



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ, ΠΛΗΡΟΦΟΡΙΚΗΣ & ΥΠΟΛΟΓΙΣΤΩΝ

**«Αυτόματη αξιολόγηση δυσκινησίας ασθενών με Πάρκινσον  
με τη χρήση μη επεμβατικών μεθόδων μηχανικής μάθησης»**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

Παρασκευή-Αντωνία Α. Θεοφίλου

A.M. 03116059

**Επιβλέπων :** Στέφανος Κόλλιας

Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2021





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ, ΠΛΗΡΟΦΟΡΙΚΗΣ & ΥΠΟΛΟΓΙΣΤΩΝ

**«Αυτόματη αξιολόγηση δυσκινησίας ασθενών με Πάρκινσον  
με τη χρήση μη επεμβατικών μεθόδων μηχανικής μάθησης»**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

Παρασκευή-Αντωνία Λ. Θεοφίλου

**Επιβλέπων :** Στέφανος Κόλλιας

Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 29<sup>η</sup> Ιουλίου 2021.

.....

Στέφανος Κόλλιας

Καθηγητής Ε.Μ.Π.

.....

Ανδρέας-Γεώργιος Σταφυλοπάτης

Καθηγητής Ε.Μ.Π.

.....

Γεώργιος Στάμου

Αν.Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2021

.....

Παρασκευή-Αντωνία Λ. Θεοφίλου

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

Copyright © Παρασκευή-Αντωνία Λ. Θεοφίλου, 2021

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

## **Ευχαριστίες**

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου, κ. Στέφανο Κόλλια για την ευκαιρία που μου έδωσε να ασχοληθώ με ένα τόσο ενδιαφέρον αντικείμενο, καθώς και για τη συνεισφορά του στην εποικοδομητική εξέλιξη των σπουδών μου. Επίσης, θα ήθελα να ευχαριστήσω θερμά τον υποψήφιο διδάκτορα κ. Γεώργιο Τσατήρη για την πολύτιμη και ουσιαστική βοήθεια που μου παρείχε κατά τη διάρκεια της εκπόνησης της διπλωματικής μου εργασίας. Τέλος, ευχαριστώ την οικογένεια μου για τη υποστήριξή της.

## Περίληψη

Η νόσος του Πάρκινσον είναι μια σοβαρή νευροεκφυλιστική διαταραχή που προκαλεί απώλεια ελέγχου των σωματικών κινήσεων. Τα συμπτώματα της νόσου καλύπτουν ένα ευρύ φάσμα διαφόρων λειτουργιών, κινητικών και διανοητικών, επηρεάζοντας τον καθημερινό τρόπο ζωής των νοσούντων. Η νόσος αυτή δεν επιδέχεται ίαση, ωστόσο μπορεί να περιοριστεί η εξέλιξη της, ανάλογα με τα στάδια της, κυρίως με φαρμακευτική αγωγή και συγκεκριμένα με τη χρήση λεβοντόπας. Γι' αυτόν τον λόγο είναι πολύ σημαντική η έγκαιρη διάγνωση της νόσου και η όσο το δυνατόν καλύτερη προσέγγιση του επιπέδου σοβαρότητας αυτής. Ωστόσο, πολλές φορές, οι ασθενείς δεν αντιμετωπίζουν μονάχα τα συμπτώματα της νόσου, αλλά και τις διαταραχές που προκαλούνται από τη μακροχρόνια χορήγηση της λεβοντόπας. Η ανίχνευση των διαφόρων συμπτωμάτων και διαταραχών βασίζεται στην κλινική εξέταση των ασθενών με την αξιολόγηση τους σύμφωνα με τις κλίμακες UPDRS και UDysRS. Γι' αυτόν τον λόγο, η παρούσα εργασία αποσκοπεί στην ανάπτυξη ενός συστήματος για την αυτόματη αξιολόγηση της δυσκινησίας που προκαλείται από τη λεβοντόπα σύμφωνα με τη μετρική UDysRS και τη χρήση μη επεμβατικών μεθόδων μηχανικής μάθησης. Χρησιμοποιούνται δεδομένα προερχόμενα από βίντεο που με την κατάλληλη επεξεργασία των χαρακτηριστικών τα οποία προκύπτουν από τα καρέ τους, εισέρχονται σε ένα απλό και συνάμα αποδοτικό CNN για να αξιολογηθούν με μεγάλη ακρίβεια βάσει των έξι τιμών της UDysRS.

**Λέξεις κλειδιά:** Νόσος Πάρκινσον, Ενιαία Κλίμακα Αξιολόγησης Δυσκινησιών (UDysRS), Μηχανική Μάθηση, Νευρωνικά Δίκτυα, Συνελκτικά Νευρωνικά Δίκτυα, Αναγνώριση Κίνησης.

## **Abstract**

Parkinson's disease is a serious neurodegenerative disorder that causes loss of control of physical movements. The symptoms of the disease cover a wide range of different motor and mental functions, affecting the daily lifestyle of patients. This disease is not curable, however its development can be limited, depending on its stages, mainly with medication and specifically with the use of levodopa. That is why it is very important to diagnose the disease early and to approach its severity level as best as possible. However, many times, patients not only experience the symptoms of the disease, but also the disorders caused by long-term administration of levodopa. The detection of various symptoms and disorders is based on the clinical examination of patients with their evaluation according to the UPDRS and UDysRS scales. Therefore, the present work aims to develop a system for the automatic assessment of levodopa-induced dyskinesia according to the UDysRS metric and the use of non-invasive machine learning methods. Video data are used and after the proper processing of the features resulting from their frames, enter a simple and at the same time efficient CNN to be evaluated with great accuracy based on the six values of UDysRS.

**Keywords:** Parkinson's Disease, Unified Dyskinesia Rating Scale (UDysRS), Machine Learning, Neural Networks, Convolutional Neural Networks, Action Recognition.

## **Περιεχόμενα**

Ευχαριστίες .....	v
Περίληψη .....	vi
Abstract .....	vii
Περιεχόμενα.....	viii
Λίστα εικόνων.....	x
Λίστα πινάκων .....	xi
Λίστα εξισώσεων .....	xii
Κεφάλαιο 1: Εισαγωγή .....	1
1.1 Ορισμός προβλήματος .....	1
1.2 Δομή εργασίας.....	3
Κεφάλαιο 2: Θεωρητικό υπόβαθρο & Συναφείς μελέτες .....	4
2.1 Τεχνητή Νοημοσύνη – Μηχανική Μάθηση – Τεχνητά Νευρωνικά Δίκτυα.....	4
2.1.1 Τεχνητή Νοημοσύνη.....	4
2.1.2 Μηχανική Μάθηση.....	4
2.1.3 Τεχνητά Νευρωνικά Δίκτυα .....	7
2.2 Αναγνώριση ανθρώπινης δραστηριότητας.....	11
2.2.1 Ορισμός .....	11
2.2.2 Αναπαράσταση της κίνησης.....	12
2.2.3 Εφαρμογές .....	23
2.3 Η Τεχνητή Νοημοσύνη στην ιατρική και υγειονομική περίθαλψη.....	25
2.4 Νόσος του Πάρκινσον.....	26
2.4.1 Η νόσος.....	26
2.4.2 Τα συμπτώματα .....	27



2.4.3 Διάγνωση και θεραπεία .....	29
2.4.4 Η σημασία της έγκαιρης και έγκυρης διάγνωσης .....	34
2.5 Η Νόσος Πάρκινσον στην Επιστήμη των Υπολογιστών .....	35
2.6 Η Νόσος Πάρκινσον στη Μηχανική Μάθηση και την Όραση Υπολογιστών ...	38
2.7 Αυτόματη αξιολόγηση της νόσου με vision based μεθόδους .....	44
Κεφάλαιο 3: Μεθοδολογία & Υλοποίηση .....	46
3.1 Δεδομένα .....	46
3.2 Προεπεξεργασία δεδομένων – Μεθοδολογία .....	47
3.3 Υλοποίηση.....	50
3.4 Εργαλεία.....	53
Κεφάλαιο 4: Πειράματα & Ανάλυση αποτελεσμάτων .....	54
4.1 Πειράματα .....	54
4.1.1 Πειράματα ως προς την αρχιτεκτονική του δικτύου .....	55
4.1.2 Πειράματα για τον έλεγχο απόδοσης του δικτύου .....	59
4.2 Ανάλυση αποτελεσμάτων .....	61
4.2.1 Αποτελέσματα αρχιτεκτονικών .....	61
4.2.2 Αποτελέσματα δοκιμών για τον έλεγχο απόδοσης του δικτύου.....	67
Κεφάλαιο 5: Συμπεράσματα & Μελλοντικές επεκτάσεις .....	75
5.1 Συμπεράσματα .....	75
5.2 Μελλοντικές προσεγγίσεις .....	75
Βιβλιογραφία .....	77

## Λίστα εικόνων

- Εικόνα 1: Νευρώνας του ανθρώπινου εγκεφάλου [8]
- Εικόνα 2: Τεχνητός νευρώνας [9]
- Εικόνα 3: Απεικόνιση τεχνητού νευρωνικού δικτύου [9]
- Εικόνα 4: Απεικόνιση συνελκτικού νευρωνικού δικτύου [9]
- Εικόνα 5: Pose estimation [25]
- Εικόνα 6: Pose estimation με χρήση του Kinect [26]
- Εικόνα 7: Convolution Pose Machines-εντοπισμός αρθρώσεων [28]
- Εικόνα 8: Αρχιτεκτονική CPMs [28]
- Εικόνα 9: Pose detection using CPMs [29]
- Εικόνα 10: Pipeline της μεθόδου OpenPose [30]
- Εικόνα 11: Αρχιτεκτονική των πολυεπίπεδων συνελκτικών δικτύων [30]
- Εικόνα 12: Real-time pose detection using OpenPose [30]
- Εικόνα 13: Στολές για motion capturing με αισθητήρες [31]
- Εικόνα 14: Selective Spatio-Temporal Interest Points σε εικόνα απεικόνισης backhand [45]
- Εικόνα 15: Διαδικασία εξαγωγής σιλουέτας από το φόντο [46]
- Εικόνα 16: Διαδικασία εξαγωγής των SRFs [47]
- Εικόνα 17: Διάγραμμα ποσοστιαίας μείωσης των νευρώνων ντοπαμίνης σε συνάρτηση με την ηλικία σε υγιή άτομα και άτομα με Πάρκινσον [74]
- Εικόνα 18: Sinograms - Παραδείγματα μετασχηματισμού Radon σε καρέ βίντεο
- Εικόνα 19: Αρχιτεκτονική 1ου Μοντέλου
- Εικόνα 20: Αρχιτεκτονική 2ου Μοντέλου
- Εικόνα 21: Αρχιτεκτονική 3ου Μοντέλου
- Εικόνα 22: Διάγραμμα απεικόνισης μετρικών MSE – RMSE του μοντέλου 2
- Εικόνα 23: Διάγραμμα απεικόνισης των μετρικών MSE - RMSE του μοντέλου 3
- Εικόνα 24: Συγκριτικό διάγραμμα απεικόνισης της μετρικής MSE για τα μοντέλα 2 και 3

## *Λίστα πινάκων*

Πίνακας 1: Φύλλο βαθμολογίας κλίμακας MDS-UPDRS [67]

Πίνακας 2: Μοντέλο σταδίων νόσου Πάρκινσον Hoehn και Yahr [68]

Πίνακας 3: Κλίμακα καθημερινών δραστηριοτήτων Schwab και England [69]

Πίνακας 4: Πίνακας βαθμολογίας UDysRS [71]

Πίνακας 5: Αποτελέσματα δοκιμών ατόμου 1 (Mean Squared Errors)

Πίνακας 6: Αποτελέσματα δοκιμών ατόμου 2 (Mean Squared Errors)

Πίνακας 7: Αποτελέσματα δοκιμών ατόμου 4 (Mean Squared Errors)

Πίνακας 8: Αποτελέσματα δοκιμών ατόμου 5 (Mean Squared Errors)

Πίνακας 9: Αποτελέσματα δοκιμών ατόμου 6 (Mean Squared Errors)

Πίνακας 10: Αποτελέσματα δοκιμών ατόμου 7 (Mean Squared Errors)

Πίνακας 11: Αποτελέσματα δοκιμών ατόμου 8 (Mean Squared Errors)

Πίνακας 12: Αποτελέσματα δοκιμών ατόμου 10 (Mean Squared Errors)

Πίνακας 13: Αποτελέσματα δοκιμών ατόμου 11 (Mean Squared Errors)

Πίνακας 14: Μέσοι όροι αποτελεσμάτων δοκιμών για όλα τα άτομα (Average mean squared errors)

## *Λίστα εξισώσεων*

Εξίσωση 1: Υπολογισμός εξόδου τεχνητού νευρώνα [9]

Εξίσωση 2: Απόσταση Mahalanobis [50]

Εξίσωση 3: Μετασχηματισμός Radon [50]

Εξίσωση 4: Μετασχηματισμός Radon με αντικατάσταση  $\rho$ ,  $\theta$  [50]

Εξίσωση 5: Πίνακας Mahalanobis [50]

Εξίσωση 6: Spatio-temporal Radon Footprints [50]

# **Κεφάλαιο 1: Εισαγωγή**

## ***1.1 Ορισμός προβλήματος***

Η νόσος του Πάρκινσον είναι μια σοβαρή προοδευτική εκφυλιστική διαταραχή του νευρικού συστήματος που προκαλεί απώλεια ελέγχου των σωματικών κινήσεων. Θεωρείται η δεύτερη σε σειρά συχνότητας ασθένεια μετά τη νόσο του Αλτσχάιμερ και αναμένεται μελλοντικά, σύμφωνα με μελέτες, να αποτελέσει τη δεύτερη συχνότερη αιτία θανάτου μετά τον καρκίνο. Στα άτομα που νοσούν από Πάρκινσον μία χημική ουσία στον εγκέφαλο, η ντοπαμίνη που συνεισφέρει στη μεταφορά των σημάτων εντός του εγκεφάλου, μειώνεται βαθμιαία. Τα συμπτώματα της νόσου καλύπτουν ένα ευρύ φάσμα διαφόρων λειτουργιών και επηρεάζουν σε μεγάλο βαθμό τον καθημερινό τρόπο ζωής των νοσούντων. Σε αυτά περιλαμβάνονται η αίσθηση κόπωσης, η απώλεια δύναμης, η χαμηλή φωνή, η υποβάθμιση της ποιότητας γραφής, ακόμη και οι ψυχικές διαταραχές. Πολλά αφορούν τη διαταραχή του κινητικού ελέγχου, με πιο συνηθισμένη τη δυσκαμψία στα άκρα, τον τρόμο ηρεμίας, τη μειωμένη ισορροπία, τη βραδυκίνησια, την αστάθεια και τα επεισόδια παγώματος (freezing). Η νόσος αυτή δεν επιδέχεται ίαση, ωστόσο μπορεί να αντιμετωπιστεί ανάλογα με τα στάδια της με φυσική και νοητική άσκηση, φαρμακευτική αγωγή ή χειρουργική επέμβαση, έτσι ώστε να περιοριστεί η εξέλιξη της. Γι' αυτόν τον λόγο είναι πολύ σημαντική η έγκαιρη διάγνωση της νόσου και η όσο το δυνατόν καλύτερη προσέγγιση του επιπέδου σοβαρότητας αυτής. Επίσης, πολλές φορές, η μακροχρόνια χορήγηση στους ασθενείς φαρμάκων, όπως η λεβοντόπα, επιφέρει διαταραχές. Όσον αφορά τη διάγνωση, δεν υπάρχει κάποια ακριβής διαδικασία βασισμένη σε κάποια καθορισμένη αιματολογική ή βιοχημική εξέταση, αλλά αυτή γίνεται μέσω της κλινικής εξέτασης των διαφόρων συμπτωμάτων του ασθενή και της αξιολόγησης της απόδοσης αυτού σε ορισμένα διαγνωστικά τεστ που επικεντρώνονται στην αναγνώριση των διαταραχών του σε ορισμένες δραστηριότητες της καθημερινότητας. Για την αξιολόγηση αυτή και την εξαγωγή συμπερασμάτων σχετικά με τη σοβαρότητα της νόσου που πάσχει ένας ασθενής χρησιμοποιούνται τα κλινικά διαγνωστικά κριτήρια που ορίζουν οι ενοποιημένες κλίμακες αξιολόγησης της νόσου UPDRS (Unified Parkinson's Disease Rating Scale) και UDysRS (Unified Dyskinesia Rating Scale).

Η μηχανική μάθηση αποτελεί έναν ταχύτητα αναπτυσσόμενο τεχνολογικό τομέα που συναντά ιδιαίτερη εξέλιξη και ενδιαφέρον στις μέρες μας. Μέσω της μηχανικής μάθησης υπάρχει η δυνατότητα δημιουργίας μοντέλων που εκπαιδεύονται πάνω σε ένα σύνολο δεδομένων, έτσι ώστε να επιτευχθεί η λειτουργία της ταξινόμησης ή της πρόβλεψης ανάλογα με το εκάστοτε πρόβλημα. Η μηχανική μάθηση και οι εφαρμογές της χρησιμοποιούνται σε μεγάλο βαθμό από διάφορους τομείς της επιστήμης και της αγοράς εργασίας. Ιδιαίτερο ενδιαφέρον παρουσιάζει η χρήση των εφαρμογών μηχανικής μάθησης σε θέματα ιατρικής φύσεως, όπως ασθένειες, με σκοπό την αποτελεσματικότερη, γρηγορότερη και εγκυρότερη διάγνωση και πρόβλεψη τους. Η ανάγκη συνδυασμού της επιστήμης της ιατρικής με αυτής των υπολογιστών απορρέει από το γεγονός ότι ο κόσμος αντιμετωπίζει πολλαπλά προβλήματα υγείας που όλο και αυξάνουν το επίπεδο δυσκολίας και πολυπλοκότητας στη διαχείριση και ίαση τους. Συνεπώς, η σύμπραξη αυτών των δύο επιστημονικών τομέων από τη μία, αποσκοπεί στη διευκόλυνση της αντιμετώπισης ασθενειών, αυξάνοντας τα επίπεδα ταχύτητας και εγκυρότητας της διάγνωσης και πρόβλεψης και από την άλλη, εξυπηρετεί στην πρόοδο της επιστήμης με την ανίχνευση νέων μεθόδων και τεχνικών αντιμετώπισης των διαφόρων ζητημάτων υγείας που προκύπτουν διαρκώς.

Η εργασία αυτή έχει ως σκοπό την ανάπτυξη ενός συστήματος μηχανικής μάθησης, έτσι ώστε για κάθε ασθενή που εξετάζεται υποβαλλόμενος σε συγκεκριμένες δραστηριότητες να γίνεται αυτόματη αξιολόγηση της δυσκινησίας που εμφανίζει δεδομένων των φαρμάκων που λαμβάνει, σύμφωνα με την μετρική UDysRS, η οποία αποτελεί ενδεδειγμένο κριτήριο αξιολόγησης της σοβαρότητας της διαταραχής αυτής στη νόσο του Πάρκινσον. Στο σύνολο δεδομένων που χρησιμοποιούμε για τον σκοπό αυτό, κάθε ασθενής έχει υποβληθεί σε δραστηριότητες που εμφανίζουν περιορισμένη απόκριση κατά τα διάφορα στάδια της νόσου, όπως είναι η κατάποση, η επικοινωνία, η κίνηση των κάτω άκρων και έχει βιντεοσκοπηθεί, ώστε να γίνει σε δεύτερο χρόνο, η αξιολόγηση των διαφόρων κινήσεων. Λαμβάνοντας, λοιπόν, από κάθε καρτέ του βίντεο τις συντεταγμένες των αρθρώσεων του εκάστοτε ανθρώπινου σώματος, τις συσχετίζουμε μεταξύ τους και έτσι στοχεύουμε στη δημιουργία ενός νευρωνικού δικτύου που μπορεί να προβλέψει σε μεγάλο ποσοστό τη μετρική δυσκινησίας του Πάρκινσον. Επομένως, με την παρούσα εργασία επιθυμούμε να παρουσιάσουμε ένα σύστημα μηχανικής μάθησης

που κάνει αυτόματα την αξιολόγηση της δυσκινησίας στους ασθενείς με Πάρκινσον, έχοντας ως κριτήριο την απόδοση αυτών στις σχετικές δραστηριότητες.

## ***1.2 Δομή εργασίας***

Η παρούσα εργασία δομείται ως εξής: το Κεφάλαιο 1 αποτελεί την εισαγωγή, το Κεφάλαιο 2 παρουσιάζει το θεωρητικό υπόβαθρο και τις συναφείς μελέτες, ενώ το Κεφάλαιο 3 αναλύει τη μεθοδολογία που εφαρμόζεται. Στο Κεφάλαιο 4 περιγράφονται τα πειράματα που πραγματοποιήθηκαν σε συνδυασμό με τις αναλύσεις των αποτελεσμάτων τους. Τέλος, στο Κεφάλαιο 5 παρουσιάζονται τα συμπεράσματα που εξήχθησαν από την έρευνα, καθώς και οι μελλοντικές προσεγγίσεις του προβλήματος που διερευνήθηκε.

## **Κεφάλαιο 2: Θεωρητικό υπόβαθρο & Συναφείς μελέτες**

### **2.1 Τεχνητή Νοημοσύνη – Μηχανική Μάθηση – Τεχνητά Νευρωνικά**

#### **Δίκτυα**

##### **2.1.1 Τεχνητή Νοημοσύνη**

Ο όρος «τεχνητή νοημοσύνη» περιγράφει τον κλάδο της επιστήμης υπολογιστών που ασχολείται με τη σχεδίαση και την υλοποίηση προγραμμάτων στοιχειώδους ευφυΐας που είναι ικανά να μιμηθούν την ανθρώπινη γνώση και συμπεριφορά ως προς την επίλυση προβλημάτων, την αντιληπτική ικανότητα μέσω της όρασης, τη μάθηση, την εξαγωγή συμπερασμάτων και την κατανόηση της φυσικής γλώσσας [1]. Στόχος της, λοιπόν, είναι η σύνθεση ευφυούς συμπεριφοράς που απορρέει από τη συλλογιστική, τη μάθηση και την προσαρμοστικότητα στο περιβάλλον [2]. Αποτελεί έναν τομέα που εξελίσσεται και συναντά μεγάλο ενδιαφέρον σε πληθώρα επιστημονικών κλάδων, όπως είναι η ψυχολογία, η φιλοσοφία, η νευρολογία, η γλωσσολογία, η πληροφορική και η επιστήμη των μηχανικών. Διακρίνεται σε δύο κατηγορίες: τη συμβολική τεχνητή νοημοσύνη και την υποσυμβολική. Η πρώτη επικεντρώνεται στην προσομοίωση της ανθρώπινης συμπεριφοράς μέσω της κατανόησης των νοητικών διεργασιών με τη χρήση ειδικών αλγορίθμων και συστημάτων γνώσης. Η δεύτερη βασίζεται στην ανθρώπινη ευφυΐα και μέσα από αριθμητικά μοντέλα προσεγγίζονται βιολογικές διεργασίες, όπως αυτή της λειτουργίας του εγκεφάλου. Τέλος, ανάλογα με το στόχο της, η τεχνητή νοημοσύνη χωρίζεται σε επιμέρους κατηγορίες, όπως είναι η μηχανική μάθηση, η επίλυση προβλημάτων, τα συστήματα γνώσης, η υπολογιστική όραση, η αναγνώριση φυσικής γλώσσας και η ρομποτική

##### **2.1.2 Μηχανική Μάθηση**

Ο όρος «μηχανική μάθηση» περιγράφει τον κλάδο της τεχνητής νοημοσύνης που ασχολείται με την επιστημονική μελέτη αλγορίθμων και στατιστικών μοντέλων που χρησιμοποιούνται από τα υπολογιστικά συστήματα, με σκοπό την εκτέλεση και



βελτίωση μιας εργασίας με τη χρήση της εμπειρίας τους. Το 1959, ο Arthur Samuel ορίζει τη μηχανική μάθηση ως «Πεδίο μελέτης που δίνει στους υπολογιστές την ικανότητα να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί» [3] [4], ενώ το 1997, ο Tom Mitchell αναφέρει ότι «Η μηχανική μάθηση είναι η μελέτη των αλγορίθμων που επιτρέπουν στα προγράμματα υπολογιστών να βελτιώνονται αυτόματα μέσω της εμπειρίας» [5]. Οι αλγόριθμοι μηχανικής μάθησης δημιουργούν μαθηματικά μοντέλα και πρότυπα από ένα σύνολο δεδομένων πάνω στο οποίο εκπαιδεύονται και έχουν ως στόχο την πρόβλεψη συμπεριφορών ή τη λήψη αποφάσεων βασιζόμενοι στην εκμάθηση τους. Πιο συγκεκριμένα, κατασκευάζουν μοντέλα από πειραματικά δεδομένα και βασιζόμενοι σε αυτά εξάγουν το τελικό αποτέλεσμα που συνιστά μια πρόβλεψη ή μία απόφαση.

Η μηχανική μάθηση συνδέεται στενά με τη μαθηματική βελτιστοποίηση, τη στατιστική και τη θεωρία πιθανοτήτων και συναντά εφαρμογές σε ένα ευρύ πεδίο εργασιών, όπου δεν είναι εφικτός ο ρητός προγραμματισμός των αλγορίθμων. Παράδειγμα εφαρμογών αποτελεί το φιλτράρισμα της ανεπιθύμητης αλληλογραφίας, η οπτική αναγνώριση χαρακτήρων, η υπολογιστική όραση και οι μηχανές αναζήτησης.

Η μηχανική μάθηση χωρίζεται σε τρεις κατηγορίες ανάλογα με τη φύση και τον στόχο του εκάστοτε προβλήματος. Σημειώνεται ότι οι κατηγορίες αυτές είναι, σε μεγάλο βαθμό, συνυφασμένες με τους τρόπους με τους οποίους μαθαίνει ο άνθρωπος. Οι τρεις κατηγορίες είναι οι εξής:

-Επιβλεπόμενη μάθηση (Supervised Learning): οι αλγόριθμοι κατασκευάζουν μια συνάρτηση απεικόνισης των δεδομένων εισόδου του συνόλου εκπαίδευσης με τις γνωστές εξόδους τους, με σκοπό την όσο το δυνατόν πιο επιτυχημένη γενίκευση της συνάρτησης αυτής, έτσι ώστε να προβλέπονται για νέα δεδομένα εισόδου τα επιθυμητά, αλλά άγνωστα αποτελέσματά τους. Συνεπώς, για την εκπαίδευση απαιτούνται δεδομένα με επισημασμένες τις σωστές απαντήσεις τους. Αυτού του είδους η μάθηση συναντάται σε προβλήματα ταξινόμησης, πρόβλεψης και διερμηνείας.

-Μη επιβλεπόμενη μάθηση (Unsupervised Learning): οι αλγόριθμοι βάσει των δεδομένων εισόδου κατασκευάζουν μοντέλα μέσω των οποίων εντοπίζονται συσχετίσεις και δημιουργούνται ομάδες δεδομένων με σκοπό την επίλυση του

προβλήματος χωρίς τη χρήση εμπειρίας κι άρα κάποιας καθοδήγησης. Συνεπώς, για την εκπαίδευση δεν απαιτούνται δεδομένα στα οποία έχουν επισημανθεί τα σωστά αποτελέσματα. Αυτού του είδους η μάθηση συναντάται σε προβλήματα, όπως ανάλυσης συσχετισμών και ομαδοποίησης

-Ενισχυτική Μάθηση (Reinforcement Learning): οι αλγόριθμοι μαθαίνουν μια στρατηγική ενεργειών μέσα από την αλληλεπίδραση τους με ένα δυναμικό περιβάλλον στο οποίο πρέπει να επιτευχθεί ένας συγκεκριμένος στόχος χωρίς να υπάρχουν ρητές οδηγίες. Αυτού του είδους η μάθηση συναντάται σε προβλήματα σχεδιασμού, όπως για παράδειγμα ο έλεγχος κίνησης ρομπότ, η οδήγηση ενός οχήματος, η βελτιστοποίηση εργασιών σε εργοστασιακούς χώρους και η αντιμετώπιση των αντιπάλων σε ηλεκτρονικά παιχνίδια.

Επίσης, τα προβλήματα της μηχανικής μάθησης χωρίζονται σε επιπλέον υποκατηγορίες ανάλογα με το είδος της πρόβλεψης που χρειάζεται να κάνουμε, δηλαδή ανάλογα με το αποτέλεσμα που επιθυμούμε να λάβουμε. Έτσι, έχουμε τα ακόλουθα είδη προβλημάτων:

-Ταξινόμησης (Classification): όταν τα δεδομένα εισόδου χωρίζονται σε δύο ή περισσότερες κλάσεις, στις οποίες το μοντέλο που εκπαιδεύεται πρέπει να τα αντιστοιχίσει σωστά. Ως έξοδος προβλέπεται μία διακριτή τιμή ως ταμπέλα μιας από τις κλάσεις του προβλήματος. Αυτού του είδους τα προβλήματα είναι χαρακτηριστικά της επιβλεπόμενης μάθησης

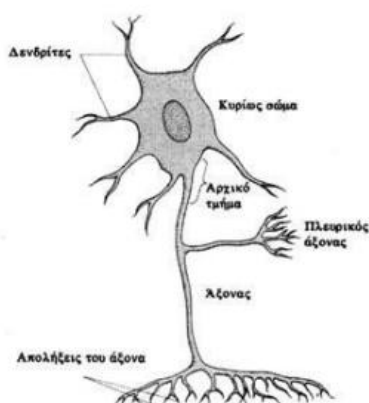
-Παλινδρόμησης (Regression): όταν το εκπαιδευμένο μοντέλο προβλέπει στην έξοδο του συνεχείς αριθμητικές τιμές. Αυτού του είδους τα προβλήματα είναι, επίσης, χαρακτηριστικά της επιβλεπόμενης μάθησης.

-Συσταδοποίησης (Clustering): όταν το σύνολο δεδομένων πρέπει να χωριστεί σε ομάδες, οι οποίες, όμως, σε αντίθεση με τα προβλήματα ταξινόμησης δεν είναι γνωστές, αλλά δημιουργούνται κατά την εκπαίδευση του μοντέλου. Αυτού του είδους τα προβλήματα είναι χαρακτηριστικά της μη επιβλεπόμενης μάθησης.

### 2.1.3 Τεχνητά Νευρωνικά Δίκτυα

Τα Τεχνητά Νευρωνικά Δίκτυα αποτελούν υπολογιστικά μοντέλα επεξεργασίας πληροφοριών, η λειτουργία των οποίων προσομοιάζει στους βιολογικούς νευρώνες του ανθρώπινου εγκεφάλου [6] [7]. Αποτελούν αναπόσπαστο μέρος της μηχανικής μάθησης και έχουν συμβάλει σημαντικά στην εξέλιξη διαφόρων εφαρμογών σε μια πληθώρα επιστημονικών κλάδων. Συνεπώς, τα Τεχνητά Νευρωνικά Δίκτυα είναι στενά συνδεδεμένα και προσομοιάζουν τη λειτουργία των Βιολογικών Νευρωνικών Δικτύων.

Τα Βιολογικά Νευρωνικά Δίκτυα είναι κυκλώματα διασυνδεδεμένων νευρώνων και αποτελούν τμήματα του νευρικού ιστού. Ο τυπικός νευρώνας του ανθρώπινου εγκεφάλου, όπως απεικονίζεται παρακάτω, απαρτίζει μεγάλα και πολύπλοκα δίκτυα του εγκεφάλου που δέχονται τις πληροφορίες του περιβάλλοντος και τις επεξεργάζονται καταλλήλως. Κάθε νευρώνας αποτελείται από το σώμα, τους δενδρίτες με τους οποίους συλλέγει πληροφορίες από άλλους νευρώνες και τον άξονα μέσω του οποίου πυροδοτεί το σήμα σε άλλους νευρώνες. Οι πληροφορίες που λαμβάνει κάθε νευρώνας αποτελούν τα σήματα εισόδου του, ενώ το σήμα που εξάγει αποτελεί την έξοδο του. Έτσι, λειτουργούν τα Βιολογικά Νευρωνικά Δίκτυα με στόχο την προσαρμογή του ανθρώπου στις αλλαγές του περιβάλλοντος.

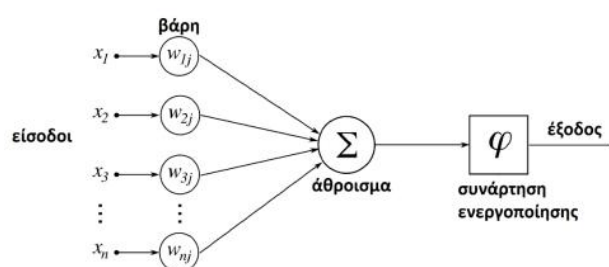


Εικόνα 1: Νευρώνας του ανθρώπινου εγκεφάλου [8]

Τα Τεχνητά Νευρωνικά Δίκτυα ακολουθούν ως πρότυπο τα Βιολογικά Νευρωνικά Δίκτυα. Η έννοια του τεχνητού νευρώνα εμφανίστηκε το 1958 από τον Rosenblatt, με το όνομα «Perceptron». Ο τεχνητός νευρώνας αποτελείται από το σώμα, λαμβάνει κάποια σήματα εισόδου και βγάζει μία έξοδο. Πιο συγκεκριμένα, σε

ένα τέτοιο δίκτυο οι εισοδοί πολλαπλασιάζονται με κάποια βάρη και τα γινόμενα αθροίζονται και τροφοδοτούν τη συνάρτηση ενεργοποίησης, ώστε να βγει κάποιο από τα αποτελέσματα 0 ή 1 που σηματοδοτούν την απενεργοποίηση ή την ενεργοποίηση του νευρώνα.

Παρακάτω απεικονίζεται ένας τεχνητός νευρώνας.



