτητες μετρήσεις CSI ανά πακέτο, αρκετές για την εκτίμηση της γωνίας άφιξης (AoA – Angle of Arrival) και την λεπτομερή ανάλυση κίνησης [5][2]. Για παράδειγμα, αν το σύστημα διαθέτει N_T κεραίες εκπομπής, N_R λήψης και K υποφορείς, ο πίνακας CSI τότε έχει διαστάσεις $N_T \times N_R \times K$ [3]. Το ενδιαφέρον πάντως είναι ότι ακόμα και τα MIMO συστήματα δεν λύνουν το πρόβλημα γενίκευσης σε νέα περιβάλλοντα. Η χωρική υπογραφή κάθε χώρου είναι μοναδική, και αναφέρεται ότι ακόμα και αλλαγή θέσης επίπλων στον ίδιο χώρο μπορεί να ρίξει την ακρίβεια αρκετά [6].

Γενικά, για ομοιόμορφη γραμμική συστοιχία (Uniform Linear Array, ULA) με δύο γειτονικές κεραίες σε απόσταση d , η διαφορά φάσης $\Delta\phi$ που μετρείται μεταξύ τους συνδέεται με τη γωνία άφιξης θ του σήματος σε σχέση με την κάθετο της συστοιχίας μέσω της σχέσης [2][3]:

$$\Delta\phi = \frac{2\pi d \sin(\theta)}{\lambda} \quad (2.19)$$

όπου λ το μήκος κύματος και d η απόσταση μεταξύ γειτονικών κεραιών (συνήθως $d = \lambda/2$). Από αυτή τη σχέση η γωνία θ προκύπτει αντιστρόφως ως $\theta = \arcsin\left(\frac{\lambda \cdot \Delta\phi}{2\pi d}\right)$ [2][3]. Στην εικόνα 2.3. παρατηρείται το συγκεκριμένο γεγονός.



Εικόνα 2.3. AoA (Angle of Arrival) [2]

2.6 Αλγόριθμοι Μηχανικής Μάθησης

Η αναγνώριση δραστηριότητας από CSI δεδομένα μπορεί να αποτελέσει και ένα πρόβλημα ταξινόμησης. Δηλαδή προκύπτει ένα διάνυσμα χαρακτηριστικών που προέρχεται από το σήμα και ζητούμενο είναι ο προσδιορισμός του αν αντιστοιχεί σε κίνηση ή απουσία κίνησης, ή σε μία από N συγκεκριμένες δραστηριότητες. Γενικά επικρατούν δύο μεγάλες οικογένειες αλγορίθμων, η παραδοσιακή (κλασική) μηχανική μάθηση και η βαθιά μάθηση [5][3]. Η διαφορά τους αφορά κυρίως τον τρόπο που αντιμετωπίζουν τα χαρακτηριστικά.

Στην παραδοσιακή προσέγγιση, ο χρήστης εξάγει χειροκίνητα ένα σύνολο χαρακτηριστικών από το σήμα (στατιστικά, φασματικά, μορφολογικά), και ένας ταξινομητής μαθαίνει να αντιστοιχίζει αυτά τα χαρακτηριστικά σε κλάσεις [5]. Στη βαθιά μάθηση, το ίδιο το δίκτυο αναλαμβάνει και την εξαγωγή χαρακτηριστικών και την ταξινόμηση, μέσα από πολλαπλά επίπεδα μη γραμμικών μετασχηματισμών [3]. Η επιλογή ανάμεσά τους εξαρτάται από τον όγκο των διαθέσιμων δεδομένων και τους υπολογιστικούς πόρους.

2.6.1 Παραδοσιακοί Αλγόριθμοι Μάθησης

Ένα δενδρικό μοντέλο χωρίζει διαδοχικά τον χώρο των χαρακτηριστικών μέσω κατωφλίων (threshold). Κάθε εσωτερικός κόμβος ρωτάει αν ένα συγκεκριμένο χαρακτηριστικό ξεπερνά κάποιο όριο, και η απάντηση κατευθύνει το επόμενο επίπεδο, μέχρι να φτάσει σε ένα «φύλλο» που αντιστοιχεί σε κάποια κλάση [5]. Το βασικό πλεονέκτημα είναι ότι μπορούν να αναλυθούν εύκολα, δηλαδή μπορεί κανείς να δει με βάση ποιο χαρακτηριστικό πήρε την απόφαση. Το Random Forest επεκτείνει αυτή την ιδέα και αντί για ένα δέντρο, εκπαιδεύει πολλά πάνω σε διαφορετικά τυχαία υποσύνολα των δεδομένων και χαρακτηριστικών, ώστε στο τέλος να ψηφίζουν όλα μαζί για την τελική απόφαση. Αυτή η στρατηγική κάνει το μοντέλο πιο σταθερό απέναντι σε θόρυβο [5][28].

Το Support Vector Machine (SVM) χωρίζει τις κλάσεις αναζητώντας το Hyperplane με το μέγιστο περιθώριο διαχωρισμού στον χώρο των χαρακτηριστικών [5]. Στη βασική του μορφή είναι γραμμικός ταξινομητής, όμως με τη χρήση μη γραμμικών πυρήνων (kernel trick), όπως ο πυρήνας RBF που χρησιμοποιείται στην παρούσα εργασία, μπορεί να διαχωρίσει και μη γραμμικά διαχωρίσιμες κλάσεις. Στο CSI HAR ο SVM χρησιμοποιήθηκε από τους πρώτους ερευνητές για ανίχνευση πτώσης (WiFall) με ακρίβεια (precision) 90% (one-class SVM, έως 94% με Random Forest) [1][33].

Ο K-Nearest Neighbors (k-NN) δεν χτίζει στην ουσία κάποιο μοντέλο. Αποθηκεύει όλα τα δείγματα εκπαίδευσης και, όταν φτάσει ένα νέο δείγμα, το ταξινομεί κοιτάζοντας τους k κοντινότερους γείτονές του [5][3]. Είναι απλός στην υλοποίηση, αν και ευαίσθητος στον θόρυβο για μικρές τιμές του k , και ταυτόχρονα απαιτητικός κατά την πρόβλεψη, αφού πρέπει να υπολογίσει την απόσταση από κάθε δείγμα εκπαίδευσης. Παρόλα αυτά σε CSI δεδομένα ανταποκρίνεται αρκετά καλά σε απλή δυαδική ταξινόμηση αλλά δυσκολεύεται καθώς αυξάνονται τα χαρακτηριστικά όπως θα παρατηρήσουμε και στο κεφάλαιο 5.

Οι αλγόριθμοι όπως το Gradient Boosting συνδυάζουν πολλούς αδύναμους ταξινομητές για να φτιάξουν έναν ισχυρότερο. Σε αντίθεση με το Random Forest, που εκπαιδεύει τα δέντρα του ανεξάρτητα και παράλληλα (bagging), το Gradient Boosting τα χτίζει διαδοχικά (boosting) δηλαδή κάθε νέο δέντρο μαθαίνει να διορθώνει τα σφάλματα που έχουν κάνει τα προηγούμενα δέντρα [5].

2.6.2 Βαθιά Μάθηση (Deep Learning)

Τα βαθιά νευρωνικά δίκτυα διαφέρουν από την παραδοσιακή προσέγγιση σε ένα κεντρικό σημείο, δεν χρειάζονται χειροκίνητη εξαγωγή χαρακτηριστικών [3] και δουλεύουν απευθείας στο ακατέργαστο ή ελάχιστα προεπεξεργασμένο σήμα και μαθαίνουν μόνα τους ποια μοτίβα είναι χρήσιμα. Παρόλα αυτά χρειάζονται, μεγάλο όγκο δεδομένων εκπαίδευσης, σημαντικούς υπολογιστικούς πόρους (συνήθως GPU), και η εκπαίδευση μπορεί να διαρκέσει ώρες ή και μέρες. [3]. Παρακάτω στην εικόνα 2.4 παρουσιάζεται η δομή ενός βαθιού νευρωνικού δικτύου.

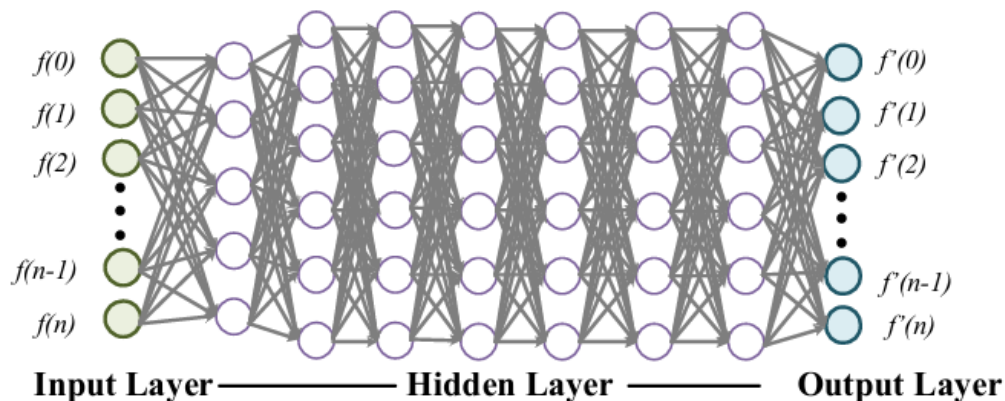
Για παράδειγμα, ένα Deep Neural Network (DNN) αποτελείται από πολλά πλήρως συνδεδεμένα επίπεδα (fully connected layers). Η έξοδος του i -οστού επιπέδου δίνεται από τον παρακάτω τύπο ως:

$$y(i) = g(W(i) \cdot x(i) + b(i)) \quad (2.20)$$

όπου W και b οι παράμετροι βαρών και bias, και g η συνάρτηση ενεργοποίησης [3].

Τα Convolutional Neural Networks (CNN) σχεδιάστηκαν αρχικά για εικόνες, αλλά πλέον εφαρμόζονται με επιτυχία και στο CSI. Το βασικό τους εργαλείο είναι το convolutional layer, που εφαρμόζει μικρά φίλτρα πάνω από την είσοδο και ανιχνεύει τοπικά μοτίβα [3][29]. Στα CSI δεδομένα υπάρχουν δύο τρόποι χρήσης. Είτε εφαρμόζεται 1D CNN απευθείας πάνω στη χρονοσειρά κάθε υποφορέα, είτε το CSI μετατρέπεται σε φασματόγραμμα (spectrogram 2D αναπαράσταση χρόνου-συχνότητας) και αντιμετωπίζεται σαν εικόνα [12].

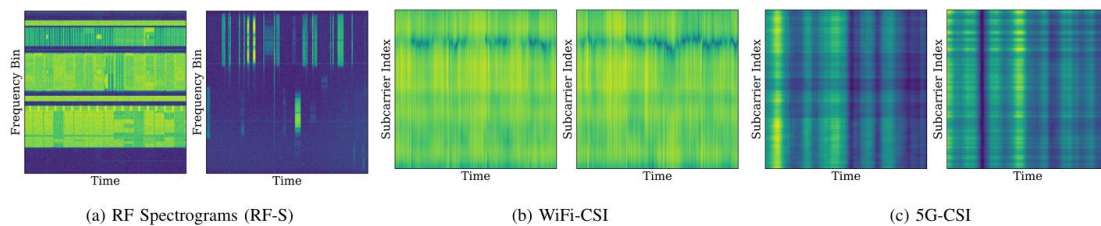
Τα Recurrent Neural Networks (RNN) επεξεργάζονται ακολουθιακά δεδομένα διατηρώντας μια εσωτερική κατάσταση (hidden state) που ενημερώνεται σε κάθε χρονικό βήμα [3]. Για το CSI αυτό έχει νόημα, μιας και το σήμα είναι από τη φύση του χρονοσειρά. Το πρόβλημα των απλών RNN είναι ότι η πληροφορία από μακρινά χρονικά βήματα χάνεται κατά την εκπαίδευση. Το LSTM (Long Short-Term Memory) από την άλλη αντιμετωπίζει το ζήτημα με τη χρήση gates (forget, input, output) που ελέγχουν τι κρατιέται στη μνήμη και τι ξεχνιέται [3]. Τέλος το Bidirectional LSTM (BLSTM) επεξεργάζεται την ακολουθία και προς τις δύο κατευθύνσεις (μνήμη μπρος και πίσω) [8].



Εικόνα 2.4. Ένα Βαθύ Νευρωνικό Δίκτυο [18].

Τα Vision Transformers (ViT) είναι η πιο πρόσφατη οικογένεια βαθιών αλγορίθμων που έχει εμφανιστεί στο CSI HAR [11][29]. Η μετάβαση σε αυτά προέκυψε κυρίως από τους περιορισμούς των προηγούμενων μοντέλων, ενώ τα συνελκτικά δίκτυα (CNNs) είναι άριστα στον εντοπισμό τοπικών χαρακτηριστικών, δυσκολεύονται να συλλάβουν συσχετίσεις σε όλη την έκταση του σήματος. Αντίστοιχα, τα αναδρομικά δίκτυα (όπως τα RNNs/LSTMs) έχουν την τάση να χάνουν πληροφορία όταν προσπαθούν να διαχειριστούν μακροπρόθεσμες εξαρτήσεις σε μεγάλες χρονικές ακολουθίες [4].

Η βασική ιδέα των ViT προέρχεται από την επεξεργασία φυσικής γλώσσας και βασίζεται στον μηχανισμό αυτο-προσοχής (self-attention). Μέσω αυτού του μηχανισμού, το δίκτυο υπολογίζει ταυτόχρονα τη βαρύτητα κάθε σημείου του σήματος σε σχέση με όλα τα υπόλοιπα. Έτσι, το μοντέλο καταφέρνει να παρατηρεί και να συσχετίζει μακρινά χωρικά και χρονικά σημεία σε μία μόνο πράξη, καθώς αντιλαμβάνεται πλήρως το πώς η κίνηση ενός ανθρώπου επηρεάζει ταυτόχρονα διαφορετικούς υποφορείς μέσα στον χρόνο [11]. Στην παρακάτω εικόνα (Εικόνα 2.5) παρουσιάζεται η μορφή των φασματογραμμάτων, δηλαδή της αναπαράστασης του σήματος στο πεδίο χρόνου–συχνότητας [11].



Εικόνα 2.5. Φασματόγραμμα (αναπαράσταση χρόνου–συχνότητας) [29].

2.6.3 Σύγκριση και Κριτήρια Επιλογής

Οι κλασικοί αλγόριθμοι έχουν χρόνους εκπαίδευσης σε δευτερόλεπτα ή λεπτά (αντί για ώρες), τα μεγέθη των μοντέλων είναι σε KB (αντί για MB ή GB), και απαιτούν σαφώς λιγότερα δεδομένα [5]. Για edge συσκευές χωρίς GPU, ή σε περιπτώσεις περιορισμένου dataset, αποτελούν την καλύτερη επιλογή με το μόνο αρνητικό ότι ο χρήστης θα πρέπει να επιλέξει ποια χαρακτηριστικά θα χρησιμοποιήσει.

Τα βαθιά δίκτυα, από την άλλη, αυτοματοποιούν την εξαγωγή χαρακτηριστικών και λειτουργούν καλύτερα με αρκετό όγκο δεδομένων [3][29]. Το αρνητικό με τα μοντέλα αυτά είναι ότι μπορούν να αποστηθίσουν πλήρως τα δεδομένα και έτσι ουσιαστικά δεν έχουν μάθει να αναγνωρίζουν μοτίβα αλλά έχουν μάθει να αποστηθίζουν (περιβάλλον, καταστάσεις) [5].

Για το συγκεκριμένο πρόβλημα της παρούσας εργασίας (binary classification σε δεδομένα ESP32 SISO, με περιορισμένο dataset και απόκριση σε πραγματικό χρόνο) η επιλογή των παραδοσιακών αλγορίθμων είναι προτιμότερη.

2.7 Θεωρητικό Υπόβαθρο Επεξεργασίας Σήματος

Πριν την υλοποίηση του Κεφαλαίου 4, αναλύεται η θεωρία της προεπεξεργασίας. Το φίλτρο Hampel στηρίζεται σε δύο εκτιμητές, τη διάμεσο και τη Median Absolute Deviation (MAD). Η λογική είναι ότι ένα spike από packet error θα μετατόπιζε έντονα τη μέση τιμή και την τυπική απόκλιση, ενώ η διάμεσος μένει σχεδόν ανεπηρέαστη. Ο συντελεστής 1,4826 κάνει την MAD συνεπή εκτιμητή της τυπικής απόκλισης για γκαουσιανά δεδομένα, με το κατώφλι αποκοπής να δίνεται στη σχέση (2.21) [9].

$$\hat{\sigma} = 1,4826 \cdot MAD, outlier \text{ αν } |x_i - m_i| > k \cdot \hat{\sigma} (k = 3) \quad (2.21)$$

Επίσης, το φίλτρο Butterworth επιλέγεται επειδή έχει μέγιστα επίπεδη απόκριση πλάτους στη ζώνη διέλευσης, χωρίς κυματώσεις που θα αλλοίωναν το σχήμα της κυματομορφής του βαδίσματος. Η απόκριση πλάτους για τάξη n και συχνότητα αποκοπής ω_c δίνεται στη σχέση (2.22). Στην υλοποίηση χρησιμοποιείται μηδενικής φάσης εφαρμογή (forward-backward, sosfiltfilt), ώστε να μην εισάγεται χρονική μετατόπιση στις κορυφές [15][3].

$$|H(j\omega)|^2 = \frac{1}{1 + \left(\frac{\omega}{\omega_c}\right)^{2n}}, \quad n = 4, \quad \omega_c = 2\pi \cdot 10 \text{ rad/s} \quad (2.22)$$

Σε μαθηματικό επίπεδο, το PCA εξάγει τα πιο σημαντικά χαρακτηριστικά των δεδομένων αναλύοντας τον πίνακα συνδιακύμανσης, όπως φαίνεται στη σχέση (2.23). Υπολογίζεται ο πίνακας C , εξάγονται οι κύριες συνιστώσες του και τελικά διατηρούνται οι 10 πιο σημαντικές, δηλαδή αυτές που συγκεντρώνουν τη μεγαλύτερη πληροφορία. Έτσι οι 114 υποφορείς συμπίεζονται σε 10 συνιστώσες που κρατούν το μεγαλύτερο μέρος της διακύμανσης [2][3][35]. Το ποσοστό που διατηρείται (από 80% σε ακινησία έως 98% σε βάδιση) αναλύεται στο Κεφάλαιο 4.

$$C = \frac{1}{N-1} X^T X, \quad C v_i = \lambda_i v_i \quad (2.23)$$

Τέλος, η κανονικοποίηση z-score φέρνει κάθε χαρακτηριστικό σε μέση τιμή 0 και τυπική απόκλιση 1. Αυτό έχει σημασία κυρίως για τους ταξινομητές που εξαρτώνται από αποστάσεις ή κλίμακα (SVM, k-NN, Logistic Regression), ώστε να μην κυριαρχούν τα χαρακτηριστικά με μεγάλες τιμές. Η προσαρμογή του scaler και του PCA γίνεται μόνο στα δεδομένα εκπαίδευσης, για να μη διαρρεύσει πληροφορία από το σύνολο ελέγχου, κάτι που εφαρμόζεται στο Κεφάλαιο 4.

2.8 Θεωρητικό Υπόβαθρο Χαρακτηριστικών και Μετρικών Αξιολόγησης

Στους κλασικούς αλγορίθμους, σε αντίθεση με τα βαθιά δίκτυα, τα χαρακτηριστικά εξάγονται χειροκίνητα. Χωρίζονται σε τρεις οικογένειες, στατιστικά πεδίου χρόνου, φασματικά (FFT) και χρονικά ή μορφολογικά. Ορισμένα έχουν συγκεκριμένη θεωρητική βάση. Οι παράμετροι Hjorth, για παράδειγμα, περιγράφουν την κινητικότητα (mobility) και την πολυπλοκότητα (complexity) ενός

σήματος, ενώ η φασματική εντροπία μετρά πόσο κανονικό ή τυχαίο είναι το φάσμα, όπως φαίνεται στη σχέση (2.24). Τα χαρακτηριστικά αυτά εφαρμόζονται αναλυτικά στο Κεφάλαιο 4 (Πίνακας 4.1) [3][5].

$$\text{Mobility} = \frac{\text{std}(x')}{\text{std}(x)}, \quad \text{Spectral Entropy} = - \sum_i p_i \log_2(p_i) \quad (2.24)$$

Από την πλευρά της αξιολόγησης, η συνολική ακρίβεια από μόνη της μπορεί να παραπλανήσει, ειδικά όταν τα σφάλματα δεν είναι συμμετρικά. Για αυτό χρησιμοποιούνται το precision και το recall, ενώ το F1-Score τα συνδυάζει ως αρμονικό μέσο, όπως φαίνεται στη σχέση (2.25). Σε ένα σύστημα HAR μια χαμένη κίνηση (False Negative) συνήθως κοστίζει περισσότερο από έναν ψευδή συναγερμό (False Positive), οπότε η ισορροπία τους μετράει [13][28].

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}, \quad F1 = 2 \frac{PR}{P + R} \quad (2.25)$$

Η καμπύλη ROC και το AUC, που αναλύονται στο Κεφάλαιο 5, δίνουν μια εικόνα ανεξάρτητη από το συγκεκριμένο κατώφλι απόφασης. Όσο για τη μεθοδολογία ελέγχου, το k-fold cross-validation μειώνει την εξάρτηση από έναν μόνο τυχαίο διαχωρισμό. Στη δική μας περίπτωση χρησιμοποιείται StratifiedGroupKFold, που κρατά ισορροπημένη την αναλογία κλάσεων και ταυτόχρονα δεν αφήνει παράθυρα της ίδιας καταγραφής να διαρρεύσουν ανάμεσα στα folds, αποτρέποντας τη μορφή data leakage που εντόπισε ο Varga [5][7].

2.9 Σύνοψη του κεφαλαίου

Σε αυτό το κεφάλαιο αναλύθηκε το θεωρητικό υπόβαθρο πάνω στο οποίο χτίζεται το υπόλοιπο της εργασίας. Αναλύθηκε πώς η τεχνολογία OFDM επιτρέπει στο Wi-Fi να μετρά την απόκριση του καναλιού ανά υποφορέα (CSI), και γιατί η πληροφορία αυτή είναι πλουσιότερη από το RSSI. Αναλύθηκαν τα φαινόμενα multipath και Doppler, καθώς και ο ρόλος των ζωνών Fresnel στο πώς η ανθρώπινη κίνηση αφήνει το αποτύπωμά της στο σήμα. Στη συνέχεια εξετάστηκαν οι περιορισμοί του εξοπλισμού, σε ένα SISO σύστημα όπως στο δικό μας δεν υπάρχει spatial diversity ούτε δυνατότητα AoA, οπότε η ανάλυση περιορίζεται αποκλειστικά στο πλάτος του σήματος.

Στις τελευταίες ενότητες έγινε η αναφορά των δύο μεγάλων κατηγοριών αλγορίθμων μάθησης. Οι κλασικοί ταξινομητές (SVM, k-NN, Decision Trees, Random Forest) οι οποίοι απαιτούν χειροκίνητη εξαγωγή χαρακτηριστικών αλλά είναι ελαφρείς και κατάλληλοι για edge συσκευές [5]. Τα βαθιά δίκτυα (CNN, LSTM, ViT) αυτοματοποιούν την εξαγωγή χαρακτηριστικών αλλά κοστίζουν σε δεδομένα και υπολογιστική ισχύ [3].

Κεφάλαιο 3ο: Βιβλιογραφική Ανασκόπηση

3.1 Εισαγωγή

Το κεφάλαιο αυτό κάνει μια ανασκόπηση της βιβλιογραφίας γύρω από την αναγνώριση ανθρώπινης δραστηριότητας (HAR) με χρήση Wi-Fi Channel State Information. Η παρουσίαση ξεκινά από τις πρώτες ιδέες της δεκαετίας του 2000, περνά στην συνέχεια στην εποχή των ερευνητικών συστημάτων που βασίστηκαν στην κάρτα Intel 5300, και φτάνει στις σημερινές υλοποιήσεις χαμηλού κόστους πάνω σε μικροελεγκτές ESP32 και πλατφόρμες όπως το Raspberry Pi. Στόχος δεν είναι να καταγραφεί απλώς η ακρίβεια κάθε συστήματος, αλλά να φανεί τι έχει αλλάξει στην πράξη, ποιο υλικό λειτουργεί έξω από το εργαστήριο, ποιοι αλγόριθμοι γενικεύουν όταν αλλάζει το άτομο ή ο χώρος, και πού εξακολουθούν να υπάρχουν αδυναμίες. Παρουσιάζονται επίσης τα πλεονεκτήματα της τεχνολογίας ως προς την ιδιωτικότητα και οι περιορισμοί που έχουν εντοπιστεί σε πρόσφατες κριτικές αναλύσεις της βιβλιογραφίας.

3.2 Η Εξέλιξη του CSI-based HAR (2007–2026)

Η ιδέα της χρήσης ασύρματων σημάτων για ανίχνευση ανθρώπινης παρουσίας χωρίς φορετές συσκευές χρονολογείται τουλάχιστον από το 2007. Ο Youssef et al. [23] εισήγαγαν τότε την έννοια Device-free Passive (DfP) localization, ένα σύστημα όπου εντοπίζει και παρακολουθεί άτομα χωρίς αυτά να φέρουν συσκευή, βασιζόμενο στις μεταβολές της ραδιοσυχνотικής διάδοσης. Η εργασία αυτή θεωρείται από τα θεμέλια του πεδίου αυτού. Τα πρώτα συστήματα στηρίζονταν αποκλειστικά στον δείκτη RSSI και αναγνώριζαν κυρίως μεγάλες αλλαγές ισχύος όταν ένα άτομο κινούνταν μεταξύ πομπού και δέκτη [23].

Λίγα χρόνια αργότερα δύο εργασίες έδειξαν τι παραπάνω μπορεί να γίνει με ασύρματα σήματα. Το WiSee (2013) εκμεταλλεύτηκε μετατοπίσεις Doppler πάνω σε σήματα OFDM για να αναγνωρίσει χειρονομίες, ενώ το WiTrack επέτρεψε παρακολούθηση κίνησης σε τρεις διαστάσεις, χρησιμοποιώντας ωστόσο εξειδικευμένο FMCW πομποδέκτη και όχι τον κλασικό εμπορικό εξοπλισμό 802.11 [1]. Το RSSI όμως ως ενιαία τιμή ανά πακέτο αποδείχθηκε ασταθές και ευαίσθητο σε multipath fading, και γι' αυτό η κοινότητα στράφηκε στην πιο πλούσια πληροφορία του Channel State Information [1].

Η μεγάλη αλλαγή ήρθε με τη διαθεσιμότητα εργαλείων εξαγωγής CSI από εμπορικές συσκευές. Το Linux 802.11n CSI Tool των Halperin et al. για το chipset Intel 5300 NIC άνοιξε τον δρόμο για νέα συστήματα που εκμεταλλεύονται την φασματική πληροφορία ανά OFDM υποφορέα [3]. Λίγο αργότερα το Atheros CSI Tool έδωσε αντίστοιχη δυνατότητα και σε άλλα chipsets. Με αυτά τα εργαλεία ως βάση δημοσιεύτηκαν πολλά αναγνωρίσιμα συστήματα [1][2][3]:

- E-eyes (2014): Αναγνώριση καθημερινών δραστηριοτήτων (ύπνος, μαγείρεμα, παρακολούθηση τηλεόρασης) χρησιμοποιώντας τη συσσωρευμένη διακύμανση των δειγμάτων CSI και Dynamic Time Warping για τις δραστηριότητες βάδισης, ενώ για τις στατικές δραστηριότητες εφαρμόστηκε Earth Mover's Distance [1][34].
- WiFall (2016–2017): Ανίχνευση πτώσης με one-class SVM και Random Forest για τον εντοπισμό μη φυσιολογικών παραθύρων CSI και την τελική ταξινόμηση, με precision 90% με one-class SVM (έως 94% με Random Forest) και χαμηλό ρυθμό ψευδών συναγεργμών (15% και 13% αντίστοιχα) [1][33].
- CARM (2015): Μοντέλο ταχύτητας CSI του Wang et al. που συσχετίζει κινήσεις σώματος με τις μεταβολές του σήματος, χρησιμοποιώντας PCA και Discrete Wavelet Transform για αποθορυβοποίηση πριν τη μοντελοποίηση με Hidden Markov Model [1][35].

Η περίοδος 2019–2024 χαρακτηρίζεται προς την κατεύθυνση των μοντέλων Βαθιάς Μάθησης. Αρχιτεκτονικές CNN, LSTM, BLSTM με attention και πιο πρόσφατα Transformer αρχίζουν να κυριαρχούν στις δημοσιεύσεις [4]. Παράλληλα όμως εντοπίζονται και τα πρώτα σοβαρά μεθοδολογικά ζητήματα όπως η διαρροή δεδομένων (data leakage) σε δημοσιευμένα συστήματα [7] και αδυναμία γενίκευσης όταν αλλάζει ο χώρος (cross-environment) [18]. Στην ίδια περίοδο η εμφάνιση χαμηλού κόστους IoT πλατφορμών (ESP32, Raspberry Pi) δημιουργεί ένα νέο υποπεδίο που εστιάζει στο edge computing HAR αντί για τη χρήση σταθερών υπολογιστών με ειδικές κάρτες δικτύου [8][9].

Σήμερα, το 2026, η βιβλιογραφία έχει απομακρυνθεί από τους αυστηρά ελεγχόμενους χώρους και επικεντρώνεται περισσότερο στο ποιο υλικό λειτουργεί σε πραγματικά περιβάλλοντα και ποιοι αλγόριθμοι γενικεύουν έξω από το εργαστήριο [4]. Μια πρόσφατη ανασκόπηση του ACM 2025 [22] κωδικοποίησε πέντε ανοιχτές προκλήσεις για ανίχνευση σε σύνθετα πραγματικά σενάρια, ανίχνευση σε συνθήκες NLOS, ανίχνευση σε συνύπαρξη πολλών χρηστών (coexistence), ανίχνευση δυναμικών ομάδων χρηστών (dynamic-group), cross-domain γενίκευση και multi-task sensing. Στην πιο πρόσφατη βιβλιογραφία εμφανίζονται επίσης Self-Supervised Learning προσεγγίσεις που προσπαθούν να μάθουν αναπαραστάσεις CSI χωρίς ετικέτες, καθώς και foundation models όπως το WavesFM που στοχεύουν σε ενιαίο μοντέλο για επικοινωνία, εντοπισμό και sensing στο πλαίσιο 6G ISAC [25][29].

3.3 Συστήματα Χαμηλού Κόστους και Υλικός Εξοπλισμός (Hardware)

Η επιλογή του υλικού καθορίζει σε μεγάλο βαθμό τι είδους CSI είναι διαθέσιμο και επομένως ποια συστήματα sensing είναι εφικτά [1][2]. Συσκευές με περισσότερες κεραίες έχουν μεγαλύτερη χωρική πληροφορία και άρα μπορούν να ξεχωρίσουν λεπτότερες κινήσεις σε σχέση με αυτές που έχουν μία κεραία. Στην πράξη όμως, η επιλογή του υλικού δεν είναι μόνο θέμα ακρίβειας αλλά εξίσου σημαντικά είναι το κόστος, η κατανάλωση ισχύος και το αν το σύστημα μπορεί να λειτουργήσει αυτόνομα ή αν απαιτεί πάντα έναν host υπολογιστή. Με βάση τον αριθμό των δεκτών και κεραιών, τα διαθέσιμα CSI εργαλεία χωρίζονται σε δύο μεγάλες κατηγορίες [8]: τα multi-link πλατφόρμες όπως το Intel 5300 NIC και το Atheros CSI Tool που υποστηρίζουν MIMO, και single-link συστήματα όπως το ESP32 με Wi-ESP και το Raspberry Pi με Nexmon, τα οποία διαθέτουν μόνο έναν δέκτη ή μία κεραία και λειτουργούν με χαμηλότερες υπολογιστικές απαιτήσεις.

3.3.1 Συστήματα Intel 5300 NIC (MIMO)

Η κάρτα Intel 5300 NIC, με τρεις κεραίες σε διάταξη 3×3 MIMO και το εργαλείο εξαγωγής CSI των Halperin et al., αποτέλεσε για σχεδόν μια δεκαετία τον κύριο εξοπλισμό για ερευνητικά συστήματα HAR [1][2]. Παρέχει θεωρητικά έως και 504 μετρήσεις CSI ανά πακέτο σε πλήρη ανάλυση (9 ζεύγη κεραιών × 56 υποφορείς), στην πράξη όμως το Linux 802.11n CSI Tool εξάγει 30 ομαδοποιημένους υποφορείς ανά ζεύγος κεραιών, δηλαδή $9 \times 30 = 270$ μετρήσεις CSI ανά πακέτο, με ρυθμό δειγματοληψίας έως και 500 πακέτα/δευτερόλεπτο.

Σε σύγκριση με τα SISO συστήματα η πληροφορία είναι σημαντικά πλουσιότερη χωρικά, αλλά αυτή η ποιότητα έχει κόστος όπου η κάρτα δεν λειτουργεί αυτόνομα και απαιτεί έναν φορητό ή σταθερό υπολογιστή με συγκεκριμένη μητρική, παλιότερη διανομή Linux και κατάλληλους drivers. Αυτό αποκλείει έτσι μια αυτόνομη, χαμηλού κόστους ανάπτυξη στο σπίτι ή σε μη ερευνητικό περιβάλλον [8].

Χαρακτηριστικό παράδειγμα αυτής της αρχιτεκτονικής είναι το TCS-Fall των Zhou et al. [13]. Πρόκειται για ένα σύστημα ανίχνευσης πτώσης που συνέλεξε εκτενές dataset από 20 εθελοντές, με 1.800 πτώσεις και 2.400 ημερήσιες δραστηριότητες. Χρησιμοποιεί τον Grouped Coefficient of Variation των πλατών CSI ως είσοδο σε ένα CNN ταξινομητή και πετυχαίνει AUC 0,999, no-error warning ratio 0,955 και correct warning ratio 0,975 ακόμη και όταν εκπαιδεύεται με δεδομένα μόλις δύο εθελοντών. Όταν στο training περιλαμβάνονται δέκα εθελοντές η απόδοση ανέρχεται στο 1,00. Παρά τις υψηλές επιδόσεις, η εγκατάσταση παραμένει μη πρακτική διότι απαιτεί δύο πλήρεις υπολογιστές, έναν στον ρόλο πομπού και έναν στον ρόλο δέκτη, κάτι που δύσκολα γίνεται αποδεκτό σε ένα οικιακό περιβάλλον [13].

3.3.2 Nexmon CSI (Raspberry Pi)

Το Nexmon CSI Extractor είναι ένα framework που επιτρέπει την εξαγωγή CSI από το ενσωματωμένο WiFi chipset του Raspberry Pi, και είναι μια ενδιαφέρουσα επιλογή ανάμεσα στο Intel 5300 και τις πλήρως ενσωματωμένες λύσεις IoT. Είναι single-link, δηλαδή χρησιμοποιεί μία κεραία λήψης, και κοστίζει σαφώς λιγότερο από τον συνδυασμό Intel 5300 με φορητό υπολογιστή [8]. Έχει όμως δύο πρακτικά μειονεκτήματα. Πρώτον, αντιμετωπίζει ορισμένα προβλήματα ποιότητας στο επίπεδο επεξεργασίας σήματος, που επηρεάζουν την ποιότητα του εξαγόμενου CSI σε σχέση με το ESP32 ή το Intel 5300 [8]. Δεύτερον, ένα Raspberry Pi καταναλώνει σημαντικά περισσότερη ισχύ από έναν μικροελεγκτή ESP32, και το κόστος του κυμαίνεται από 50 έως 100€, πολύ ακριβότερο από έναν ESP32, κάτι που το καθιστά λιγότερο ελκυστικό όταν στόχος είναι η εκτεταμένη ανάπτυξη πολλών αισθητήρων σε ένα κτίριο.

3.3.3 ESP32 Η Πλατφόρμα Νέας Γενιάς

Ο μικροελεγκτής ESP32 της Espressif είναι ένα SoC (System on a Chip) με ενσωματωμένη Wi-Fi και Bluetooth συνδεσιμότητα και αποτελεί τη νέα γενιά εξοπλισμού για Wi-Fi HAR, αξιοποιώντας την ήδη εκτεταμένη εγκατεστημένη βάση συσκευών Wi-Fi παγκοσμίως [32]. Το κόστος του ξεκινά από λίγα ευρώ, η κατανάλωση είναι ελάχιστη και χάρη σε εργαλεία ανοιχτού κώδικα όπως το Wi-ESP και το ESP32-CSI-Tool μπορεί να εξάγει CSI στο πρότυπο 802.11n ή μέσω ESPNOW σε πραγματικό χρόνο [9][10][17][30]. Σε αντίθεση με την Intel 5300, λειτουργεί αυτόνομα ως edge device και επομένως μπορεί να αναπτυχθεί ως πραγματικός αισθητήρας σε σπίτια ή κέντρα φροντίδας ηλικιωμένων. Από την άλλη, είναι single-link με λιγότερους υποφορείς ανά πακέτο και μικρότερη χωρική πληροφορία από τα MIMO συστήματα, οπότε όταν χρειάζεται διάκριση πολλών παρόμοιων κλάσεων η ακρίβεια υστερεί [17].

Μια σειρά πρόσφατων εργασιών δείχνει τι μπορεί να καταφέρει το ESP32 στην πράξη. Οι Atif et al. [8] χρησιμοποίησαν ένα ζεύγος ESP32 για την αναγνώριση πέντε δραστηριοτήτων (καμία δραστηριότητα, καθιστός, όρθιος, βάδιση, πιάσιμο αντικειμένου) σε ελεγχόμενο περιβάλλον. Το pipeline τους είναι ενδεικτικό για συστήματα ESP32 όπου παρουσιάζεται συγχρονισμός πακέτων με βάση MAC και χρονοσφραγίδα, αφαίρεση outliers με Hampel filter, Gaussian smoothing για τον υψηλόσυχο θόρυβο, κανονικοποίηση και Bidirectional LSTM ως ταξινομητής. Οι ίδιοι συνέκριναν τρεις παραλλαγές α) single-layer LSTM (84,73% sequence accuracy), β) single-layer BLSTM (85,11%) και γ) two-layer BLSTM με 100 και 50 hidden units αντίστοιχα (89,26%), επιβεβαιώνοντας ότι ακόμη και ένα ζεύγος ESP32 (SISO) παράγει αξιοποιήσιμα δεδομένα για deep learning. Σε άλλη μελέτη οι Abdurrahman & Irawan [9] κράτησαν το ίδιο hardware (ESP32-WROOM-32) και πρότειναν ένα ελαφρύτερο μοντέλο, ένα κλασικό LSTM με additive attention. Σε 100 Monte Carlo εκτελέσεις (βάσει πολλαπλών τυχαίων επαναλήψεων) πέτυχαν macro-average $0,95 \pm 0,0056$ σε τρεις δραστηριότητες, με σαφώς μεγαλύτερη σταθερότητα από ένα LSTM-1DCNN baseline ($0,89 \pm 0,0119$). Η προεπεξεργασία τους περιλάμβανε

Hampel, Savitzky-Golay και SMOTE για ισορροπία κλάσεων, και είναι ενδεικτικό ότι ακόμη και με σχετικά απλή αρχιτεκτονική η ποιότητα προεπεξεργασίας καθορίζει την τελική απόδοση. Τέλος, οι Deepthi et al. [10] συνέκριναν CNN, LSTM και υβριδικό CNN+LSTM σε συνθετικά δεδομένα ESP32, με το CNN+LSTM να φτάνει 77% ακρίβεια. Το αποτέλεσμα είναι αρκετά κάτω από τα προηγούμενα, και δείχνει ένα σημαντικό σημείο ότι η ποιότητα των δεδομένων εκπαίδευσης (πραγματικά έναντι συνθετικών) έχει συχνά μεγαλύτερη επίδραση στην απόδοση από την επιλογή της αρχιτεκτονικής.

3.4 Προσεγγίσεις Κλασικής Μηχανικής Μάθησης (Classical ML)

Πριν το Deep Learning, το CSI HAR στηριζόταν κυρίως σε παραδοσιακούς αλγορίθμους με χειροκίνητη εξαγωγή χαρακτηριστικών. Αυτές οι μέθοδοι εξακολουθούν να είναι χρήσιμες όταν τα δεδομένα είναι λίγα ή όταν η ανάπτυξη γίνεται σε edge συσκευές με περιορισμένη υπολογιστική ισχύ [5].

Στην εργασία τους όμως οι Ali et al. [5] αξιολόγησαν εννέα αλγορίθμους, έξι κλασικής μάθησης (Naive Bayes, SVM, Decision Tree, Linear Regression, k-NN, Ensemble) και τρεις βαθιάς μάθησης (CNN, RNN, DNN), πάνω σε δεδομένα CSI από δύο TP-Link Archer C7 routers σε διάταξη 3×3 MIMO με 56 υποφορείς. Η εκπαίδευση έγινε σε δύο σύνολα α) το Widar dataset (35 χιλιάδες δείγματα εκπαίδευσης, 9 χιλιάδες δοκιμές) και β) ένα τοπικά συλλεγμένο dataset (8.000/2.000). Για το απλό πρόβλημα motion detection, αλγόριθμοι όπως Decision Trees και k-NN πέτυχαν απόδοση συγκρίσιμη με DL μοντέλα, ενώ στην τελική κατάταξη Multi-Attribute Decision-Making (που στηρίζεται αποκλειστικά σε έξι μετρικές επίδοσης) προηγήθηκαν τα μοντέλα βαθιάς μάθησης, με το k-NN να αναδεικνύεται ως ο καλύτερος κλασικός ταξινομητής (4ος συνολικά) [5]. Όταν όμως το πρόβλημα γίνεται πιο σύνθετο η εικόνα αλλάζει. Οι Moghaddam et al. [28] δοκίμασαν επτά classical ML αλγορίθμους σε fine-grained αλληλεπιδράσεις πολλών ατόμων, συνδυάζοντας CSI και RSSI σε 27 χειροποίητα χαρακτηριστικά, και το Random Forest κατέληξε με 94,16% ακρίβεια, δείχνοντας ότι όταν η εξαγωγή χαρακτηριστικών γίνει με προσοχή τα παραδοσιακά μοντέλα μένουν ανταγωνιστικά. Σε διαφορετική κατεύθυνση, το TCS-Fall των Zhou et al. [13] δεν τροφοδότησε στο μοντέλο ακατέργαστο CSI αλλά τον Grouped Coefficient of Variation των πλατών, μια απλούστερη αλλά σταθερότερη στατιστική, και πέτυχε καλύτερη γενίκευση σε άγνωστα άτομα. Τέλος, το NSee των Gong et al. αλλάζει εντελώς το ερώτημα και αντί για διάκριση δραστηριοτήτων ρωτά αν υπάρχει κάποιος κοντά στον αισθητήρα, με ένα Gaussian Mixture Model φτάνει 90% ακρίβεια σε σενάριο near-field sensing. [24]

Συνοψίζοντας, σχετικά με τις παραδοσιακές μεθόδους προκύπτουν τέσσερις παρατηρήσεις:

α) Η προεπεξεργασία (Hampel, Butterworth, PCA, wavelet) είναι σχεδόν πάντα απαραίτητη, καθώς τα classical ML δεν διαθέτουν τον μηχανισμό αυτόματης εξαγωγής χαρακτηριστικών που έχουν τα DL μοντέλα.

β) Οι χρόνοι εκπαίδευσης είναι σε δευτερόλεπτα ή λεπτά αντί για ώρες, κάτι που διευκολύνει την επαναληπτική αναζήτηση υπερπαραμέτρων.

γ) Τα μοντέλα έχουν μέγεθος της τάξης των KB αντί για MB ή GB, και τρέχουν φυσιολογικά σε edge συσκευές χωρίς GPU.

δ) Απαιτούν περισσότερη δουλειά στην εξαγωγή χαρακτηριστικών, αλλά σε αντιστάθμισμα ένα σφάλμα στο pipeline εντοπίζεται πολύ πιο εύκολα από ένα αντίστοιχο σε ένα βαθύ νευρωνικό δίκτυο.

3.5 Προσεγγίσεις Βαθιάς Μάθησης (Deep Learning)

Τα τελευταία πέντε χρόνια τα μοντέλα Deep Learning κυριαρχούν στο CSI HAR, τόσο σε αριθμό δημοσιεύσεων όσο και σε αναφερόμενη ακρίβεια. Αρχιτεκτονικές CNN, LSTM και πρόσφατα Vision Transformers έχουν αναφέρει υψηλές επιδόσεις κυρίως σε ελεγχόμενες συνθήκες. Πίσω από τους αριθμούς όμως υπάρχουν τρία επαναλαμβανόμενα προβλήματα: α) διαρροή δεδομένων στο train/test split [7], β) περιορισμένα και κακώς ισορροπημένα datasets, και γ) απαιτήσεις υπολογιστικής ισχύος που δεν είναι ρεαλιστικές για live inference σε edge συσκευές [3][22][29].

3.5.1 Convolutional Neural Networks (CNN)

Τα CNN χρησιμοποιούνται για εξαγωγή χωρικών και φασματικών χαρακτηριστικών από το CSI. Η συνηθέστερη πρακτική είναι η μετατροπή του ακατέργαστου σήματος σε εικόνες, spectrograms, recurrence plots ή time-frequency maps ώστε να αξιοποιηθούν αρχιτεκτονικές που έχουν δοκιμαστεί στο πεδίο. Οι He et al. [12] σχεδίασαν το CNN-ABLSTM, έναν συνδυασμό CNN για εξαγωγή χαρακτηριστικών και attention-based bidirectional LSTM για τη χρονική μοντελοποίηση. Το σύστημα αναφέρει βελτίωση 4% στην ακρίβεια σε σύγκριση με τις βασικές μεθόδους και διατηρεί 97% επιδόσεις σε διαφορετικά περιβάλλοντα (scenes), ενώ μειώνει σημαντικά το υπολογιστικό κόστος χάρη στο attention που εστιάζει στις πιο σημαντικές χρονικές στιγμές. Σε διαφορετική εφαρμογή, οι Guo & Ho [16] χρησιμοποίησαν Branchy CNN για ταυτόχρονη ανίχνευση παρουσίας και αναγνώριση δραστηριότητας σε σενάριο καραντίνας COVID-19, με την ακρίβεια του πειράματος να αγγίζει το 98,19%. Η Branchy αρχιτεκτονική επιτρέπει πρόωμη έξοδο για τις εύκολες περιπτώσεις και βαθύτερη επεξεργασία μόνο όταν χρειάζεται, κάτι που σε edge συστήματα έχει νόημα γιατί εξοικονομεί υπολογιστικούς πόρους [16].

3.5.2 LSTM και Παραλλαγές

Τα LSTM δίκτυα ταιριάζουν φυσικά σε CSI δεδομένα επειδή διατηρούν χρονική πληροφορία, δηλαδή μπορούν να θυμούνται μέσα στον χρόνο. Οι Atif et al. [8] χρησιμοποίησαν Bidirectional LSTM σε ESP32 και αναγνώρισαν πέντε δραστηριότητες με ακρίβεια άνω του 87%, με τη βαθύτερη παραλλαγή του two-layer BLSTM να φτάνει 89,26% sequence accuracy. Οι Niu et al. [19] αντιμετώπισαν το πρόβλημα της κατεύθυνσης κίνησης συνδυάζοντας CNN με GRU και Doppler spectrograms (CSI-F), επιτυγχάνοντας direction-independent αναγνώριση έως και τα 6 μέτρα. Διαφορετικό ερώτημα έθεσαν οι Zhu et al. [20], δηλαδή εάν το σύστημα μπορεί να ξεχωρίσει άνθρωπο από κατοικίδιο ή ρομποτική σκούπα. Με ένα μόνο ζεύγος συσκευών πέτυχαν 87,09% σε άγνωστα περιβάλλοντα, αποτέλεσμα ενδιαφέρον γιατί δείχνει ότι το CSI κρατά πληροφορία για το είδος της κινούμενης οντότητας και όχι μόνο για την κίνηση καθαυτή. Παράλληλα, οι Zhuravchak et al. [21] δημοσίευσαν ένα ανοιχτό end-to-end pipeline με ακρίβεια 87,8% στο σύνολο ελέγχου (61,1% στο validation) και δείχνουν ότι η επιλογή καναλιού Wi-Fi (γειτονικά APs, παρεμβολές) μπορεί να μετατοπίσει τα αποτελέσματα κατά αρκετές μονάδες.

