



VILNIAUS GEDIMINO
TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS
ELEKTRONINIŲ SISTEMŲ KATEDRA

Monika DANILOVAITĖ

**BALSO KLOŠČIŲ BŪKLĖS VERTINIMAS
AKUSTINĖS ANALIZĖS METODAIS**


Magistro baigiamasis darbas

Elektronikos studijų programa, valstybinis kodas 6211EX050
Kompiuterizuotų elektroninių sistemų specializacija Elektronikos
inžinerijos studijų kryptis

VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS
ELEKTRONINIŲ SISTEMŲ KATEDRA

TVIRTINU

Katedros vedėjas


(parašas)

Prof. dr. D. Navakauskas

2020 m. 05 mėn. 29 d.

Monika DANILOVAITĖ

BALSO KLOŠČIŲ BŪKLĖS VERTINIMAS AKUSTINĖS ANALIZĖS METODAIS
SPEECH ANALYSIS BASED ASSESSMENT OF VOCAL FOLDS STATE

Magistro baigiamasis darbas

Elektronikos studijų programa, valstybinis kodas 6211EX050


Kompiuterizuotų elektroninių sistemų specializacija

Elektronikos inžinerijos studijų kryptis

Vadovas

doc. dr. Gintautas Tamulevičius

(pedag. vardas, moksl. laipsnis, vardas, pavardė)

 2020-05-26

(parašas) (data)

Lietuvių kalbos konsultantas

lekt. dr. Anželika Gaidienė

(pedag. vardas, moksl. laipsnis, vardas, pavardė)

(parašas) (data)

Darbas tenkina reikalavimus
Patvirtino dr. A. Gaidienė
2020-05-29

Patvirtina



Rita Raguckienė
Elektroninių sistemų katedra
Administratore

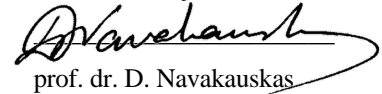
Vilnius, 2020

VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS
ELEKTRONINIŲ SISTEMŲ KATEDRA

Technologijos mokslų sritis
Elektros ir elektronikos inžinerijos mokslo kryptis
Elektronikos inžinerijos studijų kryptis
Elektronikos inžinerijos studijų programa, valst. kodas 6211EX050
Kompiuterizuotų elektroninių sistemų specializacija

TVIRTINU

Katedros vedėjas



prof. dr. D. Navakauskas

2020 m.05 mėn. 05 d.

MAGISTRO BAIGIAMOJO DARBO
UŽDUOTIS

Studentui Monika DANILOVAITĖ

Baigiamojo darbo tema: Balso klosčių būklės vertinimas akustinės analizės metodais
Speech Analysis based Assessment of Vocal Folds State
Patvirtinta 2020 m. 02 mėn. 19 d. dekanų potvarkiu Nr. 47el.
Baigiamojo darbo užbaigimo terminas 2020 m. 05 mėn. 31 d.

Darbo tikslas – Išnagrinėti ir ištirti akustinės analizės metodus balso klosčių funkcinėi būklei vertinti bei patologijai aptikti. Remiantis gautaisiais analitiniais ir kepsperimentiniais rezultatais pasiūlyti funkcinę būklę atspindinčius požymius, suformuluoti rekomendacijas jiems naudoti.

Darbo uždaviniai

1. Išnagrinėti balso klosčių pakitimų įtaką šnekos signalo savybėms. Atlikti analizės metodų, taikomų balso klosčių būklei vertinti apžvalgą ir lyginamąją analizę.
2. Suformuluoti metodiką balso klosčių būklei vertinti. Atlikti eksperimentinę metodikos patikrą.
3. Apibendrinti gautuosius analitinius ir eksperimentinius rezultatus, suformuluoti rekomendacijas balso klosčių funkcinėi būklei vertinti.

Aiškinamojo rašto turinys

Įvadas

Darbo aktualumas ir tikslas

Darbo uždaviniai

Naudoti tyrimo ir analizės metodai

Darbo naujumas ir praktinė nauda

Darbo struktūra

1. Šnekos signalo savybių ir akustinės analizės metodų analitinė apžvalga
1.1. Šnekos signalo generavimas ir savybės

- 1.2. Balso klosčių būklės įtaka šnekos signalo savybėms
- 1.3. Akustinės analizės metodai šnekos signalui analizuoti
2. Šnekos signalo spektrinėmis savybėmis grįsta balso klosčių būklės analizė
 - 2.1. Pagrindinio tono analizė
 - 2.2. Šnekos signalo spektro savybių analizė
 - 2.3. Balso klosčių būklės požymiai
 - 2.4. Balso klosčių būklės identifikavimas
3. Eksperimentinio tyrimo planas
 - 3.1. Eksperimentinių duomenų atranka
 - 3.2. Eksperimentinio tyrimo metodika
4. Balso klosčių funkcinės būklės tyrimo rezultatai
 - 4.1. Analitiniai rezultatai
 - 4.2. Eksperimentinio tyrimo rezultatai
 - 4.3. Rekomendacijos balso klosčių būklei vertinti

Apibendrinimas. Išvados

Literatūra

Priedai

A priedas. Balso klosčių būklės vertinimo analitiniai ir eksperimentiniai rezultatai

B priedas. Pranešimo 22-oje Lietuvos jaunųjų mokslininkų konferencijoje medžiaga

Baigiamojo darbo rengimo konsultantas:

(pedag. vardas, moksl. laipsnis, vardas, pavardė)

Vadovas

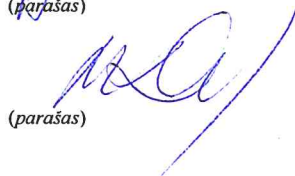


(parašas)

doc. dr. Gintautas Tamulevičius

(pedag. vardas, moksl. laipsnis, vardas, pavardė)

Užduotį gavau



(parašas)

2020 m. 05 mėn. 18 d.

Vilniaus Gedimino technikos universitetas
Elektronikos fakultetas
Elektroninių sistemų katedra

ISBN ISSN
Egz. sk.
Data-.....-.....

Antrosios pakopos studijų **Elektronikos inžinerijos** programos magistro baigiamasis darbas

Pavadinimas **Balso klosčių būklės vertinimas akustinės analizės metodu**
Autorius **Monika Danilovaitė**
Vadovas **Gintautas Tamulevičius**

Kalba: lietuvių

Anotacija

Darbo metu siekta išnagrinėti ir ištirti akustinės analizės metodus balso klosčių funkcinėi būklei vertinti ir patologijoms aptikti. Remiantis analitiniais ir eksperimentiniais rezultatais gauta, kad siūlomos funkcinę būklę atspindinčios akustinės charakteristikos leidžia aptikti patologinius moterų ir vyrų balsus vidutiniu 82,74 % klasifikavimo taiklumu vyrams ir 87,19 % moterims, kai duomenys klasifikuojami į tris klases (sveikų balsų klasė ir dvi patologijų grupių klasės). Charakteristikų imtis šiuo atveju sudaryta iš balso pagrindinio tono (f_0), I , $H2A$, $LNPSD$, HNR , NVB ir viršija aptikimo tikslumą, kai naudojamos kitokios akustinės charakteristikos. Darbe taip pat pateiktos rekomendacijos ir galimybės tolesniems tyrimams.

Prasminiai žodžiai: parametrinė akustinė analizė, balso klosčių būklė, balso klosčių patologijos, mašinių mokymasis, *SVM*, *Naive Bayes*, *DNN*, f_0 , I , $H2A$, $LNPSD$, HNR , NVB .

Vilnius Gediminas Technical University
Faculty of Electronics
Department of Electronic Systems

ISBN ISSN
Copies No.
Date-.....-.....

Master Degree Studies **Electronics Engineering** study programme Master Graduation Thesis

Title **Speech Analysis Based Assessment of Vocal Folds State**
Author **Monika Danilovaitė**
Academic supervisor **Gintautas Tamulevičius**

Thesis language: Lithuanian

Annotation

The aim of the paper is to examine and investigate speech analysis based assessment of vocal folds functional state and detection of pathologies. Based on analytical and experimental results, proposed acoustic characteristics allow detection of functional conditions with average 82.74% accuracy for men and 87.19% for women, when the acoustic data is classified into three classes (healthy voice class and two pathological voice classes). Characteristics in this case were main pitch (f_0), I , $H2A$, $LNPSD$, HNR , NVB and results while using these characteristics have higher accuracy than experiment results with other characteristics. This paper also provides recommendations and opportunities for further research.

Keywords: parametric acoustic analysis, vocal folds functional state, vocal folds pathologies, machine learning, *SVM*, *Naive Bayes*, *DNN*, f_0 , I , $H2A$, $LNPSD$, HNR , NVB .

(Baigiamojo darbo sąžiningumo deklaracijos forma)

VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

Monika Danilovaitė, 20132557

(Studento vardas ir pavardė, studento pažymėjimo Nr.)

Elektronikos fakultetas

(Fakultetas)

Elektronikos inžinerija, EKSm-18

(Studijų programa, akademinė grupė)

BAIGIAMOJO DARBO (PROJEKTO)

SĄŽININGUMO DEKLARACIJA

2020 m. gegužės 31 d.

Patvirtinu, kad mano baigiamasis darbas tema „Balso klosčių būklės vertinimas akustinės analizės metodu“ patvirtintas 2020 m. vasario 19 d. dekanato potvarkiu Nr. 47el, yra savarankiškai parašytas. Šiame darbe pateikta medžiaga nėra plagijuota. Tiesiogiai ar netiesiogiai panaudotos kitų šaltinių citatos pažymėtos literatūros nuorodose.