Εικόνα 2: Τεχνητός νευρώνας [9]

Τα  $x_i$  είναι οι εισοδοί του νευρώνα, οι οποίες πολλαπλασιάζονται με τα βάρη  $w_i$ . Το σώμα του νευρώνα αποτελείται από δύο μέρη. Το 1ο μέρος είναι ο αθροιστής ( $\Sigma$ ), ο οποίος προσθέτει τα σταθμισμένα από τα βάρη σήματα εισόδου μαζί με μία πόλωση, που συμβολίζεται με  $w_0$  ή  $b$ . Το 2ο μέρος του σώματος του νευρώνα είναι η συνάρτηση ενεργοποίησης  $\varphi$ , η οποία τροφοδοτείται με το άθροισμα και ανάλογα την τιμή που αυτό έχει και το είδος της δίνει έξοδο 0 ή 1, δηλαδή ο νευρώνας απενεργοποιείται ή ενεργοποιείται.

Η μαθηματική έκφραση είναι:

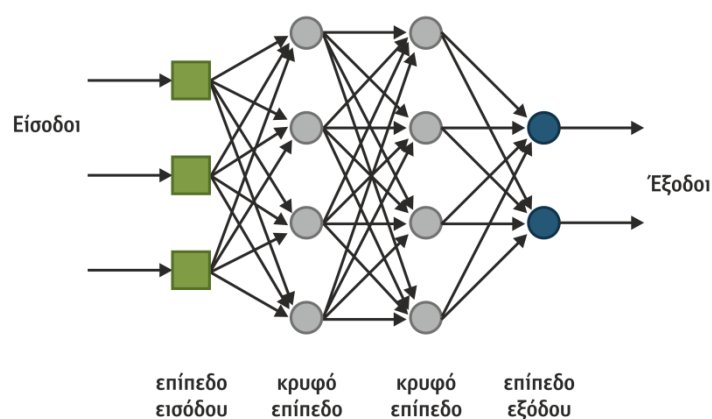
Εξίσωση 1: Υπολογισμός εξόδου τεχνητού νευρώνα [9]

$$y = f\left(\sum_{i=1}^n x_i w_i + b\right)$$

Πολλές δομές τεχνητών νευρώνων συνδυάζονται σε πολλαπλά επίπεδα δημιουργώντας ένα δίκτυο. Οι νευρώνες ενός δικτύου οργανώνονται σε διάφορα επίπεδα παράλληλης διάταξης. Τέτοιες αρχιτεκτονικές πολλών επιπέδων Perceptrons ονομάζονται Multi-Layer Perceptrons (MLP). Κάθε τέτοιο δίκτυο απαρτίζεται από το επίπεδο εισόδου που αντιστοιχεί στο αρχικό επίπεδο, το επίπεδο εξόδου που αντιστοιχεί στο τελευταίο και στα κρυφά επίπεδα που αφορούν τα ενδιάμεσα. Καθένα

από τα κρυφά επίπεδα αποτελείται από νευρώνες που συνδέονται με αυτούς του προηγούμενου και του επόμενου επιπέδου κι έτσι επιτυγχάνεται η διάδοση του σήματος. Στο επίπεδο εξόδου εμφανίζεται το αποτέλεσμα της εκμάθησης και προσαρμογής του δικτύου αυτού στα δεδομένα εισόδου, γι' αυτό και το πλήθος των διασυνδεδεμένων νευρώνων σε αυτό είναι όσες και οι επιθυμητές εξοδοί.

Ένα τέτοιο τεχνητό νευρωνικό δίκτυο (Artificial Neural Network – ANN) παρουσιάζεται παρακάτω.



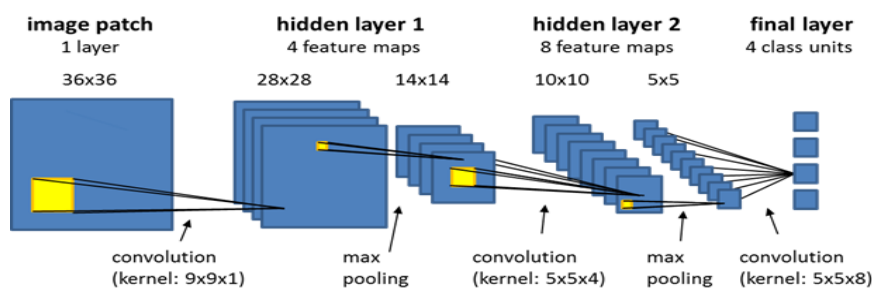
Εικόνα 3: Απεικόνιση τεχνητού νευρωνικού δικτύου [9]

Η επιτυχία των νευρωνικών δικτύων αυτών βασίζεται στην ικανότητά τους να προσαρμόζουν αυτόματα τις παραμέτρους τους στο περιβάλλον τους, καθώς και στην ικανότητα τους να μοντελοποιούν πολύπλοκες και μη γραμμικές σχέσεις μεταξύ των δεδομένων. Έτσι, τα Τεχνητά Νευρωνικά Δίκτυα μπορούν να δώσουν αποδοτικές λύσεις σε διάφορα προβλήματα αναγνώρισης προτύπων και ταξινόμησης.

Μία ειδική κατηγορία νευρωνικών δικτύων αποτελούν τα Συνελκτικά Νευρωνικά Δίκτυα (CNNs), τα οποία είναι κατάλληλα για την επεξεργασία πληροφορίας που μπορεί να αναπαρασταθεί σε μορφή πλέγματος. Τα CNNs βρίσκουν εφαρμογή στην αναγνώριση εικόνων και βίντεο, την κατηγοριοποίηση εικόνας, την ανάλυση ιατρικών εικόνων, τις χρονοσειρές και την επεξεργασία φυσικής γλώσσας. Όπως δηλώνει και το όνομα τους, βασίζονται στη μαθηματική διαδικασία της συνέλιξης

Η είσοδος της συνέλιξης είναι είτε ακατέργαστα δεδομένα, είτε κάποιος χάρτης χαρακτηριστικών που προέρχεται από προηγούμενα επίπεδα CNNs. Ένα φίλτρο, ο

λεγόμενος πυρήνας «σαρώνει» κατά πλάτος και κατά ύψος τα δεδομένα εισόδου και παράγει χαρακτηριστικά μέσα από τη συνέλιξη. Συγκεκριμένα, η είσοδος περνάει ακολουθιακά από διαδοχικά επίπεδα επεξεργασίας, προκειμένου να εξαχθεί η τελική πρόβλεψη του δικτύου. Η διαδικασία αυτή είναι σειριακή, και ακολουθείται από μία άλλη διαδικασία, για την εκπαίδευση του δικτύου, την οπισθοδιάδοση σφάλματος.



Εικόνα 4: Απεικόνιση συνελκτικού νευρωνικού δικτύου [9]

Ένα Βαθύ Συνελκτικό Νευρωνικό Δίκτυο αποτελείται από έναν αριθμό από συνελκτικά και υποδειγματοληπτικά επίπεδα, τα οποία συνήθως ακολουθούνται από πλήρως συνδεδεμένα επίπεδα.

Το συνελκτικό επίπεδο χρησιμοποιεί ένα σύνολο από φίλτρα τα οποία εντοπίζουν την παρουσία συγκεκριμένων χαρακτηριστικών ή μοτίβων που παρουσιάζονται στα δεδομένα εισόδου. Το κάθε φίλτρο «σαρώνει» κατά πλάτος και κατά ύψος την είσοδο και με τη μέθοδο της συνέλιξης δημιουργεί διαδοχικά ένα νέο πίνακα που ονομάζεται χάρτης χαρακτηριστικών (feature map). Η έξοδος του επιπέδου αυτού καθορίζεται από το μέγεθος του κάθε φίλτρου, το πλήθος των φίλτρων, το βήμα και το padding.

Το επίπεδο υποδειγματοληψίας (pooling layer) χρησιμοποιείται για τη μείωση της διάστασης του χάρτη χαρακτηριστικών. Αυτό εξυπηρετεί στην μείωση της ακρίβειας των χαρακτηριστικών κι άρα της ευαισθησίας στις διαφορές μεταξύ των δειγμάτων, ώστε με κάποιον τρόπο να αποφευχθεί η υπερεκπαίδευση. Για αυτό το λόγο, στα περισσότερα CNNs ένα συνελκτικό επίπεδο ακολουθείται από ένα επίπεδο υποδειγματοληψίας. Η υποδειγματοληψία επιτυγχάνεται είτε με την επιλογή των μέγιστων τιμών του παραθύρου βάσει συνέλιξης με το φίλτρο (Max Pooling), είτε με τον υπολογισμό του μέσου όρου των τιμών του αντίστοιχου παραθύρου του πίνακα (Average Pooling).

Το πλήρως συνδεδεμένο επίπεδο (fully connected layer/dense layer) αποτελεί ένα επίπεδο οι νευρώνες του οποίου συνδέονται με όλους τους νευρώνες του προηγούμενου επιπέδου. Αναμένει στην είσοδό του ένα μονοδιάστατο διάνυσμα αριθμών, γι' αυτό και η τρισδιάστατη έξοδος του προηγούμενου επιπέδου ισοπεδώνεται με τα λεγόμενα flatten layers. Ένα τέτοιο επίπεδο εξάγει την τελική έξοδο του δικτύου στα προβλήματα ταξινόμησης και παλινδρόμησης.

## **2.2 Αναγνώριση ανθρώπινης δραστηριότητας**

### **2.2.1 Ορισμός**

Η αναγνώριση της ανθρώπινης δραστηριότητας (κίνησης - δράσης) είναι ένας σημαντικός τομέας στον κλάδο της αναγνώρισης προτύπων και σχετίζεται με την ανίχνευση και ταξινόμηση των ανθρώπινων δραστηριοτήτων από πληροφορίες που λαμβάνονται από κάμερες ή άλλα μέσα ανίχνευσης, όπως τα επιταχυνσιόμετρα [10]. Συναντά πολλές εφαρμογές στην καθημερινή ζωή που εκτείνονται από παιχνίδια κονσόλας και άσκησης που χρησιμοποιούν κάμερες, εφαρμογές αλληλεπίδρασης ανθρώπου-υπολογιστή έως και αυτοματοποιημένη παρακολούθηση και περιβάλλοντα υποβοήθησης [11] [12]. Η συνολική κίνηση του σώματος και οι επιμέρους κινήσεις των άκρων του είναι αυτές που συμβάλλουν στην αναγνώριση της ανθρώπινης δραστηριότητας γι' αυτό και δίνεται έμφαση στην ανίχνευση της στάσης του σώματος σε κάθε μία από τις φάσεις ανάλυσης της κίνησης.

Ιδιαίτερη σημασία ερευνητικά έχει δοθεί στην όραση υπολογιστών που στοχεύει στην ανάλυση βίντεο με σκοπό τον προσδιορισμό των ενεργειών που πραγματοποιούνται σε αυτό και άρα την αναγνώριση της ανθρώπινης δραστηριότητας. Η ανίχνευση δράσης αναφέρεται στον εντοπισμό της θέσης στην οποία πραγματοποιείται μια ενέργεια σε χρόνο και χώρο από δεδομένα ακολουθιών εικόνων που δεν έχουν τμηματοποιηθεί [13]. Πληθώρα εφαρμογών που χρησιμοποιούν τα καρτέ του βίντεο σαν πηγή δεδομένων έχουν αναπτυχθεί με στόχο την πιο ακριβή και έγκυρη προσέγγιση της ανθρώπινης κίνησης.

Αξίζει να σημειωθεί ότι ένα βίντεο έχει χωρική πληροφορία που περιέχεται σε καθένα από τα καρέ του και χρονική πληροφορία που σχετίζεται με τη σειρά των καρέ. Ορισμένες ενέργειες, όπως είναι κάποια στάση του σώματος ή το τρέξιμο, πιθανότατα μπορούν να προσδιοριστούν χρησιμοποιώντας μόνο ένα καρέ, αλλά πιο σύνθετες ενέργειες που απαιτούν ταξινόμηση και διαχωρισμό μεταξύ τους, όπως το περπάτημα έναντι του τρεξίματος ή η κάμψη έναντι της πτώσης, ενδέχεται να απαιτούν περισσότερες πληροφορίες από αυτές που παρέχει ένα μόνο καρέ για τη σωστή ταυτοποίησή τους. Στις περιπτώσεις διάκρισης τέτοιου είδους ενεργειών παίζουν σημαντικό ρόλο οι χρονικές πληροφορίες μικρής ή μεγάλης διάρκειας που δίνονται από την αλληλουχία των καρέ.

Η σημασία, λοιπόν, της ανίχνευσης και αναγνώρισης της ανθρώπινης δραστηριότητας είναι μεγάλη, γι' αυτό και τα προβλήματα αυτά σχετίζονται στενά με τη μηχανική μάθηση, τα εργαλεία και τους αλγορίθμους της, για την εξαγωγή χαρακτηριστικών και την εκμάθηση αυτών με σκοπό την εξαγωγή ενός γενικευμένου αποτελέσματος και τη δημιουργία μιας πιο διευρυμένης εφαρμογής.

### **2.2.2 Αναπαράσταση της κίνησης**

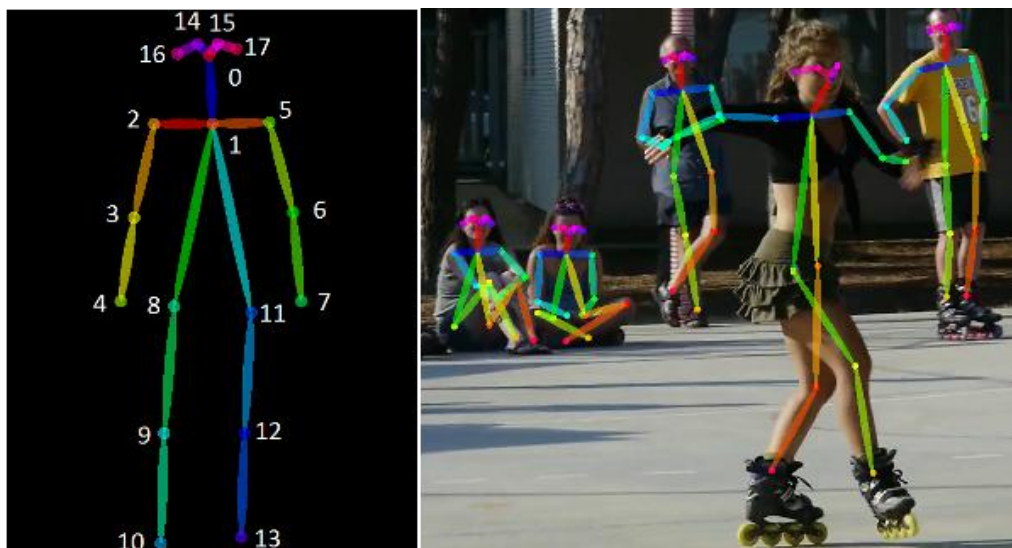
Η μοντελοποίηση της ανθρώπινης δράσης και η αναπαράσταση των χαρακτηριστικών αυτής αποτελούν σημαντικές συνιστώσες των προβλημάτων ανίχνευσης και αναγνώρισης της ανθρώπινης δραστηριότητας. Η αναπαράσταση της κίνησης και η επιλογή χαρακτηριστικών αυτής συνιστούν ένα κλασικό πρόβλημα της όρασης υπολογιστών και της μηχανικής μάθησης. Σε αντίθεση με την αναπαράσταση χαρακτηριστικών σε έναν χώρο εικόνας, η αναπαράσταση της ανθρώπινης δράσης σε ένα βίντεο όχι μόνο περιγράφει την παρουσία των ανθρώπων στον χώρο της εικόνας, αλλά επίσης εξάγει και τις αλλαγές στην εμφάνιση και τη στάση τους. Το πρόβλημα της αναπαράστασης χαρακτηριστικών επεκτείνεται από τον δισδιάστατο χώρο στον τρισδιάστατο χωροχρόνο [13]. Τα τελευταία χρόνια, έχουν προταθεί πολλά είδη μεθόδων αναπαράστασης δράσης, συμπεριλαμβανομένων τοπικών και παγκόσμιων χαρακτηριστικών που βασίζονται σε χρονικές και χωρικές αλλαγές [14] [15] [16], χαρακτηριστικά τροχιάς βασισμένα στην παρακολούθηση βασικών σημείων [17] [18], αλλαγές κίνησης βάσει πληροφοριών βάθους [19] [20] [21] και χαρακτηριστικά δράσης που βασίζονται σε ανθρώπινες στάσεις [22] [23]. Τέλος, η αναπαράσταση



κίνησης στοχεύει στην κατάλληλη μοντελοποίηση αυτής και λειτουργεί ως ένα είδος προεπεξεργασίας των δεδομένων κίνησης συμβάλλοντας έτσι στη χρήση απλούστερων δικτύων εκμάθησης κι άρα στη δημιουργία λιγότερο πολύπλοκων εφαρμογών.

### 2.2.2.1 Τεχνικές ανίχνευσης πόζας (*Pose detection - estimation*)

Η ανίχνευση ή εκτίμηση στάσης του ανθρώπινου σώματος είναι αποτέλεσμα της χρήσης ενός μοντέλου μηχανικής μάθησης για την εκτίμηση της θέσης ενός ατόμου από μία εικόνα ή ένα βίντεο με την εκτίμηση των χωρικών θέσεων των βασικών αρθρώσεων του σώματος. Με αυτόν τον τρόπο μοντελοποιείται το ανθρώπινο σώμα με βάση τον σκελετό του, δηλαδή με βάση ένα σύνολο αρθρώσεων και τον προσανατολισμό των άκρων του. Ουσιαστικά, οι αρθρώσεις - σημεία του σώματος εκφράζονται μέσω συντεταγμένων στον δισδιάστατο ή τρισδιάστατο χώρο κι έτσι μπορούν να συνδεθούν κατάλληλα μεταξύ τους με σκοπό την προσομοίωση των άκρων του σώματος.



Εικόνα 5: Pose estimation [24]

Τα τελευταία χρόνια έχουν αναπτυχθεί τέτοιου είδους αλγόριθμοι με σκοπό την αναγνώριση και κατηγοριοποίηση της ανθρώπινης κίνησης μέσω της εκτίμησης της θέσης των μελών και των αρθρώσεων του προς μελέτη ανθρώπινου σώματος.

Οι μέθοδοι που ακολουθούνται για την εκτίμηση της θέσης ποικίλουν ανάλογα με τις δυνατότητες ανίχνευσης που έχουν, καθώς και με τον τρόπο που προσεγγίζουν το πρόβλημα. Οι πρώτες μέθοδοι που εισήχθησαν εκτιμούσαν τη στάση ενός ατόμου σε μια εικόνα που περιέχει μόνο αυτό το άτομο. Αυτές οι μέθοδοι αρχικά προσδιορίζουν τα μεμονωμένα μέρη και στη συνέχεια σχηματίζουν συνδέσεις μεταξύ τους για τη δημιουργία της στάσης. Δεδομένου ότι αυτές οι μέθοδοι δεν ανταποκρίνονται σε πολλές περιπτώσεις της πραγματικής ζωής, αναπτύχθηκαν μέθοδοι εκτίμησης της θέσης πολλών ατόμων στον ίδιο χώρο. Η διαδικασία αυτή είναι αυξημένης δυσκολίας, καθώς η τοποθεσία και ο αριθμός των ατόμων σε μία εικόνα δεν είναι γνωστά.

Αξίζει να αναφερθούν ορισμένες μέθοδοι που ανταποκρίνονται στα όσα αναφέραμε παραπάνω και αποτελούν σημαντικούς ερευνητικούς σταθμούς στο πεδίο της αναγνώρισης της κίνησης.

-Kinect: Το 2010 η Microsoft κυκλοφόρησε τον αισθητήρα Kinect που κατέγραφε την ανθρώπινη στάση μέσα από εικόνες RGB και εικόνες βάθους. Ο αισθητήρας περιέχει δύο κάμερες, μια κλασική για βίντεο και μια υπερέυθρων για εικόνες, μια ειδική πηγή υπερέυθρων και τέσσερα μικρόφωνα. Περιέχει, επίσης, και hardware επεξεργασίας σήματος για την κατανόηση των δεδομένων που λαμβάνει. Οι έξοδοι που δίνουν οι επιμέρους αισθητήρες αυτοί μπορούν να συνδυαστούν και μέσω ενός προγράμματος λογισμικού να οδηγήσουν στην αναγνώριση και παρακολούθηση αντικειμένων, στον καθορισμό της κατεύθυνσης του ήχου και της απομόνωσης του θορύβου του περιβάλλοντος. Ο αισθητήρας βάθους προσφέρει τη δυνατότητα στο εργαλείο αυτό να «βλέπει» σε 3D, αφού το Kinect δημιουργεί έναν «χάρτη βάθους» της περιοχής που βρίσκεται μπροστά του. Αυτός ο χάρτης παράγεται εξ ολοκλήρου μέσα στη γραμμή αισθητήρων και στη συνέχεια μεταδίδεται από το καλώδιο USB στον κεντρικό υπολογιστή με τον ίδιο τρόπο που θα μεταφερόταν μια τυπική εικόνα κάμερας εκτός από ότι, αντί για πληροφορίες χρώματος για κάθε εικονοστοιχείο σε μια εικόνα, ο αισθητήρας μεταδίδει τιμές απόστασης. Για τον υπολογισμό των αποστάσεων χρησιμοποιείται ο προβολέας υπερέυθρων και η κάμερα υπερέυθρων που μπορεί και εντοπίζει τις μικροσκοπικές κουκίδες που έχει παράξει ο προβολέας πάνω στην εικόνα. Το Kinect, όμως, διαθέτει μια κλασική κάμερα 2D την αναπαράσταση της οποίας συγκρίνει με αυτή των υπερέυθρων, οπότε και υπολογίζεται η απόσταση κάθε σημείου από τον αισθητήρα. Όσον αφορά την αναγνώριση και παρακολούθηση

των ανθρώπων που στέκονται μπροστά στον αισθητήρα, ο υπολογιστής που λαμβάνει την εικόνα βάθους χρησιμοποιεί τη μηχανική μάθηση έχοντας εκπαιδευτεί ένα λογισμικό πάνω σε ανθρώπινες φιγούρες και εξάγει τη θέση καθενός μέρους του σώματος. Το σύστημα είναι χαμηλού κόστους και υψηλής ανάλυσης [25].

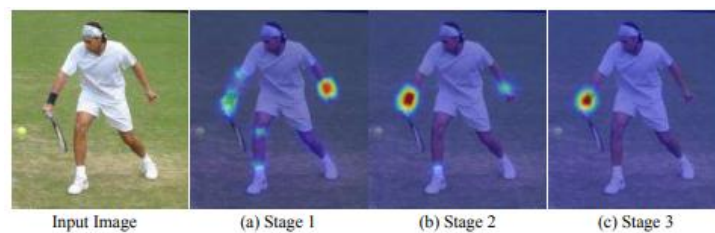


Εικόνα 6: Pose estimation με χρήση του Kinect [25]

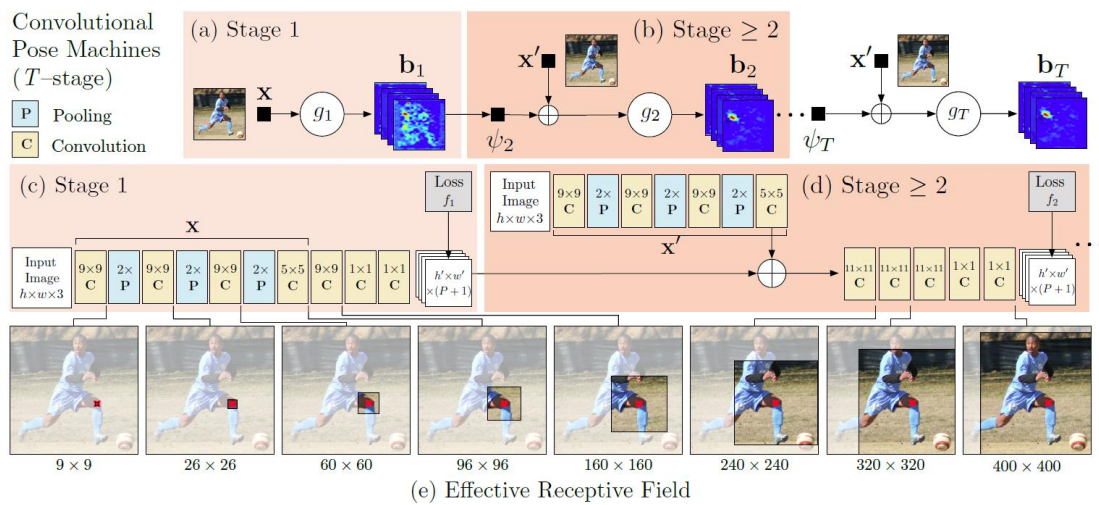
-Convolutional Pose Machines (CPMs): αποτελούνται από διαδοχικά συνελκτικά επίπεδα που εκπαιδεύονται ώστε να προβλέπουν σωστά τη θέση - τοποθεσία της κάθε άρθρωσης του ανθρώπινου σώματος σε μία φωτογραφία. Τα CPMs συνδυάζουν τα οφέλη των αρχιτεκτονικών Pose Machines με αυτά των συνελκτικών δικτύων. Τα Pose Machines χρησιμοποιούν μια ακολουθία από ταξινομητές πολλών κλάσεων που εκπαιδεύονται να προβλέπουν τη θέση των διαφόρων μερών του σώματος (αρθρώσεων) και παράγουν σε κάθε επίπεδο ένα χάρτη πεποιθήσεως σχετικά με τη θέση της εκάστοτε άρθρωσης προς αναζήτηση. Οι αρχιτεκτονικές αυτές διακρίνονται για τη μεγάλης εμβέλειας εκμάθηση εξαρτήσεων μεταξύ εικόνας και σημείων, τη στενή συσχέτιση ανάμεσα στην μάθηση και στα συμπεράσματα, καθώς και το αρθρωτό μοντέλο σχεδίασης που διαθέτουν [26]. Τα Pose Machines αντικαθίστανται τώρα από αρχιτεκτονικές βαθιάς μάθησης, όπως είναι τα συνελκτικά δίκτυα, προσδίδοντας σε αυτά την ικανότητα εξαγωγής και εκμάθησης χαρακτηριστικών απευθείας από τα δεδομένα των εικόνων. Τα CPMs, λοιπόν, διαθέτουν μία αλληλουχία συνελκτικών επιπέδων που παράγουν κάθε φορά χάρτες πεποιθήσεως (belief maps) που τροφοδοτούν το εκάστοτε επόμενο επίπεδο με μία εκφραστική μη παραμετρική κωδικοποίηση της χωρικής αβεβαιότητας της θέσης κάθε άρθρωσης με όλο και μεγαλύτερη ακρίβεια στην εκτίμηση της. Οι χάρτες που παράγονται στην έξοδο κάθε επιπέδου και δίνονται στο επόμενο επίπεδο ως είσοδοι

σε συνδυασμό με τα χαρακτηριστικά που εξάγονται από αυτούς και τα δεδομένα των εικόνων, οδηγούν στην κατάλληλη εκμάθηση των συσχετίσεων μεταξύ των διαφόρων μερών του σώματος κι άρα στην ανίχνευση της στάσης του προς μελέτη ανθρώπινου σώματος [27].

Σημειώνεται ότι τα CPMs ανιχνεύουν 14 σημεία του ανθρώπινου σώματος, αφού είναι εκπαιδευμένο με το σύνολο δεδομένων MPII, το κεφάλι, τον λαιμό, τους ώμους, τους αγκώνες, τους καρπούς, τους γοφούς, τα γόνατα και τους αστράγαλους.



Εικόνα 7: Convolution Pose Machines-εντοπισμός αρθρώσεων [27]



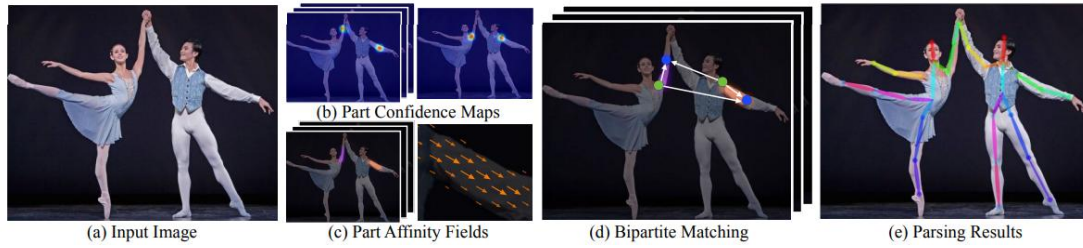
Εικόνα 8: Αρχιτεκτονική CPMs [27]



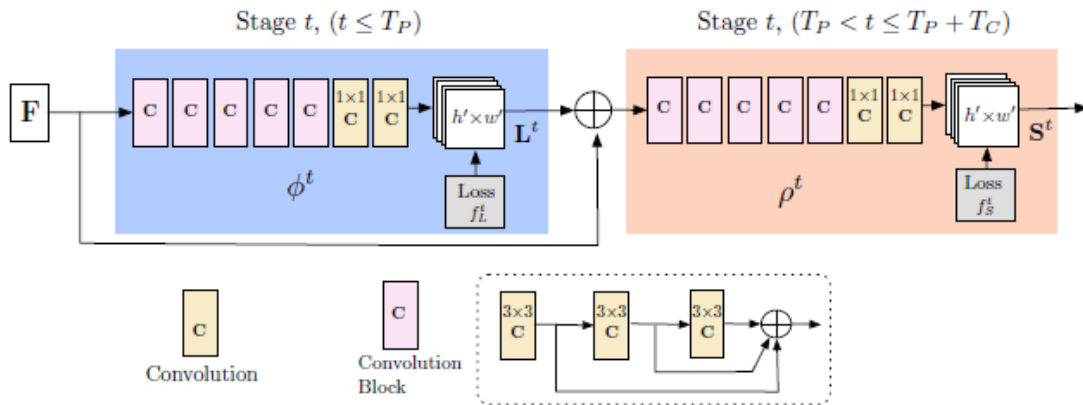
Εικόνα 9: Pose detection using CPMs [28]

-Open Pose: αποτελεί μία bottom-up προσέγγιση για την εκτίμηση σε πραγματικό χρόνο της θέσης - στάσης πολλαπλών ατόμων σε μία εικόνα. Με τη μέθοδο αυτή ανιχνεύονται οι θέσεις των διαφόρων αρθρώσεων και συνδυάζονται κατάλληλα βάσει των ζευγών αρθρώσεων, δηλαδή των άκρων που δημιουργούν αυτές, με σκοπό τη σωστή αντιστοίχισή τους στις ανθρώπινες φιγούρες που παρουσιάζονται στην εικόνα. Για αυτόν τον λόγο χρησιμοποιείται μία αρχιτεκτονική δύο κλάδων και πολλαπλών συνελικτικών επιπέδων που παίρνει ως είσοδο μία RGB εικόνα. Ο ένας κλάδος εξάγει τους χάρτες εμπιστοσύνης (confidence maps) σχετικά με την πιθανή θέση των διαφόρων αρθρώσεων, ενώ ο άλλος κλάδος εξάγει μη παραμετρικές αναπαραστάσεις, τα λεγόμενα πεδία συνάφειας μέρους (Part Affinity Fields - PAFs) που αποτελούν ένα σύνολο διανυσμάτων που κωδικοποιούν το βαθμό συσχέτισης μεταξύ των διαφόρων ανιχνεύσιμων αρθρώσεων, δηλαδή τη θέση και των προσανατολισμό των διαφόρων άκρων, με σκοπό την σωστή αντιστοίχισή τους στα άτομα που ανήκουν. Τα αποτελέσματα των δύο κλάδων συνδέονται κατάλληλα μεταξύ τους κι έτσι λαμβάνονται οι εκτιμήσεις των θέσεων των ατόμων που απεικονίζονται στην εικόνα. Με τη μέθοδο αυτή επιτυγχάνεται η ανίχνευση της ανθρώπινης στάσης σε μία εικόνα όπου ο αριθμός των ατόμων που μπορούν να εμφανιστούν σε οποιαδήποτε θέση είναι άγνωστος και οι αλληλεπιδράσεις μεταξύ ανθρώπων προκαλούν πολύπλοκες χωρικές παρεμβολές [29].

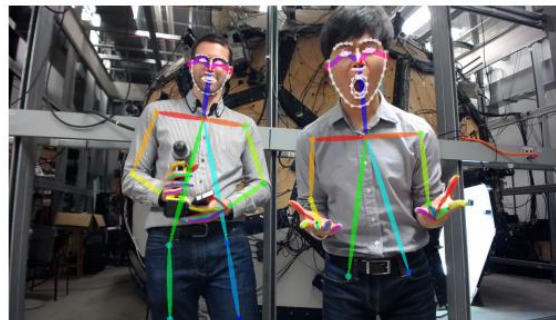
Σημειώνεται ότι το Open Pose ανιχνεύει 18 σημεία του ανθρώπινου σώματος, αφού είναι εκπαιδευμένο με το σύνολο δεδομένων COCO, το κεφάλι, τον λαιμό, τον κορμό, τους ώμους, τους αγκώνες, τους καρπούς, τους γοφούς, τα γόνατα, τους αστραγάλους και τα βασικά σημεία οριοθέτησης του προσώπου, μάτια και αυτιά.



Εικόνα 10: Pipeline της μεθόδου OpenPose [29]



Εικόνα 11: Αρχιτεκτονική των πολυεπίδων συνελκτικών δικτύων [29]



Εικόνα 12: Real-time pose detection using OpenPose [29]

Για τον εντοπισμό και την εξαγωγή των αρθρώσεων με σκοπό την αναγνώριση της ανθρώπινης δράσης χρησιμοποιούνται και μέθοδοι πιο επεμβατικοί με την έννοια ότι το υπό εξέταση άτομο πρέπει να φέρει αισθητήρες στα προς ανίχνευση σημεία. Οι μέθοδοι αυτοί καταγράφουν μέσω των αισθητήρων την κίνηση ενός αντικειμένου ή ενός ανθρώπου και μπορούν να χρησιμοποιηθούν σε πληθώρα εφαρμογών στην ψυχαγωγία, την ιατρική, τον στρατό και τον αθλητισμό.



