



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.

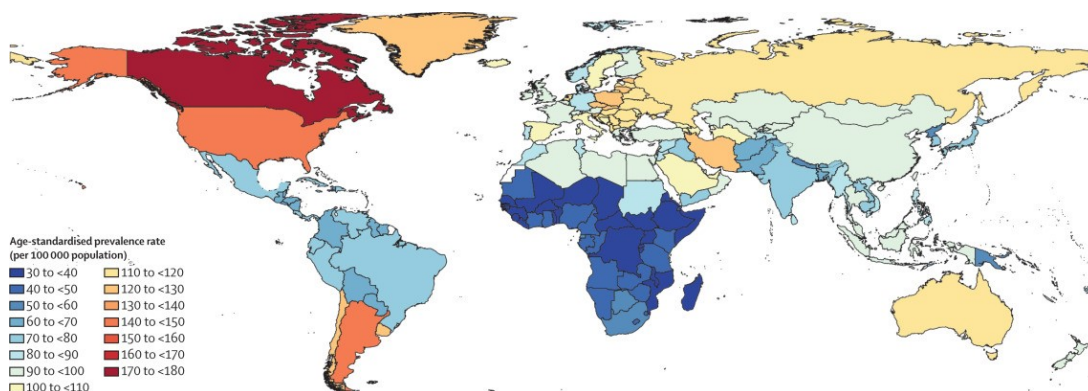
Serdecznie dziękuję ... tu ciąg dalszych podziękowań np. dla promotora, żony, sąsiada itp.

Spis treści

1. Wprowadzenie	7
1.1. Cel pracy	9
1.2. Zakres pracy	9
2. Problematyka zagadnienia	11
2.1. Znaczenie głosu w chorobie Parkinsona	11
2.2. Metody diagnozowania i monitorowania	14
2.3. Terapia osób chorych	16
3. Analiza rozwiązań do automatycznej diagnostyki choroby Parkinsona	19
3.1. Rodzaj wykorzystywanych danych	19
3.2. Metody klasyfikacji	21
3.3. Wyzwania związane z systemami automatycznej diagnostyki	24
4. Materiał i metoda badawcza	29
4.1. Materiał badawczy	29
4.2. Parametryzacja sygnału akustycznego	32
4.3. Metody klasyfikacji	32
4.4. Metody ewaluacji wyników	32
5. Wyniki badań	33
5.1. Samogłoska /a/	33
5.2. Samogłoska /e/	33
5.3. Samogłoska /i/	33
5.4. Samogłoska /o/	33
5.5. Samogłoska /u/	33
5.6. Zbiorcze podsumowanie wyników	33
6. Analiza i interpretacja wyników	35

1. Wprowadzenie

Choroba Parkinsona (ang. *Parkinson Disease, PD*) to zwyrodnieniowe schorzenie mózgu, które wiąże się z objawami ruchowymi, takimi jak spowolnienie ruchowe, drżenie, sztywność oraz zaburzenia chodu i równowagi. Ponadto może prowadzić do różnorodnych powikłań niemotorycznych, obejmujących zaburzenia poznawcze, stany psychiczne, trudności ze snem oraz dolegliwości sensoryczne, w tym ból. Początkowe objawy często rozwijają się stopniowo, nasilając się w miarę upływu czasu. Postęp choroby prowadzi do znacznego stopnia niepełnosprawności, co może wymagać wsparcia i opieki. U wielu osób zdiagnozowanych z chorobą Parkinsona występują także zmiany w sferze psychicznej i behawioralnej, takie jak trudności ze snem, depresja, problemy z pamięcią oraz uczucie przewlekłego zmęczenia.



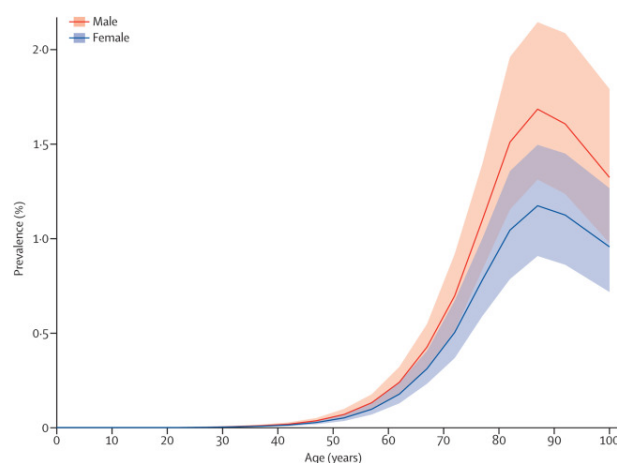
Rys. 1.1. Choroba Parkinsona na świecie [1]

Zgodnie z danymi przedstawionymi w raporcie Światowej Organizacji Zdrowia [2], choroba Parkinsona (PD) stanowi obecnie narastający problem na skalę światową. Zarówno wskaźniki niepełnosprawności, jak i zgony związane z tą chorobą rosną szybciej niż w przypadku innych zaburzeń neurologicznych.

W ciągu ostatnich 25 lat zaobserwowano podwojenie częstości występowania PD na całym świecie. Globalne szacunki na rok 2019 wskazują, że liczba osób cierpiących na PD przekroczyła 8,5 miliona, co oznacza wzrost o 81% w porównaniu z danymi z roku 2000. Jednocześnie liczba zgonów związanych z PD wyniosła 329 000, co stanowi wzrost o ponad 100% w porównaniu z rokiem 2000 [1]. 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.

PD jest istotną sprawą dotyczącą zdrowia publicznego, ponieważ jej częstotliwość występowania związana jest ze zjawiskiem starzejącego się społeczeństwa. Razem z innymi chorobami neurodegeneracyjnymi, takimi jak choroba Alzheimera, ma szansę stać się drugą zaraz za nowotworami przyczyną zgonów do 2040 roku (WHO).

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. Niektóre przypadki wydają się dziedziczne, a kilka można przypisać określonym wariantom genetycznym. Chociaż uważa się, że genetyka odgrywa rolę w chorobie Parkinsona, w większości przypadków choroba nie występuje rodzinie [3].



Rys. 1.2. Rozpowszechnienie choroby Parkinsona w zależności od wieku [1]

Mimo że każdy może być narażony na ryzyko rozwoju tej choroby, to częściej występuje ona u mężczyzn niż u kobiet, a wiek stanowi kluczowy element wpływający na ryzyko zachorowania, co można zobaczyć na Rys.1.2. Statystyki pokazują, że ryzyko zachorowania rośnie wraz z wiekiem, chociaż choroba może dotyczyć także młodszych osób (nawet w wieku 20 lat). U większości pacjentów po raz pierwszy choroba rozwija się po 60 roku życia, około 5% do 10% doświadcza jej początku przed 50 rokiem życia. Postacie choroby Parkinsona o wczesnym początku są często, choć nie zawsze, dziedziczne i niektóre formy zostały powiązane z określonymi zmianami w genach [3].

Proces diagnozowania choroby jest niezwykle złożony i czasochłonny, nie istnieje obecnie pojedyncze badanie pozwalające na postawienie diagnozy. W związku z tym poszukuje się nowych rozwiązań, które mogłyby usprawnić ten proces. Coraz częściej wykorzystuje się metody uczenia maszynowego i sztucznej inteligencji w dziedzinie medycyny. W niniejszej pracy analizowany jest aktualny stan rzeczy oraz przedstawiane jest proponowane rozwiązanie, dotyczące automatycznej diagnostyki choroby Parkinsona na podstawie głosu.

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. Pośrednim celem jest ocena skuteczności wybranych architektur konwolucyjnych sieci neuronowych (CNN) oraz analiza, która z wypowiedzanych przez pacjentów samogłosek niesie ze sobą najwięcej informacji diagnostycznej.

1.2. Zakres pracy

2. Problematyka zagadnienia

2.1. Znaczenie głosu w chorobie Parkinsona

Najbardziej widoczne oznaki i objawy choroby Parkinsona pojawiają się, gdy komórki nerwowe w zwojach podstawy mózgu, obszarze kontrolującym ruch, ulegają uszkodzeniu i/lub obumierają. Zwykle te komórki nerwowe lub neurony wytwarzają dopaminę. Kiedy neurony obumierają lub ulegają uszkodzeniu, wytwarzają mniej dopaminy, co powoduje problemy z poruszaniem się związane z chorobą. Na ten moment nie wiadomo co powoduje śmierć neuronów. Zanikają również zakończenia nerwowe, które wytwarzają norepinefrynę, główny przekaźnik chemiczny współczulnego układu nerwowego, który kontroluje wiele funkcji organizmu, takich jak tętno i ciśnienie krwi. Utrata norepinefryny może pomóc wyjaśnić niektóre cechy choroby Parkinsona związane z brakiem ruchu, takie jak zmęczenie, nieregularne ciśnienie krwi, zmniejszony ruch pokarmu w przewodzie pokarmowym i nagły spadek ciśnienia krwi, gdy osoba wstaje z pozycji siedzącej lub leżącej.

Do czterech głównych objawów choroby Parkinsona zalicza się:

- drżenie rąk, ramion, nóg, szczęki lub głowy,
- sztywność mięśni, gdy mięśnie pozostają skurczone przez długi czas,
- powolność ruchu,
- zaburzenia równowagi i koordynacji, czasami prowadzące do upadków.