Mano darbo vadovas docentas daktaras Gintautas Tamulevičius.

Kitų asmenų indėlio į parengtą baigiamąjį darbą nėra. Jokių įstatymų nenumatytų piniginių sumų už šį darbą niekam nesu mokėjęs (-usi).



(Parašas)

Monika Danilovaitė

(Vardas ir pavardė)

TURINYS

| | |
|---|-----------|
| Ivadas | 6 |
| Darbo aktualumas ir tikslas..... | 6 |
| Darbo uždaviniai | 7 |
| Naudoti tyrimo ir analizės metodai | 7 |
| Darbo naujumas ir praktinė nauda | 7 |
| Darbo struktūra..... | 8 |
| 1. Šnekos signalo savybių ir akustinės analizės metodų analitinė apžvalga | 9 |
| 1.1. Šnekos signalo generavimas ir savybės..... | 9 |
| 1.2. Šnekos signalo savybės..... | 10 |
| 1.3. Balso klosčių būklės įtaka šnekos signalo savybėms | 12 |
| 1.4. Akustinės analizės metodai šnekos signalui analizuoti | 16 |
| 2. Šnekos signalo spektrinėmis savybėmis grįsta balso sklosčių būklės analizė | 20 |
| 2.1. Pagrindinio tono analizė | 20 |
| 2.2. Šnekos signalo spektro savybių analizė..... | 23 |
| 2.3. Balso klosčių būklės požymiai..... | 26 |
| 2.4. Balso klosčių būklės identifikavimas | 28 |
| 3. Eksperimentinio tyrimo planas | 31 |
| 3.1. Eksperimentinių duomenų atranka..... | 31 |
| 3.2. Eksperimentinio tyrimo metodika..... | 37 |
| 4. Balso klosčių funkcinės būklės tyrimo rezultatai | 44 |
| 4.1. Analitiniai rezultatai | 44 |
| 4.2. Eksperimentinio tyrimo rezultatai..... | 46 |
| 4.3. Rekomendacijos balso klosčių būklei vertinti..... | 54 |
| Apibendrinimas. Išvados | 56 |
| Literatūra | 58 |
| Priedai | 63 |
| A priedas. Balso klosčių būklės vertinimo analitiniai ir eksperimentiniai rezultatai..... | 63 |
| B priedas. Pranešimo 22-oje Lietuvos jaunųjų mokslininkų konferencijoje medžiaga | 67 |

ĮVADAS

Balso sutrikimai vargina nemažą dalį žmonių – laikoma, kad tarp 3 % ir 9 % JAV gyventojų skundžiasi vienokiomis arba kitokiomis balso sutrikimų patologijomis, bet, nepaisant šios priežasties, tik nedidelė dalis siekia gydymo – vos 0,98 % iš visų registruotų pacientų gydosi dėl balso sutrikimų (Ramig ir Verdoli, 1998; Roy et al 2005). Šios pacientų elgsenos priežastimi galima laikyti diagnostinį procesą, parentą apklausa ir gydytojo apžiūra – turint gerklų problemų, tokie diagnozės būdai nėra optimalūs, nes pacientui sunku naudoti balsą, gerklų apžiūra gali kelti diskomfortą, o gydytojo vertinimas gali būti klaidingas dėl žmogiškojo faktoriaus.

Darbo aktualumas ir tikslas

Siekiant kompensuoti pirminės apžiūros ir apklausos trūkumus, diagnozuojant ligą, papildomai yra naudojami kitokie informacijos šaltiniai, nes siekiant kuo tiksliau nustatyti patologiją, verta turėti kuo daugiau skirtingų duomenų, gautų iš kelių šaltinių. Dėl to balso sutrikimai taip pat diagnozuojami remiantis informacija, surinkta naudojantis:

- ligos istorija;
- laringoskopija;
- kompiuteriniais kiekybiniais ir kokybiniais metodais (pagal *American Speech–Language–Hearing Association* rekomendacijas balso vertinimui, 2015).

Verta paminėti, kad laringoskopija yra invazinis metodas, nes naudojamas specialus prietaisas (laringoskopas) siekiant tiesiogiai apžiūrėti balso klostes – atsiranda didelė minkštųjų audinių traumos tikimybė (52,1 % pagal Chosè Mourao atliktą tyrimą (Mourao et al. 2015)), o remiantis ligos istorija yra ne visada įmanoma suformuluoti išvadas apie ligonio balso būklę, nes balso patologijos gali atsirasti netikėtai ir buvusių ligų istorija jų atsiradimo gali nepaveikti. Įvertinus prieš tai minėtų metodų trūkumus, suformuluotos išvados, kad balso klosčių būklei nustatyti, optimaliausi yra neinvaziniai diagnostikos būdai, kurių metu įmanoma įvertinti didelį kiekį informacijos. Tokie būdai yra kompiuteriniai kiekybiniai ir kokybiniai metodai, įgyvendinami programinės įrangos priemonėmis, kitaip dar vadinami akustine analize. Tačiau, siekiant kompiuterizuoti patologijų atpažinimo procesą, pastebima, kad dėl balso signalo specifikos yra sunku nubrėžti gaires, kaip įvertinti optimalią balso būklę ir pagal balso klosčių funkcinę būklę aptikti patologijas.

Darbo metu išnagrinėti ir ištirti akustinės analizės metodai balso klosčių funkcinę būklę vertinti ir patologijoms aptikti. Remiantis gautais analitiniais ir eksperimentiniais rezultatais, siekiama suformuluoti balso klosčių funkcinę būklę nusakančias charakteristikas ir suformuluoti rekomendacijas kaip jas naudoti.

Darbo uždaviniai

- Išnagrinėti balso klosčių pakitimų įtaką šnekos signalo savybėms. Atlikti analizės metodų, taikomų balso klosčių būklei vertinti apžvalgą ir lyginamąją analizę.
- Suformuluoti metodiką balso klosčių būklei vertinti. Atlikti eksperimentinę metodikos patikrą.
- Apibendrinti gautuosius analitinius ir eksperimentinius rezultatus, suformuluoti rekomendacijas balso klosčių funkcinei būklei vertinti.

Naudoti tyrimo ir analizės metodai

Darbo tikslui ir uždaviniams įgyvendinti atliekama mokslinės literatūros kokybinė analizė, kurios metu gilinamos žinios apie balso klostes, jų patologijas ir akustinės analizės metodus. Remiantis analizės duomenimis, sudaroma akustinių charakteristikų imtis. Akustinių charakteristikų imties pritaikomumas vertinamas pasitelkus eksperimentinį tyrimą. Imtis įvertinama pasitelkus kiekybinį jautrumo ir specifiškumo (angl. *sensitivity and specificity*) analizės metodą sistemos našumui vertinti.

Darbo naujumas ir praktinė nauda

Balso klosčių vertinimas akustinės analizės metodais yra įdomi sritis signalų apdorojimo ir mašinų mokymosi mokslininkams, kurie tiria diagnozavimo automatizavimo galimybes ir tiria balso pokyčius, atsiradusius dėl pasireiškusių įvairaus pobūdžio etiologinių faktorių. Balsų įrašai naudojami, nes akustinis balso signalas teikia didelį kiekį informacijos apie pacientą. Iš balso įmanoma įvertinti paciento amžių, lytį (Gupta et al. 2016), patologijas, tiesiogiai veikiančias balso klostes ir gerklas (pvz., tiesioginiai sužeidimai ir pokyčiai audiniuose (Halava et al. 2014), bet ir patologijas, kurių poveikis balso aparatui yra papildomas simptomas (pvz., išsėtinė sklerozė (Vizza et al. 2017), Parkinsono liga (Braga et al. 2018) ir kitos patologijos). Tačiau pastebima, kad dėl plataus patologijų spektro ir didelio informacijos kiekio balse nėra sutarta dėl bendros metodikos vertinant balso būklę. Pastebimas poreikis patologinių balsų akustinės analizės metodikos standartizacijai ir paprastinimui – pasiekus šiuos tikslus įmanoma taupyti laiką ir resursus (Titze, 1994).

Ši problema yra nagrinėjama Aurelijos Vegienės (2014) disertacijoje. Autorė pastebi, kad dėl balso daugiamatiškumo vertinant patologijas geriau naudotis kompleksiniais tyrimais, pagrįstais įvairiapuse informacija ir patologiją apibūdinančiais požymių rinkiniais (Vegienė, 2014). Ragulo Guptos et al. (2016) apžvalgoje taip pat pastebima, kad nėra rastas optimalus akustinių charakteristikų rinkinys, bet atkreipiamas dėmesys, kad charakteristikų rinkiniai turėtų pasižymėti kompaktiškumu ir būti lengvai interpretuojami, tai yra svarbu siekiant kurti automatizuotus klasifikatorius, gebančius greitai apdoroti informaciją ir gebančius spręsti duomenų trūkumo problemą.

Be anksčiau minėtų problemų, taip pat pastebimos rekomendacijos, kad kuriant minimalius charakteristikų rinkinius, yra svarbus bendradarbiavimas tarp medicinos specialistų ir signalų apdorojimo su mašinų mokymosi specialistais, nes skirtingų disciplinų specialistai vertina patologijas remdamiesi skirtingais metodais. Medicinoje remiamasi įvairaus tipo skalėmis (pvz., *GRBAS*), kuriose vertinamos jutiminės charakteristikos, paremtos estetiniu balso vertinimu. Dėl skirtingo patologijų vertinimo pobūdžio, atsiranda charakteristikų suliginimo problema, nes yra sunku vertinti koreliaciją tarp akustinių charakteristikų ir jutiminių, paremtų estetiniu pojūčiu.