-Motion capture with sensors (mocaps): για να καταγραφεί η κίνηση ενός ατόμου αυτό πρέπει να φοράει μία ειδική στολή που αποτελείται από αισθητήρες που αντιστοιχίζονται σε ορισμένα κρίσιμα σημεία της κίνησης του σώματος με στόχο την παρακολούθηση αυτής μέσω των εικόνων που δημιουργούνται στον υπολογιστή. Οι δείκτες που χρησιμοποιούνται στις στολές και υποδεικνύουν τα κύρια σημεία ενδιαφέροντος μπορεί να είναι αναδρομικοί αντανακλαστικοί, οι οποίοι αντανακλούν το φως που παράγεται κοντά στην κάμερα ή να εκπέμπουν φως οι ίδιοι απαιτώντας κάποια πηγή ισχύος. Μόλις σταλεί το φως από τους δείκτες, αυτό χρησιμοποιείται για τον υπολογισμό της θέσης του ανθρώπου στον τρισδιάστατο χώρο. Ωστόσο, υπάρχουν και τεχνικές καταγραφής κίνησης που δεν χρησιμοποιούν δείκτες-αισθητήρες, αλλά ευαίσθητες σε βάθος κάμερες που εντοπίζουν την κίνηση. Επίσης, ορισμένες τεχνικές δεν χρησιμοποιούν κάμερες για την καταγραφή της κίνησης, αλλά αδρανειακές μονάδες κίνησης, όπως είναι τα γυροσκόπια, τα επιταχυνσιόμετρα και τα μαγνητόμετρα [30] [31].

Οι μέθοδοι motion capture χρησιμοποιούνται κατά βάση στον κινηματογράφο, όπου οι ηθοποιοί καλούνται να υποδυθούν ρόλους επιστημονικής φαντασίας ή καρτούν, οπότε και η ανάγκη καταγραφής της κίνησης είναι απαραίτητα για να έχουν ανθρώπινη υπόσταση οι φιγούρες που δημιουργούνται.



Εικόνα 13: Στολές για motion capturing με αισθητήρες [30]

### ***2.2.2.2 Χωροχρονικά σημεία ενδιαφέροντος (Spatio-temporal interest points)***

Τα χωροχρονικά σημεία ενδιαφέροντος (Spatio-Temporal Interest Points-STIPs) αποτελούν μία στρατηγική εξαγωγής τοπικών χαρακτηριστικών από τα προς μελέτη βίντεο. Τα τοπικά αυτά χαρακτηριστικά παραμένουν αναλλοίωτα σε διάφορες παραλλαγές βίντεο, όπως είναι η περιστροφή, οι αλλαγές στην κλίμακα, οι μετασχηματισμοί, η αλλαγή του σημείου προβολής, ο θόρυβος και το κούνημα της κάμερας [32]. Τα STIPs είναι σημεία με σημαντική τοπική διακύμανση της έντασης των pixel στον χωροχρονικό τομέα του όγκου βίντεο. Επιπλέον, τα STIPs μπορεί να ανιχνευθούν απευθείας από το βίντεο για να περιγραφούν κινούμενα αντικείμενα, χωρίς την ανάγκη μοντελοποίησης φόντου και τμηματοποίησης προσκηνίου. Τα χαρακτηριστικά που εξάγονται μέσω των σημείων αυτών στοχεύουν στην ανίχνευση των χωρικών και χρονικών αλλαγών που αντιπροσωπεύουν την προς μελέτη δράση στο βίντεο [13]. Τα τελευταία χρόνια, ερευνητικές μελέτες στον τομέα της ανάλυσης βίντεο έχουν επικεντρωθεί στα STIPs δεδομένων των πλεονεκτημάτων τους και πολλοί αλγόριθμοι αναγνώρισης δράσης έχουν αναπτυχθεί έχοντας αυτά ως βασικό δομικό στοιχείο τους [33].

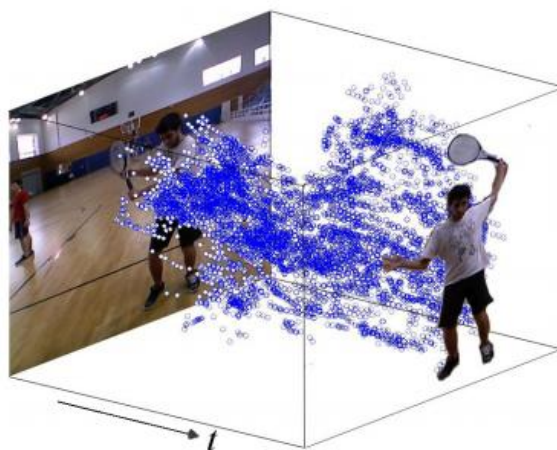
Τα STIPs χρησιμοποιούνται ευρέως σε εφαρμογές ανίχνευσης και αναγνώρισης αντικειμένων, ανάλυσης και ταξινόμησης εικόνων [34], αναγνώρισης της ανθρώπινης δράσης [35] [36] [37], αναγνώρισης προσώπων σε εικόνες και βίντεο [38], ανάλυσης βίντεο [39], παρακολούθησης βίντεο [40], σύνοψης βίντεο [41], ανάκτησης βίντεο βάσει περιεχομένου [42], εικονικής πραγματικότητας [43] και αλληλεπίδρασης ανθρώπου-υπολογιστή [44].

Όσον αφορά την αναγνώριση της ανθρώπινης δραστηριότητας, οι μέθοδοι που βασίζονται σε STIPs εντοπίζουν τη βασική περιοχή αλλαγής της κίνησης και εξάγουν τα κατάλληλα τοπικά χαρακτηριστικά, ώστε να μοντελοποιήσουν τη δράση. Καθορίζουν τις μεθόδους ανίχνευσης και διανυσματικής αναπαράστασης των χαρακτηριστικών της εξεταζόμενης περιοχής, καθώς και τους αλγόριθμους ταξινόμησης που θα χρησιμοποιηθούν στη συνέχεια. Τα σημεία ενδιαφέροντος αυτά, όπως ονομάζονται, αναφέρονται γενικά στη θέση που αλλάζει πιο δραματικά στη χωρική - χρονική διάσταση [45]. Οι δημοφιλείς μέθοδοι STIPs περιλαμβάνουν τα



σημεία χωροχρονικών χαρακτηριστικών 3D-Harris [46] και παρόμοιες ή βελτιωμένες τεχνικές [47].

Σημειώνεται ότι τα STIPs είναι συχνά edges ή Harris corners, δηλαδή σημαντικά σημεία της εικόνας που αναπαριστούν τη δράση και αλλάζουν θέση στο χρόνο.



Εικόνα 14: Selective Spatio-Temporal Interest Points σε εικόνα απεικόνισης backhand [48]

### 2.2.2.3 Εξαγωγή σιλουέτας (Silhouette extraction)

Για την ανίχνευση της παρουσίας του ανθρώπου σε ένα βίντεο μέσα από τη σιλουέτα - φιγούρα του, χρησιμοποιούνται τεχνικές κίνησης και αφαίρεσης φόντου, ώστε να διαχωριστούν τα pixels του προσκηνίου από αυτά του φόντου κι έτσι να απομονωθεί η ανθρώπινη σιλουέτα. Οι εξαγόμενες πληροφορίες σιλουέτας μπορούν να χρησιμοποιηθούν για την παρακολούθηση και την ανάλυση κυκλοφορίας, για την παρακολούθηση ανθρώπων που χρειάζονται βοήθεια, για την παρακολούθηση βίντεο, αξιοποιώντας την ταχύτητα και την απλότητα που προσφέρει η μέθοδος αυτή [49].



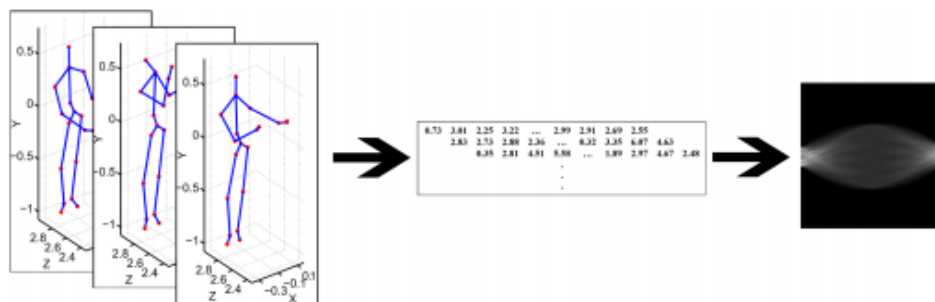
Εικόνα 15: Διαδικασία εξαγωγής σιλουέτας από το φόντο [49]

#### ***2.2.2.4 Άλλες μέθοδοι αναπαράστασης της κίνησης και της εξαγωγής χαρακτηριστικών***

Το 2017 προτάθηκε μία νέα μέθοδος εξαγωγής χαρακτηριστικών για την αναγνώριση της ανθρώπινης δράσης, επεκτείνοντας τις δυνατότητες του μετασχηματισμού Trace στον 3D τομέα. Ορίστηκε η έννοια της τρισδιάστατης μορφής του μετασχηματισμού Trace σε διακριτούς όγκους που εξάγονται από χωροχρονικές ακολουθίες εικόνας. Ο νέος αυτός μετασχηματισμός, που ονομάζεται 3D Cylindrical Trace Transform συνδιάστηκε με τα Selective Spatio-Temporal Interest Points του βίντεο, δημιουργώντας ένα σχήμα εξαγωγής τριπλών χαρακτηριστικών (όπως προκύπτουν από τους διαφόρους μετασχηματισμούς) που ονομάζεται Volumetric Triple Features, το οποίο και καταφέρνει να συλλάβει την πολύτιμη γεωμετρική κατανομή των σημείων ενδιαφέροντος στις χωροχρονικές ακολουθίες εικόνας και να δώσει έμφαση στους γεωμετρικούς τους συσχετισμούς που υποδεικνύουν τη δράση. Η μέθοδος αυτή παρέχει μία πληθώρα εξαγόμενων χαρακτηριστικών, ωστόσο η παραγόμενη αναπαράσταση διατηρείται περιεκτική. Επίσης, προσφέρει αναλλοίωτη παραμόρφωση και χρονικά ευαίσθητα χαρακτηριστικά για την ταξινόμηση των ανθρώπινων ενεργειών. Τα πειράματα σε διαφορετικά απαιτητικά σύνολα δεδομένων αναγνώρισης δράσης παρείχαν εντυπωσιακά αποτελέσματα που δείχνουν την αποτελεσματικότητα του προτεινόμενου μετασχηματισμού και του συνολικού προτεινόμενου σχήματος για το πρόβλημα της ανίχνευσης και αναγνώρισης κίνησης [48].

Το 2020 προτάθηκε μία ακόμη αξιολογη μέθοδος για την αναπαράσταση της κίνησης και την εξαγωγή των χαρακτηριστικών της. Δημιουργήθηκε ένα νέο σχήμα κωδικοποίησης της ακολουθίας της κίνησης που μετατρέπει αποτελεσματικά τις χωροχρονικές αλληλουχίες δράσης σε συμπαγείς αναπαραστάσεις χρησιμοποιώντας χαρακτηριστικά την απόσταση Mahalanobis και το μετασχηματισμό Radon [50]. Η μέθοδος αυτή αξιοποιεί το γεγονός ότι τα υψηλού επιπέδου χαρακτηριστικά ανθρώπινης πόζας και δραστηριότητας, όπως είναι τα χωροχρονικά σημεία ενδιαφέροντος και οι αρθρώσεις, μπορούν να θεωρηθούν ως πιθανοτικές κατανομές στο χωροχρόνο και στο χώρο αντίστοιχα [51]. Οι σκελετικοί σύνδεσμοι θεωρούνται έγκυρα χωροχρονικά σημεία ενδιαφέροντος για την ανάλυση της ανθρώπινης δράσης, γι' αυτό και η σχετική τους θέση με το ανθρώπινο σώμα παίζει καθοριστικό ρόλο

στον καθορισμό της κίνησης. Γι' αυτό το λόγο χρησιμοποιείται η απόσταση Mahalanobis με την οποία υπολογίζεται η απόσταση της θέσης της εκάστοτε άρθρωσης σε σχέση με τη συνολική κατανομή των αρθρώσεων του σώματος. Μια χρονική στιγμή είναι ένα καρέ του βίντεο σε μια ακολουθία δράσης, οπότε αν υπολογιστούν οι αποστάσεις αυτές για όλες τις αρθρώσεις σε όλα τα καρέ, δημιουργείται ο πίνακας Mahalanobis. Στη συνέχεια, ακολουθεί η εφαρμογή του Radon μετασχηματισμού, ώστε να περιγραφεί η ακολουθία δράσης μέχρι μια συγκεκριμένη χρονική στιγμή ως ένα χώροχρονικό χαρακτηριστικό που ονομάζεται Spatio-temporal Radon Footprint (SRFt). Αυτή η αναπαράσταση, τώρα, μπορεί να χρησιμοποιηθεί ως είσοδος σε ένα ελαφρύ συνελκτικό νευρωνικό δίκτυο, οπότε και να συνεισφέρει στον σχηματισμό αποτελεσματικών τεχνικών αναγνώρισης δραστηριότητας με τη χρήση υπερσύγχρονων τεχνικών εκτίμησης της ανθρώπινης στάσης [50].



Εικόνα 16: Διαδικασία εξαγωγής των SRFs [50]

### 2.2.3 Εφαρμογές

Η ανίχνευση και αναγνώριση της ανθρώπινης κίνησης συναντά πολλές εφαρμογές σε διάφορους τομείς της ζωής ενός ανθρώπου. Αξίζει να αναφερθούμε ενδεικτικά σε κάποιες εφαρμογές που χρησιμοποιούν ως μεθόδους αναπαράστασης και εξαγωγής χαρακτηριστικών της κίνησης, αυτές που παρουσιάσαμε παραπάνω.

Το 2017 οι Yu et al. [49], πρότειναν ένα νέο σύστημα ανίχνευσης πτώσης με βάση την μέθοδο εξαγωγής σιλουέτας, το οποίο προσφέρεται στην υγειονομική περίθαλψη για την παρακολούθηση και υποβοήθηση των ηλικιωμένων. Το σύστημα αυτό λαμβάνει στην είσοδο του βίντεο, όπου εφαρμόζεται η αφαίρεση φόντου για να

εξαχθεί η σιλουέτα του ανθρώπινου σώματος. Οι σιλουέτες που εξάγονται χρησιμοποιούνται από ένα συνελκτικό δίκτυο, το οποίο έχει εκπαιδευτεί σε σύνολα δεδομένων σιλουετών σχετικές με καθημερινές ενέργειες και στάσεις του σώματος, Το δίκτυο ταξινομεί τις διάφορες κατηγορίες στις στάσεις του σώματος, όπως κάμψη, πτώση, στάση, κάθισμα, οπότε και μπορεί να ανιχνεύσει ένα συμβάν πτώσης προς αποφυγή.

Το 2018, οι Kurose et al. [52], ανέπτυξαν μία μέθοδο ανάλυσης της κίνησης των παικτών του τένις μέσα στον αγωνιστικό χώρο με σκοπό την κατανόηση και αξιολόγηση λεπτομερειών σχετικά με τα χτυπήματα της μπάλας. Με τη μέθοδο αυτή ερευνάται η στάση του παίκτη στο βίντεο και εκτιμάται η θέση του κάθε αθλητή με χρήση του Open Pose ως τεχνική εκτίμησης πόζας. Δημιουργείται ένα διάνυσμα χαρακτηριστικών χρησιμοποιώντας τις συντεταγμένες των σχετικών θέσεων των αρθρώσεων που εκτιμήθηκαν με αυτή της λεκάνης, ώστε να μην υπάρχει εξάρτηση από το χώρο που βρίσκεται ο παίκτης. Το διάνυσμα αυτό χρησιμοποιείται για την πρόβλεψη της πιθανότητας επιτυχίας του παιχνιδιού μέσω εκπαιδευμένων ταξινομητών μη εποπτευόμενης μάθησης SVM. Η χρησιμότητα αυτής της εφαρμογής έγκειται στην ανάλυση των παιχνιδιών από την πλευρά των παικτών με σκοπό την κατανόηση της κίνησης και της αποτελεσματικότητας σε βολές των αντιπάλων τους, οπότε και αξιολογεί κατάλληλα την παρουσία ενός παίκτη στο γήπεδο.

Το 2020, οι Hendry et al. [53], ανέπτυξαν ένα σύστημα αναγνώρισης ανθρώπινης δραστηριότητας με τη χρήση δεδομένων από αισθητήρες wearables για τον ακριβή εντοπισμό των βασικών κινήσεων μπαλέτου, όπως άλματα, με σκοπό την ποσοτικοποίηση του όγκου της προπόνησης μιας χορεύτριας μπαλέτου και το συσχετισμό της με τους τραυματισμούς που αυτή υφίσταται. Το σύστημα αυτό λαμβάνει δεδομένα από έξι αισθητήρες που τοποθετούνται στο σώμα της χορεύτριας, στη σπονδυλική στήλη, στο κέντρο μάζας, στον αριστερό και δεξιό μοιρό και στην αριστερή και δεξιά κνήμη. Οι προς μελέτη χορεύτριες πραγματοποιούν μεμονωμένες φιγούρες αλμάτων και ανύψωσης ποδιών, οι οποίες και καταγράφονται σε βίντεο. Από αυτά τα δεδομένα προκύπτουν χωροχρονικά χαρακτηριστικά τα οποία εισέρχονται σε κατάλληλα συνελκτικά νευρωνικά δίκτυα, ώστε να γίνει η ταξινόμηση των μπαλετικών κινήσεων.

## ***2.3 Η Τεχνητή Νοημοσύνη στην ιατρική και υγειονομική περίθαλψη***

Ο όρος «τεχνητή νοημοσύνη στην υγειονομική περίθαλψη» περιγράφει τη χρήση αλγορίθμων και λογισμικών μηχανικής μάθησης που έχουν ως σκοπό την μίμηση της ανθρώπινης γνώσης και συμπεριφοράς στην ανάλυση, παρουσίαση και κατανόηση πολύπλοκων και απαιτητικών ιατρικών δεδομένων. Η τεχνητή νοημοσύνη, μέσω αλγορίθμων μηχανικής και βαθιάς μάθησης, συλλέγει και επεξεργάζεται δεδομένα, ώστε να εξάγει χρήσιμες πληροφορίες και κατάλληλα συμπεράσματα πρόβλεψης, τα οποία προσφέρει στον εκάστοτε χρήστη. Τα μοντέλα που χρησιμοποιούνται εκπαιδεύονται σε σύνολα δεδομένων, αναγνωρίζουν μοτίβα συμπεριφοράς μέσα σε αυτά και δημιουργούν τη δική τους λογική για την αντιμετώπιση των διαφόρων προβλημάτων που αναλαμβάνουν να επιλύσουν. Όσον αφορά τις εφαρμογές τεχνητής νοημοσύνης στον τομέα της υγείας, αυτές στοχεύουν στην ανάλυση των σχέσεων μεταξύ τεχνικών πρόληψης ή θεραπείας και αποτελεσμάτων στον ασθενή [54]. Υπάρχει μία πληθώρα τέτοιων εφαρμογών που σχετίζονται με διαδικασίες διάγνωσης, ανάπτυξης πρωτοκόλλου θεραπείας, ανάπτυξη φαρμάκων, εξατομικευμένη ιατρική, παρακολούθηση και φροντίδα ασθενών. Επίσης, οι αλγόριθμοι τεχνητής νοημοσύνης χρησιμοποιούνται για την ανάλυση μεγάλων ποσοτήτων δεδομένων μέσω ηλεκτρονικών αρχείων υγείας για την πρόληψη και τη διάγνωση ασθενειών. Πολλά ιατρικά ιδρύματα, όπως η κλινική Mayo και το κέντρο καρκίνου Memorial Sloan Kettering [55] [56], υπηρεσίες υγείας, ακόμη και μεγάλες εταιρείες τεχνολογίας, όπως η IBM [57] και η Google [58], έχουν αναπτύξει τέτοιου είδους αλγορίθμους. Τέλος, στόχος των σύγχρονων εταιρειών και νοσοκομείων είναι η ανάπτυξη λογισμικών για την όσο το δυνατόν καλύτερη διαχείριση της υγειονομικής περίθαλψης, τόσο από άποψη ασθενών όσο και ανθρώπινου δυναμικού.

Η τεχνητή νοημοσύνη, μέχρι σήμερα, έχει συμβάλει αρκετά στη δημιουργία εφαρμογών που είναι σε θέση να επεξεργαστούν ιατρικά δεδομένα, όπως, τιμές, εικόνες, βίντεο και να εξάγουν από αυτά χαρακτηριστικά, αλλά και αποτελέσματα που βοηθούν στην πρόβλεψη, ανίχνευση και καταπολέμηση μιας ασθένειας. Οι Esteva al. (2017) [59] παρουσίασαν μια υψηλής απόδοσης ταξινόμηση του καρκίνου του δέρματος στα διάφορα είδη του και ειδικά στα πιο θανατηφόρα, χρησιμοποιώντας ένα μεγάλο σύνολο δεδομένων από εικόνες με τη χρήση συνελκτικών δικτύων. Οι

Harmon et al. (2020) [60] παρουσίασαν μια σειρά από αλγορίθμους που επιτυγχάνουν σε πολύ υψηλά ποσοστά την ανίχνευση μέσω αξονικών τομογραφιών θώρακος της νόσου Covid-19. Οι Mirmomeni et al. (2020) [61] παρουσίασαν διάφορα παραδείγματα χρήσης συσκευών wearables από τον άνθρωπο που με τη συνεισφορά της τεχνητής νοημοσύνης έχουν καταλήξει σε έξυπνες συσκευές που μπορούν να συλλέξουν πληροφορίες, να ανιχνεύσουν και να προβλέψουν καταστάσεις, όπως ένα επιληπτικό επεισόδιο, τον κίνδυνο τραυματισμού, το τρέμουλο, τα επίπεδα γλυκόζης στο αίμα κι έτσι να συμβάλλουν στην καλύτερη παρακολούθηση της υγείας των ασθενών.

Τέλος, η χρήση τεχνητής νοημοσύνης στην υγειονομική περίθαλψη αποδίδει εξίσου καλά και πολλές φορές καλύτερα από τον άνθρωπο σε ορισμένες διαδικασίες, ωστόσο δεν είναι σε θέση να τον αντικαταστήσει στο ευρύ φάσμα των ιατρικών εργασιών αλλά να τον βοηθήσει.

## ***2.4 Νόσος του Πάρκινσον***

### ***2.4.1 Η νόσος***

Η νόσος αναφέρεται πρώτη φορά στη βιβλιογραφία το 1817 με τον όρο «Τρομώδης παράλυση» (An Essay on the Shaking Palsy), από τον Άγγλο γιατρό James Parkinson, από τον οποίο πήρε και το όνομά της. Η νόσος του Πάρκινσον επηρεάζει μία περιοχή στο εσωτερικό του εγκεφάλου, τη μέλαινα ουσία (substantia nigra) που αποτελείται από νευρώνες και κατηγοριοποιείται στα βασικά γάγγλια. Η επίδραση της νόσου στη μέλαινα ουσία έχει ως αποτέλεσμα τον εκφυλισμό των νευρώνων που είναι υπεύθυνοι για την παραγωγή μιας συγκεκριμένης ουσίας, της ντοπαμίνης. Η μείωση αυτή στους νευρώνες (εκφυλισμός) γίνεται προοδευτικά, μεταξύ των διαφόρων σταδίων της ασθένειας, με αποτέλεσμα η νόσος να καθίσταται ως χρόνια.

Η ντοπαμίνη είναι η οργανική χημική ουσία που παράγουν οι νευρώνες της μέλαινας ουσίας. Η ουσία αυτή λειτουργεί ως νευροδιαβιβαστής, γιατί βοηθά τα

νευρικά κύτταρα να επικοινωνούν μεταξύ τους κι άρα συμβάλλει στη μεταφορά πληροφοριών σε όλα τα μέρη του εγκεφάλου. Η ντοπαμίνη παίζει καθοριστικό ρόλο στον συντονισμό των κινήσεων, αφού μεταφέρει τα νευρικά μηνύματα και συνδέει τη μέλαινα ουσία με το ραβδωτό σώμα του εγκεφάλου, το οποίο είναι υπεύθυνο για την ομαλότητα, τον συντονισμό και τον έλεγχο των κινήσεων. Συνεπώς, ο εκφυλισμός των νευρώνων της μέλαινας ουσίας οδηγεί στη μειωμένη κι άρα ανεπαρκή έκκριση ντοπαμίνης που έχει ως αποτέλεσμα τον περιορισμένο έλεγχο και συντονισμό κινήσεων του σώματος. Ωστόσο, τα συμπτώματα της νόσου εμφανίζονται με την καταστροφή μεγάλου ποσοστού, περίπου 60% με 80%, των υπεύθυνων κυττάρων [62].

Σημειώνεται ότι η νόσος του Πάρκινσον είναι η πιο συχνή νευροεκφυλιστική ασθένεια μετά το Αλτσχάιμερ και εμφανίζεται, κυρίως, σε άτομα άνω των 60 ετών. Το ποσοστό των νοσούντων ανά τον κόσμο αυξάνεται συνεχώς, δεδομένης και της γήρανσης του πληθυσμού.

#### ***2.4.2 Τα συμπτώματα***

Τα κυριότερα και πιο γνωστά συμπτώματα της νόσου είναι:

Τρέμουλο (Τρόμος): ένα μέρος του σώματος του ασθενούς, όπως πόδι, χέρι, κεφάλι, αρχίζει να τρέμει.

Άκαμπτοι μυς: οι μυς του ασθενή, κυρίως αυτοί που είναι υπεύθυνοι για την κίνηση, είναι σφιχτοί και δεν μπορούν να συσπαθούν με αποτέλεσμα ο ίδιος να μη μπορεί να κινηθεί.

Αργή κίνηση: ο ασθενής δυσκολεύεται να κινηθεί και προσπαθεί να βαδίζει σέρνοντας τα πόδια του.

Ισορροπία: ο ασθενής δυσκολεύεται να παραμείνει όρθιος κι άρα να διατηρήσει την ισορροπία του.

Κούραση ή εξάντληση: ο ασθενής νιώθει κόπωση χωρίς να έχει προηγηθεί κάποιο εξαντλητική δραστηριότητα.

Λύπη ή κατάθλιψη: τα συναισθήματα αυτά εμφανίζονται ως επιπλέον συμπτώματα που απορρέουν από την όλη κατάσταση του ασθενή και την αδυναμία που αισθάνεται.

Δυσκοιλιότητα: η λειτουργία του εντέρου εξαρτάται από τη λειτουργία του νευρικού συστήματος, το οποίο δυσλειτουργεί κι άρα οδηγεί σε τέτοια συμπτώματα.

Δυσκολία στην ομιλία ή στην κατάποση τροφής ή ποτού: σε πιο προχωρημένο στάδιο επηρεάζονται οι λειτουργίες της μάσησης, της κατάποσης, αλλά και της ομιλίας.

Προβλήματα στην εκτέλεση περισσότερων του ενός πραγμάτων ταυτόχρονα: η αδυναμία των νευρικών κυττάρων να επεξεργαστούν επαρκώς τις πληροφορίες που λαμβάνουν δυσχεραίνει την εκτέλεση πολλαπλών εντολών [63].

Τα συμπτώματα της νόσου Πάρκινσον εμφανίζονται σταδιακά στον ασθενή και διαφέρουν από τον έναν στον άλλο. Με βάση, λοιπόν, τη βαρύτητα των συμπτωμάτων έχουν χαρακτηριστεί ορισμένα στάδια της νόσου, με σκοπό τη βελτιστοποίηση των προσφερομένων ιατρικών υπηρεσιών [64].

Υπάρχουν πέντε στάδια στη νόσο Πάρκινσον [62], τα ακόλουθα:

Στάδιο 0: Οι διαγνωστικές εξετάσεις εμφανίζουν ευρήματα, ωστόσο δεν εμφανίζεται κανένα σύμπτωμα της νόσου.

Στάδιο 1: Εμφανίζονται τα πρώτα συμπτώματα της νόσου. Σε αυτό το στάδιο επηρεάζεται μόνο η μια πλευρά του σώματος του ασθενούς.

Στάδιο 2: Η ασθένεια έχει πλέον εξελιχθεί και έχουν προσβληθεί και οι δύο πλευρές του σώματος με αποτέλεσμα να δημιουργούνται προβλήματα στο βάδισμα του ασθενούς.

Στάδιο 3: Ο ασθενής δυσκολεύεται όλο και περισσότερο να βαδίσει, να σταθεί όρθιος και χάνει συχνά την ισορροπία του.

Στάδιο 4: Ο ασθενής δεν μπορεί πλέον να βαδίσει χωρίς υποστήριξη.

Στάδιο 5: Ο ασθενής δεν μπορεί πλέον ούτε να βαδίσει, αλλά ούτε και να σταθεί κι άρα βρίσκεται καθηλωμένος είτε σε αναπηρικό αμαξίδιο είτε κλινήρης



Η εξέλιξη της νόσου Πάρκινσον γίνεται σχετικά αργά, ωστόσο τα βαριά συμπτώματα της νόσου εμφανίζονται μετά από πολλά χρόνια με αποτέλεσμα να προκαλούνται πολλά προβλήματα στη ζωή του ασθενούς [62]. Όσον αφορά την αντιμετώπιση της νόσου, η επιστημονική έρευνα έχει συμβάλει σημαντικά, δεδομένου ότι δεν υπάρχει συγκεκριμένος τρόπος διάγνωσης, καθώς και μία ορισμένη θεραπεία [64].

### **2.4.3 Διάγνωση και θεραπεία**

Η νόσος του Πάρκινσον δύσκολα ανιχνεύεται στα πρώιμα στάδια της, καθώς τα συμπτώματα που συνοδεύουν την εμφάνιση της δεν αφορούν το κομμάτι της κίνησης και μοιάζουν αρκετά με αυτά άλλων ασθενειών, όπως του Αλτσχάιμερ. Η νόσος εξελίσσεται και τα προβλήματα στην κίνηση αρχίζουν να παρουσιάζονται [65]. Αυτό σημαίνει ότι έχει εκφυλιστεί ήδη ένα πολύ μεγάλο ποσοστό νευρώνων της μέλαινας ουσίας, οπότε και η καταστολή της ασθένειας είναι ιδιαίτερης σημασίας.

Η διάγνωση της νόσου επιτυγχάνεται μέσω κλινικών εκτιμήσεων, παρατηρώντας τόσο τα κινητικά όσο και τα μη κινητικά συμπτώματα [66]. Δεν υπάρχουν ασφαλείς μέθοδοι εργαστηριακών βιοδεικτών και νευροαπεικόνισης, οπότε και η ασθένεια έχει εξελιχθεί αρκετά και τα συμπτώματα της είναι εμφανή μέχρι να μπορέσει να ανιχνευτεί με μεγάλα ποσοστά ακριβείας. Γι' αυτό το λόγο το 1987 εδραιώθηκε η Ενοποιημένη Κλίμακα Βαθμολόγησης της νόσου Πάρκινσον (Unified Parkinson's Disease Rating Scale - UPDRS) με στόχο την επιτυχή διάγνωση της νόσου από εξειδικευμένο ιατρικό προσωπικό λαμβάνοντας υπόψη ένα μεγάλο σύνολο κριτηρίων [67]. Η βαθμολόγηση στην κλίμακα αυτή προκύπτει από την αξιολόγηση ενός συνόλου δραστηριοτήτων και δοκιμών που τίθεται ο ασθενής και μπορούν να δώσουν μία ξεκάθαρη εικόνα της σοβαρότητας της νόσου [68]. Η αξιολόγηση αποτελείται από τέσσερις ενότητες και συνολικά πενήντα ερωτήσεις πολλαπλής επιλογής συσχετιζόμενες με τα κινητικά και τα μη κινητικά συμπτώματα της νόσου [69]. Οι δύο πρώτες ενότητες συμπληρώνονται από τον εξεταζόμενο και οι υπόλοιπες δύο από τον ειδικό και αφορούν κατά σειρά:

- Εμπειρίες της καθημερινής ζωής του εξεταζόμενου
- Κινητικές εμπειρίες της καθημερινής ζωής του εξεταζόμενου

- Εξέταση κινήσεων
- Επιπλοκές που εμφανίζονται κατά τις διάφορες κινήσεις.

Ακολουθεί το φύλλο βαθμολογίας στην κλίμακα UPDRS των δραστηριοτήτων που απαρτίζουν τα παραπάνω τμήματα αξιολόγησης.

