

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Matjaž Zupanič

**Avtomatsko ocenjevanje testa
tapkanja s prsti pri bolnikih s
Parkinsonovo boleznijo**

MAGISTRSKO DELO
MAGISTRSKI ŠTUDIJSKI PROGRAM DRUGE STOPNJE
RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: doc. dr. Jure Žabkar
SOMENTOR: doc. dr. Dejan Georgijev, dr. med.

Ljubljana, 2024

To delo je ponujeno pod licenco *Creative Commons Priznanje avtorstva-Deljenje pod enakimi pogoji 2.5 Slovenija* (ali novejšo različico). To pomeni, da se tako besedilo, slike, grafi in druge sestavine dela kot tudi rezultati diplomskega dela lahko prosto distribuirajo, reproducirajo, uporabljajo, priobčujejo javnosti in predelujejo, pod pogojem, da se jasno in vidno navede avtorja in naslov tega dela in da se v primeru spremembe, preoblikovanja ali uporabe tega dela v svojem delu, lahko distribuira predelava le pod licenco, ki je enaka tej. Podrobnosti licence so dostopne na spletni strani creativecommons.si ali na Inštitutu za intelektualno lastnino, Streliška 1, 1000 Ljubljana.



Izvorna koda diplomskega dela, njeni rezultati in v ta namen razvita programska oprema je ponujena pod licenco GNU General Public License, različica 3 (ali novejša). To pomeni, da se lahko prosto distribuira in/ali predeluje pod njenimi pogoji. Podrobnosti licence so dostopne na spletni strani <http://www.gnu.org/licenses/>.

ZAHVALA

Zahvaljujem se mentorju doc. dr. Juretu Žabkarju in somentorjo doc. dr. Dejanu Georgievu, dr. med. za pomoč in odzivnost. Zahvaljujem se tudi vsem udeležencem, ki so izvedli test tapkanja. Prav tako se zahvaljujem družini in svoji puncici za vzpodbudo in podporo.

Matjaž Zupanič, 2024

Bog ne kocka.

"God doesn't play dice."

— Albert Einstein

Kazalo

Povzetek

Abstract

1	Uvod	1
1.1	Parkinsonova bolezen	1
1.2	Sorodna dela	6
1.3	Motivacija	13
1.4	Prispevki	14
1.5	Struktura naloge	14
2	Tehnične metode	17
2.1	Zaznava roke in skeleta	17
2.2	Filter s končnim impulznim odzivom	18
2.3	Filter z neskončnim impulznim odzivom	19
2.4	Detekcija tapkanja	20
2.5	Sledenje zamrznitvam in prekinitvam	21
2.6	Avtokorelacija	23
2.7	Vizualizacija značiln	23
2.8	Metrike za vrednotenje	25
2.9	Izbira značiln	26
2.10	Klasifikatorji	28

KAZALO

3	Raziskovalna metodologija	33
3.1	Podatkovna zbirka	33
3.2	Predprocesiranje podatkov	35
3.3	Metoda 1	38
3.4	Metoda 2	43
3.5	Metoda 3	47
3.6	Orodje za vizualizacijo	48
4	Vrednotenje	51
4.1	Rezultati	51
5	Sklepne ugotovitve	59
5.1	Diskusija	59
5.2	Nadaljnje delo	69

Seznam uporabljenih kratic

kratica	angleško	slovensko
PTEN	phosphatase and tensin homolog deleted on chromosome 10	homologa fosfataze in tenzina izbrisana na kromosomu 10
RBR	RING-between-RING	domena RING med domeno RING
RNA	ribonucleic acid	ribonukleinska kislina
GWAS	genome-wide association studies	študije na ravni celotnega genoma
GBA	glucocerebrosidase	glukocerebrozidaze
PRKN	parkin RBR E3 ubiquitin protein ligase	parkina RBR E3 ubikvitin pro- tein ligaze
LRRK2	leucine-rich repeat kinase 2	z levcinskimi ponovitvami bogate kinaze 2
RING	really interesting new gene	zelo zanimiv nov gen
PCA	principal component analysis	metoda glavnih komponent
CNN	convolutional neural network	konvolucijska nevronska mreža
FCN	fully convolutional neural net- work	polna konvolucijska nevronska mreža
SVM	support vector machine	metoda podpornih vektorjev
K-NN	k-nearest neighbours	k-najbližjih sosedov
MLP	multi layer perceptron	večplastni perceptron
RF	random forest	naključni gozdovi

KAZALO

kratica	angleško	slovensko
MAD	mean absolute deviation	povprečno absolutno odstopanje
MAE	mean absolute error	povprečna absolutna napaka
FIR	finite response filter	filter s končnim impulznim odzivom
IIR	infinite response filter	filter z neskončnim impulznim odzivom
FFT	fast Fourier transform	hitra Fourierjeva transformacija
MDS-UPDRS	the movement disorder society-sponsored revision of the unified Parkinson's disease rating scale	revizija enotne lestvice za ocenjevanje Parkinsonove bolezni, ki jo je podprlo društvo za motnje gibanja
ANOVA	analysis of variance	analiza variance
BoostRFE	boosted recursive feature elimination	ojačana rekurzivna eliminacija značilnk

Povzetek

Naslov: Avtomatsko ocenjevanje testa tapkanja s prsti pri bolnikih s Parkinsonovo boleznijo

Parkinsonova bolezen je kronična nevrodegenerativna bolezen, ki močno poslabša kvaliteto življenja pacientov. Čakalne vrste za nevrologa so danes precej dolge, pacienti pa so v tem obdobju brez ustrezne terapije s katero bi si olajšali simptome. Zato smo razvili avtomatsko metodo za določanje stopnje motorične prizadetosti oziroma bradikinezije Parkinsonove bolezni na podlagi testa tapkanja in s tem omogočili hitrejšo postavitev diagnoze. Zbrali smo 183 video posnetkov tapkanja, posnetih kar s pametnim telefonom v vsakdanjem okolju. Videe je v 5 razredov lestvice MDS-UPDRS ocenil nevrolog. Za prepoznavo roke smo uporabili MediaPipe Hand, ki nam kot rezultat vrne časovno vrsto skeleta roke. Za klasifikacijo smo ubrali dva različna pristopa. Prvič smo iz časovne vrste skeleta roke sami sestavili značilke, enkrat strogo po lestvici MDS-UPDRS, drugič pa se te nismo strogo držali. Te značilke smo nato uporabili v klasifikatorjih in z večplastnim perceptronom dosegli 61 % točnost in 0,62 F1 vrednost. V drugem pristopu smo časovno vrsto razdalj med palcem in kazalcem uporabili neposredno v polnem konvolucijskem nevronske omrežju in dosegli 77 % točnost in 0,75 F1 vrednost. Izdelali smo še orodje za vizualizacijo tapkanja in izpis ključnih podatkov.

Ključne besede

klasifikacija, Parkinsonova bolezen, test tapkanja s prsti, strojno učenje

Abstract

Title: Automatic assessment of finger tapping in patients with Parkinson's disease

Parkinson's disease is a chronic neurodegenerative disorder that severely impairs patients' quality of life. Currently, the waiting lists for a neurologist are quite long, and during this period, patients are without adequate therapy to alleviate their symptoms. Therefore, we have developed an automatic method to determine the level of motor impairment or bradykinesia of Parkinson's disease based on a tapping test, thus enabling faster diagnosis. We collected 183 tapping videos recorded with a smartphone in everyday environments, which were assessed by a neurologist into 5 classes of the MDS-UPDRS scale. For hand detection we used MediaPipe Hand, which returns a time series of the hand skeleton. For classification, we took two different approaches. First, we constructed features from the hand skeleton time series, once strictly following the MDS-UPDRS scale, and another time not strictly adhering to it. These features were then used in classifiers and achieved 61 % accuracy and 0,62 F1 score using a multi-layer perceptron. In the second approach, we used the time series of thumb-pointer distances directly in a fully convolutional neural network achieving 77 % accuracy and 0,75 F1 score. We also created a tool for visualizing tapping and displaying key data.

Keywords

classification, Parkinson's disease, finger tapping test, machine learning

Poglavje 1

Uvod

Magistrsko delo povezuje področje računalništva in medicine. V delu je predstavljen razvoj modela za avtomatsko prepoznavo in klasifikacijo stopnje bradikinezije, na podlagi testa tapkanja s prsti, pri Parkinsonovi bolezni po standardizirani lestvici MDS-UPDRS (*angl. unified Parkinson's disease rating scale*) [1] za oceno tapkanja iz video posnetkov tapkanja s prsti. Pri razvitem modelu za klasifikacijo ni potrebe po uporabi posebnih senzorjev ali naprav, zadostuje že pametni telefon ali kamera računalnika. V sklopu magistrskega dela smo tako želeli ugotoviti kako natančno lahko iz video posnetkov tapkanja, posnetimi s pametnimi telefoni, določimo stopnjo bradikinezije, ki je vodilni simptom in znak Parkinsonove bolezni.

1.1 Parkinsonova bolezen

Parkinsonova bolezen je kronična nevrodegenerativna bolezen, katere etiologija še ni poznana. Hitrost napredovanja pa se razlikuje med posamezniki. Po pogostosti jo uvrščamo na drugo mesto nevrodegenerativnih bolezni, takoj za Alzheimerjevo boleznijo, saj z njo oboleva med 1 do 2 % populacije starejše od 65 let [2]. Samo v Sloveniji je bilo leta 2018 za Parkinsonovo boleznijo obolelih več kot 10.000 bolnikov. V Evropi pa z njo oboleva več kot 1,2 milijona bolnikov [3]. Bolezen je pogostejša med starejšo populacijo,

posledično je njena pojavnost v porastu zaradi demografskega pojava staranja prebivalstva. Kot predhodno že omenjeno, vzrok za razvoj Parkinsonove bolezni še ni znan, a se predvideva, da k nastanku prispevajo tako genetski, kot tudi okoljski dejavniki [4].

Genetski dejavniki predstavljajo vzrok za razvoj Parkinsonove bolezni pri približno 10 do 15 % obolelih [5, 6]. V zadnjem desetletju je bilo mnogo raziskav usmerjenih prav v dešifriranje genetske strukture Parkinsonove bolezni. Nove tehnologije kot so strojno učenje, sekvenciranje ribonukleinskih kislin (RNA) na ravni posameznih celic (*angl. single-cell RNA sequencing*) in visoko zmogljivostno presejanje bodo pripomogle k razumevanju genetskega tveganja za nastanek bolezni [7]. Študije GWAS (*angl. genome-wide association studies*) [8] so pokazale, da obstaja več kot 90 medsebojno neodvisnih alelov, ki so povezani s povečanim tveganjem za razvoj bolezni [7]. Izmed teh so najbolj raziskane mutacije genov sinukleina alfa (SNCA, *angl. synuclein alpha*), z levcinskimi ponovitvami bogate kinaze 2 (LRRK2, *angl. leucine-rich repeat kinase 2*), parkina RBR E3 ubikvitin protein ligaze (PRKN, *angl. parkin RBR E3 ubiquitin protein ligase*), PTEN-inducirane kinaze 1 (PINK1, *angl. PTEN-induced kinase 1*) in glukocerebrozidaze (GBA, *angl. glucocerebrosidase*) [9].

Tveganje za Parkinsonovo bolezen je povezano tudi z različnimi okoljskimi dejavniki, vendar je izjemno težko opredeliti tiste, ki so ključni za nastanek bolezni, saj so le ti mnogokrat prepleteni z genetskimi dejavniki. Med znane okoljske vplive štejemo izpostavljenost okoljskim onesnaževalcem, kot so pesticidi, težke kovine, industrijske kemikalije in onesnažen zrak [10]. Tveganje za nastanek Parkinsonove bolezni povečajo tudi poškodbe glave (npr. travmatska poškodba možganov). Prav tako pa se pojavnost bolezni med spoloma razlikuje, namreč moški so bolj dovzetni za razvoj bolezni [9]. Raziskave so pokazale, da bolezen spremlja tudi spremenjen črevesni mikrobiom [11, 12].

Bolezen se razvije kot posledica poškodbe ali propada dopaminergičnih nevronov v substanci nigra (*lat. substantia nigra*). Slednja predstavlja predel

v možganskem deblu, kjer se nahajajo živčne celice, ki proizvajajo dopamin. To je živčni prenašalec, ki opravlja številne pomembne naloge in sicer vpliva na motorične sposobnosti, zaznavanje, spanje, razpoloženje, vedenje, učenje, motivacijo, zbranost, itd. Ocenjuje se, da se značilni simptomi Parkinsonove bolezni razvijajo, ko je uničenih že preko 70 % dopaminergičnih nevronov.

Bolezen prizadene telesno gibanje in bolnika postopoma onemogoči. Običajno se diagnosticira po prvih motoričnih simptomih. Glavni simptom Parkinsonove bolezni je bradikinezija, ki se kaže kot upočasnjena gibljivost in pomanjkanje spontanega gibanja. Temu je priključen eden ali več izmed naslednjih simptomov: tremor (tresenje v mirovanju, čeprav je za Parkinsonovo bolezen značilen tudi posturalni tremor), moteni posturalni refleksi (težave z ravnotežjem) in rigidnost (togost mišic). Obolele lahko prizadenejo še številni drugi simptomi, kot so težave s spanjem, izguba vonja, kognitivni upad, itd. Približno polovica bolnikov z napredovalim stanjem te bolezni ima demenco [3]. Parkinsonovi boleznijo so priključeni tudi atipični simptomi, zato je diagnozo velikokrat težko postaviti. Eden atipičnih simptomov Parkinsonove bolezni je ataksija, ki se pogosto kaže v neusklajenosti gibanja. Ataksija je značilna za multiplo sistemsko atrofijo, ki se poleg sindroma parkinsonizma in ataksije, kaže še z znaki okvare kortikospinalnih prog in motnjami avtonomnega živčevja [13].

Predvsem v kasnejši fazi, ko so simptomi izrazitejši, bolezen močno spremeni vsakdanje življenje obolelih. Poleg fizičnih težav, kot so težave s hojo, govorom in vonjem, vpliva tudi na duševno stanje bolnikov. Obolele spremljajo spremembe v razpoloženju, nespečnost in kognitivna upočasnjenost, pogosta je tudi pojavnost anksioznih motenj in depresije [14, 15]. Trenutno ne poznamo terapije, ki bi bolezen pozdravila oziroma ustavila ali upočasnila njen napredek. Poznamo pa ustrezne terapije, ki lajšajo simptome bolezni. V začetni fazi bolezni se simptome tipično lajša z zdravilom Levodopa, ki je prekursor dopamina in se uporablja za povišanje koncentracije dopamina v možganih bolnikov. Odziv bolnikov na zdravilo je tako v začetnem stadiju bolezni običajno dober. Z napredovanjem bolezni pa se sposobnost shra-

njevanja dopamina v možganih manjša, posledično je odziv na posamezne odmerke slabši. Zaradi tega se pacientu predpišejo višji in pogostejši odmerki, kar vpliva na povečano pojavnost neželenih učinkov, kot so zgibki (diskinezije) in ostali motorični zapleti. V kasnejšem stadiju bolezni se tako predpišejo drugačne medikamentozne terapije, kot so zaviralci monoaminooksidaze tipa B, zaviralci katehol-O-metiltransferaze, amantadin in zdravila, ki posnemajo učinke dopamina. Pri pacientih, ki jim zdravila za obvladovanje gibalnih težav ne zadoščajo, se lahko izvede kirurški postopek globoke možganske električne stimulacije [14], ki lahko za nekaj let omili simptome Parkinsonove bolezni [16]. Pacienti si lahko pomagajo tudi z nemedikamentoznimi metodami kot so zdrava prehrana, spominske in jezikovne vaje, redna telesna aktivnost in tehnike za obvladovanje stresa [17].

1.1.1 Test tapkanja

Za ocenjevanje stadija Parkinsonove bolezni se uporablja standardizirana lestvica MDS-UPDRS [1], ki sestoji iz štirih delov:

- 1) nemotorične izkušnje v vsakdanjem življenju,
- 2) motorične izkušnje v vsakdanjem življenju,
- 3) pregled motoričnih sposobnosti,
- 4) zapleti zdravljenja motoričnih simptomov.

Posamezen del se ovrednoti na lestvici od 0 do 4, pri čemer stopnja 0 pomeni, da pacient ne kaže znakov Parkinsonove bolezni. Stopnja 1 nakazuje na rahle težave, stopnja 2 blage težave, stopnja 3 srednje težave, stopnja 4 pa hude težave pacienta.

Sklop za ocenjevanje motoričnih sposobnosti pacienta vključuje tudi test tapkanja s prsti [1] za ocenjevanje stopnje bradikinezije. Kriteriji ocenjevanja so prikazani v tabeli 1.1. Test se izvaja tako, da se testiranec s kazalcem dotakne palca, nato pa kazalec odmakne čim dlje od palca in se vrne v

prvotno lego. Takšen gib predstavlja cikel in se izvede od 10 do 15-krat, pri čemer mora testiranec gibe izvajati čim hitreje [18]. Če med izvajanjem tapkanja pride do motenj gibanja, le te vplivajo na končno oceno. Po lestvici MDS-UPDRS se med motnje gibanja štejejo zamrznitve oziroma prekinitve ritma tapkanja, upad amplitude in hitrosti tapkanja ter čas ob katerem pride do upada slednjih. Problem testa tapkanja predstavlja objektivnost, saj je ovrednotenje gibanja odvisno od zdravnika, ki test opazuje. Slednji mora v času izvajanja testa prepoznati vse motnje gibanja, namreč v nasprotnem primeru lahko zaradi nepozornosti poda napačno oceno. Ponovljen test ni nujno reprezentativen, saj se pacient utruje, kar prispeva k nastanku motenj gibanja, ki niso posledica bolezni, temveč le utrujenosti. Ocenjuje se, da je do 25 % kliničnih diagnoz Parkinsonove bolezni napačnih [19].

MDS-UPDRS 3.4 tapkanje s prsti	
Ocena	Navodila za točkovanje
0	- test opravi brez problemov
1	- 1–2 prekinitvi ali oklevanji ritma tapkanja - rahla upočasnitev hitrosti tapkanja - upad amplitude proti koncu sekvence 10 oziroma 15 ciklov
2	- 3–5 prekinitev ali oklevanj med tapkanjem - blaga upočasnitev hitrosti tapkanja - upad amplitude približno od sredine sekvence 10 oziroma 15 ciklov
3	- več kot 5 prekinitev med tapkanjem ali vsaj ena daljša zamrznitev gibov - zmerna upočasnitev hitrosti tapkanja - upad amplitud po prvem ciklu
4	- naloge ni zmožen opraviti ali pa jo lahko opravi z veliko težavami zaradi upočasnitev, prekinitev ali zmanjšanja amplitude in hitrosti

Tabela 1.1: Merila za točkovanje po lestvici MDS-UPDRS 3.4 ocena tapkanja s prsti. Ocenjuje se na podlagi zaporedja najmanj 10 ciklov odpiranja in zapiranja. Glavni parametri ocenjevanja so prekinitve oziroma zamrznitve ter upad amplitud in hitrosti. Tabela je povzeta po ocenjevalni lestvici Stamatakisa in sodelavcev [20].

1.2 Sorodna dela

Avtomatizirani sistemi za prepoznavo Parkinsonove bolezni so računalniški sistemi, ki uporabljajo različne tehnologije, kot so umetna inteligenca, strojno učenje, poglobljena analiza podatkov in senzorske tehnologije, za zgodnje odkrivanje in spremljanje Parkinsonove bolezni. Pri prvih avtomatiziranih sistemih za prepoznavo Parkinsonove bolezni so se uporabljali komercialni žiroskopi, pospeškovni senzorji [21, 22, 23, 24] in elektromiografija [25, 26].

Pridobljene podatke so analizirali s spektralno analizo [23], adaptivnimi Fourierjevimi modeli [24] in modeli strojnega učenja [26, 27, 28]. Omenjeno spektralno analizo omejujejo kratki časovni intervali, npr. 10-sekundni posnetek zajema le 10 do 15 ponovitev giba, kar za natančno analizo dinamike tapkanja ne zadostuje [29]. Uporabo senzorjev omejuje tudi komercialna nedostopnost in potreba po fizični nameščenosti na roko, kar lahko vpliva na test tapkanja s prsti.

Hyoseon Jeon in sodelavci [26] so signale tapkanja pridobili z uporabo nosljive naprave nameščene na kazalcu, ki jo sestavljata merilnik pospeška in žiroskop. V študiji so želeli določiti stopnjo tremorja pri Parkinsonovih bolnikih. Iz pospeškov in kotnih hitrosti so izračunali amplitude in kote. Iz vsake izmed 4 časovnih vrst so pridobili 19 značilk. Zaradi velikega števila značilk so uporabili metodo parne korelacije in analizo glavnih komponent (PCA *angl. Principal Component Analysis*). Tako so odstranili odvečne značilke in zmanjšali prostor značilk. Značilke so nato za učenje modela uporabili kot vhodne podatke v odločitvenih drevesih (*angl. decision trees*), podpornih vektorjih (SVM, *angl. Support Vector Machine*), diskriminantni analizi, naključnih gozdovih (*angl. Random Forest*) in k-najbližjih sosedih (kNN *ang. k-nearest neighbors*). Na podatkovni zbirki 85 signalov so z uporabo odločitvenih dreves dosegli največ 85 % točnost. Sadikov in sodelavci [30] so razvili model za zaznavanje bradikinezije in diskinezije pri Parkinsonovih bolnikih. V ta namen so uporabili digitalno spirografijo v kombinaciji s strojnim učenjem. Podatke so zbirali z napravo za telemetrijo z zaslonom na dotik [31], na katero so pacienti s pisalom v dominantni roki obrisali Arhimedovo spiralo. Značilke so nato pridobili iz primerjave napak med narisano in idealno Arhimedovo spiralo ter iz hitrosti risanja. Naučili so 2 linearna in 2 nelinearna klasifikatorja, a so se vsi odrezali podobno in dosegli maksimalno 86 % točnost. Avtorji so za potrditev natančnosti rezultata uporabili 10-krat ponovljeno 10-kratno navzkrižno preverjanje (*angl. 10-fold Cross-Validation*) s stratificiranim vzorčenjem.

Za namen prepoznave Parkinsonove bolezni so se raziskovalci poslužili

tudi zajema podatkov s tipkanjem na tipkovnico [32, 33]. Le to so zajeli z namenskim programom Tappy. Značilke so pridobili na podlagi časa pritiska tipke, časa med pritiskom in spustom tipke ter pritiskom naslednje tipke. Pri klasifikaciji v razred zdravih oziroma bolnih so dosegli med 92 % in 100 % (v določenih primerih) točnost oziroma priklic. Iakovakis in sodelavci [34, 35] so izvedli podobno raziskavo, vendar so tipkanje na tipkovnico zamenjali s tapkanjem po pametnem telefonu.