Tyrimuose nagrinėjamas ir kitoks požiūris į šią problemą. Vietoje akustinių charakteristikų rinkinio minimizavimo bandoma klasifikavimą atlikti naudojantis rinkiniais, sudarytais iš visų įmanomų akustinių charakteristikų, apskaičiuotų iš balso įrašo. Šio požiūrio tikslumas tiriamas Zuzanos Dankovičovos et al. (2018), kuri lygina įvairius mašinų mokymosi klasifikavimo algoritmus, apmoko skirtingo dydžio charakteristikų rinkinius. Studijoje nustatyta, kad egzistuoja optimalus charakteristikų rinkinio dydis, kuriam esant, pasiekiamas maksimalus aptikimo koeficientas.

Darbo struktūra

Įvertinus skirtingus požiūrius į akustinių charakteristikų rinkinio sudarymo problemą ir panašaus pobūdžio studijų rezultatus, darbo metu bandoma sukurti optimalų rinkinį, pasinaudojus koreliacija tarp jutiminių charakteristikų ir apskaičiuojamų akustinių charakteristikų. Rinkinys patikrinamas lyginant įvairaus pobūdžio mašininio mokymosi algoritmų rezultatus, kai algoritmai apmokomi sudarytu akustinių charakteristikų rinkiniu ir etaloniniu akustinių charakteristikų požymiu rinkiniu, sudarytu remiantis tyrimais, kurių metu akustinėmis charakteristikomis vertinama balso klosčių būklė ir būdingos patologijos. Gauti rezultatai palyginti su tyrimais, kurių metu balso klosčių patologijos yra vertinamos pasitelkus požymių mažinimo algoritmus. Darbas baigiamas rezultatų apibendrinimu ir rekomendacijomis.

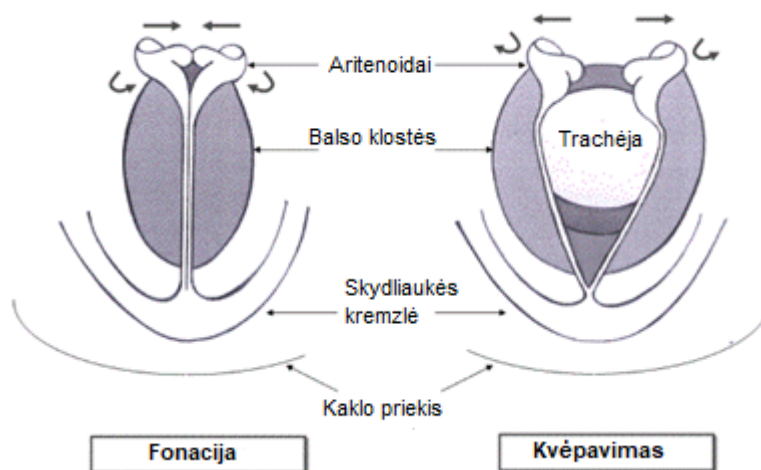
1. ŠNEKOS SIGNALO SAVYBIŲ IR AKUSTINĖS ANALIZĖS METODŲ ANALITINĖ APŽVALGA

Siekiant įvertinti balso klosčių būklę remiantis akustine analize, verta išnagrinėti balso klosčių pakitimų įtaką šnekos signalo savybėms. Šiam uždaviniui atlikti šiame skyriuje nagrinėjamas šnekos signalo generavimas ir tipinės savybės, kuriomis medicinoje apibūdinamas paciento balsas. Atlikus šią užduotį, įmanoma nustatyti, kaip tam tikros patologijos veikia balso klostes. Pagaliau atlikus analizės metodus, taikomų balso klosčių būklei vertinti apžvalgą ir lyginamąją analizę, sudaroma žinių bazė eksperimentiniam planui kurti.

1.1. Šnekos signalo generavimas ir savybės

Žmogaus kalbos aparatas, nors ir sudarytas iš tų pačių komponentų, kaip ir artimiausių rūšies giminaičių primatų (primatų, kaip ir žmonių kalbinis aparatas yra sudarytas iš plaučių, gerklės, balso dėžutės, lūpų ir liežuvio), yra unikalus. Žmogaus smegenys sugeba apdoroti ir atgaminti kalbinius garsus, o maža burna (lyginant su primatais), balso klostės ir ilgos gerklos padeda šiuos garsus artikuliuoti.

Žmogaus kalbos aparato veikimo principas yra toks: iš plaučių gerklomis aukštyn keliaujantis oro srautas virpina balso klostes, kurios raumenų susitraukimais (smegenys nerviniais impulsais reguliuoja šiuos susitraukimus) generuoja tinkamo aukščio ir tono garsą, kuriam pasiekus burnos ertmę, kalbiniai garsai yra formuojami liežuvio ir lūpomis (tuo pat metu burnos ir nosies ertmė padeda garsui rezonuoti). Iš 1.1 pav. matyti, kad balso klostės reguliuoja gerklų ertmės plotį – ertmei esant siaurai, iš plaučių keliaujantis oras stipriai virpins stygas ir generuos aukšto dažnio signalą, o balso stygomis atsipalaidavus, oro srautas pereis per gerklas be pasipriešinimo ir virpesiai nebus generuojami.



1.1 pav. Gerklų ir balso klosčių vaizdas iš viršaus (Iqbal, 2015)

Akivaizdu, kad kalbinio aparato pagrindinė dalis yra balso dėžutės (gerklų) ir balso klosčių organų sistema. Dėl šios priežasties, dauguma su balso patologijomis susijusių sutrikimų veikia būtent šiuos organus (šnekos sutrikimai, tokie kaip mikčiojimas, yra nervinės prigimties (Ward, 2006), dėl to manytina, kad balso kokybės jie itin neveiks – jei vertinamas garso pagrindinis tonas, tonas ir pan. savybės). Matyti, kad balso kokybė tiesiogiai priklauso nuo gerklų ir balso klosčių būklės, todėl šių organų būklės vertinimas yra daugelio tyrimų, susijusių su balso patologijomis objektas ir darytina prielaida, kad šių organų pokyčiai atsispindės ir apskaičiuojamose balso akustinėse charakteristikose.

1.2. Šnekos signalo savybės

Kiekvienam individualiam pacientui yra būdingos unikalios akustinių ir jutiminių charakteristikų vertės, o balso kokybė priklauso nuo tokių faktorių: balso trakto formos, gerklų anatomijos ir išmoktos dedamosios, t. y. įgimtų fizinių faktorių ir išmoktų technikų (pagal *The National Center for Voice and Speech*). Žmogaus ausimi įmanoma lengvai įvertinti kai kurias garsines charakteristikas. Šios charakteristikos yra:

- garso pagrindinis tonas ir rezonansas;
- balso tonas;
- garso intonacija;
- garso stiprumas.

Garso pagrindinis tonas yra charakteristika, apibūdinanti signalo harmonikų dažnį. Nes signalai, kuriuos žmogaus smegenys charakterizuoja kaip „aukštus“ arba „žemus“, yra kompoziciniai – sudaryti iš skirtingų dažnių harmonikų, tai virpesiai, kuriuose vyrauja aukšto dažnio komponentės, bus registruojami kaip aukšti. Dėl šios priežasties skaitmeninių signalų apdorojimo teorijoje signalo aukštis yra dar kitaip vadinamas signalo pagrindiniu dažniu f_0 . Balso klosčių dydis ir elastingumas lemia paciento f_0 ir rezonansą. Žinoma, kad vyrų balso klostės yra ilgesnės nei moterų (vyrų balso klosčių ilgis yra tarp 1,75–2,5 cm, o moterų – 1,25–1,75 cm (Titze, 1994)). Dėl šios priežasties vyrų balsams būdingi žemesni dažniai lyginant su moterimis. Balso klosčių membraninio raumens elastingumas lemia klosčių virpesių rezonansą – organizmui bręstant šie raumenys tvirtėja ir balso stygos virpės žemesniu dažniu (dėl šios priežasties vaikų balsai būna aukštesnio registro nei vyresnių žmonių).

Balso tonas yra charakteristika, apibūdinanti balso aukščio kitimą tariant kalbos žodžius (Yip, 2002). Ši savybė priklauso nuo šnekamosios kalbos ypatumų (egzistuoja tonalinės kalbos, tokios kaip kinų, kur skirtingas žodžio tarimas suteikia žodžiui skirtingą reikšmę ir netonalinės – anglų). Tono svarba suprantant kalbą sukelia problemų šnekos sintezėje. Garso intonacija yra tono charakteristikos atmaina, apibūdinanti kalbėtojo emocijas, sakinio tipą (klausiant intonacija bus kitokia nei teigiant)

ir pabrėžianti svarbias kalbėtojų šnekos signalo dalis. Tonas priklauso nuo šnekamosios kalbos specifikos, bet intonacija priklauso nuo kalbėtojo, dėl to yra suvokiama subjektyviai – ši charakteristika taip pat tampa problematiška kalbos sintezėje.

Garso stipris yra charakteristika, nusakanti akustinio signalo pernešamos energijos kiekį. Žmogaus balso stipris priklauso nuo oro kiekio plaučiuose, vokalo trakto rezonanso ir gerklų raumenų gebėjimo kontroliuoti balso stygą taip, kad šios susitraukinėtų tam tikru dažniu. Laikoma, kad stygomis susiliečiant pusei virpesių ciklo, pasiekama idealus balso stipris) (pagal *The National Center for Voice and Speech*). Akivaizdu, kad praradus balso stygų kontrolę, prarandamas gebėjimas generuoti stiprų garsą.