Zarówno objawy choroby Parkinsona jak i tempo jej postępu mogą znacząco różnić się wśród poszczególnych osób. Na wczesnym etapie choroby objawy są subtelne i kształtują się stopniowo. Często zaczynają się od jednej strony ciała lub nawet jednej kończyny. Niektórzy pacjenci doświadczają pewnych zwiastunów przed wystąpieniem charakterystycznych cech, takich jak sztywność czy drżenie. Mogą to być trudności ze snem, problemy z wypróżnianiem, utrata węchu czy także zespół niespokojnych nóg. Warto jednak zaznaczyć, że niektóre z wymienionych objawów mogą również występować w procesie naturalnego starzenia się [3].

Chociaż tempo postępu choroby Parkinsona zazwyczaj jest powolne, w końcu ma to wpływ na codzienne funkcjonowanie osoby dotkniętej tą dolegliwością. Wykonywanie zwykłych czynności, takich jak praca, prowadzenie domu czy uczestnictwo w spotkaniach towarzyskich z przyjaciółmi, może stawać się wyzwaniem.

Choroba Parkinsona jest spowodowana zaburzeniem funkcji układu nerwowego, co prowadzi do występowania objawów w różnych obszarach organizmu, w tym również w sferze głosu. Objawy zaburzeń mowy dotyczą między 60 a 80% pacjentów, a powodowane są przede wszystkim przez deficyt czynnościowy krtani, zmniejszoną pojemność życiową płuc, osłabioną pracę mięśni mimicznych oraz zmniejszony napęd mówienia [4]. Zwykle stają się widoczne w średniozaawansowanym stadium schorzenia, co oznacza, że przez długi okres mowa pozostaje relatywnie nienaruszona. Rozpoznanie tych zaburzeń może być niekiedy trudne, gdyż wymaga odróżnienia, czy powstały one na skutek samej choroby, czy też są rezultatem naturalnego procesu starzenia się organizmu.

Objawy zaburzeń mowy i głosu związane z chorobą Parkinsona nie są łatwo dostrzegalne dla osób bez specjalistycznej wiedzy w tej dziedzinie. Przeważnie zdolność rozumienia mowy pozostaje niezmieniona. Niemniej jednak, w trakcie spontanicznych wypowiedzi pacjenci zaczynają ograniczać ilość przekazywanych informacji i mogą napotykać trudności w składaniu pełnych zdań. Trudności te nie wynikają koniecznie z ubytku słownictwa, ale raczej z nieprawidłowego doboru słów. Wskazujące na podłoże chorobowe objawy obejmują między innymi [5, 6]:

- mowę powolną, monotonną i przerywaną,
- nadmierne ślinienie się,
- niewyraźną i zamazaną artykulację,
- skrócony czas fonacji
- chuchający i tremolujący głos
- spłaszczoną barwę i obniżone natężenie
- niewłaściwą koordynację mięśni nasady, które mogą być zwiotczone lub zbyt napięte,
- czasami przyspieszenie tempa wypowiedzi w jej końcowej fazie, co może utrudnić zrozumienie pacjenta.

Wymienione objawy pojawiają się w średniozaawansowanym stadium choroby i są wystarczająco wyraźne, aby mogły zostać zauważone słuchowo przez specjalistów. Niemniej jednak badania wskazują, że istnieją subtelne zmiany w głosie, które pojawiają się jeszcze wcześniej, nawet w fazie przedobjawowej [7].

W roku 2000 przeprowadzono badanie akustyczne i percepcyjne cech głosu pacjentów z chorobą Parkinsona, zależnie od stopnia nasilenia choroby [8]. W nagraniach głosowych, składających się z przedłużonej samogłoski /a/, śpiewu gamy oraz 1-minutowego monologu, stwierdzono, że głosy pacjentów z PD, zarówno we wczesnych, jak i późniejszych stadiach choroby, charakteryzowały się ograniczoną percepcyjnie zmiennością tonu i głośności, chropowatością oraz zmniejszoną głośnością.

Badanie sugerowało również, że głosy pacjentów z PD wykazywały nadmierne drganie, wysoką częstotliwość podstawową (szczególnie u mężczyzn) oraz zmniejszoną zmienność częstotliwości podstawowej (szczególnie u kobiet). Część z tych cech głosu nie wydawała się pogarszać w miarę postępu

choroby, jednak cechy takie jak oddech, monotonność i jednolitość mowy, niska głośność oraz ograniczony maksymalny zakres częstotliwości fonacyjnej były bardziej zauważalne w późniejszych stadiach choroby Parkinsona.

Podobne badanie przeprowadzone przez Gamboę i innych (1997) wykazało, że w porównaniu z grupą kontrolną, pacjenci z PD wykazywali wyższy jitter, niższy stosunek harmonicznych do szumów (H/N), mniejszą zmienność częstotliwości i intensywności mowy oraz niższy zakres fonacyjny oraz wyższą częstotliwość obecności głosu o niskim natężeniu. Wskazano również, że te cechy nie wykazywały znaczącego związku z czasem trwania choroby [9].

Mnogość objawów, które są zauważalne w głosie, motywuje do uwzględnienia ich w diagnostyce. Prowadzone są badania, które wykorzystują analizę mowy do wykrywania patologii i schorzeń związanych z narządem głosu, takich jak ostre zapalenie krtani czy porażenie nerwu krtaniowego wstecznego. Może to w przyszłości pozwolić na identyfikację problemów zdrowotnych, na przykład wśród osób pracujących głosem, jak nauczyciele, bez konieczności inwazyjnych badań gardła. Podobne podejście można zastosować do diagnozowania i monitorowania chorób neurodegeneracyjnych. Badania naukowe wskazują, że analiza głosu może stanowić podstawę dla automatycznej diagnostyki oraz monitorowania choroby Parkinsona.

Takie podejście niesie ze sobą wiele korzyści, które mogą rewolucjonizować sposób diagnozowania oraz monitorowania tej neurodegeneracyjnej choroby. Głos, będący wskaźnikiem stanu układu nerwowego i zdolności komunikacyjnych, dostarcza szeroką gamę informacji kluczowych dla procesu diagnozowania. Różnorodność parametrów akustycznych i fonacyjnych, które można zbadać, otwiera drzwi do kompleksowej oceny zmian zachodzących w organizmie pacjenta.

Wczesne objawy choroby Parkinsona często bywają trudne do wykrycia, szczególnie w standardowych badaniach klinicznych. Analiza głosu pozwala na szybką identyfikację subtelnych zmian, które pojawiają się we wczesnych fazach choroby. Ta wczesna detekcja umożliwia natychmiastową interwencję terapeutyczną, co może wpłynąć na spowolnienie progresji choroby i poprawę jakości życia pacjenta.

Analiza głosu jako narzędzie diagnostyczne wprowadza nowe perspektywy dla specjalistów zajmujących się chorobą Parkinsona. Logopedzi, foniatry i lekarze mogą wykorzystać obiektywne dane akustyczne do dokładnej oceny zmian w mowie i jakości głosu pacjenta. To również umożliwia sugerowanie odpowiednich interwencji, włącznie z dostosowaniem leczenia farmakologicznego.

Dla pacjentów oznacza to bardziej konkretne oceny ich stanu i dostosowane terapie, przyczyniające się do zwiększenia efektywności procesu leczenia. Badanie to jest szybkie, wygodne i bezpieczne, co może zachęcać pacjentów do systematycznego uczestnictwa w procesie diagnostycznym. Wprowadza to szczególną wartość dla pacjentów, którym trudno się przemieszczać.

Analiza głosu jako narzędzie diagnostyczne przy chorobie Parkinsona otwiera drzwi ku nowym, zaawansowanym metodologiom diagnozowania i leczenia. Może być wykorzystana do oceny stanu pacjenta oraz sugerowania odpowiednich interwencji, włączając zmiany w leczeniu farmakologicznym. W konsekwencji ma potencjał stania się szybkim, nieinwazyjnym wsparciem diagnostycznym i terapeutycznym, przyczyniając się do polepszenia opieki nad pacjentami dotkniętymi chorobą Parkinsona.

2.2. Metody diagnozowania i monitorowania

Diagnostyką choroby Parkinsona zajmują się neurologi i geriatrzy. Rozwój PD jest długotrwały, a w początkowych latach klinicznie niemal niewidoczny, co utrudnia wczesne rozpoznanie. Subtelne objawy często są uważane za skutek starzenia się lub błędnie diagnozowane jako inne zaburzenia neurologiczne. Kluczowym elementem w tym stadium jest dokładny wywiad, badanie fizykalne oraz identyfikacja objawów przez lekarza. Następnie diagnoza jest rozwijana poprzez badania laboratoryjne oraz obrazowe. Niestety, wyniki tych badań rzadko potwierdzają diagnozę od razu.

