



AGH

AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE

**WYDZIAŁ ELEKTROTECHNIKI, AUTOMATYKI,
INFORMATYKI I INŻYNIERII BIOMEDYCZNEJ**

Praca dyplomowa

Detekcja anomalii w sygnale mowy u osób z chorobą Parkinsona
Anomalies detection in speech signals in people with Parkinson's
disease

Autor:

Kierunek studiów:

Opiekun pracy:

Małgorzata Szwed

Inżynieria Biomedyczna

dr inż. Daria Hemmerling

Kraków, 2023

Uprzedzony o odpowiedzialności karnej na podstawie art. 115 ust. 1 i 2 ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (t.j. Dz.U. z 2006 r. Nr 90, poz. 631 z późn. zm.): „Kto przywłaszcza sobie autorstwo albo wprowadza w błąd co do autorstwa całości lub części cudzego utworu albo artystycznego wykonania, podlega grzywnie, karze ograniczenia wolności albo pozbawienia wolności do lat 3. Tej samej karze podlega, kto rozpowszechnia bez podania nazwiska lub pseudonimu twórcy cudzy utwór w wersji oryginalnej albo w postaci opracowania, artystycznego wykonania albo publicznie zniekształca taki utwór, artystyczne wykonanie, fonogram, wideogram lub nadanie.”, a także uprzedzony o odpowiedzialności dyscyplinarnej na podstawie art. 211 ust. 1 ustawy z dnia 27 lipca 2005 r. Prawo o szkolnictwie wyższym (t.j. Dz. U. z 2012 r. poz. 572, z późn. zm.): „Za naruszenie przepisów obowiązujących w uczelni oraz za czyny uchybiające godności studenta student ponosi odpowiedzialność dyscyplinarną przed komisją dyscyplinarną albo przed sądem koleżeńskim samorządu studenckiego, zwanym dalej «sądem koleżeńskim».”, oświadczam, że niniejszą pracę dyplomową wykonałem(-am) osobiście i samodzielnie i że nie korzystałem(-am) ze źródeł innych niż wymienione w pracy.

Serdecznie dziękuję ... tu ciąg dalszych podziękowań np. dla promotora, żony, sąsiada itp.

Spis treści

1. Wprowadzenie	7
1.1. Cel pracy	7
1.2. Zakres pracy	7
2. Problematyka zagadnienia w świetle literatury	9
2.1. Charakterystyka choroby Parkinsona	9
2.1.1. Etapy Choroby	10
2.1.2. Głos w chorobie Parkinsona	10
2.1.3. Problematyka choroby	10
2.2. Przegląd dostępnych rozwiązań	10
2.2.1. Rozwiązania teoretyczne	11
2.2.2. Aplikacje rzeczywiste	11
2.2.3. Problem dotychczasowych rozwiązań	11
2.2.4. Wnioski	13
3. Materiał i metoda badawcza	15
3.1. Baza danych	15
3.2. Architektury sieci	15
4. Wyniki badań	17
5. Podsumowanie	19

1. Wprowadzenie

Aktualnie prowadzonych jest wiele prac, które umożliwiłyby stworzenie narzędzia do automatycznej diagnostyki Parkinsona na podstawie sygnału głosowego.

Choroba Parkinsona uważana jest za chorobę cywilizacyjną.

1.1. Cel pracy

Celem pracy jest detekcja choroby Parkinsona na podstawie sygnału głosowego z wykorzystaniem metod uczenia maszynowego. Obejmuje to dokładny przegląd literaturowy ze szczególnym uwzględnieniem aktualnie najlepszych algorytmów dostępnych w literaturze. Na tej podstawie oceniona zostanie skuteczność różnych architektur konwolucyjnych sieci neuronowych (CNN). Ponadto dokonana zostanie analiza, która z wypowiedzianych przez pacjentów samogłosek niesie ze sobą najwięcej informacji diagnostycznej.