Zaradi napredka strojne opreme in razvoja programskih rešitev je področje računalniškega vida v kombinaciji s strojnim učenjem ponudilo dobro alternativo za prepoznavo Parkinsonove bolezni. Takšni sistemi omogočajo testiranje od doma z računalnikom oziroma pametnim telefonom, ki sta del skoraj vsakega gospodinjstva. Gre tudi za cenovno ugodno rešitev, saj ni potrebe po dodatni opremi. Lainscsek in sodelavci [29] so videe tapkanja analizirali z nelinearnimi dinamičnimi modeli, ki temeljijo na nelinearnih diferencialnih enačbah z zakasnitvijo. Predstavljen model se lahko uporabi za opredelitev in oceno ne le stopnje disfunkcije gibanja bolnikov s Parkinsonovo boleznijo, temveč tudi za opredelitev strukture gibanja pri različnih motoričnih motnjah, kot so disfunkcije gibanja zaradi možganske kapi ali razvojne motorične disfunkcije v otroštvu. Predstavljen model, ki temelji na nelinearnih diferencialnih enačbah z zakasnitvijo, rešuje problem spektralne analize. Namreč po lestvici MDS-UPDRS je za oceno potrebnih 10 do 15 ciklov, kar je premalo za natančno spektralno analizo osnovne dinamike. Struktura nelinearnih diferencialnih enačb z zakasnitvijo, ki so jo prilagodili dani časovni vrsti gibanja, je bila izbrana s pomočjo genetskih algoritmov. Iz koeficientov nelinearne diferencialne enačbe so nato sestavili deskriptor gibanja za oceno stopnje motorične prizadetosti. V povprečju so dosegli 76 % specifičnost in priklic klasifikacije v razred zdravih in obolelih testirancev. Obe skupini sta skupaj vsebovali le 13 videov tapkanja, rezultate klasifikacije pa so pridobili s k-kratnim navzkrižnim preverjanjem.

Eden prvih primerov vrednotenja stopnje motorične prizadetosti pri Parkinsonovi bolezni iz videov tapkanja z uporabo strojnega učenja so pred-

stavili Khan in sodelavci [36]. Poziciji roke na videih tapkanja so sledili s pomočjo segmentacije in gradientov. Normalizacijo podatkov so izvedli glede na višino obraza, da so izločili vpliv razdalje med kamero in roko. Značilke so pridobili iz statističnih metrik križne korelacije normaliziranih ekstremov in slabše značilke z uporabo hi kvadrat testa odstranili. Od 13 pacientov s Parkinsonovo boleznijo so pridobili 387 posnetkov in od 6 zdravih udeležencev dodatnih 84 kontrolnih posnetkov. Torej so od le 19 oseb pridobili kar 471 videov kar pomeni, da so imeli na voljo precej homogeno bazo podatkov. Dva nevrologa sta nato neodvisno ocenila zbrane videe tapkanja v razrede med 0 in 3 po lestvici UPDRS. S SVM in 10-kratnim navzkrižnim preverjanjem so dosegli 88 % točnost klasifikacije v štiri razrede po lestvici UPDRS oziroma 95 % točnost pri napovedi stopnje bradikinezije.

Williams in sodelavci [37] so vzpostavili model za zaznavo bradikinezije pri testu tapkanja. V ta namen so zbrali 70 video posnetkov tapkanja. Za sledenje skeleta roke so najprej naredili segmentacijo z modelom konvolucijske nevronske mreže (CNN, *angl. convolutional neural network*). Nato so s pomočjo polja optičnega toka določili magnitudo gibanja, ki predstavlja enodimenzionalen časovni signal. Iz tega so pridobili značilke in s pomočjo metode PCA zmanjšali prostor le teh. Značilke so uporabili za treniranje SVM, naivnega Bayesovega klasifikatorja in binarne logistične regresije. Dosegli so 80 % točnost in 86 % priklic pri napovedovanju prisotnosti bradikinezije, ki je eden glavnih pokazateljev Parkinsonove bolezni. Pri napovedovanju stopnje bradikinezije v Parkinsonovi bolezni so razreda 0 in 1 ter razrede 2, 3 in 4 po lestvici MDS-UPDRS združili in tako dosegli 63 % točnost ter 68 % priklic.

Zupančič in sodelavci [18] so razvili spletno aplikacijo za analizo tapkanja s prsti. Gibanju roke so sledili s pomočjo prileganja skeleta z Mediapipe Hands [38], ki omogoča zaznavo roke v realnem času. Kot rezultat vrne 21 točk skeleta roke, kakor je vidno na sliki 2.1. V programu so za pridobivanje podatkov uporabili točki konice palca in konice kazalca. Iz podatkov so nato izračunali hitrosti in razdalje v odvisnosti od časa, ter ustvarili grafe hitrosti in pospeška.

Li in sodelavci [39] so predstavili avtomatsko metodo vrednotenja, primerno za klasifikacijo motorične prizadetosti zaradi Parkinsonove bolezni pri majhnem številu posnetkov tapkanja. Metoda združuje časovno vrsto podatkov z nevronske mreže. Testirali so tudi klasifikacijski model za prepoznavo akcij imenovan 3DCNN (*angl. action recognition model*). Slednji model, ki kot vhod prejme kar neobdelane videe, se ni dobro odrezal. Razlog slabega delovanja modela tiči v majhnem številu podatkov in podobnosti med razredi ocen po lestvici MDS-UPDRS. Zato so model opustili. Iz trajektorij palca in kazalca so z Mediapipe Hand pridobili amplitude in hitrosti ter jih kot časovno vrsto uporabili v kombinaciji s predstavljeno polno konvolucijsko nevronske mrežo (FCN *angl. fully convolutional neural network*). Zaradi majhne množice podatkov so uporabili 5-kratno navzkrižno preverjanje in tako dosegli višjo zanesljivost rezultatov. V povprečju so dosegli 79,7 % točnost klasifikacije po lestvici MDS-UPDRS na njihovi lastni testni množici podatkov, sestavljeni iz 252 posnetkov, pridobljenih od 120 oseb. Kombinacija značilk časovne vrste in nevronske mreže se je v času izdaje članka, januarja 2022, izkazala za boljšo kot uporaba nevronskih mrež (npr. Motif-GCNs, 2s-AGCN in Three-Stream). Ta so na podatkih skeleta roke v najboljšem primeru dosegla 72,4 % točnost. Uporaba časovnih podatkov v Resnet-u se je prav tako izkazala za slabšo in dosegla 77,4 % točnost.

Adonay in sodelavci [40] so predstavili model za ločevanje oseb z ataksijo, Parkinsonovo boleznijo in zdravimi ljudmi. Podatkovna zbirka videov tapkanja je obsegala 169 primerov z ataksijo, 78 s Parkinsonovo boleznijo in 58 zdravih primerov. Pozicije roke so pridobili z nevronske mreže ResNet152. V primeru, da roke na sličici niso zaznali, so za dopolnitev uporabili linearno interpolacijo v kolikor roka ni bila zaznana v manj kot 0,5 sekunde. Časovno vrsto razdalj med konicama kazalca in palca so zaradi različnih razdalj med kamero in roko normalizirali. Da so odstranili artefakte nastale pri določanju pozicije palca in kazalca, so signale filtrirali z nizkopasovnim filtrom (*angl. low pass filter*). Nato so iz časovne vrste razdalj, lokalnih maksimumov, lo-

kalnih minimumov in iz njihovih odvodov izračunali 256 statističnih značilnk, ki so jih z metodo PCA zreducirali na 10 značilnk. Slednje so uporabili v kombinaciji z logistično regresijo in izdelali modele za razlikovanje med Parkinsonovo boleznijo, zdravo skupino in ataksijo.

Jakubowski in sodelavci [41] so sestavili binarni klasifikacijski model za razlikovanje med zdravimi in obolelimi s Parkinsonovo boleznijo. Pri učenju klasifikacijskega modela so uporabili le 4 značilke, ki so bile izbrane iz domenskega znanja o motoričnih težavah obolelih. Dve so izbrali iz domene amplitud, drugi 2 pa iz domene frekvenc, nato pa za zmanjšanje prostora značilnk uporabili latentno semantično analizo. Z uporabo SVM in k-NN (*angl. k-nearest neighbors*) so na bazi 42 posnetkov dosegli približno 81 % priklic in 0,78 F1 vrednost. Rezultate so zaradi zagotavljanja večje zanesljivosti pridobili s pomočjo metode k-kratnega navzkrižnega preverjanja.

Islam in sodelavci [42] so izdelali model za klasifikacijo v enega izmed 5 razredov lestvice MDS-UPDRS. Na bazi 489 posnetkov pridobljenih od 250 različnih oseb so z modelom strojnega učenja dosegli povprečno absolutno napako (MAE, *angl. mean absolute error*) 0,58 in 52 % točnost klasifikacije. Ocene videov tapkanja so navzkrižno preverili, saj so vidoe neodvisno ocenili trije izkušeni nevrologi. V predstavljenem modelu so zraven trajektorije gibanja palca in kazalca pridobljene z MediaPipe Hand upoštevali še gibanje zapestja in tako pridobili dodatne značilke. Slabše značilke so potem s pomočjo metode BoostRFE (*angl. boosted recursive feature elimination*) izločili. Tako so za učenje modela uporabili 22 od 65 prvotnih značilnk. Uporabili so regresijske modele (npr. support vector regression, LightGBM, XGBoost, AdaBoost) in plitve nevronske mreže z enim oziroma dvema učljivima nivojema.

Tianze Yu in sodelavci [43] so problem klasifikacije stopnje bradikinezije po lestvici MDS-UPDRS iz video tapkanja reševali s pomočjo izdelave domenskih značilnk v kombinaciji s strojnim učenjem. Glavna razlika v primerjavi z ostalimi študijami je v tem, da so izbrali specifične značilke strogo

v skladu z uveljavljenimi smernicami MDS-UPDRS. Kot klasifikacijski model so izbrali odločitvena drevesa, saj so ta lahka za interpretacijo, kar je v skladu z njihovo idejo pristopa, ki tesno sovпада s kliničnim razmišljanjem nevrologa. Njihova podatkovna zbirka je štela 75 videov, pridobljenih od 50 udeležencev. Od tega je bilo 20 udeležencev posnetih v laboratoriju, 30 pa v kliničnem okolju. V laboratoriju so za posnetke tapkanja uporabljali 8 kamer, pod različnimi koti. Razdalja med kamero in roko je bila fiksna. Snemanje je potekalo v ločljivosti 3840 x 2160 slikovnih pik s hitrostjo 15 sličic na sekundo. Izmed vseh 8 posnetkov so izbrali tistega, ki je bil najbolj jasen in kjer je bil kot med kamero in roko blizu pravega kota. Video posnetki, posneti v kliničnem okolju, so bili pridobljeni z uporabo fiksirane kamere in aplikacije, ki je udeležence usmerjala, po potrebi pa jim je pomagalo tudi osebje. Tukaj so bili posnetki zajeti v ločljivosti 1920 x 1080 slikovnih pik s hitrostjo 60 sličic na sekundo. Videe, kjer tapkanje ni bilo jasno vidno ali snemano pod pravim kotom so izločili. Da so izboljšali podatke so videe procesirali. Najprej so videom povečali resolucijo, nato so jim odstranili šum in zmanjšali zamegljenost zaradi gibanja (*angl. motion blur*), na koncu so izvedli še interpolacijo sličic (*angl. frame interpolation*). Njihove videe sta označila 2 označevalca, da so se izognili človeškim napakam in garantirali večjo natančnost oznak. Za pridobitev amplitud palca in kazalca so kot večina raziskovalcev uporabili MediaPipe. Pridobljene amplitude so normalizirali glede na velikost zapestja, nato so jih zgladili ter odstranili šum. Njihov vektor značilk je sestavljalo 15 specifičnih značilk. Od tega sta bili 2 demografski, to sta starost in spol. Pet značilk je bilo sestavljenih iz amplitud, 5 iz hitrosti, tem so dodali še zaustavitve in zamrznitve ter 1 značilko iz frekvenčnega spektra. Odločitvena drevesa so trenirali na 60 posnetkih z uporabo vseh 15 značilk in na testni množici 15 posnetkov dosegli 80 % točnost in 79 % priklic.

Alam in sodelavci [44] so za klasifikacijo stopnje bradikinezije na podlagi testa tapkanja s prsti razvili grafovsko konvolucijsko nevronske mrežo imenovano PULSAR. Zbrali so bazo 382 video posnetkov tapkanja, prido-

bljenih od različnih ljudi posnetih preko računalniške kamere. Od tega so jih 200 uporabili za učno ter testno množico. Za detekcijo roke so uporabili MediaPipe Hand in sličice, na katerih slednji ni zaznal roke, odstranili. Ker nevronske mreže tipično potrebujejo večje število podatkov za dobro delovanje, so izvedli bogatenje podatkov z uporabo horizontalnega, vertikalnega in horizontalno-vertikalnega zrcaljenja. Kot vhod so nevronske mreže podali vse točke skeleta roke MediaPipe, vidnega na sliki 2.1 in povezave med njimi. Nevronska mreža je v najboljšem primeru dosegla 81 % točnost in 0,81 F1 na testni množici ter 71 % točnost in 0,71 F1 na za sistem še nevideni skriti testni množici 182 posnetkov.

1.3 Motivacija

Parkinsonova bolezen je za paciente zelo naporna. Da jim olajšamo vsakodnevno življenje je potrebno zdravljenje. Čakalne vrste za nevrologa so zelo dolge, kakovost življenja pacientov v tem času drastično upada. Z uporabo avtomatskih metod analize bi čakalne vrste skrajšali. Avtomatska diagnoza stopnje motorične prizadetosti na podlagi testa tapkanja s prsti zaradi Parkinsonove bolezni je sicer že mogoča z uporabo senzorjev, vendar takšni sistemi niso dostopni vsem. Obstoječe rešitve na osnovi računalniškega vida v večini uporabljajo vidne pridobljene v nadzorovanem okolju pod idealnimi pogoji. Zato nas je zanimalo ali lahko oziroma kako natančno lahko določimo stopnjo motorične prizadetosti oziroma bradikinezije Parkinsonove bolezni po lestvici MDS-UPDRS iz video posnetkov tapkanja s prsti posnetih v ambulantni brez posebne opreme. Ker baze videov tapkanja niso prosto dostopne, smo vidne posneli sami. Tako so bili videi snemani v nenadzorovanem okolju, s kamero pametnega telefona pri različnih resolucijah, osvetlitvah, sličicah na sekundo, razdaljah med kamero in roko ter koti med kamero in roko. Pridobljena podatkovna zbirka je zelo heterogena, kar zahteva izdelavo robustnega sistema. V redkih raziskavah, kjer uporabljajo podobno posnete vidne, so bili rezultati zmogljivosti relativno nizki. Zato smo se odločili za gradnjo lastnega

klasifikatorja. Zaradi lažjega razumevanja modela smo se odločili poseči po gradnji značilke, ki so se ravnale po lestvici MDS-UPDRS. Dobljene rezultate smo tudi primerjali z rezultati nevronske mreže, ki avtomatsko pridobijo rezultate ter posledično ne omogočajo vpogleda v odločitev klasifikatorja. Ker gre za področje medicine, kjer avtomatski sistemi potrebujejo potrditev s strani zdravnika smo v pomoč potrditve razvili še orodje za vpogled v tapkanje, ki slednje vizualizira z grafi ključnih podatkov. Na ta način si prizadevamo zmanjšati subjektivnost pri postavitvi diagnoze in omogočiti hitrejšo diagnozo bolezni, s tem pa prispevati k izboljšanju kakovosti življenja pacienta.

1.4 Prispevki

V sklopu magistrskega dela smo vzpostavili model, ki na podlagi video posnetkov zajetih z računalnikom ali s telefonom zazna motnje v tapkanju in diagnosticira stopnjo motorične prizadetosti na podlagi testa tapkanja s prsti po lestvici MDRS-UPDRS. Izdelali smo tudi orodje za prikaz video posnetkov, ki omogoča predvajanje video posnetka, drsenje po posameznih sličicah videa in izris amplitud ter grafov hitrosti tapkanja. Na grafih so označeni cikli, zamrznitve in trenutna pozicija glede na predvajan video. Razvito orodje tako nudi podrobnejši vpogled v posneti test tapkanja in s tem omogoča ročno preverjanje ocene motorične prizadetosti na podlagi testa tapkanja s prsti po MDS-UPDRS. Za razvoj modela smo zajeli in obdelali zbirko 183 video posnetkov tapkanja s pripadajočimi ocenami nevrologa Dejana Georgieva po lestvici MDS-UPDRS.

1.5 Struktura naloge

Magistrska naloga je razdeljena na 5 poglavij. V uvodnem poglavju je predstavljena Parkinsonova bolezen ter pregled dosedanjih pristopov za prepoznavo Parkinsonove bolezni s pomočjo računalnika. V drugem poglavju so

opisane najpomembnejše metode uporabljene pri izdelavi magistrskega dela. V tretjem poglavju so opisani pristopi izgradnje klasifikatorja za klasifikacijo tapkanja iz videov tapkanja po lestvici MDS-UPDRS. Poglavje vključuje tudi opis orodja za vizualizacijo tapkanja. Rezultati klasifikacije preko različnih pristopov so predstavljeni v četrtem poglavju. V zadnjem poglavju so zapisane ugotovitve in problemi s katerimi smo se srečevali pri izgradnji klasifikatorja. Tukaj smo opisali tudi potencialne izboljšave in alternativne pristope za izgradnjo klasifikacijskega sistema.

Poglavje 2

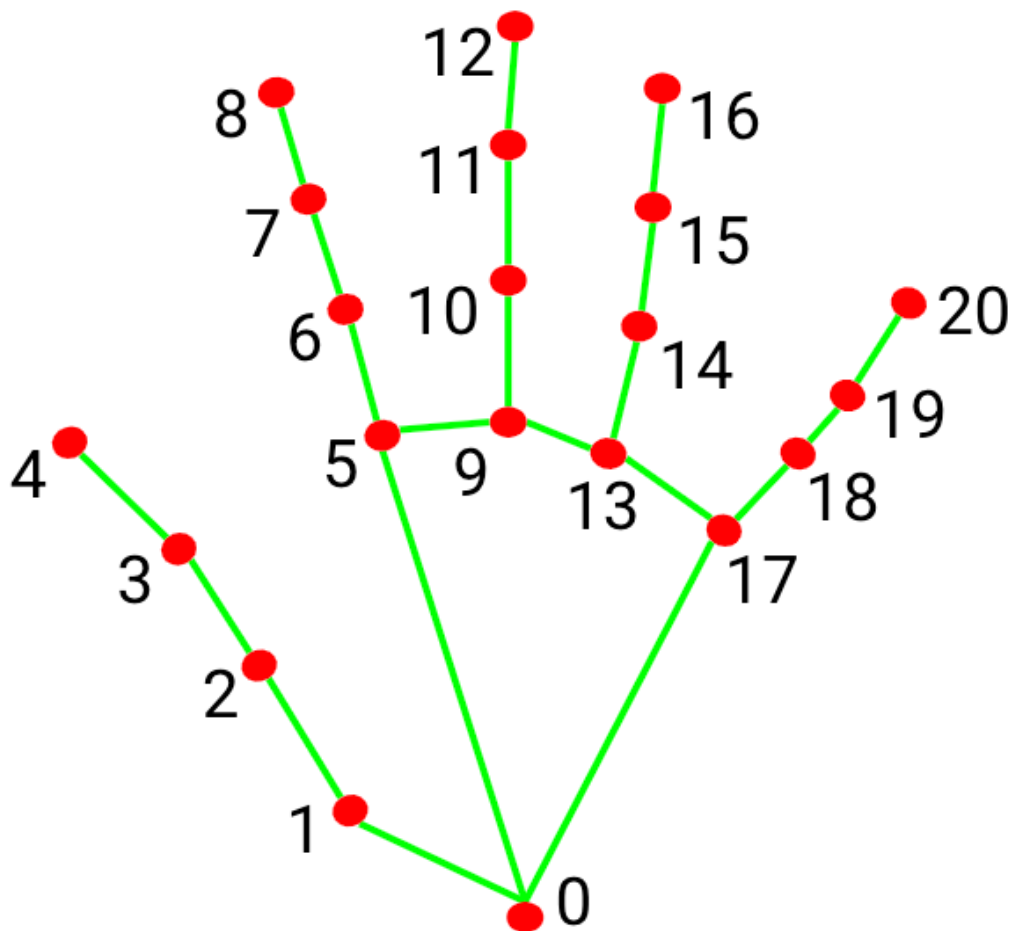
Tehnične metode

V poglavju so opisane ključne metode, ki smo jih potrebovali pri pridobivanju in obdelavi podatkov ter za izdelavo značilk in klasifikatorja stopnje bradikinezije po MDS-UPDRS. Nekateri algoritmi smo razvili sami, določene obstoječe pa smo prilagodili specifičnim zahtevam našega projekta. Ostale smo uporabili neposredno iz razpoložljivih knjižnic za Python.

2.1 Zaznava roke in skeleta

Za zaznavanje roke in prilagajanje skeleta na njo smo uporabili MediaPipe Hand. MediaPipe [45] je odprto kodno ogrodje razvito s strani Googla. Zasnovano je za področje računalniškega vida, kjer ponuja različne rešitve nalog s pomočjo strojnega učenja, kot so detekcija obraza, prepoznavanje objektov, sledenje roke in še veliko več. MediaPipe je modularne zasnove kar omogoča uporabo le željenih rešitev, v našem primeru le detekcije roke. MediaPipe Hand omogoča tudi delovanje v realnem času in je preprost za integracijo. Omogoča nam detekcijo ene ali več rok ter nam kot rezultat vrne pozicije ključnih točk roke oziroma skelet roke. Za vsako roko nam vrne skelet iz 21 ključnih točk, ki so vidne na sliki 2.1. Za potrebe naloge smo uporabili končne točke palca in kazalca ter zapestje, torej točke 0, 4 in 8 prikazane na sliki 2.1. V določenih primerih smo za izgradnjo značilk dodatno uporabili tudi

točki 3 in 7 ter samostojno tudi točko 0. Točke so vidne na sliki 2.1. Ostalih točk za izgradnjo značilk nismo uporabljali, saj so pozicije med tapkanjem precej statične. Vsaka točka T je določena s koordinatami (x, y, z) , pri čemer x predstavlja horizontalno pozicijo, y vertikalno pozicijo, in z globinsko.



Slika 2.1: Prikaz skeleta MediaPipe Hand. Točke 1 do 8 predstavljajo skelet palca in kazalca, točka 0 pa pozicijo zapestja. Slika je vzeta iz [46].