Początkowo pacjent zwykle konsultuje się z lekarzem pierwszego kontaktu, który powinien dokonać wstępnej diagnozy i skierować do neurologa. W tej fazie diagnozy przeprowadza się szczegółowy wywiad, uwzględniający rodzaj, nasilenie oraz okres występowania objawów, a także obecność chorób neurodegeneracyjnych w rodzinie. Neurolog przeprowadza kompleksowe badanie neurologiczne, identyfikując symptomy takie jak sztywność mięśni, ograniczenia w ruchu (spowolnienie, trudności w poruszaniu się), drżenia spoczynkowe (np. w głowie, palcach rąk) oraz zaburzenia postawy i równowagi (zgarbienie, niestabilność, upadki). Kolejne badania są wykonywane w celu potwierdzenia lub wykluczenia diagnozy [10, 11].

a) Badania laboratoryjne

Obecnie brak specyficznych badań laboratoryjnych krwi, które potwierdzałyby diagnozę choroby Parkinsona. Niemniej jednak, takie badania są użyteczne w wykluczaniu innych chorób o podobnym przebiegu. Wykonuje się podstawowe badania, takie jak morfologia krwi, elektrolity, poziom glukozy, TSH, próby wątrobowe, mocznik, kreatynina oraz poziom witaminy B12.

b) Badania obrazowe

Badania obrazowe głowy są przeprowadzane w celu wykluczenia innych chorób o podobnych objawach. Zalicza się do nich tomografię komputerową, ultrasonografię mózgu (USG) oraz rezonans magnetyczny głowy (MRI). Międzynarodowe kryteria rozpoznania choroby Parkinsona nie nakładają obowiązku wykonywania badań obrazowych w celu potwierdzenia diagnozy. Dostępne są również zaawansowane techniki obrazowania, takie jak PET (pozytonowa emisyjna tomografia) oraz SPECT (tomografia emisyjna pojedynczego fotonu), które pozwalają na obserwację metabolizmu w układzie pozapiramidowym. Skan DAT (skan transportera dopaminy) jest przykładem SPECT i może być sugerowany przez specjalistę. Mimo to, ostateczna diagnoza opiera się na objawach oraz wynikach badania neurologicznego. Większość pacjentów nie wymaga skanowania DAT.

c) Test z lewodopą

Test polega na podaniu pacjentowi preparatu z lewodopą. Jeśli następuje poprawa po zażyciu, istnieje wysokie prawdopodobieństwo, że pacjent rzeczywiście cierpi na chorobę Parkinsona. W przypadku braku poprawy, konieczne może być dalsze rozszerzenie diagnostyki.

d) Badania genetyczne

Choroba Parkinsona może występować rodzinnie, co skłania do rozważenia diagnostyki genetycznej u pacjenta i jego krewnych. Obecnie zidentyfikowano 12 mutacji genów, które mogą wpływać na ryzyko zachorowania na PD, należy jednak pamiętać, że badania genetyczne są kosztowne.

e) Badania węchu

Większość osób z chorobą Parkinsona (90%) doświadcza zaburzeń węchu, manifestujących się hiposomią (osłabienie węchu), także we wczesnym stadium choroby. Jednak nie obserwuje się tych zaburzeń w przypadku zaniku wieloukładowego i postępującego porażenia nadjądrowego.

f) Badania neuropsychologiczne i neuropsychiatryczne

Badania te służą identyfikacji zaburzeń poznawczych i emocjonalnych. Celem jest diagnoza łagodnych zaburzeń poznawczych, otępienia, a także zaburzeń psychotycznych, lękowych, kontroli impulsów i depresji.

Objawy przypominające chorobę Parkinsona są diagnozowane jako parkinsonizm i mogą być spowodowane różnymi zaburzeniami, takimi jak postępujące porażenie nadjądrowe, zanik wieloukładowy, drżenie samoistne, demencja z ciałami Lewy'ego, choroby naczyniowe mózgu, otępienie, reumatyzm i inne [10]. Różnicowanie tych schorzeń jest kluczowe, ponieważ leczenie i podejście terapeutyczne są różne. Chociaż ostateczną diagnozę można ustalić tylko na podstawie badania mózgu po zgonie, wcześniej zdefiniowane kryteria diagnostyczne pozwalają na dokonanie diagnozy klinicznej. Według badania z 2021 roku opisanego w publikacji [12] diagnoza klinicznie potwierdzonej choroby Parkinsona może zabierać od kilku miesięcy do kilku lat, zależnie od indywidualnych czynników oraz reakcji na terapię lewodopą. Uznaje się, że aktualnie wykorzystywane kryteria pozwalają na diagnozę z dokładnością 90%, ale trwa to średnio 2,9 roku.

Pomocnym narzędziem w diagnostyce są szeroko stosowane skale oceny choroby Parkinsona. Stanowią istotne narzędzie w monitorowaniu stanu pacjentów oraz pomagają lekarzom i opiekunom ocenić stopień nasilenia objawów ruchowych, jak również wpływ choroby na codzienne funkcjonowanie pacjenta. Popularne skale, takie jak Skala Hoehn-Yahra, Skala UPDRS (Unified Parkinson's Disease Rating Scale) oraz Skala Schwab-England, umożliwiają stosunkowo obiektywną analizę symptomów i wsparcie w podejmowaniu decyzji terapeutycznych. Dzięki tym narzędziom możliwe jest dostosowanie leczenia do indywidualnych potrzeb pacjenta oraz śledzenie skuteczności terapii na przestrzeni czasu.

Obecnie proces diagnozy jest wyjątkowo złożony i wieloetapowy. W odpowiedzi na te wyzwania, naukowcy koncentrują się na opracowaniu bardziej efektywnych narzędzi diagnostycznych. Rozwinięcie skuteczniejszych narzędzi diagnostycznych przyniesie korzyści nie tylko finansowe, ale także pozwoli na szybsze i trafniejsze udzielanie pomocy pacjentom. W nadchodzących latach, dążenie do wypracowania bardziej efektywnych metod diagnozowania choroby Parkinsona będzie kluczowym krokiem w zapewnieniu lepszej opieki zdrowotnej i poprawie jakości życia pacjentów.

2.3. Terapia osób chorych

Obecnie brak jest kuracji na chorobę Parkinsona, dlatego terapia skupia się na przywracaniu pacjentom zdolności funkcjonowania lub, w przypadkach zaawansowanych, na poprawie jakości życia. Zgodnie z aktualnym standardem medycznym, w terapii wykorzystuje się różnorodne metody, w tym leczenie farmakologiczne, głęboką stymulację mózgu oraz rehabilitację [3].

Leczenie farmakologiczne choroby Parkinsona opiera się na zwiększeniu poziomu dopaminy w mózgu, co wpływa na kontrolę objawów ruchowych i niezwiązanych z ruchem. Główną terapią jest lewodopa, która jest przetwarzana przez komórki nerwowe w dopaminę. Leczenie lewodopą często łączy się z karbidopą, która zmniejsza skutki uboczne i ilość potrzebnej lewodopy. Stosuje się też inne terapie farmakologiczne o różnych zasadach działania m.in. pobudzające produkcję dopaminy, zwiększające ilość dopaminy poprzez spowolnienie jej rozkładu, redukujące ruchy mimowolne czy zmniejszające drżenie i sztywność mięśni.

W przypadku pacjentów, u których leczenie farmakologiczne nie przynosi oczekiwanych efektów, może być rozważana Głęboka Stymulacja Mózgu (ang. *Deep Brain Stimulation* - DBS). W tym procedurze chirurgicznym lekarz implantuje elektrody w określone obszary mózgu, łącząc je z małym urządzeniem elektrycznym umieszczonym w klatce piersiowej. Poprzez bezbolesne stymulowanie konkretnych obszarów mózgu kontrolujących ruch, DBS może pomóc w zmniejszeniu wielu objawów związanych z ruchem, takich jak drżenie, spowolnienie ruchu i sztywność.

Kluczową rolę w leczeniu odgrywa rehabilitacja neurologiczna, rozpoczynając się już od momentu postawienia diagnozy. Jej wsparcie jest nieocenione w łagodzeniu zaburzeń chodu, głosu, drżenia, sztywności oraz pogorszenia funkcji umysłowych. Wśród różnorodnych terapii, znajdują się między innymi:

- zbilansowana dieta: wspiera ogólne samopoczucie pacjenta,
- ćwiczenia fizyczne: wzmacniają mięśnie, poprawiają równowagę, elastyczność i koordynację,
- masaże terapeutyczne: pomagają w redukcji napięcia mięśniowego oraz przynosi ulgę w objawach,
- joga i tai chi: wspomagają rozciąganie i elastyczność ciała, wpływając korzystnie na zdolność ruchową,
- rehabilitacja foniczna: pomaga w eliminowaniu trudności w mówieniu,
- psychoterapia: zapewnia wsparcie i umożliwia pacjentom pełne cieszenie się życiem pomimo choroby.

W ostatnim czasie na rynku pojawiło się też wiele aplikacji, które mają poprawić jakość życia osób z chorobą Parkinsona. Na przykład, *Parkinson's Central* zawierająca informacje dla pacjentów i opiekunów, obejmujące objawy, leczenie, wizyty lekarskie, zdrowy styl życia, badania i inne aspekty. Podobną aplikacją jest *Parkinson Symptom Tracker (PRO-PD App)*, zaprojektowana jako narzędzie do oceny i monitorowania nasilenia objawów choroby Parkinsona w czasie. Bazuje na wywiadzie w formie testu, który umożliwia ocenę objawów. Skala została opracowana w celu bycia wrażliwą na różne etapy choroby, charakteryzuje się dokładnością, szybkością, prostotą użycia i dostępnością.