Nors žmogaus ausimi ir programine įranga įmanoma įvertinti tokias charakteristikas, kaip garso pagrindinis tonas ir stipris ir įvertinti balso toną, intonaciją remiantis spektru, yra sunku vertinti balso estetiką, kuri yra svarbi vertinant balso kokybę. Siekiant išspręsti šią problemą, Ingo Titzė 8-ojoje Balso klosčių patologijų konferencijoje 1994 m. balandį pasiūlė jutiminių charakteristikų rinkinį. Šiomis charakteristikomis įmanoma įvertinti balso savybes. Titzės rinkinyje yra aprašomos tokios charakteristikos kaip balso afoniškumas, bifoniškumas, diplofoniškumas, oringumas, gergždimas, šiurkštumas, kimumas, įtemptumas. Be šių charakteristikų taip pat išskiriamos tokios kaip balso mirguliavimas (angl. *shimmer*), drebulys (angl. *jitter*), pulsavimas (angl. *flutter*), rezonansas, ventrikuliacija, glotiškumas ir kitos. Ne visos iš išvardytų charakteristikų yra naudojamos balso specialistų vertinant balso klostes veikiančias patologijas, nes atsiradus tam tikroms patologijoms, balsas įgyja specifines savybes. Šis rinkinys yra vertingas, nes suteikia balso specialistams sisteminių žodyną balso savybių apibrėžimui.

Medicinoje balso patologijoms vertinti dažnai naudojamos garsinio suvokimo (angl. *auditory perceptual*) skalės. Šiose skalėse vertinamos tam tikros Titzės aprašytos balso charakteristikos (autorius pristatytas rinkinys buvo kulminacija įvairių autorių anksčiau atliktų balso vertinimo tyrimų, todėl kai kurios balso vertinimo skalės buvo sukurtos anksčiau nei Titzės tyrimas). Populiariausiomis yra laikomos *GRABAS* ir *CAPE-V* skalės. *GRABAS* skalėje vertinamas balso šiurkštumas, oringumas, astenija (balso silpnumas) ir įtampa. Šios skalės turinys sudarytas remiantis *GPnotebook* duomenų baze, o *CAPE-V* įvertina balso būklę pagal šiurkštumą, oringumą, įtampą, pagrindinį toną ir stiprį (pagal *The American Speech-Language-Hearing Association* rekomendacijas, išleistas 2002 m.).

Apibendrinant galima teigti, kad atliekant akustinę analizę balso klosčių funkciniai būklėi vertinti, reikia atsižvelgti ne vien į paprastas charakteristikas, tokias kaip garso pagrindinis tonas ir stipris, bet ir sudėtingesnes (pvz., balso šiurkštumas, oringumas ar kimumas). Sudėtingesnės charakteristikos yra lengvai įvertinamos pasinaudojus garsiniu suvokimu, bet yra sunkiai aprašomos matematiniais metodais, nes jų pokyčiai atsispindi balso signalo spektre. Taip pat galima teigti, kad vertinant balso

kokybę, reikia įvertinti balso intonaciją ir toną, nes šios charakteristikos priklauso nuo paciento kalbos ir emocinės būklės – šios charakteristikos gali iškreipti tyrimo rezultatus. Dėl šios priežasties pacientų balsai turi būti įrašomi neutraliomis sąlygomis.

1.3. Balso klosčių būklės įtaka šnekos signalo savybėms

Pacientų grupės atskirtos pagal tokius kriterijus, kaip amžius, lytis ir profesinė veikla pasižymi skirtingų tipų balso patologijomis. Pavyzdžiui, tarp darbingo amžiaus žmonių (19–60 m. amžiaus) yra paplitę funkciniai balso sutrikimai, tokie kaip balso nuovargis, raumenų įtempio disfonija, diplofonija – šie balso sutrikimai diagnozuojami profesionalams, kurių darbo specifika reikalauja intensyvaus balso naudojimo (mokytojai, dainininkai ir pan.). Vyresnio amžiaus žmonėms įprastai yra diagnozuojamos patologijos, tokios kaip presbifonija, funkcinų balso sutrikimai ir net balso klosčių paralyžius (Hagen, 1996; Martins, 2015), o nepilnamečiams dažniausiai diagnozuojami balso klosčių mazgeliai (Martins, 2012; Martins, 2015). Balso sutrikimų spektras pacientų tarpe yra platus, bet įmanoma išskirti bendrinius principus, kurie lemia balso klosčių gebą kurti garsą.

Šiam uždaviniui išspręsti balso klosčių patologijos rūšiuojamos pagal jų kilmės priežastis. Žinant balso sutrikimo priežastį, įmanoma nustatyti kurią šnekos signalo generavimo aparato dalį sutrikimas veikia. Žinant šią informaciją, įvertinama kaip balso kokybė yra paveikiama vieno arba kito sutrikimo. Balso sutrikimų priežastys yra skirstomi į grupes:

- Organinės balso sutrikimų priežastys;
- Funkcinės balso sutrikimų priežastys;
- Psichogeninės balso sutrikimų priežastys.

Šios balso sutrikimų priežastys organizme pasireiškia įvairiai. Supratus balso sutrikimų kilmės priežastis, įmanoma atrinkti balso klostes veikiančias patologijas. Pastebima, kad balso sutrikimai yra sukelti kelių etiologinių faktorių (pagal *American Speech-Language-hearing association* tyrimą, atliktą remiantis Stemplo (2014), Verdolini ir Roseno (2006) tyrimais). Tyrimuose pastebima, kad ventrikuliacinė fonacija gali pasireikšti dėl kvėpavimo takų ligų arba įtempio balso naudojimo galinčio sukelti kitokio tipo disfonijas (Shouren ir Ziyou, 2003). Akivaizdu, kad akustinės analizės metodais nagrinėjant balso klosčių patologijų įtaką balso kokybei privaloma atsižvelgti ir į kitokios prigimties dažniausiai pasitaikančius kartu su tipinėmis balso klosčių ligomis balso sutrikimus, nes šie sutrikimai iškreips duomenų imties charakteristikas.

Organinės kilmės balso sutrikimai papildomai skirstomi į struktūrinius ir neurologinius sutrikimus. Bendru atveju šio pobūdžio sutrikimai veikia balso trakto ir balso klosčių audinius ir jų kontrolę. Organinio pobūdžio balso sutrikimų lentelėje (žr. 1.1 lentelę) nurodyti organinio pobūdžio sutrikimai ir jų kilmė.

1.1 lentelė. Organinio pobūdžio balso sutrikimai

| 1. Sutrikimų rūšis | 2. Sutrikimai | 3. Priežastis |
|--------------------------|---|---|
| Organiniai struktūriniai | Balso mazgeliai, Reinkės edema, glotalinė sklerozė, kvėpavimo papiloma, sarkopenija, polipai ir kt. ataugos ir pan. | Balso klosčių nukrypimai |
| | Cirotienidino arba krikityroido artritas, laringitas, ryklės refliuksas | Gerklų uždegimas |
| | Gerklų trauma | Išorinė trauma, cheminė ekspozicija ir pan. |
| Organiniai neurologiniai | Gerklų nervų paralyžius | Paralyžiuoti gerklų nervai nesiunčia impulsų raumenų kontrakcijai |
| | Spazmodinė disfonija | Spazmai kalbos aparate |
| | Parkinsono liga | Sutrikusi judesių kontrolė lemia prastą balso aparato valdymą |
| | Išsėtinė sklerozė | Pažeistos nervinės skaidulos lema kalbos sutrikimus |

Lentelėje matyti, kad labiausiai balso klostes veikia struktūrinio pobūdžio patologijos, pasireiškiančios per balso klosčių raumens formos ir valdymo sutrikimus. Įvairaus pobūdžio išaugos bendrinu atveju yra vadinamos „dainininko mazgeliais“ (angl. *singer's nodules*) (Watson, 2018). Dėl šių sutrikimų, balsas įgauna tokias jutimines balso charakteristikas: yra kimus, gergždžiantis, pavargęs, oringas, lūžinėjantis, skambantis žemesniu tonu nei įprastai. Balsui įgavus šias charakteristikas, yra ribojamas dainininko balso diapazonas – sunku pasiekti aukštesnes oktavas; ir yra pavojus, kad pacientas praras balsą. Papildomai įvairaus pobūdžio neurologiniai sutrikimai dalyvauja balso klosčių raumenų valdyme, todėl dėl netinkamo šių raumenų susitraukinėjimo, kinta generuojamo šnekos signalo savybės, todėl vertinant balso kokybę, verta atsižvelgti ir į neurologinius sutrikimus.

Daugeliu atvejų funkcinio pobūdžio balso sutrikimai atsiranda dėl įtempto balso aparato naudojimo. Šio tipo balso sutrikimais skundžiasi profesionalai: dainininkai, mokytojai ir panašaus pobūdžio profesijų žmonės. Dėl šios priežasties galima teikti, kad funkcinio pobūdžio balso sutrikimai iš visų balso klosčių sutrikimų yra dažniausiai diagnozuojami ir gydomi, nes veikia žmonės, kurių darbo pobūdis reikalauja intensyvaus balso naudojimo ir jų balso kokybė yra svarbi profesijai. Funkcinio pobūdžio balso sutrikimų lentelėje (1.2 lentelė), nurodyta informacija apie funkcinio pobūdžio balso sutrikimų priežastis (pagal *American Speech-Language-Hearing association*).