Πίνακας 1: Φύλλο βαθμολογίας κλίμακας MDS-UPDRS [69]

1.A	Πηγή πληροφοριών	<input type="checkbox"/> Ασθενής <input type="checkbox"/> Φροντιστής <input type="checkbox"/> Ασθενής + Φροντιστής	3.3b	Δυσκαμψία ΔΑΑ	
			3.3c	Δυσκαμψία ΑΑΑ	
<b>Τμήμα I</b>			3.3d	Δυσκαμψία ΔΚΑ	
1.1	Γνωστική διαταραχή		3.3e	Δυσκαμψία ΑΚΑ	
1.2	Ψευδαισθήσεις και ψύχωση		3.4a	Πλήξη δακτύλων χεριού ΔΕ	
1.3	Καταθλιπτική διάθεση		3.4b	Πλήξη δακτύλων χεριού ΑΡ	
1.4	Αγχώδης διάθεση		3.5a	Κινήσεις χεριών ΔΕ	
1.5	Απάθεια		3.5b	Κινήσεις χεριών ΑΡ	
1.6	Χαρακτηριστικά συνδρόμου Ntop_Απορρύθμισης		3.6a	Κινήσεις πρηνισμού-υπτιασμού - ΔΕ χέρι	
1.6a	Ποιος συμπληρώνει το ερωτηματολόγιο	<input type="checkbox"/> Ασθενής <input type="checkbox"/> Φροντιστής <input type="checkbox"/> Ασθενής + Φροντιστής	3.6b	Κινήσεις πρηνισμού-υπτιασμού-ΑΡ χέρι	
			3.7a	Πλήξη δακτύλων ποδιού ΔΕ	
1.7	Προβλήματα ύπνου		3.7b	Πλήξη δακτύλων ποδιού ΑΡ	
1.8	Ημερήσια υπνηλία		3.8a	Ευκινησία κάτω άκρου ΔΕ	
1.9	Πόνος & άλλες αισθήσεις		3.8b	Ευκινησία κάτω άκρου ΑΡ	
1.10	Ουρολογικά προβλήματα		3.9	Έγερση από καρέκλα	
1.11	Προβλήματα δυσκολότητας		3.10	Βάδισμα	
1.12	Ζάλη στην όρθια στάση		3.11	Πάγωμα στη βάδιση	
1.13	Κόπωση		3.12	Στατική ισορροπία	
<b>Τμήμα II</b>			3.13	Στάση	
2.1	Ομιλία		3.14	Βραδυκίνησία σώματος	
2.2	Σάλιο & σιελορροια		3.15a	Τρόμος θέσης χεριών ΔΕ	
2.3	Μάσηση και κατάποση		3.15b	Τρόμος θέσης χεριών ΑΡ	
2.4	Διαδικασία φαγητού		3.16a	Κινητικός τρόμος χεριών ΔΕ	
2.5	Ντύσιμο		3.16b	Κινητικός τρόμος χεριών ΑΡ	
2.6	Υγιεινή		3.17a	Εύρος τρόμου ηρεμίας ΔΑΑ	
2.7	Γραφή		3.17b	Εύρος τρόμου ηρεμίας ΑΑΑ	
2.8	Χόμπι & άλλες δραστηριότητες		3.17c	Εύρος τρόμου ηρεμίας ΔΚΑ	
2.9	Γύρισμα στο κρεβάτι		3.17d	Εύρος τρόμου ηρεμίας ΑΚΑ	
2.10	Τρέμουλο		3.17e	Εύρος τρόμου ηρεμίας/ χειλη-γνάθος	
2.11	Σήκωμα από το κρεβάτι		3.18	Διάρκεια τρόμου ηρεμίας	
2.12	Βάδισμα & ισορροπία			Υπήρχαν δυσκινησίες	<input type="checkbox"/> Όχι <input type="checkbox"/> Ναι
2.13	Πάγωμα			Οι κινήσεις αυτές επηρέασαν τη βαθμολογία;	<input type="checkbox"/> Όχι <input type="checkbox"/> Ναι
3a	Είναι ο ασθενής σε αγωγή;	<input type="checkbox"/> Όχι <input type="checkbox"/> Ναι		Στάδιο Hoehn & Yahr	
3b	Κλινική κατάσταση του ασθενή	<input type="checkbox"/> off <input type="checkbox"/> on	<b>Τμήμα IV</b>		
3c	Παίρνει ο ασθενής λεβοντόπα;	<input type="checkbox"/> Όχι <input type="checkbox"/> Ναι	4.1	Χρονικό διάστημα με δυσκινησίες	
3C1	Αν ναι, πόσα λεπτά από την τελευταία δόση;		4.2	Λειτουργική επίπτωση δυσκινησιών	
<b>Τμήμα III</b>			4.3	Χρονικό διάστημα σε κατάσταση OFF	
3.1	Ομιλία		4.4	Λειτουργική επίπτωση των διακυμάνσεων	
3.2	Έκφραση προσώπου		4.5	Πολυπλοκότητα των διακυμάνσεων	
3.3a	Δυσκαμψία-Αυχένας		4.6	Επώδυνη δυστονία στην κατάσταση OFF	

Ανάλογα με τις απαντήσεις του ερωτηματολογίου τα αποτελέσματα ερμηνεύονται σύμφωνα με τους παρακάτω δύο πίνακες, οι οποίοι βασίζονται στο μοντέλο σταδίων νόσου Πάρκινσον Hoehn και Yahr [70] και στην κλίμακα καθημερινών δραστηριοτήτων των Schwab και England [71] αντίστοιχα:

**Πίνακας 2: Μοντέλο σταδίων νόσου Πάρκινσον Hoehn και Yahr [70]**

Στάδιο ασθένειας	Ερμηνεία
Στάδιο 0	Κανένα σημάδι της νόσου.
Στάδιο 1	Μονομερής ασθένεια.
Στάδιο 1.5	Μονομερής ασθένεια συν αξονική εμπλοκή.
Στάδιο 2	Διμερής ασθένεια, χωρίς εξασθένιση ισορροπίας.
Στάδιο 2.5	Ήπια αμφίπλευρη ασθένεια, με ανάκτηση στη δοκιμή έλξης.
Στάδιο 3	Ήπια έως μέτρια αμφίπλευρη ασθένεια. Μερική αστάθεια, ανεξάρτητος.
Στάδιο 4	Σοβαρή αναπηρία. Ακόμα σε θέση να περπατήσει ή να σταθεί χωρίς βοήθεια.
Στάδιο 5	Χρήση αναπηρικού αμαξιδίου ή ξαπλωμένος στο κρεβάτι, εκτός αν υποβοηθείται.

**Πίνακας 3: Κλίμακα καθημερινών δραστηριοτήτων Schwab και England [71]**

Ποσοστό	Ερμηνεία
100%	Εντελώς ανεξάρτητος. Είναι σε θέση να κάνει όλες τις δουλειές χωρίς βραδύτητα ή δυσκολία. Ουσιαστικά φυσιολογικός. Δεν αντιλαμβάνεται κάποιο κώλυμα.
90%	Εντελώς ανεξάρτητος. Είναι σε θέση να κάνει όλες τις δουλειές με κάποιο βαθμό βραδύτητας, δυσκολίας και δυσλειτουργίας. Ίσως να καταλαβαίνει κάποια δυσκολία.
80%	Πλήρως ανεξάρτητος στις περισσότερες δουλειές. Πολλές διαδικασίες διαρκούν διπλάσιο χρονικό διάστημα. Έχει επίγνωση της δυσκολίας και της βραδύτητας.
70%	Δεν είναι εντελώς ανεξάρτητος. Περισσότερες δυσκολίες με κάποιες δραστηριότητες.
60%	Κάποια εξάρτηση. Μπορεί να κάνει τις περισσότερες δουλειές, αλλά εξαιρετικά αργά και με μεγάλη προσπάθεια.
50%	Περισσότερο εξαρτημένος. Δυσκολία σε όλες τις δραστηριότητες.
40%	Πολύ εξαρτημένος. Μπορεί να βοηθήσει μόνο σε κάποιες δουλειές.
30%	Χρειάζεται προσπάθεια για να πραγματοποιήσει λίγες μόνο δουλειές. Απαιτείται μεγάλη βοήθεια.
20%	Δεν μπορεί να κάνει τίποτα μόνος του. Μπορεί να προσφέρει μια μικρή βοήθεια σε ορισμένες δουλειές. Κατάσταση σοβαρή.
10%	Εντελώς εξαρτημένος, ανήμπορος.
0%	Λειτουργίες όπως της κατάποσης, ουροδόχου κύστης και εντέρου δεν λειτουργούν. Κλινήςης.

Για την αντιμετώπιση των συμπτωμάτων της νόσου χρησιμοποιείται ως φάρμακο η λεβοντόπα, η οποία αποτελεί πρόδρομο μόριο της ντοπαμίνης. Κατά τα αρχικά στάδια της νόσου, μετατρέπεται από τους εναπομείναντες νευρώνες στη μέλαινα ουσία, σε ντοπαμίνη, καλύπτοντας την έλλειψη της ουσίας αυτής, λόγω της ασθένειας. Η χορήγηση της λεβοντόπας περιορίζει τα συμπτώματα της νόσου, ωστόσο με την πάροδο των ετών παύει να δρα ή δρα σε πολύ μικρότερο βαθμό, αφού έχει μειωθεί δραματικά ο αριθμός των εναπομείναντων νευρώνων στη μέλαινα ουσία

που μπορούν να την μετατρέψουν σε ντοπαμίνη. Αυτό έχει ως αποτέλεσμα την εμφάνιση διαταραχών στους ασθενείς είτε στην κίνηση, είτε στην νόηση, οι οποίες αξιολογούνται μέσω κατάλληλης βαθμολογικής κλίμακας για το επίπεδο σοβαρότητας τους [72].

Έτσι, οι ασθενείς με Πάρκινσον δεν υποβάλλονται μονάχα σε δραστηριότητες με σκοπό την παρατήρηση και αξιολόγηση των συμπτωμάτων της νόσου. Συχνά εξετάζονται και για τις παρενέργειες που εμφανίζονται εξαιτίας της μακροχρόνιας φαρμακευτικής αγωγής τους. Για το λόγο αυτό χρησιμοποιείται η μετρική αξιολόγησης Unified Dyskinesia Rating Scale (UDysRS). Ωστόσο, στην κλίμακα αυτή συνυπολογίζεται και η δυσκινησία που προκαλείται σε περιπτώσεις που δεν λαμβάνονται συστηματικά τα φάρμακα ή αυτά δεν δρουν.

Πιο συγκεκριμένα, με τη κλίμακα αυτή αξιολογούνται οι ακούσιες κινήσεις που εμφανίζονται σε ασθενείς υπό θεραπεία. Η αξιολόγηση χωρίζεται σε δύο τμήματα: Ιστορικό (Τμήμα 1 και 2) και Αντικειμενικό (Τμήμα 3 και 4). Το 1<sup>ο</sup> τμήμα περιέχει την αξιολόγηση από τον εξεταστή των απαντήσεων που δίνει ο ασθενής σχετικά με την εμφάνιση τιναγμάτων ή περιστροφικών κινήσεων όταν λαμβάνει τα φαρμακά του και αποκαλείται On-δυσκινησία. Το 2<sup>ο</sup> τμήμα περιέχει την αξιολόγηση των απαντήσεων των ασθενών σχετικά με την εμφάνιση σπασμών και κραμπών όταν δεν λαμβάνουν τα φάρμακά τους ή αυτά δεν δρουν και αποκαλείται Off-δυστονία. Το 3<sup>ο</sup> τμήμα περιλαμβάνει την αντικειμενική αξιολόγηση της διαταραχής με τη βαθμολόγηση συνολικά της έντασης της δυσκινησίας τόσο κατά τη δυσκινησία όσο και κατά τη δυστονία. Το 4<sup>ο</sup> τμήμα περιλαμβάνει την αξιολόγηση της αναπηρίας που οφείλεται στις λειτουργικές επιπτώσεις της δυσκινησίας σε κάθε περίπτωση. Στα δύο τελευταία τμήματα η αξιολόγηση γίνεται από την παρακολούθηση του ασθενή να εκτελεί ορισμένες δραστηριότητες που ενδείκνυνται για την εξαγωγή σχετικών συμπερασμάτων [73]. Σημειώνεται ότι καθόλη τη διάρκεια της αξιολόγησης δεν πρέπει να λαμβάνονται υπόψη οι διαταραχές στην κίνηση που οφείλονται στην ασθένεια παρά μόνο όσες σχετίζονται με τη θεραπεία.

Παρατίθεται ο πίνακας βαθμολόγησης για τα δύο αυτά κύρια τμήματα που ορίζουν τη βαθμολογία στην κλίμακα UDysRS.

Πίνακας 4: Πίνακας βαθμολογίας UDysRS [73]

Ιστορικό	Βαθμός	Αντικειμενική αξιολόγηση	Βαθμός
1. Χρόνος με δυσκινησία		16. Πρόσωπο	
2. Ομιλία		17. Αυχένιας	
3. Μάσηση / κατάποση		18. Δεξιό χέρι / βραχίονας/ ώμος	
4. Διαδικασία φαγητού		19. Αριστερό χέρι /βραχίονας/ ώμος	
5. Ενδυσση		20. Κορμός	
6. Προσωπική υγιεινή		21. Δεξιό πόδι / μηρός/ ισχίο	
7. Γραφή		22. Αριστερό πόδι/ μηρός/ ισχίο	
8. Χόμπι/Δραστηριότητες		23. Επικοινωνία	
9. Βάδιση / Ισορροπία		24. Πόση	
10. Δημόσιες / κοινωνικές περιστάσεις		25. Ενδυσση	
11. Διεγερτικές καταστάσεις		26. Μετακίνηση	
12. Χρόνος με OFF δυστονία			
13. Επίδραση δυστονίας στις δραστηριότητες (όχι πόνος)			
14. Επίδραση του δυστονικού πόνου			
15. Βαρύτητα δυστονικού πόνου			
Υποσύνολο Ιστορικού (άθροισμα)		Υποσύνολο Αντικειμενικής Αξιολόγησης (άθροισμα)	
Συνολική UDysRS Βαθμολογία (Ιστορικό + Αντικειμενική αξιολόγηση)			

Η εξέλιξη της ιατρικής έχει οδηγήσει στην εύρεση νέων μεθόδων διάγνωσης της νόσου Πάρκινσον με τη βοήθεια απεικόνισης της περιοχής του εγκεφάλου που παράγεται η ντοπαμίνη, δηλαδή των νευρώνων της μέλαινας ουσίας. Οι απεικονίσεις αυτές πραγματοποιούνται με αξονικούς και μαγνητικούς τομογράφους και διακρίνονται σε: α) τομογραφία εκπομπής ποζιτρονίων (PET), β) υπολογιστική τομογραφία εκπομπής απλών φωτονίων (SPECT), γ) διακρανιακή υπερηχογραφία (TCS) του εγκεφαλικού παρεγχύματος. Τα σημαντικότερα μειονεκτήματα των μεθόδων αυτών είναι η μη έγκαιρη διάγνωση της νόσου Πάρκινσον, το υψηλό κόστος και τα ανακριβή αποτελέσματα [63]. Οι εξετάσεις αυτές δίνουν υψηλής ακρίβειας διάγνωση και διαχωρίζουν επαρκώς τους υγιείς ασθενείς, τους ασθενείς με ιδιοπαθή τρόπο κι αυτούς που πάσχουν από άλλες νευροεκφυλιστικές νόσους από αυτούς που νοσούν ακόμη κι αν βρίσκονται σε πρώιμο στάδιο της νόσου [74] [75].

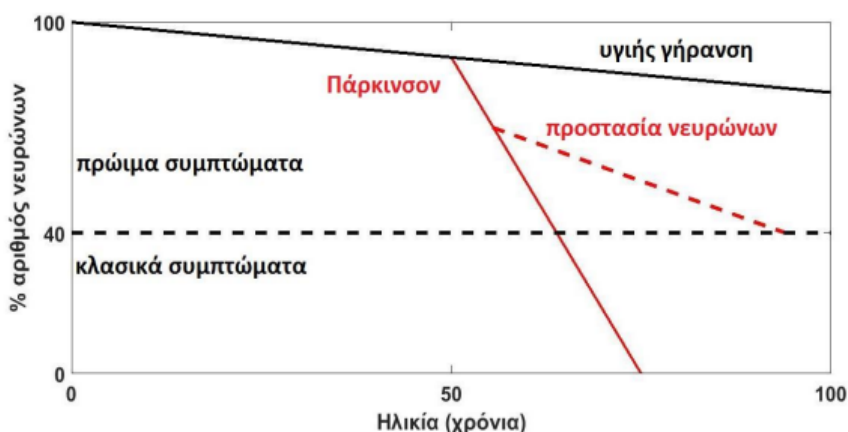
Αξίζει να σημειωθεί ότι μία κλασική μέθοδος διάγνωσης του Πάρκινσον είναι αυτή της εξέτασης της ομιλίας των ασθενών μέσα από συγκεκριμένες καταγραφές που αναλύονται για την εξαγωγή των χαρακτηριστικών της φωνής [76]. Οι ασθενείς εμφανίζουν προβλήματα στην ομιλία, οπότε και μία τέτοια εξέταση είναι αντιπροσωπευτική. Στις μέρες μας, έχει γίνει πρόοδος στον τομέα αυτό και έχουν αναπτυχθεί μέθοδοι με τη χρήση των υπολογιστών και της μηχανικής μάθησης, ώστε

να αυξηθεί η πολυπλοκότητα και οι δυνατότητες ανάλυσης των εγγραφών φωνής προς διάγνωση της νόσου.

#### 2.4.4 Η σημασία της έγκαιρης και έγκυρης διάγνωσης

Η κλινική διάγνωση του Πάρκινσον είναι ιδιαίτερα δύσκολη στα πρώτα στάδια της ασθένειας, δεδομένου ότι τα συμπτώματα δεν είναι απόλυτα εμφανή και ξεκάθαρα. Η ακρίβεια διάγνωσης σε προχωρημένα στάδια είναι περίπου 98%, ενώ σε αρχικά 67%.

Οι ασθενείς, τις περισσότερες φορές, παραβλέπουν τα αρχικά, ήπια κινητικά και μη συμπτώματα, με αποτέλεσμα η νόσος να εξελίσσεται κι ο εκφυλισμός των νευρώνων να έχει προχωρήσει σημαντικά. Στόχος της έγκαιρης διάγνωσης είναι η χορήγηση της κατάλληλης φαρμακευτικής αγωγής, ώστε να προστατευτούν οι υγιείς νευρώνες, αφού οι εκφυλισμένοι δεν αναγεννιούνται. Επομένως, η διάγνωση σε πρώιμο στάδιο είναι πολύ σημαντική για να επιβραδύνθει η εξέλιξη της νόσου και η εμφάνιση των σοβαρότερων συμπτωμάτων της, ώστε να εξασφαλιστεί μία καλύτερη ποιότητα ζωής στον νοσούντα.



Εικόνα 17: Διάγραμμα ποσοστιαίας μείωσης των νευρώνων ντοπαμίνης σε συνάρτηση με την ηλικία σε υγιή άτομα και άτομα με Πάρκινσον [77]

Για να γίνει πιο κατανοητό, στο παραπάνω διάγραμμα φαίνεται η καμπύλη ποσοστιαίας μείωσης των νευρώνων που σχετίζονται με την ασθένεια για υγιή άτομα και παρκινσονικούς ασθενείς αντίστοιχα. Τα εντονότερα συμπτώματα της ασθένειας

προκύπτουν όταν το ποσοστό των κυττάρων αυτών πέσει κάτω από 40%. Εάν όμως επιτευχθεί διάγνωση πιο νωρίς, η νευροπροστασία που παρέχουν τα φάρμακα μπορεί να προσφέρει σημαντικά οφέλη στην ποιότητα ζωής του ασθενή.

Συνεπώς, η τεχνητή νοημοσύνη μπορεί και πρέπει να συμβάλει σημαντικά στην έρευνα μεθόδων ανίχνευσης του Πάρκινσον, έτσι ώστε η διάγνωση της νόσου να γίνεται με μεγαλύτερη ακρίβεια και σε συντομότερο χρονικό διάστημα και οι γιατροί να υποβοηθούνται στη λήψη καλύτερων αποφάσεων σχετικά με τη φαρμακευτική αγωγή των νοσούντων [78].

## ***2.5 Η Νόσος Πάρκινσον στην Επιστήμη των Υπολογιστών***

Η νόσος Πάρκινσον με τα συμπτώματά της στα διάφορα στάδια της προκαλεί σοβαρά προβλήματα στην καθημερινότητα του ασθενή, δυσχεραίνοντας και υποβαθμίζοντας, πολλές φορές, την ποιότητα της ζωής του. Γι' αυτό τον λόγο και δεδομένου ότι δεν υπάρχει ακριβής θεραπεία της νόσου, έχουν αναπτυχθεί συστήματα υποβοηθούμενα από τους υπολογιστές με σκοπό την γρήγορη και αποτελεσματική διάγνωση της νόσου και κατά συνέπεια τη βελτίωση της ποιότητας ζωής των νοσούντων.

Σημαντική είναι η συνεισφορά των αισθητήρων στην προσπάθεια αυτή. Κατά καιρούς έχουν δημοσιευθεί σχετικές μελέτες που παρουσιάζουν μια πληθώρα τέτοιων εφαρμογών [79] [80]. Ενδεικτικά θα αναφερθούμε σε ορισμένες σχετικές δημοσιευμένες εφαρμογές που εξυπηρετούν στην ανίχνευση, την ανάλυση ή και την υποβοήθηση στο πλαίσιο της ασθένειας.

Το 2015, οι Tay et al. [81] ανέπτυξαν ένα φορητό ασύρματο σύστημα παρακολούθησης και βιοανάδρασης με σκοπό την ανίχνευση και αντιμετώπιση του παγώματος στο βάδισμα του ασθενή (Freezing of Gait –FoG). Οι φορητές συσκευές του συστήματος αυτού αποτελούνται από χαμηλού κόστους αισθητήρες, όπως επιταχυνσιόμετρα, γυροσκόπια, πυξίδες κι άλλους ευέλικτους αισθητήρες, καθώς και συνοδευτικά συστήματα επικοινωνίας μέσω Bluetooth και wi-fi για τη μετάδοση των δεδομένων λήψης ασύρματα σε έναν υπολογιστή. Το σύστημα αυτό επεξεργάζεται τα

δεδομένα που προέρχονται από αισθητήρες σε πραγματικό χρόνο κάνοντας ανάλυση του χρονικού βάδισης και ανιχνεύει συμβάντα FoG, οπότε και ενεργοποιεί τη βιοανάδραση ήχου και δόνησης για να αποτρέψει ή να μειώσει το πάγωμα (freezing) όταν έχει συμβεί FoG. Έτσι, οι ασθενείς με Πάρκινσον θα γνωρίζουν πότε διατρέχουν κίνδυνο πτώσης και θα επωφεληθούν από τις κατάλληλες ενδείξεις που θα τους κατευθύνουν ώστε να ακολουθήσουν τα βήματά τους μετά την εμφάνιση ενός FoG. Το σύστημα αυτό είναι κινητό και hands-free και επιτρέπει στον ασθενή να περπατά ελεύθερα για μεγάλα χρονικά διαστήματα και μεγάλες αποστάσεις.

Σε παρόμοιο πλαίσιο κινήθηκαν και οι Lorenzi et al. (2015) [82] που πρότειναν ένα ασύρματο σύστημα ανίχνευσης με ακουστικά που βασίζεται σε αδρανειακές μονάδες μέτρησης που έχουν σχεδιαστεί για μακροχρόνια παρακολούθηση συγκεκριμένων κινητικών διαταραχών. Το σύστημα αποτελείται από έναν μόνο αδρανειακό αισθητήρα που τοποθετείται πλευρικά στο κεφάλι, κοντά στο αυτί και εντοπίζει με μεγάλη ευαισθησία τα σήματα που σχετίζονται με ταλαντώσεις του κορμού. Για την ταυτοποίηση και ανάλυση των χαρακτηριστικών από τα δεδομένα που λαμβάνονται χρησιμοποιούνται κατάλληλα νευρωνικά δίκτυα. Έτσι, επιτυγχάνει την έγκαιρη ανίχνευση του παγώματος του βηματισμού (FoG), καθώς και την έγκαιρη ακουστική διέγερση στο αυτί του ασθενή με αποτέλεσμα να μειώνεται σε πολύ μεγάλο ποσοστό ο κίνδυνος πτώσης του.

Την ίδια χρονιά (2015), επίσης, οι Dong et al. [83] πρότειναν ένα ασύρματο δίκτυο αισθητήρων σώματος για την παρατήρηση και λήψη μετρήσεων σχετικά με τη δραστηριότητα των ατόμων με Πάρκινσον προκειμένου να μελετηθεί η αυθόρμητη κίνησή τους σε ένα μη παρατηρούμενο περιβάλλον. Οι μικροσκοπικοί αισθητήρες σώματος που συνδέονται με τα κάτω άκρα των ασθενών συλλέγουν δεδομένα θέσης και επιτάχυνσης με περιοδικό τρόπο και τα μεταδίδουν σε έναν κόμβο επεξεργασίας και αποθήκευσης που μπορεί να αποθηκεύσει δεδομένα και να μεταφέρει τις απαραίτητες πληροφορίες στον γιατρό μέσω δικτύου τηλεπικοινωνιών ή ασύρματου τοπικού δικτύου. Για να μετρηθούν οι αλλαγές στάσης, χρησιμοποιήθηκαν αισθητήρες κάμψης χαμηλής τάσης, αισθητήρες πίεσης, επιταχυνσιόμετρα και γυροσκόπια. Έτσι, δημιουργήθηκε μια δικτυωμένη συσκευή αισθητήρων που μπορεί να φορεθεί από ένα άτομο για να ανιχνεύσει αλλαγές στη στάση και τη βάδιση.



Το 2017, οι Eskofier et al. [84] εισήγαγαν τα έξυπνα παπούτσια που έχουν τη δυνατότητα πρόληψης, διάγνωσης, παρακολούθησης και λήψης θεραπευτικών αποφάσεων για τη νόσο με συνεχή αξιολόγηση του βηματισμού και της κινητικότητας του ασθενή. Αποτελούν φορητά υπολογιστικά συστήματα που έχουν την ικανότητα συλλογής μέσω αισθητήρων των απαραίτητων δεδομένων βάδισης που θα σταλούν στο cloud με σκοπό τη διαχείριση και αξιολόγησή τους. Επιπλέον, αποτελούν ένα σύστημα χαμηλής ισχύος κι άρα κατανάλωσης ενέργειας και αυξημένης αποδοτικότητας. Τα έξυπνα παπούτσια διαθέτουν αισθητήρες προσδιορισμού θέσης και προσανατολισμού, όπως επιταχυνσιόμετρα, γυροσκόπια, μαγνητόμετρα για την ανάλυση της βάδισης, δορυφορικά συστήματα πλοήγησης, όπως Gps, Galileo, που δίνουν χρονικές πληροφορίες και συμβάλλουν στην παρακολούθηση της κίνησης και αισθητήρες πίεσης με σκοπό την καταμέτρηση της πελματικής πίεσης. Τα δεδομένα αποθηκεύονται τοπικά και στη συνέχεια με τη χρήση Bluetooth και WiFi μεταδίδονται στο cloud για περαιτέρω επεξεργασία.

Το 2019, ο Almogren [85], παρουσίασαν ένα σύστημα απομακρυσμένης παρακολούθησης για τους ασθενείς με νόσο του Πάρκινσον. Το προτεινόμενο σύστημα συγκεντρώνει διάφορες πληροφορίες, όπως δείγματα φωνής, πληροφορίες βάδισης και ενισχύει την παρακολούθηση των συμπτωμάτων πρώιμων σταδίων της νόσου στο σπίτι. Αξιολογεί, για παράδειγμα, τις φωνητικές διαταραχές των ασθενών ή τη δυσφωνία και τη δυσκολία στην κίνηση κι έτσι επιτρέπει στους γιατρούς να εντοπίσουν τα συμπτώματα των ασθενών και το επίπεδο σοβαρότητας της νόσου. Χρησιμοποιούνται διάφορες τεχνολογίες φορητών αισθητήρων, φορητό υπολογιστικό σύστημα, διαδίκτυο και τεχνολογίες cloud computing. Έτσι, εγγυάται την αποτελεσματική χρήση των δεδομένων που συλλέγονται και την έγκυρη αξιολόγηση των πρώιμων συμπτωμάτων της νόσου.

Ακόμη, οι Dang et al. (2019) [86] πρότειναν μια μέθοδο για την εκτίμηση της κάμψης της μόνιμης σκυφτής στάσης του σώματος που προκαλείται σε ασθενείς με Πάρκινσον. Χρησιμοποιούνται επιταχυνσιόμετρα που τοποθετούνται στον λαιμό και στο πάνω μέρος της πλάτης και από τα δεδομένα που προκύπτουν υπολογίζεται η απόσταση C7-SAR που αντιστοιχεί στη γωνία κάμψης του κορμού που μετριέται μεταξύ των σπονδύλων C7 και S1. Ο στόχος αυτής της μελέτης είναι να αποσαφηνιστεί η αποτελεσματικότητα της εφαρμογής ενός φορητού αισθητήρα (στο πάνω μέρος της πλάτης ή στο λαιμό) για τη μέτρηση της σοβαρότητας της κάμψης

του σώματος στη νόσο. Αυτές οι συσκευές παρέχουν απομακρυσμένη παρακολούθηση ασθενών με Πάρκινσον και κλινική αξιολόγηση αυτών, καθώς ανιχνεύουν και ελέγχουν τη σοβαρότητα κάποιου συμπτώματος της νόσου.

## ***2.6 Η Νόσος Πάρκινσον στη Μηχανική Μάθηση και την Όραση Υπολογιστών***

Η Μηχανική Μάθηση με τις διάφορες μεθοδολογίες της συμβάλλει σημαντικά στην ανάπτυξη εφαρμογών που βοηθούν στην έγκαιρη και έγκυρη διάγνωση της νόσου του Πάρκινσον και χρησιμοποιούν λιγότερο επεμβατικές μεθόδους για την επίτευξη του σκοπού τους. Οι περισσότερες από αυτές τις εφαρμογές χαρακτηρίζονται από απλότητα στη λειτουργία τους και ευχρηστία.

Παρουσιάζονται ενδεικτικά κάποιες δημοσιευμένες ερευνητικές μελέτες στο που προτείνουν χρήσιμες εφαρμογές στο πεδίο αυτό.

Το 2015, οι Kostikis et al. [87] πρότειναν ένα σύστημα που βασίζεται στη χρήση smartphone για την ακριβή εκτίμηση του τρόμου των άνω άκρων σε ασθενείς με Πάρκινσον. Χρησιμοποιώντας ένα επιταχυνσιόμετρο τηλεφώνου και πληροφορίες γυροσκοπίου, υπολογίζεται ένα σύνολο μετρήσεων που εξυπηρετεί στον ποσοτικό προσδιορισμό των συμπτωμάτων τρόμου ενός ασθενή. Για να ταξινομηθεί σωστά το ποσοστό των ασθενών με Πάρκινσον έναντι των υγιών εθελοντών, χρησιμοποιήθηκαν τεχνικές μηχανικής μάθησης, καθιστώντας δυνατή την απομακρυσμένη αξιολόγηση της κατάστασης του ασθενούς. Το προτεινόμενο εργαλείο εφαρμόζει μια μη επεμβατική τεχνική, είναι απλό στη χρήση και έχει χαμηλό κόστος.

Κινούμενοι στο ίδιο πλαίσιο, το 2020 οι Anju et al. [88], υποστήριξαν ότι η νόσος του Πάρκινσον μπορεί να διαγνωστεί χρησιμοποιώντας smartphone στα οποία κάθε ενέργεια των ασθενών μπορεί να αναγνωριστεί. Τα δεδομένα που αφορούν τα κινητικά συμπτώματα της νόσου συλλέγονται από το πάτημα των κουμπιών, κάποια προκαθορισμένη κίνηση και την ομιλία μέσω των κατάλληλων αισθητήρων που

διαθέτουν τα κινητά τηλέφωνα. Όσον αφορά τα συμπτώματα λόγω της δραστηριότητας μνήμης, όπως διαταραχές, ελέγχονται με κάποια εικόνα στην οποία ο ασθενής καλείται να θυμηθεί το μοτίβο που ακολουθείται. Τα δεδομένα αυτά εισέρχονται σε κάποια δίκτυα μηχανικής μάθησης που έχουν εκπαιδευτεί σε αντίστοιχα σύνολα δεδομένων με σκοπό την αξιολόγηση τους και την εξαγωγή των σχετικών συμπερασμάτων. Τέλος, η εφαρμογή αυτή χρησιμοποιείται και για την αυτοαξιολόγηση των ασθενών χωρίς να απαιτείται τακτική επίσκεψη στον γιατρό.

Επίσης, το 2015, οι Paredes et al. [89] ανέπτυξαν ένα σύστημα e-Motion Capture, το οποίο είναι ένα λογισμικό που βασίζεται στο Kinect για τον υπολογισμό κινητικών (ρυθμός και μήκος βήματος) και χωροχρονικών (ταχύτητα και επιτάχυνση) παραμέτρων που επηρεάζουν στην κίνηση των ασθενών. Χρησιμοποιείται, λοιπόν, ένα σύστημα λήψης κίνησης πολλαπλών 3D καμερών για την παρακολούθηση του μοτίβου βάδισης κατά τη διάρκεια μιας δραστηριότητας περπατήματος. Τέλος, το σύστημα αυτό μετράει τις κινητικές και χωρικές-χρονικές μεταβλητές που είναι αλλάζουν κατά τη διάρκεια της νόσησης.

Το 2019, οι Buongiorno et al. [90], εφάρμοσαν ένα απλό και χαμηλού κόστους κλινικό εργαλείο που μπορεί να εξάγει χαρακτηριστικά από τη στάση και την κίνηση του σώματος με τη χρήση του αισθητήρα Microsoft Kinect v2 με σκοπό την ταξινόμηση και αξιολόγηση των ασθενών με Πάρκινσον. Το εργαλείο αυτό διερευνά τις κινητικές ικανότητες των άνω και κάτω άκρων των εξεταζόμενων μέσω της ανάλυσης του βαδίσματος, του πατήματος των δακτύλων και των ποδιών. Τα χαρακτηριστικά που εξάγονται από αυτή την ανάλυση εισάγονται σε διαφορετικούς ταξινομητές αν περίπτωση, όπως τα Support Vector Machine (SVM) και Artificial Neural Networks (ANN), που εκπαιδεύτηκαν και αξιολογήθηκαν για την καλύτερη λύση. Τα αποτελέσματα έδειξαν πως ένα τέτοιο σύστημα μπορεί να ανιχνεύσει ανεπαίσθητα συμπτώματα της νόσου κι άρα μπορεί να χρησιμοποιηθεί από ειδικούς για την αξιολόγηση και βαθμολόγηση των επιπέδων της νόσου στους ασθενείς.

Το 2019, οι Celik et al [91], πρότειναν ένα σύστημα αποφάσεων που συμβάλλει στη διάγνωση της νόσου του Πάρκινσον σε πραγματικό χρόνο χρησιμοποιώντας δεδομένα φωνής. Οι ασθενείς με Πάρκινσον καλούνται να καταγράψουν συγκεκριμένες φωνητικές δραστηριότητες, τις οποίες οι γιατροί μπορούν να αξιολογήσουν με τη μετρική UPDRS. Τα φωνητικά αυτά δεδομένα επεξεργάζονται

και εισέρχονται κατάλληλα σε δίκτυα ταξινομητών, όπως ο SVM, ώστε να γίνει η ταξινόμηση στα διάφορα επίπεδα της νόσου.