Również w dziedzinie rehabilitacji ruchowej istnieje kilka aplikacji, takich jak *Lift Pulse*, służąca do rejestrowania danych dotyczących drżenia rąk. *Delay the Disease*, czyli program fitnessu rozwijany przez OhioHealth dla osób z chorobą Parkinsona, mający na celu poprawę funkcji fizycznych i opóźnienie postępu objawów choroby. W kontekście rehabilitacji poprzez gry cyfrowe warto wspomnieć o platformie *MindMotion® GO*. Oferuje ona różnorodne aktywności i śledzenie ruchu ciała dzięki technologii śledzenia markerów. Jedną z istotnych zalet programu to zdalne monitorowanie i dostosowanie przez terapeutę, zapewniające efektywną opiekę.

Obok aplikacji rehabilitacyjnych istnieje również *Parkinson's Cognitive Research* dedykowana osobom zainteresowanym uczestnictwem w badaniach naukowych dotyczących objawów poznawczych związanych z chorobą Parkinsona. Aplikacja umożliwia analizę aspektów takich jak skupiona uwaga, percepcja wzrokowa, rozpoznawanie, pamięć krótkotrwała, krótkotrwała pamięć wzrokowa, nazewnictwo, pamięć operacyjna, elastyczność poznawcza, planowanie, czas reakcji i prędkość przetwarzania. Jej głównym celem jest wspieranie badań naukowych poprzez dostarczanie narzędzi cyfrowych do oceny i terapii poznawczej. Mimo że stanowi cenny instrument dla społeczności naukowej oraz uniwersytetów na całym świecie, to jednak ma wyłącznie charakter badawczy i nie jest przeznaczona do diagnozowania ani leczenia choroby Parkinsona.

2.3.1. Rola i zastosowanie aplikacji głosowych w poprawie jakości mowy u osób z chorobą Parkinsona

Tematem przewodnim tej pracy magisterskiej jest rola głosu w kontekście choroby Parkinsona. W związku z tym, przedstawione zostaną różnorodne aplikacje dostępne obecnie na rynku, które koncentrują się na poprawie jakości głosu w przypadku tej choroby.

Jedną z nich jest *Speak Up For Parkinson's*, przedstawiona przez Northwest Parkinson's Foundation. Aplikacja skupia się szczególnie na głośności głosu pacjenta. Oferuje dwa narzędzia do ćwiczeń: *Słowa i Zwroty*, gdzie użytkownik musi wypowiedzieć serię losowych stwierdzeń oraz *Czytanie i Konwersacja*, czyli obszar do swobodnych ćwiczeń o dłuższym czasie trwania. W obu narzędziach dostarczany jest miernik głośności oraz opinie dźwiękowe/wideo.

Inną propozycją jest aplikacja *Voice Trainer*, stworzona z myślą o osobach cierpiących na problemy z mową związane z chorobą Parkinsona, ale również jest przydatna dla profesjonalnych mówców oraz, przy wsparciu logopedy, dla osób z innymi zaburzeniami głosu lub mowy. W ramach aplikacji wyświetlane są wizualne informacje zwrotne dotyczące głośności i tonu mowy za pomocą jednej kropki na ekranie, co pozwala szybko zidentyfikować obszary wymagające poprawy. Aplikacja może być używana zarówno do ćwiczenia techniki, jak i do monitorowania toku rozmowy.

Kolejną godną uwagi aplikacją jest *Delayed Auditory Feedback (DAF)*. Przeznaczona jest dla osób z zaburzeniami mowy, które charakteryzują się szybkim tempem wypowiedzi. Aplikacja pomaga użytkownikom zwolnić tempo mówienia, co w efekcie czyni je bardziej zrozumiałymi dla innych. Działanie opóźnionej zwrotności dźwiękowej (DAF) polega na zmienionym odbiorze własnej mowy. Zakłócenie

normalnego cyklu zwrotnej informacji dźwiękowej skutkuje spowolnieniem tempa mówienia i poprawą klarowności wypowiedzi.

Aplikacja *Beats Medical Parkinson's* obejmuje rehabilitację objawów związanych z mową, chodem oraz drżeniem rąk. Została zaprojektowana tak, aby pomóc użytkownikom ćwiczyć głośnie i wyraźne mówienie, co przekłada się na większą pewność siebie podczas komunikacji. Ważnym atutem aplikacji jest możliwość otrzymywania informacji zwrotnych w czasie rzeczywistym podczas ćwiczeń. Dzięki temu użytkownicy mogą śledzić swoje postępy i dostosowywać trening do swoich potrzeb.

Aplikacja LSVT LOUD skupia się na treningu osób z chorobą Parkinsona w celu osiągnięcia bardziej naturalnego poziomu głośności podczas mówienia w codziennych sytuacjach, takich jak komunikacja w domu, pracy czy w społeczności. W kontekście teorii terapeutycznej, kluczowym aspektem tej metody jest pomoc pacjentom w "rekalibracji" ich percepcji dźwięku, aby byli świadomi własnego głosu w kontekście interakcji z innymi ludźmi. Dzięki tym treningom, osoby z chorobą Parkinsona mogą lepiej kontrolować swoją głośność mówienia i efektywniej uczestniczyć w komunikacji.

Zupełnie inną aplikacją, zapewniającą ciągłe wsparcie ze strony specjalistów jest *Teletherapy*. Oferuje ona terapię głosu dla osób z chorobą Parkinsona. Dzięki dostępowi zdalnemu pacjenci mogą oceniać swój stan oraz cele terapeutyczne, monitorować postępy i otrzymywać spersonalizowane wsparcie od logopedów. Aplikacja kieruje użytkowników przez program terapii mowy, a ćwiczenia są dostosowane do indywidualnych potrzeb i przypisywane przez logopedę. Ponadto terapeuci mają możliwość monitorowania ćwiczeń i udzielania opinii zwrotnej.

Aplikacja *Voiceitt* nie ma zastosowania w rehabilitacji, ale można znacznie poprawić komfort życia osób z chorobą Parkinsona. Pomaga tłumaczyć niewyraźne lub nietypowe dźwięki na zrozumiałą mowę. Każdy użytkownik trenuje oprogramowanie, podążając za prostymi komendami na ekranie. Aplikacja może być zintegrowana jako samodzielne ASR (*Automatic Speech Recognition*) dla osób o nietypowej mowie lub używany z innym systemem ASR jako rozszerzenie dostępności z dodatkowym interfejsem dla osób z ograniczoną kontrolą ruchową i zaburzeniami mowy. Planowane jest rozszerzenie aplikacji o napisy na spotkaniach z asystentów głosowych i sterowanie inteligentnym domem.

Powyższy przegląd ukazuje różnorodność dostępnych aplikacji, skierowanych do osób z chorobą Parkinsona. Te narzędzia cyfrowe stanowią cenną pomoc dla pacjentów, jednak istotne jest zrozumienie, że nie zastępują one profesjonalnej terapii mowy prowadzonej przez certyfikowanych logopedów. Ich rola polega na ułatwianiu ćwiczeń i treningów, ale nie mogą pełnić funkcji zastępczej dla terapii specjalistycznych. Warto zaznaczyć, że żadna z omówionych aplikacji nie posiada zdolności diagnostycznych. Ich działanie opiera się na wsparciu w ćwiczeniach, terapii i monitorowaniu postępów, ale nie mają zdolności stawiania diagnoz.

Mimo tych ograniczeń, rosnące zainteresowanie dziedziną terapii mowy w kontekście choroby Parkinsona jest obiecujące. Wspomniane aplikacje są dowodem na rozwijający się obszar badań i innowacji w tej dziedzinie. W miarę postępu technologicznego istnieje szansa na powstanie nowych, bardziej zaawansowanych produktów, które mogą jeszcze skuteczniej wspomagać osoby z chorobą Parkinsona w poprawie jakości mowy i komunikacji.

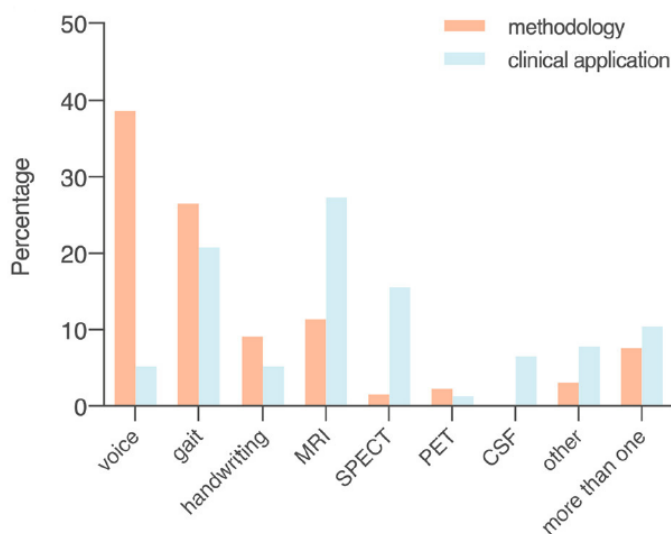
3. Analiza rozwiązań do automatycznej diagnostyki choroby Parkinsona

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 między innymi 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).