2.2 Filter s končnim impulznim odzivom

Filter s končnim impulznim odzivom (FIR filter, *angl. Finite-Impulse Response filter*) ima končni impulzni odziv. Velja za stabilen filter in lahko

ohrani linearno fazo. Za razliko od filtrov z neskončnim odzivom potrebujejo filtri s končnim odzivom več parametrov. Njihova uporabnost je obsežna, primeri uporabe so odstranjevanje šuma, izenačevanje zvoka, odstranjevanje prekrivajočih se frekvenc (npr. frekvence nad Nyquistovo frekvenco), itd. Primeri FIR filtrov so drseče povprečje, Savitzky-Golay filter, Gaussovo glajenje, itd. Filter drsečega povprečja se uporablja za zmanjšanje šuma oziroma glajenje podatkov. Filter je definiran z enačbo 2.1,

$$y[n] = \frac{1}{N} \sum_{k=0}^{N-1} x[n-k], \quad (2.1)$$

kjer je $y[n]$ izhodni filtriran signal, n je časovni indeks, perioda N je dolžina okna filtra (torej število vzorcev za izračun povprečja filtra), $x[n-k]$ pa predstavlja vhodni signal ob času $n-k$.

Njegove prednosti so, da je enostaven za razumevanje in implementacijo. Fazni zamik, ki ga vpelje s filtriranjem je odvisen od dolžine okna (periode N). Njegova slabost je, da lahko zamaskira hitre spremembe v signalu, saj jih zgladi. Prav tako mu problem predstavljajo robni primeri (začetek in konec signala), kjer ni dovolj podatkov za izračun povprečja.

2.3 Filter z neskončnim impulznim odzivom

Filter z neskončnim impulznim odzivom (IIR filter, *angl. Infinite Impulse Response filter*) je tip digitalnega filtra, ki se uporablja za spreminjanje ali izboljšanje značilnosti signala. Uporablja se pri obdelavi slik, videov, avdio podatkov in raznih signalov merilnih naprav. Z IIR filtri tako lahko odstranjujemo šum, dodamo ali odstranimo zamegljenost, izenačujemo zvok, dodamo zvočne učinke in izluščimo značilke (IIR filtri se predvsem uporabljajo v nevronske mrežah [47]). Poznamo nizkoprepustne, visokoprepustne, pasovno prepustne in pasovno neprepustne filtre z neskončnim impulznim odzivom. Filtri s končnim impulznim odzivom so od njih stabilnejši, vendar imajo večjo zakasnitev ter so procesno zahtevnejši. IIR filtri imajo neskončni

impulzni odziv zato njihov izhod ni odvisen le od trenutne vrednosti vhodnega signala ampak tudi od prejšnjih vrednosti. Posledično jim rečemo rekurzivni filtri. Primeri IIR filtrov so filtri Bessel, Butterworth [48], Chebyshev, itd.

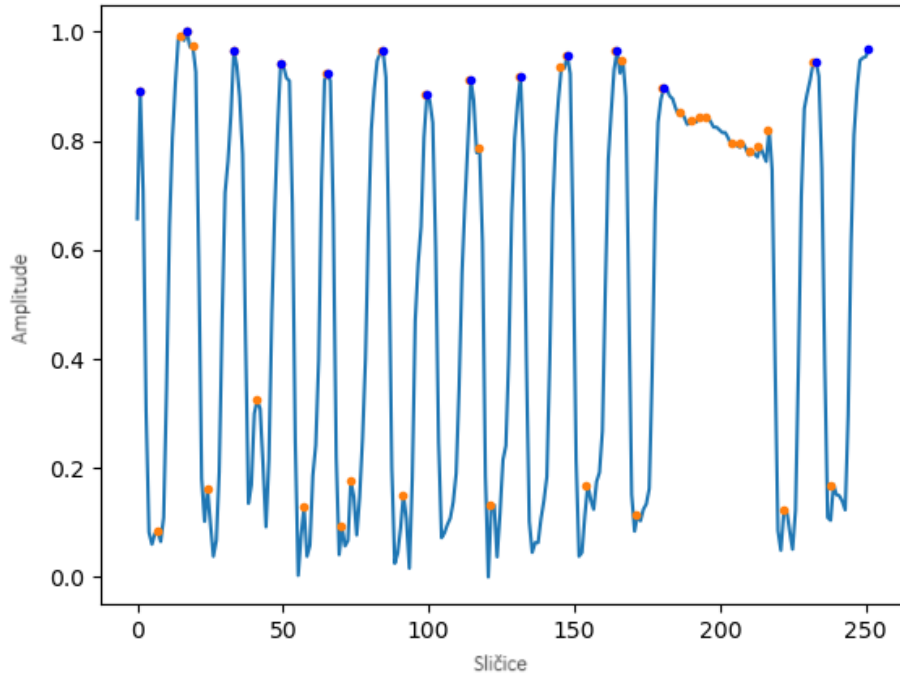
Nizko prepustni filter je posebna vrsta IIR filtra, zasnovan tako, da prepušča signale z nižjo frekvenco od frekvence rezanja, višje frekvence signala pa zaduši. Takšno filtriranje uporabljamo, da odstranimo oziroma zmanjšamo visokofrekvenčni šum, ki se v naših podatkih pojavi zaradi prileganja skeleta in pri nekaterih osebkih prisotnega tremorja. V naši nalogi smo uporabili nizkoprepusten IIR filter s frekvenco odreza določeno za vsak vhodni signal posebej. Implementirali smo ga kot nizkoprepustni Butterworth-ov filter [48]. Slednji je podan z enačbo 2.2

$$|H(j\omega)| = \frac{1}{\sqrt{1 + (\frac{\omega}{\omega_c})^{2N}}}, \quad (2.2)$$

kjer je $H(j\omega)$ magnituda frekvenčnega odziva filtra, ω je kotna frekvenca, ω_c je frekvenca odreza filtra, N pa stopnja filtra [49].

2.4 Detekcija tapkanja

Za zaznavo posameznih ciklov tapkanja je bilo potrebno zaznati ekstreme amplitud. V signalu je veliko lokalnih ekstremov, določeni se lahko pojavijo zaradi tremorja, težav pri tapkanju, ali šuma v podatkih (npr. zaradi prileganja skeleta). Posledično vsak lokalni ekstrem ne označuje enega cikla. Zato je potrebno zaznati odvečne ekstreme in jih ustrezno izločiti. Izmed lokalnih ekstremov se v okolici teh glede na definirano razdaljo poišče največji ekstrem, ostali pa se izločijo. Torej algoritem postopoma odstranjuje manj pomembne lokalne ekstreme na podlagi vrednosti teh in razdalje do sosednjih, dokler ne ostanejo le najbolj vztrajni, vidni na sliki 2.2 označeni z modro piko. Vztrajnostni prag (*angl. persistence threshold*) nadzoruje občutljivost postopka filtriranja in je bil določen kot 10 % razlike med maksimalno in minimalno vrednostjo amplitude vhodnega signala.



Slika 2.2: Modro označeni lokalni maksimumi označujejo cikle tapkanja. Oranžne lokalne maksimume algoritem odstrani.

2.5 Sledenje zamrznitvam in prekinitvam

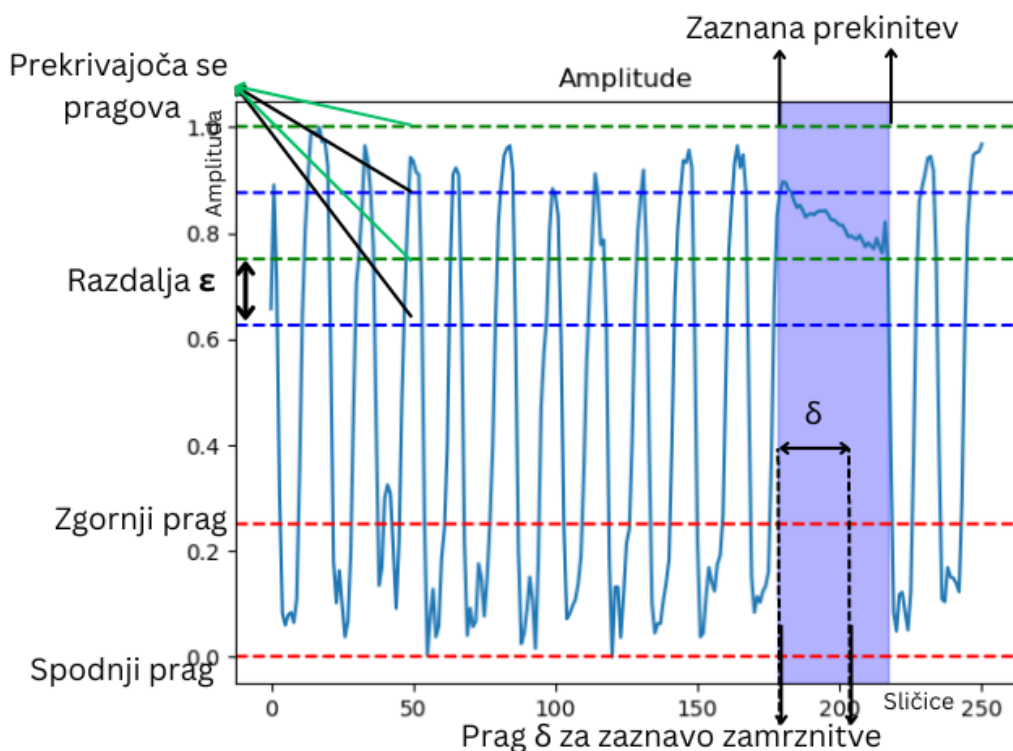
Za potrebe odkrivanja zamrznitev, oklevanj in zaustavitev smo razvili algoritem, ki na podlagi amplitud in fiksnih mej določi le te. Algoritem deluje tako, da sledi intervalom, kjer so spremembe amplitude dovolj majhne, torej med mejama $[T_1, T_2]$. V kolikor so spremembe med mejama za nek dovolj velik čas δ , algoritem označi zamrznitev oziroma zaustavitev. T_1 in T_2 predstavljata spodnjo in zgornjo mejo, ki je bila empirično določena na 25 % razlike maksimalne in minimalne vrednosti amplitud.

Algoritem še sprejme parameter n , ki pomeni število iteracij oziroma natančnost sledenja. Namreč na začetku nastavimo praga na $T_1 = \min(A)$, $T_2 = 0,25 \cdot (\max(A) - \min(A))$, kjer $\min(A)$ in $\max(A)$ pomenita minimalno in maksimalno vrednost signala amplitud. Začetna praga sta vidna na sliki

2.3, označena z rdečo črto. Nato z vsako iteracijo povečujemo oba praga za majhno vrednost ϵ (vidno na sliki 2.3), ki je definirana z enačbo 2.3.

$$\epsilon = i \cdot \left(\frac{(\max(A) - \min(A)) \cdot (1 - 0,25)}{n} \right), \quad i = 1, 2, 3, \dots, n \quad (2.3)$$

Algoritem tako za vsak korak i pridobi ustrezne intervale zamrznitev. Tako imamo več prekrivajočih se intervalov, ki jih moramo izločiti. Izločanje poteka tako, da obdržimo najdaljši interval izmed prekrivajočih se intervalov vsakega koraka i . Na sliki 2.3 lahko vidimo prekrivajoča se pragova, označena z zeleno in modro barvo, ki nam vrneto prekrivajoč interval različnih dolžin. Interval zamrznitve med modrima pragoma M zavržemo, saj je krajši.



Slika 2.3: Prikaz delovanja algoritma za zaznavo zamrznitev. S horizontalno rdečo črto sta označena začetna pragova. Z modro horizontalno črto je označen spodnji in zgornji prag M , z zeleno pa spodnji in zgornji prag Z .

Alternativa zaznave zamrznitev, prekinitev in zaustavitev bi bila uporaba hitrosti. Tukaj bi sledili intervalom, kjer je hitrost dovolj blizu 0 za določen

čas δ . Slednja možnost pa je manj robustna zaradi šuma v podatkih. Ker MediaPipe Hand pri prileganju skeleta rahlo spreminja pozicije med sličicami, čeprav roka lahko miruje, nam vpelje visokofrekvenčni šum. Ta se odrazi v višjih hitrostih, sprememba amplitud pa je dejansko majhna. Tako slednji algoritem ne bi natančno zaznal zaustavitve oziroma zamrznitve. Hitrosti so tudi same po sebi manj natančne zaradi snemanja pri relativno malo sličicah na sekundo.

2.6 Avtokorelacija

Avtokorelacija je matematična metoda za merjenje podobnosti signala z njegovo časovno zamaknjeno kopijo, torej meri linearno odvisnost signala. Omogoča nam odkrivanje ponavljajočih se vzorcev znotraj signala. Avtokorelacija je za zvezne signale definirana po enačbi 2.4,

$$R(\tau) = \int_{-\infty}^{\infty} x(t) \cdot x(t + \tau) dt, \quad (2.4)$$

za diskretne signale pa po enačbi 2.5,

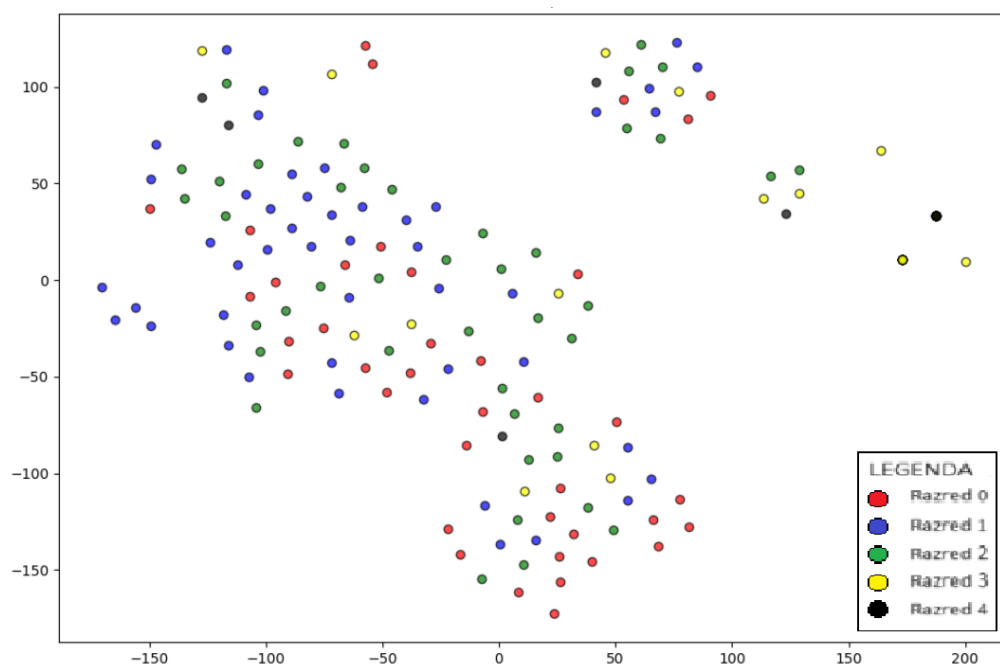
$$R(\tau) = \sum_{n=-\infty}^{\infty} x[n] \cdot x[n + \tau]. \quad (2.5)$$

Ker je dolžina izhodnega signala avtokorelacije odvisna od dolžine vhodnega signala smo iz avtokoreliranih signalov izluščili časovno neodvisne značilke.

2.7 Vizualizacija značilk

Metoda t-SNE (*angl. t-Distributed Stochastic Neighbor Embedding*) je statistična nelinearna metoda za zmanjševanje prostora podatkov, ki so jo predstavili Maaten in sodelavci [50]. Predvsem se uporablja za vizualizacijo visoko dimenzionalnih prostorov v 2D ali 3D prostoru. Za PCA, ki je linearna metoda za zmanjševanje prostora, je značilno, da v največji meri ohrani varianco podatkov, zato se osredotoča na ohranitev globalne strukture podatkov.

Metoda t-SNE je učinkovita pri ohranjanju lokalne in globalne strukture v podatkih. Zajema lokalne informacije med bližnjimi podatkovnimi točkami in globalne informacije med oddaljenimi podatkovnimi točkami. Za potrebe naše vizualizacije značilke smo uporabili implementacijo iz `sklearn.manifold` [51] in značilke vizualizirali v 2D prostoru, primer je na sliki 2.4.



Slika 2.4: Prikaz vizualizacije podatkov opisanih s 15 značilkami s t-SNE v 2 dimenzijah.

2.8 Metrike za vrednotenje

Kot metrike za vrednotenje strojnega učenja pri klasifikaciji smo izbrali:

- točnost (*angl. accuracy*),
- priklic (*angl. recall*),
- natančnost (*angl. precision*),
- metriko F1.

Slednje metrike so primerne za klasifikacijske probleme in temeljijo na tabeli napačnih klasifikacij, ki je definirana v tabeli 2.1.

	Napovedani +	Napovedani –
Dejansko +	pravilno razvrščeni + primeri <i>angl. True Positives (TP)</i>	napačno razvrščeni – primeri <i>angl. False Negatives (FN)</i>
Dejansko –	napačno razvrščeni + primeri <i>angl. False Positives (FP)</i>	pravilno razvrščeni – primeri <i>angl. True Negatives (TN)</i>

Tabela 2.1: Tabela napačnih klasifikacij (*angl. confusion matrix*). Oznaka + pomeni pozitivne primere, – pa negativne primere

Točnost je definirana po enačbi 2.6,

$$\text{Točnost} = \frac{TP + TN}{TP + TN + FP + FN}. \quad (2.6)$$

Pove nam razmerje med številom ujemanj napovedi in številom vseh primerov. Problem te metrike je, da v primeru neuravnovešenih razredov lahko dobimo dobre rezultate, čeprav zares ti niso. Zato potrebujemo še ostale metrike.

Priklic nam pove razmerje pravilno razvrščenih pozitivnih primerov med vsemi pozitivnimi primeri. Definiran je z enačbo 2.7,

$$\text{Priklic} = \frac{TP}{TP + FN}. \quad (2.7)$$

Natančnost je število pravih pozitivnih napovedi glede na vse pozitivne napovedi. Definirana je po enačbi 2.8,

$$\text{Natančnost} = \frac{TP}{TP + FP}. \quad (2.8)$$

Metrika **F-1** nam zagotavlja ravnovesje med priklicem in natančnostjo. Definirana po enačbi 2.9,

$$F1 = \frac{2 \cdot \text{Natančnost} \cdot \text{Priklic}}{\text{Natančnost} + \text{Priklic}} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}. \quad (2.9)$$

Ker se soočamo s problemom klasifikacije v več razredov, smo uporabili makro natančnost, makro priklic in makro F1. V kontekstu točnosti termin makro ni potreben, saj po definiciji že zajema celovito analizo. Makro pristop pomeni, da za vsak klasifikacijski razred posebej izračunamo metrike in jih nato povprečimo. S tem zagotovimo enakopravno obravnavo vseh razredov, neodvisno od njihove pogostosti v podatkovnem naboru. Makro natančnost izračunamo po enačbi 2.10,

$$\text{makro natančnost} = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FP_i}, \quad (2.10)$$

makro priklic po enačbi 2.11,

$$\text{makro priklic} = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FN_i}, \quad (2.11)$$

makro F1 pa enačbi 2.12 ,

$$\text{makro F1} = 2 \times \frac{\text{Makro Preciznost} \times \text{Makro Priklic}}{\text{Makro Preciznost} + \text{Makro Priklic}}. \quad (2.12)$$

2.9 Izbira značilnk

Za identifikacijo najprimernejših značilnk v našem naboru značilnk smo se odločili za uporabo metode SelectKBest [52] iz knjižnice scikit-learn [53]. Ta metoda spada med t.i. filtrirne pristope izbire značilnk, kar pomeni, da je

izbira značilke neodvisna od algoritmov strojnega učenja. Izbira značilke se izvaja pred fazo strojnega učenja in se ne prilagaja na podlagi modela, ampak se zanaša na statistične izračune za ocenjevanje in razvrščanje značilke [54]. Metoda SelectKBest [52] omogoča uporabo različnih statističnih in regresijskih cenilnih funkcij, ki ocenjujejo povezavo značilke z odvisno spremenljivko. Za klasifikacijo so relevantni slednji testi: hi kvadrat test, ANOVA F-test (*angl. F-test in one-way analysis of variance*) in ocena medsebojne informacije (*angl. mutual information*). Na podlagi pridobljenih vrednosti testne statistike metoda izbere k najboljših značilke, ki so najbolj informativne za napovedovanje izhodne spremenljivke. Hi kvadrat testa nismo izbrali, ker je bolj primeren za kategorične spremenljivke, naši podatki pa so numerični. Zato sta za našo analizo bila najbolj primerna ocena medsebojne informacije in ANOVA F-test.

Cenilna funkcija `f.classif` [55] temelji na ANOVA F-testu. Primerna je za numerične spremenljivke. Ta test preverja hipotezo o enakosti povprečnih vrednosti več skupin, kar v kontekstu izbire značilke pomeni iskanje linearnih odvisnosti med značilkami in odvisno spremenljivko. Kakovost posameznih značilke ocenimo z vrednostmi F-statistike, ki izraža razmerje med variabilnostjo povprečnih vrednosti značilke med različnimi skupinami in znotraj skupin (enačba 2.13).

$$F_{testANOVA} = \frac{\text{razlika med povprečji vzorcev}}{\text{razlika znotraj vzorcev}} \quad (2.13)$$

Visoke vrednosti F-statistike nakazujejo, da določena značilka pomembno prispeva k razlikovanju med skupinami in je zato uvrščena med k izbranih značilke za nadaljnjo analizo. Za oceno medsebojne informacije smo uporabili cenilno funkcijo `mutual_info_classif` [56]. Ocena medsebojne informacije lahko deluje za numerične in kategorične podatke ali kombinacijo teh. Temelji na konceptu vzajemne informacije med posameznimi značilkami in ciljno spremenljivko. Vzajemna informacija je mera količine informacij, ki jih vsebuje ena naključna spremenljivka o drugi. Višja medsebojna informacija pomeni, da je značilka bolj relevantna. Algoritem tako tiste z najvišjimi vrednostmi uvrsti med k najboljših.

2.10 Klasifikatorji

Pri določanju stopnje motorične prizadetosti pri Parkinsonovi bolezni preko testa tapkanja po MDS-UPDRS gre za klasifikacijski problem. Poznamo klasifikacijske in regresijske probleme. Pri regresijskih so podatki zvezni in iščemo matematično funkcijo, ki se čim bolj prilega podatkom. Slednje se imenuje regresija. V klasifikaciji pa podatkom določimo razred. V našem primeru gre za večrazredno klasifikacijo, saj imamo 5 razredov. V našem delu uporabljamo nadzorovano učenje, saj imamo znane razrede oziroma označene podatke. Poznamo še nenadzorovano in spodbujevano učenje. Opisi tipov učenj so prikazani v tabeli 2.2.