Την ίδια χρονιά (2019), οι Liu et al. [92] εισήγαγαν μία μέθοδο για τον αυτόματο ποσοτικό προσδιορισμό της βραδυκινησίας σε ασθενείς με Πάρκινσον. Για τον σκοπό αυτό μελετώνται τρία στοιχεία που σχετίζονται με τη βραδυκινησία: το πάτημα των δακτύλων, το ανοιγοκλείσιμο των χεριών και προαπλισμός των χεριών. Το σύστημα λαμβάνει στην είσοδο του ένα βίντεο με κάποια από τις παραπάνω κινήσεις και με τη χρήση τεχνικών εκτίμησης ανθρώπινης θέσης, εντοπίζονται οι περιοχές ενδιαφέροντος, ώστε να εκτιμηθούν οι θέσεις των άκρων και εξάγονται τα χαρακτηριστικά της κίνησης. Τα χαρακτηριστικά αυτά σε συνδυασμό με τις χωροχρονικές παράμετροι που προκύπτουν εισέρχονται στους εκπαιδευόμενους με επίβλεψη ταξινομητές για την πραγματοποίηση της αξιολόγησης πέντε σημείων. Επομένως, η μέθοδος μας μπορεί να παράγει αξιόπιστα αποτελέσματα αξιολόγησης για τη βραδυκινησία της Πάρκινσον με τη χρήση μίας κάμερας κι ενός υπολογιστή, προσφέροντας δυνατότητες στην πραγματοποίηση μακροχρόνιας απομακρυσμένης παρακολούθησης της κατάστασης των ασθενών.

Παράλληλα (2019), οι Sato et al. [93], πρότειναν μία μέθοδο χωρίς επίβλεψη για τον προσδιορισμό των χαρακτηριστικών και του ρυθμού βάδισης σε απλά βίντεο αναπαράστασης βήματος υγιών ατόμων και ασθενών με Πάρκινσον. Για τον σκοπό αυτό χρησιμοποιείται το Open Pose για την εξαγωγή των συντεταγμένων των αρθρώσεων που χρησιμεύουν στην ανάλυση του βηματισμού και στο υπολογισμό κατάλληλων παραμέτρων όπως αυτής του μήκους του βήματος, του χρόνου και του ρυθμού του. Οι παράμετροι αυτοί παίζουν σημαντικό ρόλο στην παρατήρηση της κίνησης ασθενών με Πάρκινσον, καθώς αυτοί εμφανίζουν χαμηλότερο ρυθμό βήματος, μικρότερο μήκος και διάρκεια, και ανεπαρκή ανύψωση των ποδιών, ισχίου και γονάτου έναντι των υγιών ατόμων. Έτσι, μπορεί να παρακολουθηθούν τα κινητικά συμπτώματα ασθενών με Πάρκινσον με τη χρήση απλών 2D καμερών οικιακής χρήσης.

Επίσης, το 2019, οι Langevin et al. [94], ανέπτυξαν την εφαρμογή PARK (Parkinson's Analysis with Remote Kinetic-tasks) που βοηθά στην ανίχνευση και παρακολούθηση της νόσου χωρίς την απαίτηση συνεχών ιατρικών επισκέψεων. Στο πλαίσιο της αξιολόγησης, το PARK υποβάλλει τους χρήστες σε έξι κινητικές

δραστηριότητες, όπως το πάτημα δακτύλων, η προέκταση και η περιστροφή της παλάμης, το τρέμουλο σε σταθερή στάση, και μία ηχητική, όπως η εκφορά παρατεταμένα του φωνήεντος «α», σύμφωνα με την κλίμακα UPDRS, οι οποίες καταγράφονται μέσω μίας webcam. Από τα βίντεο αυτά εξάγονται οι εκφράσεις των προσώπων των ατόμων, οι οποίες και συγκρίνονται με αυτές της βάσης με σκοπό την ανίχνευση διαφορών μεταξύ τους. Η ίδια λογική ακολουθείται και στην ανάλυση της κίνησης των χεριών τους. Για την εξαγωγή αυτών των χαρακτηριστικών χρησιμοποιείται ένας ειδικά διαμορφωμένος αλγόριθμος και συνεπώς επιτυγχάνεται με την πραγματοποίηση απλών δραστηριοτήτων η αυτο-παρακολούθηση της εξέλιξης της νόσου από τους ασθενείς κι όσους ίσως κινδυνεύουν.

Το 2020, οι Amirhossein et al. [95], παρουσίασαν ένα μοντέλο βαθιάς μάθησης για την αξιολόγηση της σοβαρότητας της διαταραχής της κίνησης στους ασθενείς με Πάρκινσον με τη μετρική UPDRS αξιοποιώντας την κίνηση των χεριών και το βάδισμα. Για τον σκοπό αυτό λαμβάνεται ένα βίντεο από έναν ασθενή που εκτελεί μια δοκιμή σύμφωνα με τη UPDRS, όπως το άνοιγμα και το κλείσιμο των χεριών, το μοντέλο αναλύει την κίνηση αυτή και τη διαχωρίζει από το παρασκήνιο της με την τεχνική optical flow. Ακολουθεί ένα 3D CNN που εκπαιδεύεται σε χωροχρονική πληροφορία που παρέχουν τα βίντεο με χαμηλό υπολογιστικό κόστος. Τα όρια της κίνησης υπολογίζονται και αξιοποιούνται για την καταστολή της επιρροής των αυθαίρετων κινήσεων της κάμερας σε δεδομένα βίντεο πραγματικού κόσμου.

Την ίδια χρονιά (2020), οι Lu et al. [96], πρότειναν μια μέθοδο βασισμένη στην ανάλυση βίντεο για την αυτόματη αξιολόγηση της σοβαρότητας του Πάρκινσον που σχετίζεται με διαταραχές στο βάδισμα και στη στάση του σώματος. Η μέθοδος αυτή παρατηρεί απλές εγγραφές βίντεο ατόμων, εξάγει τους τρισδιάστατους σκελετούς του σώματος (πόζες), τους παρακολουθεί στη διάρκεια του χρόνου και ταξινομεί τις κινήσεις σύμφωνα με τη μετρική βάδισης UPDRS. Τα μοντέλα εκτίμησης τρισδιάστατων θέσεων εκπαιδεύονται σε μεγάλα δημόσια σύνολα δεδομένων, ώστε να είναι αποτελεσματική η χρήση τους και αξιοποιούνται στην εξαγωγή τρισδιάστατων σκελετών (τρειςδιάστατες συντεταγμένες αρθρώσεων σώματος) από τα κλινικά δεδομένα της εφαρμογής. Τέλος, η μέθοδος αυτή συνιστά έναν αντικειμενικό βιοδείκτη της σοβαρότητας της νόσου.

Παράλληλα (2020), οι Jin et al. [97], ανέπτυξαν ένα σύστημα διάγνωσης του Πάρκινσον μέσω της ανάλυσης προσώπου. Οι ασθενείς υποφέρουν από απώλεια έκφρασης προσώπου και μειωμένη κίνηση των ματιών τους. Στην προσπάθειά τους να εκφραστούν, οι μύες τους κινούνται αργά και συχνά εμφανίζουν υπερβολικές εκφράσεις, οπότε μπορεί να γίνει διάκριση μεταξύ των εκφράσεων υγιών ατόμων και ασθενών. Εκμεταλλευόμενοι αυτές τις διαπιστώσεις, οι ερευνητές συνέλεξαν βίντεο όπου καταγράφονται εκφράσεις του προσώπου, όπως χαμόγελο, και εφαρμόζουν την τεχνική Face++ για την ανίχνευση σημείων του προσώπου από εικόνες. Με βάση αυτά τα χαρακτηριστικά γίνεται η ταξινόμηση των εκφράσεων βάσει μοντέλων νευρωνικών δικτύων, όπως είναι ο SVM, ώστε να κατηγοριοποιηθούν οι ασθενείς σε πάσχοντες και μη από τη νόσο Πάρκινσον.

Το 2021, οι Channa et al. [98], πρότειναν μια μέθοδο για την ανίχνευση και αντικειμενική εκτίμηση του τρόμου και της βραδυκινησίας των ασθενών με Πάρκινσον με τη χρήση ενός ειδικού βραχιολιού. Τα προς εξέταση άτομα πραγματοποιούν κινήσεις με τα πάνω άκρα τους που λαμβάνονται και καταγράφονται μέσω του βραχιολιού, ώστε να αξιολογηθούν βάσει της μετρικής της νόσου UPDRS. Το βραχιόλι αποτελείται από έναν μικροελεγκτή που διαβάζει σε πραγματικό χρόνο τα δεδομένα που λαμβάνει από τους αισθητήρες, όπως το επιταχυνσιόμετρο και το γυροσκόπιο και αποθηκεύει τις τιμές που εξάγει σε μία κάρτα μνήμης micro SD. Τα χαρακτηριστικά που εξάγονται, εισέρχονται σε δίκτυα μηχανικής μάθησης, όπως τον KNN για να γίνει η ταξινόμηση με σκοπό την ανίχνευση του τρόμου και της βραδυκινησίας μεταξύ ασθενών και υγιών ατόμων. Σε κλινικό περιβάλλον, ο γιατρός μπορεί να χρησιμοποιήσει τη συσκευή για να κατανοήσει γρήγορα τον τρόπο και τη βραδυκινησία των ασθενών με μεγαλύτερη ακρίβεια.

Την ίδια χρονιά (2021), οι Mehta et al. [99], παρουσίασαν ένα νέο πλαίσιο βαθιάς μάθησης που χρησιμοποιεί αναπαραστάσεις δεδομένων ανθρώπινου σώματος από βίντεο με sit-to-stand στάσεις για την αυτοματοποίηση της αξιολόγησης των επιμέρους συντελεστών της μετρικής UPDRS, της βραδυκινησίας και των διαταραχών βάδισης. Πρώτα, χρησιμοποιείται ένας ανιχνευτής ανθρώπου, εξάγει ανθρώπινες πόζες από το κάθε καρέ βίντεο και τροφοδοτεί με τις πληροφορίες αυτές ένα μοντέλο 2D pose estimation που προβλέπει τις θέσεις συντεταγμένων ανθρώπινων αρθρώσεων. Στη συνέχεια, ένα μοντέλο 3D Pose χρησιμοποιεί τις πληροφορίες της 2D πόζα και προβλέπει τις θέσεις των αρθρώσεων σε 3D χώρο.

Τέλος, οι πληροφορίες 3D pose τροφοδοτούνται στο προτεινόμενο δίκτυο για να προβλεφθούν οι βαθμολογίες UPDRS. Η προτεινόμενη αυτή τεχνολογία αλλάζει τον τρόπο με τον οποίο κατανοούμε τα κινητικά συμπτώματα παρέχοντας αυτοματοποιημένες και συνεχείς προβλέψεις για τις αξιολογήσεις κινητικών εξετάσεων που χρησιμοποιούνται συνήθως ως κλινικά σημεία αξιολόγησης που επιτρέπει τη λήψη κλινικών αποφάσεων και επιταχύνει τις σχετικές κλινικές διαδικασίες. Τέλος, προσφέρεται και για αυτο-αξιολόγηση στο σπίτι.

Αξίζει να αναφερθούμε και στην εφαρμογή KELVIN-PD [100] που επιτρέπει στους γιατρούς να καταγράψουν βίντεο και άλλα δεδομένα του ασθενή που εξετάζουν με σκοπό την αξιολόγηση του. Οι πληροφορίες περνούν σε ένα ασφαλές σύστημα λογισμικού, η πρόσβαση στο οποίο είναι πολύ εύκολη. Χρησιμοποιούνται τεχνικές εκτίμησης θέσης για την επεξεργασία και ανάλυση των καταγεγραμμένων βίντεο κι έτσι εντοπίζονται συμπτώματα βραδυκινησίας και τρόμου κατά την κίνηση του ασθενή. Καταγράφεται ο ασθενής να εκτελεί μία ορισμένη δραστηριότητα και επιλέγεται από τον γιατρό η αξιολόγησης, η οποία καταχωρείται και έτσι λαμβάνονται διάφορα στατιστικά και συγκριτικά αποτελέσματα. Στην εφαρμογή αυτή, λοιπόν, δημιουργείται ένας ηλεκτρονικός φάκελος του ασθενή μέσα στον οποίο υπάρχουν όλα εκείνα τα κλινικά στοιχεία που αντιστοιχούν στην κατάστασή του. Η εφαρμογή αυτή διατίθεται προς χρήση από κάθε κινητό μέσο με κάμερα συνδεδεμένη στο διαδίκτυο και είναι εύκολη στη χρήση και στην πρόσβαση.

Επίσης, έχουν αναπτυχθεί συστήματα που χρησιμοποιούν Βαθιά Νευρωνικά Δίκτυα (DNNs) για την ανίχνευση και τη διάγνωση της νόσου Πάρκινσον μέσα από ιατρικές εικόνες απεικόνισης του εγκεφάλου σε συνδιασμό με δημογραφικά, επιδημιολογικά και κλινικά δεδομένα που τις συνοδεύουν. Αυτού του είδους τα συστήματα συνδυάζουν Συνελκτικά Νευρωνικά Δίκτυα (CNNs) και Αναδρομικά Νευρωνικά Δίκτυα (RNNs) που με τη χρήση επιβλεπόμενης και μη μάθησης, καθώς και transfer learning μπορούν να αποδώσουν σημαντικά στην πρόβλεψη της νόσου μέσα από δεδομένα MRIs (Magnetic Resonance Images) ή DaT (Dopamine Transporters) scans [101] [102] [103] [104] [105] [106] [107]. Επίσης, οι μέθοδοι που αναπτύσσονται σε αυτές τις μελέτες μπορούν να επεκταθούν στην διάγνωση κι άλλων ασθενειών που χρησιμοποιούν αντίστοιχες απεικονιστικές μεθόδους, δεδομένου ότι μαθαίνουν κάποια συμπεριφορά ως προς τη διαχείριση τέτοιου είδους εικόνων. Τέλος, τα συστήματα αυτά υποβοηθούν σημαντικά την ιατρική πράξη με σκοπό τη

διαχείριση απαιτητικών και ακατέργαστων δεδομένων και την υψηλής ακρίβειας κατηγοριοποίηση αυτών λόγω των DNNs. συμβάλλουν σημαντικά στην ανάπτυξη συστημάτων που επιδέχονται ιατρικά δεδομένα και μπορούν να τα αξιολογήσουν χωρίς κάποια προεπεξεργασία και με αποδοτικό τρόπο έγκαιρα και έγκυρα.

## **2.7 Αυτόματη αξιολόγηση της νόσου με *vision based* μεθόδους**

Τα τελευταία χρόνια, η επιστημονική έρευνα έχει στραφεί στην ανεύρεση μεθόδων αξιολόγησης του Πάρκινσον που μπορούν να χρησιμοποιηθούν ασφαλώς για κλινικές εκτιμήσεις. Έχει δοθεί μεγάλη σημασία στην ανάπτυξη συστημάτων βασισμένων σε βίντεο με σκοπό τον εντοπισμό συμπτωμάτων σχετικών με τη νόσο. Τα συστήματα αυτά παρέχουν συνήθως μεγάλη ευχρηστία και είναι αποδοτικά, συνεισφέροντας στη συνεχή παρακολούθηση των ασθενών είτε από τους ιατρούς, είτε από τους ίδιους τους ασθενείς και προσφέροντας ταχύτητα αλλά και μεγάλη ακρίβεια στην διάγνωση συμπτωμάτων της νόσου. Τέλος, οι εφαρμογές αυτού του είδους στοχεύουν στη χρήση όλο και λιγότερο εξειδικευμένων εξαρτημάτων αποτύπωσης και ανάλυσης της κίνησης, μειώνοντας έτσι και το κόστος τους.

Ένα αξιολογικό παράδειγμα τέτοιας μεθόδου είναι η παρακάτω, η οποία αποτελεί μία από τις αρχικές απόπειρες αναγνώρισης κίνησης με σκοπό την αυτόματη αξιολόγηση συμπτωμάτων του Πάρκινσον.

Το 2018, οι Li et al. [108], πρότειναν μία μέθοδο αξιολόγησης του παρκινσονισμού και της δυσκινησίας που προκαλεί η μακροχρόνια χορήγηση λεβοντόπας σε ασθενείς με Πάρκινσον με τη χρήση μεθόδων εκτίμησης θέσης σε βίντεο. Για τον σκοπό αυτό, οι ασθενείς υποβάλλονται σε ορισμένες δραστηριότητες, όπως είναι η επικοινωνία, η κατάποση, η κίνηση των ποδιών και το πάτημα των δακτύλων, οι οποίες και βιντεοσκοποούνται για να αναλυθούν στη συνέχεια. Για την εξαγωγή των χαρακτηριστικών κίνησης και την εκτίμηση των θέσεων συγκεκριμένων μερών του σώματος χρησιμοποιείται η τεχνική των Convolutional Pose Machines. Σημειώνεται ότι οι δραστηριότητες της επικοινωνίας και της κατάποσης συμβάλλουν στην αξιολόγηση της δυσκινησίας μέσω της κλίμακας UDysRS, ενώ οι τροχιές της



κίνησης των ποδιών και η δύναμη του πατήματος των δακτύλων καθορίζουν την σοβαρότητα του Πάρκινσον βάσει της κλίμακας UPDRS. Για κάθε μία περίπτωση αξιολόγησης λαμβάνονται υπόψη τα χαρακτηριστικά ορισμένων αρθρώσεων που σχετίζονται με την εκάστοτε δραστηριότητα. Τα χαρακτηριστικά αυτά χρησιμοποιούνται για την δυαδική ταξινόμηση των ασθενών, ασθενείς με Πάρκινσον και ασθενείς με δυσκινησία λόγω χορήγησης λεβοντόπας, για την ταξινόμηση των ασθενών σε υγιείς, ασθενείς με Πάρκινσον και ασθενείς με δυσκινησία λόγω χορήγησης λεβοντόπας, καθώς και για την πρόβλεψη τιμών των μετρικών του Πάρκινσον και της δυσκινησίας που υποδεικνύουν τη σοβαρότητα της νόσου και των συμπτωμάτων. Τέλος, το σύστημα αυτό μπορεί έχει προοπτικές να αναπτυχθεί, ώστε να χρησιμοποιείται μέσω κινητών και tablet με σκοπό οι ασθενείς να αυτοαξιολογούνται συχνά και να μεταδίδουν πληροφορίες στον γιατρό τους για περαιτέρω κλινική αξιολόγηση.

Οι ίδιοι ερευνητές, την ίδια χρονιά [109], επικεντρώθηκαν στην παρουσίαση της παραπάνω μεθόδου για την αυτόματη αξιολόγηση της δυσκινησίας που προκαλείται σε ασθενείς με Πάρκινσον από τη χορήγηση λεβοντόπα. Ακολουθείται το σκεπτικό της αξιοποίησης του ποσοτικού προσδιορισμού των συμπτωμάτων της νόσου προς αξιοποίηση στην εξαγωγή συμπερασμάτων σχετικά με την αποτελεσματικότητα των θεραπειών. Τα χαρακτηριστικά που εξάγονται χρησιμοποιούνται για την ανίχνευση της έναρξης και της ύφεσης της δυσκινησίας με την αξιολόγηση του UDysRS. Για την εξαγωγή της μετρικής αυτής αξιολογούνται οι κινήσεις και τα χαρακτηριστικά συγκεκριμένων σημείων, του λαιμού, του δεξιού και αριστερού χεριού, του δεξιού και αριστερού ποδιού και ο κορμός.

Εμπνεόμενοι από την παραπάνω μεθοδολογία, ακολουθούμε κι εμείς την προσπάθεια ανάπτυξης ενός συστήματος αυτόματης αξιολόγησης της δυσκινησίας που προκαλείται από τη θεραπεία του Πάρκινσον μέσω της χρήσης απλών καμερών και μεθόδων μηχανικής μάθησης.

## **Κεφάλαιο 3: Μεθοδολογία & Υλοποίηση**

Στόχος της παρούσας εργασίας είναι η ανάπτυξη μιας μεθοδολογίας που επεξεργάζεται και αναλύει δεδομένα που προέρχονται από τα καρέ ενός βίντεο με σκοπό την αυτόματη αξιολόγηση της δυσκινησίας που πιθανόν να εμφανίζουν οι ασθενείς με Πάρκινσον στους οποίους χορηγείται ως φαρμακευτική αγωγή η δραστική ουσία λεβοντόπα.

### **3.1 Δεδομένα**

Τα δεδομένα που θα διαχειριστούμε για την υλοποίηση ενός τέτοιου συστήματος προέρχονται από το σύνολο δεδομένων της μελέτης που αναφέρθηκε στην υποενότητα 2.7 [108]. Τα δεδομένα αυτά καταγράφηκαν από το Κέντρο Κινητικών Διαταραχών του Toronto Western Hospital. Στη μελέτη συμμετείχαν εννέα άτομα, πέντε άνδρες και τέσσερις γυναίκες, με μέση ηλικία 64 έτη, οι οποίοι είχαν διάγνωση ιδιοπαθούς Πάρκινσον και ενοχλήσεις από τη χορήγηση της λεβοντόπα στο 25% της ημέρας τους τουλάχιστον. Οι ασθενείς αυτοί υποβλήθηκαν σε επιλεγμένες δραστηριότητες σύμφωνα με τις τυπικές κλίμακες αξιολόγησης του Πάρκινσον και της δυσκινησίας, UPDRS και UDysRS αντίστοιχα. Πιο συγκεκριμένα, οι εξεταζόμενοι καλούνται να περιγράψουν μια εικόνα, να συνομιλήσουν με τον εξεταστή και να ανταποκριθούν σε διανοητικά και μαθηματικά τεστ (Communication task), καθώς και να πιουν ένα ποτήρι νερό (Drinking task). Αυτές οι δύο δραστηριότητες αξιολογούνται από τους γιατρούς σύμφωνα με τη κλίμακα UDysRS. Επίσης, υποβάλλονται σε κινήσεις των ποδιών που καταδεικνύουν την ευελιξία τους (Leg agility task) και σε ασκήσεις πατήματος των ποδιών στο έδαφος για τον έλεγχο της δύναμης των δακτύλων (Toe tapping task). Από αυτές τις δραστηριότητες αξιολογούνται σύμφωνα με την κλίμακα UPDRS. Οι ενέργειες αυτές των εξεταζόμενων καταγράφονται μέσω μίας κάμερας ανάλυσης 480 × 640 ή 540 × 960 των 30 καρέ ανά δευτερόλεπτο, στην οποία στέκονται ακριβώς απέναντι, και οι αξιολογήσεις τους γίνονται κάθε 15-30 λεπτά και σε διάστημα 2-4 ωρών από τρεις ειδικούς νευρολόγους. Συνεπώς, οι εξεταζόμενοι υποβλήθηκαν σε απλές εργασίες που ενδείκνυνται για την αυτοματοποιημένη αξιολόγηση τους με τη χρήση μη επεμβατικών μεθόδων.

Τα σύνολα δεδομένων αυτά έχουν υποβληθεί σε επεξεργασία με σκοπό την εξαγωγή κατάλληλων χαρακτηριστικών. Σε όλα τα βίντεο που αναφέρθηκαν έχει εφαρμοστεί η τεχνική εκτίμησης κίνησης Convolutional Pose Machines που αποτελεί έναν αλγόριθμο βαθιάς μάθησης που έχει την ικανότητα εντοπισμού των αρθρώσεων του ανθρώπινου σώματος στα καρέ του βίντεο και αναφέρθηκε στην υποενότητα 2.2.2.1. Από κάθε καρέ εξάγονται οι θέσεις - συντεταγμένες 2D των εξής σημείων: κεφάλι, λαιμός, ώμοι, αγκώνες, καρποί, γοφοί, γόνατα και αστράγαλοι.

Εμείς επικεντρωθήκαμε στο σύνολο δεδομένων του Drinking task με σκοπό την ανάπτυξη ενός συστήματος μηχανικής μάθησης που θα προβλέπει τη βαθμολογία της δυσκινησίας στην κλίμακα UDysRS δεδομένης της χορήγησης της λεβοντόπας. Το σύνολο αυτό περιλαμβάνει 118 βίντεο που αντιστοιχούν στους ασθενείς με αναγνωριστικό 1,2,4,5,6,7,8,10,11. Κάθε βίντεο αποτελείται από ένα πλήθος καρέ στο καθένα εκ των οποίων αντιστοιχίζονται οι συντεταγμένες όλων των σημείων που ανιχνεύτηκαν από το CPM. Επίσης, για καθένα από αυτά βίντεο υπάρχει και μία αξιολόγηση στο σύνολο δεδομένων UDysRS, η οποία απαρτίζεται από έξι τιμές με κλίμακα 0-4. Οι έξι αυτές τιμές αφορούν τη βαθμολογία που δόθηκε από τους νευρολόγους για έξι σημαντικά μέρη του σώματος για την αξιολόγηση της δυσκινησίας, τον λαιμό, το δεξί χέρι (δεξιός ώμος-αγκώνας-καρπός), το αριστερό χέρι (αριστερός ώμος-αγκώνας-καρπός), τον κορμό (δεξιός-αριστερός ώμος), το δεξί πόδι (δεξιός γοφός, γόνατο, αστράγαλος), το αριστερό πόδι (αριστερός γοφός, γόνατο, αστράγαλος).

### ***3.2 Προεπεξεργασία δεδομένων – Μεθοδολογία***

Σε οποιαδήποτε διαδικασία Μηχανικής Μάθησης, η προεπεξεργασία των δεδομένων είναι το βήμα με το οποίο μετατρέπονται τα δεδομένα, ή αλλιώς κωδικοποιούνται με κατάλληλο τρόπο, έτσι ώστε να πάρουν μία μορφή την οποία η μηχανή μπορεί εύκολα να αναλύσει. Με άλλα λόγια, τα χαρακτηριστικά των δεδομένων πρέπει να μπορούν να ερμηνευθούν εύκολα από τον αλγόριθμο. Επίσης, η προεπεξεργασία των δεδομένων είναι σημαντική, καθώς μπορεί να επιλεγθούν προς εκμάθηση νευρωνικά δίκτυα απλούστερων αρχιτεκτονικών, δεδομένου ότι ήδη έχουν

εξαχθεί κάποια χαρακτηριστικά. Γι' αυτό κι εμείς ακολουθούμε μία τεχνική κωδικοποίησης των δεδομένων μας πριν την εισαγωγή τους σε κάποιο νευρωνικό δίκτυο.

Πιο συγκεκριμένα, για την προεπεξεργασία των δεδομένων ακολουθούμε το σχήμα κωδικοποίησης ακολουθίας δράσης που προτάθηκε στη μελέτη των Tsatiris et al. [50] και αναφέρθηκε στην υποενότητα 2.2.2.4. Σύμφωνα με αυτό το σύστημα οι χωροχρονικές ακολουθίες δράσης μετατρέπονται αποτελεσματικά σε συμπαγείς αναπαραστάσεις με τη χρήση χαρακτηριστικών εξαγόμενων από τον υπολογισμό της απόστασης Mahalanobis και τον μετασχηματισμό Radon.

Οι υψηλού επιπέδου αναπαραστάσεις της ανθρώπινης κίνησης, όπως είναι τα χωροχρονικά σημεία ενδιαφέροντος και οι αρθρώσεις του σώματος, μπορούν να αντιμετωπιστούν ως πιθανοτικές κατανομές στον χωροχρόνο και στον χώρο αντίστοιχα, σύμφωνα με μία μελέτη των Tsatiris et al. το 2017 [51]. Η διαπίστωση αυτή βοηθά στην ακριβή μοντελοποίηση των χαρακτηριστικών που περιγράφουν μια ακολουθία δράσης και έτσι διευκολύνονται οι τεχνικές αναγνώρισης. Έτσι, τα σημεία ενδιαφέροντος και οι θέσεις των αρθρώσεων συσχετίζονται κατάλληλα δημιουργώντας μια κατάλληλη αναπαράσταση στον χωροχρόνο της κίνησης.

Τα δεδομένα που διαχειριζόμαστε εμείς αφορούν τις συντεταγμένες σκελετικών αρθρώσεων του ανθρώπινου σώματος. Ωστόσο, αυτά θεωρείται ότι αντιπροσωπεύουν επαρκώς την ανθρώπινη δράση και αποτελούν τα χωροχρονικά σημεία ενδιαφέροντος αυτής, δεδομένου ότι περιέχουν τις απαραίτητες πληροφορίες για τη στάση του ανθρώπου, σύμφωνα με όσα αναφέρονται στη μελέτη των Wang et al. το 2017 [110]. Έτσι, η σχετική θέση μίας άρθρωσης και του ανθρώπινου σώματος, το οποίο και αντιστοιχεί στην κατανομή των σημείων των αρθρώσεων, παίζει σπουδαίο ρόλο στον ορισμό της κίνησης. Στην πράξη για την μοντελοποίηση των δεδομένων μας, επιλέγουμε σύμφωνα με τη μελέτη [50] την απόσταση Mahalanobis για τον υπολογισμό των αποστάσεων των σημείων σε σχέση με την κατανομή αυτών. Η απόσταση Mahalanobis επιλέχθηκε γιατί είναι πιο αποτελεσματική και αντιπροσωπευτική έναντι της Ευκλείδειας που υπολογίζει την απόσταση κάθε σημείου από το κέντρο της κατανομής. Ως αποτέλεσμα για όλα τα σημεία όλων των καρέ του βίντεο λαμβάνουμε έναν πίνακα, τον Mahalanobis matrix.

Η απόσταση Mahalanobis μεταξύ  $x$  και  $D$  δίνεται από την εξίσωση:

Εξίσωση 2: Απόσταση Mahalanobis [50]

$$mahal(\mathbf{x}, \boldsymbol{\mu}, S) = \sqrt{(\mathbf{x} - \boldsymbol{\mu})^T S^{-1} (\mathbf{x} - \boldsymbol{\mu})}$$

Όπου το  $x$  αντιστοιχεί σε ένα διάνυσμα της κατανομής  $D$  μέσω  $\mu$  και διακύμανση πίνακα  $S$ .

Όσον αφορά την τελική κωδικοποίηση των δεδομένων και την εξαγωγή χαρακτηριστικών από αυτά, μετά την απόσταση Mahalanobis εφαρμόζεται ο Radon μετασχηματισμός. Ο μετασχηματισμός αυτός υπολογίζει μία 2D αναπαράσταση της εικόνας σε συντεταγμένες  $\phi$  και  $\rho$  [111]. Για τον σκοπό αυτό, ανιχνεύει ευθείες γραμμές στην εικόνα και υπολογίζει ορισμένες λειτουργίες, μέσω συναρτήσεων, των τιμών της εικόνας κατά μήκος αυτών των γραμμών. Τα δεδομένα του αποκαλούνται και sinograms επειδή η έξοδος του μετασχηματισμού είναι ημιτονοειδής.

Δεδομένης μιας συνάρτησης 2D  $f(x, y)$  με  $x, y \in \mathbb{R}$ , ο μετασχηματισμός Radon  $Rf$  είναι ίσος με το ολοκλήρωμα όλων των τιμών του  $f$  κάτω από όλες τις γραμμές  $L$  που λαμβάνονται από την εκάστοτε εικόνα-πλαίσιο, που ορίζονται από τις παραμέτρους  $\rho, \theta \in \mathbb{R}$ , όπου  $\rho$  είναι η απόσταση κάθε γραμμής από την αρχή και  $\theta$  είναι η γωνία που σχηματίζεται μεταξύ της γραμμής και του οριζόντιου άξονα.

Ο μετασχηματισμός υπολογίζεται από την εξίσωση:

Εξίσωση 3: Μετασχηματισμός Radon [50]

$$R_f(\rho, \theta) = \int_L f(x, y) dL$$

Αν αντικαταστήσουμε την παραμετρική εξίσωση γραμμής  $\rho, \theta$  στην παραπάνω εξίσωση, έχουμε την εξίσωση:

$$R_f(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - \rho) dx dy$$

όπου  $\delta$  η συνάρτηση δέλτα Dirac delta, διασφαλίζοντας ότι μόνο οι τιμές της γραμμής  $f$ ,  $\rho$  και  $\theta$  θα ενσωματωθούν.

Πλεονέκτημα του μετασχηματισμού Radon αποτελεί το γεγονός ότι τα χαρακτηριστικά που εξάγει μπορούν να χρησιμοποιηθούν ως είσοδοι σε βαθιά συνελκτικά δίκτυα (CNN), έναντι των αναδρομικών (RNN, LSTM) που συνήθως επιλέγονταν στην περίπτωση των χωροχρονικών δεδομένων. Τα αναδρομικά δίκτυα έχουν υψηλές απαιτήσεις σε δεδομένα και υπολογιστικούς πόρους με αποτέλεσμα να εμφανίζουν [112]. Με την κωδικοποίηση των χωροχρονικών πληροφοριών που αναπτύχθηκε παραπάνω στοχεύουμε στη χρήση ισχυρών αρχιτεκτονικών ανίχνευσης αντικειμένων σε εικόνες που συμβάλλουν στην εξάλειψη της ανάγκης για τέτοια δίκτυα που απαιτούν εξελιγμένο υλικό.

### 3.3 Υλοποίηση

Συνδυάζουμε την παραπάνω μεθοδολογία των Spatio-temporal Radon Footprints [50] με τα δεδομένα μας, με σκοπό την κατάλληλη μοντελοποίηση αυτών για την εισαγωγή τους σε ένα νευρωνικό δίκτυο ικανό να προβλέψει τις αξιολογήσεις δυσκινησίας των ασθενών με Πάρκινσον. Τα παραπάνω εργαλεία προσδίδουν στα δεδομένα μας, τις συντεταγμένες των σκελετικών αρθρώσεων, χωροχρονική αναπαράσταση.