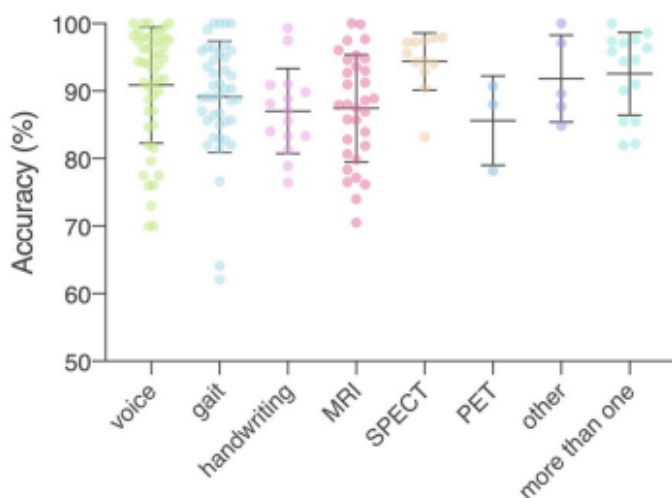
3.1. Rodzaj wykorzystywanych danych

Diagnozowanie choroby Parkinsona stanowi zadanie złożone ze względu na różnorodność objawów, które dotyczą różne aspekty funkcjonowania ciała i umysłu ludzkiego. W związku z tym, techniki uczenia maszynowego wykorzystywane w tym obszarze, także skupiają się na różnych rodzajach danych. Wśród tych źródeł informacji znajdują się wyniki badań obrazowych (np. rezonans magnetyczny - MRI, tomografia komputerowa - SPECT), które wydają się intuicyjne, biorąc pod uwagę zmiany w aktywności mózgu, które można zaobserwować. Niemniej jednak, istnieją także mniej oczywiste metody diagnozy, które budzą duże zainteresowanie w środowisku naukowym, szczególnie w początkowym stadium choroby. Przykłady to analiza głosu, ocena charakterystyki chodu oraz badanie pisma odręcznego.



Rys. 3.1. Wykorzystanie rozwiązań w opracowaniach teoretycznych i zastosowaniach klinicznych w zależności od rodzaju danych (stan na dzień 14 lutego 2020) [13]

Rysunek 3.1 ilustruje zastosowanie wymienionych metod zarówno w teorii, jak i praktyce. Metody oparte na obrazowaniu medycznym wykazują wyraźną przewagę w zastosowaniach klinicznych w porównaniu do kontekstu teoretycznego. Niemniej jednak, to pozostałe metody budzą znacznie większe zainteresowanie ze strony środowiska naukowego. Szczególnie w przypadku analizy głosu, gdzie rozbieżność między teorią a praktyką jest szczególnie znacząca. Przyczyny tego zjawiska zostaną dokładniej rozważone w dalszej części pracy. Co ciekawe, jak przedstawiono na wykresie 3.2, detekcja choroby na podstawie głosu daje bardzo wysokie wyniki, w większości analizowanych artykułów.



Rys. 3.2. Dokładność rozwiązań diagnostycznych w zależności od rodzaju danych (stan na dzień 14 lutego 2020) [13]

Opracowanie dotyczy diagnostyki opartej na analizie głosu, dlatego temat ten zostanie bliżej rozważony. Założeniem dla takich systemów jest zadanie potencjalnemu pacjentowi zadania wokalnego, może to być:

- podtrzymywane samogłoski (ang. *sustained vowels*),
- zadanie diadochokinetyczne (DDK), mogące mierzyć zdolność do wydawania serii szybkich i naprzemiennych dźwięków (syłab),
- czytanie tekstu,
- wypowiedzenie pojedynczego zdania,
- monolog.

Dotychczas nie przeprowadzono badań porównawczych dotyczących wpływu wyboru zadania wokalnego na efektywność klasyfikacji w przypadku choroby Parkinsona (PD). Jedynie w artykule [14] przeprowadzono takie porównanie, przy użyciu cech prozodii. Nie pozwala to jednak na wyciągnięcie obiektywnych wniosków, ponieważ każde z zadań wokalnych może wymagać innego podejścia i metod klasyfikacji. Niemniej jednak, autorzy publikacji [15] przeprowadzili badanie na 200 pacjentach z PD, gdzie dokonano klasyfikacji deficytów mowy na pięć poziomów nasilenia. Oceniono typ (głos, artykulacja, płynność) oraz zakres upośledzenia dla każdego poziomu, korzystając z 2-minutowego fragmentu mowy. Wyniki ukazały, że głos stanowił najczęściej występujący i bardziej nasilony deficyt we wczesnych stadiach choroby. Deficyty artykulacji i płynności pojawiały się później. Wykazano, że upośledzenie artykulacji korelowało z upośledzeniem głosu w fazie "ciężkiej", a w fazie "głębokiej" dominującą cechą była upośledzona artykulacja.

W rezultacie, w kontekście wczesnej diagnostyki choroby, artykulacja i płynność mowy nie wymagają głębokiej uwagi. Koncentrację należy skupić przede wszystkim na cechach głosowych, co sprawia, że wybór podtrzymywanych samogłosek jako zadania wokalnego wydaje się być najlepszym wyborem ze względu na ich stabilność w czasie oraz łatwość wypowiedziania przez pacjenta. To podejście może posiadać potencjał uniwersalności dla różnych języków, co oznacza, że analiza może być stosowana niezależnie od języka ojczystego pacjenta. W efekcie pozwala to na bardziej efektywne i znormalizowane diagnozowanie choroby Parkinsona.

Najczęściej występującymi w opracowaniach samogłoskami są /a/, /e/, /i/ oraz /u/. Aktualnie brakuje albo nie udało się jeszcze ustalić, która z tych samogłosek niesie najcenniejsze informacje z punktu widzenia diagnostycznego. W związku z tym, w ramach badawczej części niniejszej pracy przeprowadzone zostanie takie porównanie dla wybranych metod.

3.2. Metody klasyfikacji

W publikacji [13] dokonano przeglądu wykorzystania metod uczenia maszynowego w automatycznej diagnostyce choroby Parkinsona. Spośród 55 wybranych artykułów opublikowanych do lutego 2020 roku

49 badań osiągnęło średnią dokładność na poziomie 90,9%, wahając się od 70,0% [16, 17] do 100,0% [18, 19, 20, 21]. Eksperymenty były przeprowadzane na różnych bazach danych, co czyni je trudne do wzajemnego porównania. Wykorzystywały różne zadania wokalne, a tym samym różne podejścia do problemu.

W zależności od przyjętego podłoża diagnostycznego dobiera się odpowiednią metodę klasyfikacji. Różne metody okażą się skuteczne w przypadku analizy mowy spontanicznej w porównaniu do mowy kontrolowanej. Dlatego w tej części pracy zostaną przedstawione metody klasyfikacji związane z podtrzymywaniem samogłosek, które to zostały wybrane jako fundament badawczy w ramach niniejszego opracowania.

Najpowszechniejszym podejściem do automatycznej diagnostyki choroby Parkinsona na podstawie głosu jest tzw. tradycyjna inżynieria cech. To proces wydobywania z sygnału mowy charakterystycznych cech, takich jak formanty, jitter, współczynniki cepstralne MFCC, stosunek harmoniczny do szumów (NHR), stosunek harmoniczny do szumów (HNR), wskaźnik migotania Shimmer, częstotliwość podstawowa i inne. Chociaż to podejście jest powszechne, to istnieją także inne, bardziej zaawansowane metody.

Jedną z nich jest wykorzystanie spektrogramów jako podstawy dla procesu klasyfikacji. Ten proces polega na przekształceniu dźwięków mowy na formę wizualną w postaci spektrogramów, które prezentują zmiany w czasie i częstotliwości. Następnie, w celu klasyfikacji, można wykorzystać sieci konwulucyjne, które są specjalnie zaprojektowane do pracy z obrazami. Analizują one strukturę i wzorce w spektrogramach, umożliwiając rozróżnienie cech charakterystycznych dla różnych stanów zdrowotnych. To podejście może przyczynić się do szybszej i bardziej precyzyjnej diagnostyki poprzez analizę cech mowy, które mogą być trudne do wykrycia za pomocą innych metod.

W publikacji [22] dokonano porównania tradycyjnego podejścia opartego na inżynierii cech z nowoczesnym wykorzystaniem zaawansowanych głębokich sieci neuronowych. Wyniki wskazały, że głęboka sieć konwulucyjna AlexNet osiągnęła wrażliwość na poziomie 93,2% i swoistość na poziomie 79,5%, podczas gdy inżynieria cech dawała wrażliwość na poziomie 97,7% i swoistość na poziomie 93,2%. Autorzy sugerują, że połączenie tych podejść może zwiększyć skuteczność. Mimo że podejście oparte na tradycyjnej inżynierii cech dało lepsze wyniki, istnieje potencjał uzyskania lepszych rezultatów poprzez wykorzystanie głębszych sieci, takich jak te z rodzin VGG i ResNet.