Tip učenja	Opis učenja	Podatki	Uporaba
Nadzorovano	Nauči se preslikave med vhodi in izhodi	Označeni (z vnaprej določenimi oznakami)	Razvrščanje, regresija
Nenadzorovano	Išče skrite vzorce v podatkih	Neoznačeni (brez vnaprej določenih oznak)	Grupiranje, zmanjšanje prostora značilk
Spodbujevano	Uči se preko interakcij z okoljem in sistema nagrad	Neoznačeni (brez vnaprej določenih oznak)	Robotika, rač. igre

Tabela 2.2: Tipi strojnega učenja in njihove osnovne značilnosti.

2.10.1 Metoda podpornih vektorjev

Metoda podpornih vektorjev je algoritem, ki v več dimenzionalnem prostoru išče hiper ravnino, ki najbolj jasno loči različne razrede podatkov. Od linearne regresije se razlikuje v tem, da za iskanje hiper ravnine uporablja statistični pristop namesto verjetnostnega. Uporablja se za klasifikacijo, regresijo

in zaznavo odstopanj (*angl. outliers*). Je učinkovita rešitev za visokodimenzionalne probleme, pomnilniško učinkovita in delno razložljiva. Slabost metode podpornih vektorjev je zmožnost ločevanja le linearno odvisnih podatkov. Za reševanje nelinearnih odvisnosti pa potrebuje dodatne metode npr. preslikavo podatkov v višje dimenzije, itd. Za večrazredno klasifikacijo uporablja pristop eden proti enemu. Torej konstruira $n \cdot \frac{n-1}{2}$ klasifikatorjev, kjer n predstavlja število razredov. Nato pa s pomočjo odločitvene funkcije rezultate primerja in določi končne vrednosti.

2.10.2 Logistična regresija

Logistična regresija je statistični model primeren za binarno in linearno klasifikacijo. Deluje podobno kot metoda podpornih vektorjev, razlikuje se v tem, da za iskanje hiper ravnine uporablja verjetnostni model. Prav tako kot SVM za večrazredno klasifikacijo uporablja pristop eden proti enemu in odločitveno funkcijo, da primerja posamezne klasifikatorje. Prednost logistične regresije je, da je relativno enostavna za interpretacijo. Metoda nudi tudi dobro odpornost na šumne podatke in je sama po sebi manj podvržena problemu pretiranega prilagajanja podatkov (*angl. overfitting*). Njena slabost je, da lahko zaznava le linearne odvisnosti in da potrebuje več podatkov kot je število značilk.

2.10.3 Odločitvena drevesa

Odločitvena drevesa so neparametrična metoda nadzorovanega strojnega učenja primerna za klasifikacijo in regresijo. Ciljno spremenljivko napoveduje tako, da se naučijo preprostih odločitvenih pravil iz značilk podatkov. So lahka za interpretacijo, razumevanje ter vizualizacijo. Lahko obravnavajo kategorične in numerične ter manjkajoče značilke. Njihov problem je, da strmiijo k pretiranemu prilagajanju in tudi visoki varianci napovedi. Da se temu izognemo uporabljamo različne tehnike. To sta recimo obrezovanje odločitvenih dreves (*angl. decision tree pruning*) [57] in združevalne metode (*angl. ensemble*

methods), kot so naključni gozdovi [58].

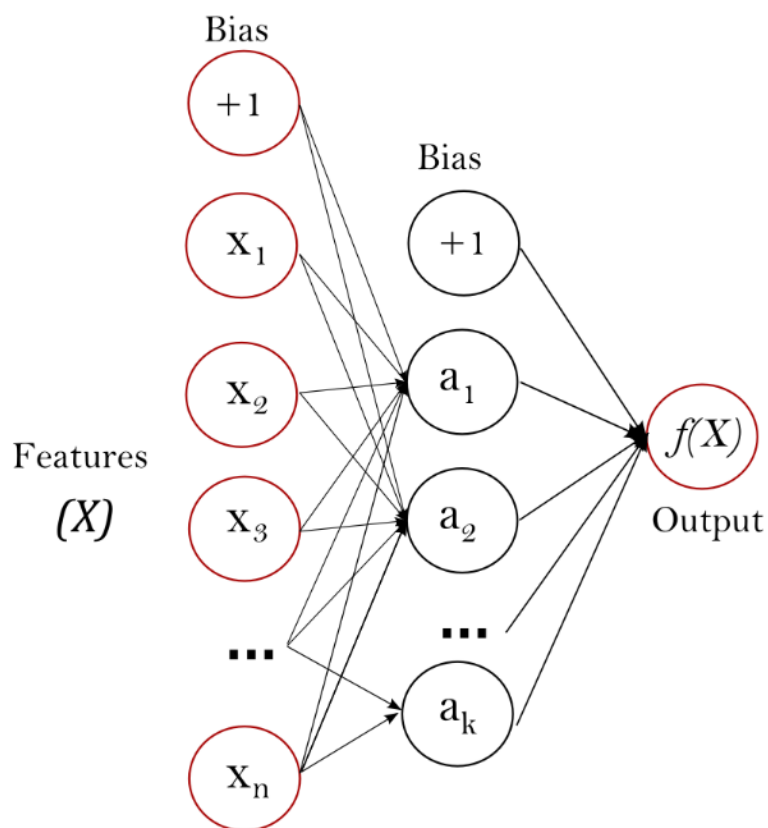
2.10.4 Naključni gozdovi

Naključni gozdovi so združevalna metoda za klasifikacijske ali regresijske probleme. Sestavljeni so iz velikega števila odločitvenih dreves [58], ki na koncu oblikujejo močan model. Združujejo izhode večih odločitvenih dreves, da dobijo le en izhod. Kot rezultat upoštevajo povprečne napovedi večih dreves pri regresijskih problemih oziroma število glasov dreves posameznih razredov pri klasifikaciji. Iz njih je lahko razbrati pomembnost posameznih značilk. So manj nagnjeni k pretiranemu prilagajanju kot odločitvena drevesa. Tipično so kompleksni in težko razložljivi ter procesno zahtevnejši.

2.10.5 Večplastni perceptron

Večplastni perceptron (MLP, *angl. multilayer perceptron*) je nadzorovana tehnika strojnega učenja. Predstavljamo si ga kot mrežo nevronov, ki so med seboj povezani, vidno na sliki 2.5. Nevroni obdelujejo podatke in jih prenašajo naprej skozi mrežo, dokler ne dosežejo izhoda. Mreža se uči tako, da prilagaja uteži povezav med umetnimi nevroni. Obdržijo se tiste uteži, ki zagotavljajo najboljši rezultat glede na izbrano metriko. Ta je ponavadi natančnost ali pa funkcija izgube.

Njegova prednost je, da se lahko nauči nelinearnih relacij med podatki, hkrati pa se model lahko uči sproti, brez potrebe po ponovnem učenju. To pomeni, da s pridobitvijo dodatnih podatkov modela ne rabimo znova modelirati, ampak ga le priučimo (*angl. on-line learning*). Njegova slabost je, da je za optimizacijo učenja potrebno nastaviti veliko parametrov (npr. število skritih slojev, nevronov, hitrosti učenja, iteracij, itd.). Prav tako je občutljiv na skalo značilk, zato je pomembno, da te ustrezno normaliziramo. Zaradi nekonveksne funkcije izgube, se lahko rezultati napovedi razlikujejo v odvisnosti od nastavitve začetnih uteži [59]. Primer arhitekture MLP je prikazan na sliki 2.5.



Slika 2.5: Prikaz arhitekture MLP z enim skritim nivojem. Features pomeni vhodne značilke, Output pomeni izhod, Bias pa vmesne uteži. Slika je vzeta iz [59].

2.10.6 Konvolucijske nevronske mreže

Konvolucijske nevronske mreže spadajo v družino globokih nevronskih mrež. Uporabljajo veliko število filtrov in zmanjševalne nivoje, da se zmanjša število parametrov. Z različno velikimi filtri mreža poizkuša iz podatkov izluščiti specifične vzorce oziroma značilke. V splošnem so konvolucijske nevronske mreže pomnilniško učinkovitejše od drugih tipov nevronskih mrež. V grobem lahko rečemo, da gre za avtomatsko metodo, saj takšne mreže same zaznajo pomembne lastnosti in odvisnosti v podatkih ter izluščijo značilke. Tipično se uporabljajo za obdelavo slik in videov, procesiranja naravnega jezika, lahko pa tudi za dvodimenzionalne časovne podatke, kar je bilo po-

kazano v večih raziskavah [60, 61, 62, 63]. Arhitekturno gledano je CNN sestavljena iz treh glavnih nivojev: začetni konvolucijski nivo (*angl. convolutional layers*), vmesni združevalni nivo (*angl. pooling layers*) in končni polno povezani nivo (*angl. fully connected layer*).

V nalogi smo uporabili polno povezano konvolucijsko nevronske mrežo, ki je po zgradbi podobna CCN-ju z razliko, da je zadnji polno povezan nivo zamenjan s konvolucijskim nivojem. Prednost tega je, da se ohranijo prostorske informacije med podatki, potrebuje pa tudi manj parametrov. Prednosti nevronske mreže CNN in FCN so, da same izluščijo odvisnosti v podatkih, pri tem pa zajamejo globalno in lokalno strukturo teh. Njihove slabosti so, da so računsko zahtevne, tipično pa za učenje potrebujejo veliko podatkov. So slabo razložljive in jih obravnavamo kot črne škatle, prav tako so tudi nagnjene k pretiranemu prilagajanju. V namen preprečevanja tega uporabljamo različne tehnike, kot so regularizacija, izpuščanje, normalizacija paketov, združevanje, zgodnje zaustavljanje, itd. Izpuščanje preprečuje pretirano prilagajanje tako, da naključno izbrane uteži v mreži povrne na začetno vrednost. Zgodnje zaustavljanje ustavi učenje modela, ko ni napredka v določenem številu iteracij in s tem prepreči slabšanje modela. Normalizacija paketov nam v nevronske mreže zraven preprečevanja pretiranega prilagajanja doda še stabilnost in omogoča hitrejše delovanje.

Poglavje 3

Raziskovalna metodologija

Metodo za klasifikacijo in vizualizacijo podatkov smo razvili v Pythonu, saj nudi dobro podporo za strojno učenje, računalniški vid in video obdelavo. V nadaljevanju poglavja je opisan pristop razvoja klasifikatorja za oceno stopnje motorične prizadetosti na podlagi testa tapkanja s prsti po lestvici MDS-UPDRS. V prvem koraku smo zajeli podatkovno zbirko in jo obdelali, da je primernejša za nadaljnje procesiranje. Tako smo pripravili podatke za gradnjo značilk. Podatkovna zbirka in njena obdelava je enaka v vseh opisanih pristopih v nadaljevanju poglavja. Slednji se razlikujejo po izdelavi značilk in načinih strojnega učenja. Izmed vseh značilk smo nato izbrali najustrežnejše in jih uporabili v različnih klasifikacijskih metodah strojnega učenja. Uporabili smo tudi polno konvolucijsko nevronske mrežo, ki ne potrebuje značilk, saj sama izlušči odvisnosti v podatkih. Za potrjevanje rezultatov smo uporabili 10-kratno navzkrižno preverjanje.

3.1 Podatkovna zbirka

Videi tapkanja so bili posneti z različnimi pametnimi telefoni. Večino videov je bilo zajetih od pacientov s Parkinsonovo boleznijo v ambulantni na Univerzitetnem kliničnem centru v Ljubljani, nekaj pa od prostovoljcev v različnih okoljih. Ker so bili posnetki zajeti v realnem okolju, ni nujno, da je bila ka-

mera postavljena pravokotno na profil palca in kazalca. Tudi osvetlitve med videi se razlikujejo, razdalje med kamero in roko pa so različne. Nekateri videi so bili posneti pri hitrosti 30 sličic na sekundo, večina pa pri 60 sličicah na sekundo. Ločljivost snemanja je bila 3840 x 2160 slikovnih pik, pri nekaterih pa 1920 x 1080 slikovnih pik. Videi so bili za potrebe analize obrezani tako, da je na njih vidna celotna roka in manj okolice. To nam omogoča hitrejše procesiranje ter izvedbo analize na manj zmogljivih sistemih, zaradi manjšega števila slikovnih pik videa. Posnetki, kjer celotna roka ni bila vidna, ter posnetki, kjer so koti snemanja močno odstopali od pravega kota so bili odstranjeni, saj so povzročali težave pri zaznavi roke. Prav tako smo odstranili posnetke udeležencev, ki niso imeli potrjene Parkinsonove bolezni vendar pa so po lestvici MDS-UPDRS dobili oceno višjo od nič. Po odstranitvi neustreznih videov smo zbrali 183 posnetkov pridobljenih od 91 različnih oseb. Stopnjo motorične prizadetosti v posnetkih je po lestvici MDS-UPDRS od 0 do 4 ocenil izkušen nevrolog. Tipično smo za vsakega udeleženca pridobili 2 video posnetka. Enega iz tapkanja leve in drugega z tapkanja desne roke. Pri določanju stopnje motorične prizadetosti na podlagi testa tapkanja se ocenjuje dominantna roka. Ker smo v sklopu magistrskega dela raziskovali natančnost klasifikacije stopnje motorične prizadetosti po MDS-UPDRS, smo lahko uporabili tapkanje obeh rok enega udeleženca. Namreč tapkanje leve in desne roke sta med seboj neodvisni.

Ocena MDS-UPDRS	0	1	2	3	4	Skupaj
Število videov	49	51	53	23	7	183
Delež %	26	28	29	13	4	100

Tabela 3.1: Število video posnetkov testa tapkanja po razredih MDS-UPDRS in njihovi deleži.

3.2 Predprocesiranje podatkov

Posnetke smo najprej obrezali, tako smo odstranili čim več slikovnih pik okolice, ki ne zajemajo območja roke. Nato smo na posnetkih pogнали MediaPipe Hand, ki v določenih primerih pozicije prstov ni natančno zaznal oziroma je zamenjal palec in sredinec, v redkih primerih posameznih sličic videa je zaznavo tudi popolnoma izpustil. Za namen izboljšanja detekcije pozicije prstov smo videom zvišali resolucijo in posameznim sličicam odstranili zamegljenost zaradi gibanja. Zamegljenost zaradi gibanja (vidno na sliki 3.1) je zamegljenost, ki jo opazimo pri hitro premikajočih se predmetih na fotografiji ali posamezni sličici videa. Do tega pride, ker se predmeti med zajemom fotografije (*angl. expose time*) ali kadra premikajo, zaslonka pa v času zajema ne uspe popolnoma zabeležiti pozicije objekta. Zaradi zamegljenosti roke oziroma prstov, jih MediaPipe ne zazna ali pa ne določi natančne pozicije skeleta. Za odstranjevanje zamegljenosti objektov zaradi gibanja smo preizkusili metodi, ki so ju predstavili Liu in sodelavci [64] ter Son in sodelavci [65].

V prvi metodi [64] so za odstranjevanje zamegljenosti zaradi gibanja razvili lahko generativno kontradiktorno mrežo (*angl. lightweight generative adversarial network*) Ghost-DeblurGAN. Opisana metoda na naših podatkih ni dajala vidnih izboljšav, zato smo jo opustili. V drugi metodi [65] so za odstranjevanje zamegljenosti zaradi gibanja uporabili informacije sosednjih sličic v videu. Tukaj v realnih podatkih naletimo na problem, da so tudi sosednje sličice zamegljene, zato njihova metoda deluje v dveh korakih. V prvem koraku se izvede učenje napovedi gibanja brez zamegljenosti, da se izboljša natančnost napovedi gibanja med zamegljenimi sličicami. V drugem delu pa gre za kompenzacijo gibanja. Namesto poravnave posnetkov preko združevanja z napovedjo gibanja, uporabljajo volumen slikovnih pik (*angl. pixel volume*), ki vsebuje kandidate jasnih slikovnih pik, da se odpravijo napake pri napovedi gibanja. Slednja metoda je v določenih primerih dala vidne izboljšave, v določenih pa je sličice poslabšala. Posledično se je zaradi slabih sličic poslabšalo sledenje skeletu z MediaPipe Hand in tudi sama klasi-

fikacija, zato odstranjevanja zamegljenosti zaradi gibanja na podatkih nismo uporabili. Primera zamegljenosti zaradi gibanja in artefaktov po odstranitvi zamegljenosti sta vidna na sliki 3.1

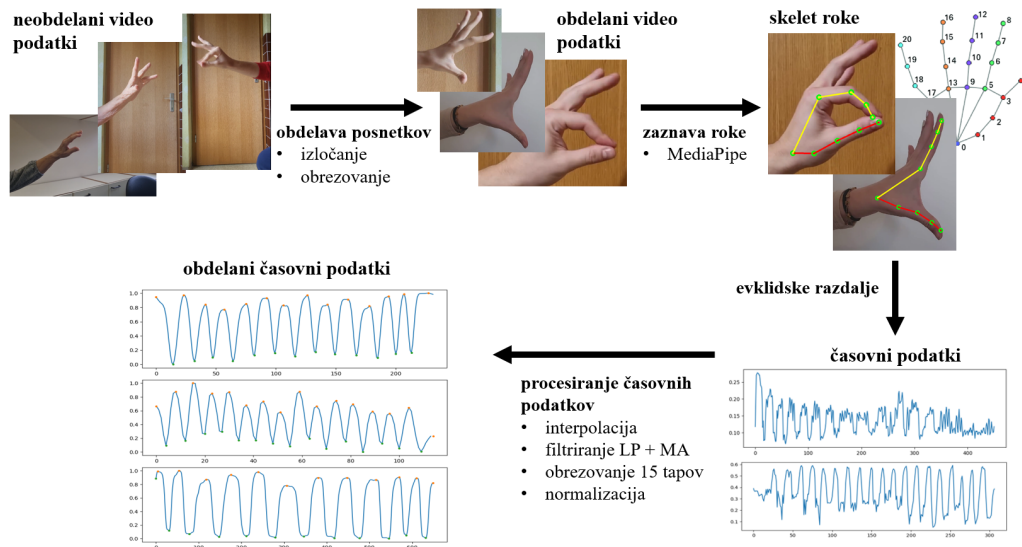


Slika 3.1: Desno je prikazana zamegljenost zaradi gibanja, ki se pojavi zaradi snemanja pri nizkem številu sličic. Levo so prikazani artefakti, ki so nastali pri odstranjevanju zamegljenosti zaradi gibanja, zato slednjega nismo izvajali.

Za zviševanje resolucije smo preizkusili orodje Video2x [66], ki resolucijo zvišuje preko strojnega učenja in nudi podporo za Nvidia grafične kartice, za hitrejšo obdelavo podatkov. Videom smo resolucijo podvojili, vendar je bil vpliv na sledenje skeleta in klasifikacijo minimalen. Zaradi dolgotrajnega procesiranja in zanemarljivih rezultatov v nadaljevanju nismo uporabljali videov z zvišano resolucijo.

V nadaljevanju smo MediaPipe Hand pognali le na neobdelanih obrezanih posnetkih. Iz dobljenega skeleta MediaPipe Hand smo po enačbi 3.1 izračunali evklidske razdalje med končnimi točkami palca in kazalca (točke 4 in 8 na sliki 2.1) in tudi med točkami zadnjih sklepov palca in kazalca (točke 3 in 7 na sliki 2.1).

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (3.1)$$



Slika 3.2: Prikaz postopka obdelave podatkov. Postopek prikazuje pridobitev obdelanih časovnih podatkov iz video posnetkov in je za vse nadaljnje metode enak.

Druga možnost bi bila, da bi namesto evklidskih razdalj med palcem in kazalcem računali kote med palcem in kazalcem. Koti imajo pred razdaljami prednost, saj so neodvisni od razdalje med kamero in roko. So pa veliko bolj občutljivi na kot snemanja. Zaradi načina snemanja je v videih prihajalo do nagiba kamere, kar pomeni, da kot med roko in kamero ni nujno 90 stopinj. Posledično so bile evklidske razdalje boljša izbira kot koti, saj so odpornejše na nagibe.

Dodali smo še linearno interpolacijo za dopolnitev manjkajočega skeleta posameznih sličic. Interpolacijo smo uporabili le v primeru, ko je Mediapipe zgrešil natanko eno sličico, torej je zaznal prejšnjo in naslednjo sličico. Za več zgrešenih sličic interpolacije nismo uporabili, da ne bi preveč priredili podatkov. Videe, kjer je Mediapipe zgrešil več sličic, smo odstranili, saj roke nismo mogli uspešno zaznati. Posledično ni bilo smiselno računati značilnk, ker bi te bile okvarjene. Dobljene razdalje oziroma amplitude smo nato zgladili z nizko prepustnim IIR filtrom. Frekvenco odreja smo določili za vsak vhodni

podatek posebej, glede na maksimalno frekvenco frekvenčnega spektra in ji prišteli del pasovne širine. Postopek je opisan z enačbo 3.2,

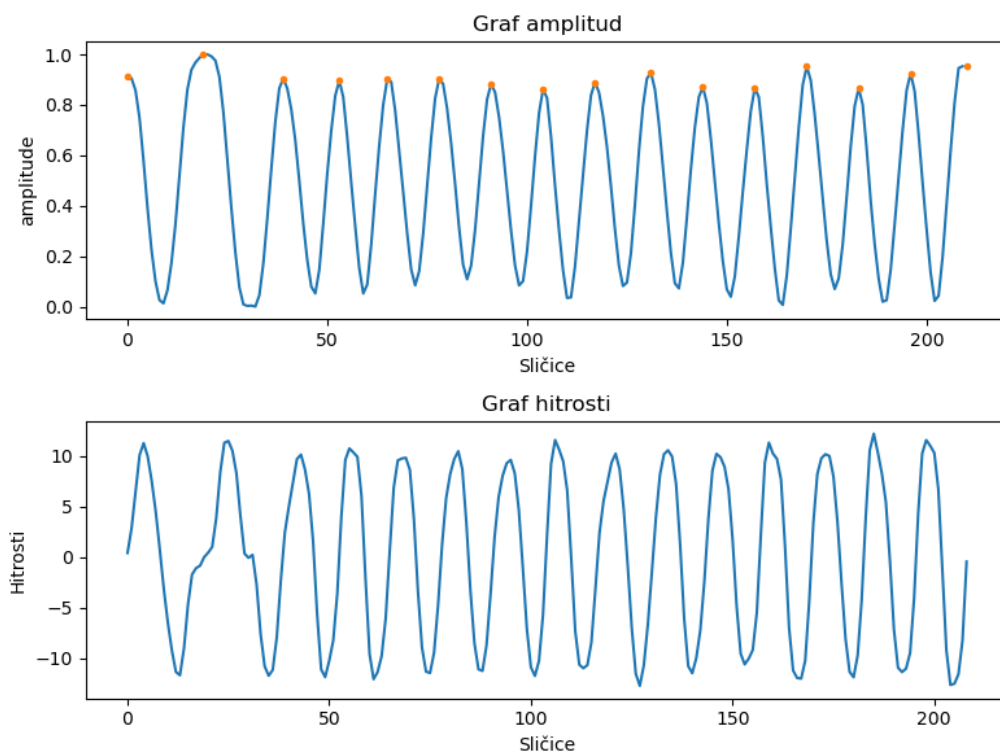
$$f_{odrez} = f_{max} + \frac{\text{pasovna širina} * k}{2}, \quad (3.2)$$

kjer k predstavlja odstotek vpliva. Slednjega smo nastavili na 0,8, ker je predstavljal dobro ravnotežje med glajenjem in ohranitvijo signala. V primeru, da smo ga nastavili na manjšo vrednost, nam je filter nekatere signale močno okvaril. Za namen določitve najvišje frekvence f_{max} smo signale pretvorili v frekvenčni spekter preko Fourierjeve transformacije (*angl. Fourier transformation*). Posledično smo zaradi šibkejšega filtriranja dodali filter drsečega povprečja z dolžino okna 5. Z njim dodatno zgladimo signal in zmanjšamo vpliv nenatančnih zaznav skeleta. MediaPipe Hands prilega skelet na sredino prsta, ki pa ga vedno ne zazna točno, zato rahlo premika skelet med sličicami, tudi če roka miruje. V podatke nam vpelje visokofrekvenčni šum. Filtriranje je pomembno, saj z njim odstranimo del visokofrekvenčnega šuma, ki nastane zaradi prileganja skeleta in tudi zaradi prisotnosti tremorja pri nekaterih pacientih. Tremor otežuje klasifikacijo po MDS-UPDRS, saj nekoliko maskira samo tapkanje, ni pa del ocenjevalne tabele 1.1. Potrebno ga je odstraniti oziroma zgladiti.