Matyti, kad funkcinio pobūdžio balso sutrikimai atsiranda dėl netinkamo balso stygų vartojimo. Balso stygos yra plona raumeninė membrana todėl dėl intensyvaus darbo yra lengvai pažeidžiamos. Pažeidus balso stygų membraną kinta balso klosčių elastingumas, o tuo tarpu ir gebėjimas vibruoti – kis generuojamų vibracijų dažnis.

1.2 lentelė. Funkcinio pobūdžio balso sutrikimai

| Sutrikimų rūšis | Sutrikimai | Priežastis |
|-----------------|---------------------------|---|
| Funkciniai | Fontrauma | Atsiranda dėl bandymo generuoti stiprius garsus (pvz., rėkiant) |
| | Raumenų įtempio disfonija | Įtampa balso aparato raumenyse lemia fonacijos sutrikimus |
| | Ventrikuliacinė fonacija | Atsiranda dėl balso klosčių standumo – nesugebėjimo vibruoti, dėl to garsas generuojamas naudojantis netikrosiomis balso stygomis |
| | Balso nuovargis | Pasireiškia dėl intensyvaus balso vartojimo |

Fontrauma yra balso klosčių sutrikimas, kurio metu balso klostės nebegali susiglausti – dėl šios priežasties kinta generuojamo garso slėgio lygis, pagrindinis tonas, harmonikų sodrumas ir harmonikos–triukšmo santykis. Esant fonotraumai, pradeda veikti kompensuojamieji mechanizmai, dėl kurių balso stygos yra verčiamos susitraukinėti kur kas stipresne jėga – dėl šios priežasties taip pat didėja maksimali srauto mažėjimo norma (angl. *maximum flow-declination rate*), nestabilus srauto amplitudė, bet generuojami akustiniai virpesiai nebūna smarkiai pakitę (Galindo et al. 2017).

Raumenų įtempio disfonija yra terminas, taikomas bendrajai sutrikimų grupei, pasireiškiančiai dėl balso stygų raumenų koordinacijos ir kvėpavimo įpročių sutrikimų. Sutrikus šiems procesams, balso stygų generuojamas akustinis signalas tampa iškreiptu. Šio tipo disfonija gali pasireikšti be balso klosčių anatominės struktūros nuokrypių (pirminė raumenų įtempio disfonija) ir tuo pat metu gali būti diagnozuojama kaip antrinė raumenų įtempio disfonija (kai balso stygose pastebimi anatominiai nuokrypiai) – šiuo atveju disfonija tampa organizmo kompensaciniu procesu, kuriuo bandoma prisitaikyti prie gerklų aparato sužeidimų. Dažniausiai pasireiškiantys simptomai, esant raumenų įtempio disfonijai, yra:

- balso nuovargis;
- nepatikimas balsas;
- žemas, šiurkštus balsas;
- lūžinėjantis balsas;

- oringas balsas;
- kimus balsas.

Be šių simptomų, taip pat pastebima, kad pacientas nesugeba dainuoti ir reikia naudoti daugiau jėgos siekiant generuoti stipresnį garsą (pagal Pistburgo universiteto Ortolaringologijos departamento duomenis).

Ventrikuliacinė fonacija pasireiškia, kai antrinės balso stygos pradeda spausti tikrąsias balso stygas. Šio proceso priežastis yra tikrųjų balso stygų nesugebėjimas vibruoti, nes yra kritiškai prarastas raumenų elastingumas – antrinės balso stygos pradeda veikti kaip kompensacinis mechanizmas, siekiant užtikrinti organizmo gebėjimą generuoti akustinį signalą. Pasireiškus ventrikuliacinei fonacijai, patologinis paciento balsas pasižymi tokiomis charakteristikomis, kaip: žemu garsu, žemu pagrindiniu tonu, šiurkštumu ir įtampa (pagal Kolumbijos universiteto Ortolaringologijos, galvos ir kaklo chirurgijos departamento straipsnį).

Balso nuovargis yra dažnas sutrikimas, atsirandantis dėl intensyvaus balso naudojimo. Šio sutrikimo atsiradimo priežastys gali būti bendras audinių nuovargis ar balso aparato raumenų nuovargis. Skirtingos nuovargio kilmės priežastys lems kokiose balso aparato srityse pasireikš skausmas (esant audinių nuovargiui, pastebimas diskomfortas gerklų zonoje, o esant raumenų nuovargiui – skausmas gerklėje). Be skausmo, balso nuovargis taip pat nustatomas pagal balso pakitimus, kurie yra:

- pastebimas oringas balsas;
- balsas tampa silpnu ir trūkinėjančiu;
- balsas tampa kimiu (pagal *PainAssist inc.* straipsnį).

Akivaizdu, kad nagrinėjant pateiktų balso stygų patologijų simptomus, galima pastebėti panašumų. 1.3 lentelėje pateiktos apibendrintos savybės (pagal Titzės įvardytas savybes), kuriomis pasižymi paciento balsas, esant sutrikimams, pateiktiems 1.1 lentelėje, ir taip pat pasireiškus „dainininko mazgeliams“.

1.3 lentelė. Balso charakteristikos pasireiškiančios esant funkciniam balso sutrikimams

| Sutrikimas | Bendrinės balso charakteristikos | Unikalios balso charakteristikos |
|---------------------------|---|--|
| Fontrauma | Kimumas Šiurkštumas Oringumas Sumažėjęs balso pagrindinis tonas Sumažėjęs balso stiprumas | Didėja maksimali srauto mažėjimo norma, nestabili srauto amplitudė |
| Raumenų įtempio disfonija | | Lūžinėjantis balsas |
| Ventrikuliacinė fonacija | | Balso įtemptumas |

| | | |
|------------------------|--|---------------------|
| Balso nuovargis | | Lūžinėjantis balsas |
| „Dainininko mazgeliai“ | | Lūžinėjantis balsas |

Įvertinus 1.3 lentelėje pateiktą informaciją, rasta, kad esant funkciniam balso sutrikimams, pastebimos charakteristikos, tokios kaip kimumas, oringumas, pakitęs balso pagrindinis tonas, stiprumas. Suformuluota išvada, kad sutrikimams būdingos unikalios balso charakteristikos (žr. 1.3 lentelę) leidžia atskirti vienus funkcinis sutrikimus nuo kitų.

Psichogeninio pobūdžio balso sutrikimai atsiranda dėl psichologinės paciento būklės. Kadangi šio tyrimo objektas yra balso klostės, o psichogeninio pobūdžio balso sutrikimai pasireiškia dėl nervų sistemos sutrikimų veiklos, šio tyrimo metu šie sutrikimai nebus plačiai nagrinėjami. Verta įvertinti psichogeninio pobūdžio sutrikimų įtaką balsui balso patologijų sąryšių. Psichologiniai sutrikimai, lemiantys psichogeninio pobūdžio šnekos signalo generavimo sutrikimus yra sekantys (pagal *American Speech-Language-Hearing association*):

- chroninio streso sutrikimai;
- nerimas;
- depresija;
- konversinė reakcija (t. y. neurologiniai sutrikimai, kuriems neįmanoma rasti fizinės priežasties – pavyzdžiui, histerinė afonija, nebylumas ir pan.).

Mikčiojimas yra tipinis balso sutrikimas, kurio atsiradimo priežastys yra psichogeninio pobūdžio. Mikčiojimą lemia stresas, nerimas ir kiti panašūs psichologiniai sutrikimai. Depresijos metu, paciento balsas pasižymi neįprastu paciento balsui tylumu, bet kitokie balso charakteristikų nuokrypiai nepasireiškia.

Įvertinus rastą informaciją, galima teigti, kad balso klostes labiausiai veikia funkciniai ir organiniai struktūriniai sutrikimai. Pasireiškus šiems sutrikimams, sutrinka balso klosčių elastingumas, pakinta jų forma ir tai atsiliepia jų gebėjimui vibruoti įprastu pacientui dažniu. Pastebėta, kad pasireiškus šioms sutrikimų grupėms, pacientų balsai įgyja tam tikras savybes, tokias kaip: kimumas, šiurkštumas, oringumas. Taip pat pastebima, kad kinta balso pagrindinis tonas (f_0) ir gebėjimas išlaikyti tam tikrą toną – balsas lūžinėja.

1.4. Akustinės analizės metodai šnekos signalui analizuoti

Atliekant signalo akustinę analizę garso signalai (jų tarpe ir šnekos signalai) analizei programinėje įrangoje yra konvertuojami į skaitmeninį signalą, todėl vykdant akustinę analizę nėra svarbios kai kurios bangų charakteristikos (pvz., bangos greitis) ir signalo akustinė analizė tampa

skaitmeninių signalų apdorojimo uždaviniu. Dėl to signalai yra analizuojami diskretaus laiko ir dažnių srityse, siekiant rasti savybes, teikiančias skirtingo tipo informaciją arba yra įvertinami pagal tam tikrus analitinius modelius.

Yra keli požiūriai į akustinę analizę: parametrinis ir neparametrinis. Remiantis neparametriniu požiūriu, garsas analizuojamas pasitelkus bangų analizę, granulinę analizę (angl. *granular analysis*) ir tiesinės prognozės metodą (Kostek, 1997) ir kitus metodus (pvz., Keisukė Imotas ir Nobutaka Onas 2016 m. Europos signalų apdorojimo konferencijoje pristatė gyvo garso analizę remiantis neparametriniu Bajeso modeliu). Remiantis parametriniu požiūriu, analizuojamas signalo dažnių spektras ir formantės arba pasirenkamos tam tikros charakteristikos iš laiko ir dažnio sričių – kaip jau minėta įvade, problema tampa charakteristikų svarbos įvertinimas (Chen, 1995; Kostek, 1995). Įvertinimas gali būti atliekamas pasitelkus statistinius arba mašinių mokymosi algoritmus (Kostek, 1995).