Większość publikacji w analizowanym obszarze dotyczy niewielkich zbiorów danych (mniej niż 50 nagrań). Wynika to z małej dostępności otwartych baz danych i może wiązać się z nieobiektywnymi wynikami. W badaniu [23], bazującym na bazie danych 135 osób, wyniki mieściły się między 81,74% a 83,91% dokładności, przy wykorzystaniu konwulucyjnych sieci neuronowych (ang. *Convolutional Neural Networks*, CNN) oraz ELM (ang. *Extreme Learning Machines*). Analiza pokazała, że większa liczba próbek wpływa na lepsze wyniki, a sieć AlexNet miała najlepszą równowagę między rozproszeniem a wydajnością.

W innym badaniu [24] przeanalizowano różne wersje ELM w celu klasyfikacji pacjentów z chorobą Parkinsona na podstawie 135 spektrogramów. Wyniki wskazują, że wielowarstwowe sieci ELM wykazują lepszą wydajność niż jednowarstwowe. Osiągnięto dokładność, oscylującą w okolicach 80%.

Ciekawą analizę opisano w publikacji [25]. Zaproponowano trzy podejścia - pierwsze oparte o transfer learning, wykorzystujące spektrogramy nagrań mowy; drugie, wykorzystujące głębokie cechy wyodrębnione ze spektrogramów mowy za pomocą klasyfikatorów uczenia maszynowego; oraz trzecie, oceniające proste cechy akustyczne nagrań również przy użyciu klasyfikatorów uczenia maszynowego. Wyniki wskazują, że druga propozycja wykazuje obiecujące rezultaty. Zaobserwowano najwyższą dokładność na poziomie 99,7% dla samogłoski *\o* oraz odczytywanego tekstu przy użyciu perceptronu wielowarstwowego. Natomiast dla głębokich cech samogłoski *\i* uzyskano dokładność wynoszącą 99,1% przy użyciu lasu losowego. Metoda bazująca na głębokich cechach wykazuje lepsze wyniki w porównaniu do prostych cech akustycznych i podejść opartych na transfer learning. Zaproponowana metodologia przewyższa istniejące techniki na zbiorze danych PC-GITA w wykrywaniu choroby Parkinsona.

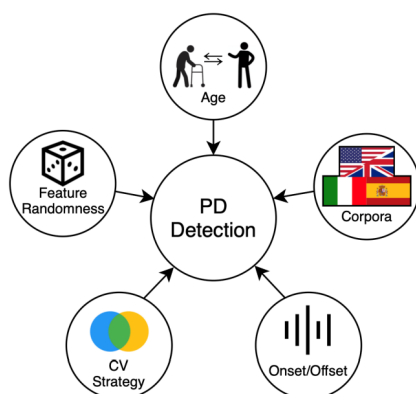
W publikacji [26] zaproponowano wykorzystanie *Spectrogram-Deep Convolutional Generative Adversarial Network* (S-DCGAN) do augmentacji próbek w celu przezwyciężenia ograniczonej ilości istniejących zestawów danych pacjentów oraz próbek głosowych. S-DCGAN generuje spektrogram wysokiej rozdzielczości poprzez zwiększenie warstw sieci, dodanie metody normalizacji spektralnej (ang. *Spectral Normalization*, SN) oraz zastosowanie strategii dopasowania cech. Wybierane są spektrogramy o wysokim podobieństwie i niskim zniekształceniu na podstawie wartości wskaźnika strukturalnej indywidualności (ang. *Structural Similarity Index*, SSIM) oraz stosunku sygnału do szumu maksymalnego (ang. *Peak Signal to Noise Ratio*, PSNR) w celu augmentacji próbek. Wykorzystano model ResNet50 z warstwą globalnego uśredniania (ang. *Global Average Pooling*, GAP) do ekstrakcji cech głosowych i skutecznej klasyfikacji. GAP eliminuje problem nadmiernego dopasowania i przyspiesza optymalizację. Ostatecznie, na zestawie danych Sakar, hybrydowy model S-DCGAN-ResNet50 osiągnął najwyższą dokładność rozpoznawania wzorca głosowego wynoszącą 91,25% oraz swoistość na poziomie 92,5%, co pozwala na precyzyjniejsze różnicowanie między pacjentami z PD a zdrowymi osobami w porównaniu z modelem DCGAN-ResNet50. Rozszerza to środowisko stosowania rozpoznawania wzorca głosowego w dziedzinie medycyny i umożliwia jego uniwersalność w różnych zestawach danych.

W badaniu [27] zaproponowano nowe podejście oparte na wcześniej wytrenowanych głębokich sieciach i długoterminowej pamięci krótkoterminowej (LSTM), wykorzystujące mel-spektrogramy. Zaproponowany model składa się z czterech etapów. W pierwszym kroku usuwany jest szum poprzez zastosowanie rozkładu modalnego wariacyjnego (VMD). W drugim etapie z sygnałów dźwiękowych ulepszonych za pomocą VMD wyodrębniane są mel-spektrogramy. W trzecim etapie wykorzystywane są wcześniej wytrenowane głębokie sieci do ekstrakcji głębokich cech z mel-spektrogramów. W tym celu używane są modele ResNet-18, ResNet-50 i ResNet-101. W ostatnim etapie proces klasyfikacji zachodzi poprzez podanie tych cech jako wejścia do modelu LSTM, który jest zaprojektowany do wyodrębniania informacji sekwencyjnych z wyekstrahowanych cech. Eksperymenty przeprowadzono na zestawie danych PC-GITA. Dokładność uzyskana za pomocą zaproponowanej metody wynosi 98.61%.

3.3. Wyzwania związane z systemami automatycznej diagnostyki

Zainteresowanie systemami do automatycznej diagnostyki choroby Parkinsona na podstawie głosu jest ogromne i wiąże się z nim duże nadzieje. Istnieje jednak duża dysproporcja pomiędzy pracą badawczą a ich wykorzystaniem w rzeczywistym środowisku Rys 3.1. Przyczyn takiego stanu rzeczy jest wiele, a większość z nich związana jest z brakiem usystematyzowanego podejścia do problemu, co utrudnia porównanie rozwiązań, a tym samym rzetelny postęp.

Ostatnie badania wykazały, że możemy wytrenować dokładne modele do wykrywania oznak PD z nagrań audio. Jednakże, istnieją rozbieżności, które są częściowo powodowane przez różnice w wykorzystywanych korpusach lub metodologii. Dlatego autorzy publikacji [28] przeprowadzili analizę, wpływu niektórych czynników na wyniki klasyfikacji. Głównym celem 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 są one najlepszym i najpopularniejszym zadaniem wokalnym w takich systemach. 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. 3.3 oraz omówiono poniżej.



Rys. 3.3. Czynniki wpływające na dokładność detekcji PD na podstawie głosu [28]

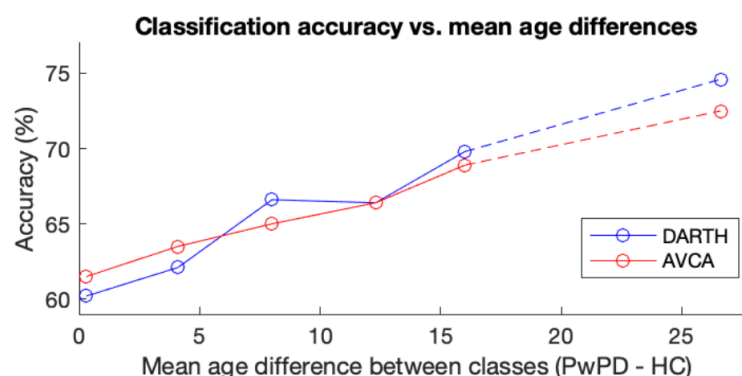
a) Pominięcie aspektu tożsamości mówcy przy konstruowaniu zbiorów treningowych i testowych

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. Drugie podejście to 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, a nie rzeczywiste biomarkery lub wzorce PD. Dlatego też rekomendowana jest technika *subject-wise split*, aby uniknąć zbyt optymistycznych wyników.

b) Niezbalansowanie klas pod względem wieku

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 pomijają informacje o ryzyku, że klasyfikator uczy się wykrywać cechy powiązane z wiekiem, zamiast rzeczywistych wzorców PD. Wyniki eksperymentów w publikacji [28] pokazują, że wraz ze wzrostem różnicy między średnim wiekiem uczestników z PD i HC, dokładność klasyfikacji konsekwentnie rośnie (Rys. 3.4). Na tej podstawie można stwierdzić, że związany z wiekiem wpływ na głos mówców może zaburzać wyniki otrzymywane przez klasyfikator. Dlatego też zaleca się zbalansowanie używanych zbiorów danych, tak aby średnia różnica wieku między tymi dwoma klasami była jak najmniejsza.



Rys. 3.4. Wykres przedstawiający zależność różnicy wieku między klasami a dokładnością klasyfikacji [28]

c) wpływ losowości cech na dokładność klasyfikacji

W publikacji [28] przeprowadzono badania analizujące wpływ losowości cech na dokładność klasyfikacji. Zamieniono cechy obliczone za pomocą DARTH-VAT na losowe liczby, zachowując etykiety i podziały. Wyniki wskazały, że nawet losowe cechy mogą prowadzić do wysokich wyników klasyfikacji (ponad 72%). Efekt ten jest bardziej widoczny w mniejszych korpusach, gdzie różnica między liczbą nagrań a wymiarowością cech ma większy wpływ na potencjalną korelację przypadkową. Badanie pokazuje, że nadmierna liczba cech w stosunku do liczby obserwacji może prowadzić do fałszywie wysokich wyników klasyfikacji nawet przy użyciu losowych cech.