Zaradi različnih razdalj med kamero in roko smo na filtriranih amplitudah izvedli min-max normalizacijo, ki vse podatke skalira na enotno skalo med 0 in 1. Signale smo tudi porezali na dolžino 15 ciklov. Te smo zaznali z algoritmom za detekcijo tapkanja opisanem v poglavju 2.4. Opisani postopek je za vse nadaljnje metode enak. V nadaljevanju poglavja so opisani različni pristopi gradnje značilk in klasifikacije.

3.3 Metoda 1

V prvi metodi je bilo narejeno večje število značilk z namenom, da pridobimo širok opis tapkanja. Iz slednjega smo s statistično metodo izbrali najustrežnejše značilke za klasifikacijo. Pri oblikovanju značilk smo imeli v mislih



Slika 3.3: Slika prikazuje graf amplitud in hitrosti pridobljenih iz obdelanih evklidskih razdalj med konicami palca in kazalca. Amplitude so osnova za izpeljavo ostalih spektrov podatkov in posledično za izračun značilk. V primeru nevronske mreže so uporabljene direktno kot vhod. Grafi so porezani na 15 ciklov, označeni z oranžno piko na grafu amplitud. Pripadajoča ocena tapkanja po MDS-UPDRS je 0.

tabelo 1.1 za ocenjevanje tapkanja po MDS-UPDRS, katere se pa nismo strogo držali. Za osnovo za izračun značilke smo uporabili evklidske razdalje med konicami prstov, med točkama 4 in 8, in evklidske razdalje med zadnjimi sklepi palca in kazalca, med točkama 3 in 7. Dodatno smo spremljali tudi absolutne premike zapestja (preko točke 0) in iz njih pridobili značilke. Točke so vidne in označene na sliki 2.1. Za uporabo razdalj med točkama 4 in 8 smo se odločili, ker smo mnenja, da nosijo največ informacij o tapkanju. Tukaj je gibanje prstov najbolj izrazito. Dodali smo še evklidske razdalje med točkama 3 in 7, saj bi lahko nudile dodatne, sicer bolj omejene informacije o tapkanju. Gibanje teh sklepov prsta je veliko manj izrazito kot gibanje konic prstov v točkah 4 in 8. Ostale točke palca in kazalca so med tapkanjem precej statične, zato jih nismo uporabili. Ker smo bili mnenja, da zdrav udeleženec tapka s čim manj gibi roke, smo se odločili spremljati gibanje zapestja preko točke 0.

Iz evklidskih razdalj konic in zadnjih sklepov, smo za vsako pridobili spekter amplitud in iz njega izpeljali dodatne 4 spektre. To so: spekter hitrosti, pospeškov, frekvenc in maksimumov amplitud. Slednje smo pridobili z algoritmom za zaznavo in filtriranje maksimumov opisanem v poglavju 2.4. Dodaten spekter še predstavljajo absolutni premiki zapestja. Iz vseh 6 spektrov smo nato pridobili različne značilke. Primer spektra hitrosti in amplitud je viden na sliki 3.3. Med značilke smo dodali še čas najdaljše zamrznitve in število zamrznitev, ki smo jih beležili iz amplitud med konicami prstov (točki 4 in 8, vidni na sliki 2.1). Algoritem za detekcijo zamrznitev je opisan v poglavju 2.5. Računali smo tudi percentile, ki nam povejo, kako je razporejena določena količina podatkov. Iz razlike 75. in 25. percentila smo izračunali kvartilno razliko, ki nam pove razpršenost srednjih 50 % podatkov.

Za ekstreme, torej maksimume amplitud smo zabeležili povprečno odstopanje lokalnih ekstremov od absolutno največjega ekstrema. Izračunali smo tudi povprečje 3 največjih odstopanj od absolutno največjega ekstrema. Za število 3 smo se odločili zaradi večje odpornosti na šum. Pri detekciji ciklov smo odvečne ekstreme odstranili in kot značilko uporabili število odstranje-

nih ekstremov. V frekvenčnem spektru smo izračunali še razpon, ki je razlika maksimalne in minimalne frekvence. Izračunali smo še dominantno frekvenco tapkanja ter ploščatost spektra, ki je kvocient geometrične in aritmetične sredine in nam pove enakomernost razporeditve energije po frekvenčnem spektru. Za boljši opis razporeditve frekvenc smo izračunali še masno središče, centroid spektra in neperiodičnost frekvenc. Masno središče in centroid frekvenc opisujeta povprečno frekvenco prispevkov energije v spektru signala, razlika pa je v njunem izračunu. Masno središče se računa izključno na magnitudah spektra in ne uporablja dejanskih frekvenc. Metoda je bolj splošna in primernejša za spektre z neenakomerno razporejenimi frekvencami. Pri izračunu centroida pa se uporabljajo frekvence. Za spremljanje upada smo naredili prileganje polinoma druge stopnje na frekvencah in dobljene parametre uporabili kot značilko, ki predstavlja naklon spektra. Določene značilke so še opisane v poglavju 3.4. Skupaj smo naredili 193 značilk, najdemo jih v tabeli 3.2.

Značilka	Podatki
povprečje	amplitude, hitrosti, pospeški, zapestje, ekstremi
mediana	amplitude, hitrosti, pospeški, zapestje, ekstremi
standardna deviacija	amplitude, hitrosti, pospeški, zapestje, ekstremi
varianca	amplitude, hitrosti, pospeški, zapestje, ekstremi
srednje absolutno odstopanje	amplitude, hitrosti, pospeški, zapestje, ekstremi
kvartilni razmik	amplitude, hitrosti, pospeški, zapestje, ekstremi
percentili 10, 25, 75, 90	amplitude, hitrosti, pospeški, zapestje, ekstremi
absolutna sploščenost	amplitude, hitrosti, pospeški, zapestje, ekstremi
absolutna poševnost	amplitude, hitrosti, pospeški, zapestje, ekstremi
koeficient variacije	amplitude, hitrosti, pospeški, zapestje
minimum	hitrosti, pospeški, zapestje
maksimum	hitrosti, pospeški, zapestje
entropija	vsi
sploščenost	vsi
poševnost	vsi
št. zamrznitev	amplitude
trajanje maks. zamrznitve	amplitude
povp. razlika od maks.	ekstremi
3 največje razlike od maks.	ekstremi
št. ekstremov med cikli	ekstremi
razpon	frekvence
moč	frekvence
dominanta frekvenca	frekvence
centroid spektra	frekvence
pasovna širina	frekvence
masno središče	frekvence
neperiodičnost	frekvence
razširitev spektra	frekvence
naklon spektra	frekvence
ploščatost spektra	frekvence

Tabela 3.2: Značilke metode 1 in spektri podatkov za pridobitev teh. Oznaka vsi pomeni uporabo spektrov amplitud, hitrosti, pospeškov, frekvenc, premikov zapestja in ekstremov, ki so lokalni maksimumi amplitud.

3.4 Metoda 2

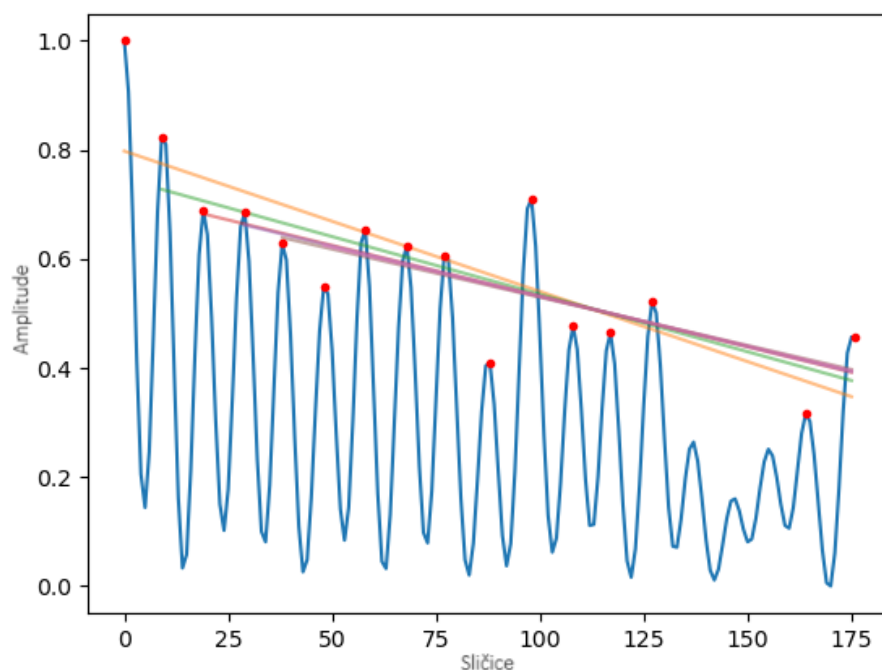
Da bi izboljšali klasifikacijo smo se lotili gradnje značilk strogo po navodilih lestvice MDS-UPDRS za ocenjevanje tapkanja s prsti. Navodila za ocenjevanje so vidna v tabeli 1.1. Za izdelavo značilk smo uporabili le evklidske razdalje med konicama palca in kazalca, torej razdalje med točkami 4 in 8 vidne na sliki 2.1. Te najbolj odražajo težave med izvajanjem testa tapkanja. Izdelane značilke lahko razdelimo v 3 sklope. Prvi sklop značilk opisuje zamrznitve in oklevanja glede na točko a po tabeli 1.1, drugi sklop značilk opisuje upad hitrosti, po točki b v tabeli 1.1. Zadnji, tretji sklop značilk pa opisuje upad amplitud, po točki c v tabeli 1.1. Primer podatkov, ki smo jih uporabili za gradnjo značilk je viden na sliki 3.3.

Za oceno zamrznitev, oklevanj in zakasnitev v točki a smo izdelali algoritem, ki sledi intervalom, kjer so spremembe amplitud dovolj majhne. Podrobnejše delovanje algoritma je opisano v poglavju 2.5. Za izbiro amplitud smo se odločili, ker so bolj odporne na šum kot hitrosti. Zaradi prilagajanja skeleta glede na sredino prsta se lahko slednje izrazi v visokih hitrostih, ki pa dejansko niso prisotne. Kot značilke smo uporabili število zamrznitev, najdaljši čas zamrznitve in povprečen čas med cikli. Slednje so vidne v tabeli 3.3.

Za oceno upada hitrosti in amplitud v točki b in c so si značilke podobne, saj v obeh primerih spremljamo upad vrednosti in kdaj se ta zgodi. Za izgradnjo značilk, ki opisujejo upad hitrosti oziroma amplitud, smo uporabili časovne podatke, avtokorelirane časovne podatke in lokalne ekstreme časovnih vrst amplitud in hitrosti. Ekstreme, ki predstavljajo cikle, smo izluščili z algoritmom za zaznavo in filtriranje ekstremov opisanem v poglavju 2.4. V primeru časovne vrste hitrosti smo uporabili maksimume in minimume, ker predstavljajo hitrost odpiranja oziroma zapiranja roke. Pri amplitudah smo uporabili le maksimume, saj nas minimumi, ki pomenijo, da sta prsta staknjena, ne zanimajo. Iz časovne vrste ekstremov smo nato izračunali značilke. Za spremljanje upada amplitud in hitrosti smo naredili več linearnih regresij na zaznanih lokalnih ekstremih. Primer linearnih re-

gresij na ne logaritmiranih ekstremih amplitud je viden na sliki 3.4. Vzeli smo absolutno vrednost ekstremov in jim prišteli majhno vrednost zaradi stabilnosti postopka. Nato smo jih logaritmirali, da smo zmanjšali njihovo varianco in posledično občutljivost na šum. Tako smo ekstreme pripravili za izvedbo linearnih regresij. Ker je v videu največ 15 ciklov, torej 16 maksimumov amplitud, smo dobili 15 regresijskih premic in njihove koeficiente uporabili kot značilke. Najprej smo naredili regresijo čez vseh 16 ekstremov, potem smo prvi ekstrem izločili in naredili regresijo na ostalih. Postopek smo ponavljali dokler nismo prišli do linearne regresije zadnjih 2 ekstremov. Na sliki 3.4 lahko vidimo, da se regresijske premice začnejo na različnih pozicijah, ki sovpadajo z ekstremiti po opisanem postopku. V primerih kjer smo imeli manj kot 15 ekstremov smo manjkajoče koeficiente dopolnili z ničelnimi vrednostmi. V primeru hitrosti smo postopek naredili ločeno za maksimume in minimume. Ker lahko pri hitrostih dobimo več kot 16 ekstremov zaradi oklevanj med tapkanjem, smo v takšnih primerih ekstreme razdelili v tretjine. Nato smo naredili preslikavo tako, da smo iz vsake tretjine izbrali 5 najmanjših koeficientov upadanja, torej 5 najbolj negativnih koeficientov. Slednje je ustrezno, saj po lestvici MDS-UPDRS iščemo v kateri tretjini tapkanja pride do upada amplitud oziroma hitrosti. Tako smo na koncu iz ekstremov amplitud in hitrosti pridobili 90 značilk iz koeficientov naklona premic in odsekov na y-osi.

Iz vseh časovnih spektrov smo izračunali še poševnost (*angl. skewness*), ki nam opiše asimetričnost verjetnostne porazdelitve. Torej nam pove, če je porazdelitev maksimumov zamaknjena v levo ali desno stran od srednje vrednosti. Za poševnost lahko rečemo, da nam nudi omejene opise upada vrednosti in je primerna kot podporna mera v kombinaciji z drugimi merami za opis upada vrednosti. Kot značilko smo uporabili tudi sploščenost (*angl. kurtosis*), ki nam pove, koliko imamo odstopanj v podatkih (*angl. outliers*). Pomembna značilka se nam je zdela tudi varianca, ki je mera razpršenosti podatkov in je kvadrat standardnega odklona. Pove nam kako daleč so posamezne vrednosti razpršene od povprečja. Dodatno smo izbrali še koeficient



Slika 3.4: Prikaz večih linearnih regresij za zaznavo upada amplitud oziroma hitrosti. Zaradi preglednosti na grafu ni označenih vseh 15 regresijskih premic. Na grafu so prikazane amplitude tapkanja udeleženca z oceno 3 po MDS-UPDRS.

variacije, ki je podobna mera kot standardni odklon le, da nam razpršenost izrazi relativno. Omogoča primerjavo podatkov zajetih iz različnih statističnih populacij.

Avtokorelacija signalov omogoča identifikacijo notranjih vzorcev in periodičnosti v podatkih. Iz avtokoreliranih hitrosti in amplitud, ter iz ostalih spektrov smo izračunali še entropijo, ki nam pove količino informacij signala. Energija avtokoreliranega signala se nanaša na skupno količino moči ali intenzivnosti, ki jo signal prenaša čez čas. Zabeležili smo tudi najvišjo vrednost avtokorelacije brez zamika 0 in njen čas, kar nam omogoča vpogled v stopnjo podobnosti med trenutnim stanjem signala in njegovimi preteklimi stanji.

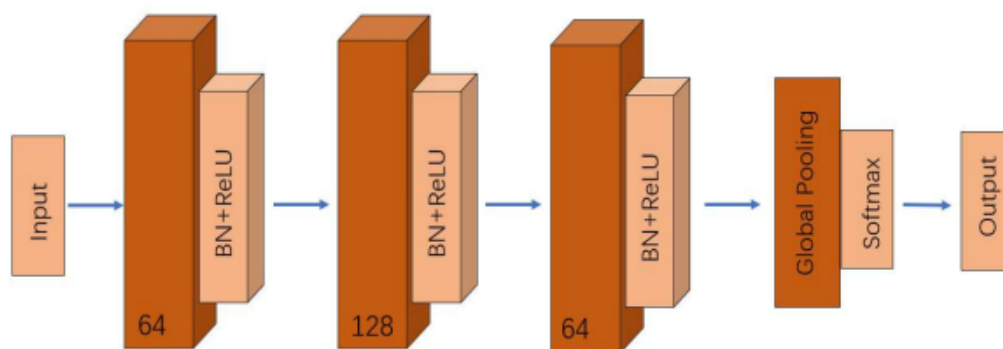
Kot značilke smo še izračunali maksimalno in povprečno spektralno gostoto avtokorelacije, ki nam daje vpogled v periodične vzorce oziroma trende v signalu. Na koncu smo izdelali 145 značilk iz spektrov hitrosti, amplitud, avtokoreliranih hitrosti, avtokoreliranih amplitud, minimumov in maksimumov hitrosti, ter maksimumov amplitud.

Značilka	Podatki
povprečje	amplitude, hitrosti, avtok., ekstremi
varianca	amplitude, hitrosti, avtok., ekstremi
koeficient variacije	amplitude, hitrosti, avtok., ekstremi
entropija	amplitude, hitrosti, avtok., ekstremi
poševnost	amplitude, hitrosti, avtok., ekstremi
sploščenost	amplitude, hitrosti, avtok., ekstremi
mediana	amplitude, hitrosti, ekstremi
standardna deviacija	amplitude, hitrosti
energija	avtok.
prvi ekstrem avtok.	avtok.
čas prvega ekstrema	avtok.
maks. (brez zamika 0)	avtok.
čas pri maks.	avtok.
maks. spektralna gostota	avtok.
povprečna spektralna gostota	avtok.
povp. čas cikla	maksimum amplitud
čas maks. zamrznitve	amplitude
št. zamrznitev	amplitude
regresijski koeficienti	ekstremi

Tabela 3.3: Prikaz izdelanih značilk za metodo 2 in spektrov časovnih vrst podatkov iz katerih smo jih dobili. Avtok. označuje avtokorelirane hitrosti in amplitude. Ekstremi označujejo 3 vrste podatkov: maksimume amplitud, maksimume hitrosti in minimume hitrosti.

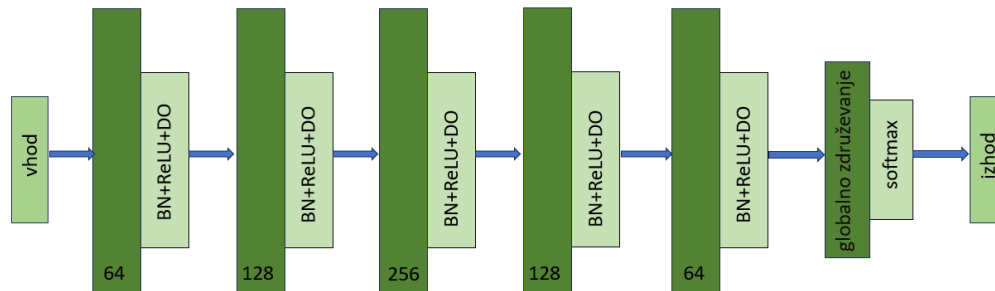
3.5 Metoda 3

V metodi 1 in 2 smo sami naredili časovno neodvisne značilke in jih uporabili v kombinaciji s strojnim učenjem. Izdelava značilk predstavlja izziv, zato smo se za primerjavo odločili preizkusiti nevronske mreže za procesiranje časovnih vrst. Takšen pristop nam omogoča, da zaobidemo oblikovanje značilk, ki zahteva dobro poznavanje problematike in je časovno zahteven. Prednost takšnega pristopa je, da kot vhod nevronske mreže podamo časovno zaporedje podatkov. Iz njih bo samo izluščilo odvisnosti oziroma značilke. Preizkusili smo polno konvolucijsko nevronske mrežo, ki so jo predstavili Li in sodelavci [39], prikazano na sliki 3.5. Slabost takšne mreže je, da je omejena na procesiranje enako dolgih vhodnih sekvenc. Ker so naše časovne vrste amplitud različno dolge, zaradi različnih dolžin videov, smo jih morali dopolniti do enotne dolžine. Proces pridobitve in obdelave podatkov do koraka dopolnitve je enak kot v metodah 1 in 2 ter je prikazan na sliki 3.2. Izbrani nevronske mreže smo dodali tudi modifikacije.



Slika 3.5: Prikaz polne konvolucijske nevronske mreže (FCN) za klasifikacijo iz časovnih vrst. Vsakemu konvolucijskemu sloju sledi sloj za normalizacijo paketov (BN) in sloj z aktivacijsko funkcijo ReLU. Global Pooling pomeni globalno združevanje, Input pomeni vhod in Output pomeni izhod. Slika je vzeta iz [39].