Pastebima, kad atliekant paciento balso kokybės įvertinimą dažnai remiamasi parametrine akustine analize. Dranšri Gundžiaivati et al. 2018 m. atliktoje apžvalgoje nagrinėjama, kokiomis charakteristikomis remiantis yra vertinama dainininkų balsų kokybė (ši profesinė grupė yra svarbi, nes dėl intensyvaus balso naudojimo šie žmonės yra funkcinų balso klosčių sutrikimų rizikos grupėje). Apžvalgoje nagrinėjami tyrimai, paskelbti *PubMed/MEDLINE*, *CINAHL* ir kt. bazėse. Autoriai pastebėjo, kad nėra apibrėžto daugiausiai informacijos teikiančio akustinių charakteristikų rinkinio, o garso įrašų turinys yra įvairus. Apžvalgoje minimų balso kokybės tyrimų duomenys yra įvairaus tipo – akustinės charakteristikos yra apskaičiuojamos remiantis dainuojančiu balsu, fonuojamais balsiais, tariamomis frazėmis ir pan. Akivaizdu, kad įrašo turinys lems skaičiuojamų charakteristikų vertes ir pritaikomumą, kai siekiama atpažinti tam tikras patologijas ar atpažinti specifines balso savybes. Remiantis apžvalgoje atliktu rekomendacijų, vertinimo, plėtros ir darbinės grupės nustatymo vertinimu (angl. *The Grades of Recommendation, Assessment, Development, and Evaluation Working Group (GRADE)*), buvo atrinkti 8 tyrimai, kurių įvertinimas yra aukščiausias (71 %). Šių tyrimų trumpa apžvalga pateikta 1.4 lentelėje.

1.4 lentelė. Tyrimų, kurių objektas yra balso kokybė, apžvalga (Gunjawate et al. 2018)

| Tyrimo autoriai | Garso įrašo turinys | Tirtos charakteristikos |
|------------------------|----------------------------|---|
| Prakup (2012) | /a/ fonacija | pagrindinis tonas (f_0), drebulys, mirguliavimas, balso stiprumas (I) |
| Guzman et al. (2013) | /a/ fonacija | harmonikos–triukšmo santykis (HNR), drebulys, mirguliavimas |

| | | |
|--|---|---|
| Echternach ir Richter (2012) | /a,i,u,o,ae/ dainavimas įvairiomis technikomis | santykinis drebulio vidurkis (<i>RAP</i>), drebulys, mirguliavimas |
| Arunachalam et al. (2014) | Skirtingais greičiais dainuojamos trys oktavos dainininkui įprastame diapazone; skaitomas tekstas skirtingais garsumo lygiais | Balso diapazonas dainuojant, disfonijos sunkumo indeksas (<i>DSI</i>) |
| Gunjawate, Aithal, Guddattu ir Bellur (2015) | /a/ balsės fonacija madhya, taar Saptak/Sthayi dialektais | pagrindinis tonas (f_0), aukščiausias pagrindinis dažnis (<i>Fhi</i>), žemiausias pagrindinis dažnis (<i>Flo</i>), balso diapazonas, drebulys, santykinis drebulio vidurkis (<i>RAP</i>), tono triukšmo koeficientas (<i>PPQ</i>), glotnintas tono triukšmo koeficientas (<i>SPPQ</i>), virtualus kontūras (vFO), triukšmo ir signalo santykis (<i>NHR</i>) |
| Awan ir Ensslen (2010) | /a/ fonacijos keičiant pagrindinį toną; kiek įmanoma tylesnė /a/ fonacija, skanduojama 1, 2, 3, 4; 2–3 s palaikoma /a/ fonacija; kiek įmanoma ilgiau palaikoma /a/ fonacija | disfonijos sunkumo indeksas (<i>DSI</i>), pagrindinis tonas (f_0), balso stiprumas (<i>I</i>), drebulys, maksimalus skambėjimo laikas |
| de Almeida Bezerra et al. (2009) | Dainuojama daina | vibrato dažnis, vibrato laipsnis |
| Omori et al. (1996) | Palaikoma /a/ fonacija; dainuojama /a/ pacientui būdingu tonu ir garsumu | dainavimo galios santykis (<i>SPR</i>) |

Matyti, kad siekiant įvertinti paciento balso kokybę, naudojami įvairaus turinio balsų įrašai ir akustinės charakteristikos. Pastebima, kad yra paplitęs balsės /a/ fonacijų naudojimas. Remiantis balsių fonacijomis išvengiama kalbos tono, intonacijos ir kalbos specifikos įtakos balsui. Šią prielaidą paremia Briuso Gerato et al. (2016) tyrimas, kuriame įvertinta ar balso įrašo turinys turi įtakos skaičiuojamoms akustinėms charakteristikoms. Studijoje rasta, kad lyginant charakteristikas, gautas iš kalbos įrašo ir charakteristikas apskaičiuotas iš balsių fonacijų įrašų, skirtumų harmonikose nerasta, kai bandoma įvertinti harmonikas ir harmonikos–triukšmo santykį (*HNR*).

Pastebima, kad tyrimuose, kuriuose nagrinėjamas balso patologijų aptikimas akustiniais parametrais, vyrauja kitoks požiūris į duomenis. Vietoje balso savybių įvertinimo remiantis specifinėmis charakteristikomis, klasifikacija atliekama pasitelkus didelį kiekį charakteristikų, kuriais būna apmokyti įvairūs klasifikatoriai. Pavyzdžiui, Dankovičovos et al. 2018 m. tyrime remtasi 1560

charakteristikų imtimi. Charakteristikos gautos įvertinus balsių /a/, /i/ ir /u/ fonacijų akustines charakteristikas. Charakteristikų kiekis buvo mažinamas pasitelkus požymių atrankos algoritmą, o atrinktos svarbiausios charakteristikos yra:

- penkių taškų amplitudės triukšmo koeficientas (*APq5*);
- absoliutusias skirtumas tarp kaimyninių periodų amplitudžių (*DDP*);
- santykinis drebulio vidurkis (*RAP*);
- minimalus spektro nuokrypis;
- melodijos-dažnio cepstraliniai koeficientai (*MFCC-6*);
- penkių taškų periodo triukšmo koeficientas (*PPQ5*);
- lokalus drebulys (*LOCAL*).

Šio tyrimo rezultatus palyginus su kituose tyrimuose naudojamomis charakteristikomis, pateiktomis 1.4 lentelėje, rasta, kad bandant automatizuoti patologijų klasifikaciją, remiantis akustinės analizės metodais, naudojamas didelis kiekis charakteristikų, o charakteristikų atranka vykdoma pasitelkus įvairaus tipo algoritmus (Selvakumari ir Radha, 2017). Skirtingi požiūriai į akstinių charakteristikų atrankos problemą atsirado, nes ši problema buvo nagrinėjama skirtingų mokslo šakų mokslininkų (balso specialistai domisi tiksliais charakteristikomis, koreliuojančiomis su balso savybėmis, o signalų ir kompiuterių mokslo specialistai remiasi dideliu duomenų kiekiu ir pažangiais matematiniais metodais informacijos optimizacijai ir klasifikacijai).

Galima teigti, kad tyrimuose vyrauja parametriniai akustinės analizės metodai, bet nepastebima, kad visuose tyrimuose būtų naudojamas tas pats charakteristikų rinkinys, tinkamas nusakyti balso klosčių būklę. Įvertinus šią informaciją, rasta, kad atliekant patologinių balsų akustinę analizę, nėra bendros metodikos, tinkamoms visoms duomenų imtims ir atvejams (Gupta, 2016). Šią prielaidą patvirtina 1.4 lentelės informacija, kurioje matyti, kad studijose, kurių tyrimo objektas yra tas pats (pacientų balso kokybė), naudojama didelė įvairovė akstinių charakteristikų. Dėl to bandant spręsti šią problemą, remiamasi įvairaus tipo metodikomis, pvz., akstiniai parametrai yra analizuojami pasitelkus įvairius savybių atrankos algoritmus ar yra įvertinami pasitelkus balso ekspertų nuorodas (Gupta, 2016). Pastebėta, kad bandant klasifikuoti balso klosčių patologijas remiantis įvairaus pobūdžio klasifikavimo algoritmais, nebandoma įvertinti jutiminių savybių (pvz., oringumo, šiurkštumo ar kimumo) ir akstinių charakteristikų koreliacijos, tokiu būdu sukuriant optimalų charakteristikų rinkinį. Dėl šios priežasties reikia įvertinti, dėl kokių priežasčių studijose nėra taikomas šis metodas.

2. ŠNEKOS SIGNALO SPEKTRINĖS SAVYBĖS GRĮSTA BALSO SKLOSČIŲ BŪKLĖS ANALIZĖ

Atlikus šnekos signalo savybių ir akustinės analizės metodų analitinę apžvalgą, pastebėta, kad balso savybėms nusakyti ir vertinti yra naudojamos įvairaus pobūdžio metodikos. Dėl šios priežasties, sukurtos įvairios skalės, naudingos balso patologijų specialistams, siekiantiems įvertinti balso disfoniją (disfonija yra bendrinis terminas, taikomas diagnozuojant balso klosčių patologijas) (pvz., *GRBAS*, *CAPE-V*). Pastebėta, kad šį darbą turi atlikti profesionalai, o tai palieka vietos žmogiškai klaidai (pavyzdžiui, Džiovano Rizo et al. (2016) studijoje nustatyta, kad apmokyto medicinos personalo diagnozė siekia 80,6 % tikslumą, jei bandoma diagnozuoti Parkinsono ligą, dėl to galima daryti prielaidą, kad medicinos personalo diagnozės tikslumas patologijoms, veikiančioms balso klostes nesiekia 100 %).