Im większa różnica między liczbą plików a wymiarem wektora cech, tym większe szanse na znalezienie cechy, która losowo koreluje z etykietami klas. To sugeruje, że osiągnięcia klasyfikacyjne powinny być analizowane w kontekście proporcji cech do próbki, aby uniknąć nadmiernie optymistycznych interpretacji wyników klasyfikacji w zastosowaniach medycznych.

d) ograniczenie losowego nadmiernego dopasowania poprzez uwzględnienie zbioru walidacyjnego

Dla mniejszych zbiorów danych, praktyką jest często używanie tylko zbiorów treningowych i testowych podczas krzyżowej walidacji. Jest to podejście, które może prowadzić do wyników zbyt optymistycznych, ponieważ wszystkie wyniki testowe są brane pod uwagę przy wyborze optymalnej konfiguracji modelu. Inną strategią jest wykorzystanie danych treningowych do oceny wytrenowanych modeli i późniejsze przetestowanie najlepszego modelu na zbiorze testowym. Niemniej jednak, to podejście może być niepraktyczne, ponieważ może prowadzić do wyników idealnych (dokładność 100%) na zbiorze treningowym, co jest niepożądane. Aby uniknąć tych problemów, proponuje się wykorzystanie dodatkowego zbioru walidacyjnego [28]. Wybierając model na podstawie wyników walidacyjnych, a następnie testując go na zbiorze testowym, można uniknąć ryzyka nadmiernego dopasowania. Dla mniejszych zbiorów danych, ta strategia może ograniczać dostępną liczbę danych treningowych, co wpływa na wydajność klasyfikacji.

e) wpływ początku i końca nagrań samogłosek na wyniki klasyfikacji

Główną różnicą między korpusami wykorzystywanymi do klasyfikacji choroby Parkinsona (PD) jest obecność fragmentów nagrania oznaczonych jako "onset" i "offset". Niektóre nagrania zawierają te segmenty, podczas gdy inne zostały ich pozbawione, aby zapewnić stabilniejszą fonację, co jest korzystne dla pewnych cech i algorytmów analizy. W celu oceny znaczenia informacji zawartych w obszarach "onset" i "offset" dla klasyfikacji, przeprowadzono eksperymenty porównawcze, wykorzystując nagrania zarówno z fragmentami przyciętymi, jak i nieprzyciętymi [28]. Wyniki tych eksperymentów ukazały, że wyeliminowanie fragmentów początkowych i końcowych wpłynęło negatywnie na dokładność klasyfikacji. To wskazuje na to, że obszary te zawierają istotne informacje artykulacyjne, które mają znaczenie dla procesu klasyfikacji.

f) eksperymenty międzykorporowe a zdolności generalizacyjne

Większość badań dotyczących diagnozowania choroby Parkinsona na podstawie głosu opiera się na jednym, lub ewentualnie kilku (wykorzystywanych niezależnie) korpusach mowy. W tym kontekście często pomija się badanie zdolności klasyfikatorów do ogólnego zastosowania. W artykule [28] przeprowadzono międzykorporowe eksperymenty na bazach danych w językach włoskim i hiszpańskim w celu przetestowania zdolności ogólnych modeli. Skuteczność tych modeli różniła się w zależności od języka zbioru testowego. Może to wynikać z odmiennej różnorodności nagrań, co ma wpływ na stabilność modelu. Drugim możliwym wyjaśnieniem jest to, że głos osób z chorobą Parkinsona może być w różnym stopniu obciążony objawami choroby w zależności od

języka ojczystego lub stopnia zaawansowania choroby. Innymi słowy, w zależności od użytego zbioru danych objawy mogą być nasilone w różny sposób i konieczne jest wzięcie tego pod uwagę tak by zdolności generalizacyjne modelu były jak najwyższe.

Identyfikacja i świadomość wpływu powyższych czynników, pozwala na dostosowaniu przeprowadzanych eksperymentów tak, aby uniknąć wyników zbyt optymistycznych. Usystematyzowanie podejścia do analizy głosu pod kątem diagnostyki choroby Parkinsona przyczyni się do możliwości obiektywnego porównania istniejących i nowych rozwiązań. Tym samym przyspieszy to postęp w tej dziedzinie i uzyskanie optymalnego rozwiązania, które mogłoby zostać wykorzystane w rzeczywistym środowisku.

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.

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:

- zróżnicowanie wzorców mowy: osoby z chorobą Parkinsona mogą różnić się w sposób, w jaki zmiany w mowie wpływają na ich głos. To zróżnicowanie utrudnia stworzenie uniwersalnego modelu, który działałby skutecznie dla wszystkich pacjentów.
- wpływ zmiennych czynników: wpływ na mowę mogą mieć różne czynniki, takie jak zmęczenie, stres czy otoczenie akustyczne. Te zmienne mogą wprowadzać zakłócenia w analizie mowy i utrudniać jednoznaczną diagnozę.
- potrzeba dużej bazy danych: aby stworzyć dokładny system detekcji, konieczne jest posiadanie dużej bazy danych głosów osób z i bez choroby Parkinsona. Uzyskanie takiej bazy danych, która odzwierciedla różnorodność pacjentów i warunki środowiskowe, może być wyzwaniem. Większość publikacji opiera się na bazach danych zawierających około 50 nagrań, co nie jest wystarczająco reprezentatywną próbą.
- wczesne wykrycie i subtelne objawy: wczesne stadia choroby Parkinsona często manifestują się subtelnie, a różnice w mowie mogą być trudne do zauważenia. To może prowadzić do błędnych diagnoz lub niskiej skuteczności systemu.
- weryfikacja i walidacja: Aby narzędzie diagnostyczne oparte na mowie było skuteczne, musi być poddane rygorystycznym testom w rzeczywistych warunkach klinicznych. Weryfikacja i walidacja takiego systemu to skomplikowany proces.
- ograniczenia technologiczne: pomimo postępów w technologii analizy mowy, istnieją nadal ograniczenia w dokładności i precyzji takich systemów. Może to prowadzić do wyników fałszywie pozytywnych lub negatywnych.

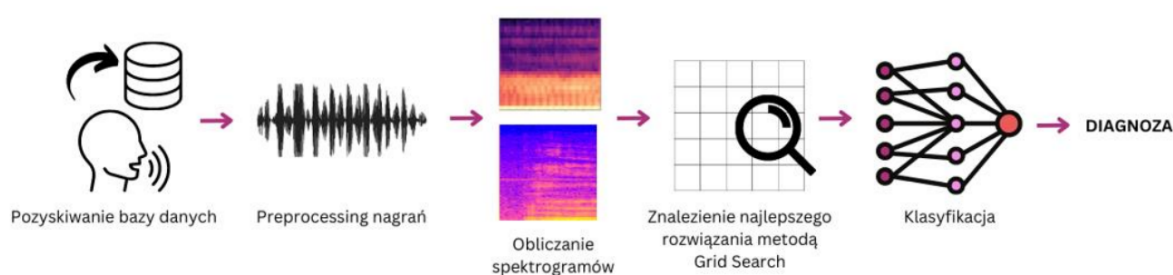
- aspekty etyczne i prywatność: wykorzystywanie danych głosowych do diagnozowania chorób podnosi kwestie związane z prywatnością i etyką. Konieczne jest zagwarantowanie odpowiednich zabezpieczeń danych i zgody pacjentów na wykorzystanie ich informacji w celach medycznych.

Mimo tych wyzwań, prace nad wykorzystaniem analizy mowy do diagnozowania choroby Parkinsona są obiecujące i mogą przyczynić się do poprawy jakości życia pacjentów oraz usprawnienia procesu diagnozy i leczenia. Jednakże przed stworzeniem skutecznego narzędzia diagnostycznego opartego na głosie jest jeszcze wiele pracy do wykonania

W niniejszej pracy podjęta zostanie próba implementacji takiego rozwiązania. Uwzględnione zostaną wszystkie z rekomendacji przedstawionych w artykule [28].

4. Materiał i metoda badawcza

W ramach pracy przeprowadzone zostały badania dotyczące klasyfikacji choroby Parkinsona. Głównym celem było opracowanie efektywnego modelu klasyfikacji binarnej, który może rozróżniać osoby chore na chorobę Parkinsona od osób zdrowych na podstawie analizy sygnałów mowy. Przyjęte podejście opiera się na wykorzystaniu metod przetwarzania sygnałów mowy oraz uczenia maszynowego. Najpierw zebrano dane, w tym nagrania głosowe osób z chorobą Parkinsona (PD) oraz osób zdrowych (HC, ang. *Healthy Controls*). Następnie przeprowadzono analizę przy użyciu różnych ustawień spektrogramów i melspektrogramów. Obejmuje to zmienne parametry takie jak rozmiar okna i przesunięcie okna. Dodatkowo, przeprowadzono badania dotyczące różnych architektur konwolucyjnych sieci neuronowych. Celem jest zbadanie, które architektury sieci i ustawienia spektrogramów dają najlepsze wyniki w klasyfikacji choroby Parkinsona dla poszczególnych samogłosek. Przeprowadzona analiza porównawcza pozwala na lepsze zrozumienie wpływu tych czynników na skuteczność klasyfikacji. W rezultacie, możliwe będzie ustalenie optymalnych ustawień i architektur dla klasyfikacji choroby Parkinsona na podstawie analizy sygnałów mowy. Ponadto zbadano wpływ rozszerzenia zbioru danych o dodatkowe nagrania pochodzące od tych samych osób.