3.5.3 Vision Transformers και Foundation Models

Στην τελευταία διετία αρχίζουν να εμφανίζονται προσεγγίσεις βασισμένες σε Transformer. Οι Luo et al. [11] σύγκριναν πέντε αρχιτεκτονικές ViT (vanilla ViT, SimpleViT, DeepViT, SwinTransformer, CaiT) σε δύο public datasets, το UT-HAR και το NTU-Fi HAR. Το CaiT πέτυχε 98,78% στο UT-HAR

και 98,2% στο NTU-Fi, δείχνοντας ότι Transformer αρχιτεκτονικές μπορούν να αποδώσουν σε CSI που έχει μετασχηματιστεί σε spectrogram (Εικόνα 2.5). Πιο ενδιαφέρον όμως είναι ότι το SwinTransformer επιτυγχάνει συγκρίσιμη ακρίβεια στο UT-HAR με τις λιγότερες παραμέτρους από όλα τα μοντέλα, κάτι που δείχνει ότι ο όγκος του μοντέλου από μόνος του δεν εγγυάται πάντα καλύτερη απόδοση [11]. Σε ακόμη μια πιο πρόσφατη κατεύθυνση, μοντέλα όπως το WavesFM [29] προτείνουν τη χρήση foundation models που εκπαιδεύονται σε τεράστιους όγκους ασύρματων δεδομένων χωρίς ετικέτες και στη συνέχεια προσαρμόζονται με ελάχιστο fine-tuning σε επιμέρους εργασίες (HAR, εντοπισμός, classification). Η λογική είναι η ίδια με αυτή που ακολουθήθηκε στα LLM και τα ViT στην όραση δηλαδή μια μεγάλη προεκπαιδευμένη βασική αρχιτεκτονική (pretrained backbone) και πολλές μικρές downstream εργασίες. Παράλληλα, η ανασκόπηση των Radwan et al. [25] τεκμηριώνει ότι Self-Supervised Learning τεχνικές (contrastive, masked modeling) αποτελούν μια αποτελεσματική λύση στο πρόβλημα της έλλειψης επισημασμένων (labeled) δεδομένων στο CSI HAR.

3.6 Πλεονεκτήματα: Ιδιωτικότητα και Ασφάλεια

Ένα από τα σημαντικότερα θετικά του Wi-Fi sensing έναντι άλλων τεχνολογιών είναι η προστασία της ιδιωτικότητας. Σε αντίθεση με τις κάμερες, οι Wi-Fi αισθητήρες δεν καταγράφουν εικόνες ή βίντεο, και επομένως δεν υπάρχει σημαντικός κίνδυνος παρακολούθησης αναγνωρίσιμων χαρακτηριστικών προσώπου ή σώματος [1][4]. Πιο συγκεκριμένα διακρίνονται τα εξής σημεία:

α) Φυσική ιδιωτικότητα: Τα σήματα CSI δεν αποκαλύπτουν άμεσα την ταυτότητα του ατόμου ούτε στοιχεία της εμφάνισής του και ακόμη και αν υποκλαπούν δεν παρέχουν αναγνωρίσιμη πληροφορία με την κλασική έννοια [6]. Παρόλα αυτά η ασφάλεια του δεν είναι αυτονόητη, καθώς η μελέτη των Zhang et al. [6] παρατήρησε ότι όταν δεδομένα CSI από διαφορετικά περιβάλλοντα συγκεντρώνονται για κοινή εκπαίδευση μπορεί να αποκαλυφθεί πληροφορία τοποθεσίας ή ακόμη και βιομετρική. Πρότειναν λοιπόν μια λύση βασισμένη στον Johnson-Lindenstrauss μετασχηματισμό που είναι differentially private, σε συνδυασμό με adversarial learning για την παραγωγή ανεξάρτητων από το περιβάλλον (environment-invariant) αναπαραστάσεων, και ανέφεραν βελτιώσεις 2,18% και 1,24% σε ακατέργαστο (raw) CSI και 5,71% και 1,55% σε wavelet features σε δύο πραγματικά περιβάλλοντα [6].

β) Λειτουργία σε κάθε συνθήκη: Τα Wi-Fi σήματα διαπερνούν τοίχους, λειτουργούν στο σκοτάδι και δεν επηρεάζονται από ομίχλη ή καπνό, κάτι που είναι αδύνατο για κάμερες [1]. Αυτό κάνει το CSI HAR ιδιαίτερο για σενάρια όπως ανίχνευση πτώσης στο μπάνιο, παρακολούθηση ύπνου ή λειτουργία σε υπονοδωμάτιο, όπου μια κάμερα θεωρείται απαράδεκτη παρέμβαση στο χώρο.

γ) Καμία wearable συσκευή: Δεν απαιτείται από τον χρήστη να φοράει αισθητήρα, κάτι σημαντικό για ευάλωτες ομάδες (ηλικιωμένοι, ασθενείς, παιδιά) που είτε ξεχνούν τη συσκευή είτε τη βγάζουν για λόγους άνεσης [8][13].

3.7 Ανοιχτές Προκλήσεις: Data Leakage και Cross-Environment

Παρά τα παραπάνω θετικά, η βιβλιογραφία περιέχει αρκετά αποτελέσματα που πρέπει να αντιμετωπιστούν με επιφύλαξη. Η πιο γνωστή κριτική ανάλυση είναι αυτή του Varga [7], που εξέτασε ένα δημοσιευμένο σύστημα CNN με αναφερόμενη ακρίβεια 99,9% και εντόπισε σοβαρό πρόβλημα data leakage στο train/test split. Συγκεκριμένα, ο διαχωρισμός γινόταν τυχαία σε επίπεδο παραθύρων ή

εικόνων CSI, οπότε δείγματα από τον ίδιο συμμετέχοντα κατέληγαν τόσο στο training όσο και στο test set. Το μοντέλο είχε δηλαδή «δει» το ίδιο άτομο κατά την εκπαίδευση, και η υψηλή ακρίβεια στο test ήταν στην ουσία αναγνώριση γνώριμων προσώπων, και όχι γενίκευση. Όταν ο Varga ξαναέκανε το πείραμα με διαχωρισμό με βάση το υποκείμενο (subject-based split) στα ίδια datasets, το precision έπεσε σε 23,4% στο WiAR και 47,4% στο Widar3.0 [7].

Η δεύτερη μεγάλη πρόκληση είναι η γενίκευση σε νέο περιβάλλον. Τα μοντέλα CSI HAR εκπαιδεύονται σε δεδομένα ενός συγκεκριμένου χώρου (δωμάτιο, γραφείο, διάδρομος) και η χωρική υπογραφή του CSI που μαθαίνει το μοντέλο εξαρτάται από:

- τη γεωμετρία του δωματίου (διαστάσεις, σχήμα),
- τα υλικά των τοίχων (τσιμέντο, γύψος, γυαλί),
- την επίπλωση και τα αντικείμενα μέσα στον χώρο,
- τη σχετική θέση πομπού και δέκτη.

Όταν το μοντέλο μεταφερθεί σε νέο δωμάτιο το «background» του CSI αλλάζει ριζικά και τα χαρακτηριστικά που εξήγαγε το μοντέλο δεν έχουν πλέον την ίδια σημασία [6][18]. Στη μελέτη CrossSense των Zhang et al. [18] τα τρέχοντα συστήματα αναγνώρισης βάδισης (gait identification) έπεφταν από υψηλή ακρίβεια σε 20% όταν δοκιμάζονταν σε νέο περιβάλλον. Με τη χρήση roaming model και mixture-of-experts προσέγγισης, η ακρίβεια ανέβαινε ξανά σε πάνω από 80% για βάδιση (gait) και 90% για χειρονομίες (gestures). Παρόμοια, η μελέτη TCS-Fall αναφέρει ότι, σε προηγούμενα συστήματα χωρίς ειδική μέριμνα για cross-individual γενίκευση, η ακρίβεια μπορούσε να πέσει από 93,5% σε 60% όταν άλλαζε το άτομο που πραγματοποιούσε τις δραστηριότητες σε σχέση με αυτό που χρησιμοποιήθηκε κατά την εκπαίδευση [13]. Αυτό αναδεικνύει το θεμελιώδες πρόβλημα το οποίο υποδεικνύει ότι δεν αρκεί η γενίκευση σε νέο χώρο, αλλά απαιτείται και γενίκευση σε νέα άτομα, κάτι που είναι δύσκολο να επιτευχθεί ταυτόχρονα.

3.8 Συγκριτικός Πίνακας Βιβλιογραφίας

Ο ακόλουθος πίνακας συνοψίζει τα κυριότερα συστήματα που έχουν πραγματοποιηθεί.

Μελέτη	Hardware	Κλάσεις	Μέθοδος	Ακρίβεια	Real-time	Cross-Env
Abdurrahman & Irawan (2025) [9]	ESP32 (SISO)	3	LSTM-Attention	95%	Όχι	Όχι

Μελέτη	Hardware	Κλάσεις	Μέθοδος	Ακρίβεια	Real-time	Cross-Env
Deepthi et al. (2025) [10]	ESP32 (SISO)	3	CNN+LSTM	77%	Όχι	Όχι
Atif et al. (2024) [8]	ESP32 (SISO)	5	BLSTM	>87%	Ναι	Όχι
Zhou et al. (2024) [13]	Intel 5300 (MIMO)	2	CNN + GCoV	AUC=0,999	Ναι	Ναι*
Luo et al. (2024) [11]	Intel 5300 (MIMO)	6-7	ViT (CaiT)	SOTA	Όχι	Όχι
He et al. (2023) [12]	Intel 5300 (MIMO)	7	CNN-ABLSTM	97%	Ναι*	Ναι*
Ali et al. (2023) [5]	TP-Link 3x3 (MIMO)	2	SVM/RF/DL	~97%	Όχι	Όχι
Moghaddam et al. [28]	MIMO	7	RF	94,16%	Όχι	Όχι
Zhang et al. (2023) [6]	Intel 5300 (MIMO)	6	Adversarial DL	Βελτίωση 2-5,7%	Όχι	Ναι
Mattela et al. (2022) [14]	Redpine RS9116 (SISO)	2	LSTM Autoenc.	97-99%	Ναι	Ναι*
Hao et al. (2020) [15]	Atheros NIC (MIMO)	5+	RBM	85,4%	Ναι	Ναι*
Niu et al. (2024) [19]	MIMO	6+	CNN+GRU(CSI-F)	>SOTA	Όχι	Μερικά

Μελέτη	Hardware	Κλάσεις	Μέθοδος	Ακρίβεια	Real-time	Cross-Env
Zhu et al. (2024) [20]	SISO (1 ζεύγος)	2	Deep Learning (A-ACF)	87,09% unseen env	Ναι	Ναι
Zhuravchak et al. (2022) [21]	Commercial NIC	7	InceptionTime / LSTM	87,8% (test)	Όχι	Όχι
Zhang et al. (MobiCom'18) [18]	MIMO/RS SI	gait / gesture	Transfer Learning	20%→>80% cross-site	Όχι	Ναι
Παρούσα εργασία	ESP32-C6 (SISO)	2	Gradient Boosting	96,8–97,2%	Ναι	Ιδανικά μόνο

Πίνακας 3.1. Συγκριτικός πίνακας ανίχνευσης ανθρώπινης δραστηριότητας, Wi-Fi CSI HAR. (*: μερική υποστήριξη ή υπό προϋποθέσεις, σύμφωνα με τους συγγραφείς των αντίστοιχων εργασιών.)

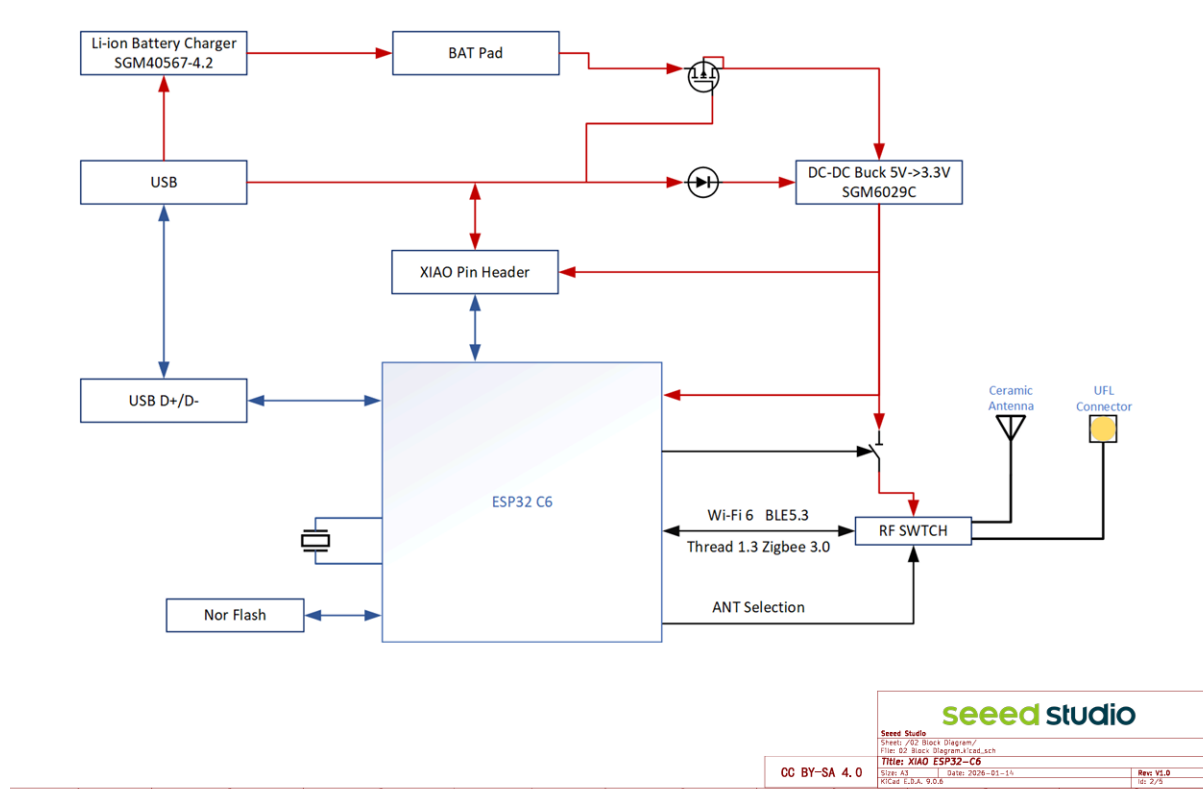
Κεφάλαιο 4ο: Σχεδιασμός και Υλοποίηση Συστήματος

Το παρόν κεφάλαιο παρουσιάζει αναλυτικά τον σχεδιασμό και την υλοποίηση του συστήματος σχετικά με την ανίχνευση ανθρώπινης παρουσίας χρησιμοποιώντας την τεχνολογία Wi-Fi CSI. Θα αναλυθεί η υλική διάταξη καθώς και το firmware εξαγωγής CSI. Έπειτα αναλύεται η προεπεξεργασία του σήματος, την εξαγωγή χαρακτηριστικών και στο τέλος τους αλγορίθμους μηχανικής μάθησης που χρησιμοποιήθηκαν.

4.1 Αρχιτεκτονική Συστήματος

Το σύστημα, υλοποιείται με την αρχιτεκτονική SISO (Single Input Single Output), δηλαδή διαθέτει έναν πομπό (Transmitter, Tx) ο οποίος εκπέμπει συνεχώς πακέτα Wi-Fi, ενώ ο δέκτης παράλληλα (Receiver, Rx) λαμβάνει τα πακέτα και τα εξάγει σε πραγματικό χρόνο σε έναν πίνακα CSI. Και ο πομπός και ο δέκτης είναι ταυτόσημες πλακέτες Seeed Studio XIAO ESP32-C6, ενσωματώνοντας την τεχνολογία (SoC - System on a Chip) ESP32-C6FH4 της εταιρείας Espressif Systems [30]. Σε σχέση με τις MIMO διατάξεις (π.χ. 3×3 της Intel 5300 NIC), η SISO δεν προσφέρει χωρική πληροφορία και γωνία άφιξης (AoA), επιτρέπει όμως το χαμηλό κόστος ανάπτυξης χωρίς εξειδικευμένες κάρτες δικτύου [9][8]. Τα SISO συστήματα έχουν αποδειχτεί ότι επαρκούν για δυαδική ταξινόμηση Motion vs Idle καθώς η χρονική ενεργειακή υπογραφή κίνησης ενός ανθρώπινου σώματος είναι διακριτά ανιχνεύσιμη [3][9].

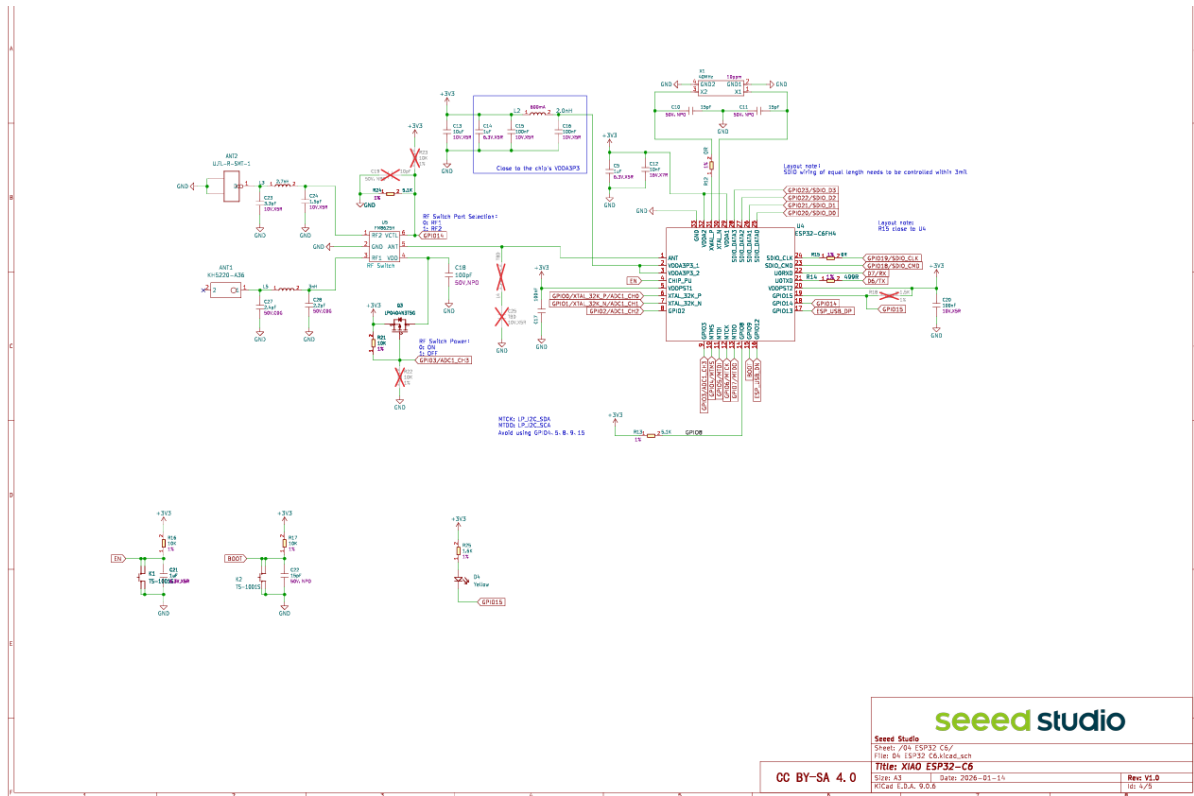
Η συνολική αρχιτεκτονική περιλαμβάνει τέσσερα στάδια: (α) Το πρώτο στάδιο είναι η επικοινωνία και η δημιουργία του καναλιού και του CSI με τη χρήση του firmware και στους δύο μικροελεγκτές ESP32-C6 [30]. Το δεύτερο στάδιο είναι το στάδιο επικοινωνίας όπου αφορά την μεταφορά των δεδομένων σε μορφή CSV ή και TXT μέσω σειριακής θύρας UART/ JTAG στον υπολογιστή. Το τρίτο στάδιο αφορά την επεξεργασία του σήματος με την χρήση κατάλληλου κώδικα Python (pipeline) και εξαγωγής χαρακτηριστικών, και τέλος υπάρχει το στάδιο ταξινόμησης ή αλλιώς ML (Machine Learning) με τη χρήση της βιβλιοθήκης sklearn όπου είναι εφικτή η χρήση των αλγορίθμων της και η αξιολόγησή τους.



Εικόνα 4.1. Αρχιτεκτονική της πλακέτας σε υψηλό επίπεδο.[31]

Η παραπάνω εικόνα 4.1 δείχνει την αρχιτεκτονική της πλακέτας σε υψηλό επίπεδο (χωρίς τα ηλεκτρονικά εξαρτήματα).

- Ο πυρήνας (ESP32-C6): Επικοινωνεί με την εξωτερική μνήμη (Nor Flash) όπου αποθηκεύεται ο κώδικας.
- Η Διαδρομή της Τροφοδοσίας (Κόκκινες γραμμές): Το ρεύμα έρχεται από τη θύρα USB, περνάει από το κύκλωμα φόρτισης μπαταρίας (Li-ion Battery Charger) και καταλήγει σε έναν μετατροπέα DC-DC (Buck converter 5V->3.3V).
- Το τμήμα RF (Κάτω δεξιά): Αυτό είναι το πιο σημαντικό κομμάτι για τη συλλογή δεδομένων του καναλιού. Παρατηρείται ένας διακόπτης στα δεξιά του κεντρικού chip με την ονομασία RF SWITCH στον οποίο επιλέγεται αν το σήμα θα πάει στην ενσωματωμένη κεραμική κεραία (Ceramic Antenna) ή στο εξωτερικό βύσμα (UFL Connector).

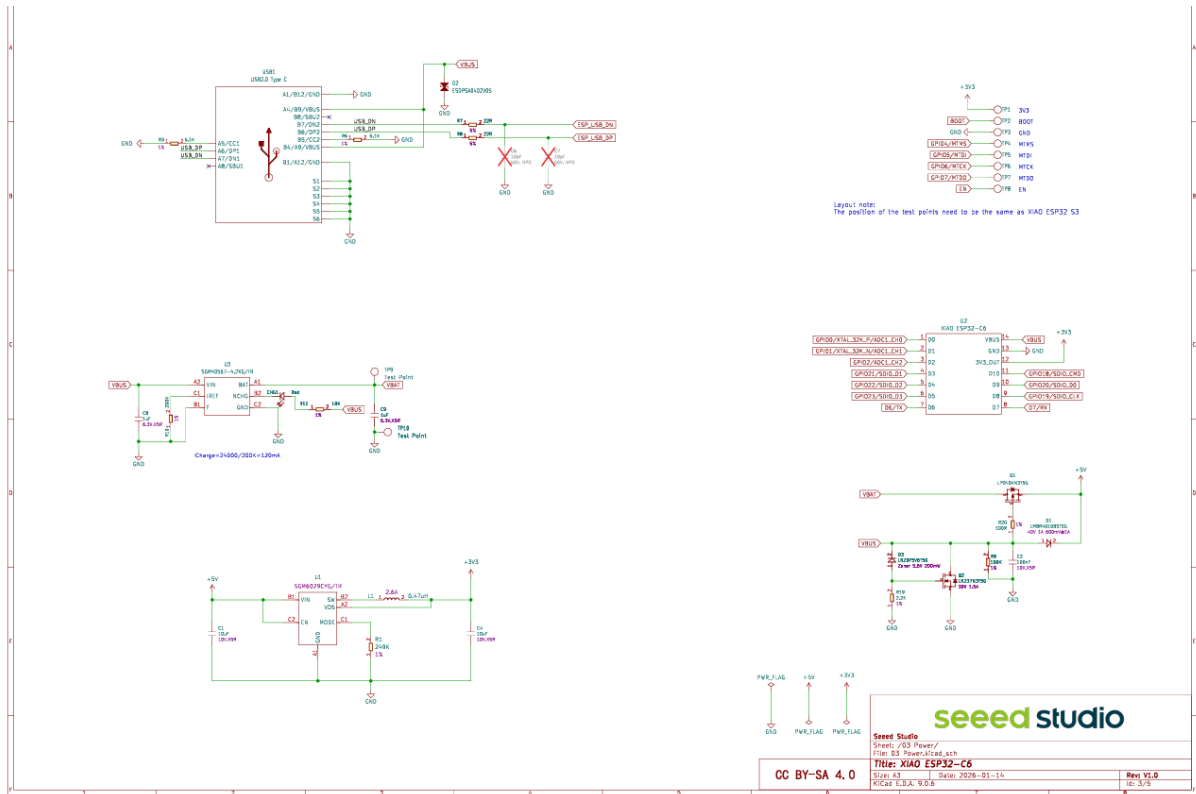


Εικόνα 4.2. Ο Μικροελεγκτής και το Κύκλωμα RF.[31]

Εδώ, στην εικόνα 4.2, παρατηρείται το πλήρες κύκλωμα του μικροελεγκτή και όλες τις εξόδους και εισόδους του (τα GPIOs, τις συνδέσεις για τη μνήμη SPI, κ.λπ.). Το Ρολόι (πάνω δεξιά από το chip) περιλαμβάνει έναν κρύσταλλο στα 40 MHz. Ο Διακόπτης Κεραίας (Αριστερά): Το εξάρτημα U5 είναι ο διακόπτης RF που παρουσιάστηκε στην προηγούμενη εικόνα και ελέγχεται από το pin GPIO14.

Τέλος η εικόνα 4.3 έχει να κάνει με την τροφοδοσία και τη διασύνδεση και αφορά αποκλειστικά το πώς η πλακέτα παίρνει το ρεύμα και πώς επικοινωνεί με τον υπολογιστή.

- Η Θύρα USB (Πάνω αριστερά): Το βύσμα USB Type-C. Παρέχει την τάση (VBUS) και τα δεδομένα (D+ / D-).
- Ο Μετατροπέας Τάσης (Κάτω αριστερά - U1): Αυτό είναι το chip (SGM6029C) που παίρνει τα 5V από το USB (ή τα 3.7V-4.2V από την μπαταρία) και τα κατεβάζει ακριβώς στα 3.3V (+3V3) που χρειάζεται το σύστημα.



Εικόνα 4.3. Τροφοδοσία και Διασύνδεση (Power & USB).[31]

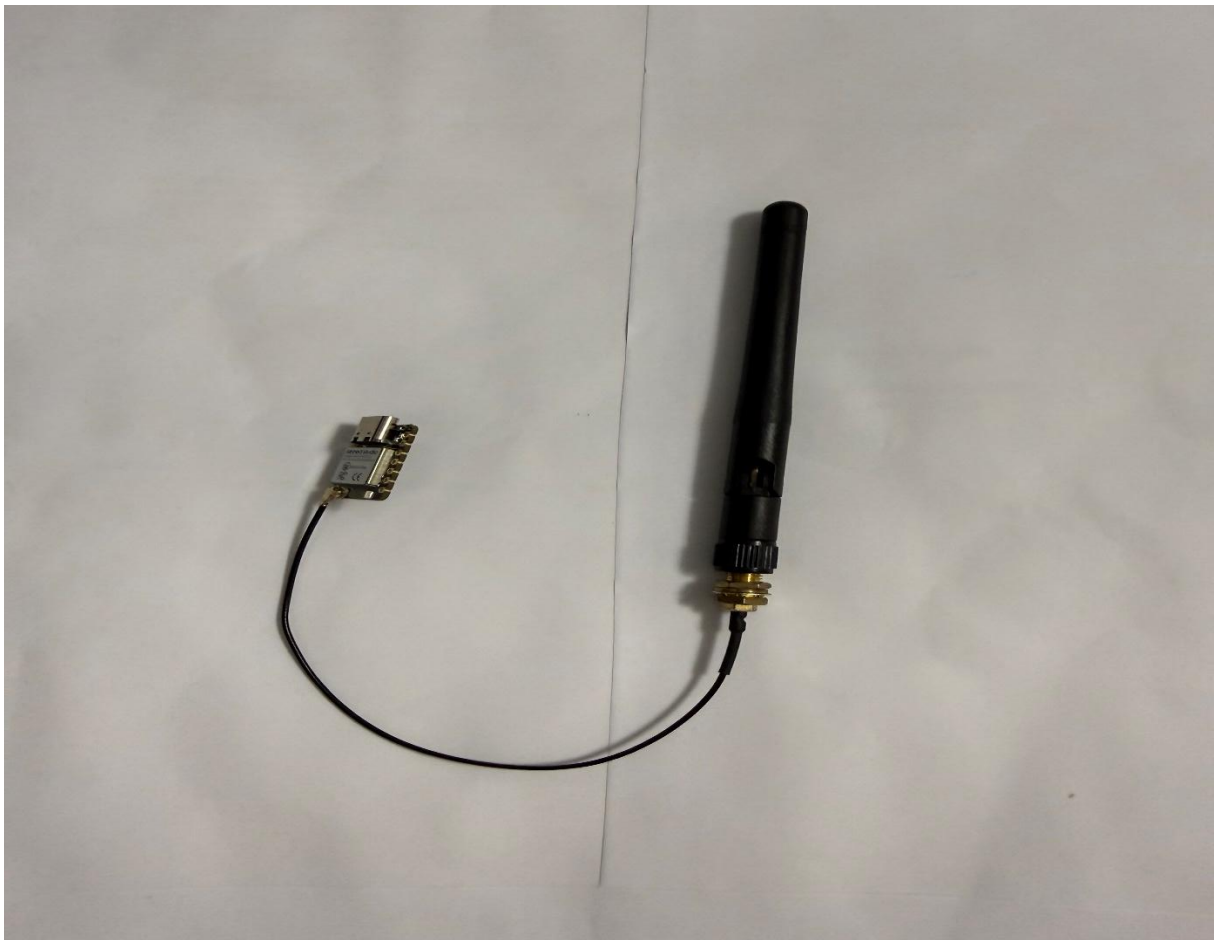
4.1.1 AGC Gain Locking

Ένα από τα πιο κρίσιμα σημεία που υλοποιήθηκε είναι το Automatic Gain Control (AGC) locking. Το AGC είναι στην ουσία ένα κύκλωμα στο δέκτη που αυτόματα προσαρμόζει το κέρδος του ενισχυτή λήψης ώστε να διατηρήσει το λαμβανόμενο σήμα σε ένα ικανοποιητικό επίπεδο. Για την κανονική λειτουργία επικοινωνίας του Wi-Fi η προσαρμογή αυτή είναι επιθυμητή και δεν «κλειδώνεται»· για την εξαγωγή CSI όμως αποτελεί πρόβλημα, καθώς οι αυτόματες μεταβολές του κέρδους αλλοιώνουν το πλάτος του CSI με τρόπο που δεν οφείλεται στο κανάλι, καθιστώντας τα δείγματα μη συγκρίσιμα μεταξύ τους και εισάγοντας ψευδείς «κινήσεις» [30].

Για να αντιμετωπιστεί αυτό, το firmware ρυθμίζει το AGC gain σε μια σταθερή τιμή κατά την αρχικοποίηση του μικροελεγκτή όταν τροφοδοτηθεί με ρεύμα και το διατηρεί κλειδωμένο για όλη τη διάρκεια της λειτουργίας. Έτσι με αυτόν τον τρόπο εξασφαλίζεται ότι οποιαδήποτε αλλαγή γίνεται στο CSI amplitude προέρχεται από πραγματική αιτία (π.χ. κίνηση) και όχι το ίδιο το υλικολογισμικό. Το κλειδωμα υλοποιείται με ρυθμίσεις χαμηλού επιπέδου του RF front-end (force-gain) κατά την αρχικοποίηση, η δυνατότητα αυτή δεν είναι επίσημα τεκμηριωμένο API της ίδιας εταιρίας Espressif, αλλά καθιερωμένη πρακτική της κοινότητας εργαλείων εξαγωγής CSI για ESP32.

4.1.2 Δυνατότητες Wi-Fi 6 (802.11ax), ο λόγος μη αξιοποίησής τους, και μέγεθος του Μικροελεγκτή

Το ESP32-C6 είναι ο πρώτος μικροελεγκτής της Espressif που έχει την τεχνολογία Wi-Fi 6 στα συστήματα χαμηλού κόστους [30]. Παρόλα αυτά το παρόν σύστημα χρησιμοποιεί το πρωτόκολλο ESPNOW, το οποίο δεν επιτρέπει τη χρήση του Wi-Fi 6. Έτσι χρησιμοποιείται το πρωτόκολλο ESPNOW, με Wi-Fi 4 και το CSI να εξάγεται σε λειτουργία 802.11n (HT40) [30]. Όσον αφορά το μέγεθος του XIAO της Seeed Studio είναι (περίπου 21 x 17,8 mm) [31]. Αυτό επιτρέπει την χρήση του αισθητήρα μέσα σε κοινές οικιακές συσκευές (π.χ. σε μια έξυπνη πρίζα ή έναν λαμπτήρα). Στην παρακάτω εικόνα παρουσιάζεται το πραγματικό μέγεθος του μικροελεγκτή (Εικόνα 4.4).



Εικόνα 4.4. Μικροελεγκτής ESP32-C6

4.1.3 Πλακέτα Seeed Studio XIAO ESP32-C6 και Εξωτερική Κεραία

Η επιλογή εξωτερικής κεραίας ελέγχεται από το RF switch chip (FM8625H) και στο pin του μικροελεγκτή GPIO14. Αναλυτικότερα, το WIFI_ANT_CONFIG με VCTL=LOW επιλέγεται το RF1 (εσωτερική κεραία), ενώ με VCTL=HIGH επιλέγεται το RF2 (εξωτερική). Στο firmware της παρούσας εργασίας ορίζεται το WIFI_ANT_CONFIG=GPIO14 και μια εξωτερική κεραία 3 dBi τύπου dipole πανκατευθυντικής (omnidirectional) ακτινοβολίας [30][31]. Οι 3 dBi dipole κεραίες παρέχουν πανκατευθυντική κεραία 360° στο οριζόντιο επίπεδο, η οποία είναι ιδανική για ανίχνευση κίνησης σε τυχαία κατεύθυνση εντός δωματίου [9].

Ο κώδικας αρχικοποίησης της κεραίας περιλαμβάνει:

```
// Configure GPIO for antenna switch

gpio_config_t io_conf = {

    .pin_bit_mask = (1ULL << ANT_SEL_PIN),

    .mode = GPIO_MODE_OUTPUT,

};

gpio_config(&io_conf);

// Select external antenna

gpio_set_level(ANT_SEL_PIN, 1);
```

4.2 Firmware Εξαγωγής CSI

Το firmware αναπτύχθηκε στο ESP-IDF (Espressif IoT Development Framework) με λειτουργικό σύστημα FreeRTOS. Η αρχιτεκτονική firmware διαχωρίζει τις εργασίες: α) ο πομπός (Tx) μεταδίδει συνεχώς πακέτα Wi-Fi σε καθορισμένο ρυθμό ενώ β) ο δέκτης (Rx) δέχεται τα πακέτα εξαγωγής του CSI και μέσω ουράς (Queue) στο task τα δίνει στον υπολογιστή. Στο παράρτημα Α παρουσιάζεται αντίστοιχα ο κώδικας υλοποίησης.

4.2.1 Ρυθμός Δειγματοληψίας και FreeRTOS Ουρά

Ο ρυθμός δειγματοληψίας ορίζεται ακριβώς σε SAMPLING_RATE = 100,0 Hz (config.py). Στα 100 Hz, κάθε δείγμα CSI αντιστοιχεί σε χρονικό βήμα $\Delta t = 10$ ms, το οποίο είναι ικανό να καταγράψει κίνηση ανθρώπινου σώματος, όπως είναι ο βηματισμός (0,5–3 Hz) και η αναπνοή (0,2–0,5 Hz) [1][19].

Στον δέκτη, κάθε νέο πλαίσιο (frame) CSI που έρχεται από το Wi-Fi driver μπαίνει σε μια ουρά χωρητικότητας CSI_QUEUE_SIZE = 30. Με αυτόν τον τρόπο, αν η διεργασία εκτύπωσης καθυστερήσει για λίγο, τα πλαίσια δεν χάνονται αλλά περιμένουν στην ουρά μέχρι να διαβαστούν.

Στην περίπτωση που για κάποιο λόγο η ουρά γεμίσει (Queue Overflow), το σύστημα δεν κολλάει. Απλώς τυπώνει ένα προειδοποιητικό μήνυμα WARNING: X CSI frames dropped! και συνεχίζει κανονικά τη λειτουργία του, ώστε να μην χαθεί η πραγματική (real-time) συμπεριφορά ακόμη και υπό φορτίο [30]. Το αντίστοιχο απόσπασμα κώδικα βρίσκεται στο Παράρτημα Β.

4.2.2 Δομή και Περιεχόμενο CSI Πλαισίου

Το ESP32-C6 είναι σε λειτουργία 802.11n HT40 (40 MHz). Στην λειτουργία αυτή το firmware εξάγει ακατέργαστα δεδομένα για 128 υποφορείς (128 ζεύγη real/imaginary, ένα ζεύγος ανά υποφορέα), εκ των οποίων 14 είναι null (guard bands + DC). Η αφαίρεσή τους γίνεται στο Python pipeline, αφήνοντας 114 ενεργούς υποφορείς: 108 δεδομένων και 6 pilot. Κάθε τιμή εκφράζεται ως ζεύγος (real, imaginary) ακεραίων 8-bit, και το συνολικό μέγεθος φόρτωσης CSI ανά πλαίσιο είναι $128 \times 2 = 256$ bytes, όπως φαίνεται και στο δείγμα δεδομένων παρακάτω (buffer 256 τιμών, με μηδενικά στους guard και στην DC). Από αυτές λοιπόν, εξάγεται το πλάτος ανά υποφορέα ως:

$$A(i) = \sqrt{H_{real(i)}^2 + H_{imag(i)}^2}, i = 1, \dots, 114 \quad (4.1)$$

Στην παρούσα εργασία χρησιμοποιείται το πλάτος του CSI, καθώς είναι λιγότερο ευαίσθητο σε τυχαίες μεταθέσεις φάσης που οφείλονται σε CFO (Carrier Frequency Offset) και SFO (Sampling Frequency Offset) που αναφέρθηκαν και στο κεφάλαιο 2 [9][19]. Μετά την αφαίρεση των null υποφορέων, κάθε δείγμα CSI καταλήγει σε διάνυσμα 114 τιμών πλάτους, τα οποία αποθηκεύονται σε CSV/TXT ως όνομα αρχείου την ετικέτα κλάσης (label).

Στα αρχεία αυτά κάθε γραμμή που περιέχει έγκυρα CSI δεδομένα ξεκινά με την λέξη 'CSI_DATA' και στην συνέχεια από τα ακατέργαστα bytes. Το αρχείο `csi_parser.py` διαβάζει αυτά τα αρχεία, εξάγει τις γραμμές CSI_DATA, και φτιάχνει το CSI.

Η μορφή των ακατέργαστων δεδομένων CSI που παρέχει το ESP-IDF είναι μια σειρά από signed integers που αναπαριστούν τα φανταστικά (imaginary) και πραγματικά (real) μέρη. Συγκεκριμένα, για κάθε υποφορέα, το buffer περιέχει δύο διαδοχικούς αριθμούς: το imaginary component και το real component. Η ανακατασκευή του μιγαδικού αριθμού γίνεται ως εξής:

```
for i in range(0, len(data), 2):
```

```
    imag = data[i]
```

```
    real = data[i+1]
```

```
    complex_sample = real + 1j * imag
```

Παρακάτω παρουσιάζεται ένα δείγμα από τα δεδομένα αυτά από το αρχείο `walk_activity_45_room1_vasilis_1777806476.txt`.

```
CSI_DATA,373127,1a:00:00:00:00:00,-47,11,-
```

```
97,17,23,11,7055570,51CSI_DATA,373613,1a:00:00:00:00:00,-47,11,-
```

```
97,17,23,11,11910713,51,0,256,0,"[0,0,0,0,0,0,0,0,0,0,0,0,19,1,21,3,22,2,20,5,21,6,21,9,21,11,18,14,2  
2,13,19,16,19,17,19,19,18,19,19,16,23,19,20,19,18,16,23,15,23,11,24,6,25,13,26,8,27,10,26,8,27,8
```

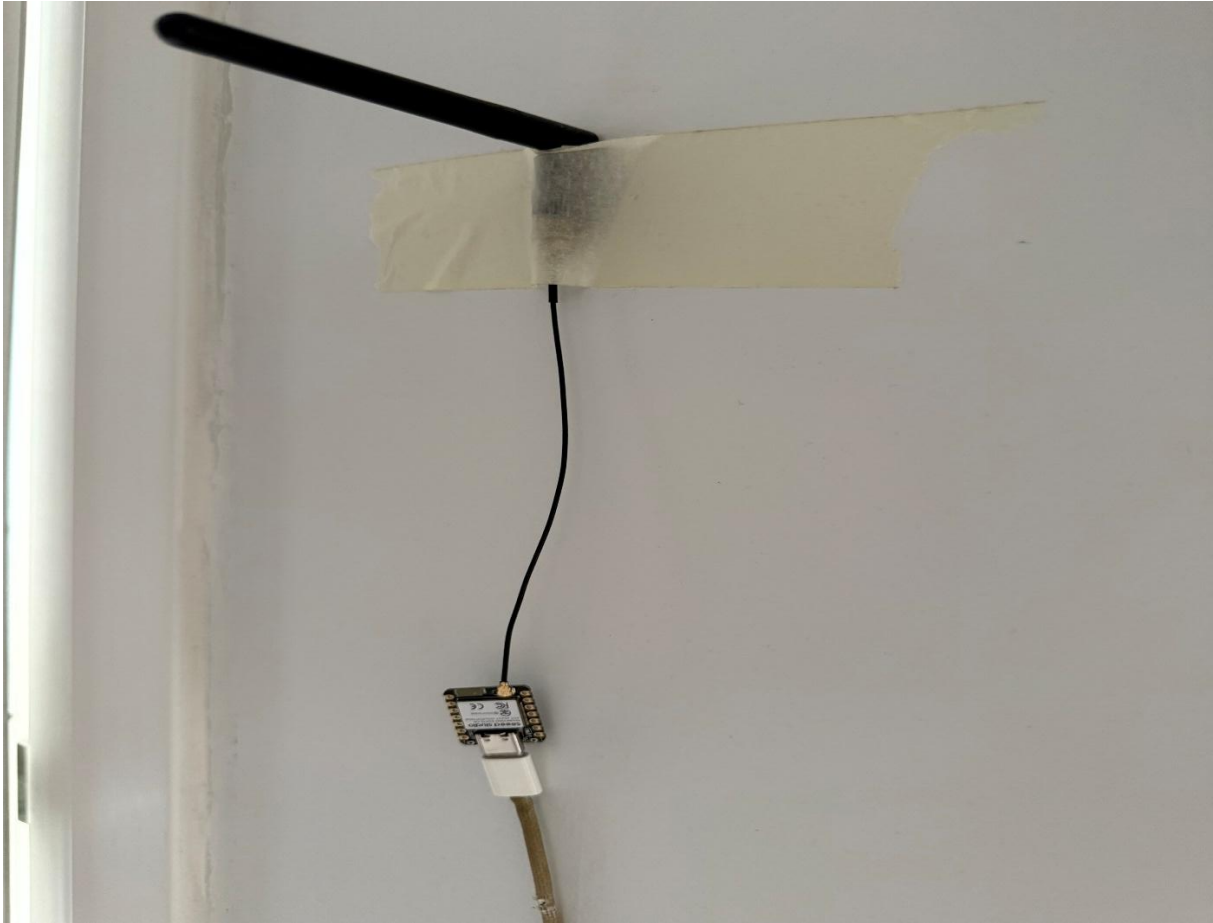
,27,8,25,6,27,9,25,10,23,7,25,7,24,3,24,4,23,2,26,4,23,3,23,6,22,6,21,5,22,1,23,8,23,7,23,4,24,3,23,3,21,3,22,0,24,4,22,3,24,5,22,5,23,3,22,3,23,5,22,1,24,1,21,0,0,0,0,0,21,1,21,0,18,1,19,2,17,1,17,0,16,-2,16,-1,15,0,15,1,14,0,14,4,16,0,14,1,13,-1,13,-1,13,-1,12,1,12,0,11,-3,10,-2,12,-1,11,-3,12,-1,11,-1,11,-3,11,-1,10,-4,12,-2,11,-2,11,-2,11,-1,10,-4,9,-2,11,-2,10,-1,10,0,10,1,11,1,11,0,9,3,9,1,8,2,8,2,9,4,8,0,7,3,8,3,5,3,4,3,5,4,4,3,4,3,3,4,4,1,3,3,1,4,0,0,0,0,0,0,0,0,0,0,0]

4.3 Πειραματικό Περιβάλλον και Χωρική Διάταξη

Τα πειράματα πραγματοποιήθηκαν σε εσωτερικό χώρο (indoor) πολυκατοικίας, ενός συνηθισμένου δωματίου με τοίχους, έπιπλα και ηλεκτρονικές συσκευές. Η παρουσία αντικειμένων εντός του χώρου εισάγει μόνιμες ανακλάσεις (static multipath) που ωστόσο παραμένουν σταθερές κατά τη διάρκεια κάθε καταγραφής και επομένως φιλτράρονται αποτελεσματικά από την προεπεξεργασία. Η συχνότητα 2,4 GHz χρησιμοποιείται αρκετά χωρίς να απαιτείται άδεια, για αυτόν τον λόγο σε πραγματικά περιβάλλοντα λόγω των πολλών εκπομπών από άλλες συσκευές δημιουργούνται εξωτερικές παρεμβολές δυσκολεύοντας την ακεραιότητα του συστήματος.

Οι δύο πλακέτες XIAO ESP32-C6 τοποθετήθηκαν σε σταθερά σημεία στον χώρο, ο Tx σε ένα άκρο του δωματίου και ο Rx στο άλλο σε ύψος περίπου 1,0-1,7 m (τοποθέτηση πάνω σε τραπέζι, κολλημένο πάνω σε τοίχο), χωρίς την μετατόπιση τους μεταξύ των καταγραφών. Η διάταξη σχεδιάστηκε ώστε να καλύπτει τρία σενάρια: (α) άμεση οπτική επαφή (Line-of-Sight, LoS) όταν ο άνθρωπος κινείται μεταξύ των δύο μικροελεγκτών πάνω στον κεντρικό άξονα Tx–Rx· (β) (partial NLoS) όταν ο άνθρωπος βρίσκεται πλαγίως, έως ~1 m από τον κεντρικό άξονα LoS· και (γ) πλήρης αποκλεισμός οπτικής επαφής, όπου υπήρχε ανταπόκριση μόνο εντός του δωματίου. Σε μεγάλες αποστάσεις NLoS ο δέκτης δεν λαμβάνει αξιόπιστα αποτελέσματα, όπως διαπιστώθηκε κατά τις δοκιμές.

Σχετικά με την απόσταση T_x-R_x , εξετάστηκαν αποστάσεις μεταξύ 2 και 5 m καθώς και σε μέγιστη απόσταση 10-14 μέτρων. Διευκρινίζεται ότι το βασικό σύνολο των 100 καταγραφών συλλέχθηκε εντός του δωματίου (80 καταγραφές-room1 και 20-living room), σε αποστάσεις T_x-R_x 2–5 m. Η απόσταση των 14 m αφορά ξεχωριστή δοκιμή εμβέλειας σε μεγαλύτερο χώρο περιλαμβάνοντας το σαλόνι, ένα δωμάτιο και τον διάδρομο του σπιτιού και δεν συμμετέχει στα αποτελέσματα του Κεφαλαίου 5. Σε LoS συνθήκες, η ανίχνευση κίνησης είναι αξιόπιστη σε όλο το εύρος αποστάσεων, ενώ σε NLoS η ακρίβεια μειώνεται εκτός και εάν ο άνθρωπος βρίσκεται εντός 1 m μακριά από τον κεντρικό άξονα LoS. Τα συμπεράσματα αυτά ισχύουν για τις δύο περιπτώσεις αποστάσεων. Τα πειράματα πραγματοποιήθηκαν από έναν μόνο άνθρωπο, σε φυσιολογικές συνθήκες δωματίου χωρίς άλλα κινούμενα αντικείμενα ή πρόσωπα κατά τη διάρκεια των καταγραφών. Στις παρακάτω εικόνες (Εικόνα 4.5 και Εικόνα 4.6) παρουσιάζεται η τοποθέτηση του πομπού και του δέκτη, αντίστοιχα, στον χώρο του σπιτιού.