1.2. Zakres pracy

- detekcja Parkinsona
- testowanie różnych architektur
- ocena różnych samogłosek

2. Problematyka zagadnienia w świetle literatury

2.1. Charakterystyka choroby Parkinsona

Choroba Parkinsona (PD) jest zwyrodnieniowym schorzeniem mózgu związanym z objawami motorycznymi (spowolnienie ruchowe, drżenie, sztywność, zaburzenia chodu i równowagi) oraz szeroką gamą powikłań niemotorycznych (zaburzenia poznawcze, zaburzenia psychiczne, zaburzenia snu oraz ból i inne zaburzenia sensoryczne). Zaburzenia motoryczne, takie jak dyskinezy (ruchy mimowolne) i dystonie (bolesne mimowolne skurcze mięśni) przyczyniają się do ograniczenia mowy, mobilności i ograniczeń w wielu dziedzinach życia. Postęp tych objawów powoduje wysoki stopień niepełnosprawności i konieczność opieki. U wielu osób z PD w trakcie trwania choroby rozwija się również demencja.

Ryzyko zachorowania na PD rośnie wraz z wiekiem, chociaż choroba może dotyczyć także młodszych osób. Bardziej narażeni są mężczyźni niż kobiety. Przyczyna PD nie jest znana, ale uważa się, że powstaje w wyniku złożonej interakcji pomiędzy czynnikami genetycznymi i narażeniem na czynniki środowiskowe, takie jak pestycydy, rozpuszczalniki i zanieczyszczenia powietrza przez całe życie.

Według raportu Światowej Organizacji Zdrowia[1] w skali globalnej niepełnosprawność i zgony z powodu PD rosną szybciej niż w przypadku jakichkolwiek innych zaburzeń neurologicznych. Częstość występowania PD podwoiła się w ciągu ostatnich 25 lat. Globalne szacunki w 2019 roku wykazały ponad 8,5 miliona osób z PD. Obecne szacunki sugerują, że w 2019 roku PD spowodowała 5,8 miliona lat życia z niepełnosprawnością, co stanowi wzrost o 81% od 2000 roku, i spowodowała 329 000 zgonów, co stanowi wzrost o ponad 100% od 2000 roku.

[DODACĆ WYKRES]

W Polsce z chorobą Parkinsona zmagają się około 100 tys. pacjentów, z czego około 20% jest już w stadium zaawansowanym według informacji przekazywanych przez Fundację Chorób Mózgu. Ponadto co roku w naszym kraju wykrywanych jest ok. 8 tys. nowych zachorowań. Nowe zachorowania nadal skorelowane są z wiekiem, średnia wieku chorych wynosi 60 lat, niestety wzrasta odsetek chorych wśród osób młodych (nawet w wieku 20 lat).

2.1.1. Etapy Choroby

2.1.2. Głos w chorobie Parkinsona

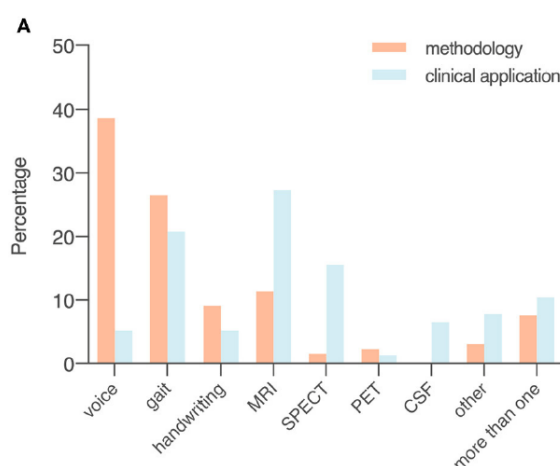
2.1.3. Problematyka choroby

2.2. Przegląd dostępnych rozwiązań

Diagnoza PD jest powszechnie oparta na obserwacjach lekarskich i ocenie objawów klinicznych, w tym charakterystyce różnorodnych objawów ruchowych. Rosnąca liczba zachorowań i obniżenie wieku osób będących w grupie ryzyka, skutkuje wzrostem zainteresowania dotyczącym narzędzi, które ułatwiłyby zarówno codzienne funkcjonowanie pacjentów jak i pracę lekarzy. Tradycyjne metody diagnostyczne mogą być obciążone subiektywizmem ponieważ opierają się na ocenie ruchów, które są czasami subtelne dla ludzkiego oka i dlatego trudne do sklasyfikowania, co może przyczynić się do błędnej diagnozy. Ponadto wczesne objawy niemotoryczne PD mogą być łagodne oraz spowodowane wieloma innymi schorzeniami. Dlatego też rozpoznanie tej choroby na wczesnym etapie stanowi wyzwanie.