Rys. 4.1. Schemat przyjętej metody badawczej [opracowanie własne]

4.1. Materiał badawczy

Materiałem badawczym w niniejszej pracy magisterskiej są nagrania głosowe samogłosek: /a/, /e/, /i/, /o/ oraz /u/. Baza danych obejmuje nagrania osób zdrowych oraz z chorobą Parkinsona.

Tabela 4.1. Charakterystyka bazy danych [opracowanie własne]

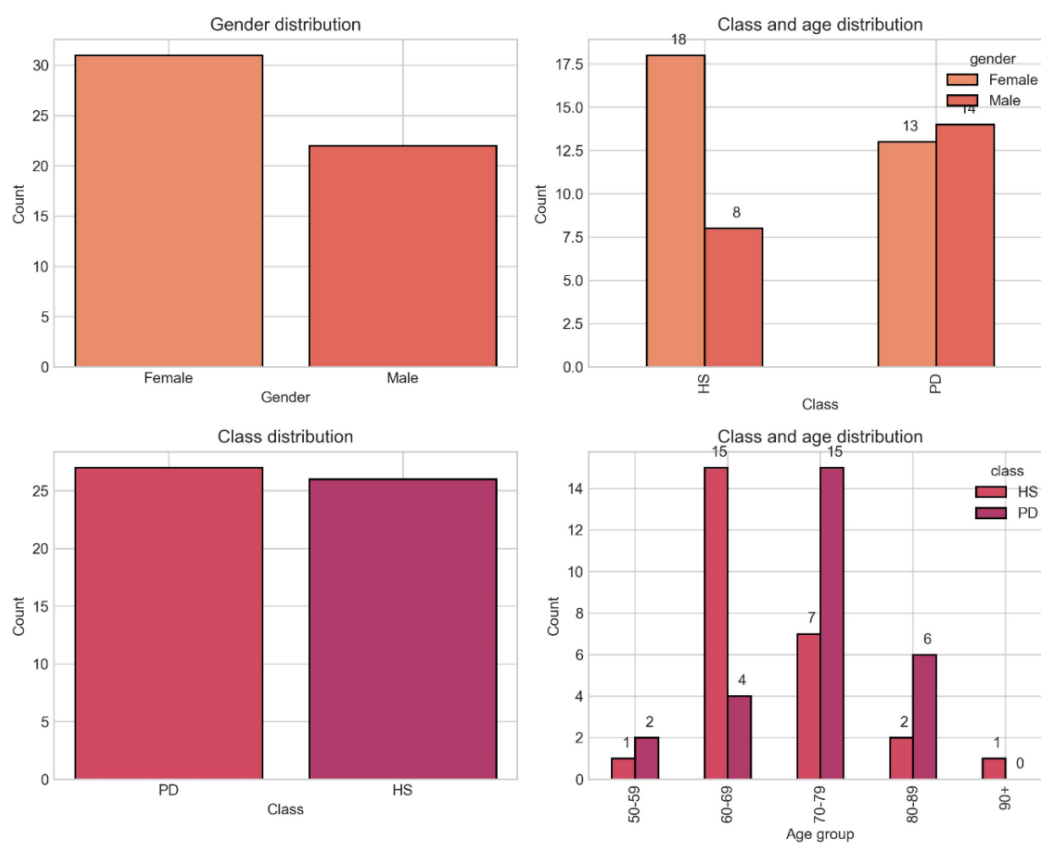
Kategoria	Osoby zdrowe (HC)	Osoby chore (PD)	Razem
Liczba osób	26	27	53
Średnia wieku	60,88 ± 7,98	64,49 ± 8,49	62,68 ± 8,43
Liczba kobiet	18	13	31
Liczba mężczyzn	8	14	22

[Opis źródła nagrań osób chorych] Autorka pracy wykonała nagrania głosu u osób z chorobą Parkinson’a w Krakowskim Szpitalu Specjalistycznym im. Jana Pawła II. Nagrania obejmowały wypowiedzenie samogłosek /a/, /e/, /i/, /o/ oraz /u/ w przedłużonej intonacji przez 27 pacjentów. Pierwszy pomiar występował wtedy, kiedy u pacjenta zdiagnozowano zupełne wysycenie leków łagodzących objawy choroby. Następnie lekarz podawał lek lewodopy. Kolejne pomiary wykonano po 30, 60, 120 i 180 minutach od spożycia leku. Przed każdym pomiarem, lekarz neurolog wykonywał badanie i opisywał stan pacjenta w skali UPDRS.

Nagrania osób zdrowych zostały zebrane w ramach pracy przy wykorzystaniu aplikacji Easy Voice Recorder, która jest programem do nagrywania dźwięku. Przed przystąpieniem do pozyskiwania danych zapoznano się z charakterystyką problemu. Na podstawie literatury wyróżniono płeć, wiek oraz język ojczysty jako czynniki, które mogą wpływać na wyniki klasyfikacji. Zbiór danych powinien być zrównoważony pod kątem tych aspektów, tak by nie dopuścić do sytuacji, gdy model uczy się wzorców nie związanych z chorobą Parkinsona. Ustalono protokół nagrywania, wykluczając osoby poniżej 50 roku życia, palące oraz ze zdiagnozowaną lub podejrzaną chorobą wpływającą na aparat mowy lub korę mózgową (np. choroba Parkinsona, epilepsja, padaczka). Pozyskiwano nagrania jedynie od osób, dla których językiem ojczystym jest polski. W tabeli 4.1 przedstawiono informacje dotyczące bazy danych.

Starano się zachować jak najbardziej zbliżone proporcje wieku i płci pomiędzy grupą pacjentów a grupą porównawczą. Jednak ze względu na specyfikę samodzielnego zbierania danych, nie udało się osiągnąć pełnej zgodności w tym zakresie. Mimo to zapewniono różnorodny zbiór danych, obejmujący zarówno kobiety, jak i mężczyzn w różnych grupach wiekowych. Drobne różnice w liczbie próbek w poszczególnych przedziałach wiekowych nie powinny mieć istotnego wpływu na wyniki klasyfikacji. Szczegółowa charakterystyka zbioru danych została przedstawiona na Rys. 4.2

Każda osoba kwalifikująca się do badania otrzymała zadanie trzykrotnego wypowiedzenia samogłosek /a/, /e/, /i/, /o/ oraz /u/ w odstępach czasowych, utrzymując dźwięk przez co najmniej 2 sekundy. Aby zapewnić jednakowe warunki nagrywania, wyeliminowano hałasy pochodzące z otoczenia oraz wykorzystano pomieszczenia o podobnej akustyce. Wszystkie nagrania zostały zarejestrowane z częstotliwością próbkowania 44 kHz. Czas trwania nagrań samogłosek wynosił od 2 do 5 sekund. Samogłoski były nagrywane trzykrotnie dla każdej z osób.



Rys. 4.2. Charakterystyka bazy danych [opracowanie własne]

Nagrania zostały dokładnie przeanalizowane. Usunięto nagrania zbyt krótkie oraz te, które nie spełniały kryteriów dotyczących jakości. Otrzymana baza danych nadal zawierała nagrania od 27 osób chorych oraz od 26 osób zdrowych, zmieniła się jedynie liczebność nagrań dla poszczególnych samogłosek.

Jednym z głównych celów niniejszej pracy jest przeprowadzenie analizy porównawczej samogłosek pod kątem ich przydatności w klasyfikacji choroby Parkinsona. Aby zagwarantować wiarygodność wyników, niezwykle istotne jest utrzymanie jak najbardziej zbliżonych warunków eksperymentalnych. Kluczowym aspektem tego zagadnienia jest odpowiednie dostosowanie zbiorów danych, ponieważ może mieć znaczący wpływ na ostateczne rezultaty analizy. Początkowo dysponowano zbiorem, w którym na każdą osobę zdrową przypadały trzy nagrania każdej z samogłosek, a na osobę cierpiącą na chorobę Parkinsona jedno. Jednak po oczyszczeniu bazy danych, liczby te uległy zmianie. W związku z tym, podjęto decyzję o ograniczeniu zbioru w taki sposób, aby dla każdej samogłoski uwzględniona była identyczna liczba nagrań pochodzących od poszczególnych osób. Ostateczny skład wybranej bazy danych przedstawiono w tabeli 4.1.2 oraz na wykresie 4.1.3.

4.2. Parametryzacja sygnału akustycznego

4.3. Metody klasyfikacji

4.4. Metody ewaluacji wyników

5. Wyniki badań

5.1. Samogłoska /a/

5.2. Samogłoska /e/

5.3. Samogłoska /i/

5.4. Samogłoska /o/

5.5. Samogłoska /u/

5.6. Zbiorcze podsumowanie wyników

6. Analiza i interpretacja wyników