Εικόνα 4.5. Τοποθέτηση του Πομπού στο χώρο της δοκιμής εμβέλειας στα 14 μέτρα, απόσταση από τον Δέκτη (LoS).



Εικόνα 4.6. Τοποθέτηση του Δέκτη στο χώρο της δοκιμής εμβέλειας στα 14 μέτρα, απόσταση από τον Πομπό (LoS).

4.4 Πρωτόκολλο Συλλογής Δεδομένων

Στη συλλογή δεδομένων ακολουθήθηκε το πρωτόκολλο ESPNOW που παρέχει η ίδια η εταιρεία λόγω της ταχείας επικοινωνίας μεταξύ των μικροελεγκτών [30]. Κάθε καταγραφή αντιστοιχεί σε μία συνεχή καταγραφή μιας κλάσης δραστηριότητας διάρκειας περίπου ~ 60 δευτερολέπτων, παράγοντας δηλαδή $100 \text{ Hz} \times 60 \text{ s} \approx 6.000$ ακατέργαστα πλαίσια (Frames) CSI ανά καταγραφή (session). Στην πράξη, ο αριθμός πλαισίων ανά καταγραφή μπορεί να διαφέρει ελαφρώς (π.χ. 6.003 όπως φαίνεται στην Εικόνα 4.29) λόγω αναπόφευκτης μεταβλητότητας κατά την διαδικασία της συλλογής.

Ορίστηκαν δύο συνολικές κλάσεις δραστηριότητας (LOGGER_ACTIVITY_PRESETS, config.py): Walk (βάδιση, ο άνθρωπος βαδίζει εντός του δωματίου με φυσιολογικό ρυθμό, χωρίς ακαριαίες κινήσεις, καλύπτοντας τον συνολικό χώρο) και Idle / No Activity (ακίνησία, ο άνθρωπος παραμένει ακίνητος σε φυσιολογικές συνθήκες (όχι σε στάση αγάλματος) σε σχεδόν όλα τα σημεία του χώρου είτε όρθιος είτε καθιστός, χωρίς σκόπιμη κίνηση). Συλλέχθηκαν συνολικά 50 καταγραφές ανά κλάση, δηλαδή 100 καταγραφές συνολικά, παράγοντας ονομαστικά $50 \times 6.000 = 300.000$ ακατέργαστα πλαίσια ανά κλάση και περίπου 600.000 πλαίσια στο σύνολο. Οι καταγραφές πραγματοποιήθηκαν σε περισσότερες από μία από τις διατάξεις όπως αναφέρεται και στο κεφάλαιο 4.3 από έναν συνδυασμό

αποστάσεων (2–5 m) και σεναρίων LoS/μερικού NLoS, κυρίως εντός του ίδιου χώρου (room1), με ένα μικρό υποσύνολο καταγραφών σε δεύτερο χώρο (σαλόνι).

Κάθε session/καταγραφή αντιμετωπίζεται ως ένα μοναδικό αρχείο ώστε να πραγματοποιείται έλεγχος cross-validation, αποτρέποντας τη διαρροή δεδομένων κατά την εκπαίδευση (Data Leakage) [7]. Για περισσότερες λεπτομέρειες σχετικά με την υλοποίηση του κώδικα γίνεται παραπομπή στο Παράρτημα C.

4.5 Σύνολο Δεδομένων και Επαύξηση

4.5.1 Ακατέργαστα Δεδομένα και Παράθυρα sliding window

Από τα 600.000 ακατέργαστα πλαίσια CSI, εφαρμόζεται τεμαχισμός σε (sliding window) μετά την προεπεξεργασία. Οι παράμετροι παραθύρου είναι: μέγεθος παραθύρου $W = 100$ δείγματα (= 1 δευτερόλεπτο στα 100 Hz) και βήμα $S = 50$ δείγματα (50% επικάλυψη, PIPELINE_STEP_SIZE = 50). Με αυτές τις παραμέτρους προκύπτουν περίπου 118 παράθυρα ανά καταγραφή, δηλαδή περίπου 11.800 παράθυρα συνολικά πριν την επαύξηση. Τέλος, το παράθυρο 1 δευτερολέπτου επιλέχθηκε ώστε να επιτυγχάνεται ταχεία πρόβλεψη [1][19].

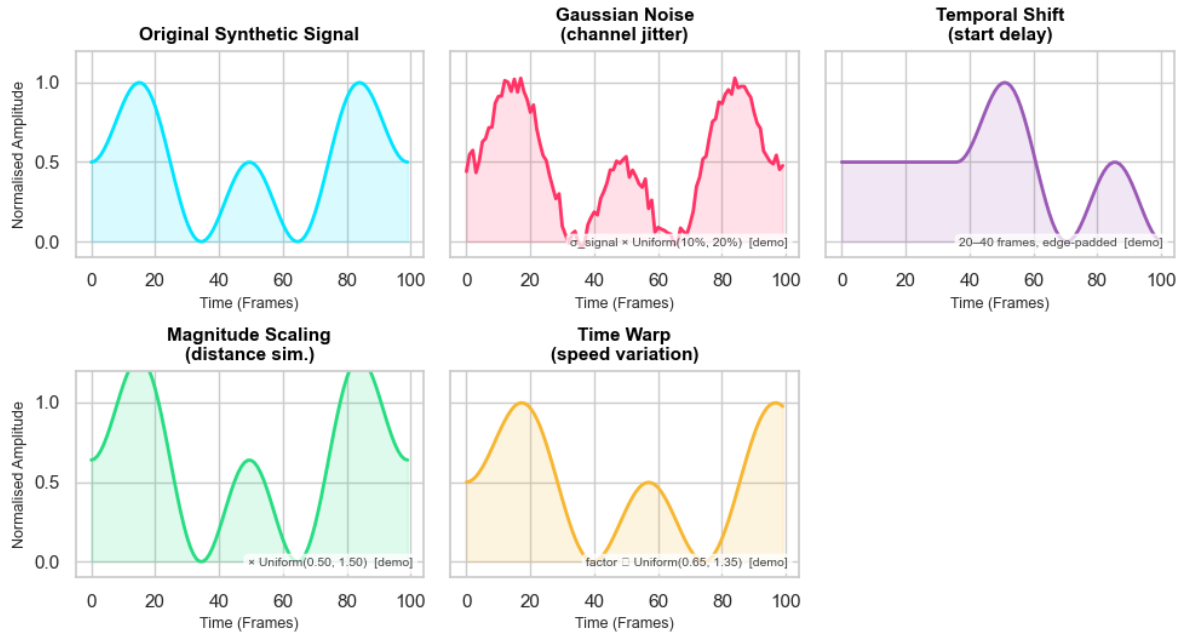
4.5.2 Επαύξηση Δεδομένων (Data Augmentation)

Για την αύξηση της ποικιλομορφίας του training set και τη μείωση της υπερεκπαίδευσης λόγω του μικρού όγκου δεδομένων (Overfitting), εφαρμόστηκε τεχνική επαύξησης δεδομένων [9] ώστε για κάθε αρχικό παράθυρο εκπαίδευσης να παράγονται ($N_AUGMENTS = 4$) 4 επιπλέον παράθυρα μέσω τυχαίων μετασχηματισμών (Gaussian θορύβου, κλίμακα πλάτους, χρονική ολίσθηση, χρονική παραμόρφωση). Το αποτέλεσμα είναι να έχουν πολλαπλασιαστεί τα δεδομένα επί 5 (1 πρωτότυπο + 4 επαυξημένα) χωρίς την προσθήκη άλλου υλικού καταγραφής.

Ένα από τα κρίσιμα σημεία είναι ότι το cross-validation εκτελείται αποκλειστικά σε πραγματικά, μη επαυξημένα παράθυρα, η επαύξηση εφαρμόζεται μόνο κατά την τελική εκπαίδευση του μοντέλου στο πλήρες σύνολο εκπαίδευσης (80%) που αξιολογείται στο ανεξάρτητο σύνολο ελέγχου, ποτέ στα δεδομένα επικύρωσης ή ελέγχου. Αυτή η μέθοδος αποτρέπει τη μορφή διαρροής δεδομένων (Data Leakage) που προκαλείται όταν οι επαυξημένες εκδοχές δεδομένων συμπεριλαμβάνονται στην εκπαίδευση, ακόμα και έμμεσα [7]. (Παράρτημα D)

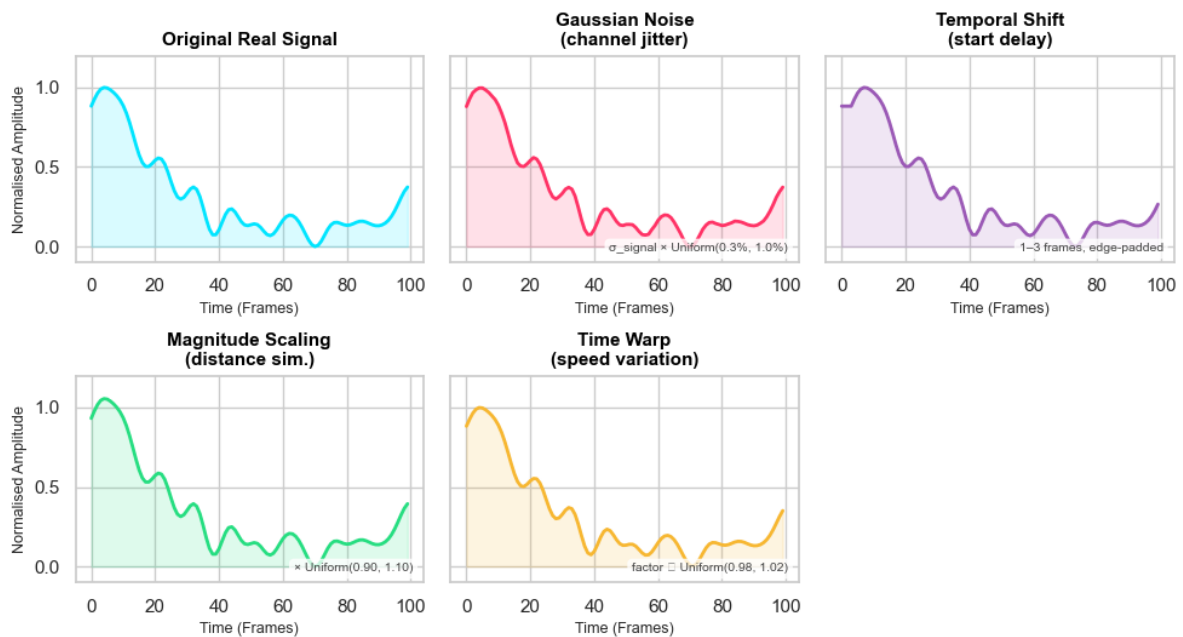
Στις παρακάτω εικόνες (4.7 έως 4.9) επιβεβαιώνεται πως εφαρμόζεται πραγματικά η επαύξηση δεδομένων στα πραγματικά αρχεία. Αν και οι πραγματικές τιμές με τις επαυξημένες έχουν απειροελάχιστες διαφορές και δεν είναι εμφανείς στο ανθρώπινο μάτι, παρακάτω παρουσιάζεται η επαύξηση των δεδομένων σε πραγματικά δεδομένα, καθώς και ένα demo με ακραίες τιμές επαύξησης ώστε να επιβεβαιωθεί πραγματικά το αποτέλεσμα της τεχνικής αυτής.

Data Augmentation — Synthetic CSI Signal | class: walk_activity
Exaggerated parameters for thesis visualization — training values unchanged

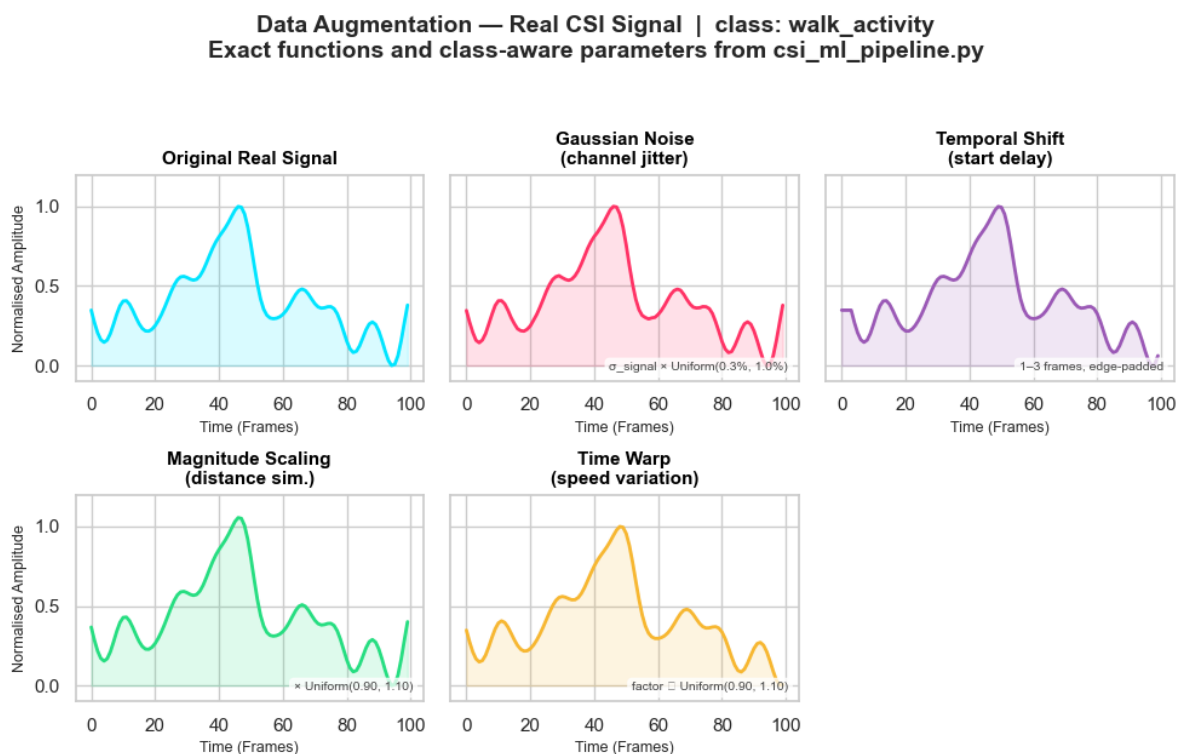


Εικόνα 4.7. Επαύξηση δεδομένων σε μη πραγματικά δεδομένα

Data Augmentation — Real CSI Signal | class: no_activity
Exact functions and class-aware parameters from csi_ml_pipeline.py



Εικόνα 4.8. Επαύξηση δεδομένων σε πραγματικά δεδομένα στο αρχείο
no_activity_11_room1_vasilis_1777973010.txt



Εικόνα 4.9. Επαύξηση δεδομένων σε πραγματικά δεδομένα στο αρχείο
walk_activity_45_room1_vasilis_1777806476.txt

4.6 Προεπεξεργασία Σήματος

Το στάδιο προεπεξεργασίας αποτελείται από 6 συνολικά στάδια, τα τέσσερα πρώτα εφαρμόζονται ανά υποφορέα CSI ανεξάρτητα, ενώ τα δύο τελευταία (PCA και κανονικοποίηση) λειτουργούν στο σύνολο των υποφορέων. Το πρώτο στάδιο είναι το στάδιο αφαίρεσης των κενών υποφορέων (Null Subcarrier Removal), στη συνέχεια χρησιμοποιείται το φίλτρο Hampel (Hampel Filter), έπειτα εφαρμόζεται ένα βαθυπερατό φίλτρο Butterworth (Low-Pass Filter), χρονική διαφοροποίηση (Temporal Difference), χρήση PCA για μείωση περιττής πληροφορίας και τέλος η κανονικοποίηση δεδομένων [9][19][21].

Η επιλογή (pipeline) προεπεξεργασίας ποικίλλει σημαντικά στη βιβλιογραφία. Δεν υπάρχει βέλτιστη λύση και κάθε εφαρμογή έχει τις δικές της απαιτήσεις. Παρ' όλα αυτά, κάποια μοτίβα επαναλαμβάνονται.

Το φίλτρο Butterworth εμφανίζεται σε πολλές εργασίες. Οι Hao et al. [15] χρησιμοποιούν Butterworth σε συνδυασμό με wavelet για την απόρριψη των ακραίων τιμών (outliers) σε σύστημα αναγνώρισης πολύπλοκων κινήσεων πολεμικών τεχνών.

Οι Abdurrahman & Irawan [9] χρησιμοποιούν Hampel φίλτρο σε συνδυασμό με Savitzky-Golay. Αυτός ο συνδυασμός αποδίδει σταθερά αποτελέσματα σε ESP32-based pipeline και χρησιμεύει ως αναφορά για τη δική μας επιλογή Hampel + Butterworth.

Ορισμένες εργασίες εφαρμόζουν Principal Component Analysis (PCA) ή Independent Component Analysis (ICA) για μείωση πληροφορίας [14]. Αντίθετα, άλλες αποφεύγουν εντελώς την εξαγωγή χαρακτηριστικών και τροφοδοτούν απευθείας τα ακατέργαστα δεδομένα (raw) CSI σε μοντέλα βαθιάς μάθησης CNN ή LSTM [15][16].

4.6.1 Αφαίρεση Null Subcarriers

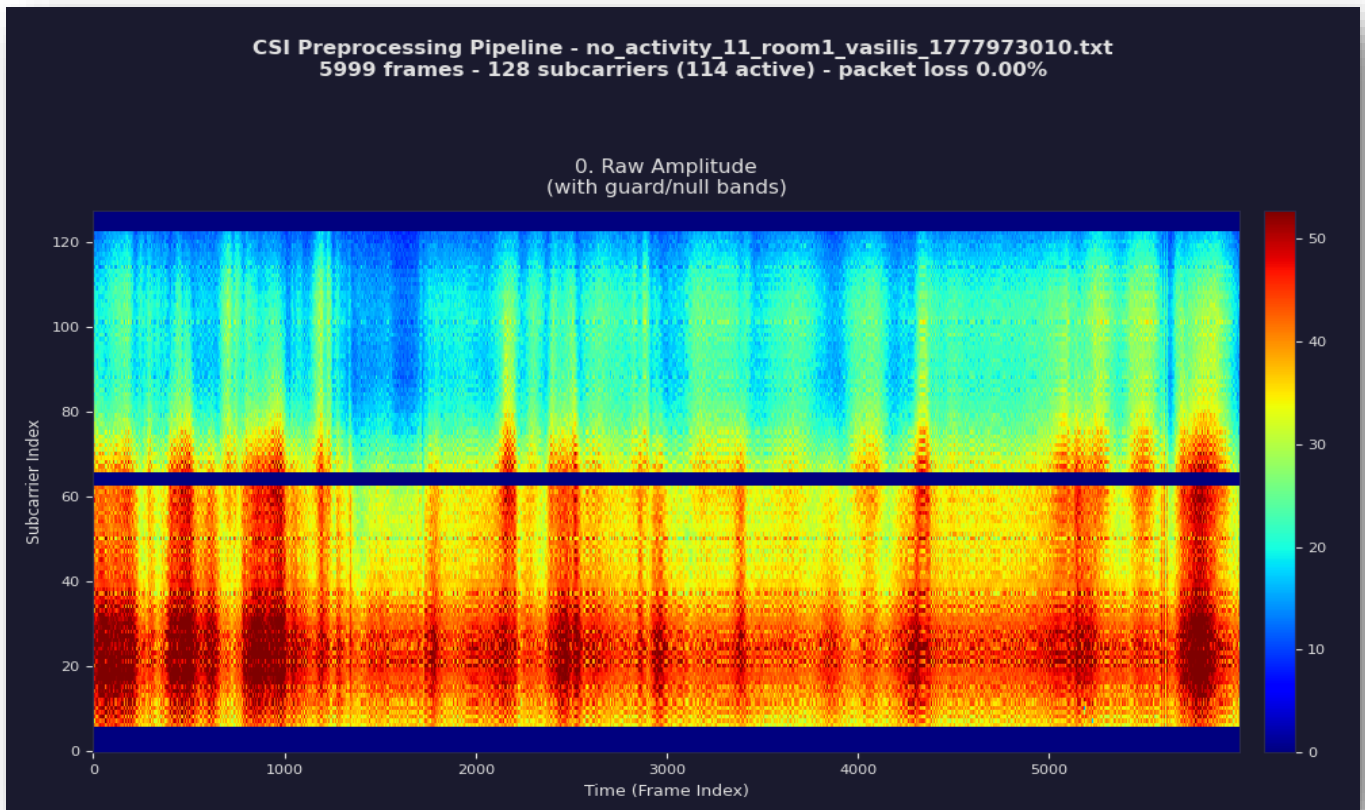
Όπως αναφέρθηκε προηγουμένως, σε ένα σύστημα HT40 από τους 128 συνολικούς υποφορείς μόνο 114 είναι ενεργοί. Οι υπόλοιποι 14 είναι null subcarriers που χρησιμοποιούνται ως guard bands ή αντιστοιχούν στην DC και δεν μεταφέρουν πληροφορία [1][3]. Αυτοί οι null subcarriers έχουν τυπικά μηδενικό ή σχεδόν μηδενικό πλάτος στο CSI.

Επομένως, η αφαίρεση των null subcarriers επιτυγχάνεται με τη δημιουργία μιας μάσκας (mask) που προσδιορίζει ποιοι υποφορείς είναι ενεργοί (ποιοι δηλαδή δεν έχουν μηδενική τιμή). Στην δική μας περίπτωση, οι null subcarriers βρίσκονται στα άκρα του φάσματος και γύρω από τη DC (κεντρική συχνότητα) (Εικόνα 4.10 έως 4.17).

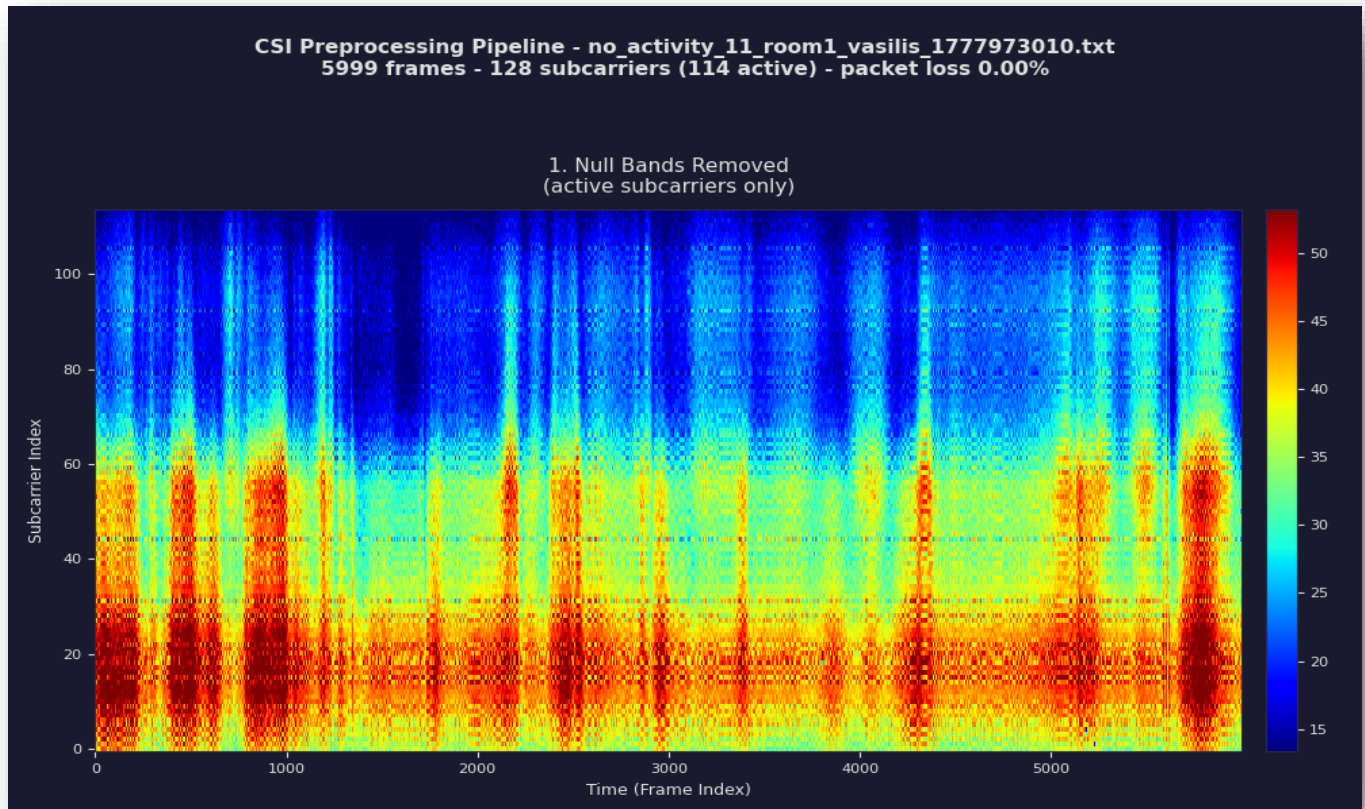
Κατά την υλοποίηση, η μάσκα υπολογίζεται αυτόματα με βάση το πλάτος των CSI δεδομένων. Συγκεκριμένα, διατηρούνται οι υποφορείς των οποίων το μέσο πλάτος υπερβαίνει ένα μικρό σταθερό κατώφλι ($1e-3$). Οι null υποφορείς εμφανίζουν μηδενικό πλάτος, σε σαφή αντίθεση με τους ενεργούς, οπότε ο διαχωρισμός είναι ξεκάθαρος και η μάσκα καταλήγει ακριβώς στους 114 ενεργούς υποφορείς του προτύπου. Παρακάτω παρουσιάζεται ένα απόσπασμα από τον κώδικα υλοποίησης.

```
amplitude = np.abs(csi_complex)
mean_amplitude = np.mean(amplitude, axis=0)
threshold = 1e-3 # σταθερό κατώφλι μη-μηδενικού πλάτους
active_mask = mean_amplitude > threshold
csi_active = csi_complex[:, active_mask]
```

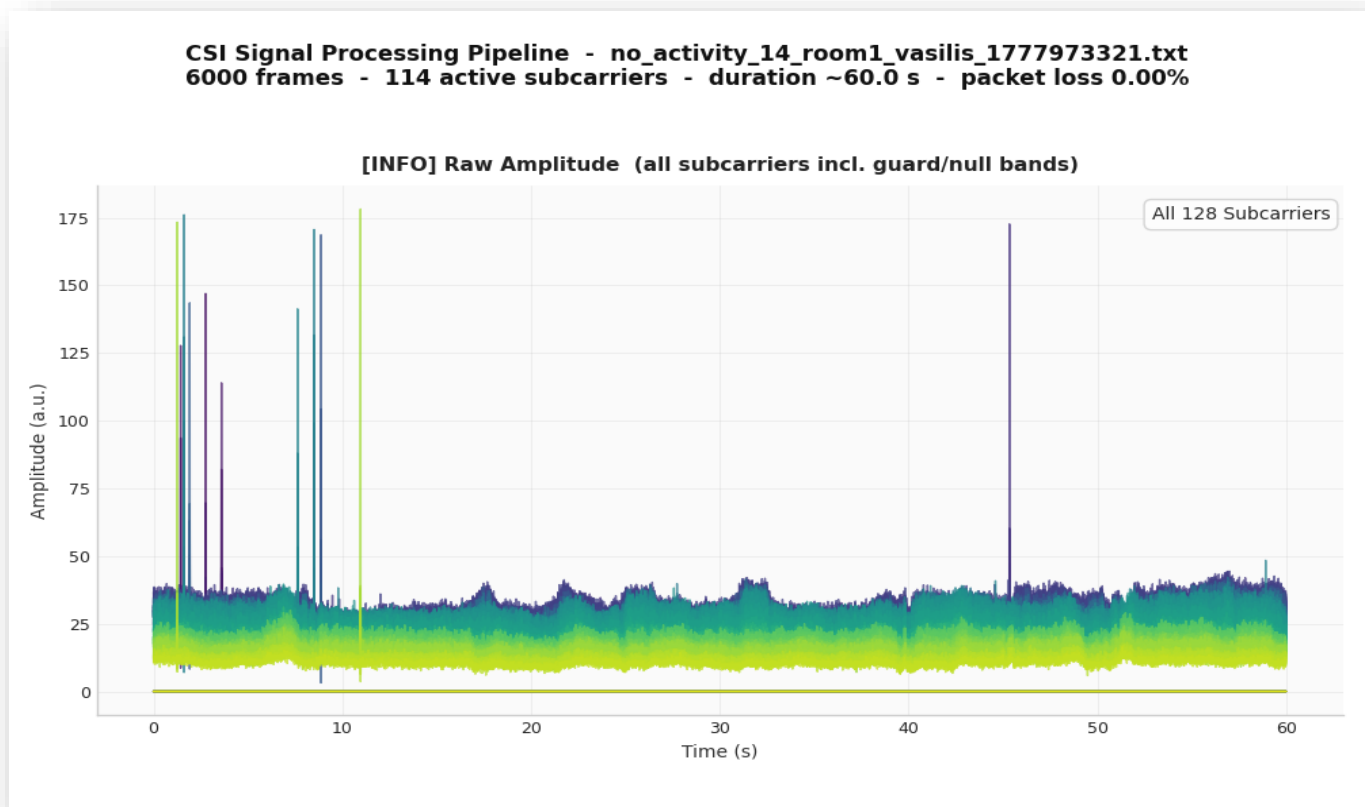
Στα παρακάτω διαγράμματα συγκρίνεται ο τρόπος αφαίρεσης των ακραίων τιμών, καθώς και παρουσιάζονται οι κυματομορφές των 2 κλάσεων (Εικόνα 4.10 έως 4.17).



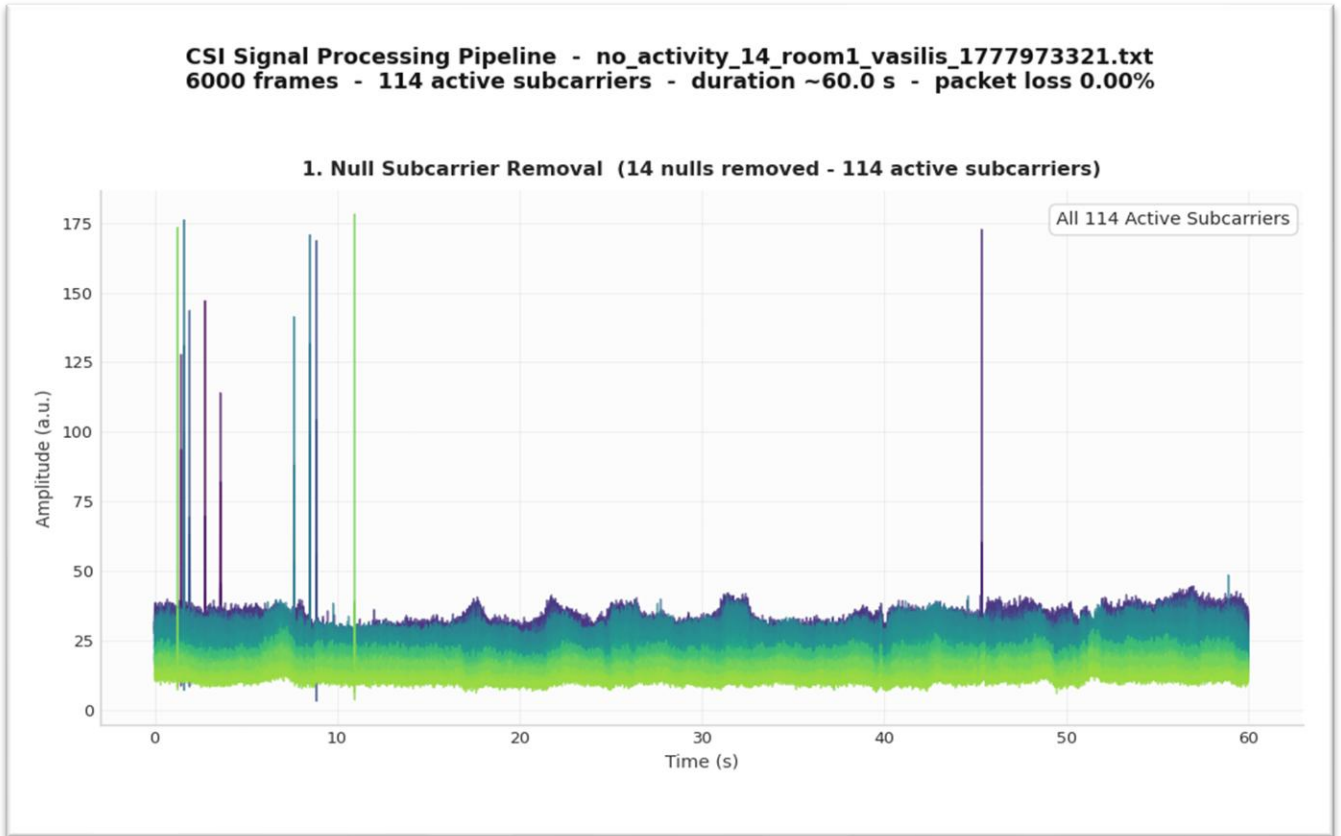
Εικόνα 4.10. Το ακατέργαστο σήμα από το αρχείο no_activity_11_room1_vasilis_1777973010.txt (Heatmap Diagram).



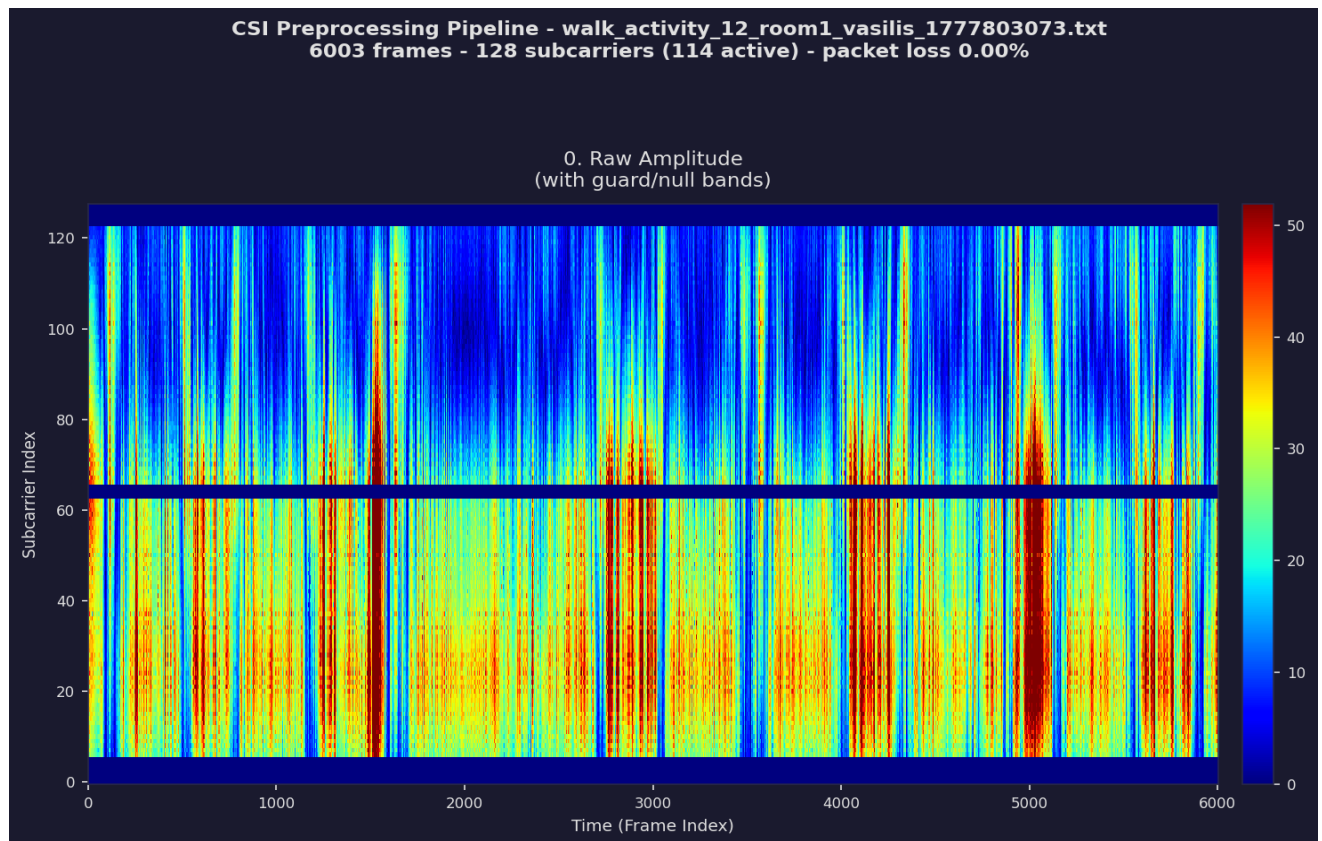
Εικόνα 4.11. Αφαίρεση μηδενικών τιμών από το αρχείο
no_activity_11_room1_vasilis_1777973010.txt (Heatmap Diagram).



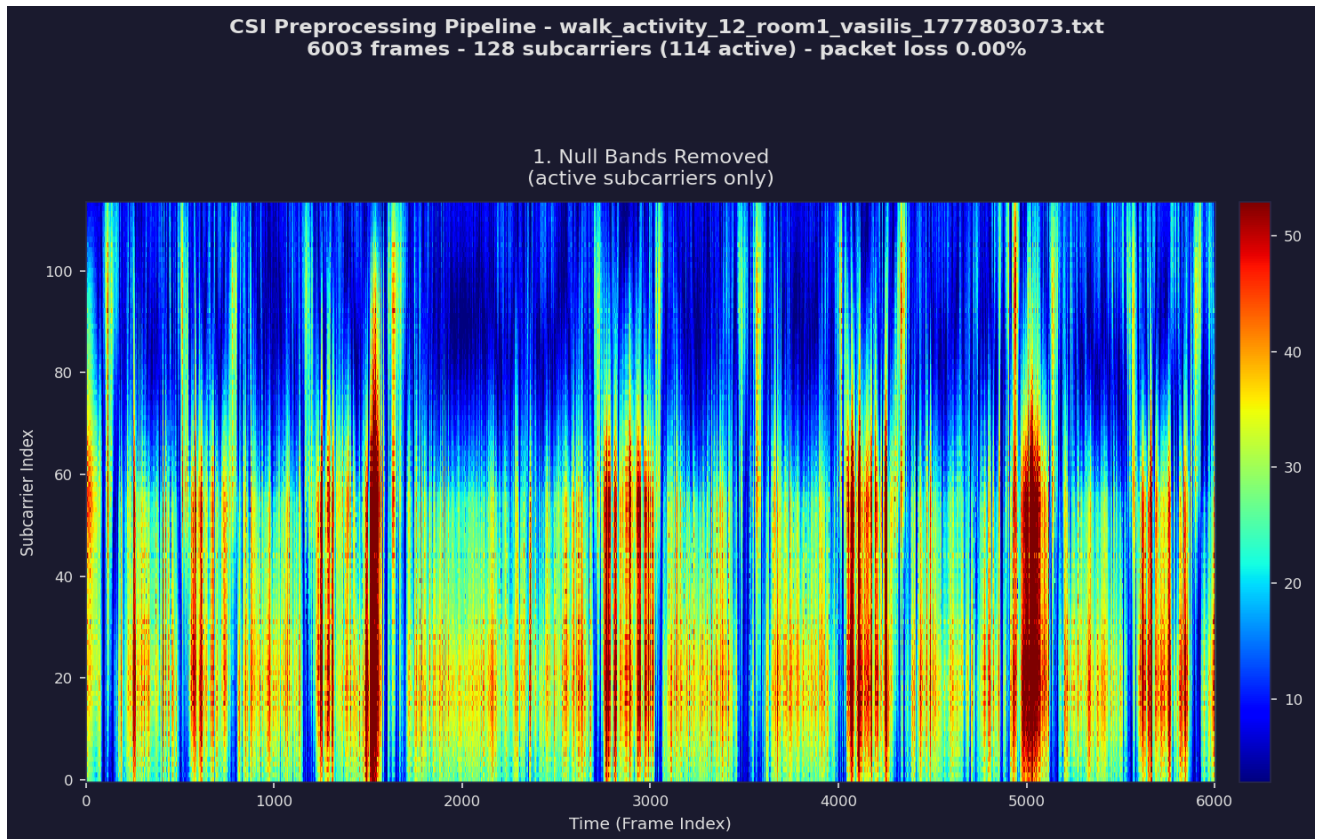
Εικόνα 4.12. Το ακατέργαστο σήμα από το αρχείο no_activity_14_room1_vasilis_1777973321.txt από 128 υποφορείς σε 114 (Plot Diagram).



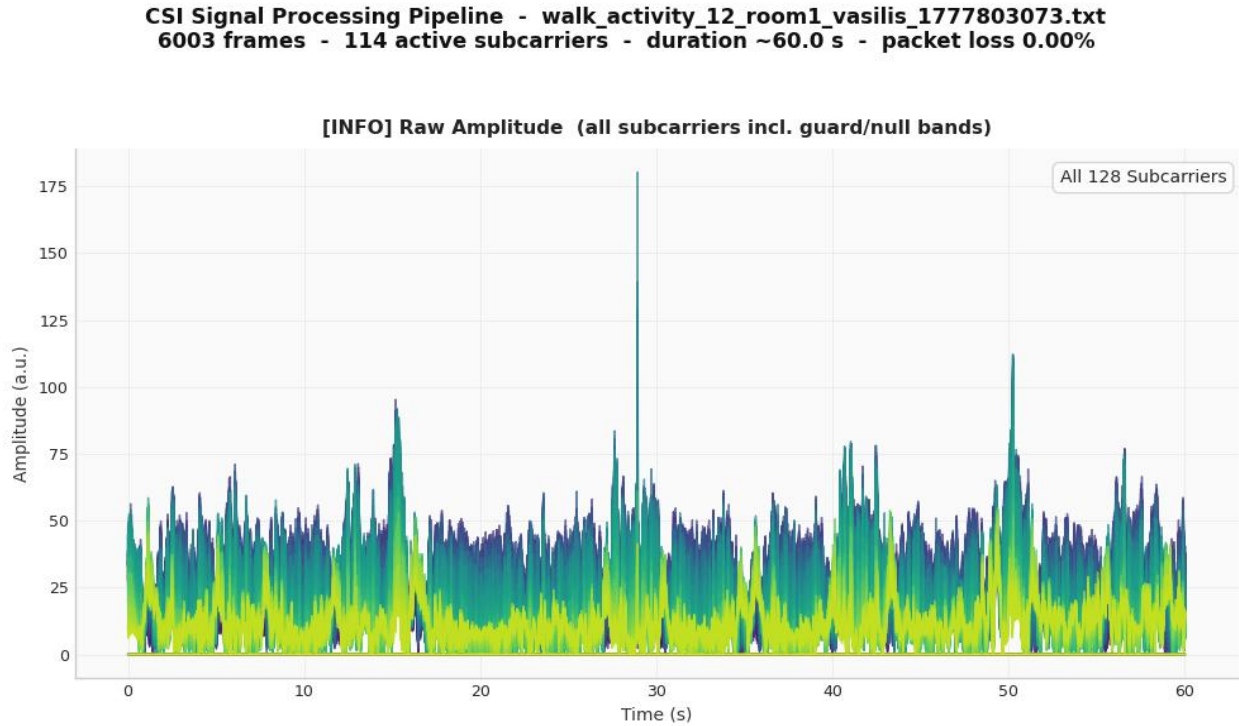
Εικόνα 4.13. Αφαίρεση μηδενικών τιμών από το αρχείο no_activity_14_room1_vasilis_1777973321.txt από 128 υποφορείς σε 114.



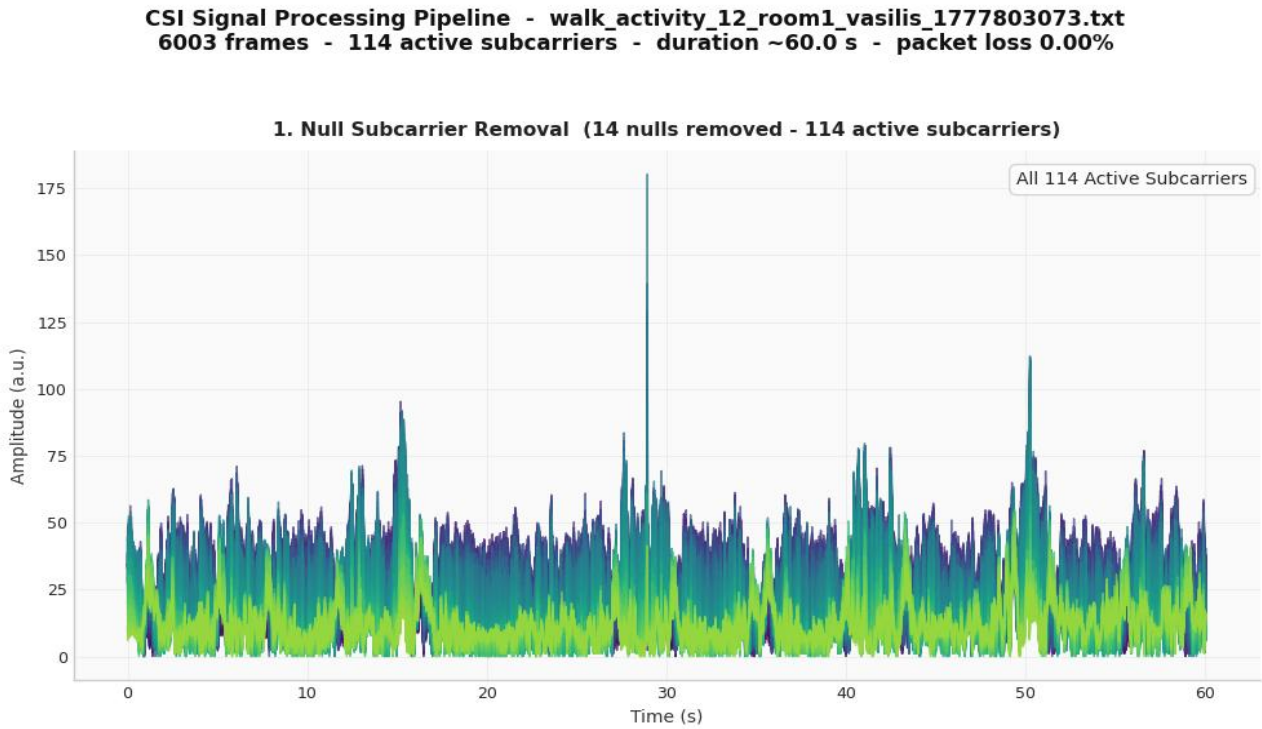
Εικόνα 4.14. Το ακατέργαστο σήμα από το αρχείο walk_activity_12_room1_vasilis_1777803073.txt (Heatmap Diagram)



Εικόνα 4.15. Αφαίρεση μηδενικών τιμών από το αρχείο walk_activity_12_room1_vasilis_1777803073.txt (Heatmap Diagram)



Εικόνα 4.16. Το ακατέργαστο σήμα από το walk_activity_12_room1_vasilis_1777803073.txt από 128 υποφορείς σε 114.



Εικόνα 4.17. Αφαίρεση μηδενικών τιμών από το αρχείο walk_activity_12_room1_vasilis_1777803073.txt από 128 υποφορείς σε 114.