Μελετώντας τη μεθοδολογία διαπιστώσαμε ότι η σχέση μεταξύ της θέσης μιας άρθρωσης και των υπόλοιπων τμημάτων του σώματος ενός ατόμου, σε ένα συγκεκριμένο χρονικό σημείο, μπορεί να κωδικοποιηθεί αντιμετωπίζοντας το σύνολο των αρθρώσεων ως κατανομή και υπολογίζοντας την απόσταση Mahalanobis, σύμφωνα με την εξίσωση 2, μεταξύ αυτής και της συγκεκριμένης άρθρωσης. Στην

περίπτωσή μας, μια χρονική στιγμή είναι ένα καρέ σε μια ακολουθία δράσης ή βίντεο. Έτσι, εάν επεκτείνουμε την παραπάνω εξίσωση σε όλες τις αρθρώσεις ενός καρέ και σε όλα τα καρέ μιας ακολουθίας δράσης, μπορούμε να διατυπώσουμε έναν πίνακα Mahalanobis  $M$ , του οποίου οι τιμές για την άρθρωση  $j$  και το καρέ  $f$  μπορούν να υπολογιστούν χρησιμοποιώντας την ακόλουθη εξίσωση:

Εξίσωση 5: Πίνακας Mahalanobis [50]

$$M(j, f) = mahal(\mathbf{x}_f^j, \boldsymbol{\mu}_f, S_f)$$

όπου  $\mathbf{x}_f^j$  είναι ο φορέας συντεταγμένων της άρθρωσης  $j$  στο καρέ  $f$ , οριζόμενος στα  $R^2$  ή  $R^3$ , το  $\boldsymbol{\mu}_f$  είναι ο μέσος όρος όλων των φορέων συντεταγμένων αρθρώσεων στο καρέ  $f$  (ουσιαστικά το κέντρο) και το  $S_f$  είναι ο πίνακας συνδιακύμανσης της κατανομής αρθρώσεων στο καρέ  $f$ .

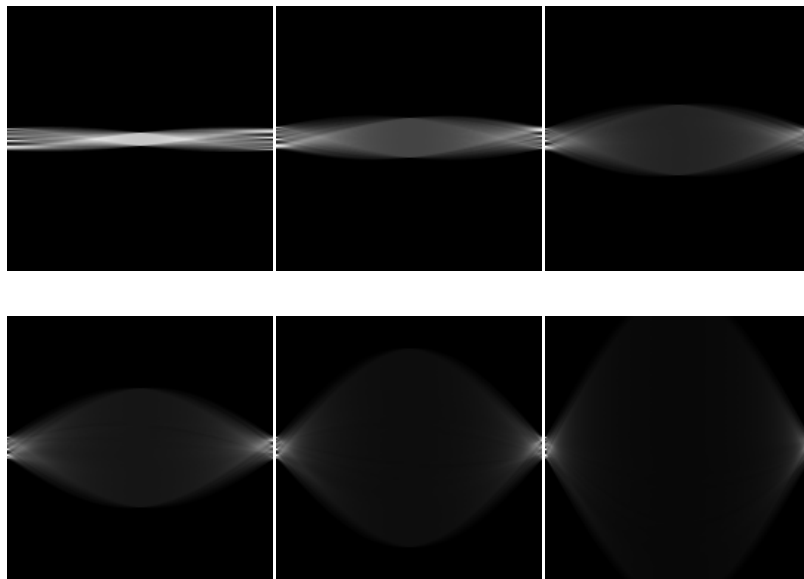
Ο πίνακας Mahalanobis που διατυπώθηκε στην εξίσωση 5 είναι μια πλήρης αναπαράσταση 2D της ακολουθίας δράσης στο χώρο και το χρόνο. Με ανάλυση  $J \times t$ , όπου  $J$  είναι ο συνολικός αριθμός αρθρώσεων και  $t$  είναι η χρονική στιγμή (αριθμός καρέ που έχουν περάσει) από την αρχή της ακολουθίας, κάθε γραμμή του πίνακα έχει τις αποστάσεις Mahalanobis κάθε άρθρωσης στο αντίστοιχο καρέ. Ωστόσο, η αναπαράσταση αυτή δε διαθέτει έναν τρόπο συσχέτισης των πληροφοριών κάθε μεμονωμένου καρέ σε ένα συμπαγές χαρακτηριστικό που θα περιγράψει την ακολουθία δράσης από την αρχή έως τη χρονική στιγμή  $t$ . Γι' αυτό τον λόγο, υπολογίζουμε τον μετασχηματισμό Radon της 2D συνάρτησης  $M$ , δηλαδή του πίνακα Mahalanobis, αντικαθιστώντας την αναπαράσταση  $M(j, f)$  μέσω της εξίσωσης 4. Έτσι, εξάγεται ένα 2D χωροχρονικό χαρακτηριστικό που περιγράφει την ακολουθία δράσης έως τη χρονική στιγμή  $t$  και αποτελεί το επονομαζόμενο Spatio-temporal Radon Footprint (SRFt) της δράσης έως τη χρονική στιγμή  $t$  που εξετάστηκε στην προαναφερθείσα μελέτη [50] και υπολογίζεται σύμφωνα με την παρακάτω εξίσωση:

Εξίσωση 6: Spatio-temporal Radon Footprints [50]

$$SRF_t(\rho, \theta) = \int_0^{t-1} \int_0^{J-1} M(j, f) \delta(j \cos \theta + f \sin \theta - \rho) dj df$$

Πρέπει να ισχύει ότι  $t > 2$ , τόσο για το  $M$ , ώστε να είναι μια 2D συνάρτηση του  $j$  και  $f$ , όσο και για την αναπαράσταση, ώστε να κωδικοποιεί ελάχιστες χρονικές πληροφορίες. Στην πράξη, ανάλογα με την προς μελέτη ενέργεια, ενδέχεται να χρειαστεί να περάσουν κάποια καρέ πριν αρχίσουν να υπολογίζονται τα SRF για να υπάρχει πληροφορία προς αναπαράσταση. Αυτή την πρακτική ακολουθούμε κι εμείς ξεκινώντας την εξαγωγή των χαρακτηριστικών αφού περάσουν τα πρώτα πέντε καρέ. Επίσης, θα πρέπει να υποστηρίζει ότι το  $F > 2$ , το οποίο είναι προφανές, τόσο για τον προαναφερθέντα λόγο όσο και επειδή είναι αδύνατο να εξακριβωθεί η φύση μιας δραστηριότητας με ανεπαρκείς κοινές πληροφορίες.

Τα χαρακτηριστικά που εξάγονται για κάθε βίντεο από την παραπάνω διαδικασία αποθηκεύονται ως ημιτονοειδείς αναπαραστάσεις (sonograms) σε αρχεία της μορφής NPZ. Με αυτόν τον τρόπο για κάθε καρέ των βίντεο παράγεται ένα είδος 2D εικόνας που περιέχει όλη την πληροφορία που παράχθηκε από την προεπεξεργασία των δεδομένων. Το πλεονέκτημα αυτής της αναπαράστασης έγκειται στο ότι διατηρεί τις πραγματικές τιμές που περιέχει ο πίνακας Mahalanobis και δεν αποτελεί ένα μέσω visualization όπως οι εικόνες τύπου png.



Εικόνα 18: Sonograms - Παραδείγματα μετασχηματισμού Radon σε καρέ βίντεο



Δεδομένου, λοιπόν, με το ότι τα χαρακτηριστικά που εξήχθησαν αναπαρίστανται από εικόνες, μπορούμε να χρησιμοποιήσουμε συνελκτικά δίκτυα με σκοπό την εξαγωγή επιπλέον χαρακτηριστικών και άρα την όσο το δυνατόν αποτελεσματικότερη εκμάθηση και πρόβλεψη.

Συνεπώς, τα αρχικά μας δεδομένα που αποτελούνται από συντεταγμένες  $x$  και  $y$  των σκελετικών αρθρώσεων του ανθρώπινου σώματος για κάθε καρέ καθενός καταγεγραμμένου βίντεο, υποβάλλονται σε μία μέθοδο χωροχρονικής αναπαράστασης που βασίζεται στον υπολογισμό της απόστασης Mahalanobis και του μετασχηματισμού Radon που δημιουργεί ισχυρά και αντιπροσωπευτικά μοντέλα 2D αναπαράστασης. Έτσι, εξάγονται χαρακτηριστικά της κίνησης, τα οποία αποτελούν το νέο, κατάλληλα διαμορφωμένο, σύνολο δεδομένων βάσει του οποίου θα επιλεγεί το συνελκτικό νευρωνικό δίκτυο που θα ακολουθήσει.

### ***3.4 Εργαλεία***

Για την υλοποίηση του συστήματός μας χρησιμοποιήσαμε τη γλώσσα προγραμματισμού Python3 και κάποιες βιβλιοθήκες αυτής. Συγκεκριμένα χρησιμοποιήσαμε την NumPy, τη Scikit-learn, την Scipy, την Tensorflow και την Keras.

## **Κεφάλαιο 4: Πειράματα & Ανάλυση αποτελεσμάτων**

### **4.1 Πειράματα**

Το πρόβλημα της αυτόματης αξιολόγησης της δυσκινησίας των ασθενών με Πάρκινσον που προσπαθούμε να επιλύσουμε, ανήκει στην κατηγορία προβλημάτων παλινδρόμησης (regression problems). Τα αποτελέσματα, λοιπόν, που θέλουμε να πάρουμε είναι πραγματικές τιμές. Πιο συγκεκριμένα, επιζητούμε την πρόβλεψη έξι τιμών της κλίμακας UDysRS για καθένα βίντεο που αντιστοιχίζεται στην βαθμολογία των έξι τμημάτων του σώματος που καθορίζουν τη δυσκινησία. Συνεπώς, η έξοδος του νευρωνικού μας δικτύου είναι ένα διάνυσμα (λίστα) έξι τιμών.

Υπενθυμίζουμε ότι στο σύνολο δεδομένων που επεξεργαζόμαστε κάθε εξεταζόμενος ασθενής έχει καταγραφεί να πραγματοποιεί τη δραστηριότητα της κατάποσης νερού (Drinking task) αρκετές φορές σε διαφορετικές χρονικές στιγμές. Γι' αυτόν τον σκοπό, φροντίσαμε να έχουμε αντιστοιχίσει κάθε βίντεο στο άτομο που ανήκει, καθώς και στην αξιολόγηση που έχει λάβει, ώστε κάθε σύνολο εξαγόμενων χαρακτηριστικών που προκύπτει από τα βίντεο λόγω της προεπεξεργασίας να έχει γνωστή προέλευση.

Πριν τον ορισμό κάποιου νευρωνικού δικτύου, ορίζουμε τα σύνολα εκείνα που θα αποτελέσουν την είσοδο του για την εκπαίδευση και την επικύρωση. Εμείς ακολουθήσαμε το πρωτόκολλο επικύρωσης Leave-One-Person-Out Cross-validation (LOPOCV), επιλέγοντας κάθε φορά τα βίντεο ενός ατόμου για την επικύρωση και των υπόλοιπων οκτώ για την εκπαίδευση. Κάθε άτομο εκτελεί μια ενέργεια με μοναδικό τρόπο σε σχέση με κάθε άλλο, οπότε το να διαμορφωθούν έτσι τα σύνολα εκπαίδευσης και επικύρωσης εξυπηρετεί στο γεγονός ότι το δίκτυο δεν υπάρχει περίπτωση να εκπαιδευτεί από δεδομένα του ατόμου με το οποίο θα εξεταστεί η απόδοσή του. Αν και αποτελεί μία ακριβή υπολογιστικά διαδικασία, προσφέρει όμως αξιόπιστη και αμερόληπτη εκτίμηση του εκάστοτε μοντέλου. Ακόμη, μπορεί να είναι ένα αυστηρό πρωτόκολλο, ωστόσο ανταποκρίνεται πλήρως στην πραγματικότητα και τη φύση του προβλήματος, γι' αυτό και επιλέγεται. Επίσης, το πρωτόκολλο αυτό ενδείκνυται για τις περιπτώσεις εκείνες, όπως και τη δική μας, που το σύνολο δεδομένων δεν είναι πολύ μεγάλο και υπάρχει ανάγκη για μία καλή εκτίμηση

απόδοσης μοντέλου. Συνεπώς, η τεχνική αυτή έναντι της απλής διάσπασης του συνόλου αποφεύγει την πιθανή υπερεκπαίδευση και άρα την προκατειλημμένη εκτίμηση απόδοσης του μοντέλου.

Επίσης, το σύνολο εκπαίδευσης πριν εισαχθεί στο νευρωνικό δίκτυο ανακατεύεται ως προς τη σειρά με τη χρήση της συνάρτησης shuffle. Η ενέργεια αυτή είναι απαραίτητη, καθώς τα χαρακτηριστικά των καρτέ των βίντεο δεν παρατίθενται με τη σειρά κι έτσι το μοντέλο διαχειρίζεται κάθε καρτέ σαν να διαθέτει μία δραστηριότητα μεμονωμένη όντας αυτόνομο από τα υπόλοιπα. Το κάθε καρτέ αναπαριστά την κίνηση που έχει γίνει μέχρι εκείνη τη χρονική στιγμή που του αντιστοιχεί, οπότε και είναι λογικό να μπορούμε να διαχειριστούμε τα καρτέ ανεξάρτητα από την αλληλουχία που ανήκουν. Ακόμη, δεδομένου ότι χρησιμοποιούμε batch size στην εκπαίδευση των δικτύων μας, στην περίπτωση που δεν προηγηθεί κάποιο ανακάτεμα του συνόλου εκπαίδευσης, μπορεί κάθε batch να περιέχει την ακολουθία καρτέ ενός βίντεο. Με αυτόν τον τρόπο η έννοια των batches θα έπαυε να αντιστοιχεί σε μια ανεξάρτητη αναπαράσταση του συνόλου δεδομένων κι άρα η εκπαίδευση θα ήταν λιγότερο αντικειμενική ως προς τα δεδομένα της και η απόδοση λιγότερο καλή φυσικά. Συνεπώς, το δίκτυο εκπαιδεύεται με το να εξάγει χαρακτηριστικά από ανεξάρτητα μεταξύ τους καρτέ, με αποτέλεσμα να επιτυγχάνεται καλύτερη συμπεριφορά του δικτύου.

Για να πειραματιστούμε και να εξάγουμε κάποια πρώτα συμπεράσματα που μας βοήθησαν να αποφανθούμε σχετικά με την απόδοση ενός μοντέλου, κρατήσαμε εκτός εκπαίδευσης τα βίντεο του ατόμου 1. Εκπαιδεύσαμε με τα βίντεο των ατόμων 2,4,5,6,7,8,10,11 και επικυρώσαμε με αυτά του 1. Αυτή η τακτική μας εξυπηρέτησε, καθώς θεωρήσαμε ότι τα πειράματα με επικύρωση μονάχα στο άτομο 1 αντιπροσωπεύουν κατά πολύ την πραγματική εικόνα απόδοσης του εκάστοτε μοντέλου, οπότε και δεν αναλωθήκαμε σε περιττές δοκιμές σε περιπτώσεις πειραμάτων που εμφάνιζαν κακή απόδοση.

#### ***4.1.1 Πειράματα ως προς την αρχιτεκτονική του δικτύου***

Το μοντέλο νευρωνικού δικτύου που θα ολοκληρώσει το σύστημα μας, αποφασίζουμε να το δημιουργήσουμε εμείς συνδυάζοντας ακολουθιακά διάφορα

επίπεδα. Πρώτον, επιλέγουμε να χρησιμοποιήσουμε 2D συνελκτικά επίπεδα (Conv2D layers) δεδομένου ότι στο μοντέλο μας θα εισαχθεί μία 2D εικόνα μορφής NPZ. Επίσης, δίκτυα με τέτοιου είδους επίπεδα προορίζονται κατά βάση για την ανάλυση και εξαγωγή χαρακτηριστικών από εικόνες. Έπειτα, επιλέγουμε τα επίπεδα υποδειγματοληψίας (Max Pooling layers) για τη μείωση της διάστασης των επιμέρους εξόδων που γίνονται οι εισοδοί των επόμενων επιπέδων και τα πλήρως διασυνδεδεμένα επίπεδα (Dense layers) για την παραγωγή της τελικής πλέον εξόδου του δικτύου.

Ξεκινάμε τα πειράματά μας με σκοπό να καταλήξουμε σε μία καλή και αποδοτική αρχιτεκτονική νευρωνικού δικτύου. Στα προς δοκιμή δίκτυα χρησιμοποιούμε δύο συνελκτικά επίπεδα, δύο Max Pooling, ένα Flatten και δύο Dense. Εκπαιδεύουμε το εκάστοτε δίκτυο μας με loss function το mean squared error (mse) που αποτελεί κλασική μετρική απόδοσης στα προβλήματα regression, optimizer τον adam που ενδείκνυται για τέτοια προβλήματα και metrics μαζί με το mean squared error (mse) και το root mean squared error (rmse). Κάθε φορά, δοκιμάζουμε να αλλάζουμε κάποιες παραμέτρους των δικτύων, όπως το πλήθος των νευρώνων τους ή της εκπαίδευσης, όπως το batch size και οι εποχές.

Επίσης, πραγματοποιείται και η επικύρωση για το σύνολο των δεδομένων του ατόμου 1, ωστόσο αυτή μας δίνει μία συνολική τιμή για τις δύο μετρικές mse και rmse, που προκύπτει από το testing που γίνεται για κάθε καρέ που περιέχεται στον πίνακα του test set. Εμάς μας ενδιαφέρει η επικύρωση να γίνεται ανά βίντεο, όπως ακριβώς έχουμε λάβει και τα δεδομένα μας (αφού σε κάθε βίντεο αντιστοιχεί μία αξιολόγηση κι ένα άτομο). Όμως, στην παρούσα περίπτωση που δοκιμάζουμε αρχιτεκτονικές δικτύων για να καταλήξουμε σε κάποιο αξιόλογο και αποδοτικό μοντέλο για να συνεχίσουμε τα πειράματά μας, αρκούμαστε σε αυτές τις δύο ενδεικτικές τιμές mse και rmse, αφού αντικατοπτρίζουν καλά την απόδοση του δικτύου. Αυτό συμβαίνει γιατί κάθε καρέ ενός βίντεο έχει ως ετικέτα την αξιολόγηση που έχει γίνει σε όλο το βίντεο και επίσης κάθε καρέ αντιμετωπίζεται στην εκπαίδευση σαν ανεξάρτητο που αναπαριστά κάποια μεμονωμένη δράση (λόγω του shuffling που έχει προηγηθεί). Έτσι, κάθε καρέ στο σύνολο επικύρωσης μπορεί να ελεγχθεί ανεξάρτητα και να προβλεφθεί η αξιολόγηση του με τον ίδιο τρόπο.

Σημειώνεται ότι στα προβλήματα regression οι μετρικές mse και rmse που χρησιμοποιούνται, πρέπει να είναι σε τιμή όσο το δυνατόν μικρότερες και κάτω από το 1 για να θεωρηθεί ότι ένα δίκτυο μπορεί να αξιολογήσει καλά με μεγάλο ποσοστό επιτυχίας.

Παρατίθενται, στη συνέχεια, κάποιες αρχιτεκτονικές δικτύων που δοκιμάστηκαν και ελέχθηκαν για την απόδοσή τους:

### 1<sup>ο</sup> Μοντέλο

```
Model: "sequential"
-----
Layer (type)                Output Shape              Param #
-----
conv2d (Conv2D)              (None, 178, 178, 16)      160
max_pooling2d (MaxPooling2D) (None, 89, 89, 16)        0
conv2d_1 (Conv2D)            (None, 87, 87, 32)        4640
max_pooling2d_1 (MaxPooling2 (None, 43, 43, 32)        0
flatten (Flatten)            (None, 59168)              0
dense (Dense)                 (None, 64)                 3786816
dense_1 (Dense)               (None, 6)                  390
-----
Total params: 3,792,006
Trainable params: 3,792,006
Non-trainable params: 0
```

Εικόνα 19: Αρχιτεκτονική 1ου Μοντέλου

Στο παραπάνω μοντέλο επιλέγονται Conv2D επίπεδα με μικρό πλήθος νευρώνων 16 και 32 αντίστοιχα, kernel size = (3,3) και φυσικά συνάρτηση ενεργοποίησης relu, Max Pooling επίπεδα με pool size = (2,2) και Dense με βάθος 64 και relu και τελικά με 6 νευρώνες για την εξαγωγή του αποτελέσματος.

## 2<sup>ο</sup> Μοντέλο

```
Model: "sequential"
Layer (type)                Output Shape                Param #
-----
conv2d (Conv2D)              (None, 178, 178, 32)       320
max_pooling2d (MaxPooling2D) (None, 89, 89, 32)         0
conv2d_1 (Conv2D)            (None, 87, 87, 64)         18496
max_pooling2d_1 (MaxPooling2 (None, 43, 43, 64)         0
flatten (Flatten)            (None, 118336)             0
dense (Dense)                (None, 256)                30294272
dense_1 (Dense)              (None, 6)                  1542
-----
Total params: 30,314,630
Trainable params: 30,314,630
Non-trainable params: 0
```

Εικόνα 20: Αρχιτεκτονική 2ου Μοντέλου

Στο παραπάνω μοντέλο επιλέγονται Conv2D επίπεδα με μικρό πλήθος νευρώνων 32 και 64 αντίστοιχα, kernel size = (3,3) και φυσικά συνάρτηση ενεργοποίησης relu, Max Pooling επίπεδα με pool size = (2,2) και Dense με βάθος 256 και relu και τελικά με 6 νευρώνες για την εξαγωγή του αποτελέσματος.

## 3<sup>ο</sup> Μοντέλο

```
Model: "sequential"
Layer (type)                Output Shape                Param #
-----
conv2d (Conv2D)              (None, 178, 178, 32)       320
max_pooling2d (MaxPooling2D) (None, 89, 89, 32)         0
dropout (Dropout)           (None, 89, 89, 32)         0
conv2d_1 (Conv2D)            (None, 87, 87, 64)         18496
max_pooling2d_1 (MaxPooling2 (None, 43, 43, 64)         0
dropout_1 (Dropout)          (None, 43, 43, 64)         0
flatten (Flatten)            (None, 118336)             0
dense (Dense)                (None, 256)                30294272
dropout_2 (Dropout)          (None, 256)                0
dense_1 (Dense)              (None, 6)                  1542
-----
Total params: 30,314,630
Trainable params: 30,314,630
Non-trainable params: 0
```

Εικόνα 21: Αρχιτεκτονική 3ου Μοντέλου

Στο παραπάνω μοντέλο επιλέγονται Conv2D επίπεδα με μικρό πλήθος νευρώνων 32 και 64 αντίστοιχα, kernel size = (3,3) και φυσικά συνάρτηση ενεργοποίησης relu, Max Pooling επίπεδα με pool size = (2,2) και Dense με βάθος 256 και relu και τελικά με 6 νευρώνες για την εξαγωγή του αποτελέσματος. Επίσης, μετά από κάθε επίπεδο Max Pooling τοποθετείται ένα Dropout επίπεδο με rate 0.25. Ακόμη, πριν το τελικό επίπεδο Dense τοποθετείται ένα Dropout με rate 0.5. Με τα Dropouts αποσκοπούμε στη γενίκευση και επιτάχυνση του δικτύου. Ουσιαστικά στο μοντέλο αυτό χρησιμοποιείται η αρχιτεκτονική του 2<sup>ο</sup> μοντέλου έχοντας προστεθεί τα Dropout επίπεδα.

#### **4.1.2 Πειράματα για τον έλεγχο απόδοσης του δικτύου**

Κάθε καρέ ενός βίντεο αναπαριστά την κίνηση που έχει γίνει μέχρι τη χρονική στιγμή που του αντιστοιχεί και περιέχει τις σχετικές πληροφορίες. Δεδομένου αυτού, στο σύνολο εκπαίδευσης κάθε καρέ θεωρείται αυτόνομο και ανεξάρτητο από την ακολουθία καρέ που ανήκει, οπότε κι επιτυγχάνεται πιο αμερόληπτη εκπαίδευση.

Κατά την εκπαίδευση του δικτύου επιλέγουμε κυλιόμενα το μοντέλο μας να εκπαιδεύεται με τα βίντεο οκτώ εκ των εννέα ατόμων και να επικυρώνεται με τα βίντεο του ενός που δεν χρησιμοποιήθηκαν προηγουμένως. Αυτό συμβαίνει δεδομένου ότι κάθε άνθρωπος κινείται με μοναδικό τρόπο από κάθε άλλον, οπότε οι αναπαραστάσεις του συνοδεύονται κι από κάποια χαρακτηριστικά τα οποία ταυτοποιούν μοναδικά τον καθένα. Έτσι, η χρήση του πρωτοκόλλου Leave-One-Person-Out Cross-Validation (LOPOCV) είναι απαραίτητη για να εξασφαλιστεί ότι το δίκτυο δεν μαθαίνει χαρακτηριστικά που αφορούν το «άγνωστο» άτομο που θα κληθεί να αξιολογήσει κι άρα η εκτίμηση του είναι λιγότερο κατευθυνόμενη.

Ωστόσο, σε πρώτη φάση ο έλεγχος του δικτύου γίνεται ανά καρέ των βίντεο του συνόλου. Εμάς μας ενδιαφέρει η έλεγχος κι άρα αξιολόγηση της αποδοτικότητας του δικτύου ανά βίντεο, όπως ακριβώς γίνεται και στην πραγματικότητα, καθώς οι γιατροί βαθμολογούν τη δυσκινησία των ασθενών στους οποίου χορηγείται μακροχρόνια λεβοντόπα ανά δραστηριότητα, δηλαδή ανά βίντεο. Για να ελέγξουμε με σωστό τρόπο την απόδοση του δικτύου εφαρμόζοντας το LOPOCV κάνουμε την αξιολόγηση διαφορετικά από αυτή που ορίζει η συνάρτηση evaluate του keras,

δηλαδή ανά καρτέ. Η πρόβλεψη της βαθμολογίας των έξι τιμών του UDysRS πρέπει να γίνεται για κάθε βίντεο. Γι' αυτό το λόγο, τα δεδομένα από τις ημιτονοειδείς αναπαραστάσεις κάθε καρτέ ενός βίντεο εισάγονται στη συνάρτηση predict του keras με σκοπό να γίνουν οι προβλέψεις για καθένα καρτέ και στη συνέχεια να υπολογιστεί ο μέσος όρος αυτών των προβλέψεων για καθένα βίντεο. Επίσης, για κάθε τέτοιο μέσο όρο υπολογίζεται το squared error σε σχέση με την δοσμένη αξιολόγηση των γιατρών, οπότε και υπολογίζεται το σφάλμα για κάθε μία τιμή εκ των έξι βαθμολογιών. Αφού υπολογιστούν οι μέσοι όροι και κυρίως τα σφάλματα των προβλέψεων για όλα τα βίντεο του εκάστοτε ατόμου του συνόλου επικύρωσης, υπολογίζεται το mean squared error για το άτομο αυτό ως ο μέσος όρος όλων των επιμέρους σφαλμάτων. Αυτό γίνεται δεδομένου ότι κάθε άτομο κινείται διαφορετικά από κάθε άλλο, οπότε το mse που παρουσιάζει παίζει σημαντικό ρόλο στην αξιολόγηση της αποτελεσματικότητας και της συμπεριφοράς του δικτύου. Με αυτόν τον τρόπο λαμβάνουμε μία εξάδα τιμών που αντιστοιχεί στο mse της κάθε μίας εκ των βαθμολογιών του UDysRS και είναι πολύ αντιπροσωπευτική της απόδοσης του δικτύου, καθώς μπορούμε με πιο πολλές ενδείξεις να αποφανθούμε περί της ευστοχίας των προβλέψεων. Έτσι, γίνεται πραγματική εφαρμογή του συστήματος κι άρα καλύτερη και πιο κατατοπιστική αξιολόγηση του.

Δεδομένου ότι η παραπάνω διαδικασία εφαρμόζεται σύμφωνα με το πρωτόκολλο LOPOCV σε καθένα από τα εννέα άτομο, λαμβάνουμε μία συνολική εικόνα της αποτελεσματικότητας του δικτύου μέσω του υπολογισμού του average mean squared error. Αυτό προκύπτει από τον μέσο όρο των επιμέρους mse που έχουν υπολογιστεί σε κάθε περίπτωση του LOPOCV, οπότε και παίρνουμε ένα διάνυσμα έξι τιμών που μας δείχνει το μέσο σφάλμα στην πρόβλεψη της βαθμολογίας σε καθένα από τα έξι αυτά μέρη του σώματος που συνεισφέρουν στην ανίχνευση της δυσκινησίας λόγω της λεβοντόπα.

Παράλληλα με τα παραπάνω, μέσω των πειραμάτων μας προσπαθούμε να εξακριβώσουμε αν κάποιο υποσύνολο από τα καρτέ ενός βίντεο παίζει σημαντικότερο ρόλο στην πιο ακριβή πρόβλεψη των βαθμολογιών. Όσο εξελίσσεται η κίνηση, τα καρτέ του βίντεο αποτυπώνουν περισσότερη πληροφορία, ωστόσο αυτό δε σημαίνει ότι με την χρήση ορισμένων μόνο καρτέ έναντι άλλων μπορεί να επιτευχθεί καλύτερη εκτίμηση των αποτελεσμάτων. Επίσης, πολλές φορές όσο προχωράει ένα βίντεο στα χαρακτηριστικά που αποτυπώνονται στα καρτέ του προστίθεται θόρυβος, οπότε τα



σφάλματα μεγαλώνουν και γι' αυτό και θα χρειαζόταν να σταματήσει νωρίτερα η πρόβλεψη. Λαμβάνοντας υπόψην όλα τα παραπάνω, προσπαθούμε να μελετήσουμε τη συμπεριφορά των διαφόρων καρτέ ανά τμήματα του βίντεο. Έτσι, αποφασίζουμε να εφαρμόσουμε την προαναφερθείσα τεχνική εξαγωγής πρόβλεψεων και σφαλμάτων για να ελέγξουμε ανά τμήματα του βίντεο, δηλαδή ανά κάποιο αριθμό καρτέ, τα αποτελέσματα που λαμβάνουμε και κατά πόσο αυτά μπορούν να οδηγήσουν σε κάποια γενίκευση σχετικά με τη σημασία των καρτέ κατά την εξέλιξη της κίνησης.

Συνεπώς, υπολογίζουμε τους μέσους όρους, τα επιμέρους squared errors και το mean squared error του εκάστοτε ατόμου για τις εξής δοκιμές:

- 1<sup>η</sup> Δοκιμή: 1<sup>ο</sup> καρτέ κάθε βίντεο (ουσιαστικά έχει κωδικοποιήσει τα πρώτα 5 καρτέ του πραγματικού βίντεο βάσει της προεπεξεργασίας που έχουμε κάνει)
- 2<sup>η</sup> Δοκιμή: πρώτο 25% των καρτέ κάθε βίντεο
- 3<sup>η</sup> Δοκιμή: πρώτο 50% των καρτέ κάθε βίντεο
- 4<sup>η</sup> Δοκιμή: πρώτο 75% των καρτέ κάθε βίντεο
- 5<sup>η</sup> Δοκιμή: 100% των καρτέ κάθε βίντεο
- 6<sup>η</sup> Δοκιμή: τελευταίο 50% των καρτέ κάθε βίντεο

Στην παρακάτω ενότητα θα αναλύσουμε τα αποτελέσματα των δοκιμών που πραγματοποιήσαμε, ώστε να εξάγουμε τα σχετικά συμπεράσματα μας.

## ***4.2 Ανάλυση αποτελεσμάτων***

### ***4.2.1 Αποτελέσματα αρχιτεκτονικών***

Για να εξάγουμε συμπεράσματα σχετικά με την αποδοτικότητα των μοντέλων που δοκιμάζουμε, παρατηρούμε τις μετρικές mean squared error και root mean squared error καθόλη τη διάρκεια της εκπαίδευσης και φυσικά κατά το στάδιο της

επικύρωσης. Επικεντρωνόμαστε, αρχικά, στα αποτελέσματα που εξάγονται από τη δοκιμή των μοντέλων όταν εκτός εκπαίδευσης μένουν τα βίντεο του ατόμου 1 και στη συνέχεια ανάλογα με το αν χρειαζόμαστε περισσότερα στοιχεία για να οδηγηθούμε σε κάποια απόφαση εφαρμόζουμε πλήρως το πρωτόκολλο LOPOCV, ώστε να έχουμε πιο πλήρη εικόνα της συμπεριφοράς των παρακάτω δικτύων στο σύνολο του προβλήματος.

### 1° Μοντέλο:

Η εκπαίδευση του μοντέλου αυτού γίνεται με batch size = 8 ή 16 και epochs = 50.

Το δίκτυο αυτό με τα υπάρχοντα δεδομένα φαίνεται να μην εκπαιδεύεται, καθώς οι τιμές των μετρικών mse και rmse δεν αλλάζουν κατά το πέρασμα των εποχών.

Αν αυξήσουμε το batch size και γίνει 32, κάτι που έχει σημασία στην εκπαίδευση ενός δικτύου, λαμβάνουμε αποτελέσματα, ωστόσο από αυτά καταλαβαίνουμε ότι το παρόν δίκτυο δεν είναι αποδοτικό, δεν εμφανίζει καλά αποτελέσματα ούτε στην εκπαίδευση ούτε στην επικύρωση.

### 2° Μοντέλο:

Η εκπαίδευση γίνεται με batch size = 32 και epochs = 50.

Το δίκτυο αυτό φαίνεται να εκπαιδεύεται καλά, καθώς οι τιμές των μετρικών mse και rmse φθίνουν με το πέρασμα των εποχών, φτάνοντας σε πολύ χαμηλές τιμές. Έτσι, παίρνουμε χαμηλές τιμές μετρικών για το training και σχετικά καλές για το testing, όμως οριακά πάνω από το 1, στην περίπτωση που το άτομο 1 δεν συμμετέχει στην εκπαίδευση.

Αποτελέσματα:

❖ Training: --mse: **0.024171382188796997**

--rmse: **0.15547145903110504**

❖ Testing: --mse: **1.0245941877365112**

--rmse: 1.012222409248352

Δοκιμάζουμε να μειώσουμε τις εποχές στις 35 για να ελέγξουμε τη συμπεριφορά του δικτύου, ωστόσο τα αποτελέσματα τόσο στο training όσο και στο testing είναι χαμηλότερα της προηγούμενης δοκιμής.