Kaip pastebėta atlikus analitinę apžvalgą, balso savybių įvertinimas remiantis akustinės analizės metodais nėra lengvas uždavinys, nes egzistuoja didelis kiekis įvairaus pobūdžio akustinių charakteristikų, kurių svarbumą ir koreliaciją su balsinėmis savybėmis yra sunku įvertinti. Dėl šios priežasties, skyriuje išsamiau nagrinėjamos svarbiausios akustinės charakteristikos ir patologinio balso, atsiradusio dėl balso klosčių sutrikimų spektrinės savybės. Įvertinus gautą informaciją, įmanoma įvertinti koreliacijos tarp jutiminių savybių, pasireiškiančių atsiradus balso klosčių patologijoms ir akustinių charakteristikų tikslumą. Koreliacijos vertės gaunamos atlikus medicininių tyrimų, kurių tikslas yra akustinių charakteristikų ir jutiminių balsinių savybių sąryšio tyrimas ir vertinimas. Galiausiai, remiantis studijų apžvalga, įvertinami ir atrenkami geriausi matematiniai metodai balso klosčių būklės duomenų vertinimui.

2.1. Pagrindinio tono analizė

Balso pagrindinis tonas f_0 yra svarbi charakteristika, apibūdinanti paciento balso specifiką ir gebą rezonuoti tam tikromis formantėmis. Yra žinoma (žr. 1.2 poskyrį), kad balso pagrindinis tonas priklauso nuo paciento balso klosčių fizinių parametrų ir elastingumo (dėl to įmanoma atskirti vyrų ir moterų balsus ir pat įvertinti paciento amžių – žr. 1.2 poskyrį), o įvairios balso klostes veikiančios patologijos lemia klosčių gebėjimą vibruoti pacientui būdingu dažniu (žr. 1.3 poskyrį) – dėl šios priežasties tyrimuose atliekant akustinę analizę balso klosčių būklės vertinimui yra itin svarbi pagrindinio tono analizė (žr. 1.4 poskyrį).

Idealiu atveju, signalo pagrindinis tonas yra apskaičiuojamas remiantis dažnio atvirkštine priklausomybe nuo signalo periodo (t_s), bet šis metodas nėra tinkamas apskaičiuoti balsinio signalo pagrindinį toną, nes tokio tipo akustinis signalas neturi pastovaus periodo ir nagrinėjant balsinį signalą, reikia įvertinti fonacijų energijos kintamumą. Dėl šios priežasties, sukurti įvairaus pobūdžio

metodai pagrindinio dažnio apskaičiavimui. Kai kurie metodai įvertinti Alaino de Čevigné ir Hideki Kavaharos (2002) tyrime, kuriame autoriai siūlo pagrindinio tono (f_0) įvertinimą remiantis modifikuotu autokoreliacijos metodu. Autokoreliacija yra sistemos analizės funkcija, skirta aptikti periodiškumams ir pasikartojimams signalė. Autoriai lygina šį modifikuotą autokoreliacijos metodą su jau įgyvendintais akustinės analizės programinėje įrangoje metodais. Kavahara ir Čevigné tyrime teigia, kad siūlomas metodas pasižymi mažiausia bendrąja paklaida, bet Sofijos Strombergson (2016) apžvalgoje, kurioje yra vertinami pagrindinio tono įvertinimo metodai, suformuluotos išvados, kad Kavaharos metodas nėra populiarus, o daugiausiai naudojamas yra programinėje įrangoje Praat įgyvendintas Paulo Boersmos 1994 m. siūlytas algoritmas, kurio pagrindas yra autokoreliacija (Kavaharos metodu rėmėsi 5 publikacijos, o Boersmos metodu paremta Praat programa 80). Autorė pastebi, kad Boersma metodas pasižymi geresniais rezultatais, kai garsiniame signalė aptinkami tylos segmentai (pvz., trūkinėjant balsui), o palyginant gauta, kad F_0 kadro paklaida (angl. *F0 Frame Error (FFE)*) naudojant Kavaharos metodą yra kur kas didesnė – 35,38 %, kai Praat programoje naudojamas metodas pasižymi 6,20 % paklaida (šio tipo paklaida yra svarbi, nes yra tinkama bendriniam našumui įvertinti (Chu ir Alwan, 2009)). Įvertinus apžvalgoje gautą informaciją, nuspręsta naudotis Boersma metodu, siekiant įvertinti balso pagrindinį toną f_0 .

Įprastai autokoreliacija skaičiuojama lyginant signalo kadro koreliaciją su supančias kadrais:

$$R_n(k) = \sum_{m=0}^{N-1-k} [x(n+m)w(m)][x(n+m+k)w(k+m)], \quad (2.1)$$

kur $R_n(k)$ yra n -tojo segmento, priklausančio diskretizuotam signalui $x(n)$, trumpalaikė autokoreliacija; $w(m)$ yra langinė funkcija. Boersma metodas buvo sukurtas siekiant išspręsti problemą, atsiradusią dėl balso fonacijų apdirbimo langinėmis funkcijomis siekiant pašalinti triukšmų įtaką pagrindinio tono apskaičiavimui. Formulėje signalo kadras yra dauginamas iš langinės funkcijos. Autorius pastebėjo, kad apskaičiuojant autokoreliaciją iš langine funkcija apdirbto signalo, autokoreliacijos funkcija turi būti padalinta iš langinės funkcijos autokoreliacijos. Atlikus šią transformaciją neprarandami svarbūs žemesni dažniai – kitokiu atveju, vietoje F_0 formato, yra aptinkamas F_1 , t. y. neteisingai aptinkamas signalo pagrindinis tonas (žr. veiksmų seką 2.1 pav.).

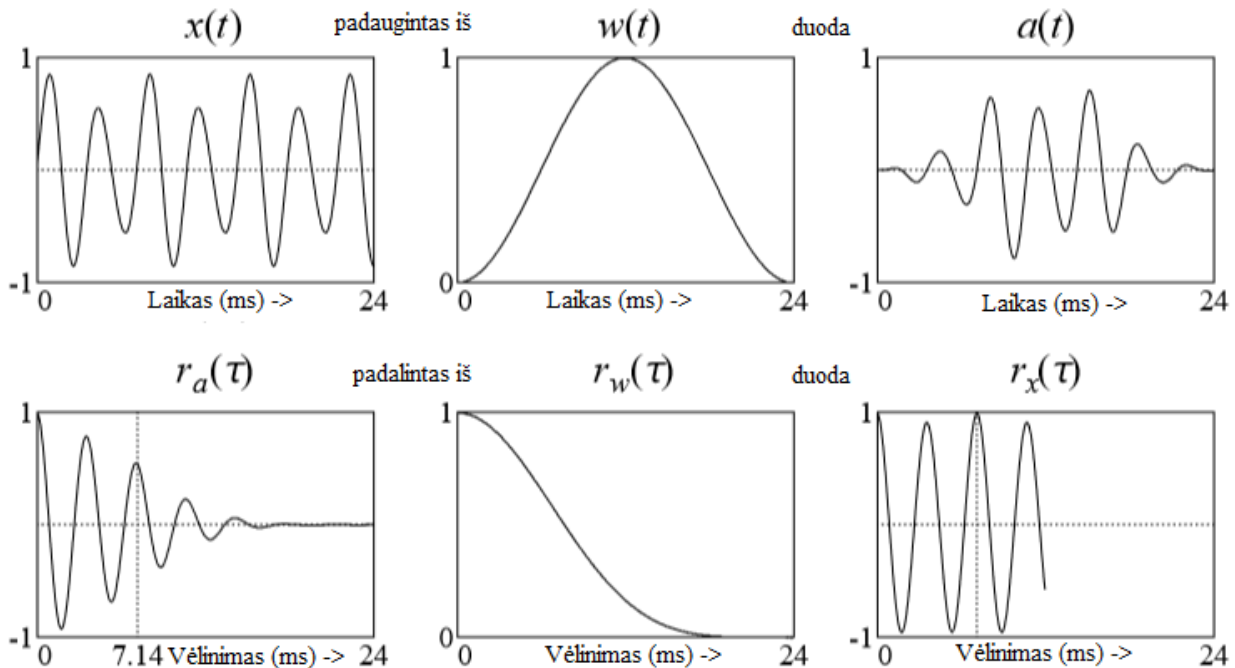
2.1 paveiksle matyti, kad garsinį signalą $x(t)$ padauginus iš langinės funkcijos $w(t)$ gaunama $a(t)$. Pagal šią funkciją apskaičiuojama autokoreliacija $r_a(\tau)$. Įvertinus langinės funkcijos autokoreliaciją $r_w(\tau)$, diskretizuota trumpalaikė autokoreliacijos funkcija $r_x(\tau)$ yra:

$$r_x(\tau) \approx r_a(\tau)/r_w(\tau). \quad (2.2)$$

Įvertinus Boersma'os tyrimą, gauta ši formulė, kuria įvertinamas balso pagrindinis tonas:

$$f_0 = \frac{1}{\frac{\text{lagmax}}{f_s}}, \quad (2.3)$$

kur f_0 – balso pagrindinis tonas, Hz; f_s – balso signalo diskretizavimo dažnis, Hz; lag_{max} – vėlinimo vertė, kai autokoreliacija yra maksimali, ms – gauta iš $r_x(\tau)$. 2.1 pav. matyti, kad tikrasis maksimumas po langinės funkcijos įvertinimo yra pasiekiamas esant 7,14 ms vėlinimui (čia aptinkama F_0 formantė).