4.6.2 Φίλτρο Hampel — Αφαίρεση Ακραίων Τιμών

Το φίλτρο Hampel αφαιρεί ακραίες τιμές (outliers) από το πλάτος κάθε υποφορέα. Για κάθε θέση i ενός κυλιόμενου παραθύρου υπολογίζεται η διάμεσος m_i του παραθύρου και ένας εύρωστος εκτιμητής της τυπικής απόκλισης μέσω της διαμέσου των απόλυτων αποκλίσεων (Median Absolute Deviation, MAD). Παρακάτω παρουσιάζεται ο τύπος υπολογισμού του φίλτρου [9][21]:

$$\sigma_{HAM} = 1.4826 \cdot \text{median}(|x_j - m_i|), \quad j \in \text{window} \quad (4.2)$$

Κάθε δείγμα x_i για το οποίο $|x_i - m_i| > 3 * \sigma_{HAM}$ αντικαθίσταται από τη διάμεσο m_i . Η χρήση διαμέσου και MAD αντί για την μέση τιμή και την τυπική απόκλιση προσδίδει ανθεκτικότητα (robust) στα ακραία δείγματα, τα οποία και αφαιρούνται. Το αποτέλεσμα είναι η αφαίρεση ακραίων τιμών (spikes) όπως παρατηρούνται στις προηγούμενες εικόνες καθώς και στις παρακάτω εικόνες που οφείλονται συνήθως σε παροδικές παρεμβολές ή packet errors [9][19].

Παρακάτω παρουσιάζεται ο κώδικας υλοποίησης σε γλώσσα python που αφαιρεί τις ακραίες τιμές στο πλάτος του σήματος.

```
amplitude = np.abs(csi_complex)

df = pd.DataFrame(amplitude)

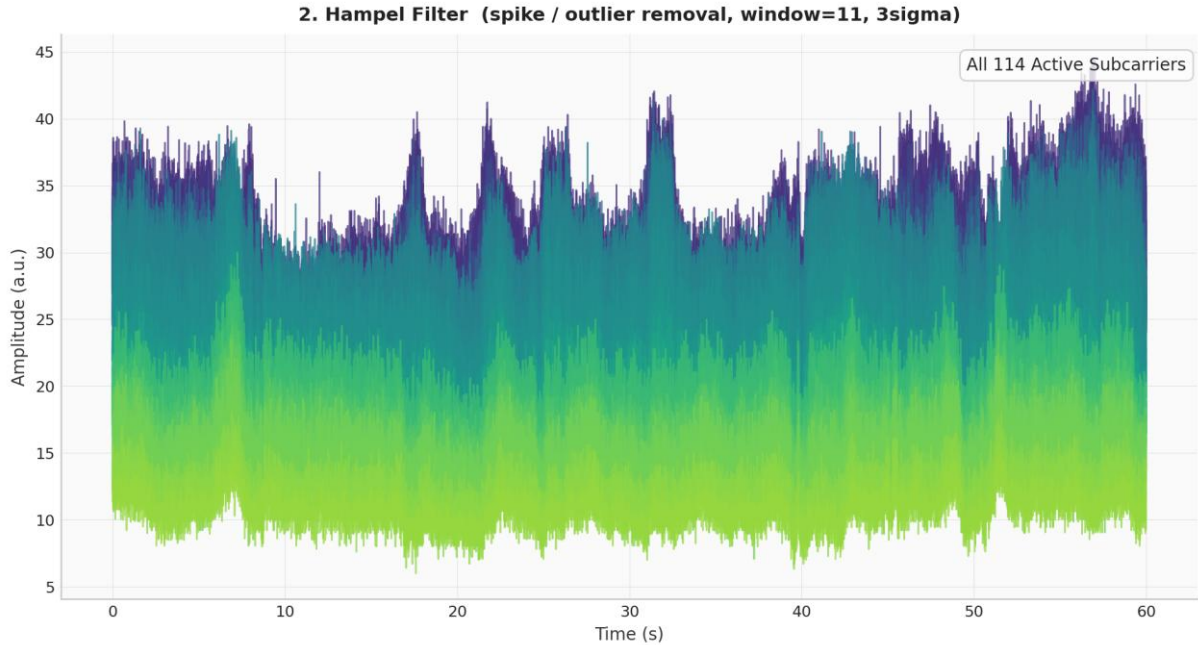
rolling_median = df.rolling(window=window_size, center=True).median()
rolling_mad = df.rolling(window=window_size, center=True).apply(
    lambda x: np.median(np.abs(x - np.median(x)))
)

threshold = k * 1.4826 * rolling_mad
outliers = np.abs(df - rolling_median) > threshold
df[outliers] = rolling_median[outliers]

Μετά τον καθαρισμό του amplitude, ανακατασκευάζεται η μιγαδική μήτρα CSI (μόνο για λόγους
συμβατότητας με τα επόμενα στάδια του pipeline, η φάση δεν αξιοποιείται ως χαρακτηριστικό):

csi_filtered = amplitude_cleaned * np.exp(1j * phase)
```

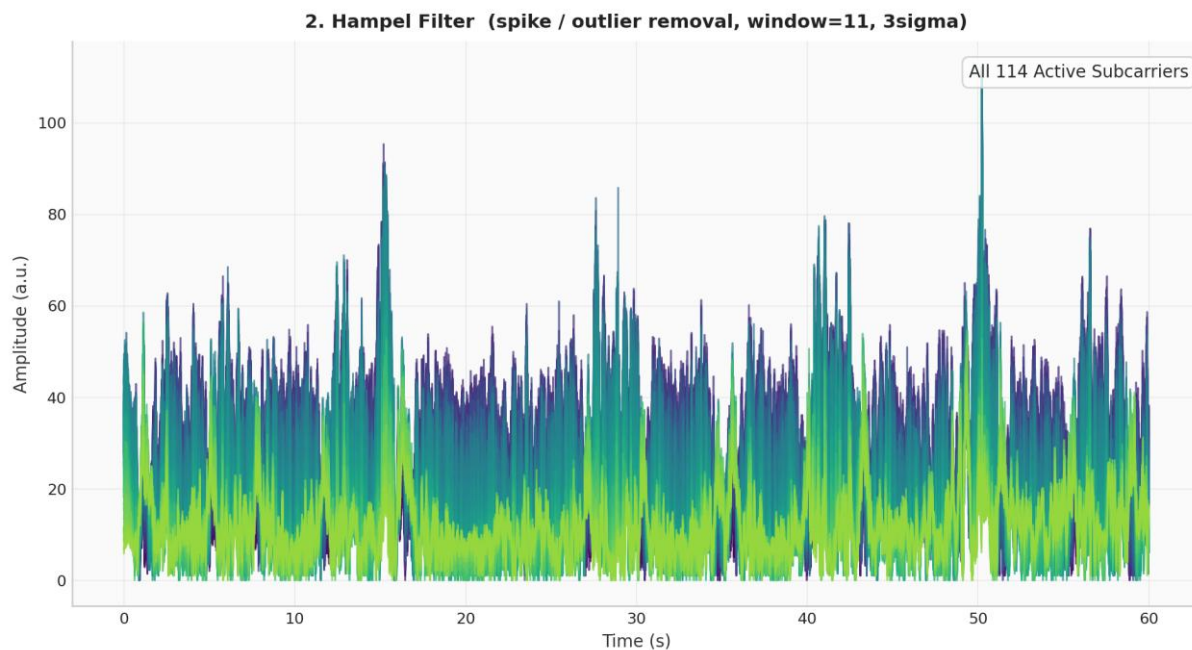
CSI Signal Processing Pipeline - no_activity_14_room1_vasilis_1777973321.txt
6000 frames - 114 active subcarriers - duration ~60.0 s - packet loss 0.00%



Εικόνα 4.18. Εφαρμογή του Φίλτρου Hampel στο αρχείο
no_activity_14_room1_vasilis_1777973321.txt.

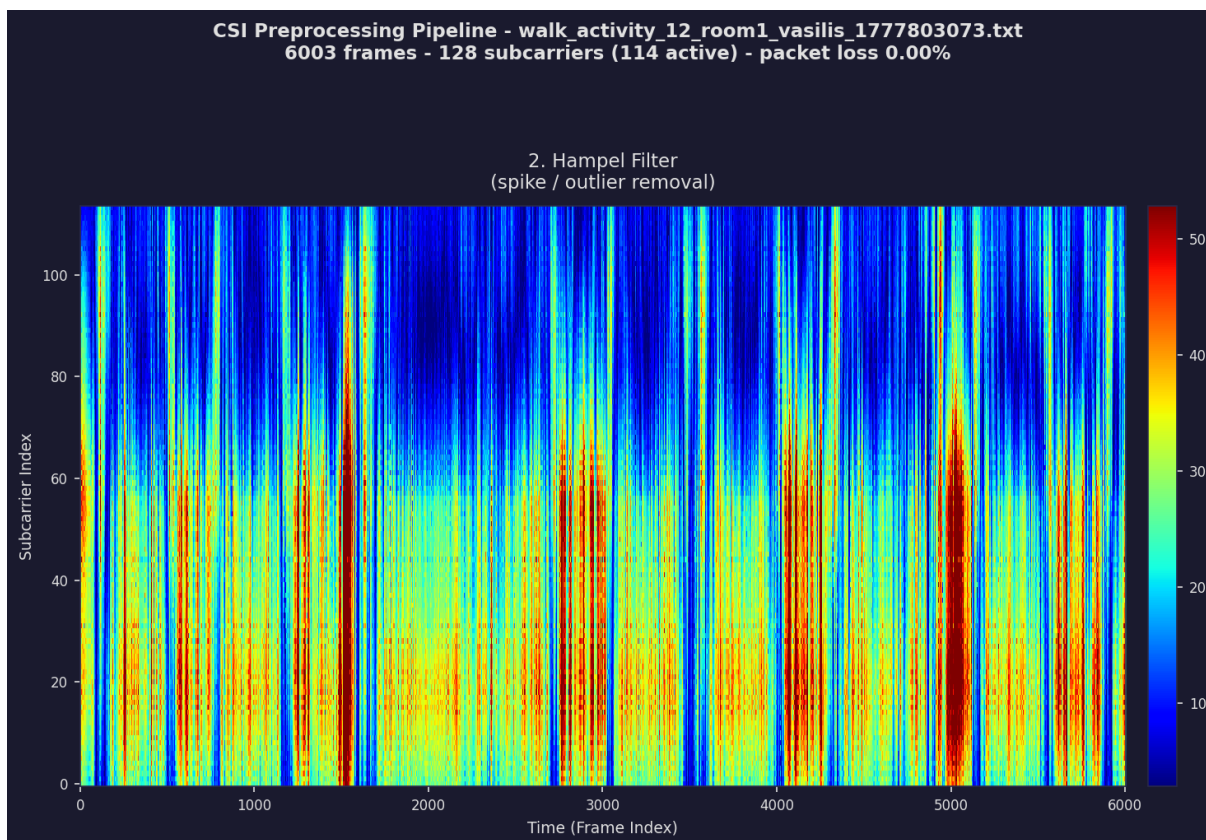
Συγκρίνοντας με την προηγούμενη εικόνα (Εικόνα 4.13) του αρχείου με την τωρινή παρατηρείται ότι το πλάτος στο ακατέργαστο σήμα έφτανε τιμές έως και 175 λόγω των ακραίων τιμών, ενώ τώρα έχει περιοριστεί σημαντικά ανάμεσα στο 40 με 45. Με αυτό τον τρόπο η ανάλυση επικεντρώνεται στην πραγματική ουσία του σήματος. Τέλος, παρουσιάζεται συνολικά η πληροφορία και των 114 ενεργών υποφορέων. Αντίστοιχα ισχύει ακριβώς το ίδιο και για την κλάση walk στην παρακάτω εικόνα (Εικόνα 4.19).

CSI Signal Processing Pipeline - walk_activity_12_room1_vasilis_1777803073.txt
6003 frames - 114 active subcarriers - duration ~60.0 s - packet loss 0.00%

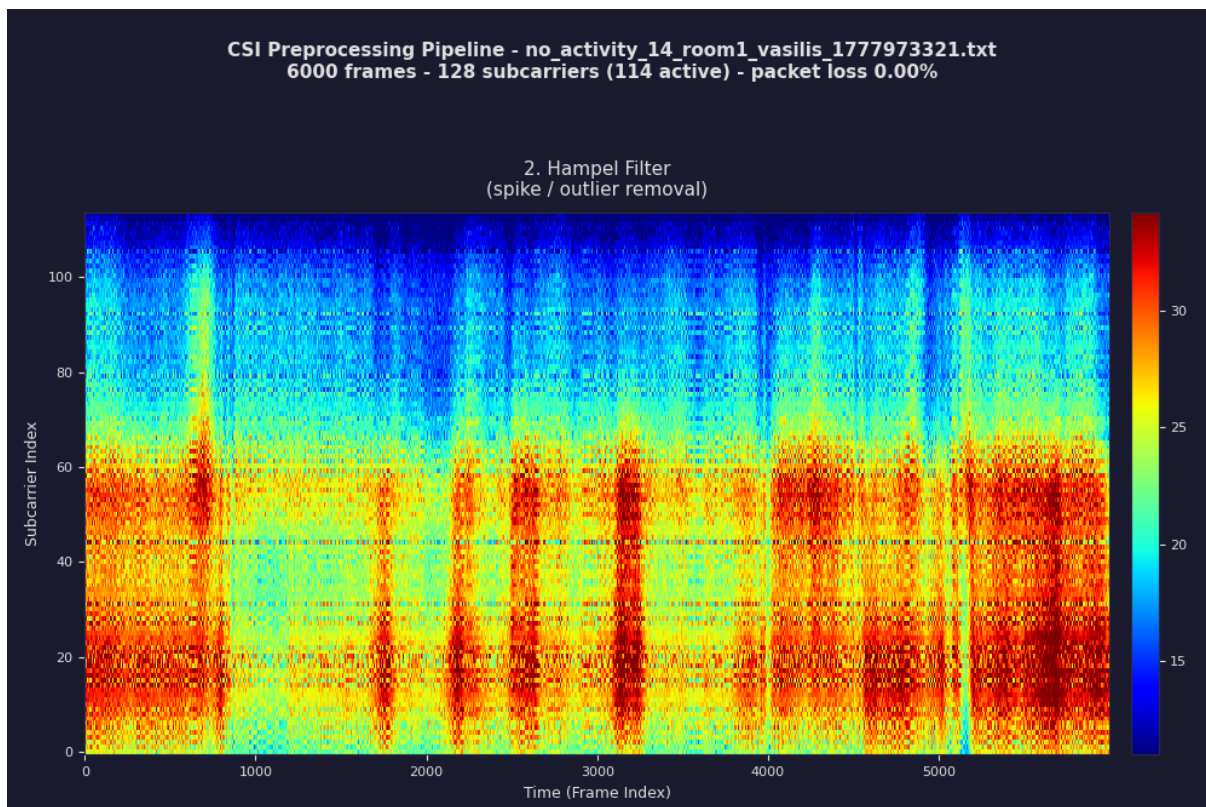


Εικόνα 4.19. Εφαρμογή του Φίλτρου Hampel στο walk_activity_12_room1_vasilis_1777803073.txt.

Παρατηρώντας τις εικόνες (Εικόνα 4.20, Εικόνα 4.21) και σε σύγκριση με την κλάση idle, εδώ εντοπίζονται περισσότερες διακυμάνσεις του σήματος και περισσότερο θόρυβο. Το αποτέλεσμα του διαγράμματος αυτού είναι φυσιολογικό καθώς το περπάτημα του ανθρώπου διαταράσσει τους υποφορείς. Τέλος, παρατηρείται ότι οι υποφορείς με δείκτες κοντά στο 15, στο 30 και στο 50 έχουν επηρεαστεί πολύ περισσότερο από τους άλλους.



Εικόνα 4.20. Εφαρμογή του Φίλτρου Hampel στο walk_activity_12_room1_vasilis_1777803073.txt (Heatmap Diagram),



Εικόνα 4.21. Εφαρμογή του Φίλτρου Hampel στο no_activity_14_room1_vasilis_1777973321.txt (Heatmap Diagram).

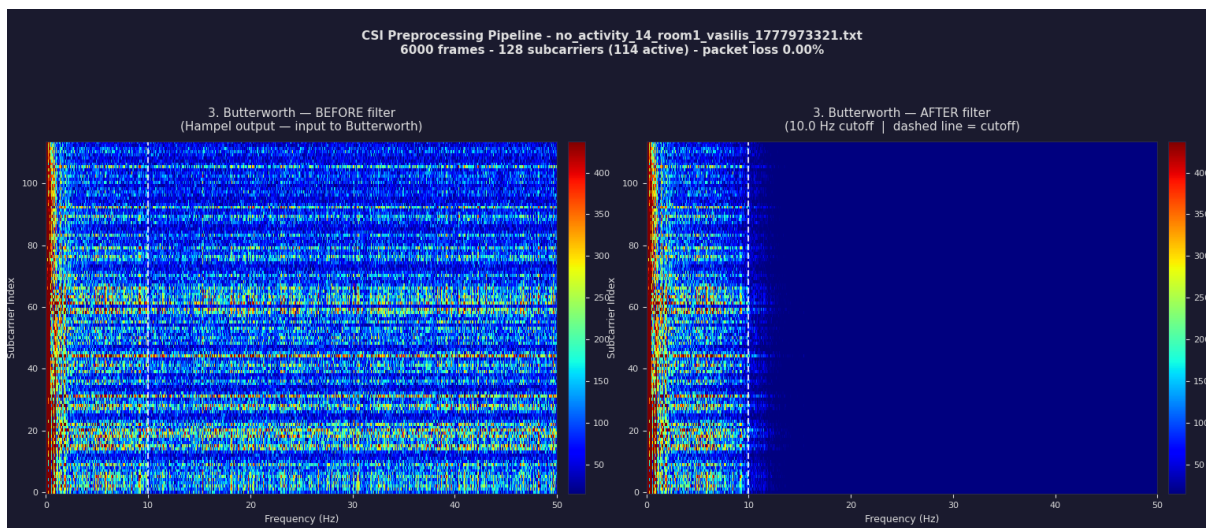
4.6.3 Φίλτρο Butterworth Χαμηλής Συχνότητας

Στο στάδιο αυτό, εφαρμόζεται το φίλτρο Butterworth (`scipy.signal.butter`, `order=4`) με συχνότητα αποκοπής στα $f_c = 10,0$ Hz [3], [21] (`cutoff_frequency = 10.0` Hz στο αρχείο `config.py`). Ο σχεδιασμός Butterworth επιλέχθηκε ώστε να επικεντρωθεί η ανάλυση στο εύρος ζώνης στο οποίο εμφανίζεται η ανθρώπινη κίνηση [3][19].

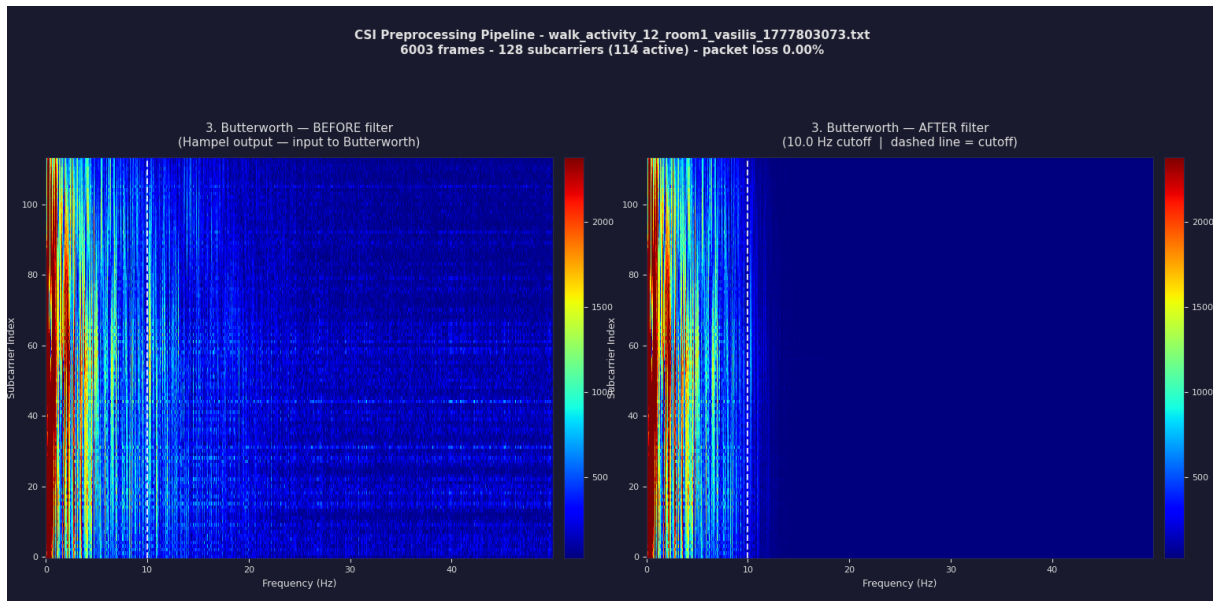
Πιο συγκεκριμένα, η αποκοπή στα 10 Hz επιλέχθηκε ώστε να ανιχνεύονται κινήσεις όπως ο βηματισμός (0,5–3 Hz), η αναπνοή (0,2–0,5 Hz), Doppler shift κτλ. Παράλληλα με την αποκοπή αυτή αφαιρείται και περαιτέρω θόρυβος στις πιο υψηλές συχνότητες καθώς εκεί παρατηρείται ηλεκτρονικός θόρυβος και παρεμβολές RF [1][19].

Το φίλτρο εφαρμόζεται ξεχωριστά σε κάθε υποφορέα (κάθε στήλη στο `amplitude`):

```
amplitude_filtered = np.zeros_like(amplitude)
for i in range(amplitude.shape[1]):
    amplitude_filtered[:, i] = sosfiltfilt(sos, amplitude[:, i])
```



Εικόνα 4.22. Εφαρμογή του Φίλτρου Butterworth στο no_activity_14_room1_vasilis_1777973321.txt (Heatmap Diagram).



Εικόνα 4.23. Εφαρμογή του Φίλτρου Butterworth στο walk_activity_12_room1_vasilis_1777803073.txt (Heatmap Diagram).

Αυτό που φαίνεται αμέσως στα δύο διαγράμματα (Εικόνα 4.22 και 4.23) είναι πόσο διαφορετική φασματική υπογραφή έχει το CSI στις δύο καταστάσεις, όπως αποτυπώνεται στα φασματικά διαγράμματα που συνοδεύουν την εφαρμογή του φίλτρου Butterworth. Στο διάγραμμα walk, η ενέργεια είναι μαζεμένη στις πολύ χαμηλές συχνότητες (έντονο κόκκινο), κάτω από 5 Hz, με τη χρωματική κλίμακα έντασης (magnitude) να φτάνει 2000+ και ελάχιστη παρουσία πιο πάνω. Στο no-activity διάγραμμα η ενέργεια απλώνεται σε όλο το εύρος μέχρι τα 50 Hz, χωρίς καμία ξεκάθαρη χαμηλόσυχη κορύφωση όπως στο walk, και οι τιμές είναι σαφώς μικρότερες, της τάξης των 400. Δηλαδή παρατηρείται θόρυβος σε όλο το φάσμα αντί για ένα σήμα με ξεκάθαρη μορφή. Μετά το low-pass στα 10 Hz, το walk παραμένει σχεδόν αναλλοίωτο αφού η χρήσιμη πληροφορία ήταν εξ αρχής εκεί κάτω, ενώ στο idle κόβεται μεγάλο μέρος του φάσματος και μένουν μόνο μικρές διακυμάνσεις από την πολυδιαδρομική διάδοση (multipath). Η επιλογή του cutoff στα 10 Hz δικαιολογείται εδώ διότι, η μετατόπιση Doppler του βαδίσματος στα 2,4 GHz μπορεί να φτάσει τα 16–24 Hz για ταχύτητες 1–1,5 m/s ($\Delta f = 2v \cdot f_c/c$), ωστόσο η υψηλή τελική ακρίβεια ταξινόμησης υποδεικνύει ότι η πληροφορία που διατηρείται κάτω από τα 10 Hz επαρκεί για τη δυαδική διάκριση (Εικόνες 4.22 και 4.23). Επιπλέον, όπως φαίνεται στα φασματικά διαγράμματα (Εικόνες 4.22–4.23), η ενέργεια του βαδίσματος συγκεντρώνεται κάτω από ~5 Hz, ενώ η χρονική διαφοροποίηση που εφαρμόζεται στη συνέχεια τονίζει ανεξάρτητα τις ταχείες μεταβολές της κίνησης, έτσι η ζώνη κάτω των 10 Hz διατηρεί την απαραίτητη για τη δυαδική διάκριση πληροφορία και η επιλογή του cutoff αποτελεί συμβιβασμό υπέρ της χαμηλής καθυστέρησης. Βέβαια, η πιο ενδιαφέρουσα παρατήρηση δεν είναι τόσο τι αφαιρεί το φίλτρο, όσο ότι η ίδια η μορφή του φάσματος που διαφέρει τόσο πολύ ανάμεσα στις δύο κλάσεις.

4.6.4 Χρονική Διαφοροποίηση (Temporal Differencing)

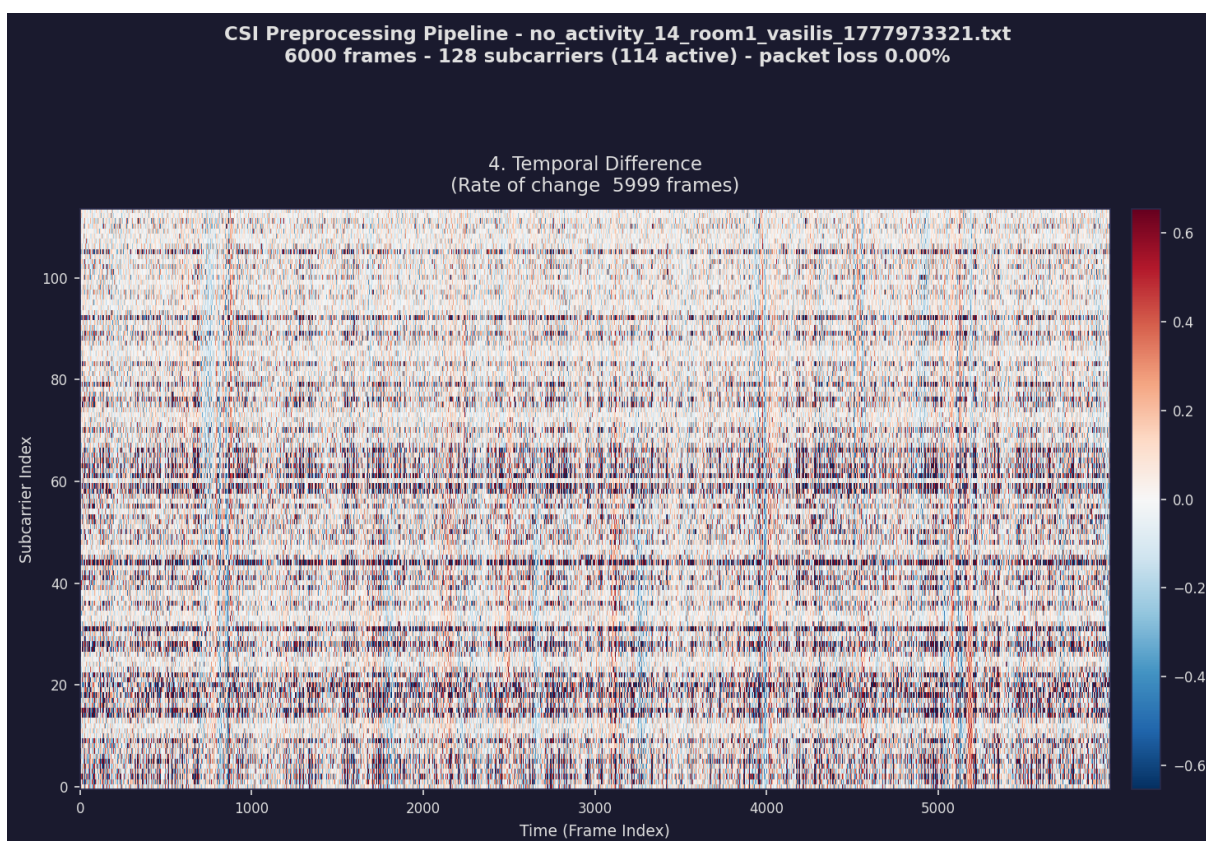
Η χρονική Διαφοροποίηση (temporal differencing) είναι μια τεχνική που υπολογίζει τη διαφορά μεταξύ διαδοχικών frames. Αυτό επιτρέπει την εστίαση κυρίως στην κίνηση και την αφαίρεση του θορύβου που βρίσκεται στο παρασκήνιο του χώρου. Ο μαθηματικός τύπος είναι ο εξής [3]:

$$\Delta X[t] = X[t] - X[t - 1] \quad (4.3)$$

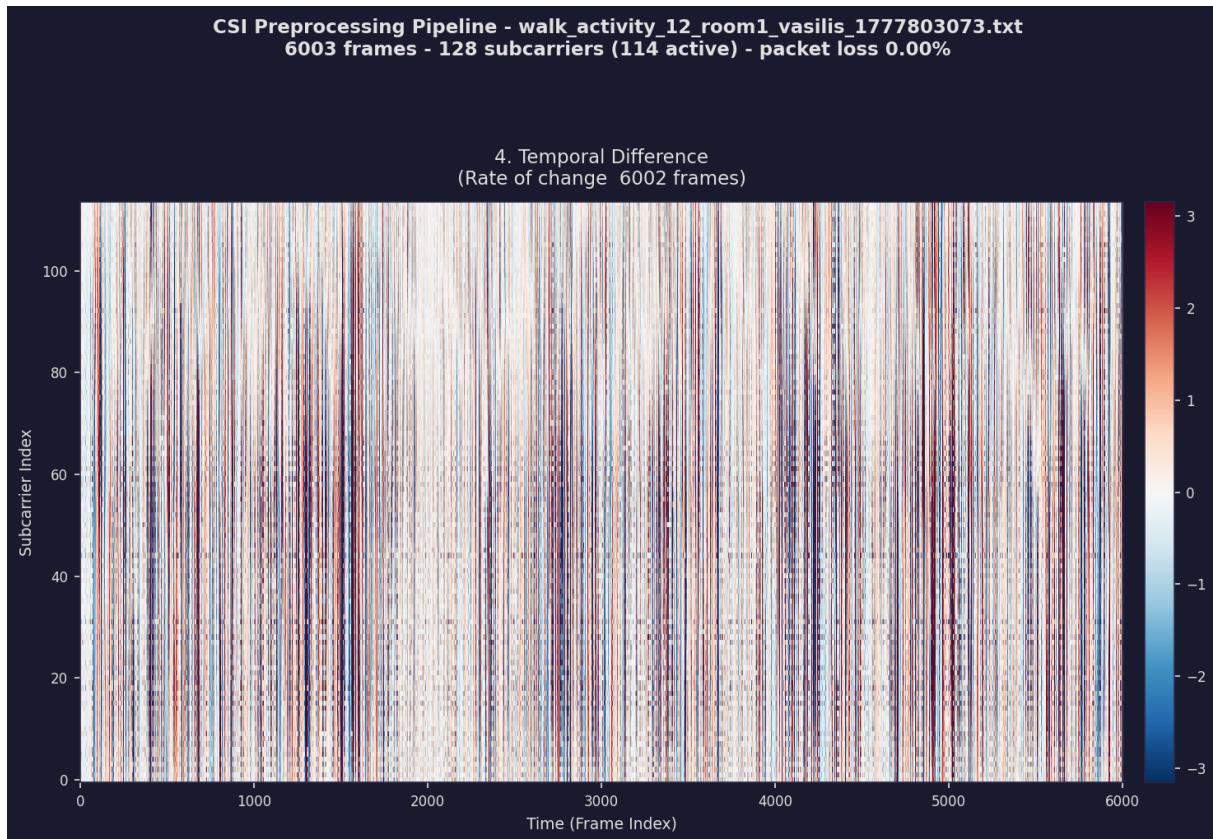
Η υλοποίηση είναι απλή προγραμματιστικά και υλοποιείται ως εξής:

```
amplitude_diff = np.diff(amplitude_no_background, axis=0)
```

Η χρονική διαφοροποίηση λειτουργεί ως ένα high-pass filter στο χρονικό πεδίο, απομακρύνοντας δηλαδή αργές μεταβολές και εστιάζοντας σε γρήγορες αλλαγές [9][19]. Σε ακινησία το $\Delta X[t]$ θα είναι κοντά στο μηδέν, ενώ σε μια σκηνή με κίνηση (Walk) θα εμφανίσει μεγάλες τιμές. Αυτό διευκολύνει αρκετά τη διαφοροποίηση μεταξύ Idle και walk κλάσεων ακόμα και πριν από την εξαγωγή χαρακτηριστικών [9][19].



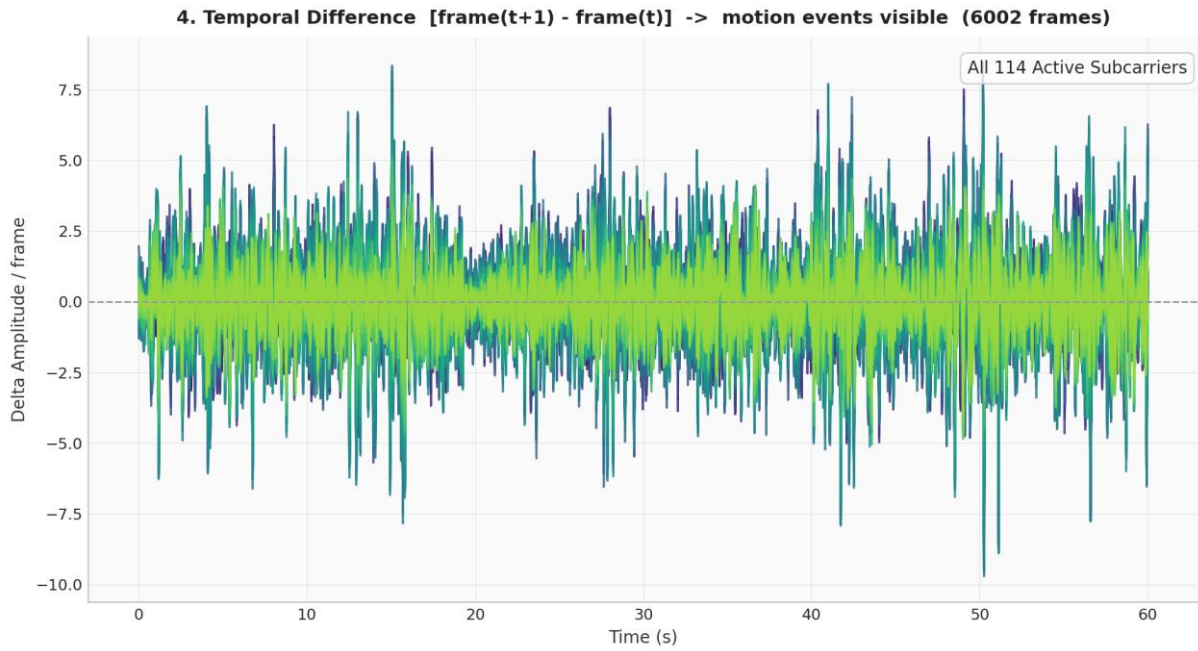
Εικόνα 4.24. Χρήση του Temporal Differencing στο no_activity_14_room1_vasilis_1777973321.txt (Heatmap Diagram).



Εικόνα 4.25. Χρήση του Temporal Differencing στο walk_activity_12_room1_vasilis_1777803073.txt (Heatmap Diagram).

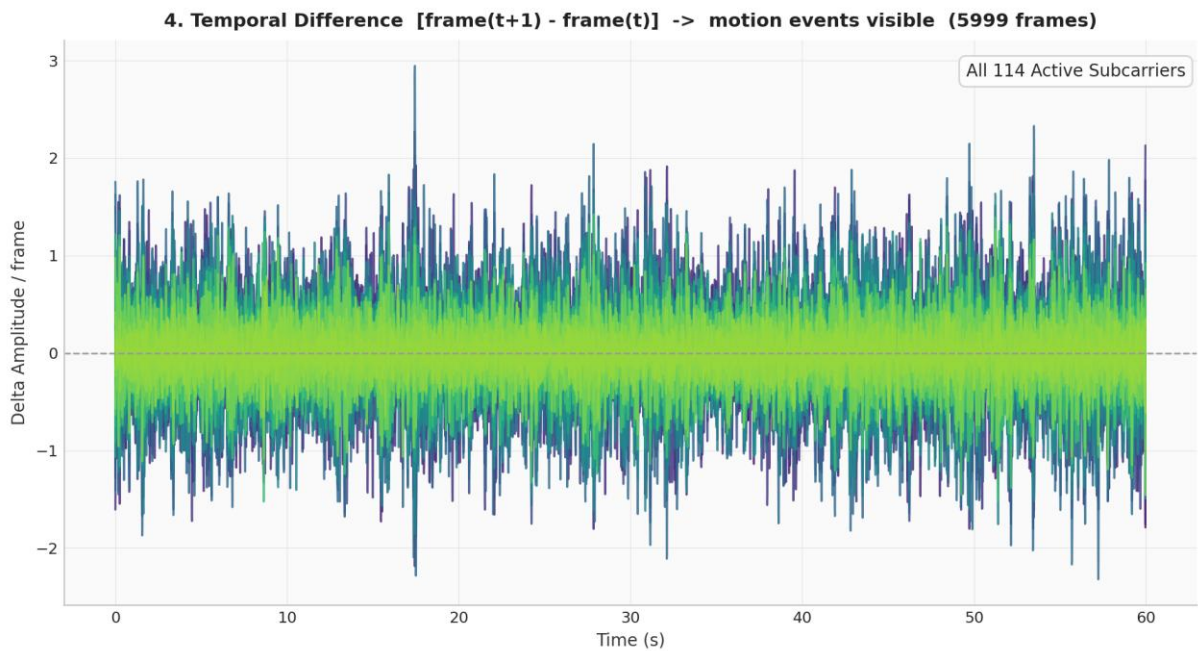
Όπως συμπεραίνεται στο πρώτο διάγραμμα (Εικόνα 4.24), υπάρχει σταθερό επίπεδο ησυχίας καθ' όλη την χρονική διάρκεια των 60 δευτερολέπτων. Η διαταραχή στον χώρο είναι ελάχιστη και παρατηρείται θόρυβος με πλάτος από $-0,6$ έως $0,6$. Αντίθετα στο δεύτερο διάγραμμα (Εικόνα 4.25) παρατηρείται διαταραχή (spikes) σχεδόν σε κάθε χρονική περίοδο καθώς και την ένταση να μεγαλώνει από -3 έως και 3 . Στις εικόνες 4.26 και 4.27 υπάρχει η δυνατότητα να διακριθεί ακόμα καλύτερα με το ανθρώπινο μάτι η κίνηση.

CSI Signal Processing Pipeline - walk_activity_12_room1_vasilis_1777803073.txt
6003 frames - 114 active subcarriers - duration ~60.0 s - packet loss 0.00%



Εικόνα 4.26. Χρήση του Temporal Differencing στο walk_activity_12_room1_vasilis_1777803073.txt.

CSI Signal Processing Pipeline - no_activity_14_room1_vasilis_1777973321.txt
6000 frames - 114 active subcarriers - duration ~60.0 s - packet loss 0.00%



Εικόνα 4.27. Χρήση του Temporal Differencing στο no_activity_14_room1_vasilis_1777973321.txt

Με μια πρώτη ματιά παρατηρώντας τα 2 plot διαγράμματα (Εικόνα 4.26 και 4.27) δεν διακρίνεται μεγάλη διαφορά διότι τα διαγράμματα έχουν μεγεθύνει ανάλογα την κλίμακα τους. Παρατηρώντας όμως την εικόνα 4.26 σχετικά με την κλάση walk ότι το πλάτος (amplitude) φτάνει από το -10 έως το 7,5. Επίσης, μέσα στον χρόνο διακρίνονται έντονα spikes τα οποία προκαλούνται από την κίνηση. Στην εικόνα 4.27 σχετικά με την κλάση idle, η κλίμακα είναι αρκετά μικρότερη και το πλάτος κυμαίνεται από -2,5 έως 3. Τα spikes που παρατηρούνται εδώ μπορεί να οφείλεται κυρίως σε θόρυβο και εξωτερικούς παράγοντες, σε ένα ιδανικό περιβάλλον θα παρατηρούνταν ένα καθαρό και ομοιόμορφο διάγραμμα.

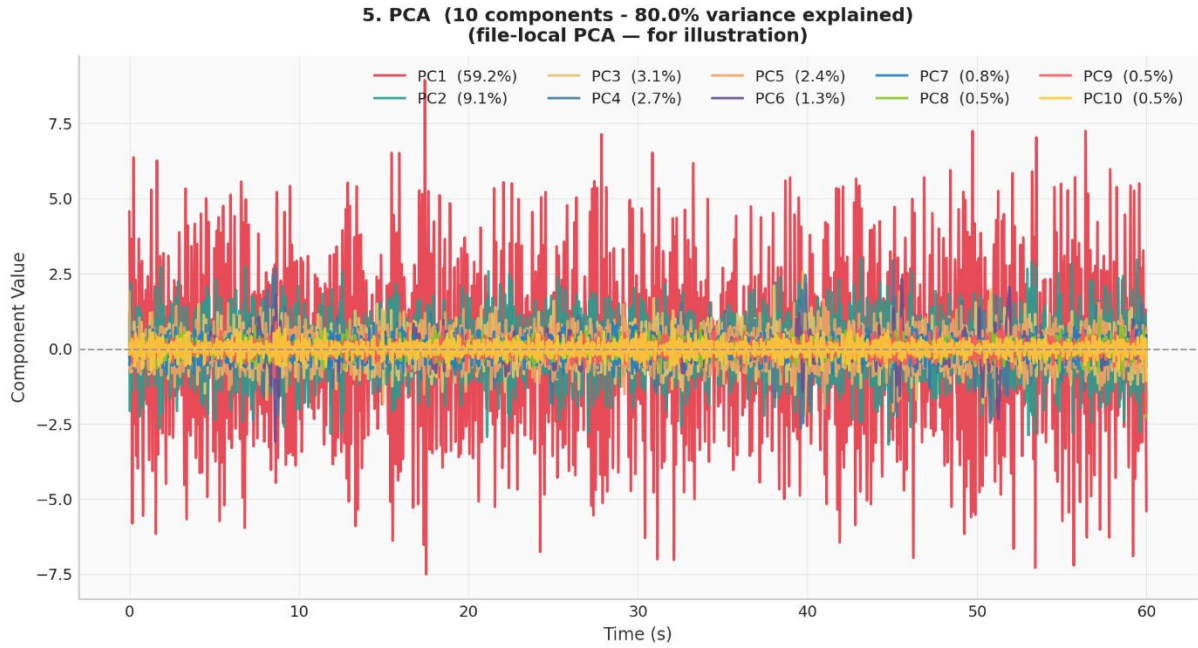
4.6.5 Ανάλυση Κύριων Συνιστωσών (PCA)

Στην παρούσα υποενότητα αναλύεται η εφαρμογή του PCA πάνω στα δεδομένα του πίνακα των 114 ενεργών υποφορέων, διατηρώντας τα $N_PCA_COMPONENTS = 10$ κύρια συστατικά. Το PCA υπολογίζεται αποκλειστικά πάνω στα δεδομένα εκπαίδευσης και εφαρμόζεται (transform) χωρίς επαναπροσαρμογή στα δεδομένα ελέγχου (testing), αποτρέποντας τη διαρροή πληροφορίας (Data Leakage) [7]. Στα δεδομένα, τα 10 κύρια συστατικά συλλαμβάνουν από περίπου 80% (καταγραφές Idle) έως 98% (καταγραφές Walk) της συνολικής διακύμανσης (βλ. Εικόνες 4.28–4.29). Έτσι με αυτόν τον τρόπο μειώνεται η υπάρχουσα πληροφορία από τους 114 υποφορείς σε μόλις 10 κύριες συνιστώσες [8], [19].

Μια σημαντική αλλαγή με τη χρήση του PCA είναι ότι, δεν υπάρχουν πλέον 114 ενεργοί υποφορείς αλλά μόλις 10 συνιστώσες [3][19]. Όπως διαπιστώνεται και στον τίτλο του διαγράμματος (Εικόνα 4.29), διατηρήθηκε το 98,2% της πληροφορίας από τους 114 ενεργούς υποφορείς. Επίσης, στον άξονα y (Component Value) το καινούργιο «πλάτος» έχει αντικατασταθεί. Όσο μεγαλύτερη είναι η τιμή του (θετική ή αρνητική), τόσο πιο έντονη είναι η μεταβολή που έχει καταγράψει η συγκεκριμένη συνιστώσα σε εκείνο το δευτερόλεπτο. Στο παρόν παράδειγμα η πρώτη συνιστώσα PC1 (κόκκινη γραμμή) έχει καταγράψει το 71,4% της πληροφορίας, αντίστοιχα το PC2 το 15% κτλ. Από αυτό συμπεραίνεται ότι το PC1 αποτύπωσε τις έντονες αυξομειώσεις που προκάλεσε το περπάτημα του ανθρώπου. Αντίστοιχα τα PC2 και PC3 αποτύπωσαν και αυτά τις αλλαγές δευτερεύουσας σημασίας (μικρότερες ή πιο σύνθετες ανακλάσεις του σήματος μέσα στο δωμάτιο). Τέλος οι συνιστώσες PC8 έως PC10 είναι πρακτικά στο μηδέν πράγμα που σημαίνει ότι αυτές οι συνιστώσες έχουν συγκεντρώσει κυρίως τον θόρυβο και τη μη χρήσιμη πληροφορία.

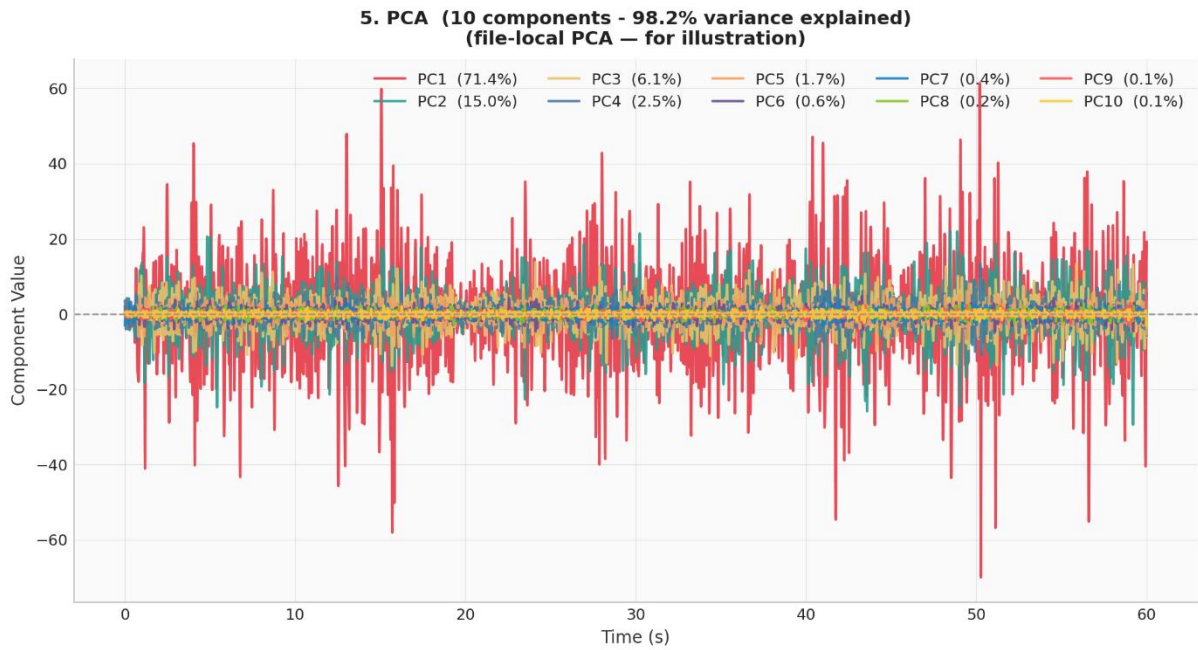
Στην εικόνα 4.28, το PCA κατάφερε να συλλέξει το 80% της πληροφορίας του σήματος. Συγκρίνοντας τα δύο διαγράμματα (Εικόνα 4.29) παρατηρείται ότι το PC1 έχει διακυμάνσεις από -70 έως 60 ενώ στην κλάση idle από -7,5 έως 7,5. Τέλος, στο περπάτημα παρατηρείται 98,2% αθροιστικά στις δέκα συνιστώσες (με 71,4% στην πρώτη), ενώ στην ακινησία το 80% και αυτό οφείλεται στο ότι κατά την κίνηση υπάρχει ένα κυρίαρχο γεγονός, ενώ κατά την ακινησία το περιβάλλον είναι τυχαίο και αποτελείται από θόρυβο.