Povečali smo velikost nevronske mreže z dodatnima konvolucijskima slo-



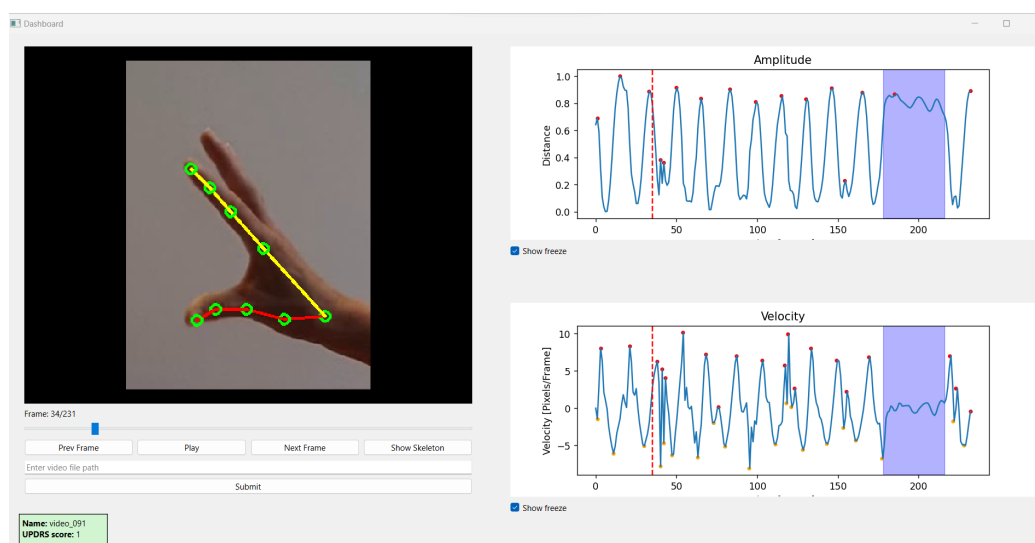
Slika 3.6: Prikaz dopolnjene polne konvolucijske nevronske mreže (FCN) za klasifikacijo iz časovnih vrst. Vsakemu konvolucijskemu sloju sledi sloj za normalizacijo paketov (BN), sloj z aktivacijsko funkcijo ReLU in sloj za izpuščanje (DO).

jema. Prvi dodan sloj ima 256 filtrov, drugi pa 128 filtrov. Da smo se izognili pretiranemu prilagajanju smo dodali zgodnje zaustavljanje in na koncu vsakega od petih konvolucijskih slojev, za aktivacijskimi sloji, dodali še sloje za izpuščanje. Arhitektura mreže je prikazana na sliki 3.6. Časovno vrsto amplitud in hitrosti, primer je viden na sliki 3.3, smo dopolnili do enotne dolžine z uporabo dopolnjevanja na koncu z vrednostjo 0. Nevronsko mrežo smo prilagodili tako, da sprejema dvodimenzionalen vhod v nasprotju z originalnim enodimenzionalnim. Ta sprememba omogoča hkratno uporabo amplitud in hitrosti. Mreži je bilo dodano še zgodnje zaustavljanje. Pri učenju mreže smo obdržali najboljši model glede na natančnost napovedi v testni množici.

3.6 Orodje za vizualizacijo

Za namen vizualizacije podatkov smo pripravili grafični uporabniški vmesnik za vizualizacijo videov tapkanja in grafov amplitud ter hitrosti pridobljenih iz skeleta MediaPipe Hand. Razvili smo ga s pomočjo odprto kodnih knjižnic PyQT [67] in OpenCV [68]. Orodje omogoča prikaz in nalaganje poljubnega video posnetka. Vgrajen ima video predvajalnik, ki omogoča zaustavitev videa, sprehajanje skozi sličice videa in možnost prilagajanja skeleta na video

v realnem času. Na grafu amplitud in hitrosti so označeni cikli tapkanja in trenutna pozicija glede na predvajan video. Na grafih se lahko označijo tudi zamrznitve oziroma oklevanja pri tapkanju. Orodje nam je olajšalo razvoj značilk, saj nam vizualizira ključne podatke in omogoča nazoren vpogled v tapkanje. Hkrati ponuja dobro osnovo kot podporno orodje pri določanju stopnje motorične prizadetosti Parkinsonovih bolnikov na podlagi testa tapkanja po lestvici MDS-UPDRS. Prikaz orodja in njegovih funkcij je viden na sliki 3.7.



Slika 3.7: Aplikacija za vizualizacijo tapkanja. Na levi strani se izriše video z možnostjo prileganja skeleta. Pod videom je predvajalnik, ki omogoča gibanje skozi video. Spodaj levo so osnovni podatki o videu. Na desni sta izrisana grafa hitrosti in amplitud tapkanja z vizualizacijo trenutne pozicije v videu in označenimi zamrznitvami ter cikli tapkanja.

Poglavje 4

Vrednotenje

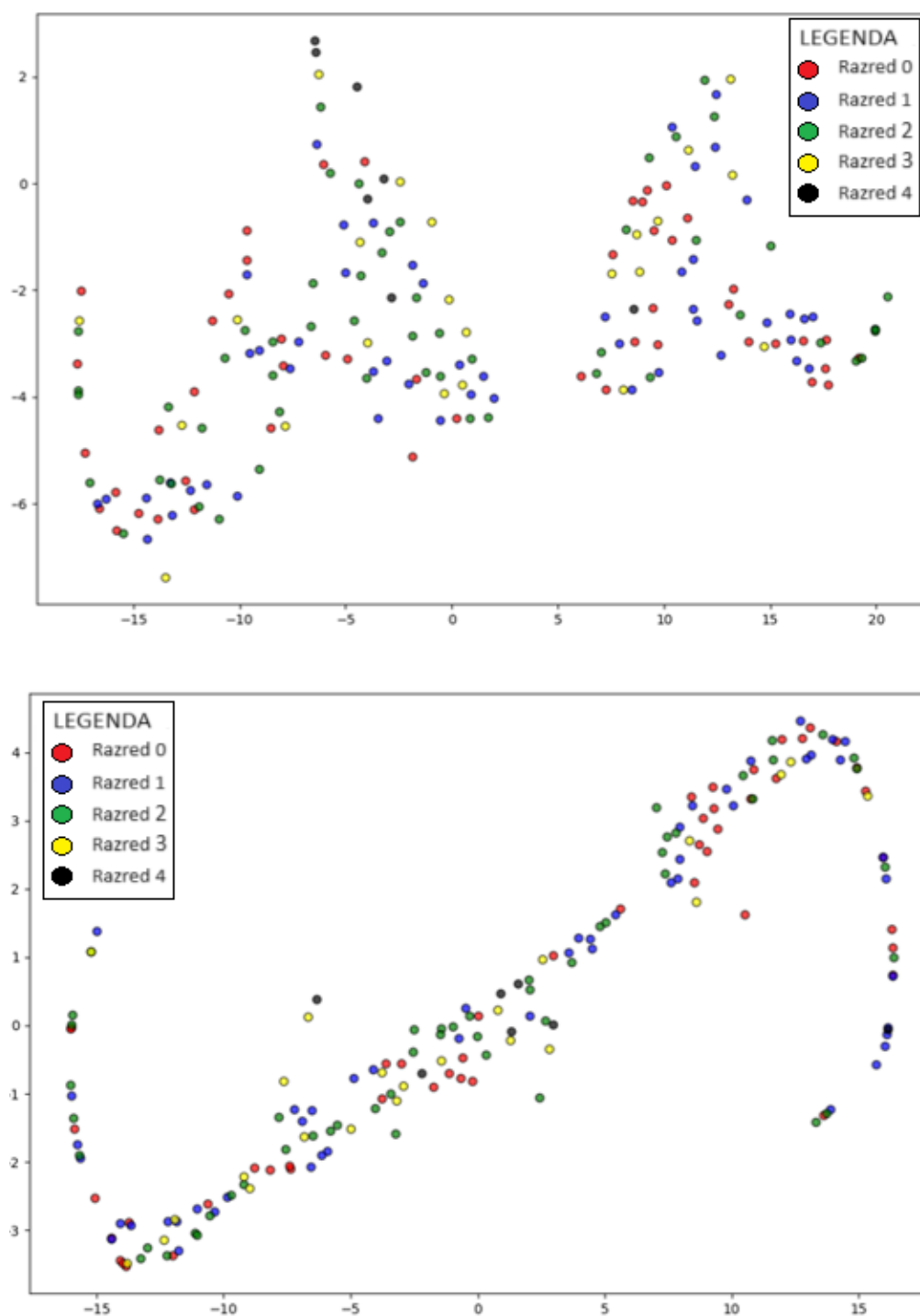
V poglavju so opisani rezultati pridobljeni z različnimi pristopi. V vlogi klasifikatorjev smo na pridobljenih značilkah preizkusili metodo podpornih vektorjev, naključne gozdove in večplastni perceptron. Neposredno na filtriranih časovnih podatkih smo trenirali tudi polno konvolucijsko nevronske mrežo. Za natančnejše rezultate smo uporabili 10-kratno navzkrižno preverjanje.

4.1 Rezultati

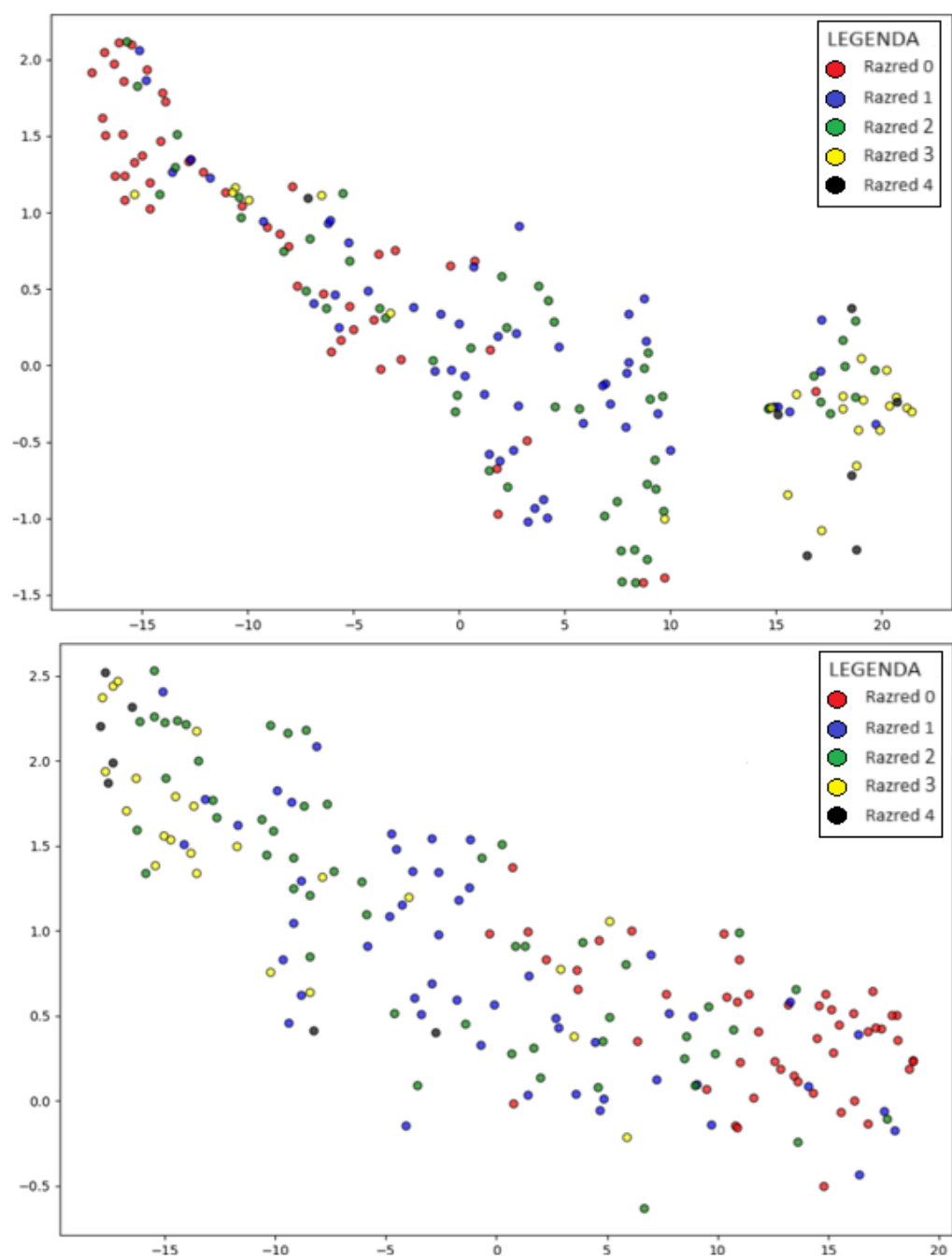
Za učenje in preverbo delovanja modela smo imeli na voljo 183 video posnetkov tapkanja, ki so bili ocenjeni v razrede od 0 do 4 po MDS-UPDRS. Število podatkov v posameznih razredih in njihovi deleži so vidni v tabeli 3.1. Da bi zagotovili večjo zanesljivost in objektivnost ocen uspešnosti modela, smo uporabili tehniko 10-kratnega navzkrižnega preverjanja. Ta pristop je pomemben, saj smo imeli na voljo relativno majhen nabor podatkov. Tako smo v vsakem od desetih korakov 90 % podatkov namenili za učenje modela, ostali 10 % pa za testiranje zmogljivosti modela. Tak pristop nam tudi preprečuje pretirano prilagajanje modela na učne podatke. Rezultati 10-kratnega navzkrižnega preverjanja so predstavljeni kot povprečje rezultatov vseh desetih korakov, vidni so v tabeli 4.3. Pri metodah 1 in 2 kjer smo naredili veliko časovno neodvisnih značilk, smo za izbiro najboljših upora-

bili metodo `SelectKBest` [52] iz knjižnice `scikit-learn` [53]. Slednji korak je pomemben, saj večje število značilk ne zagotavlja boljšega modela. Slabe značilke nam poslabšajo rezultate, ker ne nosijo veliko uporabnih informacij, v model pa vpeljejo višjo kompleksnost in povečanje časa procesiranja zaradi večih dimenzij podatkov. Značilke obeh metod smo tudi vizualizirali s `t-SNE`, primer je viden na sliki 4.1. Pri vizualizaciji podatkov z uporabo vseh značilk se podatki različnih razredov precej prekrivajo in je težko določiti centroide posameznih razredov. Na sliki 4.2, kjer so podatki predstavljeni s 15 najboljšimi značilkami, vidimo, da so pri obeh metodah prekrivanja razredov nekoliko manjša. Opazimo, da tukaj nekoliko lažje določimo centroide posameznih razredov. So pa prisotna prepletanja predvsem med sosednjimi razredi, npr. med 2 in 3. Pri izbiri najboljših značilk z metodo `SelectKBest` [52] smo kot cenilno funkcijo uporabili `f_classif` [55], ki temelji na ANOVA F-testu in `mutual_info_classif` [56], ki ocenjuje kakovost značilk na podlagi količine medsebojne informacije.

Metoda `SelectKBest` [52] nam vrne k najboljših značilk. Parameter k je potrebno določiti, zato smo za izbiro velikosti k uporabili vrednosti med 1 in številom značilk in kot končen rezultat vzeli le najboljše rezultate glede na k . Rezultati pridobljeni preko cenilne funkcije `f_classif` [55] so vidni v tabeli 4.1. Rezultate pridobljene s cenilno funkcijo `mutual_info_classif` [56] pa vidimo v tabeli 4.2. Ocena medsebojne informacije se je v splošnem odrezala bolje za obe metodi.



Slika 4.1: Vizualizacija podatkov s t-SNE v 2 dimenzijah. Podatki so predstavljeni z uporabo vseh značilk metode 1 zgoraj in metode 2 spodaj.



Slika 4.2: Vizualizacija podatkov s t-SNE v 2 dimenzijah. Podatki so predstavljeni z uporabo 15 najboljših značilk, izbranih s funkcijo `f_classif` [55], metode 1 zgoraj in metode 2 spodaj.

Izmed treh izbranih klasifikatorjev se je metoda podpornih vektorjev v povprečju odrezala najslabše. Je pa slednja najboljše delovala pri manjšem številu značilk kot ostala klasifikatorja. Večplastni perceptron je tipično uporabil precej večje število značilk za doseganje dobrih rezultatov. Najboljše rezultate za posamezno metodo in cenilno funkcijo smo dobili z uporabo naključnih gozdov oziroma večplastnega perceptrona, rezultati so vidni v tabelah 4.1 in 4.2.

Z metodo 1 smo dosegli nekoliko boljše rezultate kot z metodo 2. V najboljšem primeru smo z metodo 1 dosegli 61 % točnost in 0,62 F1, kar je za 1 % oziroma 0,07 več kot najboljši primer metode 2. Boljši rezultat pripisujemo večjemu naboru značilk, ki bolje opišejo naše podatke. Absolutno najboljši rezultat smo tako dosegli preko cenilne funkcije `mutual_info_classif` za [56] izbiro značilk metode 1 in uporabe večplastnega perceptrona v vlogi klasifikatorja. V metodi 1 in 2 je pri vseh klasifikatorjih bilo približno 83 % oziroma 73% od napačnih napovedi takšnih, ki so se od referenčne ocene razlikovale za natanko 1 razred.

Polna konvolucijka nevronska mreža, ki so jo predstavili Li in sodelavci [39] se je odrezala bolje kot uporaba ročno izluščenih značilk v klasifikatorjih. Ta nevronska mreža (vidna na sliki 3.5) je na časovni vrsti amplitud dosegla 72 % klasifikacijsko točnost in 0,62 F1 vrednost z uporabo 10-kratnega navzkrižnega preverjanja, ostale metrike so vidne v tabeli 4.3. V namen boljše klasifikacije smo nevronske mreže modificirali za dvodimenzionalen vhod, jo razširili ter ji dodali hitro zaustavljanje in sloje za izpuščanje. Na koncu smo obdržali najboljše uteži glede na točnost v testni množici. Po dopolnitvi mreže, slednja je vidna na sliki 3.6, smo opazili izboljšanje klasifikacijskih metrik. Dosegli smo 77 % točnost in 0,75 F1 vrednost, vsi rezultati so vidni v tabeli 4.3. Boljše delovanje je pričakovano in se lahko pripiše več dejavnikom. Prvič, zaradi integracije dodatnih podatkov, ki združujejo hitrosti in amplitude, je mreža zdaj opremljena z bogatejšim naborom informacij za procesiranje. Drugič, uporaba večjega števila slojev omogoča mreži, da se uči in prepozna bolj kompleksne vzorce in odvisnosti v podatkih. Te izboljšave

skupaj prispevajo k boljši zmogljivosti in natančnosti modela.

Model	Točnost %	F1	Natančnost %	Priklic %
MLP 2	54	0,52	64	55
SVM 2	54	0,48	67	54
RF 2	60	0,54	67	56
MLP 1	53	0,60	61	52
SVM 1	51	0,55	54	53
RF 1	58	0,60	61	52

Tabela 4.1: Primerjava rezultatov različnih klasifikatorjev pridobljenih z 10-kratnim navzkrižnim preverjanjem na značilkah dobljenih preko metode 1 in metode 2. Za izbiro najboljših značilk je bila uporabljena metoda Select-Kbest [52] s cenilno funkcijo **f_classif** [55]. Kratica RF označuje naključne gozdove, SVM metodo podpornih vektorjev, MLP pa večplastni perceptron. Število 1 oziroma 2 pri modelu označuje izbrano metodo gradnje značilk.

Model	Točnost %	F1	Natančnost %	Priklic %
MLP 2	54	0,52	63	53
SVM 2	60	0,54	67	57
RF 2	60	0,55	67	56
MLP 1	61	0,62	67	58
SVM 1	52	0,60	56	50
RF 1	57	0,61	62	56

Tabela 4.2: Primerjava rezultatov različnih klasifikatorjev pridobljenih z 10-kratnim navzkrižnim preverjanjem na značilkah dobljenih preko metode 1 in metode 2. Za izbiro najboljših značilk je bila uporabljena metoda SelectKbest s cenilno funkcijo **mutual_info_classif** [56]. Kratica RF označuje naključne gozdove, SVM metodo podpornih vektorjev, MLP pa večplastni perceptron. Število 1 oziroma 2 pri modelu označuje izbrano metodo gradnje značilk.

Model	Točnost %	F1	Natančnost %	Priklic %
FCN	72	0,62	84	63
FCN +	77	0,75	88	75
Metoda 1	61	0,62	67	58
Metoda 2	60	0,55	67	58

Tabela 4.3: Rezultati strojnega učenja pridobljeni z 10-kratnim navzkrižnim preverjanjem. Oznaka FCN je nevronska mreža na sliki 3.5. Oznaka FCN + označuje razširjeno mrežo vidno na sliki 3.6. Zraven sta še najboljša rezultata metode 1 in metode 2.

Poglavje 5

Sklepne ugotovitve

V poglavju so ovrednoteni glavni rezultati in odločitve, ki so nas pripeljale do rezultatov, vse od procesa zbiranja podatkov do gradnje in učenja klasifikatorjev. Na koncu poglavja so opisane možne izboljšave in alternativni pristopi.

5.1 Diskusija

Prvi korak pri gradnji sistema za klasifikacijo stopnje motorične prizadetosti pri Parkinsonovi bolezni iz videov tapkanja je zbiranje podatkov. V našem primeru so bili videi zbrani v različnih okoljih posneti s pametnimi telefoni. Ti posnetki so se razlikovali po osvetlitvi, hitrosti zajema slik na sekundo, ločljivosti, kotih snemanja in razdaljah med kamero in roko. To je dodatno otežilo razvoj klasifikatorja in vpeljalo potrebo po robustnem pristopu, vidnem na sliki 3.2. Za zagotavljanje višje klasifikacijske točnosti modelov bi rabili bolj kakovostne posnetke. Za pridobitev le teh bi potrebovali sistem, ki bi snemal roke na enotni oddaljenosti od kamere, vedno pod pravim kotom in pri visokem številu sličic na sekundo. S takšnim sistemom bi primarno izboljšali zaznavo skeleta z MediaPipe Hand in zmanjšali število slabih detekcij. Tako bi naši podatki bili manj šumni. Prav tako bi bile razdalje med prsti bolj pravilne, saj jih nagib kamere ne bi okvaril. Razdalje med

kamero in roko bi zato bile fiksne in ne bi bilo potrebe po normalizaciji. Iz skeleta roke bi lahko uporabili kote, namesto evklidskih razdalj med prsti, ker so bolj odporni na različne razdalje med kamero in roko. Kotov sicer nismo uporabili saj so veliko bolj občutljivi na nagibe med kamero in roko, takšnih posnetkov pa je v naši podatkovni zbirki zaradi načina snemanja veliko. Normalizacija podatkov je sicer v našem primeru potrebna, saj izniči učinek različnih razdalj. Vpelje pa nam dodatne težave, saj podatke, kjer pacient med tapkanjem prstov ni zmožen široko razpreti, skalira na interval, ki je enakovreden zdravim osebam. Slednje je problem predvsem pri pacientih, ocenjenih s stopnjo 4 po MDS-UPDRS, ki niso zmožni narediti niti enega celega cikla z maksimalnim razponom. Ker smo normalizacijo delali glede na maksimalen razpon med palcem in kazalcem so normalizirane amplitude posledično višje, kot je njihova dejanska vrednost in iz njih težje vidimo upad oziroma tapkanje z zelo majhno amplitudo.