Nie da się nie zauważyć, że sztuczna inteligencja oraz nowoczesne technologie coraz częściej stają się integralną częścią systemu ochrony zdrowia. Wspierają lekarzy podczas diagnozy oraz wyboru sposobu leczenia pacjenta, a także pozwalają na monitorowanie choroby. Aby rozwiązać trudności i udoskonalić procedury diagnozowania oraz oceny PD, wdrożono metody uczenia maszynowego do klasyfikacji PD i osób zdrowych lub pacjentów z podobnymi objawami klinicznymi (np. zaburzeniami ruchu lub innymi zespołami parkinsonowskimi).

W przeglądzie publikacji na temat wykorzystania uczenia maszynowego do diagnostyki PD do 2020 roku wyróżniono 209 artykułów[2].



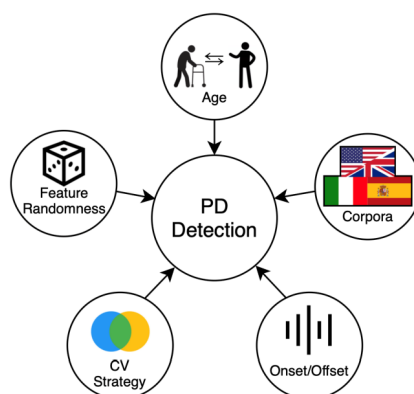
Rys. 2.1. Wykres przedstawiający rozkład rodzaju danych na których bazowały systemy ML do diagnostyki PD[2] (stan na dzień 14 luty 2020)

2.2.1. Rozwiązania teoretyczne

2.2.2. Aplikacje rzeczywiste

2.2.3. Problem dotychczasowych rozwiązań

Ostatnie badania wykazały, że możemy wytrenować dokładne modele do wykrywania oznak PD z nagrań audio. Jednakże, istnieją rozbieżności pomiędzy badaniami i mogą być spowodowane, częściowo, przez różnice w wykorzystywanych korpusach lub metodologii. Dlatego w [3] przeprowadzono analizę, wpływu niektórych czynników na wyniki klasyfikacji. Głównym celem artykułu była ich identyfikacja oraz stworzenie zasad, które w przyszłości pozwolą usystematyzować stan wiedzy w tej dziedzinie. W badaniach skupiono się na przedłużonych samogłoskach (ang. *sustained vowels*), ponieważ jak wykazano wcześniej [...]. Przeprowadzone eksperymenty wykazały, że nieuwzględnione zmienne w metodologii, projekcie eksperymentalnym i przygotowaniu danych prowadzą do zbyt optymistycznych wyników w badaniach nad automatyczną detekcją PD. Czynniki, które zidentyfikowano jako przyczyniające się do zbyt optymistycznych wyników klasyfikacji przedstawiono na Rys. 2.2 oraz omówiono poniżej.



Rys. 2.2. Czynniki wpływające na dokładność detekcji Parkinsona na podstawie głosu według analizy przeprowadzonej w [3]