CSI Signal Processing Pipeline - no_activity_14_room1_vasilis_1777973321.txt
6000 frames - 114 active subcarriers - duration ~60.0 s - packet loss 0.00%



Εικόνα 4.28. Χρήση του PCA στο no_activity_14_room1_vasilis_1777973321.txt.

CSI Signal Processing Pipeline - walk_activity_12_room1_vasilis_1777803073.txt
6003 frames - 114 active subcarriers - duration ~60.0 s - packet loss 0.00%



Εικόνα 4.29. Χρήση του PCA στο walk_activity_12_room1_vasilis_1777803073.txt.

4.6.6 Κανονικοποίηση Δεδομένων

Το τελικό στάδιο της προεπεξεργασίας είναι η κανονικοποίηση των δεδομένων. Διασφαλίζει ότι όλα τα χαρακτηριστικά έχουν παρόμοια κλίμακα, κάτι που είναι σημαντικό για τους αλγορίθμους μηχανικής μάθησης.

Χρησιμοποιείται το StandardScaler από το scikit-learn που κάνει z-score normalization με τον παρακάτω μαθηματικό τύπο [3][5]:

$$z = (x - \mu) / \sigma \quad (4.4)$$

όπου μ είναι ο μέσος όρος και σ η τυπική απόκλιση για κάθε χαρακτηριστικό. Μετά την κανονικοποίηση, κάθε χαρακτηριστικό έχει μέσο όρο 0 και τυπική απόκλιση 1.

Προγραμματιστικά δίνεται ο παρακάτω σχετικά με την κανονικοποίηση στην γλώσσα Python:

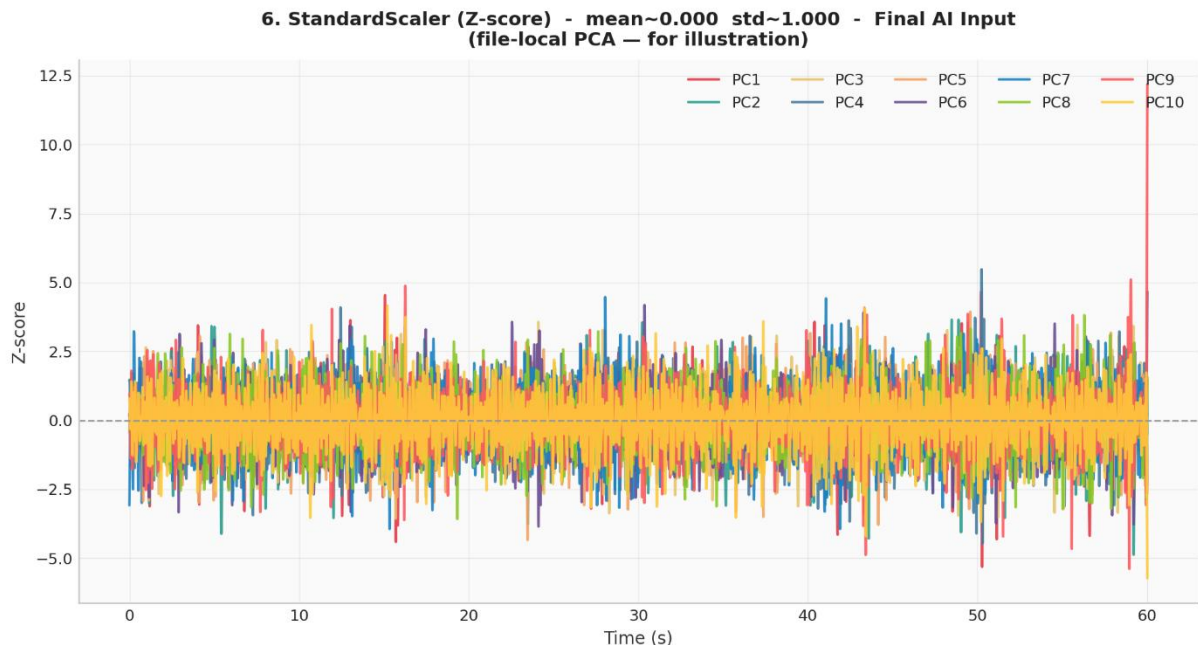
```
from sklearn.preprocessing import StandardScaler

scaler = StandardScaler()

amplitude_scaled = scaler.fit_transform(amplitude_pca)
```

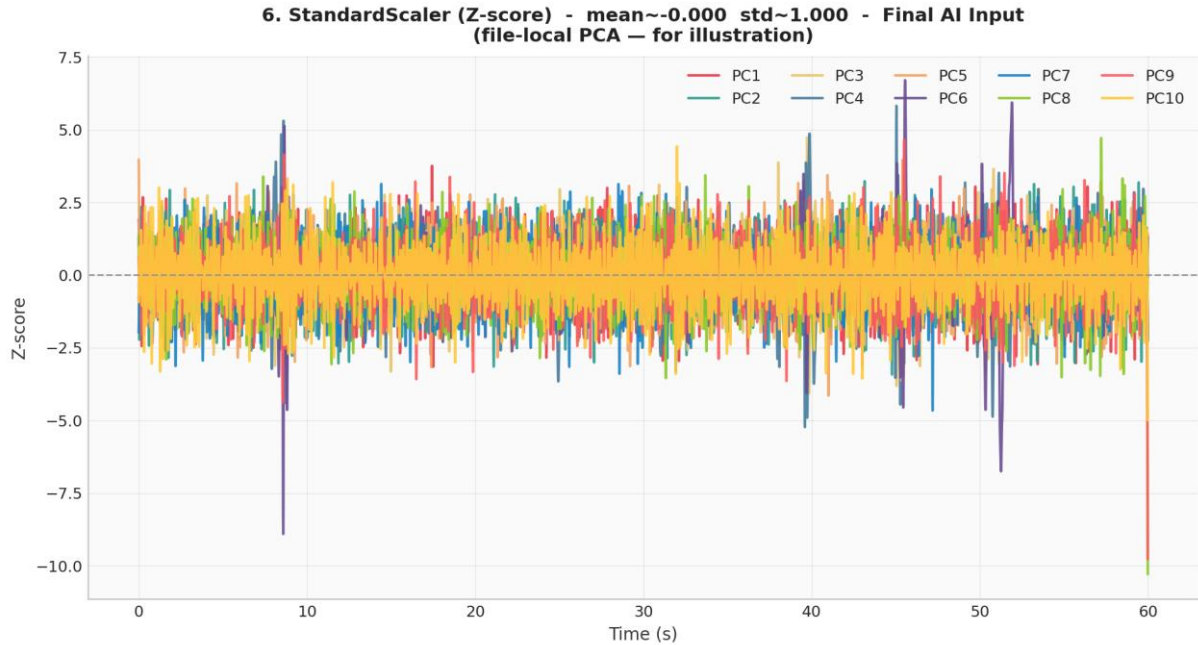
Το παραπάνω απόσπασμα είναι ενδεικτικό. Στην πραγματική ροή εκπαίδευσης το StandardScaler και το PCA προσαρμόζονται (fit) αποκλειστικά στα δεδομένα εκπαίδευσης κάθε fold μέσω ενός sklearn Pipeline και στη συνέχεια εφαρμόζονται (transform) στο σύνολο ελέγχου, ώστε να μην διαρρεύσει καμία πληροφορία από τα δεδομένα ελέγχου προς την εκπαίδευση (data leakage) [5][7].

CSI Signal Processing Pipeline - walk_activity_12_room1_vasilis_1777803073.txt
6003 frames - 114 active subcarriers - duration ~60.0 s - packet loss 0.00%



Εικόνα 4.30. Κανονικοποίηση δεδομένων στο walk_activity_12_room1_vasilis_1777803073.txt

CSI Signal Processing Pipeline - no_activity_14_room1_vasilis_1777973321.txt
6000 frames - 114 active subcarriers - duration ~60.0 s - packet loss 0.00%



Εικόνα 4.31. Κανονικοποίηση δεδομένων στο no_activity_14_room1_vasilis_1777973321.txt

Αυτές οι δύο εικόνες (Εικόνα 4.30 και 4.31) δείχνουν ότι η κανονικοποίηση (Z-score normalization) [5] έφερε τα δεδομένα του περπατήματος (walk_activity) και της ακινησίας (no_activity) στην ίδια ακριβώς κλίμακα, με μέση τιμή (mean) το 0 και τυπική απόκλιση (std) το 1. Ενώ προηγουμένως το περπάτημα είχε πολύ μεγάλες τιμές (π.χ. από -70 έως 60) και το idle πολύ μικρές, τώρα και τα δύο γραφήματα κυμαίνονται στον ίδιο άξονα Y μεταξύ του -5 και +5. Αυτό το στάδιο είναι κρίσιμο για τα μοντέλα μηχανικής μάθησης, διότι μπορούν πολύ εύκολα να μπερδευτούν και να μην εστιάσουν στα μοτίβα κίνησης αλλά στο πόσο μεγάλοι είναι οι αριθμοί.

4.7 Εξαγωγή Χαρακτηριστικών

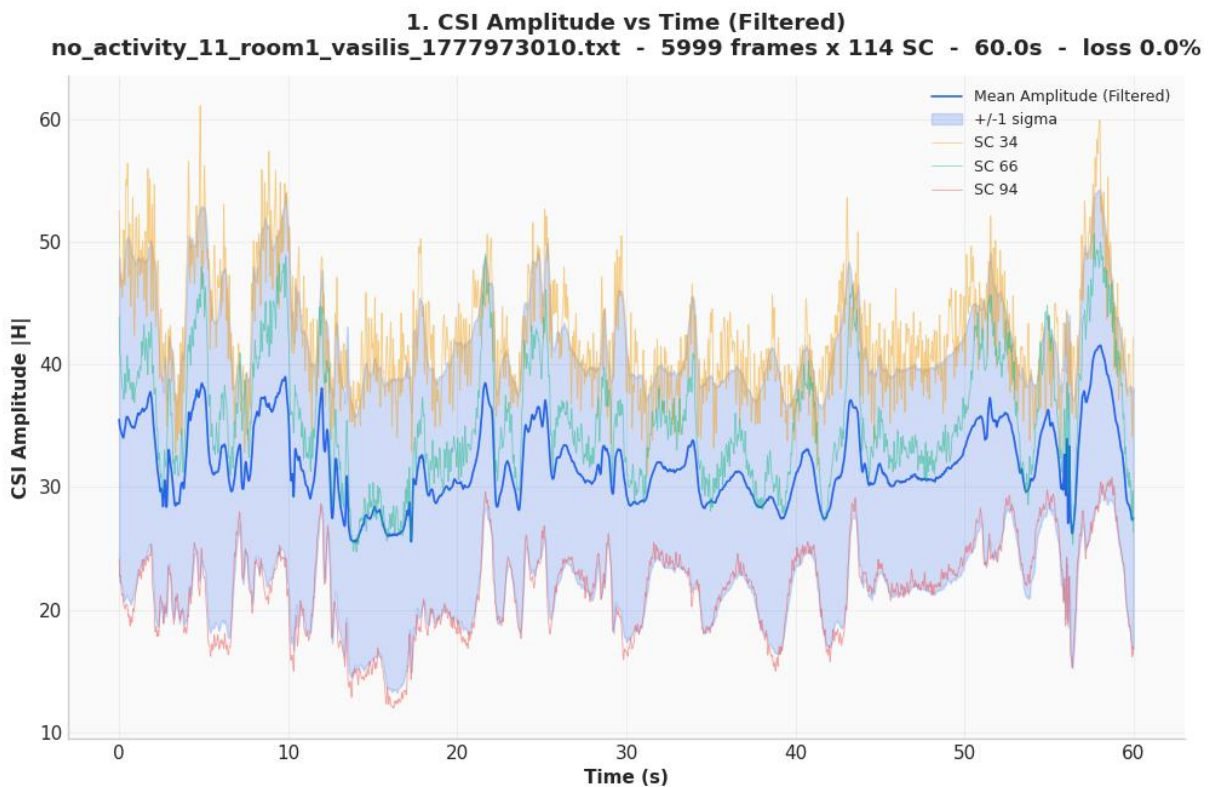
Από κάθε επεξεργασμένο παράθυρο 100×10 εξάγονται 25 χαρακτηριστικά ανά PCA συνιστώσα, δίνοντας συνολικά $25 \times 10 = 250$ χαρακτηριστικά ανά παράθυρο [3], [5]. Παρακάτω συνοψίζεται ο πίνακας με όλα τα χαρακτηριστικά και την λειτουργία τους.

Τα χαρακτηριστικά οργάνωνονται σε τρεις κατηγορίες:

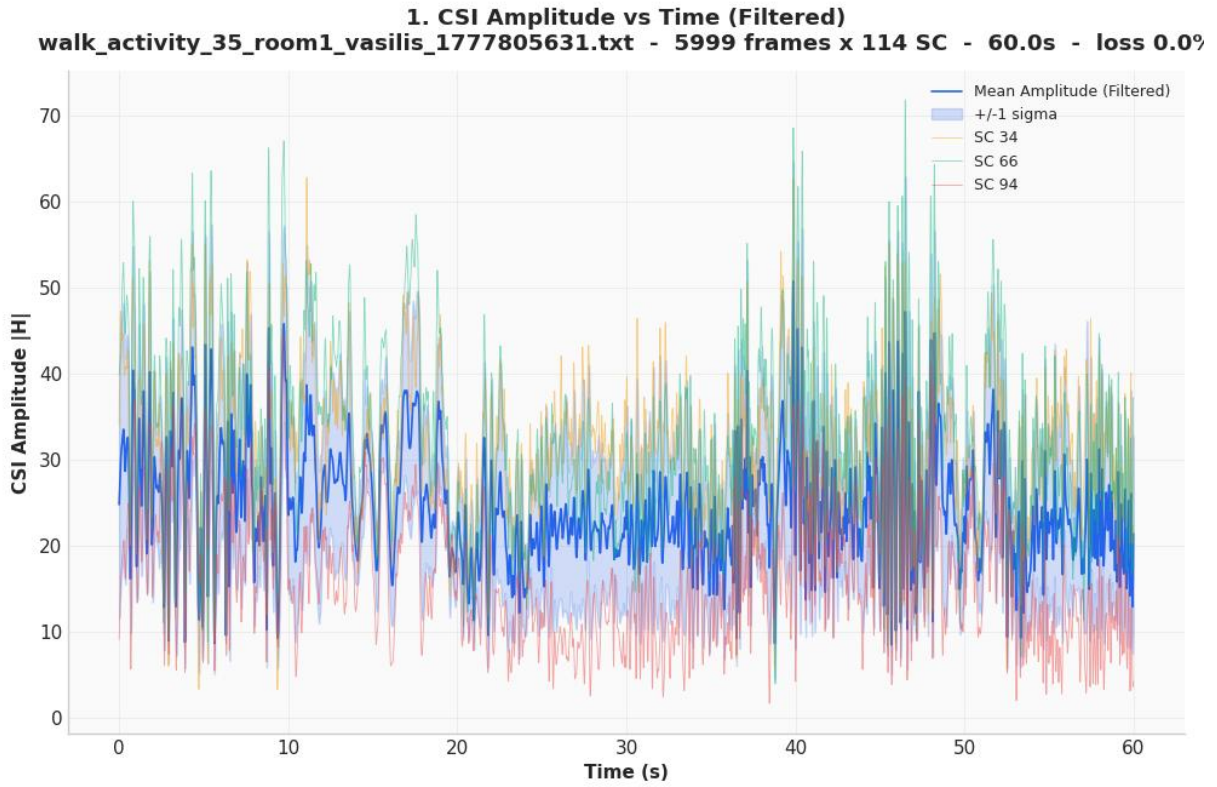
1. Στατιστικά Χαρακτηριστικά Πεδίου Χρόνου
2. Χαρακτηριστικά Πεδίου Συχνотήτων
3. Χαρακτηριστικά Αυτοσυσχέτισης και Μορφολογίας

4.7.1 Στατιστικά Χαρακτηριστικά Πεδίου Χρόνου

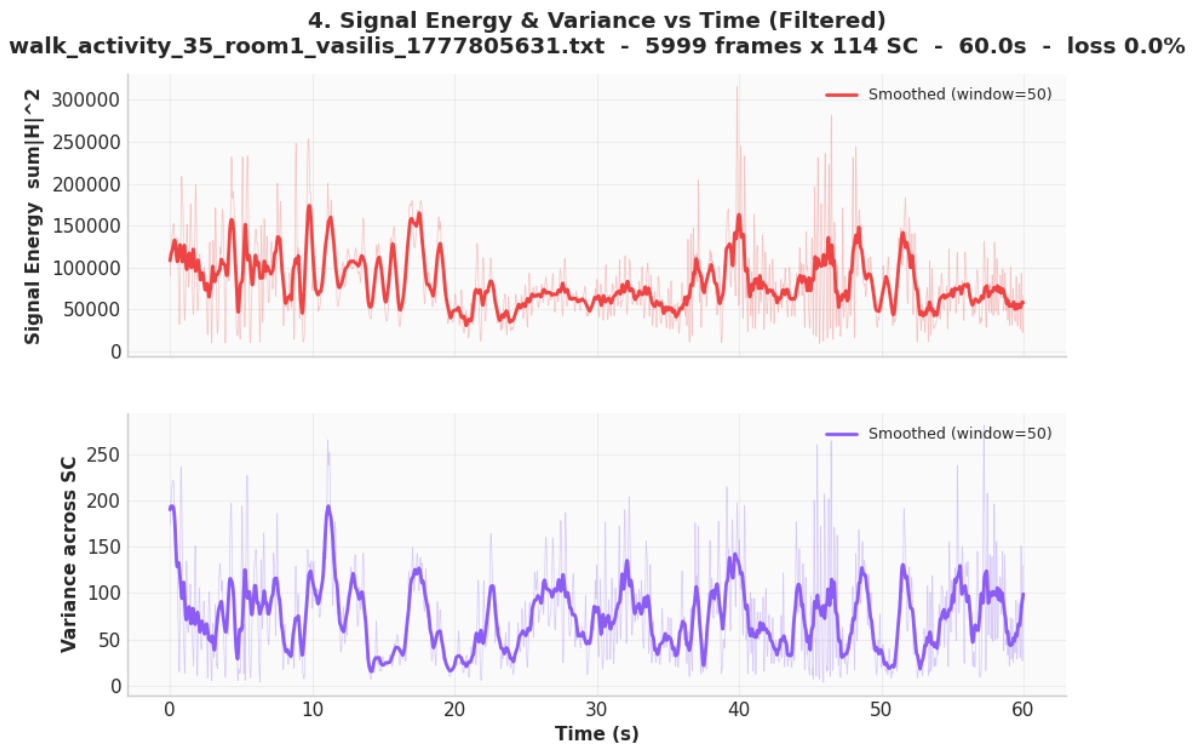
Τα στατιστικά χαρακτηριστικά πεδίου χρόνου [3], [5] είναι η μέση τιμή (mean), τυπική απόκλιση (std), μέγιστο (max), ελάχιστο (min), εύρος τιμών (range), διάμεσος (median), ενέργεια ($\text{energy} = \sum x^2/N$), ασυμμετρία (skewness), υπερβολική κύρτωση (excess_kurtosis), ποσοστό αλλαγής προσήμου (Zero-Crossing Rate, ZCR) και αναλογία παλμών (impulse_ratio = max/mean). Η skewness και η κύρτωση παρέχουν πληροφορία για τη μη-Γκαουσιανή κατανομή του σήματος. Ο λόγος είναι ότι το περπάτημα μπορεί να εμφανίζει διαφορετική κύρτωση από την ηρεμία που μπορεί να εμπεριέχει θόρυβο [5].



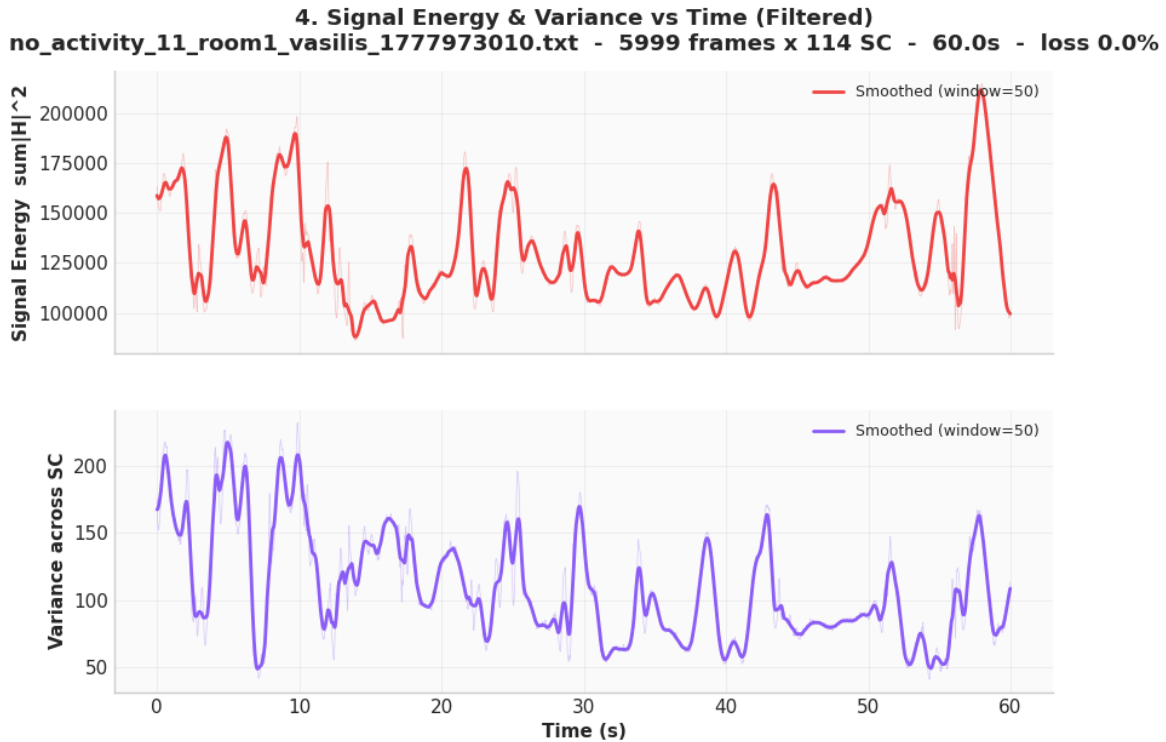
Εικόνα 4.32. Ανάλυση αρχείου no_activity_11_room1_vasilis_1777973010.txt Amplitude vs Time



Εικόνα 4.33. Ανάλυση αρχείου walk_activity_35_room1_vasilis_1777805631.txt Amplitude vs Time



Εικόνα 4.34. Signal Energy και Variance αρχείου walk_activity_35_room1_vasilis_1777805631.txt



Εικόνα 4.35. Signal Energy και Variance αρχείου no_activity_11_room1_vasilis_1777973010.txt

Οι εικόνες αυτές 4.34 και 4.35 παρέχουν την ενέργεια του σήματος (στατιστικό χαρακτηριστικό) αλλά ταυτόχρονα δείχνουν και την διακύμανση στο πεδίο της συχνότητας. Για λόγους απλότητας επιλέχθηκε η τοποθέτηση των 2 εικόνων αυτών στο κεφάλαιο των στατιστικών χαρακτηριστικών. Όπως παρατηρείται, υπάρχουν μεγαλύτερες διακυμάνσεις των τιμών όταν υπάρχει κίνηση. Το πλάτος στην εικόνα 4.32 κινείται ήρεμα με τα subcarriers να ακολουθούν σχετικά σταθερές παράλληλες πορείες, ενώ στην εικόνα 4.33 το εύρος ανοίγει με τα subcarriers να διασταυρώνονται και οι κορυφές γίνονται πιο απότομες και πυκνές. Το ίδιο παρατηρείται και στην εικόνα 4.34 και 4.35 όπου στο walk η μεταβλητότητα είναι μεγαλύτερη, ενώ στην εικόνα 4.35 οι διακυμάνσεις είναι πιο αργές.

4.7.2 Χαρακτηριστικά Πεδίου Συχνότητων

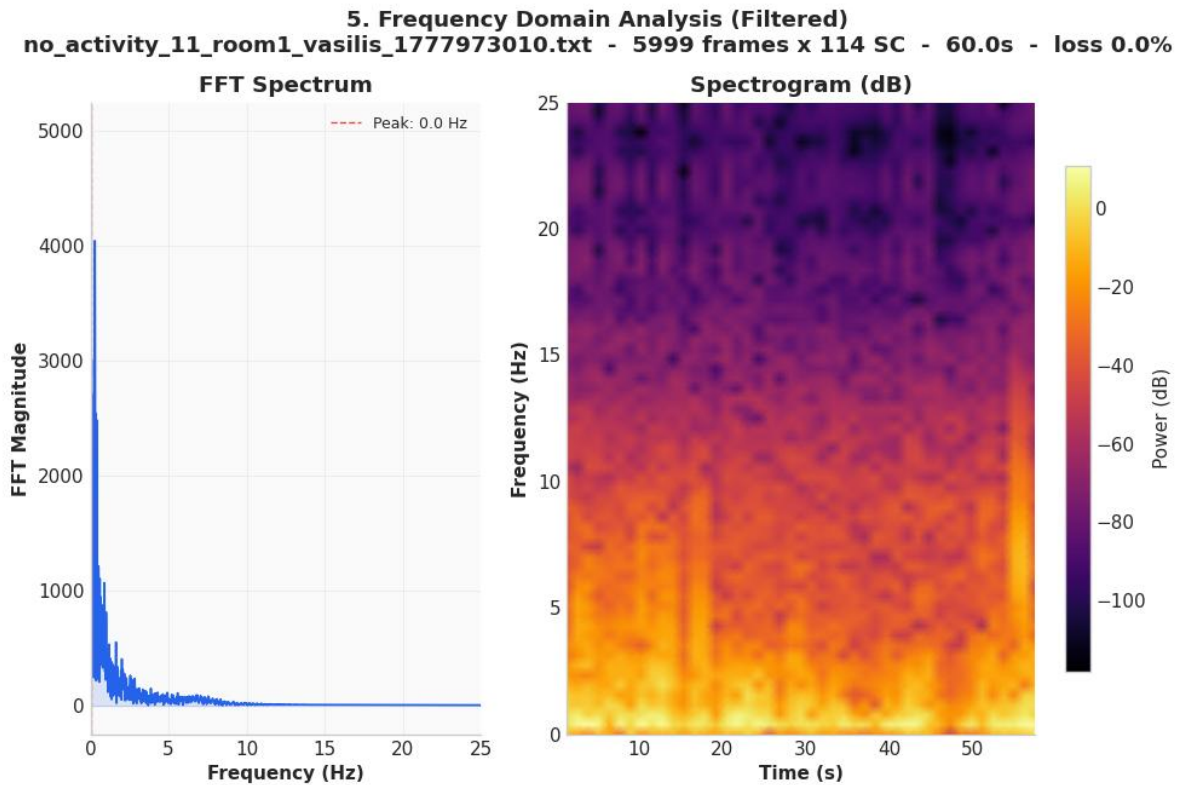
Ο Μετασχηματισμός Fourier (FFT) επιτρέπει την ανάλυση του σήματος στο πεδίο των συχνοτήτων [3], εξάγοντας χαρακτηριστικά όπως:

- Φάσμα Ισχύος (Power Spectral Density - PSD): Δείχνει την κατανομή ισχύος σε διαφορετικές συχνότητες. Διαφορετικές δραστηριότητες παράγουν διαφορετικά φάσματα.
- Κεντρική Συχνότητα (Centroid Frequency): Η σταθμισμένη μέση συχνότητα του φάσματος.
- Εντροπία Φάσματος: Μέτρο της αβεβαιότητας στη φασματική κατανομή[3].

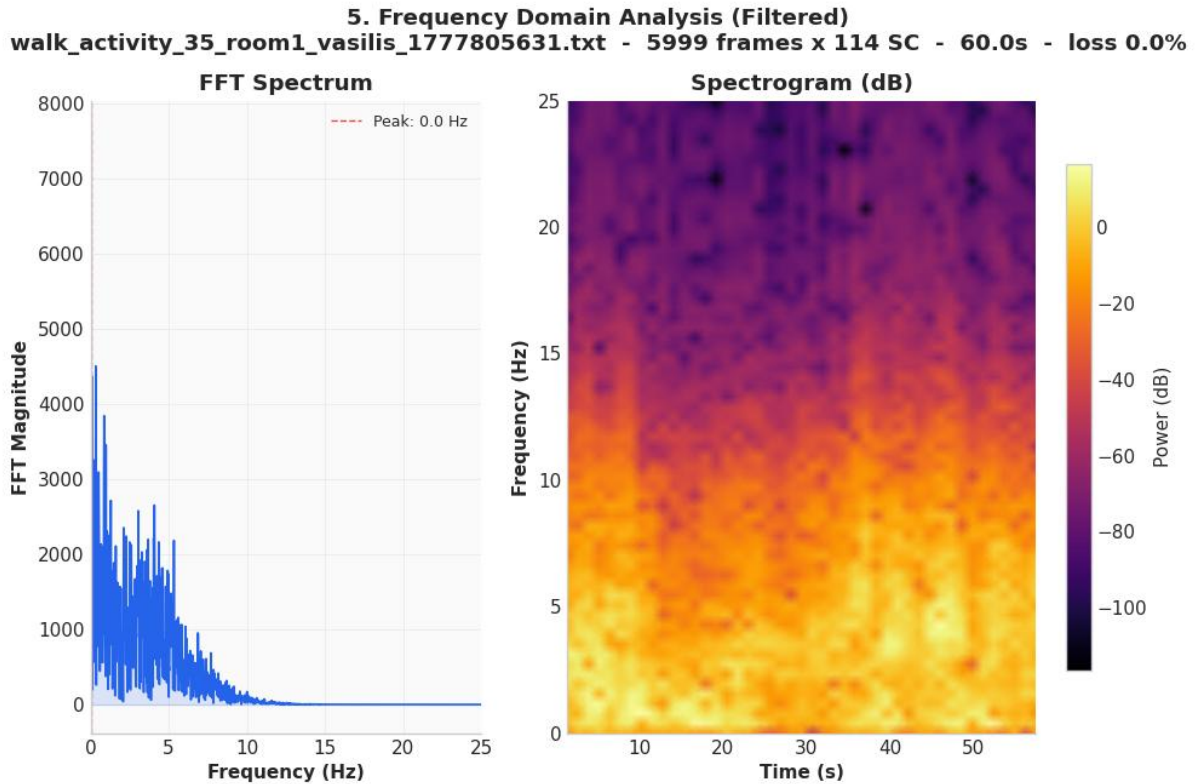
Ο Βραχύχρονος Μετασχηματισμός Fourier (STFT) έχει παρόμοια λειτουργία με το FFT και μπορεί να δημιουργήσει φασματογράμματα (spectrograms) τα οποία δείχνουν πώς το φάσμα εξελίσσεται στο χρόνο. Άρα είναι εφικτό μέσα στον χρόνο να χαρτογραφηθεί η συμπεριφορά μιας κλάσης. Επιπλέον,

στην κατηγορία αυτή υπάρχει η μέση τιμή φάσματος (fft_mean), την τυπική απόκλιση φάσματος (fft_std), τον δείκτη κορυφαίας συχνότητας (fft_peak_idx), την φασματική εντροπία (spectral_entropy), την αναλογία ισχύος ζώνης βηματισμού, το φασματικό κεντροειδές (spectral_centroid) και τέλος την προεξοχή κορυφής (peak_prominence) [1][19].

Τα παρακάτω γραφήματα βρίσκονται στο πεδίο της συχνότητας (Εικόνα 4.36 και Εικόνα 4.37) για την κλάση walk καθώς και για την κλάση idle. Όπως παρατηρείται και στις 2 κλάσεις πάνω από τα 10 Hz το φάσμα είναι πρακτικά μηδενικό, άρα το χαμηλοπερατό φίλτρο λειτουργεί σωστά. Επίσης, στο πεδίο της συχνότητας παρατηρούνται στην κλάση walk διακυμάνσεις από τις συχνότητες 0–10 Hz ενώ στην κλάση idle παρατηρείται ότι είναι συγκριτικά πιο ήρεμη.



Εικόνα 4.36. Ανάλυση του αρχείου no_activity_11_room1_vasilis_1777973010.txt στο πεδίο της συχνότητας



Εικόνα 4.37. Ανάλυση του αρχείου walk_activity_35_room1_vasilis_1777805631.txt στο πεδίο της συχνότητας.

4.7.3 Χαρακτηριστικά Αυτοσυσχέτισης και Μορφολογίας

Η τρίτη κατηγορία συμπεριλαμβάνει χαρακτηριστικά τα οποία έχουν σχέση με τον χρόνο και τον ρυθμό κίνησης. Περιλαμβάνονται η κορυφή αυτοσυσχέτισης (autocorr_peak), κυρίαρχη υστέρηση αυτοσυσχέτισης (autocorr_dominant_lag), κινητικότητα Hjorth, την πολυπλοκότητα Hjorth, το μήκος κυματομορφής, η διάρκεια έξαρσης δραστηριότητας (burst_duration) και τέλος την αναλογία ανόδου-καθόδου (rise_fall_ratio). [5]. Στον παρακάτω πίνακα 4.1 παρουσιάζονται συνοπτικά όλα τα χαρακτηριστικά τα οποία χρησιμοποιούνται.

#	Χαρακτηριστικό (Feature)	Ομάδα	Διαγνωστική αξία
1	Mean	Στατ.	Επίπεδο σήματος
2	Std	Στατ.	Ένταση / διακύμανση κίνησης

#	Χαρακτηριστικό (Feature)	Ομάδα	Διαγνωστική αξία
3	Max	Στατ.	Ανώτατο επίπεδο πλάτους
4	Min	Στατ.	Κατώτατο επίπεδο πλάτους
5	Range (max–min)	Στατ.	Εύρος μεταβολής
6	Median	Στατ.	Κεντρική τάση (ανθεκτική σε ακραίες τιμές/spikes)
7	Signal Energy	Στατ.	Συνολική ενέργεια σήματος στο πεδίο του χρόνου
8	Skewness	Στατ.	Ασυμμετρία (διαφοροποίηση τύπου κίνησης)
9	(Excess) Kurtosis	Στατ.	Παρουσία απότομων events (πολύ υψηλό σε πτώσεις)
10	Zero-crossing rate	Στατ.	Περιοδικότητα (εναλλαγές πρόσημου - ρυθμός βημάτων)
11	Impulse ratio	Στατ.	Σχέση μέγιστης κορυφής προς μέση τιμή (αναγνώριση κρούσεων)
12	FFT Mean (Mean power)	FFT	Μέση ενέργεια στο ενεργό (low-pass) φάσμα
13	FFT Std (Spectral variance)	FFT	Εξάπλωση ενέργειας στο φάσμα
14	Dominant freq. (Peak idx)	FFT	Κύριος ρυθμός κίνησης
15	Spectral entropy	FFT	Κανονικότητα / Πολυπλοκότητα φάσματος

#	Χαρακτηριστικό (Feature)	Ομάδα	Διαγνωστική αξία
16	Spectral centroid	FFT	Κέντρο βάρους φάσματος
17	Gait band energy ratio	FFT	Συγκέντρωση ενέργειας (0,5–3Hz) ταυτοποίηση βαδίσματος
18	Peak prominence	FFT	Οξύτητα κύριας συχνότητας (καθαρότητα επαναλαμβανόμενης κίνησης)
19	Autocorr peak	Temporal	Μέγιστη ισχύς περιοδικότητας της κίνησης
20	Autocorr dominant lag	Temporal	Εκτίμηση διάρκειας κύκλου κίνησης (π.χ. χρόνος ενός βήματος)
21	Signal mobility (Hjorth)	Temporal	Ταχύτητα μεταβολής (velocity) του σήματος
22	Signal complexity (Hjorth)	Temporal	Πολυπλοκότητα / τραχύτητα κυματομορφής (ακανόνιστες κινήσεις)
23	Waveform length	Temporal	Συνολική "διαδρομή" (αθροιστικό μέτρο κινητικής δραστηριότητας)
24	Burst duration	Temporal	Διάρκεια έντονης κινητικής δραστηριότητας (peaks)
25	Rise/Fall ratio	Temporal	Συμμετρία ανόδου προς κάθοδο της κύριας κορυφής (δυναμική πτώσης)

Πίνακας 4.1. Χαρακτηριστικά ανά κύρια συνιστώσα ($25 \times 10 = 250$ διαστάσεις)

Δεν είναι όλα τα subcarriers εξίσου ευαίσθητα στην ανθρώπινη δραστηριότητα. Ορισμένα subcarriers μπορεί να βρίσκονται σε σημεία μηδενισμού (null points) και να μην παρέχουν χρήσιμη πληροφορία. Γενικά η επιλογή των subcarriers (subcarrier selection) επιτρέπει τη μείωση της πληροφορίας των

δεδομένων διατηρώντας τα πιο χρήσιμα χαρακτηριστικά, χωρίς έτσι να χάνεται ουσιαστική πληροφορία.

4.8 Αλγόριθμοι Ταξινόμησης

Χρησιμοποιούνται 6 παραδοσιακοί αλγόριθμοι μηχανικής μάθησης, ώστε να συγκριθούν ως προς τη δομή και τον τρόπο σκέψης τους. Συγκεκριμένα χρησιμοποιούνται ένα γραμμικό μοντέλο (Logistic Regression), ένας kernel-based ταξινομητής (SVM με πυρήνα RBF), δενδρικά μοντέλα (Random Forest, Gradient Boosting, Extra Trees), και ο αλγόριθμος k-NN. Για όλους τους αλγορίθμους η εκπαίδευση πραγματοποιείται με επανξημένα δεδομένα (data augmentation) και η αξιολόγηση ελέγχεται σε κάθε γύρο (fold) [5]. Για λόγους αναπαραγωγιμότητας, οι κύριες υπερπαραμέτροι κάθε αλγορίθμου αναφέρονται στις περιγραφές που ακολουθούν, όπου δεν αναφέρεται ρητά κάποια τιμή, χρησιμοποιούνται οι προεπιλεγμένες τιμές του scikit-learn.

Gradient Boosting (sklearn.GradientBoostingClassifier, `n_estimators=100`, `learning_rate=0.1`, `max_depth=3`): Ο Αλγόριθμος αυτός είναι παρόμοιος με το Random forest με την διαφορά ότι το Gradient Boosting μαθαίνει από τα λάθη του. Πιο συγκεκριμένα δομεί ένα σύνολο δέντρων αποφάσεων και σε κάθε νέο δένδρο μαθαίνει να διορθώνει τα λάθη (residuals) του προηγούμενου, ελαχιστοποιώντας σταδιακά τη συνάρτηση απώλειας (deviance/log-loss). Η ιδιαιτερότητα του είναι ότι αποδίδει αποτελεσματικά σε μη γραμμικά προβλήματα [5].

Random Forest (sklearn.RandomForestClassifier, `n_estimators=200`, `max_depth=15`, `min_samples_leaf=2`, `class_weight='balanced'`): Είναι ένα σύνολο δέντρων αποφάσεων με τυχαία επιλογή των χαρακτηριστικών ανά κόμβο (random feature subsampling) και η τελική απόφαση του προκύπτει από ψηφοφορία πλειοψηφίας των δέντρων. Η ιδιαιτερότητα του αλγορίθμου αυτού είναι ότι είναι ανεκτικός στο overfitting ειδικά σε datasets με μικρά δεδομένα [5][28].

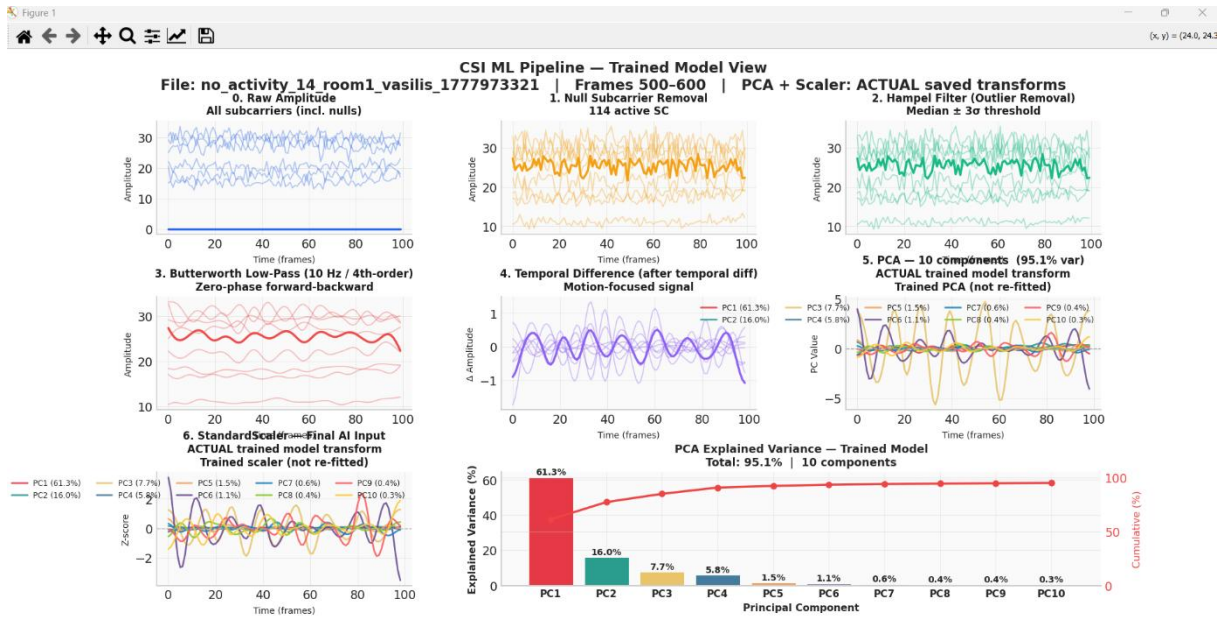
Extra Trees (sklearn.ExtraTreesClassifier, `n_estimators=200`, `max_depth=None`, `class_weight='balanced'`): Παρόμοιο με Random Forest, με τη διαφορά ότι τα σημεία διαχωρισμού επιλέγονται εντελώς τυχαία (random thresholds) αντί για βέλτιστα [5].

Logistic Regression (sklearn.LogisticRegression, `class_weight='balanced'`): Ουσιαστικά, ο αλγόριθμος προσπαθεί να διαχωρίσει τα δεδομένα γραμμικά μεταξύ διαφορετικών καταστάσεων, όπως για παράδειγμα μεταξύ της κίνησης και της ηρεμίας [5].

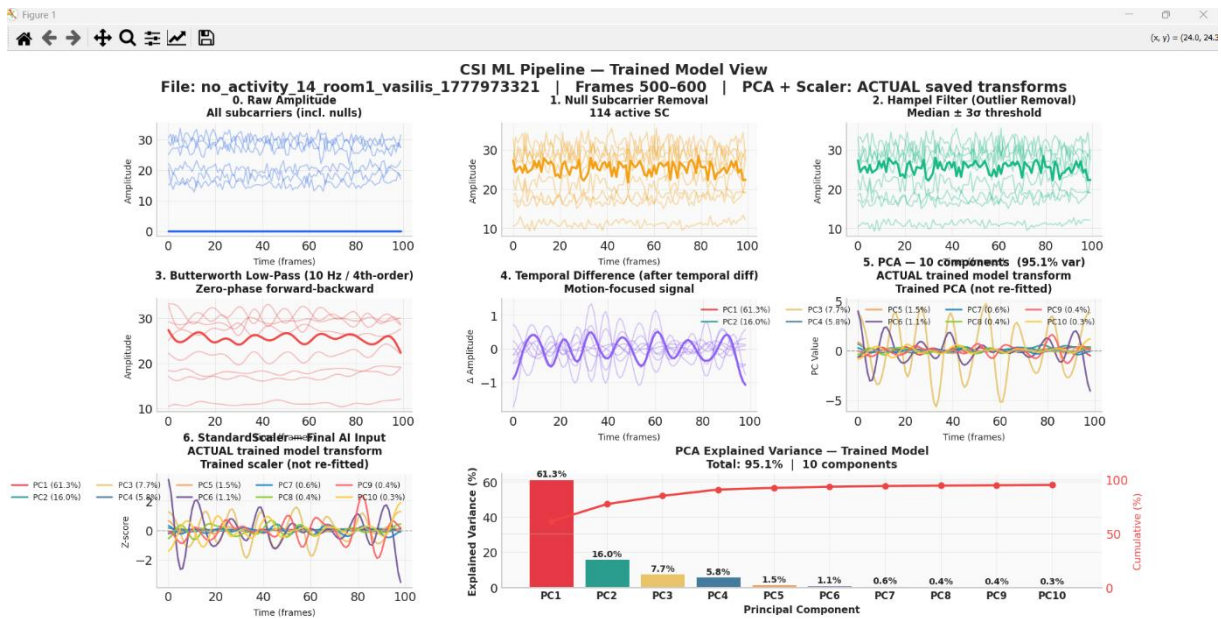
k-Nearest Neighbors (sklearn.KNeighborsClassifier, `n_neighbors=5`, `weights='distance'`, ευκλείδεια μετρική): Σε αντίθεση με άλλα μοντέλα, ο αλγόριθμος αυτός δεν εξάγει κάποιον μαθηματικό κανόνα κατά την εκπαίδευση, αλλά λειτουργεί με βάση τη λογική της ομοιότητας. Για παράδειγμα όταν το σύστημα κληθεί να ταξινομήσει ένα άγνωστο δείγμα υπολογίζει την απόστασή του με βάση όλα τα δείγματα του συνόλου όπου εκπαιδεύτηκε [5].

Support Vector Machine (sklearn.SVC, `kernel='rbf'`, `C=10`, `gamma='scale'`, `class_weight='balanced'`): Ο αλγόριθμος αυτός αναζητά τη γραμμική διαφορά που χωρίζει τις δύο κλάσεις με το μεγαλύτερο δυνατό περιθώριο (μέγιστο δυνατό διαχωρισμό) ανάλογα τα χαρακτηριστικά. Όταν οι κλάσεις δεν είναι γραμμικά διαχωρίσιμες, εφαρμόζεται kernel function (στην παρούσα υλοποίηση RBF) που προβάλλει τα δεδομένα σε υψηλότερη διάσταση όπου ο διαχωρισμός γίνεται εφικτός [5]. Σε σχέση με τα δενδρικά μοντέλα, το SVM δεν προσφέρει εύκολα τη δυνατότητα κατανόησης του τρόπου σκέψης του ως προς τα χαρακτηριστικά οπότε χάνεται η δυνατότητα ερμηνείας του τι έλαβε υπόψη του το μοντέλο για να αποφασίσει.