Resolucija snemanja video posnetkov nam ni predstavljala težav, saj je bila dovolj velika, roke pa so posledično bile lepo vidne. Sicer smo preizkusili zviševanje resolucije, ki smo jo zvišali za 200 %, vendar je bil rezultat zanamrljiv. Časi procesiranja so se močno povečali, zato zviševanja resolucije nismo uporabili. Težave nam je tudi predstavljalo snemanje pri relativno nizkih sličicah na sekundo, naši posnetki so bili snemani pri 30 oziroma 60 sličicah na sekundo. Zato je prišlo do zamegljenosti zaradi gibanja roke, kar je poslabšalo prileganje skeleta MediaPipe Hand, ki je v določenih primerih slabo določil sredino prstov, zgrešil zaznavo konic prstov in občasno tudi zamenjal kazalec in sredinec. Zamegljenost zaradi gibanja smo poizkušali zmanjšati z uporabo namenskih nevronske mreže, vendar neuspešno. Slednje so v nekaterih sličicah uspešno zmanjšale zamegljenost, drugod pa so dodale artefakte. Takšen primer lahko vidimo na sliki 3.1. Ker je bil vpliv odstranjevanja zamegljenosti zaradi gibanja na klasifikacijo minimalen oziroma negativen, se teh metod v nadaljevanju nismo poslužili. Zaradi nizke frekvence zajema sličic na sekundo je časovna vrsta hitrosti in pospeškov premalo natančna. To bi lahko reševali z interpolacijo sličic, ki pa je nismo

uporabili saj smo mnenja, da bi umetno ustvarjeni podatki prikrili težave v tapkanju. Menimo, da bi zato zajem videov s profesionalno opremo, z možnostjo zajema vsaj 240 sličic na sekundo izločil zamegljenost prstov zaradi hitrega gibanja in omogočil natančnejše časovne vrste. Posledično bi bila zaznava roke boljša, značilke pridobljene iz hitrosti in tudi amplitud pa bi bile kvalitetnejše. V tapkanju gre za hitro gibanje, ki ga nismo zmožni natančno zajeti s frekvenco snemanja pri 30 ali 60 sličicah na sekundo. Takšna oprema bi izboljšala klasifikacijo, vendar takšni sistemi niso splošno dostopni in v skladu z našo tezo.

Iz obstoječih detektorjev rok (npr. MMPose Hand [69], OpenPose [70]) smo izbrali MediaPipe, ker gre za robusten model, ki se dobro obnese pod različnimi osvetlitvami, ozadji in koti med kamero in roko. Alternativno bi lahko sami zgradili model za detekcijo skeleta roke preko segmentacije z nevronske mreže v kombinaciji s sledenjem optičnega toka ali pa natrenirali kakšno namensko mrežo za delo s slikami, recimo Resnet 50 [71, 72] ali YOLOv7 [73, 74]. Za slednje se nismo odločili saj smo mnenja, da so obstoječi detektorji rok bolj dovršeni in preprostejši za integracijo.

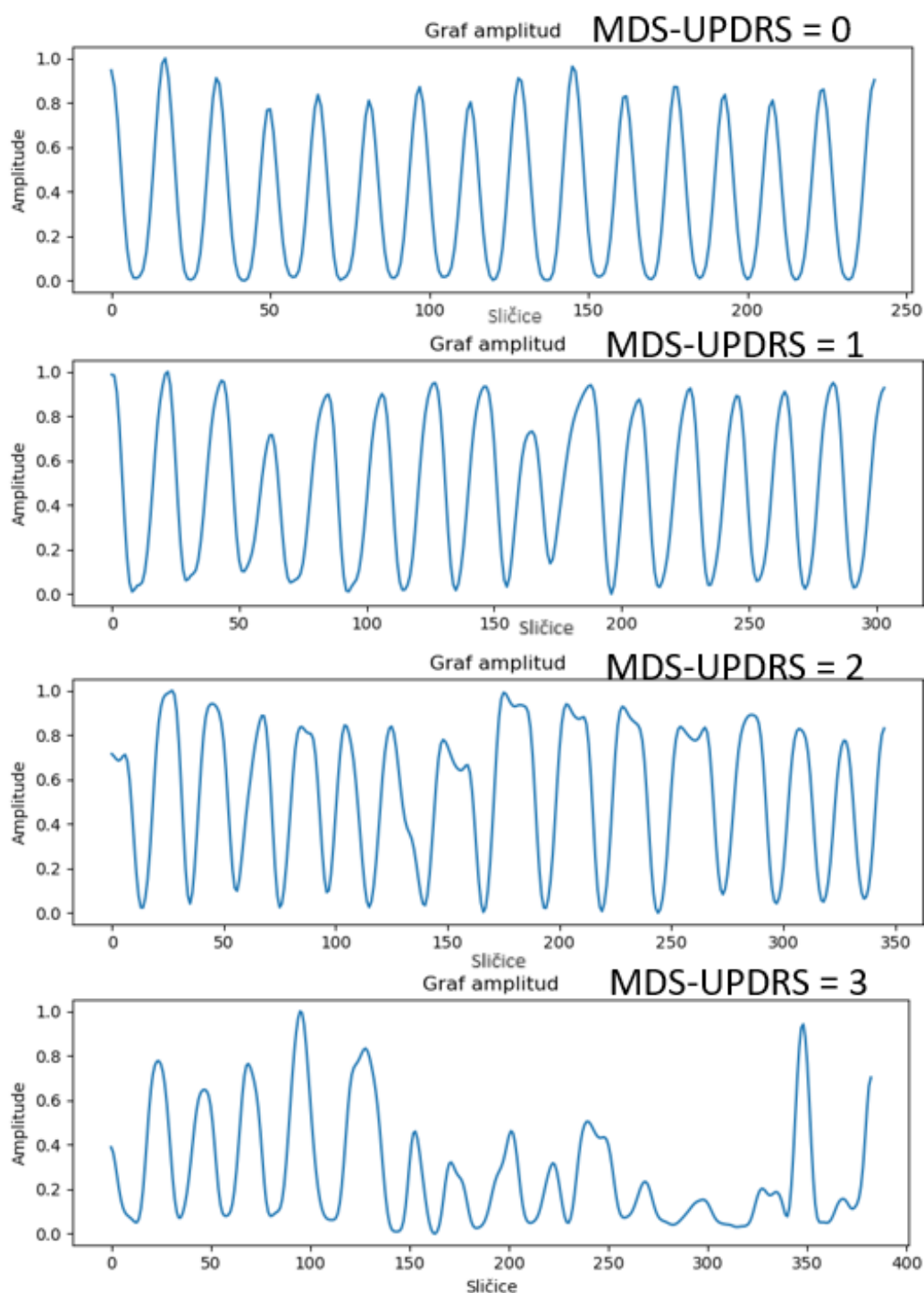
Nenatančne zaznave skeleta roke predstavljajo šum v podatkih, kateremu je potrebno prišteti še tremor. Tremor otežuje zaznavo bradikinezije, saj maskira tapkanje, sam pa ni del ocene tapkanja po lestvici MDS-UPDRS, vidne v tabeli 1.1. Za izločanje šuma v podatkih in tremorja smo uporabili digitalne filtre. Filtriranje nam je povzročalo težave, saj je popačilo določene signale tapkanja. Zato smo na koncu uporabili šibkejšo filtriranje, ki je šum uspelo odstraniti le do neke mere, je pa posledično signal manj popačilo. Za kombinacijo nizko pasovnega filtra in drsečega povprečja smo se odločili, saj ponuja dobro razmerje med glajenjem in popačitvijo signala. Razmišljali smo tudi o uporabi valovnih transformacij za odstranjevanje šuma (*angl. wavelet denoise*). Ta metoda sicer omogoča večjo prilagodljivost pri ločevanju šuma od koristnega signala na različnih frekvenčnih pasovih. Za njo se nismo odločili, saj ni tipičen nizkoprepustni filter in pri filtriranju lahko odstrani pomembne frekvence. Uporaba nizkoprepustnega filtra nam omogoča, da se

osredotočimo na odstranjevanje visokofrekvenčnega šuma, ki je tipično prisoten zaradi prileganja skeleta ali prisotnega tremorja, medtem ko drseče povprečje pomaga pri dodatnem glajenju signala brez večje izgube koristnih informacij. Tak pristop nam zagotavlja boljšo kontrolo nad ravnotežjem med odstranjevanjem šuma in ohranjanjem originalnih signalov. Konstantno prilagajanje frekvence odreza v nizkoprepustnem filtru za vsak video posebej omogoča prilagodljivost našega pristopa na heterogene podatke. Tako bolje filtriramo šum iz vsakega posnetka posebej, glede na njegove specifične lastnosti. To je še posebej pomembno v primerih, kjer standardizirani pristopi ne zagotavljajo zadovoljivih rezultatov zaradi raznolikosti v podatkih, kot so razlike v hitrosti tapkanja, amplitud in prisotnosti šuma različnih vrst.

Pri vizualizaciji značilk s t-SNE na slikah 4.1 in 4.2 lahko vidimo prepletenost razredov. Opazimo, da je prepletenost med sosednjimi razredi višja. Razred 1 in 3 sta veliko manj prepletena kot recimo razred 2 in 3, podobno velja za ostale sosednje razrede. Na sliki 5.1 lahko vidimo, da so razlike v amplitudah med sosednjimi razredi precej majhne. Zaradi majhnih razlik v amplitudah in hitrosti med sosednjimi razredi je za razlikovanje teh razredov potrebna natančna analiza časovnih vrst. Razlogov za podobnost med razredi je več. Prvič so razlike med sosednjimi razredi tapkanja majhne, posledično na končno oceno lahko vplivajo že relativno manjši upadi hitrosti oziroma amplitud. Pri uporabi metode 1 in 2 se je približno 83 % oziroma 73 % napačnih klasifikacij od referenčnih ocen tapkanj razlikovalo za natanko 1 razred. K manjšemu razlikovanju med razredi pripomoreta tudi normalizacija, ki zamaskira tapkanja z majhnimi amplitudami, in kot snemanja, ki lahko popači dejanske razdalje med prsti. Namreč tapkanje s konstantno majhno amplitudo ne pomeni, da udeleženec nima motoričnih težav, na grafih pa bodo podatki posledično izgledali kot, da jih nima. Eden izmed dejavnikov za prepletanje razredov je tudi možnost pojava napak v ocenah videov, saj so sosednji razredi precej prepleteni. Na voljo smo imeli le enega izkušenega ocenjevalca. Zaradi majhnih razlik med posameznimi razredi MDS-UPDRS se lahko hitro zgodi, da zaradi utrujenosti ocenjevalec napačno oceni stopnjo

motorične prizadetosti. Za zagotavljanje natančnosti ocen bi bilo potrebno imeti vsaj 3 izkušene ocenjevalce, ki bi podatke neodvisno ocenili, nato pa bi jih navzkrižno preverili. Takšne ocene bi omogočale natančnejše rezultate klasifikacije. Dodaten vzrok prepletenosti sosednjih razredov bi lahko bil, da izdelane značilke ne opišejo tapkanja dovolj nazorno. Zaradi zelo majhnih razlik med sosednjimi razredi z izdelavo časovno neodvisnih značilk le te zanemarimo oziroma izpustimo. Na sliki 4.2, kjer smo uporabili 15 najboljših značilk, lahko vidimo, da so razredi nekoliko manj prepleteni kot pri uporabi vseh značilk (slika 4.1), kar nakazuje, da so določene značilke neinformativne za ločevanje med razredi. Ker je t-SNE tehnika za zmanjševanje dimenzionalnosti podatkov, lahko vpelje napake v sami vizualizaciji. Kakovost vizualizacije je odvisna od izbire parametrov, če te izberemo slabo, dobimo zavajajoč prikaz podatkov. Pri vizualizaciji t-SNE izkrivi tudi globalno strukturo podatkov in določene skupine zgosti oziroma razprši, kar ustvari napačen vtis gostote podatkov.

Za natančno oceno modela bi potrebovali neodvisno testno množico, ki je model še ni videl. Namreč pri učenju klasifikatorja se dogaja, da ga ne hote optimiziramo glede na testno množico, ki pa ni nujno reprezentativna realnim podatkom. Zaradi majhne količine podatkov si tega nismo mogli privoščiti, smo pa zato uporabili 10-kratno navzkrižno preverjanje. S tem smo zagotovili realnejše ocene modela in omilili pretirano prilagajanje glede na testno množico saj je ta za vsak korak navzkrižnega preverjanja drugačna. Metoda 10-kratnega navzkrižnega preverjanja nam zmanjšuje varianco meritev, ki nastane zaradi delitve podatkov na testno in učno množico. Model se uči na učnem naboru in se preveri na testnem naboru, kar zagotavlja, da se vsak vzorec iz izvirnega nabora podatkov 1-krat pojavi v testnem naboru in 9-krat v učnem naboru. V kategoriji video posnetkov označenih s stopnjo 4 po lestvici MDS-UPDRS bi potrebovali več podatkov. Namreč naša baza je štela le 7 posnetkov kategorije 4, kar je premalo za dobro učenje in klasifikacijo.



Slika 5.1: Grafi amplitud tapkanja za razrede MDS-UPDRS od 0 do 3. Grafi vsebujejo največ 15 ciklov. Razlike med sosednjimi razredi znajo biti majhne.

Večja količina značilke ne vodi nujno do izboljšanja modela saj slabe značilke poslabšajo model. To lahko vidimo na sliki 4.2, kjer z izbiro petnajstih najboljših značilk dosežemo nekoliko večje razlikovanje med razredi kot z vsemi značilkami vidnimi na sliki 4.1. Za izbiro najbolj relevantnih značilk smo uporabili metodo SelectKBest [52] in v njej preizkusili dve različni cenilni funkciji. Izbira najboljših značilk preko ocene medsebojne informacije se je odrazila v boljših rezultatih kot ANOVA F-test (f.classif [55]). ANOVA F-test je primeren le za numerične podatke in je računsko manj zahteven. Metoda izbire značilk na podlagi ocene medsebojne informacije (mutual.info.classif [56]) je primerna za izbiro numeričnih in tudi kategoričnih značilk, ki pa jih sicer mi nimamo. Ocena medsebojne informacije za izbiro značilk se je odrezala bolje v obeh metodah izgradnje značilk ne glede na izbran klasifikator, kar lahko vidimo v tabelah 4.1 in 4.2. Menimo, da zato, ker zajema tudi nelinearne odvisnosti v podatkih in je neodvisna od porazdelitev podatkov. Namreč v ANOVA F-testu se predpostavlja normalna porazdelitev podatkov. Časi procesiranja pri uporabi ocene medsebojne informacije so se povečali zaradi višje računske zahtevnosti metode.

Absolutno najboljši rezultat glede na vse 4 metrike na izdelanih značilkah smo dosegli z uporabo metode 1, rezultat je viden v tabeli 4.2. Za ostale klasifikatorje in cenilne funkcije je metoda 2 konstantno delovala slabše glede na F1 metriko, se pa je v poprečju odrezala malenkost bolje v točnosti, natančnosti in priklicu. Zaradi uporabe makro metrik to nakazuje, da je metoda 1 delovala bolj uravnoteženo skozi vse razrede, kot metoda 2, ki je v določenem razredu dosegla slabo uspešnost. Najboljši rezultat je bil dosežen z uporabo večplastnega perceptrona, ki je v metodi SelectKBest [52] potreboval precej višje število značilk kot metoda podpornih vektorjev in naključni gozdovi. Sklepamo, da je to posledica zmožnosti generiranja kompleksnih nelinearnih vzorcev, ki jo omogoča arhitektura modela. To nakazuje, da MLP bolje izkoristi večji nabor značilk za izgradnjo natančnejšega modela, medtem ko sta metoda podpornih vektorjev in naključni gozdovi bolj omejena ali preobčutljiva na povečanje števila značilk, kar vodi do pretiranega prilaga-

janja ali nezmožnosti prilagajanja tako kompleksnih funkcij. Torej metoda 1 se je odrezala bolje kot metoda 2. Tudi v napačno ocenjenih napovedih je bilo pri metodi 2 približno 10 % več takšnih, ki so se od referenčne vrednosti razlikovale za več kot 1 razred.

V okviru metode 3, kjer smo uporabili polno konvolucijsko nevronske mreže (FCN), smo opazili znatno izboljšanje rezultatov. Dosegli smo 16 % višjo točnost in za 0,13 višjo F1 metriko, rezultati so vidni v tabeli 4.3. To izboljšanje pripisujemo značilnosti nevronske mreže, da se uči in identificirajo zapletene vzorce znotraj podatkov. Zmožnost prileganja kompleksnih funkcij, je pomembna, saj se soočamo s podatki visoke dimenzionalnosti in kompleksnosti. Izboljšanje rezultatov pripisujemo še sposobnosti polne konvolucijske nevronske mreže, da izlušči tako lokalne kot globalne odvisnosti iz časovnih vrst. To se doseže s konvolucijskimi filtri različnih velikosti, ki mreži omogočajo, da zajame in interpretira informacije na več lestvicah. Nasprotno, tradicionalni pristopi, ki se zanašajo na ročno izdelane časovno neodvisne značilke, pogosto spregledajo te povezave, kar vodi v manj natančno klasifikacijo. Zajem globalnih odvisnosti je pomemben, npr. upad hitrosti se ne zgodi v trenutku ampak se dogaja skozi več sličic. Hkrati je z izdelavo časovno neodvisnih značilk težko zajeti vse pomembne lokalne, globalne in ostale odvisnosti v podatkih. Zaradi teh značilnosti se je FCN izkazal superiorno v zmogljivosti analize tapkanja na račun daljšega učenja in potrebe po zmogljivejšem sistemu.

Naših klasifikacijskih metrik ne moremo direktno primerjati s sorodnimi deli, saj standardizirana podatkovna zbirka za testiranje zmogljivosti klasifikatorjev tapkanja ne obstaja. Baze podatkov, ki se uporabljajo v sorodnih delih pa so med seboj preveč različne, saj vsebujejo različno število podatkov. Hkrati je pomembno tudi število podatkov v vsakem razredu. Nekatera sorodna dela uporabljajo tudi sisteme kamer in ne le 1 kamere za pridobitev podatkov, zato ti niso enakovredni našim. Vseeno pa bomo v nadaljevanju razložili in argumentirali kako se naše delo primerja s sorodnimi deli glede na izbrano metodologijo.

Khan in sodelavci [36] so z uporabo SVM in ročno izdelanih značilk na bazi 471 videov tapkanja dosegli 88 % točnost pri napovedovanju stopnje motorične prizadetosti. V primerjavi z našim delom je to precej višji rezultat. Sklepamo, da je višji rezultat posledica homogene podatkovne zbirke. Namreč vseh 471 videov tapkanja je bilo zajetih le od 19 različnih oseb. Takšna homogena zbirka ne zajame različnih situacij in motenj tapkanja, model tako ne bo dobro generaliziral saj se bo naučil le kako tapkajo posamezniki, ne pa kako tapka splošna populacija. Drug razlog višje natančnosti je tudi, da so klasificirali le v 4 od 5 razredov lestvice, namreč njihova baza podatkov razreda 4 ni vsebovala.

Islam in sodelavci [42] so dosegli približno 52 % klasifikacijsko točnost klasifikacije v 5 razredov MDS-UPDRS. Eden izmed razlogov za nižjo točnost v primerjavi s sorodnimi deli bi lahko bila večja in bolj heterogena podatkovna zbirka. Tapkanja znotraj istega razreda se lahko precej razlikujejo. Njihova podatkovna zbirka je štela 489 videov tapkanja, pridobljenih od 250 različnih oseb. Potrebno je še opozoriti, da je razred z oceno 4 po MDS-UPDRS štel le 5 primerov. Za namen izgradnje modela so naredili 65 značilk in za učenje uporabili le 23 najboljših. V primerjavi z našim delom kjer smo uporabili širši nabor značilk in dosegli višjo točnost, je mogoče, da njihov nabor značilk ni tako celovito opisal podatkovne zbirke.

Yu in sodelavci [43] so imeli na voljo bazo 75 videov posnetkov tapkanja, ki so bili snemani z večimi kamerami, da čim boljše zajamejo pravi kot. Na testni množici 15 posnetkov so dosegli 80 % točnost in priklic 79 %. Testna baza je relativno majhna, še posebej če upoštevamo, da imamo v podatkih 5 razredov. Uporabili so tudi interpolacijo sličic, kar ni dobra ideja saj z umetno ustvarjenimi podatki izgubimo informacije o tapkanju oziroma dodamo nerealne podatke. To lahko negativno vpliva na zanesljivost klasifikacije. Slednjo so izvedli z izgradnjo odločitvenega drevesa. Listi njihovega drevesa so imeli med 1 in 7 vzorcev, večina pa 2 do 3 vzorce. Listi z malo vzorci nakazujejo na pretirano prilagajanje, visoko varianco in pomanjkanje generalizacije, kar vodi v nezanesljive napovedi in slabo učinkovitost modela pri

splošni uporabi na novih podatkih.

Nevronska mreža, ki so jo predstavili Li in sodelavci [39] je na naših podatkih dosegla 72 % točnost v primerjavi z 80 % točnostjo, ki jo navajajo avtorji. Sklepamo, da je razlog nekoliko slabše klasifikacije naših podatkov v kompleksnosti in heterogenosti le teh, manjši podatkovni bazi in uporabi 10-kratnega navzkrižnega preverjanja napram 5-kratnemu. Li in sodelavci so imeli na voljo več kot 40 % večjo bazo podatkov, ki pa je bila tudi precej neuravnotežena, namreč večina podatkov je pripadala razredom 0, 1 in 2, razreda 3 in 4 pa sta bila majhna. Podobna je tudi razporeditev naših podatkov po razredih, vidna v tabeli 3.1. Pomembno je še izpostaviti, da konvolucijske nevronske mreže potrebujejo večje količine podatkov za učinkovito učenje. Naša baza podatkov je bila razmeroma majhna, kar lahko pojasni nekoliko slabše rezultate. Smo pa z dopolnitvijo nevronske mreže izboljšali rezultate napovedi in dosegli 77 % točnost. Z dodajo dodatnih konvolucijskih slojev smo mreži omogočili prilagajanje kompleksnejših funkcij, s sloji za izpuščanje pa omejili pretirano prilagajanje. Hkrati je mreža kot vhod dobila amplitude in hitrosti, kar omogoča izluščitev dodatnih odvisnosti med posameznimi videi.