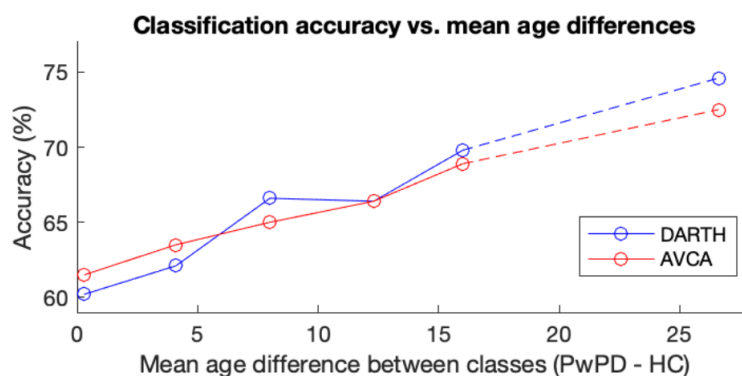
a) Wpływ tożsamości mówcy

W przypadku, gdy w zbiorze danych znajduje się kilka nagrań od tego samego mówcy można postąpić na dwa sposoby. Pierwszy z nich to podział według podmiotów (ang. *subject-wise split*) polegający na tym, że nagrania od tej samej osoby znajdują się albo w zbiorze treningowym albo testowym - nigdy w obu na raz. W niektórych opracowaniach można spotkać też drugie podejście, czyli podział według rekordów (ang. *record-wise split*), gdzie nagrania są losowo dzielone do zbiorów lub intencjonalnie używa się nagrań od tej samej osoby zarówno w zbiorze testowym jak i treningowym. Okazuje się, że podejście typu *record-wise* prowadzi do wyższej dokładności niż *subject-wise split*, jeśli pozostałe założenia pozostają identyczne. Prawdopodobnie wynika to z faktu, że klasyfikator nastawia się na wykrywanie unikalnych informacji indywidualnych, reprezentowanych przez współczynniki takie jak MFCC i PLP, a nie rzeczywiste biomarkery lub wzorce PD.

Dlatego też rekomendowana jest technika *subject-wise split*, aby uniknąć zbyt optymistycznych wyników.

b) Wpływ różnicy wieku między klasami

W literaturze można znaleźć prace wykorzystujące zbiory danych, w których średni wiek mówców w klasie osób chorych na PD różni się od średniego wieku w klasie osób zdrowych o ponad 5 lat. Autorzy zapewniają o wysokiej skuteczności swoich rozwiązań, jednak nie rozważają oni sytuacji, w której klasyfikator uczy się wykrywać cechy powiązane z wiekiem, zamiast rzeczywistych wzorców PD. Aby zbadać wpływ średniej różnicy wieku między dwiema klasami, w [3] testowano różne przesunięcia rozkładu wieku. Wraz ze wzrostem różnicy między średnim wiekiem uczestników z PD i HC, dokładność klasyfikacji konsekwentnie rosła (Rys. 2.3). Na tej podstawie można stwierdzić, że związany z wiekiem wpływ na głos mówców może zaburzać wyniki otrzymywane przez klasyfikator. Dlatego też zaleca się zbilansowanie używanych zbiorów danych, tak aby średnia różnica wieku między dwoma klasami była jak najmniejsza.



Rys. 2.3. Wykres przedstawiający zależność różnicy wieku między klasami a dokładnością klasyfikacji [3]

c) Wpływ losowości cech na dokładność klasyfikacji

Im większa różnica między liczbą plików a wymiarem wektora cech, tym większe szanse na znalezienie cech, która losowo koreluje z etykietami klas.

d) Łagodzenie losowego nadmiernego dopasowania przy użyciu danych programistycznych

e) Wpływ rozpoczęcia i przesunięcia samogłosek na wyniki klasyfikacji

f) Eksperymenty międzykorporowe

g) Analiza cech

Nie są to jednak wszystkie czynniki, które zaburzają obiektywność wyników. Konieczna jest dyskusja na temat nowych kompleksowych linii bazowych dla prowadzenia eksperymentów w automatycznym wykrywaniu PD na podstawie fonacji, a także innych ogólnych zastosowań przetwarzania mowy.

2.2.4. Wnioski

Prace nad automatyczną detekcją Parkinsona na podstawie głosu trwają już od dłuższego czasu. Jednak wciąż brakuje systemu, który mógłby zostać uznany jako wystarczająco niezawodne narzędzie diagnostyczne. Wśród problemów, które ograniczają rzeczywiste wykorzystanie takich systemów wyróżnia się:

- pierwszy punkt
- drugi punkt
- trzeci punkt

3. Materiał i metoda badawcza

3.1. Baza danych

3.2. Architektury sieci

4. Wyniki badań

5. Podsumowanie