Κεφάλαιο 4



Εικόνα 4.38. Πρακτική εφαρμογή όλων των βημάτων για την εισαγωγή τους στα μοντέλα μηχανικής μάθησης. Αρχείο no_activity_14_room1_vasilis_1777973321.txt



Εικόνα 4.39. Πρακτική εφαρμογή όλων των βημάτων για την εισαγωγή τους στα μοντέλα μηχανικής μάθησης. Αρχείο walk_activity_12_room1_vasilis_1777803073.txt

Τα διαγράμματα στην Εικόνα 4.38 και στην Εικόνα 4.39 δείχνουν ουσιαστικά πώς λειτουργεί το σύστημα στην πράξη. Η μεγάλη διαφορά με τα προηγούμενα διαγράμματα είναι ότι πλέον δεν παρουσιάζεται ολόκληρη η διάρκεια του αρχείου. Αντίθετα, απομονώνεται ένα μικρό χρονικό "παράθυρο" 100 καρέ (π.χ. frames 500-600), ακριβώς όπως κάνει το σύστημα για να αποφασίσει ακαριαία. Επιπλέον, στα στάδια του PCA και του StandardScaler, το σύστημα δεν υπολογίζει ξανά τις πράξεις από την αρχή αλλά εφαρμόζει κατευθείαν τις έτοιμες ρυθμίσεις που "έμαθε" κατά την εκπαίδευσή του. Συνολικά, φαίνεται πώς τα αρχικά, γεμάτα θόρυβο δεδομένα καθαρίζονται και "πακετάρονται" σε μια πολύ πιο απλή μορφή. Έτσι, το μοντέλο παίρνει κατευθείαν την ουσία της πληροφορίας για να κάνει την ταξινόμηση (classification), χωρίς να σπαταλάει υπολογιστική δύναμη.

4.9 Γραφικό Περιβάλλον Σε Πραγματικό Χρόνο

Το γραφικό περιβάλλον της εφαρμογής (Εικόνα 4.40) λειτουργεί ως ένα dashboard πραγματικού χρόνου, με σκοπό να δίνει στον χρήστη μια συνολική εικόνα των μετρήσεων και της ανάλυσης του CSI. Στο κεντρικό τμήμα της οθόνης βρίσκεται το διάγραμμα ζωντανού σήματος (Live Signal), όπου αποτυπώνεται η χρονική εξέλιξη του amplitude ανά frame. Επίσης, από κάτω εμφανίζεται η ισχύς ανά υποφορέα (Subcarrier Power), δίνοντας εικόνα για την κατανομή της ενέργειας στο φάσμα.

Στη δεξιά πλευρά της διεπαφής υπάρχει το πλαίσιο πρόβλεψης δραστηριότητας. Τη συγκεκριμένη στιγμή φαίνεται η ταμπέλα «IDLE/NO ACTIVITY», με τη βεβαιότητα του μοντέλου στο 76,9% και τον δείκτη κίνησης (Motion Intensity) στο 2%, δηλαδή μια ποσοτική ένδειξη της κινητικότητας στον χώρο. Δίπλα εμφανίζονται και τα στατιστικά της σύνδεσης, η ποιότητα γραμμής (Link Quality: Excellent, Δείκτης ποιότητας σύνδεσης: 99,3) και ο ρυθμός απώλειας πακέτων στο 1%. Παράλληλα, ένας πίνακας πρόσφατης δραστηριότητας κρατά τις τελευταίες προβλέψεις με χρονική σήμανση και τα αντίστοιχα ποσοστά βεβαιότητας.

Στο κατώτατο τμήμα της οθόνης συνοψίζονται οι βασικές παράμετροι λειτουργίας, δηλαδή ο συνολικός αριθμός frames (4.447), χρόνος λειτουργίας (με βάση το άνοιγμα της εφαρμογής) και το μοντέλο που χρησιμοποιείται (Random Forest). Το Random Forest επιλέχθηκε ως προεπιλεγμένο μοντέλο του dashboard επειδή προσφέρει σχεδόν ισοδύναμη ακρίβεια με τον Gradient Boosting (βλ. Κεφάλαιο 5), με ευκολότερη ερμηνεία και ανθεκτικότητα. Η συγκεκριμένη οργάνωση της διεπαφής επιλέχθηκε ώστε ο χρήστης να βλέπει σε μία οθόνη τόσο τη ροή των μετρήσεων όσο και την κατάσταση του συστήματος, χωρίς να χρειάζεται να μεταβαίνει σε διαφορετικά παράθυρα.



Εικόνα 4.40. Αρχικό dashboard της εφαρμογής.

Η καρτέλα Signal View (Εικόνα 4.41) εστιάζει σε δύο διαγράμματα που δίνουν μια πιο τεχνική εικόνα του σήματος CSI. Πιο συγκεκριμένα στο πάνω μέρος της οθόνης φαίνεται η κατανομή ισχύος ανά υποφορέα (Subcarrier Power Distribution). Επιπλέον, διακρίνονται δύο χαρακτηριστικές κορυφές, μία στους υποφορείς 30 έως 45 και μία στους 85 έως 100, κάτι αναμενόμενο για ένα κανάλι Wi-Fi 802.11. Επιπλέον, στο κάτω μέρος εμφανίζεται η κυματομορφή του ζωντανού σήματος (Live Signal Waveform, Amplitude) και παρατηρείται ότι είναι σταθερό και ομαλό κατά μήκος των frames, που σημαίνει ότι δεν υπήρχε σημαντική κίνηση στον χώρο τη στιγμή της καταγραφής.

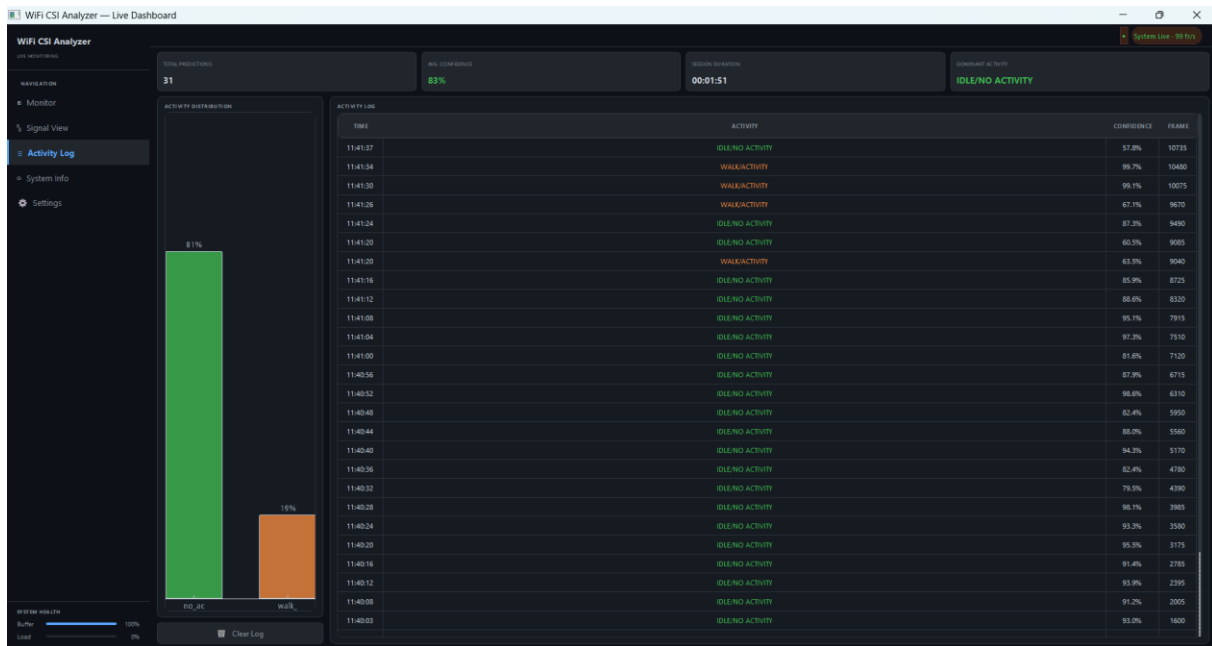
Τέλος, στο πάνω μέρος της διεπαφής εμφανίζονται και οι βασικές μετρικές πραγματικού χρόνου, δηλαδή ο ρυθμός επεξεργασίας frames (102,1 fps), το latency (16,0 ms), η απώλεια πακέτων (0%), η μέση ισχύς σήματος (-16,4 dB) και η διακύμανση (variance: 0,0001). Σημειώνεται ότι οι δείκτες αυτοί (Link Quality, δείκτης ποιότητας, ισχύς) υπολογίζονται εσωτερικά από την εφαρμογή ως σχετικές μετρικές ποιότητας πάνω στα δεδομένα CSI και δεν αποτελούν βαθμονομημένες μετρήσεις RF σε dB/dBm. Συνολικά όλα αυτά τα χαρακτηριστικά δίνουν στον χρήστη μια καθαρή ποσοτική εικόνα της ποιότητας και της σταθερότητας του καναλιού κατά τη λήψη των δεδομένων.



Εικόνα 4.41. Καρτέλα Signal View της εφαρμογής.

Από την άλλη, η καρτέλα Activity Log (Εικόνα 4.42) λειτουργεί ως ιστορικό αρχείο της δραστηριότητας, καθώς κρατά χρονολογικά όλες τις προβλέψεις που έγιναν κατά τη διάρκεια μιας συνεδρίας. Στο πάνω μέρος της οθόνης συνοψίζονται τέσσερα βασικά στατιστικά, δηλαδή ο συνολικός αριθμός προβλέψεων, η μέση εμπιστοσύνη του μοντέλου (77%), η διάμεση διάρκεια συνεδρίας (00:03:23) και η κυρίαρχη κατηγορία δραστηριότητας (IDLE/NO ACTIVITY).

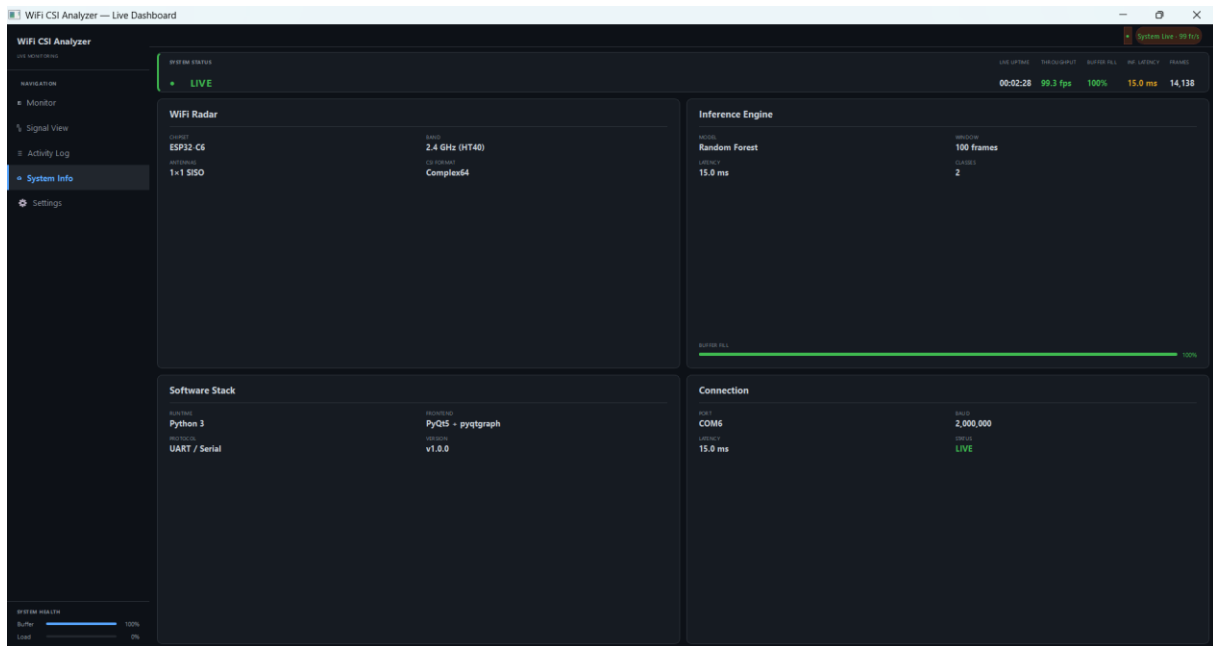
Στα αριστερά εμφανίζεται το γράφημα κατανομής δραστηριότητας (Activity Distribution), όπου στη συγκεκριμένη περίπτωση φαίνεται ότι σχεδόν όλες οι προβλέψεις (100%) αντιστοιχούν σε κατάσταση ηρεμίας, χωρίς δηλαδή ανίχνευση βάρδισης. Παράλληλα, ο κεντρικός πίνακας καταγραφής παρουσιάζει αναλυτικά, για κάθε χρονική στιγμή, την κατηγορία πρόβλεψης, το ποσοστό εμπιστοσύνης και τον αριθμό των επεξεργασμένων frames. Με αυτόν τον τρόπο γίνεται εφικτή η αναδρομική ανάλυση της συμπεριφοράς του συστήματος, οπότε μπορούν να αναλυθούν τα αποτελέσματα.



Εικόνα 4.42. Καρτέλα Activity Log.

Η καρτέλα System Info (Εικόνα 4.43) δίνει μια συγκεντρωτική εικόνα της διαμόρφωσης του συστήματος, οργανωμένη σε τέσσερις ενότητες. Η πρώτη, με τίτλο WiFi Radar, περιγράφει το υλικό λήψης, δηλαδή τον μικροελεγκτή ESP32-C6 με κεραία 1×1 SISO, τη συχνότητα 2,4 GHz (HT40) και τη μορφή του CSI.

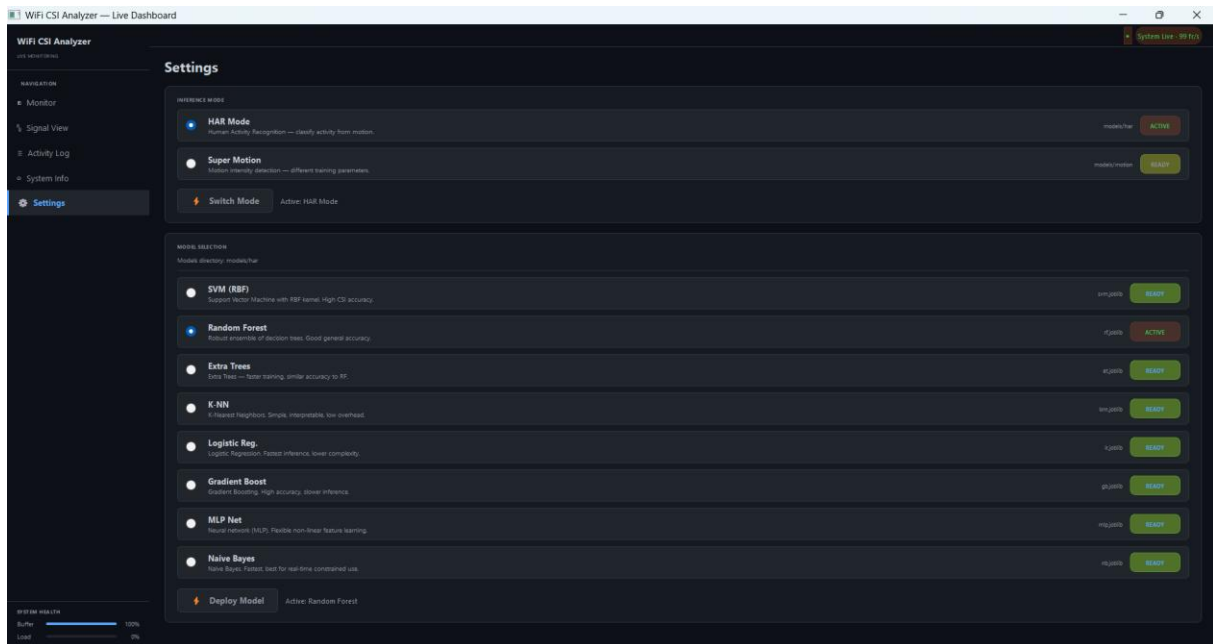
Αμέσως μετά, στην ενότητα Inference Engine αναφέρονται οι παράμετροι του ταξινομητή, ο οποίος βασίζεται στον τρέχοντα αλγόριθμο Random Forest με παράθυρο 100 frames, δύο κλάσεις εξόδου και καθυστέρηση 16,0 ms. Στη συνέχεια, το Software Stack επιβεβαιώνει το περιβάλλον λογισμικού, που στηρίζεται σε Python 3 με διάυλο επικοινωνίας UART/Serial και frontend υλοποιημένο σε PyQt5/pyqtgraph. Τέλος, η ενότητα Connection επιβεβαιώνει την ενεργή σύνδεση μέσω της θύρας USB Serial/JTAG (CDC) στο COM6, με ονομαστικό ρυθμό 2.000.000 baud και καθυστέρηση 16,0 ms.



Εικόνα 4.43. Καρτέλα πληροφοριών του Τρέχοντος Συστήματος.

Η καρτέλα Settings (Εικόνα 4.44) δίνει στον χρήστη τη δυνατότητα να διαμορφώσει τις παραμέτρους λειτουργίας του συστήματος σε δύο επίπεδα. Στην ενότητα Inference Mode υπάρχουν δύο διαθέσιμες λειτουργίες, ο HAR Mode (Human Activity Recognition), που είναι ενεργός και αναλαμβάνει την κατηγοριοποίηση της ανθρώπινης δραστηριότητας, και ο Super Motion Mode, ο οποίος εστιάζει στην ανίχνευση έντασης κίνησης με διαφορετικές παραμέτρους, δηλαδή μικρότερα παράθυρα στα 0,4s ώστε η κίνηση να ανιχνεύεται σχεδόν ακαριαία.

Στη συνέχεια, η ενότητα Model Selection παρουσιάζει οκτώ διαθέσιμα μοντέλα μηχανικής μάθησης, συγκεκριμένα τα SVM (RBF), Random Forest, Extra Trees, k-NN, Logistic Regression, Gradient Boost, MLP Net και Naive Bayes. Ο χρήστης μπορεί να επιλέξει οποιοδήποτε από αυτά και να το αναπτύξει στο σύστημα μέσω του κουμπιού Deploy Model. Στη συγκεκριμένη περίπτωση το ενεργό μοντέλο είναι ο Random Forest, κάτι που επιβεβαιώνει άλλωστε και την επιλογή που φάνηκε στις προηγούμενες καρτέλες. Οι αλγόριθμοι MLP και Naive Bayes υπάρχουν και λειτουργούν αλλά δεν αναφέρονται στην παρούσα εργασία.



Εικόνα 4.44. Ρυθμίσεις Μοντέλων Συστήματος.

Ένα από τα ζητήματα που προκύπτουν σε real-time inference είναι ότι οι στιγμιαίες προβλέψεις από μόνες τους εμφανίζουν αρκετή αστάθεια. Για τον λόγο αυτό, στο dashboard ενσωματώθηκαν δύο μηχανισμοί εξομάλυνσης που λειτουργούν συμπληρωματικά.

Ο πρώτος είναι ένας προσαρμοστικός εκθετικός κινητός μέσος (Adaptive Exponential Moving Average, EMA) που εφαρμόζεται πάνω στο διάνυσμα πιθανοτήτων εξόδου του μοντέλου. Αντί για σταθερό συντελεστή, ο παράγοντας εξομάλυνσης α ζυγίζεται δυναμικά με βάση την τρέχουσα μέγιστη πιθανότητα της πρόβλεψης. Πρακτικά, οι προβλέψεις υψηλής βεβαιότητας περνούν γρηγορότερα στον ιστορικό μέσο, ενώ οι πιο αβέβαιες αποκτούν μικρότερο βάρος. Ο δεύτερος μηχανισμός είναι η υστέρηση ετικέτας (label hysteresis), δηλαδή η κατηγορία που εμφανίζεται στη διεπαφή δεν αλλάζει με την πρώτη νέα πρόβλεψη, παρά μόνο αφού η ίδια ετικέτα επιβεβαιωθεί για έναν ελάχιστο αριθμό διαδοχικών κύκλων (hyst_count).

Παράλληλα, το σύστημα υπολογίζει σε πραγματικό χρόνο και έναν δείκτη έντασης κίνησης (Motion Intensity). Αυτός προκύπτει από την τυπική απόκλιση του πλάτους σήματος ως προς τον κυλιόμενο μέσο του ($\sqrt{\text{variance} / \text{mean_amp} \times 4}$) και εξομαλύνεται εκθετικά με συντελεστή 0,03. Με αυτόν τον τρόπο η ανίχνευση κίνησης γίνεται ανεξάρτητα από τον ταξινομητή, οπότε ουσιαστικά υπάρχει και ένα δεύτερο σημείο επιβεβαίωσης. Οι παραπάνω επιλογές έγιναν με στόχο να περιοριστούν οι ψευδείς εναλλαγές κατηγορίας και να σταθεροποιηθεί η συμπεριφορά της διεπαφής σε συνθήκες θορύβου καναλιού. Βέβαια, στη λειτουργία super motion τα παραπάνω δεν εφαρμόζονται, παρότι χρησιμοποιούνται τα ίδια μοντέλα, καθώς εκεί η διαφορά βρίσκεται στις παραμέτρους εκπαίδευσης και στην αμεσότητα που απαιτείται από τη συγκεκριμένη λειτουργία.

Πέρα από τους μηχανισμούς εξομάλυνσης, η αρχιτεκτονική του συστήματος στηρίζεται σε ορισμένες ακόμη τεχνικές επιλογές που εξυπηρετούν την αποδοτική επεξεργασία σήματος σε πραγματικό χρόνο. Η λήψη προβλέψεων γίνεται μέσω κυλιόμενου παραθύρου (sliding window) σταθερού πλήθους frames,

το οποίο μετακινείται ανά τακτό βήμα (step), οπότε προκύπτει επικαλυπτόμενη δειγματοληψία χωρίς να επιβαρύνεται υπολογιστικά το σύστημα. Τα εισερχόμενα frames αποθηκεύονται σε κυκλικό buffer (circular deque) σταθερής χωρητικότητας, οπότε η κατανάλωση μνήμης παραμένει σταθερή ανεξάρτητα από τη διάρκεια της συνεδρίας. Τέλος, ο υπολογισμός των χαρακτηριστικών και η εκτέλεση του ταξινομητή γίνονται σε ξεχωριστή διεργασία (subprocess) μέσω του multiprocessing, ώστε το γραφικό περιβάλλον να μην καθυστερεί από τις πιο απαιτητικές υπολογιστικές λειτουργίες.

Κεφάλαιο 5ο: Αποτελέσματα και Αξιολόγηση

Το κεφάλαιο αυτό παρουσιάζει τα αποτελέσματα της πειραματικής αξιολόγησης. Ξεκινά με τα αποτελέσματα στο σύνολο ελέγχου (hold-out) για κάθε αλγόριθμο ξεχωριστά, μαζί με τη συντηρητική εκτίμηση γενίκευσης από το cross-validation, ώστε να φανεί πώς συμπεριφέρεται το κάθε μοντέλο μόνο του πριν τη σύγκριση. Στη συνέχεια, τα μοντέλα συγκρίνονται απευθείας μεταξύ τους βάσει των βασικών μετρικών αξιολόγησης. Η ενότητα 5.2 εξετάζει ποια χαρακτηριστικά του σήματος CSI συνέβαλαν περισσότερο στην ταξινόμηση, ώστε να γίνει κατανοητό τι "βλέπει" πραγματικά κάθε αλγόριθμος στα δεδομένα. Οι πίνακες σύγχυσης στην ενότητα 5.3 δείχνουν αναλυτικά πού κάνει λάθη κάθε μοντέλο, δηλαδή αν ταξινομεί λανθασμένα την κίνηση με την ακινησία και αντίστροφα. Στην συνέχεια στην ενότητα 5.4 πραγματοποιείται ανάλυση χρησιμοποιώντας την μετρική AUC μέσω των καμπυλών ROC. Τέλος, το κεφάλαιο 5.5 συζητά συνολικά τα ευρήματα και τις πρακτικές τους επιπτώσεις.

5.1 Αξιολόγηση και Αποτελέσματα ανά Αλγόριθμο

Ο πιο απλός τρόπος να αξιολογηθεί ένα μοντέλο είναι να χωριστούν τα δεδομένα μία φορά σε δεδομένα εκπαίδευσης και σε δεδομένα δοκιμής train/test. Το πρόβλημα όμως με αυτή τη διαδικασία είναι ότι ο αριθμός ακρίβειας που θα εξαχθεί εξαρτάται από το πώς διαχωρίστηκαν τα δεδομένα (split). Με ένα τυχερό split μπορεί ο αλγόριθμος να βγάλει τα καλύτερα αποτελέσματα, ενώ με ένα άτυχο τα χειρότερα. Το cross-validation λύνει ακριβώς αυτό το πρόβλημα. Διαχωρίζει τα δεδομένα σε 5 κομμάτια (folds) και έπειτα τα εκπαιδεύει 5 φορές, κάθε φορά με διαφορετικό κομμάτι ως test, και έτσι προκύπτει ο μέσος όρος.

Το Stratified φροντίζει κάθε κομμάτι (fold) να έχει την ίδια αναλογία κλάσεων με το σύνολο των δεδομένων. Για παράδειγμα εάν υπάρχει αναλογία 50/50 walk/idle, κάθε fold κρατάει αυτή την αναλογία και δεν βγαίνει κατά λάθος ένα fold 80/20.

Το Group είναι το πιο σημαντικό κομμάτι διότι σημαίνει ότι τα παράθυρα από την ίδια καταγραφή μένουν όλα μαζί, είτε στο train είτε στο test, ποτέ μοιρασμένα. Αν δεν συνέβαινε αυτό, το μοντέλο θα έβλεπε στην εκπαίδευση παράθυρα από την ίδια καταγραφή που μετά θα τα έβλεπε και στην δοκιμή (test), και θα μάθαινε ουσιαστικά το συγκεκριμένο σήμα εκείνης της καταγραφής αντί για το γενικό μοτίβο του βαδίσματος, το λεγόμενο data leakage.

Τέλος, το Precision δείχνει από όσα ταξινομήθηκαν ως walk πόσα ήταν όντως walk, ενώ το Recall (Ανάκληση) λειτουργεί αντίστροφα και δείχνει από όσα ήταν walk πόσα εντοπίστηκαν. Επειδή ακριβώς και τα δύο είναι αντίθετα, χρησιμοποιείται το F1-Score, όπου τα συνδυάζει για την καλύτερη εικόνα του αποτελέσματος. Στην παρούσα εργασία τα δύο σχήματα συνδυάζονται ιεραρχικά: τα δεδομένα χωρίζονται πρώτα, σε επίπεδο καταγραφής, σε 80% εκπαίδευσης και 20% ελέγχου (hold-out). Το StratifiedGroupKFold cross-validation ($k = 5$) εκτελείται αποκλειστικά εντός του 80% της εκπαίδευσης και χρησιμεύει για τη σύγκριση και επιλογή των αλγορίθμων ως εκτίμηση γενίκευσης. Το cross-validation εκτελείται εξ ολοκλήρου σε πραγματικά, μη επαυξημένα παράθυρα (η επαύξηση κρατείται εκτός του CV), ενώ η επαύξηση χρησιμοποιείται μόνο στην τελική εκπαίδευση του μοντέλου στο πλήρες σύνολο εκπαίδευσης, που αξιολογείται στο σύνολο ελέγχου. Οι τελικές μετρικές των Πινάκων 5.1 και 5.2 υπολογίζονται στο αποκλεισμένο σύνολο ελέγχου (2.189 παράθυρα), στο οποίο δεν εφαρμόζεται επαύξηση και το οποίο δεν συμμετέχει σε καμία φάση εκπαίδευσης ή επιλογής. Ενδεικτικά, οι μέσες ακρίβειες του 5-fold cross-validation (mean $\pm$ std) ήταν: Random Forest 95,79%

$\pm 2,46$, Gradient Boosting $95,65\% \pm 2,48$, Extra Trees $95,41\% \pm 2,88$, SVM $95,15\% \pm 2,83$, Logistic Regression $95,01\% \pm 3,21$ και k-NN $92,96\% \pm 4,42$. Στο cross-validation κορυφαίος εμφανίζεται ο Random Forest (95,79%), ενώ στο τελικό σύνολο ελέγχου πρώτος είναι ο Gradient Boosting (97,17%). Πρώτον, το cross-validation είναι η συντηρητική εκτίμηση γενίκευσης που υπολογίζεται αποκλειστικά σε πραγματικά, μη επαυξημένα παράθυρα εντός του 80% της εκπαίδευσης, ενώ τα τελικά μοντέλα εκπαιδεύονται στο πλήρες επαυξημένο σύνολο. Από την αυξημένη αυτή ποσότητα δεδομένων ο Gradient Boosting, που χτίζει διαδοχικά δέντρα διορθώνοντας τα σφάλματα των προηγούμενων, τείνει να επωφελείται περισσότερο. Δεύτερον, οι διαφορές των πέντε κορυφαίων αλγορίθμων στο cross-validation (95,01%–95,79%) είναι μικρότερες από τις αντίστοιχες τυπικές αποκλίσεις τους ($\pm 2,5$ έως $\pm 3,2$ μονάδες), οπότε η σειρά κατάταξης μεταξύ τους δεν είναι στατιστικά σημαντική. Τρίτον, το hold-out αποτελεί έναν μοναδικό διαχωρισμό, ενώ το cross-validation είναι μέσος όρος πέντε διαφορετικών. Για τους λόγους αυτούς, ο Gradient Boosting παρουσιάζεται ως ο κορυφαίος αλγόριθμος με βάση την τελική, αξιολόγηση στο αποκλεισμένο σύνολο ελέγχου, με τους Random Forest και Extra Trees να ακολουθούν ουσιαστικά ισοδύναμα. Παρακάτω, στον πίνακα 5.1, σημειώνεται η απόδοση του κάθε αλγορίθμου στο αποκλεισμένο σύνολο ελέγχου.

Αλγόριθμος	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Gradient Boosting	97,17	96,48	97,89	97,17
Random Forest	96,98	96,72	97,25	96,98
Extra Trees	97,12	96,56	97,71	97,12
Logistic Regression	96,85	96,45	97,25	96,85
SVM	96,80	96,71	96,88	96,80
k-NN	93,56	94,81	92,12	93,44

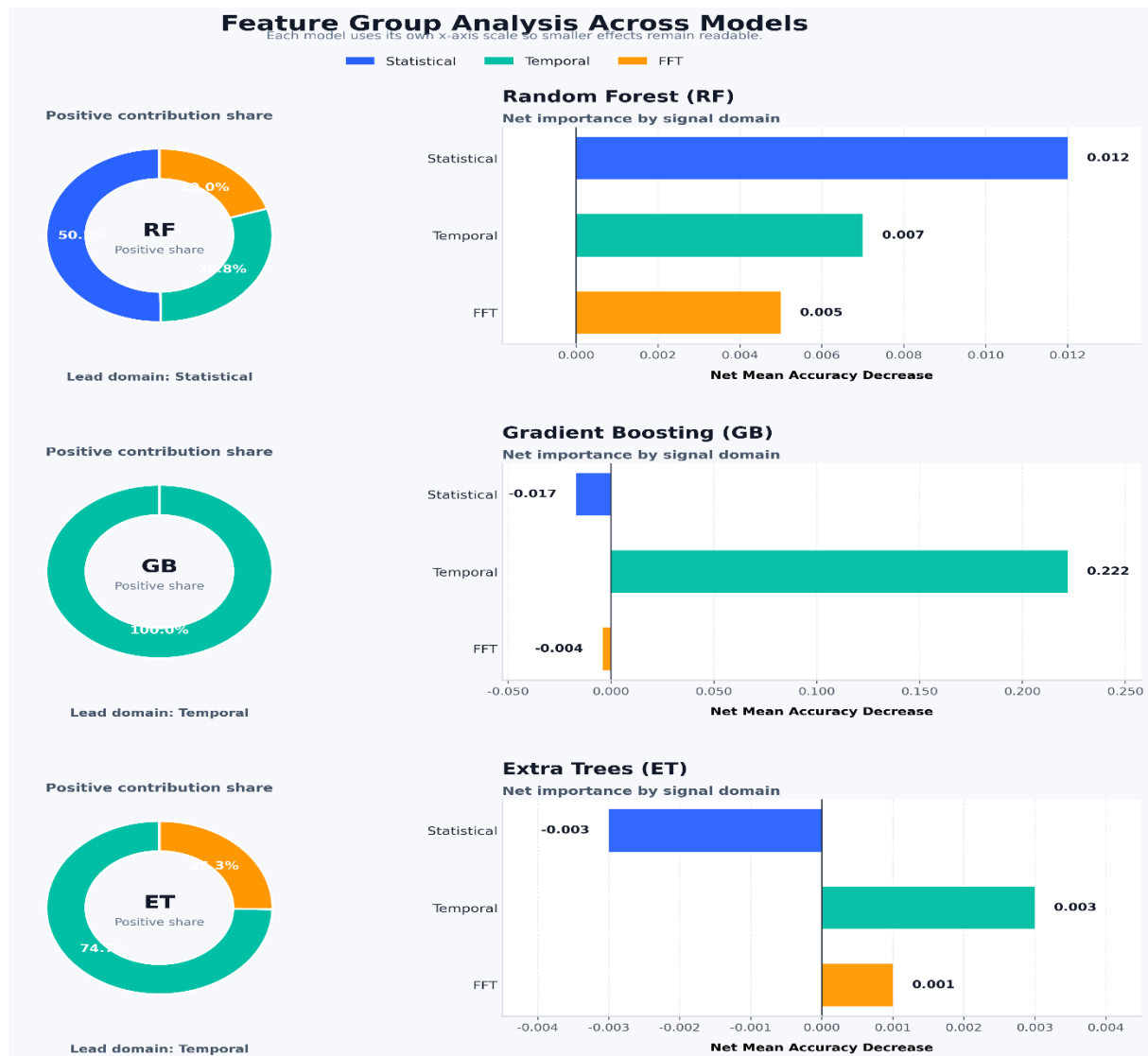
Πίνακας 5.1: Μετρικές απόδοσης στο αποκλεισμένο σύνολο ελέγχου (hold-out 20% των καταγραφών, 2.189 παράθυρα) για τους έξι αλγορίθμους ML.

5.2 Ανάλυση Σημαντικότητας Χαρακτηριστικών

Ορισμένοι αλγόριθμοι όπως το SVM λειτουργούν ως μη ερμηνεύσιμα μοντέλα («black-box»). Δέχονται δηλαδή τα χαρακτηριστικά (features), εκτελούν την ταξινόμηση, χωρίς να υπάρχει εύκολος τρόπος να

διαπιστωθεί τι ακριβώς αξιοποίησε το μοντέλο για να αποφασίσει. Τα δένδρικά μοντέλα από την άλλη όπως το Random Forest και το Gradient Boosting είναι διαφορετικά, μιας και από τη φύση τους και καταγράφουν πόσες φορές χρησιμοποιήθηκε κάθε χαρακτηριστικό (feature) για να γίνει ο διαχωρισμός (split). Για παράδειγμα εάν το waveform_length αναδεικνύεται πρώτο, προκύπτει ότι το Random Forest αποφασίζει κυρίως με βάση το πόσο έντονη μεταβλητότητα παρουσιάζει το σήμα στον χρόνο, εάν από την άλλη αναδεικνυόταν πρώτο κάτι που δεν έχει λογική ερμηνεία, αυτό θα σήμαινε πιθανότατα πρόβλημα στα δεδομένα. Βέβαια, το SVM όμως δεν διαθέτει ενσωματωμένη μετρική σημαντικότητας οπότε το permutation importance, ως μέθοδος ανεξάρτητη μοντέλου, θα μπορούσε να εφαρμοστεί και σε αυτό, ωστόσο στην παρούσα ανάλυση εφαρμόστηκε στα δένδρικά μοντέλα, οπότε για το SVM η αξιολόγηση βασίζεται στα αποτελέσματα των υπόλοιπων μοντέλων ώστε να κριθεί αν τα χαρακτηριστικά (features) που επιλέχθηκαν έχουν νόημα.

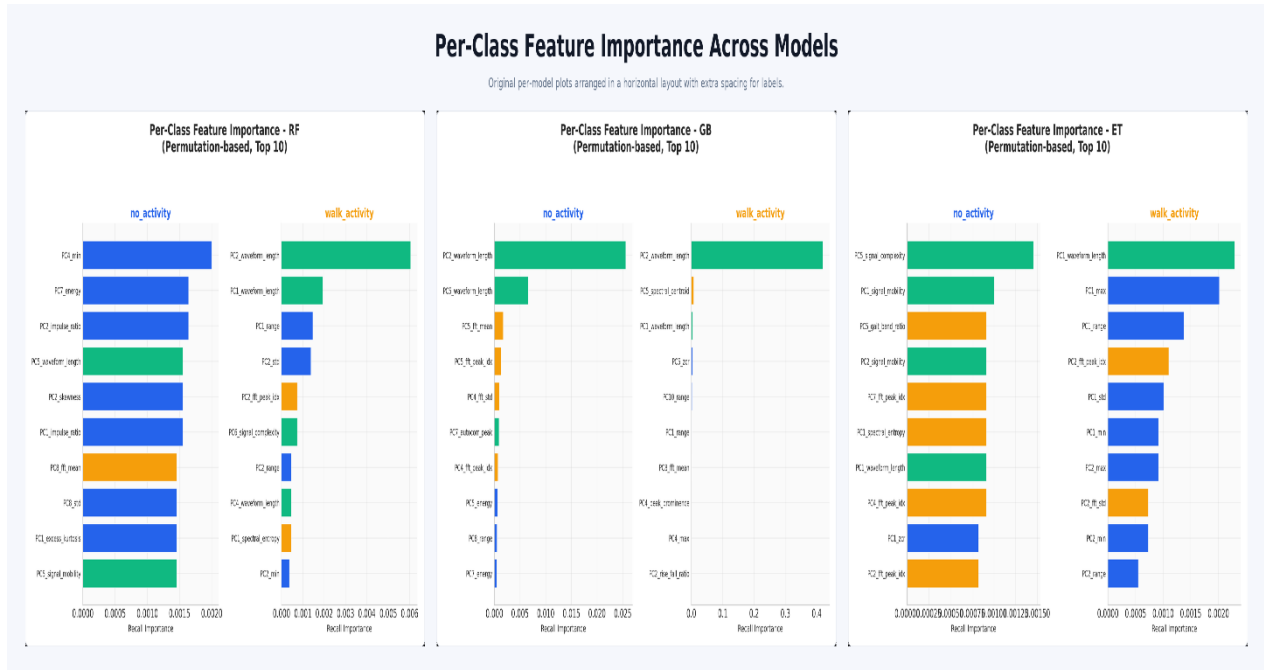
Παρακάτω παρουσιάζεται η σημαντικότητα των χαρακτηριστικών των 3 αλγορίθμων Random Forest, Gradient Boosting και Extra Trees στις εικόνες 5.1 και 5.2. Η σημαντικότητα στην Εικόνα 5.1 υπολογίζεται με permutation importance, δηλαδή ως η μέση μείωση της ακρίβειας όταν οι τιμές ενός χαρακτηριστικού ανακατεύονται τυχαία, γι' αυτό ορισμένες τιμές εμφανίζονται αρνητικές (το ανακάτεμα του συγκεκριμένου χαρακτηριστικού δεν μειώνει, ή και οριακά βελτιώνει, την ακρίβεια).



Εικόνα 5.1. Ανάλυση χαρακτηριστικών των τριών αλγορίθμων.

Παρατηρώντας την εικόνα 5.1 το ενδιαφέρον εδώ είναι ότι οι τρεις αλγόριθμοι δεν συμφωνούν σχετικά με ποια χαρακτηριστικά είναι καλύτερα για την αναγνώριση. Ο RF βασίζεται κυρίως στα στατιστικά χαρακτηριστικά (statistical features 0.012), με τα χρονικά (temporal) και τα συχνοτικά (FFT) να ακολουθούν, δηλαδή κάνει χρήση και των τριών κατηγοριών αρκετά ισορροπημένα. Ο GB από την άλλη στηρίζεται σχεδόν αποκλειστικά στα χρονικά (temporal 0.222), με τα στατιστικά (statistical) και τα συχνοτικά FFT να βγαίνουν και αρνητικά, που σημαίνει ότι όχι μόνο δεν τα χρησιμοποιεί, αλλά η παρουσία τους ίσως να συγχέει λίγο το μοντέλο. Το ET συμφωνεί με το GB ως προς την επιλογή των χρονικών, αλλά σε πολύ μικρότερη κλίμακα.

Δηλαδή, αυτό σημαίνει ότι τα χρονικά χαρακτηριστικά (temporal features), δηλαδή αυτά που μετρώνε πώς αλλάζει το σήμα στον χρόνο, είναι πιο σταθερά και χρήσιμα σε όλα τα μοντέλα. Τα στατιστικά φαίνεται να βοηθούν τον RF αλλά όχι τους άλλους δύο, άρα μάλλον δεν είναι τόσο κρίσιμα.



Εικόνα 5.2. Ανάλυση χαρακτηριστικών με βάση τις κλάσεις (walk, idle) των τριών αλγορίθμων.

Αυτό που αμέσως παρατηρείται στην εικόνα 5.2 είναι ότι και στους τρεις αλγορίθμους το waveform_length κυριαρχεί για την κλάση walk_activity, με διαφορά από τα υπόλοιπα features. Αντίθετα, για το no_activity τα πράγματα είναι πιο κατανεμημένα, με τις PCA συνιστώσες και χρονικά χαρακτηριστικά να μοιράζονται τη σημαντικότητα χωρίς κάποιο να ξεχωρίζει καθαρά. Στον αλγόριθμο GB τα πράγματα αλλάζουν και το μόνο κύριο σημαντικό χαρακτηριστικό είναι το waveform_length για την κλάση walk, ενώ από κάτω τα υπόλοιπα είναι πολύ μικρότερα. Ο RF και ο ET είναι πιο ισορροπημένοι, με περισσότερα features να συνεισφέρουν παράλληλα.

Αυτό επιβεβαιώνει όσα φάνηκαν και στα διαγράμματα του σήματος. Το βάδισμα αναγνωρίζεται κυρίως από το πόσο γρήγορα αλλάζει το σήμα στον χρόνο, που είναι ακριβώς αυτό που μετράει το waveform_length. Για την ηρεμία δεν υπάρχει ένα τέτοιο χαρακτηριστικό γνώρισμα, οπότε τα μοντέλα συνδυάζουν πολλά features για να την αναγνωρίσουν.

5.3 Πίνακες Σύγκρισης

Ο πίνακας σύγκρισης είναι ο πιο άμεσος τρόπος διερεύνησης του πού ακριβώς αστοχεί ένας ταξινομητής. Η συνολική ακρίβεια απλά δίνει έναν αριθμό, χωρίς όμως να φανερώνει αν τα λάθη πέφτουν περισσότερο στη μία κλάση ή είναι μοιρασμένα και στις δύο. Ο πίνακας δείχνει ακριβώς αυτή τη λεπτομέρεια, δείχνοντας δηλαδή τις προβλέψεις του μοντέλου με τις πραγματικές ετικέτες.

Για το δυαδικό πρόβλημα Walk (κλάση 1) vs Idle (κλάση 0) προκύπτουν τέσσερις περιπτώσεις. Όταν ένα δείγμα Walk ταξινομηθεί σωστά, καταγράφεται ως True Positive (TP), ενώ ένα Idle που χαρακτηρίζεται σωστά είναι True Negative (TN). Όταν το μοντέλο ερμηνεύσει μια κατάσταση ακινησίας ως Walk τότε καταγράφεται False Positive (FP), δηλαδή ψευδή ένδειξη κίνησης, και όταν ένα Walk χαρακτηριστεί ως Idle τότε καταγράφεται ως False Negative (FN), δηλαδή μια πραγματική κίνηση που πέρασε απαρατήρητη. Η διάκριση αυτή είναι σημαντική για το παρόν σύστημα. Για παράδειγμα ένα FP σημαίνει ότι ο ταξινομητής αντιλαμβάνεται την κίνηση εκεί που δεν υπάρχει, ουσιαστικά ένα false alarm ενώ ένα FN, αντίθετα, σημαίνει ότι μια πραγματική κίνηση πέρασε απαρατήρητη. Στις περισσότερες εφαρμογές HAR οι δύο αυτοί τύποι σφάλματος δεν είναι το ίδιο, και ο πίνακας είναι αυτός που επιτρέπει τον διαχωρισμό τους.

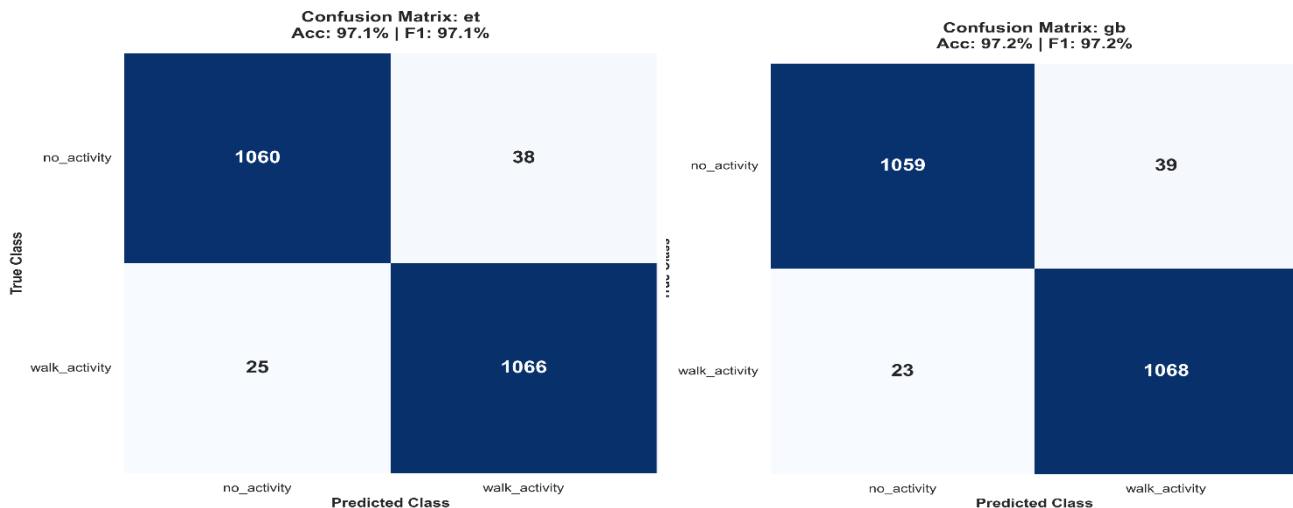
Οι τιμές του Πίνακα 5.2 αφορούν τα παράθυρα του συνόλου ελέγχου, δηλαδή το 20% των καταγραφών (TEST_RATIO = 0.2), με τον διαχωρισμό να γίνεται πάντα σε επίπεδο καταγραφής. Από τα περίπου 11.800 παράθυρα του συνόλου δεδομένων που αναλογούν στο σύνολο ελέγχου περίπου 2.360 παράθυρα, επειδή όμως ο πραγματικός αριθμός frames ανά καταγραφή αποκλίνει ελαφρώς από τον ονομαστικό (βλ. Κεφάλαιο 4.4), η προεπεξεργασία περικόπτει ορισμένα δείγματα. Τα παράθυρα δημιουργούνται ξεχωριστά ανά καταγραφή με απόρριψη των ημιτελών τελικών παραθύρων, το τελικό σύνολο ελέγχου περιλαμβάνει 2.189 παράθυρα (1.091 της κλάσης Walk και 1.098 της κλάσης Idle). Σημειώνεται ότι στο σύνολο ελέγχου χρησιμοποιούνται αποκλειστικά πραγματικά, μη επανυζημένα παράθυρα.

Αλγόριθμος	TP (Walk→Walk)	FP (Idle→Walk)	FN (Walk→Idle)	TN (Idle→Idle)
Gradient Boosting	1068	39	23	1059
Random Forest	1061	36	30	1062
Extra Trees	1066	38	25	1060
Logistic Regression	1061	39	30	1059
SVM	1057	36	34	1062
k-NN	1005	55	86	1043

Πίνακας 5.2: Πίνακες σύγχυσης ανά αλγόριθμο. (Κλάση 0 = Idle, Κλάση 1 = Walk, σύνολο ελέγχου 2.189 παράθυρα).

Παρακάτω εξετάζονται συγκεκριμένα οι πίνακες σύγχυσης (Εικόνα 5.3 έως 5.5) των αλγορίθμων.