Αποτελέσματα:

❖ Training: --mse: 0.04909951984882355

--rmse: 0.22158412635326385

❖ Testing: --mse: 1.3196065425872803

--rmse: 1.1487414836883545

Ακόμη, ελέγχουμε τη συμπεριφορά του δικτύου αυξάνοντας τις εποχές στις 100, προσπαθώντας να επιτύχουμε καλύτερα αποτελέσματα. Όσον αφορά το training παίρνουμε λίγο καλύτερες τιμές των μετρικών από την περίπτωση των 50 εποχών, αλλά όσον αφορά το testing εκεί οι τιμές είναι χειρότερες. Από τα συγκεκριμένα αποτελέσματα συμπαιρένουμε ότι το δίκτυο υπερεκπαιδεύεται (overfitting) στην περίπτωση των 100 εποχών γι' αυτό και δεν είναι τόσο αποδοτικό.

Αποτελέσματα:

❖ Training: --mse: 0.007799153681844473

--rmse: 0.0883127972483635

❖ Testing: --mse: 1.2178373336791992

--rmse: 1.103556513786316

Καταλήγουμε στα παραπάνω συμπεράσματα, όχι μόνο από τη σύγκριση των mse και rmse του training και του testing, αλλά κι από την απόδοση που φαίνεται να έχει η πραγματική εφαρμογή των μοντέλων σε κάποια βίντεο του test set. Συνεπώς, μπορούμε να αξιολογήσουμε επαρκώς την αποδοτικότητα και συμπεριφορά των μοντέλων.

### 3<sup>ο</sup> Μοντέλο:

Η εκπαίδευση γίνεται με batch size = 32 και epochs = 50.

Το δίκτυο αυτό φαίνεται να εκπαιδεύεται καλά, καθώς οι τιμές των μετρικών mse και rmse φθίνουν με το πέρασμα των εποχών, φτάνοντας σε αρκετά χαμηλές τιμές. Έτσι, παίρνουμε χαμηλές τιμές μετρικών για το training και αρκετά καλές για το testing.

Αποτελέσματα:

```
❖ Training: --mse: 0.11253499239683151
--rmse: 0.3354623019695282
❖ Testing: --mse: 0.7983666062355042
--rmse: 0.8935135006904602
```

Ακόμη, ελέγχουμε τη συμπεριφορά του δικτύου αυξάνοντας τις εποχές στις 100, προσπαθώντας να επιτύχουμε καλύτερα αποτελέσματα. Όσον αφορά το training παίρνουμε λίγο καλύτερες τιμές των μετρικών από την περίπτωση των 50 εποχών, αλλά όσον αφορά το testing εκεί οι τιμές είναι χειρότερες. Από τα συγκεκριμένα αποτελέσματα συμπαιρένουμε ότι το δίκτυο υπερεκπαιδεύεται (overfitting) στην περίπτωση των 100 εποχών γι' αυτό και δεν είναι τόσο αποδοτικό.

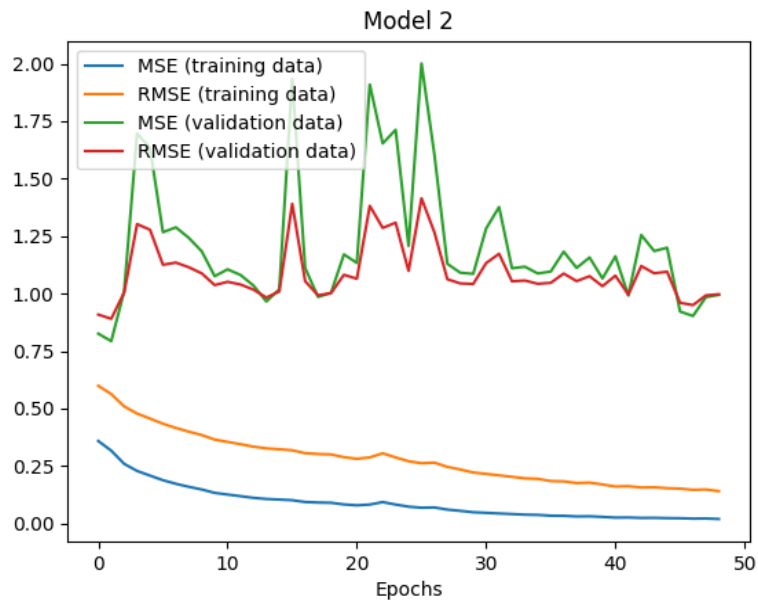
Αποτελέσματα:

```
❖ Training: --mse: 0.08525906503200531
--rmse: 0.29199162125587463
```

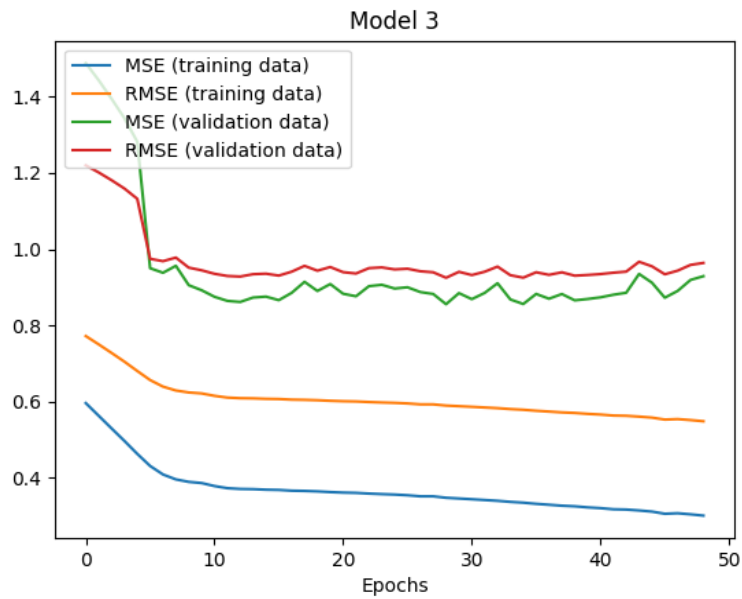
❖ Testing: --mse: 1.1132409572601318

--rmse: 1.0551023483276367

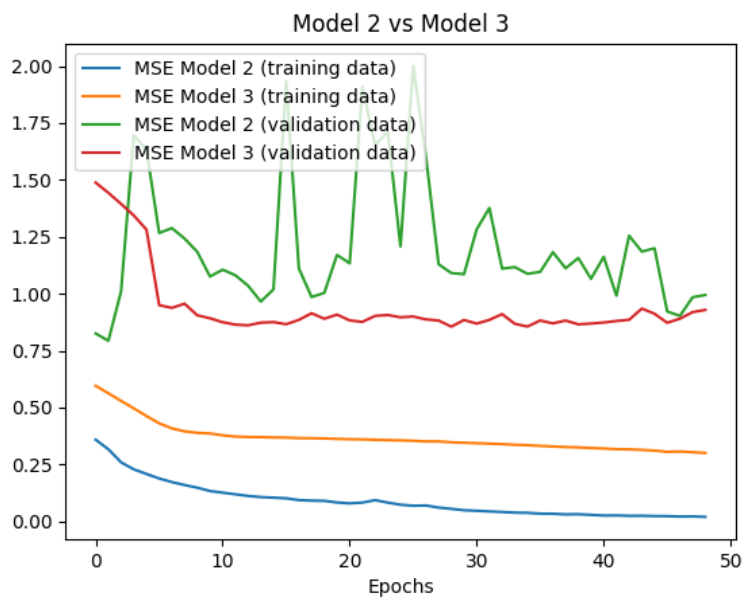
Στη συνέχεια ακολουθούν διαγράμματα των δύο καλύτερων μοντέλων (μοντέλο 1 και 2) με σκοπό τη σύγκριση της πορείας της εκπαίδευσης και βασικά της επικύρωσης τους, ώστε να επιλεγθεί το καλύτερο δυνατό ως προς την αποτελεσματικότητα των εκτιμήσεων του. Γι' αυτό το λόγο, απεικονίζονται οι μετρικές mse και rmse κατά την εξέλιξη των διαδικασιών training και validation.



Εικόνα 22: Διάγραμμα απεικόνισης μετρικών MSE – RMSE του μοντέλου 2



Εικόνα 23: Διάγραμμα απεικόνισης των μετρικών MSE - RMSE του μοντέλου 3



Εικόνα 24: Συγκριτικό διάγραμμα απεικόνισης της μετρικής MSE για τα μοντέλα 2 και 3

Συγκρίνοντας, την απόδοση των μοντέλων 2 και 3, συμπεραίνουμε ότι το μοντέλο 3 με τα επίπεδα Dropouts φαίνεται να έχει λίγο καλύτερες τιμές mse και rmse στο testing για το άτομο 1. Αυτό το συμπέρασμα επιβεβαιώνεται αν κοιτάξουμε και τις προβλέψεις που γίνονται για καθένα από τα βίντεο του ατόμου 1, καθώς και τις προβλέψεις των 6 τιμών του UDysRS στο σύνολο τους με την εφαρμογή του

πρωτοκόλλου LOPOCV. Συνεπώς, επιλέγουμε να συνεχίσουμε τις πειραματικές μας δοκιμές με το μοντέλο 3.

Επομένως, η αρχιτεκτονική του 3<sup>ου</sup> μοντέλου σε συνδιασμό με batch size = 32 και epochs = 50, δίνει πολύ καλά αποτελέσματα στην επικύρωση και κατά βάση στην πραγματική του εφαρμογή για κάθε βίντεο του συνόλου επικύρωσης έχοντας επιλύσει με τα Dropout layers το πρόβλημα της πιθανής υπερεκπαίδευσης. Εμάς μας ενδιαφέρει η αξιολόγηση για κάθε βίντεο κι όχι για κάθε καρέ. Επίσης, παρόλου που αποτελεί ένα μεγάλο δίκτυο σε βάθος και οι παράμετροι του είναι γύρω στις 30.000.000, παραμένει ένα δίκτυο με μετρημένα στον αριθμό επίπεδα, που δεν έχει υψηλές απαιτήσεις κι άρα μας ικανοποιεί αν το κρατήσουμε σαν κύριο κορμό προσπαθώντας να διερευνήσουμε πειράματα ως προς το testing που αφορούν περισσότερο τη φύση του προβλήματος που κάθε φορά αντιμετωπίζεται.

#### ***4.2.2 Αποτελέσματα δοκιμών για τον έλεγχο απόδοσης του δικτύου***

Για να εξάγουμε συμπεράσματα σχετικά με τα αποτελέσματα που λαμβάνουμε από τις διάφορες δοκιμές παρατηρούμε τη συμπεριφορά των προβλέψεων στα διάφορα προς εξέταση βίντεο. Δοκιμάζουμε, αρχικά, όλα τα σχετικά πειράματα σε ένα άτομο (συγκεκριμένα για το 1<sup>ο</sup>), για να έχουμε μία γενική εικόνα και στη συνέχεια τα εφαρμόζουμε σύμφωνα με το πρωτόκολλο LOPOCV για να πάρουμε πιο ασφαλή συμπεράσματα. Έτσι, παρατηρούμε τις όποιες αλλαγές στην ποιότητα των αποτελεσμάτων.

Αξιολογούμε τις παρακάτω δοκιμές ως προς τη συμπεριφορά που φαίνεται να έχουν ανά βίντεο κι όχι στο σύνολο της εφαρμογής, καθώς εξετάζουμε αν υπάρχει κάποιος γενικός κανόνας που να υποδεικνύει ότι ορισμένο πλήθος καρέ βίντεο θα πρέπει να συμμετέχει στις προβλέψεις. Επίσης, η βαθμολογία ορίζεται για ένα βίντεο, συνεπώς πρέπει να εξετάσουμε αν στην εξέλιξη ενός τέτοιου βίντεο υπάρχουν σημεία που παίζουν μεγαλύτερο ή μικρότερο ρόλο στο τελικό αποτέλεσμα. Γι' αυτό τον λόγο, στηριζόμαστε στις παρατηρήσεις μας ανά βίντεο και πρόβλεψη.

Επισημαίνεται ότι κατά τις δοκιμές αυτές εξάγουμε συμπεράσματα παρατηρώντας τις τιμές της μετρικής mean squared error, η οποία είναι

χαρακτηριστική τέτοιων προβλημάτων, για τις έξι βαθμολογίες της κλίμακας UDysRS.

### 1<sup>η</sup> Δοκιμή:

Στη δοκιμή αυτή αξιολογούμε τις προβλέψεις που λαμβάνουμε για καθένα βίντεο στο 1<sup>ο</sup> καρέ της αναπαράστασης της κίνησης του. Παρατηρούμε ότι στην πλειονότητα των βίντεο και κατ' επέκταση των ατόμων, τα squared errors είναι αρκετά μεγάλα. Ωστόσο, υπάρχουν και οι περιπτώσεις εκείνες όπου οι προβλέψεις εμφανίζονται να είναι πολύ κοντινές στις πραγματικές αξιολογήσεις. Παρόλα αυτά, δεν υπάρχει κάποια παρατήρηση που θα μπορούσε να γενικευτεί ασφαλώς για όλες τις περιπτώσεις σχετικών βίντεο.

### 2<sup>η</sup> Δοκιμή:

Στη δοκιμή αυτή αξιολογούμε τις προβλέψεις που λαμβάνουμε για καθένα βίντεο στο πρώτο 25% των καρέ τους. Ομοίως, παρατηρούμε ότι στην πλειονότητα των βίντεο και κατ' επέκταση των ατόμων, τα squared errors είναι μεγάλα. Ωστόσο, υπάρχουν και οι περιπτώσεις εκείνες όπου οι προβλέψεις εμφανίζονται να είναι πολύ κοντινές στις πραγματικές αξιολογήσεις. Επίσης, μεταξύ των έξι τιμών πρόβλεψης, παρατηρούμε ότι άλλες εμφανίζουν μεγαλύτερες τιμές σφαλμάτων κι άλλες μικρότερες χωρίς αυτό να αποτελεί κανόνα και να επαναλαμβάνεται, αλλά και χωρίς πάντα να υπάρχουν κάποια μέρη του σώματος, στα οποία αντιστοιχίζονται πάντα οι χαμηλές ή οι υψηλές τιμές. Επομένως, δεν υπάρχει κάποια παρατήρηση που θα μπορούσε να γενικευτεί με ασφάλεια για όλες τις περιπτώσεις σχετικών βίντεο.

### 3<sup>η</sup> Δοκιμή:

Στη δοκιμή αυτή αξιολογούμε τις προβλέψεις που λαμβάνουμε για καθένα βίντεο στο πρώτο 50% των καρέ τους. Στην πλειονότητα των βίντεο και κατ' επέκταση των ατόμων, παρατηρούμε ότι τα squared errors φθίνουν σε σχέση με τα προηγούμενα κι άρα εμφανίζουν πιο χαμηλές τιμές. Ωστόσο, αυτό δεν γενικεύεται, καθώς υπάρχουν



περιπτώσεις που τα σφάλματα είναι μεγαλύτερα, στο σύνολο των έξι τιμών ή και κατά περίπτωση αυτών.

#### 4<sup>η</sup> Δοκιμή:

Στη δοκιμή αυτή αξιολογούμε τις προβλέψεις που λαμβάνουμε για καθένα βίντεο στο πρώτο 75% των καρτέ τους. Τα σφάλματα πλέον φθίνουν εμφανώς και στις περισσότερες περιπτώσεις παρουσιάζουν τιμές πολύ μικρότερες από τις προηγούμενες. Ωστόσο, και πάλι, αυτό δεν είναι κανόνας, καθώς υπάρχουν εκείνα τα βίντεο οι προβλέψεις των οποίων δεν είναι και τόσο καλές, οπότε και τα σφάλματα διαμορφώνονται αντίστοιχα. Επίσης, δεν υπάρχει μια συγκεκριμένη εικόνα για το καθένα από τα έξι μέρη του σώματος βάσει των αξιολογήσεων τους, καθώς ανάλογα το βίντεο αυτές προσεγγίζουν ή όχι τόσο καλά τις σωστές βαθμολογίες. Αυτό που παραδεχόμαστε, όμως, είναι ότι το πρώτο 75% της κίνησης, ουσιαστικά, ενός βίντεο θα μπορούσε να χρησιμοποιηθεί για να δώσει σχετικά ασφαλή συμπεράσματα τόσο ως προς την απόδοση του δικτύου, όσο και ως προς την πρόβλεψη του ορθού αποτελέσματος.

#### 5<sup>η</sup> Δοκιμή:

Στη δοκιμή αυτή αξιολογούμε τις προβλέψεις που λαμβάνουμε για καθένα βίντεο στο 100% των καρτέ τους. Σχεδόν σε όλα τα βίντεο και κατ'επέκταση σε όλα τα άτομα που εξετάζονται, τα σφάλματα φθίνουν κατά πολύ, παρουσιάζοντας χαμηλότερες τιμές από όλες τις προηγούμενες δοκιμές, ωστόσο πολλές φορές αρκετά κοντινές με αυτές της δοκιμής 4. Ωστόσο, υπάρχουν και αυτές οι λίγες περιπτώσεις που εμφανίζουν σφάλματα μεγαλύτερα της δοκιμής 4, είτε στο σύνολο τους είτε σε κάποιες εκ των έξι τιμών. Ούτε σε αυτή τη δοκιμή μπορούμε να διακρίνουμε κάποια συγκεκριμένη συμπεριφορά των προβλεπόμενων τιμών. Συνεπώς, το μόνο που μπορούμε να ισχυριστούμε είναι ότι στην περίπτωση που αξιολογούνται όλα τα καρτέ ενός βίντεο, λαμβάνουμε στο σύνολο των τιμών καλύτερες προβλέψεις από κάθε άλλη δοκιμή.

## 6<sup>η</sup> Δοκιμή:

Στη δοκιμή αυτή αξιολογούμε τις προβλέψεις που λαμβάνουμε για καθένα βίντεο στο τελευταίο 50% των καρτέ τους. Στο σύνολο της η δοκιμή αυτή δίνει πολύ καλά αποτελέσματα mean squared error ανά άτομο και average mean squared error για την εφαρμογή, ωστόσο στα επιμέρους squared errors που αφορούν καθένα βίντεο παρουσιάζουν άλλοτε πολύ χαμηλά σφάλματα κι άλλοτε πολύ υψηλά. Συνεπώς, δεν μπορούμε να στηριχτούμε σε αυτή τη δοκιμή αξιολόγησης, αφού δεν είναι εξίσου αποδοτική για όλα τα βίντεο, γεγονός που μας ενδιαφέρει δεδομένου ότι το σύστημά μας αξιολογεί βίντεο.

Παρακάτω παρατίθενται οι πίνακες με τα σφάλματα που προκύπτουν από την εφαρμογή των δοκιμών σε καθένα άτομο ξεχωριστά σύμφωνα με το LOPOCV. Τα αποτελέσματα των πινάκων αυτών είναι ενδεικτικά της συμπεριφοράς των διαφόρων δοκιμών που αναλύσαμε παραπάνω. Παρόλοπου μας βοηθούν να διακρίνουμε την αποτελεσματικότητα των δοκιμών και να παρακολουθούμε την τάση των τιμών, εμείς για να εξάγουμε τα συμπεράσματά μας παρατηρήσαμε την πορεία των προβλεπόμενων τιμών και των σφαλμάτων τους για καθένα από τα βίντεο που αποτελούν κάθε φορά το σύνολο ελέγχου (testing set). Γι' αυτόν τον λόγο, δεν παρατηρούμε στους μέσους όρους κάθε ατόμου όσα παρατηρήσαμε στην επιμέρους ανάλυση των βίντεο.

Επίσης, παρατίθεται κι ο πίνακας με τους μέσους όρους των σφαλμάτων για όλα τα άτομα. Οι τιμές αυτές είναι ενδεικτικές της αποδοτικότητας του δικτύου για την κάθε δοκιμή, αλλά και συνολικά. Ωστόσο, και σε αυτήν την περίπτωση πρέπει να λάβουμε υπόψην τη συμπεριφορά του δικτύου σε καθένα άτομο και σε καθένα βίντεο αυτού, για να εξάγουμε τα συμπεράσματά μας.

**Πίνακας 5: Αποτελέσματα δοκιμών ατόμου 1 (Mean Squared Errors)**

	Neck	R-arm	L-arm	Trunk	R-leg	L-leg
<b>1<sup>η</sup> Δοκιμή</b>	1.4080297515	1.6652153205	0.1854136803	0.3295019501	0.3398991304	0.3215156474
<b>2<sup>η</sup> Δοκιμή</b>	1.5676770052	1.8694257355	0.2319084816	0.4922582880	0.3552561092	0.2429898376
<b>3<sup>η</sup> Δοκιμή</b>	1.4900427964	1.7914183535	0.2490902741	0.4958503417	0.4203605641	0.3163781375
<b>4<sup>η</sup> Δοκιμή</b>	1.5516006076	1.9012699685	0.2657075133	0.5212240992	0.3970826256	0.2614511351
<b>5<sup>η</sup> Δοκιμή</b>	1.4931278569	2.0095156949	0.2809842202	0.5069341873	0.3575543448	0.1967878869
<b>6<sup>η</sup> Δοκιμή</b>	1.5065010744	2.2463227974	0.3483103429	0.5636571079	0.3513296529	0.1570186083

Πίνακας 6: Αποτελέσματα δοκιμών απόμον 2 (Mean Squared Errors)

	Neck	R-arm	L-arm	Trunk	R-leg	L-leg
1 <sup>η</sup> Δοκιμή	0.4697857601	0.2169524848	0.1749454745	0.6550984796	0.5045804292	0.3889376914
2 <sup>η</sup> Δοκιμή	0.5237591105	0.2597158654	0.2105379249	0.7545756035	0.5800647073	0.4554084460
3 <sup>η</sup> Δοκιμή	0.5351020880	0.2684648073	0.2153790679	0.7926600260	0.6005681597	0.4715748539
4 <sup>η</sup> Δοκιμή	0.5408983328	0.2751277799	0.2163650081	0.8029908476	0.6015470001	0.4739087556
5 <sup>η</sup> Δοκιμή	0.5469174122	0.2795773615	0.2157412391	0.8027427484	0.5979000857	0.4696816017
6 <sup>η</sup> Δοκιμή	0.5596364681	0.2911602889	0.2173387874	0.8133723252	0.5961580101	0.4692761792

Πίνακας 7: Αποτελέσματα δοκιμών απόμον 4 (Mean Squared Errors)

	Neck	R-arm	L-arm	Trunk	R-leg	L-leg
1 <sup>η</sup> Δοκιμή	0.4431890349	0.1914055322	0.2061629772	0.2999387999	0.4700767488	0.2311551505
2 <sup>η</sup> Δοκιμή	0.4431888353	0.1914054406	0.2061630881	0.2999388892	0.4700768350	0.2311549137
3 <sup>η</sup> Δοκιμή	0.4431896073	0.1914058737	0.2061629313	0.2999387704	0.4700771873	0.2311551105
4 <sup>η</sup> Δοκιμή	0.4431895357	0.1914057305	0.2061628788	0.2999387303	0.4700769678	0.2311552849
5 <sup>η</sup> Δοκιμή	0.4431893909	0.1914052468	0.2061628496	0.2999384151	0.4700765043	0.2311554907
6 <sup>η</sup> Δοκιμή	0.4431896282	0.1914058991	0.2061629220	0.2999387634	0.4700772023	0.2311551398

Πίνακας 8: Αποτελέσματα δοκιμών απόμον 5 (Mean Squared Errors)

	Neck	R-arm	L-arm	Trunk	R-leg	L-leg
1 <sup>η</sup> Δοκιμή	0.2867046218	0.2754612796	0.1886904849	0.3364131445	0.1737063485	0.2005038386
2 <sup>η</sup> Δοκιμή	0.2867045456	0.2754610993	0.1886904871	0.3364131089	0.1737063696	0.2005038737
3 <sup>η</sup> Δοκιμή	0.2867045439	0.2754610378	0.1886904213	0.3364128527	0.1737063189	0.2005038857
4 <sup>η</sup> Δοκιμή	0.2867045532	0.2754615753	0.1886905168	0.3364134580	0.1737063031	0.2005038895
5 <sup>η</sup> Δοκιμή	0.2867045432	0.2754619056	0.1886907091	0.3364137488	0.1737063895	0.2005038078
6 <sup>η</sup> Δοκιμή	0.2867045426	0.2754610731	0.1886904311	0.3364129128	0.1737063306	0.2005038861

Πίνακας 9: Αποτελέσματα δοκιμών απόμον 6 (Mean Squared Errors)

	Neck	R-arm	L-arm	Trunk	R-leg	L-leg
1 <sup>η</sup> Δοκιμή	0.4408658905	0.2392418507	0.3647496403	0.2548867941	0.4547284727	0.2816047138
2 <sup>η</sup> Δοκιμή	0.4837623650	0.2337494645	0.5705801540	0.3950129318	0.4723909438	0.4361062402
3 <sup>η</sup> Δοκιμή	0.5948075317	0.3167972793	0.5793538143	0.4451852489	0.4690840909	0.4489084154
4 <sup>η</sup> Δοκιμή	0.5634310527	0.3052772081	0.5405089727	0.3733862506	0.4777898554	0.4049405624
5 <sup>η</sup> Δοκιμή	0.5557320352	0.3095249395	0.4936385577	0.3335810247	0.4652579870	0.3531020317
6 <sup>η</sup> Δοκιμή	0.5333240370	0.3183570509	0.4360921108	0.2687606622	0.4741261575	0.3077693874

Πίνακας 10: Αποτελέσματα δοκιμών ατόμου 7 (Mean Squared Errors)

	Neck	R-arm	L-arm	Trunk	R-leg	L-leg
1 <sup>η</sup> Δοκιμή	0.9090643247	0.8270572719	1.0669997125	1.4662591802	1.1502244961	1.7551872967
2 <sup>η</sup> Δοκιμή	0.8056490825	0.5784757943	0.5078275458	0.6735016779	0.6347502503	0.7696281123
3 <sup>η</sup> Δοκιμή	0.8945361647	0.6806179321	0.4226042686	0.5834071803	0.4799831846	0.6153266318
4 <sup>η</sup> Δοκιμή	0.8916213074	0.6779563293	0.3373860869	0.4797177432	0.3833732665	0.4888396558
5 <sup>η</sup> Δοκιμή	0.7875758782	0.5838900556	0.2845900825	0.3972094853	0.3332034392	0.4290705676
6 <sup>η</sup> Δοκιμή	0.6953737716	0.5035203848	0.1917442641	0.2731759229	0.2229544589	0.2940235196

Πίνακας 11: Αποτελέσματα δοκιμών ατόμου 8 (Mean Squared Errors)

	Neck	R-arm	L-arm	Trunk	R-leg	L-leg
1 <sup>η</sup> Δοκιμή	0.2106507474	0.2857693927	0.1956855698	0.4509268703	0.2209236232	0.2364052962
2 <sup>η</sup> Δοκιμή	0.2106507898	0.2857693711	0.1956856556	0.4509268538	0.2209236171	0.2364052330
3 <sup>η</sup> Δοκιμή	0.2106506838	0.2857694491	0.1956856922	0.4509269199	0.2209233266	0.2364052631
4 <sup>η</sup> Δοκιμή	0.2106509410	0.2857694798	0.1956856427	0.4509268885	0.2209230988	0.2364054307
5 <sup>η</sup> Δοκιμή	0.2106510477	0.2857691089	0.1956853152	0.4509267621	0.2209230559	0.2364058512
6 <sup>η</sup> Δοκιμή	0.2106507088	0.2857694762	0.1956856951	0.4509269123	0.2209232894	0.2364052706

Πίνακας 12: Αποτελέσματα δοκιμών ατόμου 10 (Mean Squared Errors)

	Neck	R-arm	L-arm	Trunk	R-leg	L-leg
1 <sup>η</sup> Δοκιμή	0.5984893945	0.3017685525	0.2170748381	0.4260930874	0.2814449860	0.2578818397
2 <sup>η</sup> Δοκιμή	0.5984891531	0.3017685578	0.2170748357	0.4260927896	0.2814449860	0.2578818059
3 <sup>η</sup> Δοκιμή	0.5984890218	0.3017690204	0.2170750508	0.4260932801	0.2814449860	0.2578821354
4 <sup>η</sup> Δοκιμή	0.5984889593	0.3017689701	0.2170747903	0.4260934615	0.2814452065	0.2578820632
5 <sup>η</sup> Δοκιμή	0.5984889461	0.3017686438	0.2170745485	0.4260928469	0.2814456387	0.2578816402
6 <sup>η</sup> Δοκιμή	0.5984890096	0.3017690507	0.2170750644	0.4260933843	0.2814449860	0.2578821565

Πίνακας 13: Αποτελέσματα δοκιμών ατόμου 11 (Mean Squared Errors)

	Neck	R-arm	L-arm	Trunk	R-leg	L-leg
1 <sup>η</sup> Δοκιμή	0.7035864834	0.5466147792	0.8818350178	1.2549275896	0.5334906892	1.4964815652
2 <sup>η</sup> Δοκιμή	0.7035866604	0.5466147728	0.8818349098	1.2549275765	0.5334907275	1.4964816601
3 <sup>η</sup> Δοκιμή	0.7035865571	0.5466150034	0.8818351941	1.2549273443	0.5334908301	1.4964815318
4 <sup>η</sup> Δοκιμή	0.7035862169	0.5466149657	0.8818353721	1.2549270865	0.5334909459	1.4964819857
5 <sup>η</sup> Δοκιμή	0.7035856506	0.5466147902	0.8818353413	1.2549276442	0.5334906396	1.4964822005
6 <sup>η</sup> Δοκιμή	0.7035865444	0.5466149947	0.8818352158	1.2549272885	0.5334908411	1.4964816055

Πίνακας 14: Μέσοι όροι αποτελεσμάτων δοκιμών για όλα τα άτομα (Average mean squared errors)

	Neck	R-arm	L-arm	Trunk	R-leg	L-leg
1 <sup>η</sup> Δοκιμή	0.6078184454	0.5054984960	0.3868397106	0.6082273218	0.4587861027	0.5744081155
2 <sup>η</sup> Δοκιμή	0.6248297275	0.5047095668	0.3567003425	0.5648497466	0.4135671718	0.4807289025
3 <sup>η</sup> Δοκιμή	0.6396787772	0.5175909729	0.3506529683	0.5650446627	0.4055154054	0.4749573294
4 <sup>η</sup> Δοκιμή	0.6433523896	0.5289613341	0.3388240869	0.5495131739	0.3932705855	0.4501743069
5 <sup>η</sup> Δοκιμή	0.6251080846	0.5315030829	0.3293780959	0.5343074292	0.3815064538	0.4301190086
6 <sup>η</sup> Δοκιμή	0.6152728649	0.5511534462	0.3203260926	0.5208072533	0.3693567699	0.4056128614

Αν και οι δοκιμές 4 και 5 συχνά εμφανίζουν κοντινές τιμές, θεωρούμε ότι είναι καλύτερο από απόψη αποτελεσμάτων να εξεταστεί το σύνολο των καρτέ ενός βίντεο, καθώς κάθε χρονική στιγμή συνεισφέρει με τα χαρακτηριστικά της στην τελική πρόβλεψη. Άρα, εμείς θεωρούμε ότι την καλύτερη πρόβλεψη τη λαμβάνουμε όταν δοθεί ολόκληρο το βίντεο ενός ατόμου προς αξιολόγηση (δοκιμή 5).

Επίσης, η δοκιμή 6 φαίνεται να έχει καλές τιμές, σύμφωνα με τους παραπάνω πίνακες, αλλά όπως σχολιάσαμε παραπάνω, εμφανίζει πολλές φορές πολύ μικρά και πολύ μεγάλα σφάλματα στις προβλέψεις των επιμέρους βίντεο.

Αξίζει να σημειώσουμε, ότι τα αποτελέσματα μας είτε σε προβλέψεις είτε σε σφάλματα παρουσιάζουν μία ανομοιομορφία, δηλαδή δεν ακολουθούν κάποιο κανόνα είτε στο σύνολο τους, είτε ανά ξεχωριστή τιμή. Αυτό προφανώς ισχυροποιεί το συμπέρασμά μας σχετικά με τη χρήση όλων των καρτέ ενός βίντεο.

Ένα ακόμη σχόλιο που μπορούμε να κάνουμε είναι ότι κάθε άνθρωπος είναι ξεχωριστός, οπότε εκτελεί κάποια δραστηριότητα - κίνηση με μοναδικό τρόπο, αλλά και η εκάστοτε ασθένεια τον επηρεάζει επίσης με μοναδικό τρόπο. Αυτό σημαίνει ότι δεν μπορούμε να προβλέψουμε κατά πόσο κάποιο εκ των έξι μερών του σώματος που αξιολογείται συνεισφέρει περισσότερο ή λιγότερο ή κατά την πρόβλεψη του συστήματος εμφανίζει καλή ή κακή εκτίμηση.

Ακόμη, αξίζει να σημειωθεί ότι ορισμένα εκ των βίντεο που συμπεριλήφθηκαν τόσο στο σύνολο εκπαίδευσης όσο και στο σύνολο ελέγχου μπορεί να είναι κουνημένα ή να εμφανίζουν κάποιες σκιές που δεν ευνοούν τις αξιολογήσεις. Ωστόσο, πάντα θα υπάρχουν τέτοιες περιπτώσεις δεδομένου το σύστημα προορίζεται για καθημερινή χρήση και δεν απαιτεί εξειδικευμένο εξοπλισμό. Ωστόσο, το σύστημά μας πρέπει να είναι πάντα σε θέση να αξιολογήσει όσο καλύτερα μπορεί και τέτοιες περιπτώσεις. Έτσι, αυτές οι περιπτώσεις επηρεάζουν τα αποτελέσματα που παίρνουμε, αυξάνοντας τα σφάλματα.