Če primerjamo ročno oblikovane značilke v kombinaciji s klasifikacijskimi modeli in polno konvolucijsko nevronske mrežo, vidimo, da se je slednje odrezalo precej bolje. Z nevronske mreže se tudi izognemo koraku oblikovanja značilk, ki zahteva dobro poznavanje problema. Nevronske mreže so nerazumljive, saj jim kot vhod podamo časovne podatke, iz katerih avtomatsko izluščijo odvisnosti, ki so nam neznane. Nevronske mreže so se sicer na večih področjih izkazala za superiorne. Tipično pa potrebujejo zmogljivejšo strojno opremo. Izdelano orodje za vizualizacijo tapkanja in izpis ključnih podatkov nam omogoča nazoren vpogled v tapkanje. Omogoča nam prikaz skeleta skozi vsako sličico videa in drsenje skozi posamezne sličice videa. Predstavi nam tudi zamrznitve in oklevanja na grafih amplitud in hitrosti ter označuje trenutno pozicijo na njih, kar je v pomoč pri določanju stopnje motorične prizadetosti iz testa tapkanja.

5.2 Nadaljnje delo

V prihodnje bi dopolnili bazo podatkov. V kategoriji oseb z oceno 4 smo imeli le 7 video posnetkov. Tako bi lahko del podatkov namenili za neodvisno nevideno testno množico. S tem bi bolj objektivno ovrednotili klasifikator. Skrite testne množice trenutno zaradi majhnega števila podatkov nismo mogli narediti. Pri ocenjevanju zmogljivosti klasifikatorjev upoštevamo le napovedi, ki se popolnoma ujemajo z referenčno vrednostjo. V praksi pogosto prihaja do razlike v oceni za 1 razred, to je posledica subjektivne narave lestvice MDS-UPDRS, kot tudi utrujenosti ocenjevalca in podobnosti med nekaterimi sosednjimi razredi. Zato bi bilo pri ocenjevanju zmogljivosti klasifikatorjev smiselno upoštevati napovedi, ki se razlikujejo za 1 razred. To bi dosegli z uporabo uteži nastavljenih glede na razlikovanje napovedi od referenčnih vrednosti razredov. Preizkusili bi še kakšen drug model detektorja roke in primerjali rezultate klasifikacije. Tako bi videli, če se drugi detektorji roke odrežejo bolje.

Zanimivo bi bilo preizkusiti še druge modele nevronske mreže za procesiranje časovnih vrst, saj bi tako dobili boljšo sliko, kako se ti primerjajo s klasifikacijo iz ročno izluščenih značilnic in klasifikacijo dopolnjene polne konvolucijske nevronske mreže na sliki 3.6. Smiselno bi bilo uporabiti tudi grafovske nevronske mreže, ki bi kot vhod prejele celoten skelet roke. V delu večinoma uporabljamo le evklidske razdalje med konicama palca in kazalca, v primeru metode 1, ko smo uporabili še dodatne točke skeleta, pa izgubimo relacije med njimi. Takšne nevronske mreže tudi same izluščijo odvisnosti v podatkih oziroma značilke, ki jih je ročno težko optimalno izbrati. V orodje za vizualizacijo tapkanja bi dodali še oznake na mestu, kjer sistem oceni, da pride do upada hitrosti oziroma amplitud, saj bi te bile v pomoč pri ocenjevanju tapkanja. Orodje bi pripravili tudi za spletno uporabo.

Literatura

- [1] C. G. Goetz, B. C. Tilley, S. R. Shaftman, G. T. Stebbins, S. Fahn, P. Martinez-Martin, W. Poewe, C. Sampaio, M. B. Stern, R. Dodel, B. Dubois, R. Holloway, J. Jankovic, J. Kulisevsky, A. E. Lang, A. Lees, S. Leurgans, P. A. LeWitt, D. Nyenhuis, C. W. Olanow, O. Rascol, A. Schrag, J. A. Teresi, J. J. van Hilten, N. LaPelle, Movement disorder society-sponsored revision of the unified parkinson's disease rating scale (mds-updrs): scale presentation and clinimetric testing results, *Movement disorders: official journal of the Movement Disorder Society* 23 (15) (2008) 2129–2170.
- [2] C. Schulte, T. Gasser, Genetic basis of parkinson's disease: inheritance, penetrance, and expression, *The Application of Clinical Genetics* 4 (2011) 67–80. doi:10.2147/TACG.S11639.
- [3] Pomen zgodnjega prepoznavanja in zdravljenja parkinsonove bolezni ter njene destigmatizacije, <https://www.nijz.si/sl/pomen-zgodnjega-prepoznavanja-in-zdravljenja-parkinsonove-%25bolezn-ter-n-jene-destigmatizacije>, dostopano: 3-3-2024 (2019-04-11).
- [4] L. V. Kalia, A. E. Lang, Parkinson's disease, *The Lancet* 386 (2015). doi:10.1016/S0140-6736(14)61393-3.
- [5] S. Bandres-Ciga, M. Diez-Fairen, J. J. Kim, A. B. Singleton, Genetics of parkinson's disease: An introspection of its journey towards precision medicine, *Neurobiology of Disease* 137 (2020) 104782. doi:https://doi.org/10.1016/j.nbd.2020.104782.

-
- [6] The Michael J. Fox Foundation, Parkinson's genetics, <https://www.michaeljfox.org/news/parkinsons-genetics>, accessed: (29.1.2024) (2024).
- [7] C. Blauwendraat, M. A. Nalls, A. B. Singleton, The genetic architecture of parkinson's disease, *Lancet Neurol* 19 (2) (2019) 170–178.
- [8] E. Uffelmann, Q. Q. Huang, N. S. Munung, J. de Vries, Y. Okada, A. R. Martin, H. C. Martin, T. Lappalainen, D. Posthuma, Genome-wide association studies, *Nature Reviews Methods Primers* 1 (2021) 1–21. doi:10.1038/s43586-021-00056-9.
- [9] B. R. Bloem, M. S. Okun, C. Klein, Parkinson's disease, *The Lancet* 397 (10291) (2021) 2284–2303. doi:[https://doi.org/10.1016/S0140-6736\(21\)00218-X](https://doi.org/10.1016/S0140-6736(21)00218-X).
URL <https://www.sciencedirect.com/science/article/pii/S014067362100218X>
- [10] B. R. D. Miranda, S. M. Goldman, G. W. Miller, J. T. Greenamyre, E. R. Dorsey, Preventing parkinson's disease: An environmental agenda, *J. Parkinsons. Dis.* 12 (1) (2022) 45–68.
- [11] A. Keshavarzian, S. J. Green, P. A. Engen, R. M. Voigt, A. Naqib, C. B. Forsyth, E. Mutlu, K. M. Shannon, Colonic bacterial composition in parkinson's disease, *Mov. Disord.* 30 (10) (2015) 1351–1360.
- [12] F. Scheperjans, V. Aho, P. A. B. Pereira, K. Koskinen, L. Paulin, E. Pekkonen, E. Haapaniemi, S. Kaakkola, J. Eerola-Rautio, M. Pohja, E. Kinnunen, K. Murros, P. Auvinen, Gut microbiota are related to parkinson's disease and clinical phenotype, *Mov. Disord.* 30 (3) (2015) 350–358.
- [13] G. Franco, G. Lazzeri, A. Di Fonzo, Parkinsonism and ataxia, *Journal of the Neurological Sciences* 433 (2022) 120020. doi:<https://doi.org/10.1016/j.jns.2021.120020>.

URL <https://www.sciencedirect.com/science/article/pii/S002510X21027167>

- [14] S. Sveinbjornsdottir, The clinical symptoms of parkinson's disease, *Journal of Neurochemistry* 139 (S1) (2016) 318–324. doi:<https://doi.org/10.1111/jnc.13691>.
- [15] M. Petrou, B. A. Dwamena, B. R. Foerster, M. P. MacEachern, N. I. Bohnen, M. L. Müller, R. L. Albin, K. A. Frey, Amyloid deposition in parkinson's disease and cognitive impairment: A systematic review, *Movement Disorders* 30 (7) (2015) 928–935. arXiv:<https://movementdisorders.onlinelibrary.wiley.com/doi/pdf/10.1002/mds.26191>, doi:<https://doi.org/10.1002/mds.26191>.
URL <https://movementdisorders.onlinelibrary.wiley.com/doi/abs/10.1002/mds.26191>
- [16] A. L. Benabid, Deep brain stimulation for parkinson's disease, *Current Opinion in Neurobiology* 13 (6) (2003) 696–706. doi:<https://doi.org/10.1016/j.conb.2003.11.001>.
URL <https://www.sciencedirect.com/science/article/pii/S0959438803001739>
- [17] D. Bega, C. Zadikoff, Complementary & alternative management of parkinson's disease: An evidence-based review of eastern influenced practices, *Journal of Movement Disorders* 7 (2014) 57 – 66.
URL <https://api.semanticscholar.org/CorpusID:18357111>
- [18] F. Zupančič, G. Žagar, D. Georgiev, J. Žabkar, Spletna aplikacija za analizo tapkanja s prsti, *Proceedings of the 8th Student Computing Research Symposium (SCORES'22)* (1) (2022) 41–44. doi:[10.51939/scores22](https://doi.org/10.51939/scores22).
- [19] N. P. S. Bajaj, V. Gontu, J. Birchall, J. Patterson, D. G. Grosset, A. J. Lees, Accuracy of clinical diagnosis in tremulous parkinsonian patients:

- a blinded video study, *Journal of Neurology, Neurosurgery & Psychiatry* 81 (11) (2010) 1223–1228. doi:10.1136/jnnp.2009.193391.
- [20] J. Stamatakis, J. Ambroise, J. Crémers, H. Sharei, V. Delvaux, B. Macq, G. Garraux, Finger tapping clinimetric score prediction in parkinson's disease using low-cost accelerometers, *Comput. Intell. Neurosci.* 2013 (2013) 717853.
- [21] A. Y. Meigal, S. M. Rissanen, M. P. Tarvainen, S. D. Georgiadis, P. A. Karjalainen, O. Airaksinen, M. Kankaanpää, Linear and nonlinear tremor acceleration characteristics in patients with parkinson's disease, *Physiological measurement* 33 (3) (2012) 395.
- [22] J. P. Giuffrida, D. E. Riley, B. N. Maddux, D. A. Heldman, Clinically deployable kinesi technology for automated tremor assessment, *Movement Disorders* 24 (5) (2009) 723–730. doi:10.1002/mds.22445.
- [23] A. Salarian, H. Russmann, C. Wider, P. R. Burkhard, F. J. G. Vingerhoets, K. Aminian, Quantification of tremor and bradykinesia in parkinson's disease using a novel ambulatory monitoring system, *IEEE Transactions on Biomedical Engineering* 54 (2) (2007) 313–322. doi:10.1109/TBME.2006.886670.
- [24] C. N. Riviere, S. G. Reich, N. V. Thakor, Adaptive fourier modeling for quantification of tremor, *Journal of Neuroscience Methods* 74 (1) (1997) 77–87. doi:10.1016/S0165-0270(97)02263-2.
- [25] M. M. Sturman, D. E. Vaillancourt, D. M. Corcos, Effects of aging on the regularity of physiological tremor, *Journal of neurophysiology* 93 (6) (2005) 3064–3074.
- [26] H. Jeon, W. Lee, H. Park, H. J. Lee, S. K. Kim, H. B. Kim, B. Jeon, K. S. Park, Automatic classification of tremor severity in parkinson's disease using a wearable device, *Sensors (Basel)* 17 (9) (Sep. 2017).

-
- [27] O. Martinez-Manzanera, E. Roosma, M. Beudel, R. W. K. Borgemeester, T. van Laar, N. M. Maurits, A method for automatic and objective scoring of bradykinesia using orientation sensors and classification algorithms, *IEEE Transactions on Biomedical Engineering* 63 (5) (2015) 1016–1024.
- [28] G. Rigas, A. T. Tzallas, M. G. Tsipouras, P. Bougia, E. E. Tripoliti, D. Baga, D. I. Fotiadis, S. G. Tsouli, S. Konitsiotis, Assessment of tremor activity in the parkinson's disease using a set of wearable sensors, *IEEE Transactions on Information Technology in Biomedicine* 16 (3) (2012) 478–487. doi:10.1109/TITB.2011.2182616.
- [29] C. Lainscsek, P. Rowat, L. Schettino, D. Lee, D. Song, C. Letellier, H. Poizner, Finger tapping movements of parkinson's disease patients automatically rated using nonlinear delay differential equations, *Chaos: An Interdisciplinary Journal of Nonlinear Science* 22 (1) (2012). doi:10.1063/1.3683444.
- [30] A. Sadikov, J. Žabkar, M. Možina, V. Groznik, D. Nyholm, M. Memedi, Feasibility of spirography features for objective assessment of motor symptoms in parkinson's disease, in: J. H. Holmes, R. Bellazzi, L. Sacchi, N. Peek (Eds.), *Artificial Intelligence in Medicine*, Springer International Publishing, Cham, 2015, pp. 267–276.
- [31] J. Westin, M. Dougherty, D. Nyholm, T. Torgny Grothh, A home environment test battery for status assessment in patients with advanced parkinson's disease, *Computer Methods and Programs in Biomedicine* 98 (1) (2010) 27–35.
- [32] W. R. Adams, High-accuracy detection of early parkinson's disease using multiple characteristics of finger movement while typing, *PLOS ONE* 12 (11) (2017) 1–20. doi:10.1371/journal.pone.0188226.

- [33] A. B. Peachap, D. Tchiotsop, V. Louis-Dorr, D. Wolf, Detection of early parkinson's disease with wavelet features using finger typing movements on a keyboard, *SN Applied Sciences* 2 (10) (2020).
- [34] D. Iakovakis, S. Hadjidimitriou, V. Charisis, S. Bostantzopoulou, Z. Katsarou, L. J. Hadjileontiadis, Touchscreen typing-pattern analysis for detecting fine motor skills decline in early-stage parkinson's disease, *Scientific Reports* 8 (1) (2018).
- [35] D. Iakovakis, S. Hadjidimitriou, V. Charisis, S. Bostantzopoulou, Z. Katsarou, L. Klingelhoefer, H. Reichmann, S. B. Dias, J. A. Diniz, D. Trivedi, K. R. Chaudhuri, L. J. Hadjileontiadis, Motor impairment estimates via touchscreen typing dynamics toward parkinson's disease detection from data harvested in-the-wild, *Frontiers ICT* 5 (2018).
- [36] T. Khan, D. Nyholm, J. Westin, M. Dougherty, A computer vision framework for finger-tapping evaluation in parkinson's disease, *Artificial Intelligence in Medicine* 60 (2013) 27–60. doi:10.1016/j.artmed.2013.11.004.
- [37] S. Williams, S. D. Relton, H. Fang, J. Alty, R. Qahwaji, C. D. Graham, D. C. Wong, Supervised classification of bradykinesia in parkinson's disease from smartphone videos, *Artificial Intelligence in Medicine* 110 (2020) 101966. doi:https://doi.org/10.1016/j.artmed.2020.101966.
- [38] Mediapipe hands, dostopano: 30-11-2023 (11-04-2019).
URL <https://google.github.io/mediapipe/solutions/hands.html>
- [39] Z. Li, lu kang, M. Cai, X. Liu, Y. Wang, J. Yang, An automatic evaluation method for parkinson's dyskinesia using finger tapping video for small samples, *Journal of Medical and Biological Engineering* 42 (3) (2022) 351–363. doi:10.1007/s40846-022-00701-y.

-
- [40] A. Nunes, N. Kozhemiako, C. Stephen, J. Schmahmann, S. Khan, A. S. Gupta, Automatic classification and severity estimation of ataxia from finger tapping videos, *Frontiers in Neurology* 12 (2022).
- [41] J. Jakubowski, A. P. Chromik, J. Chmielinska, M. Nojszewska, A. K. Pruszczyk, Application of imaging techniques to objectify the finger tapping test used in the diagnosis of parkinson's disease, *Bulletin of the Polish Academy of Sciences. Technical Sciences* 71 (2023) art. no. e144886. doi:10.24425/bpasts.2023.144886.
- [42] M. S. Islam, W. Rahman, A. Abdelkader, S. Lee, P. T. Yang, J. L. Purks, J. L. Adams, R. B. Schneider, E. R. Dorsey, E. Hoque, Using ai to measure parkinson's disease severity at home, *npj Digital Medicine* 6 (156) (2023). doi:10.1038/s41746-023-00905-9.
- [43] T. Yu, K. W. Park, M. J. McKeown, Z. J. Wang, Clinically informed automated assessment of finger tapping videos in parkinson's disease, *Sensors* 23 (22) (2023). doi:10.3390/s23229149.
- [44] M. Z. U. Alam, M. S. Islam, E. Hoque, M. S. Rahman, Pulsar: Graph based positive unlabeled learning with multi stream adaptive convolutions for parkinson's disease recognition (2023). doi:10.48550/ARXIV.2312.05780.
- [45] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C. L. Chang, M. G. Yong, J. Lee, W. T. Chang, W. Hua, M. Georg, M. Grundmann, Mediapipe: A framework for building perception pipelines, *ArXiv abs/1906.08172* (2019). doi:https://doi.org/10.48550/arXiv.1906.08172Focustolearnmore.
- [46] M. Marais, D. Brown, J. Connan, A. Bobby, An evaluation of hand-based algorithms for sign language recognition, in: *2022 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*, 2022, pp. 1–6. doi:10.1109/icABCD54961.2022.9856310.

-
- [47] B. Kuznetsov, J. D. Parker, F. Esqueda, Differentiable iir filters for machine learning applications, in: Proc. Int. Conf. Digital Audio Effects (eDAFx-20), 2020, pp. 297–303.
- [48] S. Butterworth, On the theory of filter amplifiers, *Experimental Wireless and the Wireless Engineer* 7 (1930) 536–541.
- [49] R. G. Mello, L. F. Oliveira, J. Nadal, Digital butterworth filter for subtracting noise from low magnitude surface electromyogram, *Computer Methods and Programs in Biomedicine* 87 (1) (2007) 28–35. doi:<https://doi.org/10.1016/j.cmpb.2007.04.004>.
- [50] L. van der Maaten, G. Hinton, Visualizing data using t-sne, *Journal of Machine Learning Research* 9 (nov) (2008) 2579–2605, pagination: 27.
- [51] scikit learn, Tsne, dostopano: 20-02-2024.
URL <https://scikit-learn.org/stable/modules/generated/sklearn.manifold.TSNE.html>
- [52] scikit learn, Select k best, dostopano: 28-02-2024.
URL https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.SelectKBest.html
- [53] scikit learn, scikit-learn, dostopano: 20-05-2024.
URL <https://scikit-learn.org/stable/>
- [54] D. Kavya, Optimizing performance: Selectkbest for efficient feature selection in machine learning, dostopano: 7-3-2024 (2023).
URL <https://medium.com/@Kavya2099/optimizing-performance-selectkbest-for-efficient-feature-selection-in-machine-learning-3b635905ed48#:~:text=SelectKBest%20uses%20statistical%20tests%20like,in%20the%20final%20feature%20subset.>
- [55] scikit learn, $f_{classif}$, dostopano: 20-05-2024.
URL <https://scikit-learn.org/stable/modules/generated/sklearn>

`arn.feature_selection.f_classif.html#sklearn.feature_selection.f_classif`

- [56] scikit learn, *mutual_info_classif*, dostopano: 20-05-2024.
URL https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.mutual_info_classif.html#sklearn.feature_selection.mutual_info_classif
- [57] scikit learn, Decision trees, dostopano: 7-3-2024.
URL <https://scikit-learn.org/stable/modules/tree.html>
- [58] scikit learn, Random forest classifier, dostopano: 7-3-2024.
URL <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>
- [59] scikit learn, Neural network models, dostopano: 7-3-2024.
URL https://scikit-learn.org/stable/modules/neural_networks_supervised.html
- [60] A. Taherkhani, G. Cosma, T. M. McGinnity, A deep convolutional neural network for time series classification with intermediate targets, *SN Computer Science* 4 (6) (2023). doi:10.1007/s42979-023-02159-4.
- [61] Z. Wang, W. Yan, T. Oates, Time series classification from scratch with deep neural networks: A strong baseline, 2017, pp. 1578–1585. doi:10.1109/IJCNN.2017.7966039.
- [62] J. Wang, S. Li, W. Ji, T. Jiang, B. Song, A t-cnn time series classification method based on gram matrix, *Scientific Reports* 12 (1) (2022). doi:10.1038/s41598-022-19758-5.
- [63] Z. Wang, W. Yan, T. Oates, B. Zhao, H. Lu, S. Chen, J. Liu, D. Wu, Convolutional neural networks for time series classification, *Journal of Systems Engineering and Electronics* 28 (1) (2017) 162–169. doi:10.21629/JSEE.2017.01.18.

-
- [64] Y. Liu, A. Haridevan, H. Schofield, J. Shan, Application of ghost-deblurgan to fiducial marker detection, in: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2022, pp. 6827–6832. doi:10.1109/IROS47612.2022.9981701.
- [65] H. Son, J. Lee, J. Lee, S. Cho, S. Lee, Recurrent video deblurring with blur-invariant motion estimation and pixel volumes, ACM Trans. Graph. 40 (5) (2021). doi:10.1145/3453720.
- [66] k4yt3x, Video 2x, dostopano: 16-02-2024 (2024).
URL <https://github.com/k4yt3x/video2x>
- [67] Pyqt, dostopano: 7-4-2024.
URL <https://wiki.python.org/moin/PyQt>
- [68] Opencv, dostopano: 7-4-2024.
URL <https://opencv.org/>
- [69] A. Sengupta, F. Jin, R. Zhang, S. Cao, mm-pose: Real-time human skeletal posture estimation using mmwave radars and cnns, IEEE Sensors Journal 20 (17) (2020) 10032–10044. doi:10.1109/jsen.2020.2991741.
- [70] Z. Cao, G. Hidalgo, T. Simon, S. E. Wei, Y. Sheikh, Openpose: Realtime multi-person 2d pose estimation using part affinity fields (2019). arXiv:1812.08008.
- [71] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition (2015). arXiv:1512.03385.
- [72] S. Ameli, Hand detection using transfer learning, dostopano: 7-3-2024.
URL <https://github.com/sameli74/Hand-Detection>
- [73] C. Y. Wang, A. Bochkovskiy, H. Y. M. Liao, Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors (2022). arXiv:2207.02696.

[74] Hand detection - yolov7, dostopano: 7-3-2024.

URL <https://github.com/nuwandda/yolov7-hand-detection>