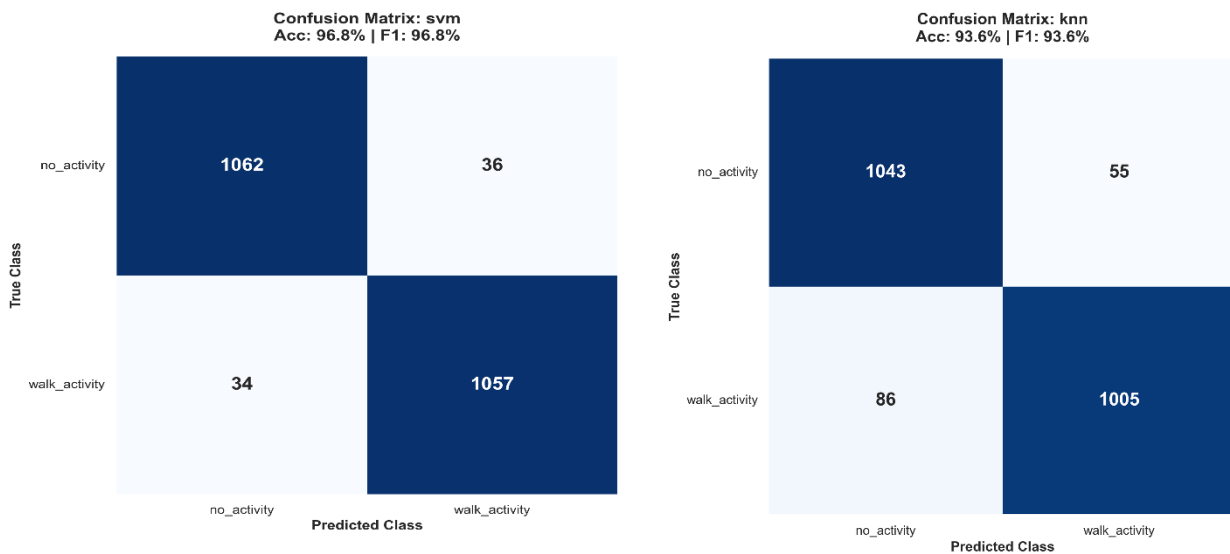
Ξεκινώντας από το Gradient Boosting (gb) στην εικόνα 5.3, παρατηρούνται 1059 σωστά no_activity και 1068 σωστά walk_activity, με συνολική ακρίβεια (accuracy) 97,2%. Τα λάθη πάντως δεν είναι συμμετρικά, με 39 FP (idle που πέρασαν για walk) απέναντι σε 23 FN (walk που χάθηκαν). Επιπλέον, ο ταξινομητής έχει μια ελαφριά τάση να προβλέπει λανθασμένα το walk σε δύσκολες περιπτώσεις. Στο Extra Trees (et) η εικόνα είναι σχεδόν ίδια, με 1060 σωστά Idle και 1066 σωστά Walk, με 38 FP και 25 FN και ακρίβεια 97,1%. Και τα δύο μοντέλα είναι δενδρικά (tree-based), οπότε εντοπίζουν τις ίδιες αλληλεπιδράσεις χαρακτηριστικών και πιθανότατα αποτυγχάνουν σε παρόμοια δείγματα. Το ίδιο μοτίβο ασυμμετρίας (περισσότερα FP από FN) επαναλαμβάνεται, πράγμα που δείχνει ότι δεν πρόκειται για θόρυβο αλλά για μια τάση του μοντέλου πάνω στα ίδια τα δεδομένα. Σε ένα HAR (Human Activity Recognition) σύστημα συνήθως είναι προτιμότερη μια χαμένη ανίχνευση κίνησης παρά ένας ψευδής συναγερμός (false alarm), οπότε η παρουσία λίγο υψηλότερου recall στο Walk (97,9% για το gb, 97,7% για το et) έναντι του precision (96,5% και 96,6% αντίστοιχα) είναι καλύτερο. Συμπεραίνοντας, και τα 2 μοντέλα κάνουν παρόμοια λάθη (39 vs 38 FP, 23 vs 25 FN) που σημαίνει ότι ίσως μερικά σύνολα δειγμάτων μπερδεύουν και τα δύο συστήματα.



Εικόνα 5.3. Πίνακες Σύγχυσης των αλγορίθμων Extra Trees και Gradient Boosting.

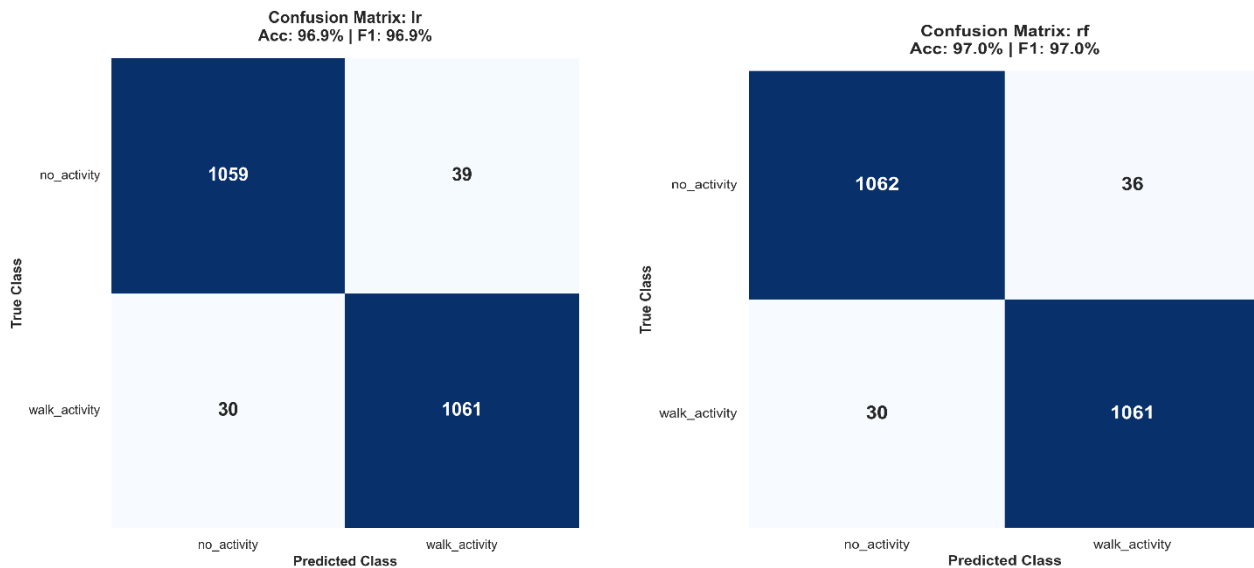
Στο SVM (Εικόνα 5.4) παρατηρείται κάτι ενδιαφέρον σε σύγκριση με τα προηγούμενα. 1062 σωστά Idle και 1057 σωστά Walk, με 36 FP και 34 FN και accuracy 96,8%. Η ακρίβεια εδώ είναι λίγο

χαμηλότερη από τα δενδρικά μοντέλα, αλλά αυτό που ξεχωρίζει είναι η συμμετρία των λαθών. Σχεδόν ίδια τα FP και FN, πράγμα που δεν συνέβαινε στα δενδρικά μοντέλα. Το SVM όμως δεν αντιλαμβάνεται την χρήση των χαρακτηριστικών με τον ίδιο τρόπο που τα αξιοποιεί ένα δενδρικό μοντέλο (tree ensemble), οπότε η μικρή προτίμηση προς την κλάση Walk που έδιναν τα gb/et εδώ χάνεται. Το k-NN, αντίθετα, είναι πιο πίσω με 1043 σωστά Idle και 1005 σωστά Walk, με 55 FP και 86 FN και ακρίβεια (accuracy) 93,6%. Πρώτον, ο συνολικός αριθμός λαθών σχεδόν διπλασιάζεται σε σχέση με τα προηγούμενα μοντέλα (141 έναντι ~62-70). Δεύτερον, και πιο ενδιαφέρον, η ασυμμετρία αντιστρέφεται, με τα FN να είναι περισσότερα από τα FP και άρα να χάνονται περισσότερα Walk παρά να παρερμηνεύονται τα idle.



Εικόνα 5.4. Πίνακες Σύγχυσης των αλγορίθμων SVM και k-NN.

Το Logistic Regression από την άλλη δίνει 1059 σωστά Idle και 1061 σωστά Walk, με 39 FP και 30 FN, και ακρίβεια 96,9%. Το ότι ένα γραμμικό μοντέλο τα καταφέρνει τόσο καλά οφείλεται μάλλον στην ποιότητα της προεπεξεργασίας και στα 250 χαρακτηριστικά. Το μοτίβο $FP > FN$ επανεμφανίζεται, αν και είναι πιο ήπιο από τα gb/et. Στην συνέχεια το Random Forest (rf) δίνει 1062 σωστά Idle και 1061 σωστά Walk, με 36 FP και 30 FN, και ακρίβεια (accuracy) 97,0%. Το RF φαίνεται να έχει παρόμοια FN με το lr, αλλά λίγο λιγότερα FP. Ουσιαστικά είναι το πιο ισορροπημένο μοντέλο από τα τρία, και η σταθερότητα των αποτελεσμάτων επιβεβαιώνει ότι το σήμα στα χαρακτηριστικά είναι ισχυρό.



Εικόνα 5.5. Πίνακες Σύγχυσης των αλγορίθμων Random Forest και Logistic Regression.

Αξιολογώντας τα αποτελέσματα, η πρώτη διαπίστωση είναι ότι πέντε από τα έξι μοντέλα συγκλίνουν, μεταξύ του 96,8% και του 97,2%. Η διαφορά ανάμεσα στον καλύτερο (gb, 97,2%) και τον χειρότερο από αυτούς τους πέντε (svm, 96,8%) είναι 0,4. Επίσης, ο k-NN είναι ο μοναδικός αλγόριθμος που αποκλίνει, με 93,6% και 141 συνολικά λάθη όταν τα υπόλοιπα κυμαίνονται μεταξύ 62 και 70.

5.4 Αξιολόγηση μέσω Καμπυλών ROC και Μετρικής AUC

5.4.1 Η ROC καμπύλη

Ένας ταξινομητής που καλείται να αποφασίσει αν σε ένα παράθυρο σήματος υπάρχει κίνηση δεν παράγει συνήθως απάντηση της μορφής «ναι/όχι», αλλά εκφράζει μια πιθανότητα μεταξύ 0 και 1 [13][28]. Η τελική απόφαση εξαρτάται από το όριο (threshold) που τίθεται πάνω σε αυτή την πιθανότητα. Έτσι σε ένα χαμηλό όριο εντοπίζει σχεδόν όλες τις πραγματικές κινήσεις αλλά ταυτόχρονα παράγει και αρκετούς ψευδείς συναγερμούς. Αντίθετα σε ένα υψηλό όριο περιορίζονται τα λάθη αυτής της κατηγορίας αλλά αφήνει ωστόσο εκτός και αληθινές κινήσεις.

Η ROC καμπύλη αποτυπώνει ακριβώς αυτή την ισορροπία, διατρέχοντας όλες τις δυνατές τιμές του ορίου. Στον οριζόντιο άξονα τοποθετείται το ποσοστό λάθος συναγερμών (False Positive Rate), δηλαδή οι περιπτώσεις όπου το μοντέλο σημείωσε κίνηση χωρίς αυτή να υπάρχει. Στον κάθετο άξονα τοποθετείται το ποσοστό σωστών εντοπισμών (True Positive Rate), δηλαδή το πόσες από τις πραγματικές κινήσεις αναγνωρίστηκαν σωστά [13][28]. Η ιδανική καμπύλη ανεβαίνει αμέσως προς την πάνω αριστερή γωνία, γεγονός που αντιστοιχεί σε σχεδόν πλήρη ανίχνευση με ελάχιστα λάθη. Η διαγώνια γκρι γραμμή χρησιμεύει ως αναφορά τύχης, δηλαδή ως η απόδοση ενός μοντέλου που αποφασίζει στα τυφλά [13][28].

5.4.2 Το AUC

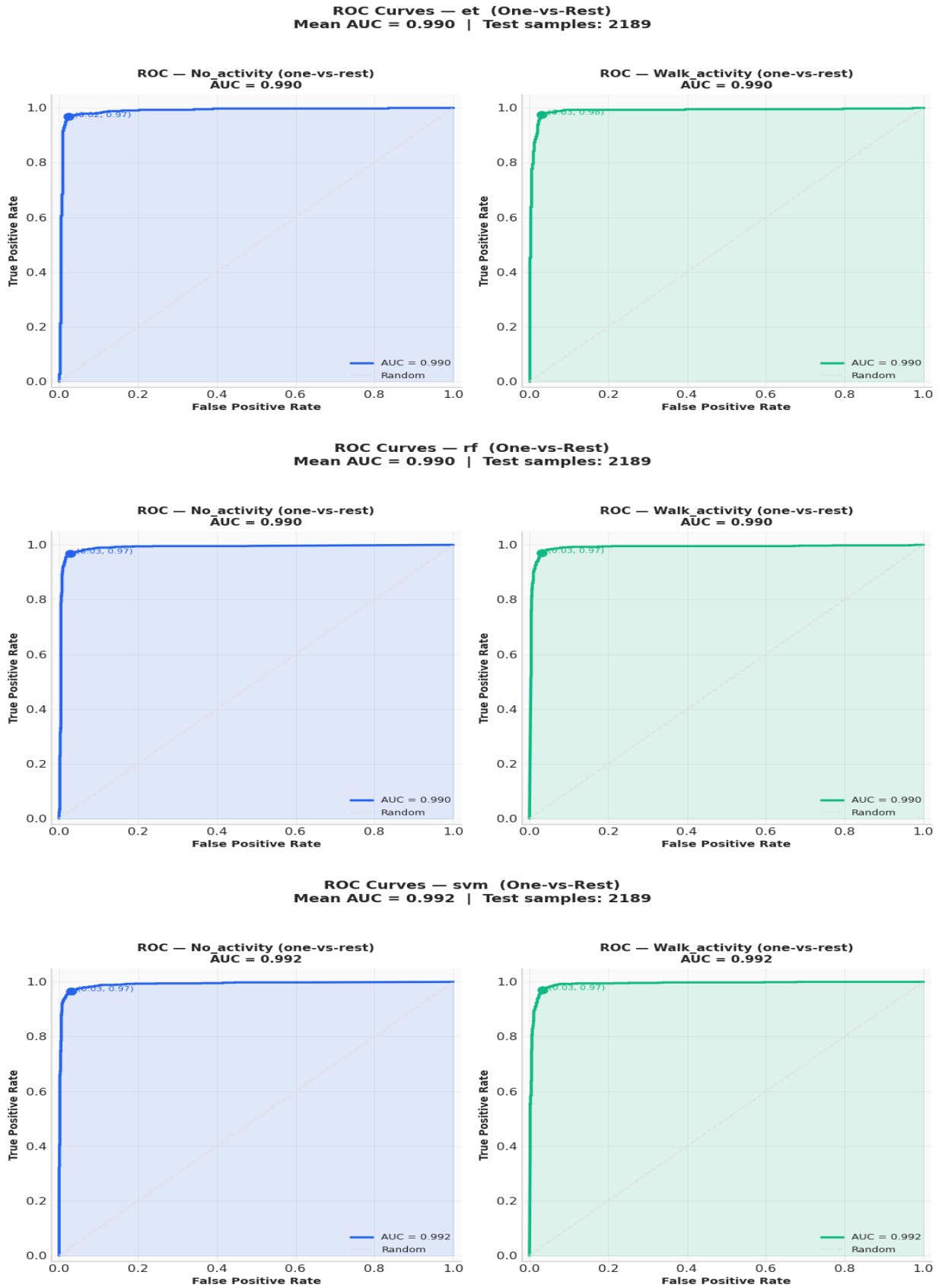
Το AUC (Area Under the Curve) αντιστοιχεί στο εμβαδόν κάτω από τη ROC καμπύλη. Εφόσον οι δυο άξονες κυμαίνονται από 0 ως 1, το AUC είναι επίσης ένας αριθμός μέσα σε αυτό το διάστημα. Όσο πιο κοντά στη μονάδα βρίσκεται, τόσο καλύτερα ο ταξινομητής διαχωρίζει τις δυο κλάσεις. Η τιμή 0,5 αντιστοιχεί σε μια τυχαία πρόβλεψη, ενώ τιμές κοντά στο 0,80 θεωρούνται μέτριες ενώ οι τιμές γύρω στο 0,90 θεωρούνται καλές και πάνω από 0,95 πραγματικά υψηλές. Το πλεονέκτημα του AUC είναι ότι δεν εξαρτάται από την επιλογή ενός συγκεκριμένου ορίου, δίνοντας δηλαδή μια συνολική εικόνα του πόσο διαχωρίσιμες είναι οι δυο κλάσεις στον χώρο που έχει μάθει το μοντέλο [13][28].

Το σημείο που σημειώνεται πάνω στην καμπύλη, με τιμές της μορφής (0,03, 0,97), είναι το λεγόμενο operating point. Αντιστοιχεί στο όριο που χρησιμοποιείται στην πραγματική λειτουργία του ταξινομητή και μεταφράζει την επιλογή αυτή σε συγκεκριμένα ποσοστά, στη συγκεκριμένη περίπτωση 3% λάθος συναγερμούς έναντι 97% σωστών εντοπισμών [13][28].

5.4.3 Αξιολόγηση Αποτελεσμάτων

Στην παρακάτω εικόνα (Εικόνα 5.6.) και οι τρεις ταξινομητές δίνουν σχεδόν ταυτόσημη εικόνα. Το ET και το RF κυμαίνονται γύρω στο 0,990 και το SVM αυξάνεται ελάχιστα στο 0,992. Σε όλες τις περιπτώσεις η καμπύλη φτάνει γρήγορα κοντά στο TPR=1 με χαμηλό FPR, δηλαδή το μοντέλο σχεδόν αμέσως αντιλαμβάνεται όλες τις πραγματικές κινήσεις χωρίς εσφαλμένους συναγερμούς. Τα σημεία λειτουργίας γύρω στο (0.02–0.03, 0.97–0.98) δείχνουν ότι στο πραγματικό σημείο λειτουργίας το σύστημα κρατάει 97–98% σωστούς εντοπισμούς με μόλις 2–3% ψεύτικους. Πρακτικά, αυτό σημαίνει ότι από 100 παράθυρα κίνησης χάνονται το πολύ 2–3, ενώ στα παράθυρα ηρεμίας τα σφάλματα είναι σπάνια.

Η συμμετρία ανάμεσα στην No_activity (μπλε) και την Walk_activity (πράσινη) είναι αναμενόμενη, καθώς πρόκειται για ένα δυαδικό (Binary problem) πρόβλημα, οπότε ό,τι σωστό πιάνει στη μία κλάση είναι αυτόματα σωστό και στην άλλη. Το ότι τρεις τόσο διαφορετικοί αλγόριθμοι (δύο tree-based και ένας kernel-based) καταλήγουν στο ίδιο σημείο υποδεικνύει ότι το βάρος δεν το σηκώνει ο ταξινομητής αλλά τα ίδια τα χαρακτηριστικά (features). Η επιλογή χαρακτηριστικών (feature space) και η προεπεξεργασία διαχωρίζει ήδη τις δυο κλάσεις, και ο εκάστοτε αλγόριθμος απλά κάνει το επόμενο βήμα.

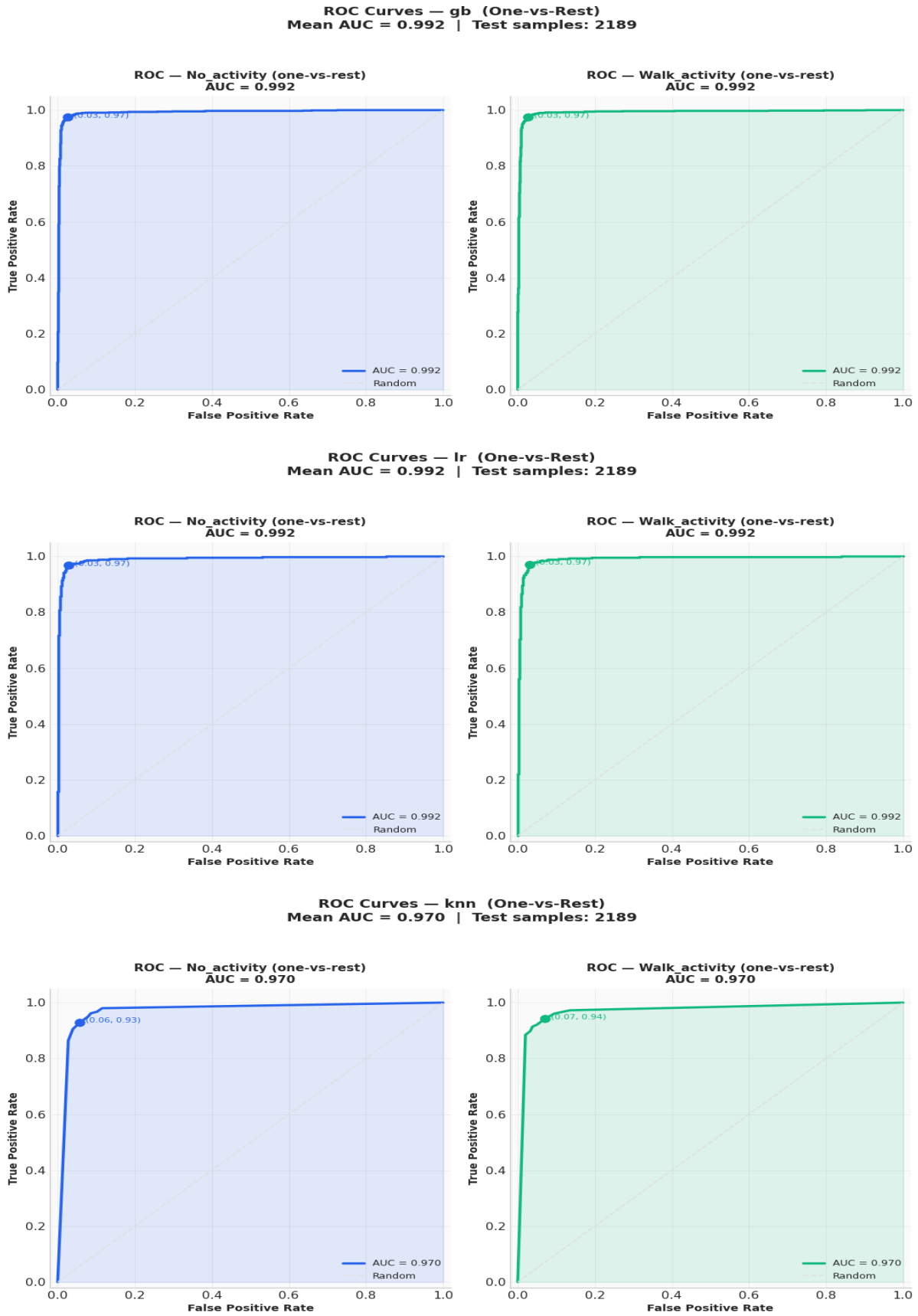


Εικόνα 5.6. Καμπύλες ROC (One-vs-Rest) για τους ταξινομητές Extra Trees, Random Forest και SVM στο σύνολο ελέγχου.

Στην παρακάτω εικόνα (Εικόνα 5.7) βρίσκονται δύο μοντέλα που φτάνουν στο 0,992 δηλαδή στο ίδιο επίπεδο με το SVM της πρώτης εικόνας (Εικόνα 5.6), και ένα που ξεχωρίζει προς τα κάτω, ο k-NN ο οποίος φτάνει στο 0,970. Το ότι ένας απλός γραμμικός ταξινομητής όπως η Logistic Regression φτάνει στο ίδιο επίπεδο με το Gradient Boosting αποτελεί από μόνο του χαρακτηριστικό εύρημα. Σημαίνει ότι μετά την προεπεξεργασία και το PCA, οι δύο κλάσεις είναι σχεδόν γραμμικά διαχωρίσιμες στο νέο χώρο, οπότε δεν απαιτείται κάποιο πιο πολύπλοκο μοντέλο για τον διαχωρισμό τους.

Ο k-NN είναι η πιο ενδιαφέρουσα εξαίρεση. Παρατηρείται ότι η καμπύλη του είναι πιο «γωνιώδης» σε σχέση με τις άλλες και δεν είναι ομαλή. Αυτό συμβαίνει επειδή ο k-NN δεν παράγει πραγματικές συνεχείς πιθανότητες, αλλά μάλλον δίνει διακριτά κλάσματα του τύπου «3 στους 5 γείτονες προβλέπουν κίνηση», οπότε υπάρχουν λίγα διαφορετικά thresholds που μπορούν να δοκιμαστούν. Το σημείο λειτουργίας του (0,06–0,07, 0,93–0,94) είναι λίγο χειρότερο από τα υπόλοιπα, με σχεδόν τριπλάσιο ποσοστό λάθος συναγεμίων. Συνολικά πάντως ακόμα και αυτός ο αλγόριθμος είναι πάνω από το 0,97 σε AUC, οπότε δεν αποτυγχάνει, απλώς είναι ο πιο αδύναμος από όλους τους υπόλοιπους.

Συνοψίζοντας τα δύο σχήματα, προκύπτει ότι 5 από τους 6 ταξινομητές συμπίπτουν σχεδόν στο ίδιο σημείο γύρω στο 0,990–0,992. Όταν τόσοι διαφορετικοί αλγόριθμοι, με τόσο διαφορετική λογική, συγκλίνουν στην ίδια απόδοση, το πιθανότερο είναι ότι έχει πλέον προσεγγιστεί ένα όριο όπου τα ίδια τα δεδομένα και τα χαρακτηριστικά (features) αναλαμβάνουν σημαντικότερο ρόλο από τον εκάστοτε αλγόριθμο. Τέλος η επιλογή του τελικού μοντέλου για την εργασία δεν καθορίζεται από την ακρίβεια αλλά από άλλα κριτήρια, όπως η ταχύτητα inference, το μέγεθος του μοντέλου και το πόσο εύκολα ερμηνεύεται.



Εικόνα 5.7. Καμπύλες ROC (One-vs-Rest) για τους ταξινομητές Gradient Boosting, Logistic Regression και k-NN στο σύνολο ελέγχου.

5.5 Μελέτη Απομόνωσης Σταδίων (Ablation Study)

Μέχρι εδώ κάθε στάδιο της προεπεξεργασίας δικαιολογήθηκε θεωρητικά και με βάση τα διαγράμματα του σήματος. Αυτό όμως δεν λέει πόσο πραγματικά μετράει το καθένα στην τελική ακρίβεια. Για να εκτιμηθεί αυτό ποσοτικά, πραγματοποιήθηκε μια συστηματική μελέτη ablation όπου διατηρείται όλο το pipeline σταθερό και κάθε φορά αλλάζει ένα μόνο στοιχείο, μετρώντας πάντα στο ίδιο αποκλεισμένο σύνολο ελέγχου. Συνολικά δοκιμάστηκαν 24 διαφορετικές διαμορφώσεις και στα έξι μοντέλα. Η πλήρης διάταξη (baseline) δίνει 97,1% για το Random Forest και 97,2% για το Gradient Boosting, και αυτές οι τιμές χρησιμοποιούνται ως σημείο αναφοράς για κάθε σύγκριση. Οι τιμές baseline της μελέτης ablation προέρχονται από συγκεκριμένη εκτέλεση με σταθερό seed και ενδέχεται να διαφέρουν κατά $\pm 0,1-0,2$ μονάδες από τις αντίστοιχες τιμές του Πίνακα 5.1, χωρίς αυτό να μεταβάλλει τα συμπεράσματα.

Το πιο καθαρό εύρημα αφορά το φίλτρο Butterworth όπου όταν αφαιρείται εντελώς, τα περισσότερα μοντέλα καταρρέουν. Πιο συγκεκριμένα το SVM, το k-NN και το Gradient Boosting πέφτουν στο 49,8–49,9%, δηλαδή ουσιαστικά προβλέπουν στην τύχη για ένα δυαδικό πρόβλημα, το Logistic Regression στο 66,2% και το Extra Trees στο 86,8%. Το μόνο μοντέλο που διατηρεί την απόδοσή του είναι το Random Forest, στο 96,2%. Πρακτικά, χωρίς το χαμηλοπερατό φιλτράρισμα ο υψηλόσυχνος θόρυβος κυριαρχεί στα χαρακτηριστικά, και τα μοντέλα που στηρίζονται σε αποστάσεις ή σε περιθώρια διαχωρισμού δεν εντοπίζουν μια σαφή συγκεκριμένη δομή. Το Random Forest, με τα τυχαία υποσύνολα χαρακτηριστικών και την ψηφοφορία των δέντρων, είναι αρκετά πιο ανθεκτικό στον θόρυβο ώστε να εντοπίσει ακόμη τη χρήσιμη δομή. Το πιο ενδιαφέρον σημείο εδώ είναι ότι το Gradient Boosting, αν και δενδρικό σαν το Random Forest, καταρρέει κι αυτό, αυτό μάλλον οφείλεται στο ότι χτίζει διαδοχικά δέντρα πάνω στα σφάλματα και έτσι προσαρμόζεται ταυτόχρονα και στον θόρυβο. Συνεπώς, το φίλτρο δεν είναι ένα ακόμη βήμα καθαρισμού αλλά το στάδιο με τον σημαντικότερο ρόλο στον διαχωρισμό.

Από την άλλη, η συχνότητα αποκοπής από μόνη της δεν αλλάζει πολλά γύρω από την επιλεγμένη τιμή. Στα 5 Hz το Random Forest ανεβαίνει οριακά στο 97,3%, στα 10 Hz (baseline) μένει στο 97,1%, ενώ στα 30 Hz πέφτει στο 96,1% και στα 50 Hz στο 96,9%. Η διαφορά είναι μικρή, οπότε τα 10 Hz παραμένουν μια καλή επιλογή, αν και τα δεδομένα δείχνουν ότι ένα ακόμη χαμηλότερο cutoff δεν θα έβλαπτε. Άλλωστε η ενέργεια του βαδίσματος βρίσκεται ούτως ή άλλως κάτω από αυτή τη ζώνη. Επιπλέον, αντίστοιχα παρουσιάζονται η ακρίβεια και η συμπεριφορά των άλλων μοντέλων στον πίνακα 5.5.

Τα υπόλοιπα στάδια επηρεάζουν την ακρίβεια πολύ λιγότερο. Ο αριθμός των συνιστωσών PCA, από 5 έως 20, κινεί το Random Forest μέσα σε ένα στενό εύρος μεταξύ 96,8–97,2%, οπότε οι 10 συνιστώσες που επιλέχθηκαν δεν είναι κρίσιμη παράμετρος, απλώς μια καλή ισορροπία. Το μέγεθος του παραθύρου έχει λίγο μεγαλύτερη επίδραση για παράδειγμα στα 50 δείγματα η ακρίβεια πέφτει στο 96,4%, ενώ στα 150 ανεβαίνει στο 97,4%. Το μεγαλύτερο παράθυρο βοηθάει λιγάκι, κοστίζει όμως σε καθυστέρηση πρόβλεψης, και γι' αυτό κρατήθηκε το παράθυρο των 100 δειγμάτων (1 δευτερόλεπτο). Τέλος, το βήμα (step) μεταξύ 25, 50 και 75 δειγμάτων αφήνει την ακρίβεια σχεδόν αμετάβλητη, γύρω στο 96,8–97,4%.

Δύο στάδια συμπεριφέρονται διαφορετικά από το αναμενόμενο: α) η χρονική διαφοροποίηση και β) η επαύξηση. Πιο συγκεκριμένα, η χρονική διαφοροποίηση (temporal differencing), όταν αφαιρεθεί, αφήνει το Random Forest στο 97,4% και το Extra Trees στο 97,3%, δηλαδή δεν επηρεάζει τα δενδρικά μοντέλα. Στα μοντέλα που εξαρτώνται από κλίμακα όμως πέφτει αισθητά, με το SVM να κατεβαίνει στο 94,1% και το Logistic Regression στο 95,3%. Κάτι ανάλογο ισχύει και για την επαύξηση δεδομένων, χωρίς augmentation το Random Forest μένει στο 97,2%, ουσιαστικά ίδια τιμή με το αρχικό (baseline). Με άλλα λόγια, στην επίδοση εντός του συνόλου ελέγχου η επαύξηση δεν βελτιώνει την

ακρίβεια. Το αποτέλεσμα αυτό είναι αναμενόμενο, καθώς το σύνολο ελέγχου προέρχεται από την ίδια κατανομή με το σύνολο εκπαίδευσης καθώς είναι αρκετά δύσκολο η αξία των πραγματικών δεδομένων να αντικατασταθεί από τα εικονικά δεδομένα και να αποδώσουν με την ίδια ακρίβεια [10][18]. Σε ένα cross-environment σενάριο το αποτέλεσμα ενδέχεται να διαφέρει, ωστόσο αυτό δεν επιβεβαιώνεται από τα τρέχοντα δεδομένα.

Στην περίπτωση όπου αφαιρεθούν ταυτόχρονα το φίλτρο, η διαφοροποίηση και η επαύξηση, η απόδοση υποβαθμίζεται σημαντικά, με το SVM η ακρίβεια του να πέφτει στο 49,8%, το Logistic Regression στο 61,9%, ενώ το Random Forest κρατιέται στο 92,7% και το Gradient Boosting στο 94,3%. Δοκιμάστηκαν επίσης μόνο η διατήρηση των χαρακτηριστικών του πεδίου χρόνου, και ξεχωριστά μόνο των φασματικών, και στις δύο περιπτώσεις το Random Forest παρέμεινε στο 96,9%. Και οι δύο οικογένειες χαρακτηριστικών είναι σε μεγάλο βαθμό πλεονάζουσες, αφού η καθεμία από μόνη της φτάνει σχεδόν στην ίδια ακρίβεια με το πλήρες σύνολο των 250 χαρακτηριστικών. Παρακάτω, στους πίνακες 5.3, 5.4 και 5.5, παρουσιάζονται τα αποτελέσματα του πειράματος συνοπτικά.

Διαμόρφωση (μεταβολή ενός στοιχείου)	Test Acc — RF (%)	Μεταβολή vs baseline
Baseline (πλήρες pipeline)	97,1	—
Filter cutoff = 5 Hz	97,3	+0,2
Filter cutoff = 30 Hz	96,1	−1,0
Filter cutoff = 50 Hz	96,9	−0,2
Χωρίς φίλτρο Butterworth	96,2	−0,9
PCA = 5	96,8	−0,3
PCA = 15	97,2	+0,1
PCA = 20	97,2	+0,1
Παράθυρο = 50	96,4	−0,7
Παράθυρο = 150	97,4	+0,3
Βήμα = 25	96,9	−0,2
Βήμα = 75	96,8	−0,3
Χωρίς Temporal Diff	97,4	+0,3
Χωρίς Augmentation	97,2	+0,1
Augmentation: μόνο noise	96,9	−0,2

Διαμόρφωση (μεταβολή ενός στοιχείου)	Test Acc — RF (%)	Μεταβολή vs baseline
Augmentation: μόνο shift	96,9	−0,2
Augmentation: μόνο scale	97,2	+0,1
Augmentation: μόνο time_warp	96,9	−0,2
Μόνο χαρακτηριστικά χρόνου	96,9	−0,2
Μόνο φασματικά χαρακτηριστικά	96,9	−0,2
Χωρίς Zero-Crossing Rate	96,9	−0,2
Χωρίς Diff + Augment	97,2	+0,1
Χωρίς Filter + Augment	94,2	−2,9
Χωρίς Diff + Augment + Filter	92,7	−4,4

Πίνακας 5.3. Μελέτη ablation με βάση το Random Forest. Κάθε γραμμή μεταβάλλει ένα στοιχείο του pipeline κρατώντας τα υπόλοιπα σταθερά (ακρίβεια στο αποκλεισμένο σύνολο ελέγχου).

Μοντέλο	Test Acc με φίλτρο (%)	Test Acc χωρίς φίλτρο (%)
Random Forest	97,1	96,2
Extra Trees	97,1	86,8
Gradient Boosting	97,2	49,9
Logistic Regression	96,9	66,2
SVM (RBF)	96,8	49,8
k-NN	93,44	49,8

Πίνακας 5.4. Επίδραση της αφαίρεσης του φίλτρου Butterworth ανά μοντέλο. Χωρίς το χαμηλοπερατό φιλτράρισμα, μόνο το Random Forest παραμένει αξιόπιστο.

Στον παρακάτω πίνακα (Πίνακας 5.5) παρουσιάζονται συνολικά όλα τα αποτελέσματα όλων των αλγορίθμων και πώς το καθένα επηρεάστηκε σε κάθε περίπτωση.

Διαμόρφωση	SVM (RBF)	RF	ET	GB	LR	k-NN
Baseline (πλήρες pipeline)	96,8	97,1	97,1	97,2	96,9	93,44
Filter cutoff = 5 Hz	96,9	97,3	97,2	97,7	97,2	90,8
Filter cutoff = 30 Hz	95,0	96,1	96,2	96,5	95,2	94,0
Filter cutoff = 50 Hz	97,4	96,9	96,9	97,5	97,3	95,9
Χωρίς φίλτρο Butterworth	49,8	96,2	86,8	49,9	66,2	49,8
PCA = 5	95,6	96,8	97,1	96,9	96,9	92,8
PCA = 15	97,0	97,2	96,9	97,5	97,3	90,2
PCA = 20	96,9	97,2	96,9	97,5	96,8	88,0
Παράθυρο = 50	95,8	96,4	96,3	96,8	96,6	88,5
Παράθυρο = 150	96,7	97,4	97,2	97,4	96,6	92,1
Βήμα = 25	96,5	96,9	97,0	97,4	97,1	90,5
Βήμα = 75	96,0	96,8	96,9	96,8	96,0	90,2
Χωρίς Temporal Diff	94,1	97,4	97,3	96,8	95,3	92,3
Χωρίς Augmentation	96,6	97,2	97,0	97,2	96,9	92,1
Augmentation: μόνο noise	96,5	96,9	97,1	97,0	96,8	90,0

Διαμόρφωση	SVM (RBF)	RF	ET	GB	LR	k-NN
Augmentation: μόνο shift	96,7	96,9	97,1	97,1	96,9	90,4
Augmentation: μόνο scale	96,7	97,2	97,0	97,1	96,9	89,9
Augmentation: μόνο time_warp	96,7	96,9	96,9	97,1	96,8	91,2
Μόνο χαρακτηριστικά χρόνου	96,5	96,9	97,1	97,0	97,1	90,0
Μόνο φασματικά χαρακτηριστικά	96,6	96,9	97,0	97,4	96,9	91,2
Χωρίς Zero-Crossing Rate	96,5	96,9	97,1	97,3	96,9	90,6
Χωρίς Diff + Augment	94,8	97,2	97,3	97,3	95,7	93,4
Χωρίς Filter + Augment	49,8	94,2	49,8	49,8	52,9	49,9
Χωρίς Diff + Augment + Filter	49,8	92,7	93,1	94,3	61,9	80,5

Πίνακας 5.5. Πλήρης μελέτη ablation με ακρίβεια (%) στο αποκλεισμένο σύνολο ελέγχου ανά μοντέλο, για κάθε μεταβολή ενός στοιχείου του pipeline.

Το Extra Trees είναι ουσιαστικά ο πιο κοντινός συγγενής του Random Forest και συμπεριφέρεται αντίστοιχα σε όλη τη μελέτη. Μένει σταθερά στο 96,3–97,3% σχεδόν παντού, και η μόνη φορά που ξεχωρίζει είναι στην αφαίρεση του φίλτρου, όπου πέφτει στο 86,8%. Δεν καταρρέει δηλαδή όπως τα υπόλοιπα, αλλά ούτε κρατιέται τόσο καλά όσο το Random Forest (96,2%). Είναι δηλαδή το δεύτερο πιο ανθεκτικό μοντέλο, λογικό αφού μοιράζεται την ίδια λογική των τυχαίων δέντρων με ψηφοφορία.

Το Gradient Boosting είναι μια πιο ενδιαφέρουσα περίπτωση. Στον ίδιο χώρο (In-domain) δίνει τις υψηλότερες τιμές από όλα, φτάνοντας 97,5% στα PCA=15 και PCA=20, οπότε όταν το pipeline είναι ολόκληρο είναι σταθερά στην κορυφή. Μόλις όμως αφαιρεθεί το φίλτρο, καταρρέει στο 49,9%, δηλαδή προβλέπει στην τύχη, παρόλο που είναι κι αυτό δενδρικό μοντέλο σαν το Random Forest. Πιθανότατα αυτό θα οφείλεται στο ότι χτίζει διαδοχικά δέντρα πάνω στα σφάλματα των προηγούμενων, οπότε χωρίς το φιλτράρισμα αρχίζει να προσαρμόζεται στον ίδιο τον θόρυβο όπως αναφέρθηκε και πριν. Με άλλα λόγια, η ίδια ιδιότητα που το κάνει ακριβές με καθαρά δεδομένα το κάνει και αρκετά ευάλωτο όταν η είσοδος του είναι θορυβώδης.

Το SVM και το Logistic Regression ακολουθούν μια κοινή λογική. Το SVM μένει στο 96,8% baseline, αλλά πέφτει αισθητά όταν αφαιρεθεί η χρονική διαφοροποίηση (94,1%), κάτι που δείχνει ότι το συγκεκριμένο βήμα βοηθάει κυρίως τα μοντέλα αυτού του τύπου και όχι τα δενδρικά. Το Logistic Regression, αν και απλό γραμμικό μοντέλο, κρατιέται ιδιαίτερα κοντά στα υπόλοιπα (96,9%), πράγμα που δείχνει ότι μετά το PCA οι δύο κλάσεις είναι σχεδόν γραμμικά διαχωρίσιμες. Στην αφαίρεση του φίλτρου και τα δύο υποφέρουν, με το SVM να πέφτει στο 49,8% και το Logistic Regression στο 66,2%. Το μοντέλο δηλαδή κρατάει λίγη πληροφορία ακόμη και μέσα στον θόρυβο, ενώ το SVM χάνει εντελώς το περιθώριο διαχωρισμού.

Το k-NN είναι σταθερά ο πιο αδύναμος από όλους. Ξεκινάει ήδη χαμηλά με βάση (baseline) 93,4%, αρκετά πίσω από τους υπόλοιπους αλγόριθμους, και είναι ιδιαίτερα ευαίσθητος στη διάσταση του χώρου. Στο PCA=20 πέφτει στο 88,0% και στο μικρό παράθυρο των 50 δειγμάτων στο 88,5%, κάτι το οποίο είναι αναμενόμενο αφού ο αλγόριθμος στηρίζεται σε αποστάσεις και δυσκολεύεται όσο αυξάνονται οι διαστάσεις ή μικραίνει το πλαίσιο της κίνησης. Χωρίς φίλτρο φτάνει κι αυτός στο 49,8%, δηλαδή πρόβλεψη στην τύχη. Συνολικά, ο k-NN λειτουργεί, αλλά δείχνει ξεκάθαρα ότι για αυτό το πρόβλημα η απλή λογική της ομοιότητας δεν αρκεί όσο οι υπόλοιποι ταξινομητές.

Παρατηρώντας και τα πέντε αυτά μοντέλα μαζί, μπορεί να εξαχθεί ένα κοινό συμπέρασμα. Με ολόκληρο το pipeline συγκλίνουν όλα γύρω στο 96–97%, εκτός από τον k-NN. Η πραγματική διαφορά τους φαίνεται μόνο όταν αφαιρεθεί κάποιο στάδιο, και κυρίως το φίλτρο, όπου στην περίπτωση αυτή χωρίζονται καθαρά σε όσα αντέχουν τον θόρυβο και σε όσα όχι. Τον καθοριστικό ρόλο, δηλαδή, τον έχει η προεπεξεργασία, και ο εκάστοτε ταξινομητής απλώς αξιοποιεί τα δεδομένα τα οποία του εισάγονται. Το RF είναι το πιο ανθεκτικό μοντέλο διατηρώντας την ακρίβεια του υψηλά, κάνοντάς το ακόμα πιο κατάλληλο για τις πραγματικές συνθήκες ενός περιβάλλοντος.

5.6 Αξιολόγηση Γενίκευσης σε Νέο Περιβάλλον (Cross-Environment)

Όλα τα προηγούμενα αποτελέσματα βρίσκονται εντός του χώρου εκπαίδευσης (in-domain), δηλαδή η εκπαίδευση και ο έλεγχος μοιράζονται τον ίδιο χώρο. Το πιο δύσκολο ερώτημα για ένα σύστημα CSI είναι αν γενικεύει όταν αλλάζει το περιβάλλον, κάτι που αναπτύχθηκε ως ανοιχτό πρόβλημα στο Κεφάλαιο 3. Για τον έλεγχο αυτό, χρησιμοποιήθηκε μια διάταξη Leave-Environment-Group-Out (LEGO) στην οποία η εκπαίδευση γίνεται αποκλειστικά στον έναν χώρο και ο έλεγχος αποκλειστικά στον άλλο. Το dataset περιλαμβάνει δύο χώρους, το room1 με 80 καταγραφές (40 ανά κλάση) και το σαλόνι (livroom) με 20 καταγραφές (10 ανά κλάση).

Όταν η εκπαίδευση γίνεται στο μεγάλο σύνολο room1 και ο έλεγχος στο σαλόνι, τα μοντέλα συμπεριφέρονται σταθερά. Το Random Forest φτάνει 88,3%, το Gradient Boosting 88,0%, το Extra Trees 87,8%, το SVM 87,2% και το Logistic Regression 87,0%, με μόνη εξαίρεση το k-NN που πέφτει στο 75,7%. Αντίστροφα, με εκπαίδευση στο μικρό σύνολο livroom και έλεγχο στο room1, τα αποτελέσματα γίνονται πιο ασταθή, κάτι που φαίνεται και από τις μεγάλες τυπικές αποκλίσεις του cross-validation (της τάξης του $\pm 8-9\%$). Εκεί το Extra Trees δίνει το καλύτερο αποτέλεσμα με 90,3%, ενώ το Logistic Regression κατεβαίνει στο 80,0%. Όμως η αστάθεια αυτή είναι λογική, αφού 20 καταγραφές από έναν μόνο χώρο είναι λίγες για να μάθει το μοντέλο τις απαραίτητες ιδιότητες ώστε μετέπειτα να μπορέσει να τις εφαρμόσει και σε διαφορετικό χώρο. Επομένως, για ακόμα πιο ανθεκτικά αποτελέσματα θα έπρεπε να υπάρχει η ίδια αναλογία των καταγραφών του σαλονιού συγκριτικά με το room1.

Συνολικά, η ακρίβεια σε νέο περιβάλλον κινείται γύρω στο 84–90% για τα ανθεκτικά μοντέλα, έναντι του ~97% που πετυχαίνουν εντός περιβάλλοντος (in-domain). Η πτώση των επτά έως δεκατριών μονάδων, ανάλογα με το μοντέλο και την κατεύθυνση της μεταφοράς, ποσοτικοποιεί ακριβώς το χάσμα γενίκευσης που στα προηγούμενα κεφάλαια περιγραφόταν μόνο ιδανικά. Δεν είναι κατάρρευση, διότι το σύστημα εξακολουθεί να ξεχωρίζει κίνηση από ακινησία σε άγνωστο χώρο, παρόλα αυτά η απόδοση πέφτει αισθητά, όπως άλλωστε αναμενόταν και από τη βιβλιογραφία.[18]

Πείραμα	Μοντέλο	CV (%)	Test Acc (%)	F1 Macro (%)
Εκπαίδευση room1 → Έλεγχος livroom	RF	97,0 ± 1,8	88,3	88,1
Εκπαίδευση room1 → Έλεγχος livroom	GB	97,2 ± 1,4	88,0	87,8
Εκπαίδευση room1 → Έλεγχος livroom	ET	96,6 ± 2,5	87,8	87,7
Εκπαίδευση room1 → Έλεγχος livroom	SVM (RBF)	95,8 ± 2,5	87,2	86,9
Εκπαίδευση room1 → Έλεγχος livroom	LR	97,1 ± 1,7	87,0	86,7
Εκπαίδευση room1 → Έλεγχος livroom	k-NN	92,7 ± 4,0	75,7	73,6
Εκπαίδευση livroom → Έλεγχος room1	ET	81,7 ± 8,8	90,3	90,2
Εκπαίδευση livroom → Έλεγχος room1	k-NN	77,5 ± 7,4	85,2	85,2
Εκπαίδευση livroom → Έλεγχος room1	RF	80,0 ± 9,7	84,0	83,7
Εκπαίδευση livroom → Έλεγχος room1	SVM (RBF)	81,3 ± 9,3	83,0	82,7
Εκπαίδευση livroom → Έλεγχος room1	GB	80,2 ± 8,6	82,3	81,9
Εκπαίδευση livroom → Έλεγχος room1	LR	77,2 ± 5,6	80,0	79,6

Πίνακας 5.6. Αξιολόγηση cross-environment (LEGO). Εκπαίδευση σε έναν χώρο, έλεγχος στον άλλο.