Επιπλέον, τα βίντεο που μελετάμε δεν έχουν συγκεκριμένη διάρκεια και καρτέ, αλλά ποικίλουν στο μέγεθος. Αυτό, λοιπόν, καθιστά δύσκολη την εξαγωγή κάποιου κανόνα ανά τμήματα του βίντεο, αφού σε κάποια βίντεο θα αντιστοιχούν περισσότερα καρτέ ανά τμήμα από ότι σε άλλα. Γι' αυτό κι εμείς μέσω των

παρατηρήσεων μας αποφαινόμαστε ότι δεν υπάρχει κανόνας για το πόσα καρτέ να χρησιμοποιήσουμε για να γίνει μία σωστή πρόβλεψη.

Τέλος, το σύστημα μας λαμβάνει βίντεο συγκεκριμένων δραστηριοτήτων από τις οποίες μπορεί να αξιολογήσει την κλινική κατάσταση του ασθενή ως προς τη δυσκινησία που προκαλείται από τη δράση της λεβοντόπας. Στο πλαίσιο αυτό, κάθε κίνηση που πραγματοποιείται μεταξύ των εντοπιζόμενων αρθρώσεων παίζει ρόλο και μπορεί να συμβάλει καθοριστικά στην επικείμενη αξιολόγηση. Έτσι, δεν μπορούμε να αποκλείσουμε τη χρήση τμημάτων του συνόλου των καρτέ ειδικά όταν αυτά αντιστοιχούν στις τελευταίες χρονικές στιγμές. Επομένως, υποστηρίζεται τόσο πρακτικά όσο και θεωρητικά το συμπέρασμά μας ότι η χρήση όλου του συνόλου των καρτέ ενός βίντεο θα δώσει πιο ακριβή αξιολόγηση.

Λαμβάνοντας υπόψην όλα τα παραπάνω, συμπεραίνουμε ότι ένα τέτοιο σύστημα είναι εξαιρετικά χρήσιμο, καθώς διακρίνεται για την αποδοτικότητα και αποτελεσματικότητα του στην πλειονότητα των εξεταζόμενων βίντεο. Αυτό επιβεβαιώνεται κι από το average mean squared error που εξάγεται για κάθε μία εκ των έξι τιμών της αξιολόγησης για το σύνολο του προβλήματος με την εφαρμογή του πρωτοκόλλου LOPOCV και απεικονίζεται στον Πίνακα 14 παραπάνω. Συγκεκριμένα, λαμβάνουμε την ακόλουθη λίστα των average mse για κάθε μία από τις έξι προβλεπόμενες τιμές για την είσοδο του 100% των καρτέ καθενός βίντεο στο δίκτυο μας: [0.6251080846, 0.5315030829, 0.3293780959, 0.5343074292, 0.3815064538, 0.4301190086]. Από το αποτέλεσμα αυτό φαίνεται ότι το δίκτυο μας δίνει πολύ χαμηλές τιμές σφαλμάτων για το παρόν πρόβλημα κι άρα ανταποκρίνεται με τον καλύτερο δυνατό τρόπο σε αυτό.

## **Κεφάλαιο 5: Συμπεράσματα & Μελλοντικές επεκτάσεις**

### **5.1 Συμπεράσματα**

Στην παρούσα διπλωματική εργασία μελετήθηκε η συμβολή της μηχανικής μάθησης στην αξιολόγηση δυσκινησίας ασθενών με Πάρκινσον κατά τη δράση της φαρμακευτικής τους αγωγής με λεβοντόπα. Για τον σκοπό αυτόν, αναπτύχθηκε ένα σύστημα που δέχεται, αναλύει, κωδικοποιεί και επεξεργάζεται δεδομένα που λαμβάνονται από τη βιντεοσκόπηση συγκεκριμένων προβλεπόμενων δραστηριοτήτων των ασθενών με στόχο την όσο το δυνατόν καλύτερη και ακριβέστερη αξιολόγηση αυτού του είδους της δυσκινησίας υπό τις έξι τιμές της κλίμακας UDysRS.

Το σύστημα μας βασίζεται σε δεδομένα που έχουν προκύψει από την ανάλυση βίντεο, τα οποία καταγράφονται από μία απλή κάμερα, με τη χρήση μη επεμβατικών αλγορίθμων εκτίμησης πόζας και ανίχνευσης κίνησης. Αυτά επεξεργάζονται και κωδικοποιούνται με τη χρήση της απόστασης σημείου-κατανομής Mahalanobis και του μετασχηματισμού Radon με στόχο την είσοδο τους σε σχετικά απλό και συνάμα αποδοτικό νευρωνικό δίκτυο που θα επιτύχει τη σωστή αξιολόγηση. Ένα τέτοιο νευρωνικό δίκτυο είναι το συνελκτικό δίκτυο του μοντέλου 3 που παρουσιάσαμε και με τη χρήση 10 επιπέδων καταφέρνει να αποδώσει με τον καλύτερο τρόπο στο σύνολο των καρέ ενός βίντεο ως προς την αξιολόγηση του. Συνεπώς, προτείναμε ένα ολοκληρωμένο σύστημα αξιολόγησης δυσκινησίας που βασίζεται στη χρήση απλών οπτικών δεδομένων και σε μεθόδους μηχανικής μάθησης που δεν έχουν επεμβατικό χαρακτήρα.

### **5.2 Μελλοντικές προσεγγίσεις**

Το παρόν σύστημα έχει δυνατότητες εξέλιξης και μπορεί να επεκταθεί μελλοντικά με διάφορους τρόπους συνεισφέροντας σημαντικά στην κλινική αξιολόγηση ενός ασθενή με Πάρκινσον σύμφωνα με την κλίμακα UDysRS και υποβοηθώντας τους γιατρούς να πραγματοποιήσουν τέτοιου είδους εκτιμήσεις.

Αρχικά, μία πρόταση είναι να χρησιμοποιηθούν αρχιτεκτονικές νευρωνικών δικτύων με χαμηλή υπολογιστική πολυπλοκότητα, όπως είναι αυτή του Mobile net, ώστε να εξεταστεί κατά πόσο μία τέτοια αρχιτεκτονική μπορεί να φέρει καλά αποτελέσματα εκτιμήσεων. Επίσης, σε μία τέτοια περίπτωση το σύστημα μας μπορεί να εξελιχθεί, ώστε να γίνει μία εφαρμογή κινητού τηλεφώνου που θα μπορούσε να χρησιμοποιηθεί ακόμη και για αυτοαξιολόγηση από μέρους των ασθενών.

Μία άλλη πρόταση βασίζεται στη διακριτοποίηση εύρους τιμών ανά βαθμολογία στην κλίμακα UDysRS, με σκοπό την ενίσχυση της αποτελεσματικότητας μιας εκτίμησης. Με αυτόν τον τρόπο όταν γίνεται μία εκτίμηση, θα υπάρχει ένα επιτρεπτό εύρος τιμών βάσει της κλίμακας που θα επιτρέπει τον υπολογισμό του ποσοστού ευστοχίας των αξιολογήσεων ανά περίπτωση. Έτσι, η κάθε εκτίμηση δεν έχει μόνο διάσταση τιμής και σφάλματος (mse), αλλά και ακρίβειας (accuracy), κάτι που βοηθά τον αξιολογητή να αποφανθεί αν το σύστημα υποβοήθησης είναι κοντά ή όχι στην πραγματική κλινική κατάσταση του ασθενή. Συνεπώς, στο πρόβλημα μας από την αξιολόγηση της μετρικής mse, μπορούμε να περάσουμε στην αξιολόγηση της μετρικής accuracy.

Επίσης, μία ακόμη ενδιαφέρουσα πρόταση αποτελεί η χρήση επιμέρους νευρωνικών δικτύων για καθένα από τα έξι μέρη του σώματος που αξιολογούνται με τη UDysRS με σκοπό την εξαγωγή μεμονωμένου αποτελέσματος για το καθένα. Με αυτόν τον τρόπο ενισχύεται η μελέτη της συνεισφοράς κάθε μέρους του σώματος και επιλέγονται τα κατάλληλα δίκτυα με στόχο την καλύτερη δυνατή εκτίμηση όλων των τιμών. Ακόμη, μέσω αυτής της προσέγγισης θα μπορούσε να επεκταθεί το πρόβλημα μας και να γίνει από πρόβλημα παλινδρόμησης (regression), πρόβλημα ταξινόμησης (classification) με σκοπό την ανίχνευση του επιπέδου σοβαρότητας των διαταραχών που οφείλονται στη μακροχρόνια χορήγηση λεβοντόπας. Έτσι, μπορούν να ταξινομούνται τα βίντεο των ασθενών ανάλογα με τις βαθμολογίες που πετυχαίνουν σε καθεμία από τις έξι τιμές αξιολόγησης.

Συνεπώς, η παρούσα εργασία συμβάλλει με τον τρόπο της στην υποβοήθηση της ιατρικής πράξης ασχολούμενη με την αξιολόγηση των επιπτώσεων που επιφέρει η νόσος του Πάρκινσον, η οποία αφορά ένα πολύ μεγάλο μέρος του ανθρώπινου πληθυσμού.



## Βιβλιογραφία

- [1] Ι. Βλαχάβας, Π. Κεφαλάς, Ν. Βασιλειάδης, Φ. Κόκκορας, Ι. Σακελλαρίου, Τεχνητή Νοημοσύνη - Δ' Έκδοση, Εκδόσεις Πανεπιστημίου Μακεδονίας, 2020.
- [2] Stuart Russel, Peter Norvig, Τεχνητή Νοημοσύνη, μια σύγχρονη προσέγγιση, Prentice Hall, 2003.
- [3] P. Simon, Too Big to Ignore: The Business Case for Big Data, Wiley, 2013, p. 89.
- [4] John R.; Bennett, Forrest H.; Andre, David; Keane, Martin A., «Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming,» σε *Artificial Intelligence in Design '96*, Springer, 1996, p. 151–170.
- [5] Tom M. Mitchell, McGraw-Hil, Machine Learning Definition, 1997, p. 1.
- [6] B. Farley, W. Clark, Simulation of self-organizing systems by digital computer, τόμ. 4, Trans. IRE Prof. Gr. Inf. Theory, 1954, pp. 76-84.
- [7] N. Rochester, J. H. Holland, L. H. Haibt, and W. L. Duda, Tests on a cell assembly theory of the action of the brain, using a large digital computer, τόμ. 2, IRE Trans. Inf. Theory, 1956, pp. 80-93.
- [8] Ε. Νταρούις, «ΔΙΑΓΝΩΣΗ ΤΗΣ ΝΟΣΟΥ ΠΑΡΚΙΝΣΟΝ ΜΕΣΩ ΒΑΔΙΣΗΣ ΜΕ ΤΕΧΝΙΚΕΣ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ,» ΕΜΠ, 2020.
- [9] S. Haykin, Νυρωνικά Δίκτυα και Μηχανική Μάθηση (3η Έκδοση), Παπασωτηρίου, 2020.
- [10] G. Caridakis, K. Karpouzis, M. Wallace, L. Kessous, N. Amir, «Multimodal users affective state analysis in naturalistic interaction,» *Journal on Multimodal User Interfaces*, τόμ. 3, pp. 49-66.
- [11] J. Nijhar, N. Bianchi-Berthouze, G. Boguslawski , «Does Movement Recognition Precision Affect the Player Experience in Exertion Games?,» *Proceedings of INTETAIN*, p. 73–82, 2011.
- [12] N. Bianchi-Berthouze, K. Isbister, « Emotion and Body-Based Games: Overview and Opportunitie,» σε *K. Karpouzis, G. N. Yannakakis (eds.)*,

*Emotion in Games: Theory and Praxis* , Springer, p. 235–255.

- [13] Zhang, H.-B., Zhang, Y.-X., Zhong, B., Lei, Q., Yang, L., Du, J.-X., Chen, D.-S., «A Comprehensive Survey of Vision-Based Human Action Recognition Methods,» *Sensors*, αρ. 1005, 2019.
- [14] Dawn, D.D., Shaikh, S.H., «A comprehensive survey of human action recognition with spatio-temporal interest point (STIP) detector,» *Vis.Comput.* , τόμ. 32, p. 289–306, 2016.
- [15] Nguyen, T.V., Song, Z., Yan, S.C., «STAP: Spatial-Temporal Attention-Aware Pooling for Action Recognition,» *IEEE Trans. Circ. Syst. Video Technol.*, τόμ. 25, p. 77–86, 2015.
- [16] Shao, L., Zhen, X.T., Tao, D.C., Li, X.L., «Spatio-Temporal Laplacian Pyramid Coding for Action Recognition,» *IEEE Trans. Cybern.*, τόμ. 44, p. 817–827, 2014.
- [17] Burghouts, G.J., Schutte, K., ten Hove, R.J.M., van den Broek, S.P., Baan, J., Rajadell, O., van Huis, J.R., van Rest, J., Hanckmann, P., Bouma, H., et al, «Instantaneous threat detection based on a semantic representation of activities, zones and trajectories,» *Signal Image Video Process*, τόμ. 8, p. 191–200, 2014.
- [18] Wang, H., Schmid, C., «Action recognition with improved trajectories,» σε *ICCV*, Sydney, NSW, Australia, 2013.
- [19] Yang, X., Tian, Y.L., «Super Normal Vector for Activity Recognition Using Depth Sequences,» σε *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 2014.
- [20] Ye, M., Zhang, Q., Wang, L., Zhu, J., Yang, R., Gall, J., «A survey on human motion analysis from depth data,» σε *Dagstuhl 2012 Seminar on Time-of-Flight Imaging: Sensors, Algorithms, and Applications and Workshop on Imaging New Modalities, GCPR 2013*, Saarbrucken, Germany, 2012.
- [21] Oreifej, O., Liu, Z., «HON4D: Histogram of Oriented 4D Normals for Activity Recognition from Depth Sequences,» σε *Computer Vision and Pattern Recognition*, Portland, OR, USA, 2013.
- [22] Li, M., Leung, H., Shum, H.P.H., «Human action recognition via skeletal and depth based feature fusion,» σε *Motion in Games 2016*, Burlingame, CA, USA, 2016.
- [23] Yang, X., Tian, Y.L., «Effective 3D action recognition using EigenJoints,» *J. Vis. Commun. Image Represent*, τόμ. 25, pp. 2-11, 2014.

- [24] [Ηλεκτρονικό]. Available: [https://miro.medium.com/max/1200/1\\*Q8sk8FwMIYuLoLmptsT-MA.png](https://miro.medium.com/max/1200/1*Q8sk8FwMIYuLoLmptsT-MA.png).
- [25] Microsoft, Start Here! Learn the Kinect API, 2012.
- [26] V. Ramakrishna, D. Munoz, M. Hebert, J. Bagnell, and Y. Sheikh., «Pose Machines: Articulated Pose Estimation via Inference Machines,» *ECCV*, 2014.
- [27] Shih-En Wei, Varun Ramakrishna, Takeo Kanade, Yaser Sheikh, «Convolutional Pose Machines,» *CVPR*, 2016.
- [28] [Ηλεκτρονικό]. Available: <https://image.slidesharecdn.com/160716cvcpm-160717060432/95/convolutional-pose-machines-3-638.jpg?cb=1503023208>.
- [29] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh, «OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,» *CVPR*, 2019.
- [30] [Ηλεκτρονικό]. Available: <http://www.audiomotion.com/blog/what-is-motion-capture.html>.
- [31] [Ηλεκτρονικό]. Available: <https://www.mo-sys.com/what-is-motion-capture-and-how-does-it-work/>.
- [32] C. Sun, I. N. Junejo, M. Tappen, H. Foroosh, «Exploring sparseness and self-similarity for action recognition,» *IEEE Transactions on Image Processing*, τόμ. 24, pp. 2488-2501, 2015.
- [33] Y. Li, R. Xia, Q. Huang, W. Xie and X. Li, «Survey of Spatio-Temporal Interest Point Detection Algorithms in Video,» *IEEE Access*, τόμ. 5, pp. 10323-10331, 2017.
- [34] X. Li, J. Shi, Y. Dong and D. Tao, «A survey on scene image classification,» *Scientia Sinica Inf.*, τόμ. 45, αρ. 7, p. 827, 2015.
- [35] C. Wang and C. Dong, «Spatial-temporal words learning for crowd behavior recognition,» *Int. J. Sci. Eng. Invest.*, τόμ. 1, αρ. 3, pp. 1-6, 2012.
- [36] Rapantzikos, Konstantinos and Tsapatsoulis, Nicolas and Avrithis, Yannis and Kollias, Stefanos, «Spatiotemporal saliency for video classification,» *Signal Processing: Image Communication*, τόμ. 24, αρ. 7, pp. 557--571, 2007.
- [37] Rapantzikos, Konstantions and Avrithis, Yannis and Kollias, Stefanos, «Spatiotemporal saliency for event detection and representation in the 3D wavelet domain: potential in human action recognition,» σε *Proceedings of the 6th ACM international conference on Image and video retrieval*, 2007, pp. 294-

-301.

- [38] Tsapatsoulis, Nicolas and Kollias, Stefanos, «Face detection in color images and video sequences,» σε *2000 10th Mediterranean Electrotechnical Conference. Information Technology and Electrotechnology for the Mediterranean Countries. Proceedings. MeleCon 2000*, 2000.
- [39] Rapantzikos, Konstantinos and Tsapatsoulis, Nicolas and Avrithis, Yannis and Kollias, Stefanos, «Bottom-up spatiotemporal visual attention model for video analysis,» *IET Image Processing*, τόμ. 1, αρ. 2, pp. 237--248, 2007.
- [40] Y. Li, W. Liu and Q. Huang, «Traffic anomaly detection based on image descriptor in videos,» *Multimedia Tools Appl.*, τόμ. 75, αρ. 5, pp. 2487-2505, 2016.
- [41] R. Laganière, R. Bacco, A. Hocevar, P. Lambert, G. Païs and B. E. Ionescu, «Video summarization from spatio-temporal features,» σε *Proc. ACM Workshop Video Summarization*, 2008.
- [42] C. Chattopadhyay and A. K. Maurya, «Multivariate time series modeling of geometric features of spatio-temporal volumes for content based video retrieval,» *Int. J. Multimedia Inf. Retr.*, τόμ. 3, αρ. 1, pp. 15-28, 2014.
- [43] W. Zhou, H. Hu and P. J. Smith, «Recent advances in camera tracking and virtual reality,» *Proc. Inf. Technol. Proc. Int. Symp. Inf. Technol. (ISIT)*, pp. 14-16, 2014.
- [44] E. Rosten and T. Drummond, «Machine learning for high-speed corner detection,» σε *Computer Vision-ECCV*, Berlin, Germany.
- [45] Dawn, D.D., Shaikh, S.H., « A comprehensive survey of human action recognition with spatio-temporal interest point (STIP) detector,» *Vis. Comput.*, τόμ. 32, p. 289–306, 2016.
- [46] Laptev, I., «On space-time interest points,» *Int. J. Comput. Vis.*, τόμ. 64, p. 107–123, 2005.
- [47] Chakraborty, B., Holte, M.B., Moeslund, T.B., Gonzalez, J. , «Selective spatio-temporal interest points,» *Comput. Vis. Image Underst.*, τόμ. 116, p. 396–416, 2012.
- [48] G. Goudelis, G. Tsatiris, K. Karpouzis and S. Kollias, «3D cylindrical trace transform based feature extraction for effective human action classification,» σε *2017 IEEE Conference on Computational Intelligence and Games (CIG)*, 2017.

- [49] Miao Yu, Liyun Gong, and Stefanos Kollias, «Computer vision based fall detection by a convolutional neural network,» σε *19th ACM International Conference on Multimodal Interaction (ICMI '17)*, 2017.
- [50] Tsatiris G., Karpouzis K., Kollias S., «A Compact Sequence Encoding Scheme for Online Human Activity Recognition in HRI Applications,» σε *Proceedings of the 21st EANN (Engineering Applications of Neural Networks) 2020 Conference. EANN 2020*, 2020.
- [51] Tsatiris, G., Karpouzis, K., Kollias, S., «Variance-based shape descriptors for determining the level of expertise of tennis players,» σε *2017 9th International Conference*, 2017.
- [52] R. Kurose, M. Hayashi, T. Ishii and Y. Aoki, «Player pose analysis in tennis video based on pose estimation,» σε *2018 International Workshop on Advanced Image Technology (IWAIT)*, 2018.
- [53] Danica Hendry, Kevin Chai, Amity Campbell, Luke Hopper, Peter O’Sullivan, Leon Straker S, «Development of a Human Activity Recognition System for Ballet Tasks,» *Sports Medicine*, τόμ. 6, αρ. 10, 2020.
- [54] C. E, «Guide to medical informatics, the Internet and telemedicine,» Chapman & Hall, Ltd, 1997.
- [55] P. B, «Artificial Intelligence Is Almost Ready for Business,» Massachusetts General Hospital, 2015.
- [56] Bahl M, Barzilay R, Yedidia AB, Locascio NJ, Yu L, Lehman CD, «High-Risk Breast Lesions: A Machine Learning Model to Predict Pathologic Upgrade and Reduce Unnecessary Surgical Excision,» *Radiology*, τόμ. 286, p. 810–818, 2018.
- [57] L. L, «Here's How IBM Watson Health is Transforming the Health Care Industry,» *Fortune*, 2016.
- [58] B.-B. S, «NHS using Google technology to treat patients,» 2016.
- [59] Esteva, Andre, Kuprel, Brett, Novoa, Roberto A., Ko, Justin, Swetter, Susan M., Blau, Helen M., Thrun, Sebastian, «Dermatologist-level classification of skin cancer with deep neural networks,» *Nature*, τόμ. 542, αρ. 7639, pp. 115-118, 2017.
- [60] Harmon, S.A., Sanford, T.H., Xu, S. et al., «Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets,» *Nat Commun*, τόμ. 11, αρ. 4080, 2020.

- [61] Mahtab Mirmomeni, Timothy Fazio, Stefan von Cavallar, Stefan Harrer, «From wearables to THINKables: artificial intelligence-enabled sensors for health monitoring,» σε *Wearable Sensors (Second Edition)*, Academic Press, 2021, pp. 229-256.
- [62] Karunanithi, D & Rodrigues, P, *Diagnosis of Disease through Voice Recordings using Artificial Neural Networks*, 2017.
- [63] Κ. Βεκρελλής, «Νόσος του Πάρκινσον: Νεότερες εξελίξεις στη διάγνωση και θεραπεία».
- [64] Υ. Υγείας, «Παγκόσμια Ημέρα κατά της ασθένειας του Πάρκινσον,» 2012. [Ηλεκτρονικό]. Available: <https://www.moh.gov.gr/articles/news/1133-pagkosmia-hmera-kata-ths-astheneias-toy-parkinson>.
- [65] Μεντενόπουλος Γ, Μπούρας Κ. , *Η νόσος του Parkinson*, Αθήνα: UNIVERSITY STUDIO PRESS, 2008.
- [66] J. Kassubek, «Diagnostic procedures during the course of Parkinson's Disease,» *Basal Ganglia*, pp. 15-18, 2014.
- [67] Fahn S & Elton R., «Unified Parkinson's disease Rating Scale – UPDRS,» 1987.
- [68] Τσίντου, Μ, Δαλαμάγκας, Κ., «Νόσος Parkinson,» 2011.
- [69] Christopher G. Goetz et al., «MDS-UPDRS- Revision of the Unified Parkinson's Disease Rating Scale,» International Parkinson and Movement Disorder Society, 2015.
- [70] «Clinical effectiveness and cost-effectiveness of physiotherapy and occupational therapy versus no therapy in mild to moderate Parkinson's disease: a large pragmatic randomised controlled trial,» [Ηλεκτρονικό]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK379751/>.
- [71] «Parkinson's Disease and Occupational Therapy,» [Ηλεκτρονικό]. Available: <https://parkinsons-info.weebly.com/updrs.html>.
- [72] Richard A. Harvey, Pamela C. Champe, *Φαρμακολογία (3η έκδοση)*, Εκδόσεις Παρισιάνου, 2007, pp. 103-105.
- [73] «Rating-Scales/UDysRS\_Greek\_Official\_Translation,» [Ηλεκτρονικό]. Available: [https://www.movementdisorders.org/MDS-Files1/Education/Rating-Scales/UDysRS\\_Greek\\_Official\\_Translation.pdf](https://www.movementdisorders.org/MDS-Files1/Education/Rating-Scales/UDysRS_Greek_Official_Translation.pdf).

- [74] Vlaar, A.M.M., et al, «Meta-analysis of the literature on diagnostic accuracy of SPECT in parkinsonian syndromes,» *BMC neurology*, τόμ. 7, αρ. 1, pp. 27-33, 2007.
- [75] Tang, C.C., et al., «Differential diagnosis of parkinsonism: a metabolic imaging study using pattern analysis,» *The Lancet Neurology*, τόμ. 9, αρ. 2, pp. 149-158, 2010.
- [76] Rahn, D.A, Chou, M, Jiang, J.J. and Zhang, Y., «Phonatory Impairment in Parkinson's disease: Evidence from Nonlinear Dynamic Analysis and Perturbation Analysis,» 2007.
- [77] Δ. Πετσάνη, «Συναισθηματική ανάλυση κειμένου και ανάλυση εικόνας προσώπου και η εφαρμογή τους στην πρόιμη διάγνωση της νόσου του Πάρκινσον: η εφαρμογή iPrognosis,» ΑΠΘ, 2017.
- [78] M. Ene, «Neural network-based approach to discriminate healthy people From those with Parkinson's disease,» 2008.
- [79] Clayton R. Pereira, Danilo R. Pereira, Silke A.T. Weber, Christian Hook, Victor Hugo C. de Albuquerque, João P. Papa, «A survey on computer-assisted Parkinson's Disease diagnosis,» *Artificial Intelligence in Medicine*, τόμ. 95, pp. 48-63, 2019.
- [80] Jinee Goyal, Padmavati Khandnor, Trilok Chand Aseri, «Classification, Prediction, and Monitoring of Parkinson's disease using Computer Assisted Technologies: A Comparative Analysis, Engineering Applications of Artificial Intelligence,» τόμ. 96, 2020.
- [81] Tay A, Yen SC, Lee PY, Wang CY, Neo A, Phan SW, Yogaprakash K, Liew S, Au WL., «Freezing of gait (FoG) detection for Parkinson disease,» σε *10th Asian control conference*, 2015.
- [82] Lorenzi P, Rao R, Romano G, Kita A, Serpa M, Filesi F, Irrera F, Bologna M, Suppa A, Berardelli A., «Smart sensors for the recognition of specific human motion disorders in Parkinson's disease,» σε *IEEE 6th international workshop on advances in sensors and interfaces*, 2015.
- [83] Dong Z, Gu H, Wan Y, Zhuang W, Rojas-Cessa R, Rabin E., «Wireless body area sensor network for posture and gait monitoring of individuals with Parkinson's disease,» σε *IEEE 12th international conference on networking, sensing and control*, 2015.
- [84] Eskofier, B.M., Lee, S.I., Baron, M., Simon, A., Martindale, C.F., Gaßner, H., Klucken, J., «An Overview of Smart Shoes in the Internet of Health Things:

- Gait and Mobility Assessment in Health Promotion and Disease Monitoring,» *Appl. Sci.*, τόμ. 7, αρ. 986, 2017.
- [85] A. Almogren, «An automated and intelligent Parkinson disease monitoring system using wearable computing and cloud technology,» *Cluster Comput*, τόμ. 22, p. 2309–2316, 2019.
- [86] Dang, Q.K., Seo, H.G., Pham, D.D., Chee, Y. , «Wearable Sensor Based Stopped Posture Estimation in Simulated Parkinson’s Disease Gaits,» *Sensors*, τόμ. 19, αρ. 223, 2019.
- [87] Kostikis N, Hristu-Varsakelis D, Arnaoutoglou M, Kotsavasiloglou C., «A smartphone-based tool for assessing parkinsonian hand tremor,» *IEEE J Biomed Health Inform*, τόμ. 19, αρ. 6, 2015.
- [88] P. Anju, A. Varghese, A. Roy, S. Suresh, E. Joy and R. Sunder, «Recent Survey on Parkinson Disease Diagnose using Deep Learning Mechanism,» σε *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, 2020.
- [89] Arango Paredes, J. D., Muñoz, B., Agredo, W., Ariza-Araújo, Y., Orozco, J. L., & Navarro, A, «A reliability assessment software using Kinect to complement the clinical evaluation of Parkinson's disease,» σε *IEEE Engineering in Medicine and Biology Society*, 2015.
- [90] Buongiorno D, Bortone I, Cascarano GD, Trotta GF, Brunetti A, Bevilacqua V., «A low-cost vision system based on the analysis of motor features for recognition and severity rating of Parkinson's Disease,» σε *BMC Med Inform Decis Mak.*, 2019.
- [91] E. Celik, S. I. Omurca, «Improving Parkinson's Disease Diagnosis with Machine Learning Methods,» σε *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, 2019.
- [92] Yu Liu, Jiansheng Chen, Chunhua Hu, Yu Ma, Dongyun Ge, Suhua Miao, Youze Xue, Luming Li, «Vision-Based Method for Automatic Quantification of Parkinsonian Bradykinesia,» *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, τόμ. 27, αρ. 10, 2019.
- [93] Sato K, Nagashima Y, Mano T, Iwata A, Toda T, «Quantifying normal and parkinsonian gait features from home movies: Practical application of a deep learning-based 2D pose estimator,» *PLOS ONE*, τόμ. 14, αρ. 11, 2019.



- [94] Raina Langevin, Mohammad Rafayet Ali, Taylan Sen, Christopher Snyder, Taylor Myers, E. Ray Dorsey, Mohammed Ehsan Hoque, «The PARK Framework for Automated Analysis of Parkinson's Disease Characteristics.» *ACM Interact. Mob. Wearable Ubiquitous Technol*, τόμ. 3, αρ. 2, p. 22, 2019.
- [95] Amirhossein Dadashzadeh, Alan Whone , Michal Rolinski , Majid Mirmehdi, «Exploring Motion Boundaries in an End-to-End Network for Vision-based Parkinson's Severity Assessment,» 2020.
- [96] Lu M, Poston K, Pfefferbaum A, Sullivan EV, Fei-Fei L, Pohl KM, Niebles JC, Adeli E., «Vision-based Estimation of MDS-UPDRS Gait Scores for Assessing Parkinson's Disease Motor Severity,» σε *Med Image Comput Comput Assist Interv*, 2020.
- [97] Jin, B., Qu, Y., Zhang, L., & Gao, Z., «Diagnosing Parkinson Disease Through Facial Expression Recognition: Video Analysis,» *Journal of medical Internet research*, τόμ. 22, αρ. 7, 2020.
- [98] Channa A, Ifrim RC, Popescu D, Popescu N., «A-WEAR Bracelet for Detection of Hand Tremor and Bradykinesia in Parkinson's Patients,» *Sensors (Basel)*, τόμ. 21, αρ. 3, 2021.
- [99] Deval Mehta, Umar Asif, Tian Hao, Erhan Bilal, Stefan von Cavallar, Stefan Harrer, Jeffrey Rogers, «Towards Automated and Marker-Less Parkinson Disease Assessment: Predicting UPDRS Scores Using Sit-Stand Videos,» σε *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2021.
- [100] [Ηλεκτρονικό]. Available: <https://machinemedicine.com/kelvin-pd/>.
- [101] Kollias, Dimitrios and Tagaris, Athanasios and Stafylopatis, Andreas and Kollias, Stefanos and Tagaris, Georgios, «Deep neural architectures for prediction in healthcare,» *Complex \& Intelligent Systems*, τόμ. 4, αρ. 2, pp. 119-131, 2018.
- [102] D. Kollias, M. Yu, A. Tagaris, G. Leontidis, A. Stafylopatis and S. Kollias, «Adaptation and contextualization of deep neural network models,» σε *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2017.
- [103] Tagaris, Athanasios and Kollias, Dimitrios and Stafylopatis, Andreas and Tagaris, Georgios and Kollias, Stefanos, «Machine learning for neurodegenerative disorder diagnosis—survey of practices and launch of benchmark dataset,» *International Journal on Artificial Intelligence Tools*, τόμ. 27, αρ. 3, 2018.

- [104] Wingate, James and Kollia, Ilianna and Bidaut, Luc and Kollias Stefanos, «Unified deep learning approach for prediction of Parkinson's Disease,» *IET Image Processing*, τόμ. 14, αρ. 10, pp. 1980--1989, 2020.
- [105] Tagaris, Athanasios and Kollias, Dimitrios and Stafylopatis, Andreas, «Assessment of Parkinson's disease based on deep neural networks,» σε *International Conference on Engineering Applications of Neural Networks*, Springer, 2017, pp. 391--403.
- [106] I. Kollia, A. -G. Stafylopatis and S. Kollias, «Predicting Parkinson's Disease using Latent Information extracted from Deep Neural Networks,» σε *2019 International Joint Conference on Neural Networks (IJCNN)*, 2019.
- [107] Kollias, D and Bouas, N and Vlaxos, Y and Brillakis, V and Seferis, M and Kollia, I and Sukissian, L and Wingate, J and Kollias, S., «Deep Transparent Prediction through Latent Representation Analysis,» *arXiv preprint arXiv:2009.07044*, 2020.
- [108] Michael H. Li, Tiago A. Mestre, Susan H. Fox, Babak Taati, «Vision-based assessment of parkinsonism and levodopa-induced dyskinesia with pose estimation,» *Journal of NeuroEngineering and Rehabilitation*, τόμ. 15, 2018.
- [109] Michael H. Lia, Tiago A. Mestre, Susan H. Foxc, Babak Taati, «Automated assessment of levodopa-induced dyskinesia: Evaluating the responsiveness of video-based features,» *Parkinsonism and Related Disorders*, 2018.
- [110] Wang, F., Jiang, M., Qian, C., Yang, S., Li, C., Zhang, H., Wang, X., Tang, X., «Residual attention network for image classification,» σε *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [111] S. R. Deans, *The Radon Transform and Some of Its Applications*, Krieger Publishing Company, 1983.
- [112] A.A. Liu, Y.T. Su, P.P. Jia, Z. Gao, T. Hao, Z.X. Yang, «Multiple/single view human action recognition via part-induced multitask structural learning,» *IEEE transactions on cybernetics*, τόμ. 45, αρ. 6, 2015.