Η αντίστοιχη αξιολόγηση cross-subject (Leave-One-Subject-Out) δεν ήταν εφικτή, καθώς όλες οι καταγραφές προέρχονται από έναν μόνο συμμετέχοντα. Επομένως, απαιτεί δεδομένα από τουλάχιστον δύο διαφορετικά άτομα, κάτι που καταγράφεται στους περιορισμούς και στις κατευθύνσεις μελλοντικής έρευνας του Κεφαλαίου 6. Παρόλα αυτά, κατά την real time λειτουργία του συστήματος της παρούσας εργασίας παρατηρήθηκε ότι το σύστημα μπορούσε να λειτουργήσει και να προβλέψει σωστά, με την εκτέλεση από διαφορετικά άτομα (2 στο σύνολο, ένα κάθε φορά) από το άτομο στο οποίο εκπαιδεύτηκε.

5.7 Υπολογιστική Απόδοση και Συμπεριφορά Πραγματικού Χρόνου

Όλη η προηγούμενη υποενότητα αφορούσε την ακρίβεια ταξινόμησης. Για ένα σύστημα όμως που προορίζεται να τρέχει σε πραγματικό χρόνο, ωστόσο, η ακρίβεια από μόνη της δεν αρκεί. Χρειάζεται και κατά πόσο το σύστημα προλαβαίνει να επεξεργαστεί ένα παράθυρο πριν φτάσει το επόμενο, και κατά πόσο «βαρύ» είναι το κάθε μοντέλο όταν αποθηκευτεί. Για τον σκοπό αυτό μετρήθηκε η υπολογιστική απόδοση και των έξι ταξινομητών, με τα αποτελέσματα να συνοψίζονται στον Πίνακα 5.7 και στην εικόνα 5.8.

Η μέτρηση έγινε σε φορητό υπολογιστή (AMD Ryzen 7, Windows 10, Python 3.11, scikit-learn 1.8), με 20 επαναλήψεις προθέρμανσης πριν από κάθε μέτρηση και 200 επαναλήψεις χρονομέτρησης ανά μοντέλο. Ο χρόνος πρόβλεψης (inference) απομονώνει αποκλειστικά την κλήση του predict, ενώ ο συνολικός χρόνος ανά παράθυρο (E2E) περιλαμβάνει και την εξαγωγή χαρακτηριστικών. Το διαθέσιμο περιθώριο πραγματικού χρόνου ορίζεται από το sliding window, δηλαδή 50 frames στα 100 Hz, που αντιστοιχεί περίπου σε 500 ms πριν ξεκινήσει το επόμενο παράθυρο. Η στήλη «Χρήση real-time» δείχνει πόσο από αυτή την προθεσμία καταναλώνει το κάθε μοντέλο στη χειρότερη περίπτωση (p95).

Αλγόριθμος	Χρόνος εκπαίδευσης Train (s)	Μέγεθος μοντέλου (KB)	Μέσος χρόνος πρόβλεψης (ms)	Χειρότερος χρόνος πρόβλεψης (ms)	Συνολικός χρόνος ανά παράθυρο (ms)	Χρήση real-time (p95)
Gradient Boosting	106,91	263,76	0,40	0,57	13,96	2,79
Logistic Regression	0,23	9,06	0,34	0,54	15,70	3,14
SVM (RBF)	11,40	6.286,86	1,13	1,66	16,09	3,22
k-NN	0,03	39.204,97	8,29	9,97	24,80	4,96
Random Forest	0,79	3.669,71	33,78	48,74	66,04	13,21
Extra Trees	0,32	2.632,20	38,58	51,18	72,87	14,57

Πίνακας 5.7. Υπολογιστική απόδοση των έξι ταξινομητών.

Η πρώτη παρατήρηση αφορά το περιθώριο. Παρατηρείται ότι ακόμη και ο πιο αργός ταξινομητής, το Extra Trees, χρειάζεται στη χειρότερη περίπτωση 72,87 ms, δηλαδή το 14,57% της διαθέσιμης προθεσμίας. Αυτό σημαίνει ότι το σύστημα προλαβαίνει το επόμενο παράθυρο σχεδόν επτά φορές, ενώ τα ελαφρύτερα μοντέλα όπως ο Gradient Boosting μένουν κάτω από 14 ms και αφήνουν περιθώριο πάνω από 35 φορές. Ο ισχυρισμός για λειτουργία σε πραγματικό χρόνο, που μέχρι τώρα στηριζόταν στη θεωρητική σύγκριση classical ML έναντι Deep Learning της ενότητας 2.6.3 [5], επιβεβαιώνεται έτσι και πάνω στα δικά μας μετρήσιμα νούμερα.

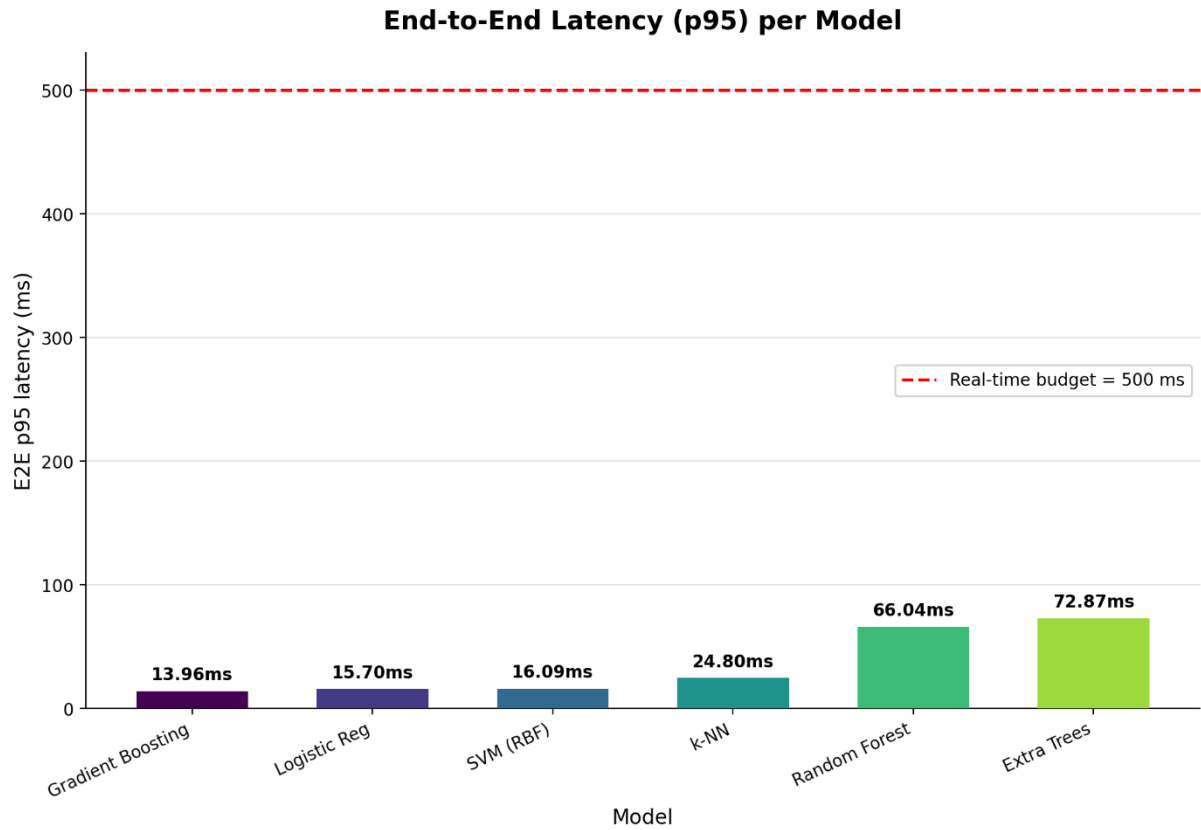
Η κατάταξη επίσης ως προς τον χρόνο πρόβλεψης, έρχεται αντίθετη με τη διαίσθηση που αφήνουν τα προηγούμενα αποτελέσματα. Τα δενδρικά ensembles, δηλαδή ο Random Forest και το Extra Trees, που ήταν από τα πιο ανθεκτικά μοντέλα στη μελέτη ablation, εμφανίζονται εδώ ως τα πιο αργά, με 33,78 ms και 38,58 ms αντίστοιχα. Πιο συγκεκριμένα, κάθε πρόβλεψη πρέπει να διασχίσει εκατοντάδες πλήρως αναπτυγμένα δέντρα και να συγκεντρώσει τις ψήφους τους. Αντίθετα, ο Gradient Boosting, που βγήκε πρώτος σε ακρίβεια, κάνει την ταχύτερη πρόβλεψη από όλα (0,40 ms). Το Logistic Regression επίσης είναι εξίσου γρήγορο (0,34 ms), κάτι αναμενόμενο για ένα γραμμικό μοντέλο που στην ουσία εκτελεί ένα γινόμενο διανυσμάτων.

Ο Gradient Boosting όμως κρύβει ένα μειονέκτημα (trade-off). Χρησιμοποιεί 106,91 δευτερόλεπτα στην εκπαίδευση, μακράν τον υψηλότερο χρόνο από όλους, αφού τα δέντρα χτίζονται διαδοχικά το ένα πάνω στο σφάλμα του προηγούμενου και η διαδικασία δεν εκτελείται παράλληλα όπως στον Random Forest. Πρόκειται όμως για την εκτέλεση του μία φορά εκτός σύνδεσης (offline). Οπότε για ένα σύστημα που εκπαιδεύεται μία φορά και στη συνέχεια τρέχει χωρίς κανένα πρόβλημα, θεωρείται κατάλληλο, μιας και αυτό που μετράει στην καθημερινή λειτουργία είναι ο χρόνος πρόβλεψης και όχι ο χρόνος εκπαίδευσης.

Η εικόνα αλλάζει ξανά στο μέγεθος του αποθηκευμένου μοντέλου, και εδώ προκύπτει το πιο χρήσιμο σημείο. Το Logistic Regression καταλαμβάνει 9,06 KB, ενώ ο Gradient Boosting μένει επίσης χαμηλά στα 263,76 KB. Από την άλλη, ο SVM φτάνει τα 6,29 MB επειδή αποθηκεύει τα support vectors, και ο k-NN στα 39,2 MB. Ο k-NN δεν χτίζει κανονικό μοντέλο αλλά κρατάει ολόκληρο το σύνολο εκπαίδευσης ως «μοντέλο», και γι' αυτό άλλωστε έχει και μηδενικό χρόνο εκπαίδευσης (0,03 s). Τέλος, το Random Forest και το Extra Trees βρίσκονται ενδιάμεσα, στα 3,67 MB και 2,63 MB.

Η διάσταση του μεγέθους αποκτά σημασία για την μελλοντική μεταφορά του μοντέλου απευθείας στον μικροελεγκτή (TinyML), που αναφέρεται στις κατευθύνσεις του Κεφαλαίου 6. Ο ESP32-C6 διαθέτει περίπου 512 KB SRAM και 4 MB flash. Λόγω του περιορισμού αυτού, μόνο το Logistic Regression και ο Gradient Boosting χωράνε άνετα, το Random Forest και το Extra Trees θα ήταν οριακά, ενώ ο SVM και κυρίως ο k-NN αποκλείονται εντελώς. Έτσι, ο Random Forest που αναδείχθηκε ως ο πιο ανθεκτικός ταξινομητής στη μελέτη ablation δεν είναι απαραίτητα και η καλύτερη επιλογή αν χρειαστεί να ενσωματωθεί μέσα στον μικροελεγκτή.

Ξεχωριστή αναφορά αξίζει ο k-NN, ο οποίος επιβεβαιώνει και από αυτή την πλευρά γιατί μένει πίσω. Πέρα από τη χαμηλότερη ακρίβεια που καταγράφηκε στις προηγούμενες ενότητες (93,6%), είναι ταυτόχρονα και ο χειρότερος για ανάπτυξη σε πραγματικό σύστημα, με 39,2 MB μέγεθος και 8,29 ms inference, καθώς κάθε πρόβλεψη απαιτεί υπολογισμό αποστάσεων από όλα τα αποθηκευμένα δείγματα. Η αδυναμία του δηλαδή δεν είναι μόνο στατιστική αλλά και υπολογιστική.



Εικόνα 5.8. Συνολικός χρόνος ανά παράθυρο (p95) ανά ταξινομητή.

Ένα τελευταίο σημείο αφορά το πού βρίσκεται πραγματικά το «βάρος» της προεπεξεργασίας. Η εξαγωγή χαρακτηριστικών κοστίζει περίπου 12,70 ms κατά μέσο όρο, για όλα τα μοντέλα. Για τους ταχύτερους ταξινομητές (Gradient Boosting, Logistic Regression, SVM) ο χρόνος αυτός είναι μεγαλύτερος από το ίδιο το inference, που σημαίνει ότι το πραγματικό πρόβλημα δεν είναι ο ταξινομητής αλλά το στάδιο φιλτραρίσματος και υπολογισμού των 250 χαρακτηριστικών. Μόνο στα δενδρικά μοντέλα το inference ξεπερνά την εξαγωγή χαρακτηριστικών. Πιο συγκεκριμένα, αν κάποιος ήθελε να επιταχύνει το σύστημα, για τα ελαφριά μοντέλα θα έπρεπε να δουλέψει πάνω στην προεπεξεργασία εξαγωγής και όχι στην επιλογή αλγορίθμου. Τέλος, το E2E των ελαφρών μοντέλων

(περίπου 14 με 16 ms) συμφωνεί με την καθυστέρηση των 16 ms που αναφέρθηκε για το dashboard στην ενότητα 4.9, κάτι που επιβεβαιώνει ότι η μέτρηση αποτυπώνει ορθά την πραγματική συμπεριφορά του συστήματος.

Συνοψίζοντας, η ανάλυση δείχνει ότι και οι έξι ταξινομητές ικανοποιούν με μεγάλη άνεση την απαίτηση πραγματικού χρόνου, αλλά διαφέρουν σε μέγεθος και σε κόστος πρόβλεψης. Ο Gradient Boosting ξεχωρίζει ως η πιο ισορροπημένη επιλογή, συνδυάζοντας κορυφαία ακρίβεια με ταχύτατη πρόβλεψη, ενώ ο Random Forest παραμένει προτιμότερος στην περίπτωση όπου προτεραιότητα έχει η ανθεκτικότητα και όχι το ελάχιστο μέγεθος.

5.8 Συζήτηση Αποτελεσμάτων

Συνοψίζοντας, δοκιμάστηκαν έξι αλγόριθμοι με StratifiedGroupKFold cross-validation ($k=5$), όπου το Group εξασφαλίζει ότι δεν υπάρχει data leakage μεταξύ train και test, το CV εκτελέστηκε εντός του συνόλου εκπαίδευσης για την επιλογή μοντέλων, ενώ όλες οι αναφερόμενες μετρικές υπολογίζονται στο αποκλεισμένο σύνολο ελέγχου (20% των καταγραφών). Πέντε από τους έξι αλγορίθμους κυμαίνονται μεταξύ 96,8% και 97,2%, με τον Gradient Boosting να βγαίνει πρώτος (97,2%) και τον SVM τελευταίος αυτής της ομάδας (96,8%), διαφορά δηλαδή μόλις 0,4%. Ο k-NN είναι σαφώς ξεχωριστός από τα υπόλοιπα, με ακρίβεια 93,6% και 141 συνολικά λάθη όταν τα άλλα κυμαίνονται μεταξύ 62 και 70. Από την ανάλυση της σημαντικότητας των χαρακτηριστικών φαίνεται ότι το waveform_length κυριαρχεί για την κλάση walk σε όλα τα δενδρικά μοντέλα, αφού μετράει ακριβώς πόσο έντονα μεταβάλλεται το σήμα στον χρόνο. Για την κλάση idle, από την άλλη, δεν υπάρχει ένα τέτοιο κυρίαρχο χαρακτηριστικό και τα μοντέλα συνδυάζουν πολλά features παράλληλα. Οι πίνακες σύγκρισης επιβεβαιώνουν ότι τα δενδρικά μοντέλα κάνουν περισσότερα FP (False Positive) από FN (False Negative) επειδή το βάδισμα έχει το waveform_length ως ξεκάθαρο χαρακτηριστικό γνώρισμα, ενώ στην ηρεμία δεν έχει αντίστοιχο, οπότε εκεί τα μοντέλα αστοχούν πιο εύκολα. Αντίθετα ο SVM δεν εκμεταλλεύεται αυτές τις αλληλεπιδράσεις, οπότε τα λάθη του βγαίνουν σχεδόν ισόποσα (36 FP, 34 FN).

Συνολικά, το σύστημα αναγνώρισης δραστηριότητας επιτυγχάνει ακρίβεια έως 97,2% με τον Gradient Boosting, ενώ επίσης οι υπόλοιποι αλγόριθμοι κυμαίνονται από 96,8%–97,2%, γεγονός που υποδεικνύει ότι το πρόβλημα μπορεί να διαχωριστεί καλά με βάση τα χαρακτηριστικά χωρίς να είναι καθοριστική η επιλογή συγκεκριμένου αλγορίθμου. Τέλος, με βάση τους αλγορίθμους και το πραγματικό περιβάλλον όπου δοκιμάστηκε το σύστημα προκύπτουν ορισμένα συμπεράσματα. Το σύστημα λειτούργησε και σε συνθήκες NLoS (πίσω από τοίχους), με αξιόπιστη απόδοση κυρίως σε μικρές αποστάσεις (εντός ~ 1 m), όπως αναλύεται παρακάτω. Συνολικά, οι καμπύλες ROC επιβεβαιώνουν την εικόνα που είχε ήδη διαμορφωθεί από τις μετρικές ακρίβειας. Πέντε από τους έξι ταξινομητές κινούνται σε AUC της τάξης του 0,990–0,992, με τον k-NN να υπολείπεται οριακά στο 0,970. Η σύγκλιση τόσο διαφορετικών αλγορίθμων στο ίδιο επίπεδο απόδοσης υποδεικνύει ότι το όριο τίθεται από τα ίδια τα δεδομένα και την επιλογή των χαρακτηριστικών, και όχι από την επιλογή του μοντέλου.

Συνοπτικά στα πειράματα και τους ελέγχους που πραγματοποιήθηκαν στο χώρο του σπιτιού προκύπτουν ορισμένες παρατηρήσεις. Για παράδειγμα, όταν ο άνθρωπος βρίσκεται σε απόσταση εντός ενός μέτρου ακόμα και σε NLOS γύρω από τον μικροελεγκτή καθώς και στην ευθεία επικοινωνίας (LoS) το σύστημα αντιλαμβάνεται την κίνηση. Επιπλέον, δοκιμάστηκε σε μέγιστη απόσταση 14 μέτρων όπου και στην περίπτωση αυτή παρατηρήθηκαν τα ίδια ακριβώς αποτελέσματα με ίσως μερικές καθυστερήσεις ενός δευτερολέπτου όπως αναφέρθηκε και στο κεφάλαιο 4. Παράλληλα, δοκιμάστηκε

με διαφορετικά άτομα και στα δύο σενάρια, χωρίς να παρατηρηθεί αισθητή διαφορά στη συμπεριφορά του συστήματος, καθώς και με πολλαπλά άτομα μέσα σε ένα δωμάτιο έως 3 όπου το σήμα ήταν πιο ασταθές. Παρόλα αυτά υπήρχε και κάποια κίνηση η οποία ανιχνευόταν σε σπάνιες περιπτώσεις όπου το σύστημα την ταξινομούσε λάθος για ένα δευτερόλεπτο και επαναφερόταν στην αρχική του πρόβλεψη (idle). Τέλος, η καταγραφή των κινήσεων πραγματοποιήθηκε κυρίως σε ένα δωμάτιο και εξετάστηκε στο ίδιο το δωμάτιο καθώς και χώρους όπως το σαλόνι και άλλα παρόμοια δωμάτια.

Με βάση τα παραπάνω, οι ενδείξεις δίνουν ότι η δυαδική ταξινόμηση λειτουργεί αξιόπιστα και σε χώρους και συνθήκες πέρα από αυτές της συλλογής δεδομένων. Η πλήρης ποσοτική επιβεβαίωση της γενίκευσης σε άγνωστα περιβάλλοντα και άτομα (cross-environment, cross-subject) παραμένει ωστόσο αντικείμενο μελλοντικής αξιολόγησης, όπως σημειώνεται και στους περιορισμούς της εργασίας στο κεφάλαιο 6.

Κεφάλαιο 6ο: Συμπεράσματα και Μελλοντική Επέκταση

Το παρόν κεφάλαιο συνοψίζει τα κύρια ευρήματα της εργασίας, ξεκαθαρίζει τους περιορισμούς και προτείνει συγκεκριμένες κατευθύνσεις για μελλοντική έρευνα και επέκταση του συστήματος.

6.1 Σύνοψη της Εργασίας

Η εργασία αντιμετώπισε το πρόβλημα ανίχνευσης ανθρώπινης κίνησης σε κλειστό χώρο, χρησιμοποιώντας αποκλειστικά την πληροφορία Channel State Information (CSI) του Wi-Fi σήματος, χωρίς κάμερες, ραντάρ ή φορέσιμες συσκευές. Το σύστημα υλοποιήθηκε με την αρχιτεκτονική SISO, με δύο πλακέτες Seeed Studio XIAO ESP32-C6 κόστους 20€. Η επιλογή αυτή δείχνει ότι η ανίχνευση κίνησης δεν απαιτεί εξειδικευμένο ή ακριβό εξοπλισμό. [1][19][24]

Η δομή της εργασίας καλύπτει συνολικά έξι κεφάλαια. Το Κεφάλαιο 1 διατύπωσε το κίνητρο, τους στόχους στο πλαίσιο εφαρμογών έξυπνης κατοικίας, ασφάλειας και υγείας. Το Κεφάλαιο 2 παρείχε το θεωρητικό υπόβαθρο, για τις τεχνολογίες CSI, OFDM, φαινόμενο Doppler κτλ. Το Κεφάλαιο 3 έκανε μια βιβλιογραφική επισκόπηση, σχετικά με το τι έχει επιτευχθεί στο χώρο του Wi-Fi HAR [3][5]. Το Κεφάλαιο 4 περιέγραψε αναλυτικά τον σχεδιασμό και την υλοποίηση, από το hardware (ESP32-C6FH4 SoC, firmware FreeRTOS 100 Hz), την προεπεξεργασία σήματος (Hampel → Butterworth → PCA) και την εξαγωγή 250 χαρακτηριστικών. Τέλος, το Κεφάλαιο 5 παρουσίασε τα αποτελέσματα για τους έξι αλγορίθμους και τη συγκριτική τους ανάλυση [10][11].

6.2 Κύρια Συμπεράσματα

Το κεντρικό συμπέρασμα της εργασίας είναι ότι η δυαδική ανίχνευση Walk/Idle είναι εφικτή με αξιόπιστο τρόπο, χρησιμοποιώντας μόνο δύο ESP32-C6. Το χαμηλό κόστος του υλικού κάνει το σύστημα αξιοποιήσιμο για πρακτικές εφαρμογές ασφάλειας και διαχείρισης ενέργειας σε έξυπνα κτίρια [1][5][19]. Επίσης το φίλτρο Hampel σε συνδυασμό με το Butterworth και το PCA (10 συνιστώσες) λειτούργησε καλά για τον καθαρισμό και τη συμπίεση των 114 υποφορέων CSI [13]. Τέλος από τους έξι αλγορίθμους, οι Gradient Boosting, Random Forest, και Extra Trees υπερτερούν οριακά έναντι του γραμμικού Logistic Regression, του SVM με πυρήνα RBF και του instance-based k-NN, με τη διαφορά μεταξύ των πέντε πρώτων να περιορίζεται στο 0,4%. Επιπλέον η ανάλυση σημαντικότητας των χαρακτηριστικών υποδεικνύει την σημασία τους στους αλγορίθμους μηχανικής μάθησης [5][10].

6.3 Περιορισμοί

Ένας κρίσιμος περιορισμός είναι η ίδια η διάταξη 1Tx 1Rx που παρέχει μόνο ένα κανάλι CSI χωρίς χωρική πληροφορία και χωρίς γωνία άφιξης, πράγμα που περιορίζει τον εντοπισμό θέσης του ανθρώπου καθώς και τη διάκριση πολλαπλών ατόμων και των λεπτομερών κινήσεων όπως sit, fall, fast walking, stand up κτλ. Συγκεκριμένα για την δυαδική ταξινόμηση παρουσίας το ένα κανάλι αποδεικνύεται επαρκές [1][2]. Επιπλέον, τα δεδομένα (dataset) περιορίζονται σε 50 καταγραφές ανά κλάση. Παρά την επαύξησή τους ($\times 5$) που αυξάνει τον αριθμό εκπαιδευτικών παραθύρων, η ποικιλομορφία πραγματικών δεδομένων παραμένει περιορισμένη και δεν την αντικαθιστά σε καμία περίπτωση. Εξίσου σημαντικό είναι ότι όλα τα ποσοτικά αποτελέσματα προέρχονται από ένα άτομο και κυρίως από έναν χώρο (με ένα μικρό υποσύνολο καταγραφών σαλονιού)· η ακρίβεια της τάξης του 97% αποτελεί επομένως βασική (in-domain) επίδοση, ενώ η αποφυγή data leakage μέσω της στρατηγικής StratifiedGroupKfold εξασφαλίζει ότι η εν λόγω επίδοση δεν είναι τεχνητά διογκωμένη.

6.4 Κατευθύνσεις Μελλοντικής Έρευνας

Η παρούσα εργασία μπορεί να βελτιωθεί και να επεκταθεί με αρκετούς τρόπους. Ο πιο άμεσος είναι η επέκταση του dataset με πολλαπλά άτομα και χώρους [13][17]. Επιπλέον, η πολλαπλή αναγνώριση κίνησης (Walk, Run, Sit, Stand, Fall) και ιδιαίτερα η ανίχνευση πτώσης, που έχει άμεση αξία σε εφαρμογές ασφάλειας ηλικιωμένων με την προϋπόθεση να χρησιμοποιηθούν συστήματα MIMO [1][5]. Ταυτόχρονα, η μεταφορά του μοντέλου απευθείας στον ESP32-C6 (TinyML) θα μπορούσε να επιτρέψει την αυτόνομη λειτουργία του μικροελεγκτή χωρίς χρήση εξωτερικού υπολογιστή. Όσον αφορά την ιδιωτικότητα, οι τεχνικές όπως adversarial learning και differentially private μετασχηματισμοί μπορούν να διασφαλίσουν ότι το σύστημα ανιχνεύει κίνηση χωρίς να αποκαλύπτει πληροφορία ταυτότητας ή τοποθεσίας [6].

Στα επόμενα χρόνια αναμένεται η εξέλιξη της τεχνολογίας του Wi-Fi CSI η οποία θα συνδέεται με μια καινούργια τεχνολογία που πρόκειται να ενσωματωθεί με το 6G, το Integrated Sensing and Communication (ISAC) [27][29]. Στο ISAC τα ίδια σήματα θα χρησιμοποιούνται ταυτόχρονα και για επικοινωνία και για ανίχνευση, χωρίς πρόσθετο hardware. Οι νέες δυνατότητες του 6G αναμένεται να φέρουν βελτιώσεις τόσο στο υλικό (Massive MIMO) όσο και στο λογισμικό (ενσωματωμένη AI). Για την ανίχνευση ανθρώπινης δραστηριότητας, αυτό σημαίνει μεγαλύτερο εύρος ζώνης (sub-THz, mmWave) και περισσότερους CSI υποφορείς μέσω massive MIMO, σε επίπεδο λεπτομέρειας που σήμερα απαιτεί εξειδικευμένο radar [27][29]. Συνοψίζοντας, αυτό θα μετατοπίσει το Wi-Fi CSI HAR από εξειδικευμένο εξοπλισμό προς μια υποδομή που υπάρχει ήδη [27][29].

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] J. Liu, H. Liu, Y. Chen, Y. Wang, and C. Wang, “Wireless Sensing for Human Activity: A Survey,” *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1629–1645, 2020, doi: 10.1109/COMST.2019.2934489.
- [2] Z. Wang, K. Jiang, Y. Hou, W. Dou, C. Zhang, Z. Huang, and Y. Guo, “A Survey on Human Behavior Recognition Using Channel State Information,” *IEEE Access*, vol. 7, pp. 155986–156024, 2019, doi: 10.1109/ACCESS.2019.2949123.
- [3] Y. Ma, G. Zhou, and S. Wang, “WiFi Sensing with Channel State Information: A Survey,” *ACM Computing Surveys*, vol. 52, no. 3, art. 46, 2019, doi: 10.1145/3310194.
- [4] I. Ahmad, A. Ullah, and W. Choi, “WiFi-Based Human Sensing With Deep Learning: Recent Advances, Challenges, and Opportunities,” *IEEE Open Journal of the Communications Society*, vol. 5, pp. 3595–3623, 2024, doi: 10.1109/OJCOMS.2024.3411529.
- [5] M. Ali, P. Hendriks, N. Popping, S. Levi, and A. Naveed, “A Comparison of Machine Learning Algorithms for Wi-Fi Sensing Using CSI Data,” *Electronics*, vol. 12, no. 18, art. 3935, 2023, doi: 10.3390/electronics12183935.
- [6] L. Zhang, W. Cui, B. Li, Z. Chen, M. Wu, and T. S. Gee, “Privacy-Preserving Cross-Environment Human Activity Recognition,” *IEEE Transactions on Cybernetics*, vol. 53, no. 3, pp. 1765–1775, 2023, doi: 10.1109/TCYB.2021.3126831.
- [7] D. Varga, “Critical Analysis of Data Leakage in WiFi CSI-Based Human Action Recognition Using CNNs,” *Sensors*, vol. 24, no. 10, art. 3159, 2024, doi: 10.3390/s24103159.
- [8] M. Atif, B. Yoo, and H. Ko, “Device Free Wireless Sensing based Human Activity Recognition Using Commercial Off-the-Shelf IoT Single-Board Computers,” in *Proc. IEEE Int. Conf. Consumer Electronics (ICCE)*, 2024, doi: 10.1109/ICCE59016.2024.10444143.
- [9] M. T. Abdurrahman and A. Irawan, “Stable Human Activity Recognition Using ESP32-Collected CSI Data with LSTM-Attention,” in *Proc. IEEE Int. Conf. Intelligent Computing (ICIC)*, 2025, doi: 10.1109/ICIC68054.2025.11309497.
- [10] S. A. Deepthi, K. Deepthi, K. Prakash, K. L. Samhitha, and B. Anoohya, “WiFi-CSI based Human Activity Recognition using ESP32,” in *Proc. IEEE AESPC*, 2025, doi: 10.1109/AESPC67542.2025.11326701.
- [11] F. Luo, S. Khan, B. Jiang, and K. Wu, “Vision Transformers for Human Activity Recognition Using WiFi Channel State Information,” *IEEE Internet of Things Journal*, vol. 11, no. 17, pp. 28111–28122, 2024, doi: 10.1109/JIOT.2024.3375337.

- [12] Z. He, X. Zhang, Y. Wang, Y. Lin, G. Gui, and H. Gacanin, “A Robust CSI-Based Wi-Fi Passive Sensing Method Using Attention Mechanism Deep Learning,” *IEEE Internet of Things Journal*, vol. 10, no. 19, pp. 17490–17499, 2023, doi: 10.1109/JIOT.2023.3275545.
- [13] Z. Zhou, Z. Liu, Y. Liu, Y. Zhao, J. Wang, B. Zhang, Y. Xia, X. Zhang, and S. Li, “TCS-Fall: Cross-individual fall detection system based on channel state information and time-continuous stack method,” *Digital Health*, vol. 10, 2024, doi: 10.1177/20552076241259047.
- [14] G. Mattela, M. Tripathi, and C. Pal, “A Novel Approach in WiFi CSI-Based Fall Detection,” *SN Computer Science*, vol. 3, art. 214, 2022, doi: 10.1007/s42979-022-01111-2.
- [15] Z. Hao, Y. Duan, X. Dang, and T. Zhang, “CSI-HC: A WiFi-Based Indoor Complex Human Motion Recognition Method,” *Mobile Information Systems*, vol. 2020, art. 3185416, 2020, doi: 10.1155/2020/3185416.
- [16] J. Guo and I. W. H. Ho, “CSI-Based Efficient Self-Quarantine Monitoring System Using Branchy Convolution Neural Network,” in *Proc. IEEE 8th World Forum on Internet of Things (WF-IoT)*, 2022, pp. 1–6, doi: 10.1109/WF-IoT54382.2022.10152190.
- [17] M. Atif, S. Muralidharan, H. Ko, and B. Yoo, “Wi-ESP: A tool for CSI-based Device-Free Wi-Fi Sensing (DFWS),” *Journal of Computational Design and Engineering*, vol. 7, no. 5, pp. 644–656, 2020, doi: 10.1093/jcde/qwaa048.
- [18] J. Zhang, Z. Tang, M. Li, D. Fang, P. Nurmi, and Z. Wang, “CrossSense: Towards Cross-Site and Large-Scale WiFi Sensing,” in *Proc. ACM MobiCom*, New Delhi, India, 2018, pp. 305–320, doi: 10.1145/3241539.3241570.
- [19] J. Niu, X. He, B. Fang, G. Han, X. Wang, and J. He, “CSI-F: A Human Motion Recognition Method Based on Channel-State-Information Signal Feature Fusion,” *Sensors*, vol. 24, no. 3, art. 862, 2024, doi: 10.3390/s24030862.
- [20] G. Zhu, B. Wang, W. Gao, Y. Hu, C. Wu, and K. J. R. Liu, “WiFi-Based Robust Human and Non-human Motion Recognition With Deep Learning,” in *Proc. IEEE Int. Conf. Pervasive Computing and Communications Workshops (PerCom Workshops)*, 2024, pp. 769–774, doi: 10.1109/PerComWorkshops59983.2024.10502413.
- [21] A. Zhuravchak, O. Kapshii, and E. Pournaras, “Human Activity Recognition based on Wi-Fi CSI Data. A Deep Neural Network Approach,” *Procedia Computer Science*, vol. 198, pp. 59–66, 2022, doi: 10.1016/j.procs.2021.12.211.
- [22] Z. Wei, W. Chen, S. Ning, W. Lin, and J. Zhao, “A Survey on WiFi-based Human Identification: Scenarios, Challenges, and Current Solutions,” *ACM Transactions on Sensor Networks*, vol. 21, no. 1, 2025, doi: 10.1145/3708323.

- [23] M. Youssef, M. Mah, and A. Agrawala, “Challenges: Device-free Passive Localization for Wireless Environments,” in *Proc. ACM MobiCom*, Montréal, QC, Canada, 2007, pp. 222–229, doi: 10.1145/1287853.1287880.
- [24] L. Gong, C. Xiang, X. Fan, T. Wu, C. Chen, M. Yu, and W. Yang, “Device-free near-field human sensing using WiFi signals,” *Personal and Ubiquitous Computing*, vol. 26, pp. 461–474, 2022, doi: 10.1007/s00779-020-01385-4.
- [25] A. Y. Radwan, M. Yildirim, N. Hasanzadeh, H. Tabassum, and S. Valaee, “A Tutorial-cum-Survey on Self-Supervised Learning for Wi-Fi Sensing: Trends, Challenges, and Outlook,” *arXiv:2506.12052*, 2025, doi: 10.48550/arXiv.2506.12052.
- [26] X. Wang, X. Wang, and S. Mao, “RF Sensing in the Internet of Things: A General Deep Learning Framework,” *IEEE Communications Magazine*, vol. 56, no. 9, pp. 62–67, 2018, doi: 10.1109/MCOM.2018.1701277.
- [27] K. Wu, Z. Wang, S.-L. Chen, J. A. Zhang, and Y. J. Guo, “ISAC: From Human to Environmental Sensing,” *IEEE Journal of Selected Topics in Electromagnetics, Antennas and Propagation*, vol. 1, no. 1, pp. 84–98, 2025, doi: 10.1109/JSTEAP.2025.3606464.
- [28] M. G. Moghaddam, A. A. N. Shirehjini, and S. Shirmohammadi, “A WiFi-Based Method for Recognizing Fine-Grained Multiple-Subject Human Activities Using CSI and RSSI Feature Fusion,” *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–13, 2023, doi: 10.1109/TIM.2023.3289547.
- [29] A. Aboulfotouh, E. Mohammed, and H. Abou-Zeid, “6G WavesFM: A Foundation Model for Sensing, Communication, and Localization,” *IEEE Open Journal of the Communications Society*, vol. 6, pp. 6792–6807, 2025, doi: 10.1109/OJCOMS.2025.3600616.
- [30] Espressif Systems, “ESP32-C6 Series Datasheet,” v1.1, 2024. [Online]. Available: https://www.espressif.com/sites/default/files/documentation/esp32-c6_datasheet_en.pdf
- [31] Seeed Studio, “XIAO ESP32-C6 v1.0 Schematic Diagram,” Rev. V1.0, Seeed Technology Co., Ltd., Shenzhen, China, Jan. 2026. [Online]. Available: https://files.seeedstudio.com/wiki/SeeedStudio-XIAO-ESP32C6/XIAO-ESP32-C6_v1.0_SCH_PDF_24028.pdf
- [32] Wi-Fi Alliance, “The state of connectivity: Wi-Fi momentum in 2024,” 2024. [Online]. Available: <https://www.wi-fi.org/beacon/the-beacon/the-state-of-connectivity-wi-fi-momentum-in-2024>
- [33] Y. Wang, K. Wu, and L. M. Ni, “WiFall: Device-Free Fall Detection by Wireless Networks,” *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 581–594, 2017, doi: 10.1109/TMC.2016.2557792.

- [34] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, “E-eyes: Device-free Location-oriented Activity Identification Using Fine-grained WiFi Signatures,” in *Proc. ACM MobiCom*, 2014, pp. 617–628, doi: 10.1145/2639108.2639143.
- [35] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, “Understanding and Modeling of WiFi Signal Based Human Activity Recognition,” in *Proc. ACM MobiCom*, 2015, pp. 65–76, doi: 10.1145/2789168.2790093.

ΠΑΡΑΡΤΗΜΑ Α : Κώδικας υλοποίησης Εξαγωγής CSI

Ορισμός ρυθμού (csi_send/main/app_main.c:31):

```
#define CONFIG_SEND_FREQUENCY 100  
#define SEND_PERIOD_US      (1000000 / CONFIG_SEND_FREQUENCY)
```

Payload με magic header + sequence (csi_send/main/app_main.c:43):

```
typedef struct __attribute__((packed)) {  
    uint32_t magic;  
    uint32_t sequence_id;  
} radar_payload_t;
```

Κύριος βρόχος αποστολής (csi_send/main/app_main.c:269):

```
while (true) {  
    payload.sequence_id = sequence_id++;  
    ret = esp_now_send(peer.peer_addr,  
        (const uint8_t *)&payload,  
        sizeof(payload));  
    /* Deterministic timing — αφαιρεί τον χρόνο εκτέλεσης */  
    next_send_us += SEND_PERIOD_US;  
    int64_t sleep_us = next_send_us - esp_timer_get_time();  
    if (sleep_us > 0) usleep((useconds_t)sleep_us);  
    else next_send_us = esp_timer_get_time(); /* drift reset */  
}
```

Rx — CSI callback → Queue → Print task

Εγκατάσταση CSI callback (csi_recv/main/app_main.c:335):

```
ESP_ERROR_CHECK(esp_wifi_set_csi_rx_cb(wifi_csi_rx_cb, NULL));
```

```
ESP_ERROR_CHECK(esp_wifi_set_csi(true));
```

Callback (producer) — copy & queue push (csi_rcv/main/app_main.c:143):

```
/* Τρέχει σε WiFi task context — πρέπει να είναι ΓΡΗΓΟΡΟ.
```

```
Μόνο: check, copy, queue push. Ποτέ print. */
```

```
static void wifi_csi_rx_cb(void *ctx, wifi_csi_info_t *info) {  
    //... MAC check, magic header parse...  
    csi_queue_item_t item = {.rx_id = rx_id, rx_ctrl = info->rx_ctrl,... };  
    memcpy(item.buf, info->buf, item.len);  
  
    /* Non-blocking push — atomic count αν γεμάτο */  
    if (xQueueSend(s_csi_queue, &item, 0) != pdTRUE)  
        atomic_fetch_add(&s_drop_count, 1);  
}
```

Δημιουργία Queue και task (csi_rcv/main/app_main.c:363):

```
s_csi_queue = xQueueCreate(CSI_QUEUE_SIZE, sizeof(csi_queue_item_t));
```

```
xTaskCreate(csi_print_task, "csi_print", 8192, NULL, 5, NULL);
```

Print task (consumer) (csi_rcv/main/app_main.c:222):

```
/* Εδώ επιτρέπεται ets_printf — δεν μπλοκάρει WiFi context. */
```

```
static void csi_print_task(void *arg) {  
    csi_queue_item_t item;  
    while (true) {  
        if (xQueueReceive(s_csi_queue, &item, portMAX_DELAY) != pdTRUE)  
            continue;  
        //... format & ets_printf(print_buffer)...  
    }  
}
```

ΠΑΡΑΡΤΗΜΑ Β : Κώδικας υλοποίησης Ρυθμού Δειγματοληψίας

csi_send/main/app_main.c:30 — ρυθμός αποστολής:

```
#define CONFIG_SEND_FREQUENCY 100 // 100 πακέτα/δευτ.  
#define SEND_PERIOD_US (1000000 / CONFIG_SEND_FREQUENCY) // 10 ms/πακέτο
```

csi_recv/main/app_main.c:39 — μέγεθος ουράς:

```
#define CSI_QUEUE_SIZE 30
```

csi_recv/main/app_main.c:363 — δημιουργία ουράς:

```
s_csi_queue = xQueueCreate(CSI_QUEUE_SIZE, sizeof(csi_queue_item_t));
```

csi_recv/main/app_main.c:211 — non-blocking push από callback (WiFi context):

```
if (xQueueSend(s_csi_queue, &item, 0) != pdTRUE)  
    atomic_fetch_add(&s_drop_count, 1);
```

csi_recv/main/app_main.c:232 — blocking receive από print task:

```
if (xQueueReceive(s_csi_queue, &item, portMAX_DELAY) != pdTRUE)  
    continue;
```

ΠΑΡΑΡΤΗΜΑ C : Πρωτόκολλο Συλλογής Δεδομένων

config.py:240 — ορισμός κλάσεων, διάρκειας session και αναμονής σταθεροποίησης:

```
LOGGER_ACTIVITY_PRESETS: dict[str, dict] = {  
    "walk": {  
        "output_dir": "datasets/walk_activity",  
        "duration": 60, # 100 Hz × 60 s = 6.000 πλαίσια  
        "wait": 5, # 5 s αναμονή πριν εκκίνηση  
        "label_prefix": "walk_activity",  
    },  
    "idle": {  
        "output_dir": "datasets/no_activity",
```

```

    "duration": 60,
    "wait": 5,
    "label_prefix": "no_activity",
},
}

```

csi_logger.py:207 — αντίστροφη μέτρηση σταθεροποίησης (επιβεβαίωση σύνδεσης):

```

if args.wait > 0:
    print(f"\n[WAIT] Starting in {args.wait} seconds... (Get into position!)")
    for i in range(args.wait, 0, -1):
        print(f" {i}...")
        time.sleep(1)
    print("[START] GO! (Recording started)")

```

csi_logger.py:260 — αυτόματη διακοπή στο τέλος του session:

```

if args.duration > 0 and (now - start_time) >= args.duration:
    print(f"\n[OK] Auto-stop reached ({args.duration} seconds).")
    break

```

ΠΑΡΑΡΤΗΜΑ D : Επαύξηση Δεδομένων (Data Augmentation)

config.py:75 — ορισμός τεχνικών και πλήθους augmentations:

```
AUGMENTATION_TECHNIQUES = ["noise", "shift", "scale", "time_warp"]
```

```
N_AUGMENTS = 4
```

csi_ml_pipeline.py:107 — Gaussian noise (κλιμακούμενο ως προς std σήματος):

```
def _aug_noise(window, rng, class_label=None):
```

```
    signal_std = np.std(window)
```

```
    noise_level = rng.uniform(0.003, 0.01) * (signal_std if signal_std > 1e-6 else 1.0)
```

```
    return window + rng.normal(0, noise_level, window.shape)
```

csi_ml_pipeline.py:124 — temporal shift (padding-based, χωρίς κυκλικές ασυνέχειες):

```
def _aug_shift(window, rng, class_label=None):
```

```
    shift_steps = int(rng.integers(1, 4))
```

```
    direction = rng.choice([-1, 1])
```

```
    if direction == 1:
```

```
        pad = np.repeat(window[0:1], shift_steps, axis=0)
```

```
        return np.vstack([pad, window[:-shift_steps]])
```

```
    else:
```

```
        pad = np.repeat(window[-1:], shift_steps, axis=0)
```

```
        return np.vstack([window[shift_steps:], pad])
```

csi_ml_pipeline.py:142 — magnitude scaling (physics-aware ανά κλάση):

```
def _aug_scale(window, rng, class_label=None):
```

```
    scale = rng.uniform(0.9, 1.1) # ±10% για walk/idle
```

```
    return window * scale
```

csi_ml_pipeline.py:155 — time warp με reflect padding:

```
def _aug_time_warp(window, rng, class_label=None):
```

```
    T = window.shape[0]
```

```
    factor = rng.uniform(0.9, 1.1) # walk: ±10%
```

```
    src_indices = np.linspace(0, T - 1, T) / factor
```

```

# Reflect padding — αποφεύγει flat tails στα όρια παραθύρου
overflow = src_indices > (T - 1)
src_indices[overflow] = 2 * (T - 1) - src_indices[overflow]
src_indices = np.clip(src_indices, 0, T - 1)
warped = np.empty_like(window, dtype=np.float32) # προδέσμευση πίνακα εξόδου
for c in range(window.shape[1]):
    warped[:, c] = np.interp(src_indices, np.arange(T), window[:, c])
return warped.astype(np.float32)

```

csi_ml_pipeline.py:195 — κεντρική συνάρτηση: στοχαστική σύνθεση 1-2 τεχνικών ανά augment:

```

def augment_window(window, n_augments=config.N_AUGMENTS,
                    techniques=None, seed=None, class_label=None):
    rng = np.random.default_rng(seed)
    for _ in range(n_augments):
        n_to_apply = int(rng.integers(1, 3)) # 1-2 τεχνικές ανά augment
        chosen = rng.choice(safe_techs, size=min(n_to_apply, len(safe_techs)), replace=False)
        aug = window.copy()
        for tech in chosen:
            aug = _AUG_FN_MAP[tech](aug, rng, class_label=class_label)
        augmented_windows.append(aug.astype(np.float32))
    return augmented_windows

```

ΠΑΡΑΡΤΗΜΑ Ε : Πλήρης Πηγαίος Κώδικας και Αποθετήριο (GitHub)

Ολόκληρος ο πηγαίος κώδικας που αναπτύχθηκε στα πλαίσια της παρούσας διπλωματικής εργασίας είναι διαθέσιμος ανοιχτά στο ακόλουθο αποθετήριο του GitHub:

<https://github.com/vasilistzanis/Wifi-CSI-Human-Detection>.

Το αποθετήριο περιλαμβάνει, το firmware για τους μικροελεγκτές ESP32-C6 (συλλογή και μετάδοση CSI), τα scripts προεπεξεργασίας και εξαγωγής χαρακτηριστικών σε Python, την υλοποίηση και εκπαίδευση των μοντέλων μηχανικής μάθησης, καθώς και το γραφικό περιβάλλον πραγματικού χρόνου (dashboard).