2.1 pav. Autokoreliacijos įvertinimas remiantis garso signalu, apdorotu langine funkcija (Boersma, 1993)

Iki langinės funkcijos autokoreliacijos įvertinimo ši informacija nebuvo aptikta, todėl neteisingai spėjama, kad maksimalus f_0 pasiekiamas esant 3,57 ms vėlinimui (pirmojo formantės). Boersma gavo šias vertes naudodamas *Hanning* langinę funkciją ir sinusoidinį signalą, sudarytą iš dviejų dažnių formančių. Autorius pastebi, kad siekiant geriausių rezultatų, remiantis atrankos algoritmais signalo maksimumai turėtų būti įvertinami tam tikrame dažnių ruože (pvz., *Brent* algoritmu iš Press ir kitų 1989 tyrimo). Dėl šios priežasties, autorius *Praat* programoje įgyvendino formančių atranką remdamasis maksimalaus dažnio ir svorio koeficientų poromis kiekvienam kadrai – dažnių ruožo režiai tampa reikalingi pagreitinti algoritmo atrankos darbą, nes praleidžiami dažniai, netelpantys į dažnių ruožo režius.

Atlikus balso pagrindinio tono (f_0) tyrimų analizę, rasta, kad šiai charakteristikai apskaičiuoti yra sukurta įvairių metodų, nes tenka įvertinti balso signalo periodo (t_s) nepastovumą ir formančių energijos kitimą. Nustatyta, kad dažniausiai yra naudojamas Boersma 1994 m. pasiūlytas autokoreliacijos metodas, kuris yra pranašesnis už kitų autorių siūlomus, nes yra patvaresnis (veikia taip pat gerai tiek žemiems, tiek aukštiems dažniams – Boersma 1994) ir atsparesnis signalo triukšmams ir pat gerai veikia esant tylos segmentams įrašė. Įvertinus Boersma metodo specifiką,

suformuluota išvada, kad bandant nustatyti balso pagrindinį toną (f_0), privalu įvertinti balsų dažnių ruožų skirtumus, atsiradusius dėl vyrų ir moterų balso klosčių fizinių parametrų. Autoriaus rekomendacijose *Praat* programai, minima, kad vyrų balsams geriausia nustatyti 75–300 Hz dažnių ribas, o moterų 100–500 Hz. Matyti, kad balsų pagrindinio tono dažnių ruožai persidengia, dėl šios priežasties, vertinant paciento balsą, reikia atkreipti dėmesį į lytį, nes neįprastai žemi, disfoniški moteriški balsai gali būti vertinami kaip sveiki vyriški balsai.

2.2. Šnekos signalo spektro savybių analizė

Šnekos signalo spektras teikia daug informacijos apie balsą. Naudojantis šnekos signalo spektrogramomis įmanoma vizualiai atskirti garso formantes ir tariamus garsus (žr. 2.2 pav. matyti pažymėtos formantės ir tariami balsiai – spektrogramos skiriasi) ir specifinės dažnių srities charakteristikos padeda nusakyti garso signalo melodiją, ritmą ar natas. Akivaizdu, kad pasireiškus balso klosčių pokyčiams, balso pokyčiai turi atsispindėti ir signalo spektre (žr. 2.3 pav.). Nors šiuos pokyčius yra įmanoma vertinti naudojantis matematiniais metodais (pvz., tiesinės prognozės metodą signalo spektro voko aptikimui), tyrimuose, kurių objektas yra balso klosčių sutrikimai, dažniau naudojamos parametrinės akustinės savybės, skaičiuojamos pasitelkus balso signalo išraiška laiko srityje (kaip minėta 1.4 poskyryje).

Šnekos signalo spektro savybėms apskaičiuoti praktikoje signalas yra transformuojamas iš laiko srities į dažnių sritį, pasitelkus greitąją Furjė transformaciją diskretiesiems signalams (pagal Boersma instrukciją *Praat* programai):

$$X(f) = \int_0^T x(t)e^{-2\pi ift} dt \quad (2.3)$$

kur $X(f)$ – signalo spektras; $x(t)$ – signalo funkcija laiko srityje. *Praat* programoje, atsižvelgus į atminties resursų apribojimus, dažninės komponentės yra skaičiuojamos žingsniu Δf :

$$\Delta f = \frac{1}{N\Delta t}, \quad (2.4)$$

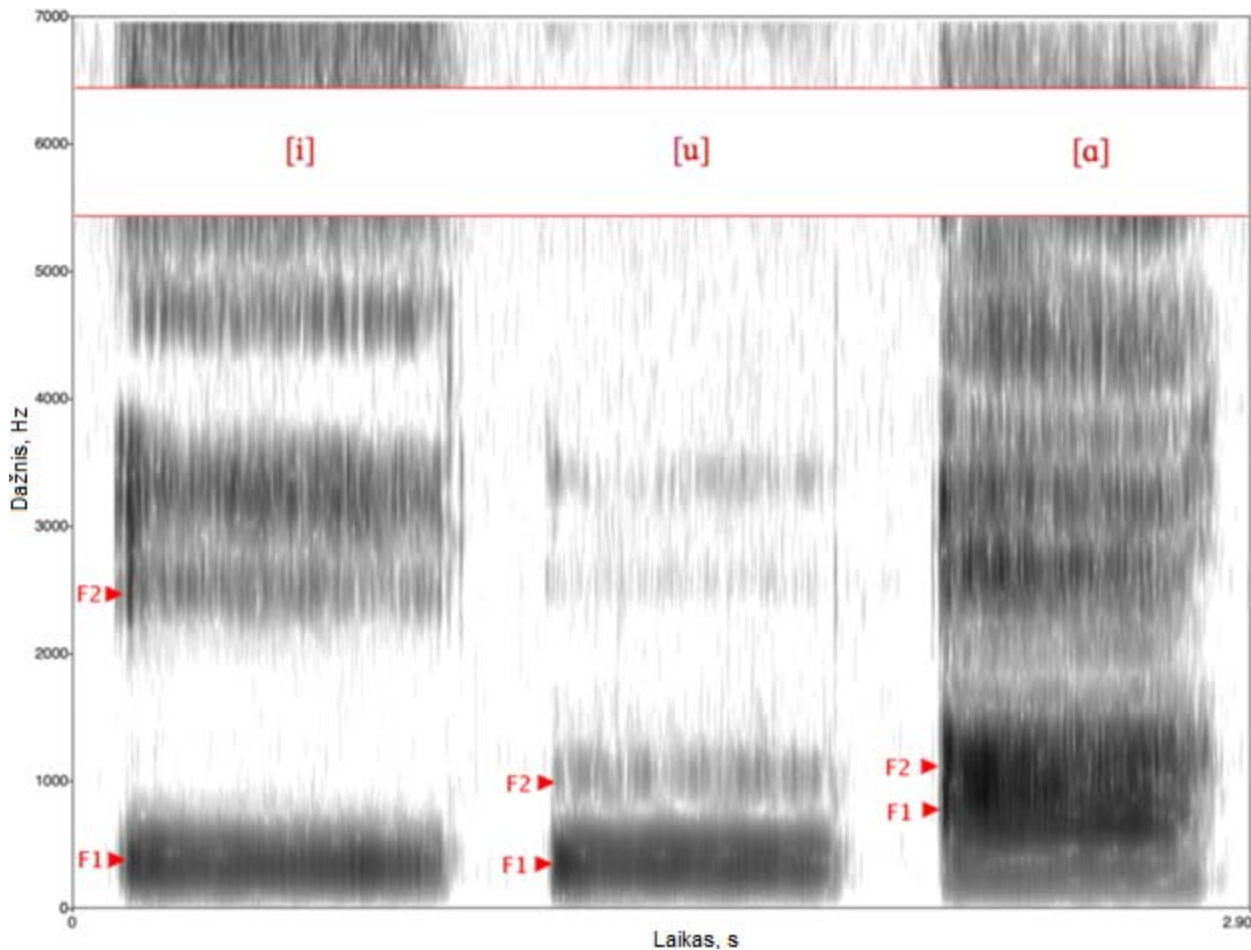
kur Δf – žingsnis; N – diskrečiųjų imčių skaičius; Δt – diskretizavimo žingsnis. Žinant signalo spektro funkciją, įmanoma apskaičiuoti signalo spektrines charakteristikas.

Tyrimuose dažnai naudojamos spektrinės charakteristikos yra sekančios: aukščiausias pagrindinis dažnis (F_{hi}), žemiausias pagrindinis dažnis (F_{lo}), įvairios dažninės komponentės (formančių analizė), spektro centroidė, spektro srautas, spektro tankis, spektrinis vokas ir pan.

Spektrinė centroidė yra charakteristika, nurodanti aplink kokį dažnį pasiskirsto spektro amplitudė, arba kitaip, kur yra signalo spektro „gravitacijos centras“. Dėl šios priežasties, šnekos signale atsiradus tylos segmentams, jų spektrinė amplitudė bus lygi 0, o skardžių garsų spektrinė centroidė bus didesnės vertės nei dusliųjų. Spektrinė centroidė skaičiuojama pagal formulę:

$$SP = \frac{\sum_{m=0}^{N-1} (m)X(m)}{\sum_{m=0}^{N-1} X(m)}, \quad (2.5)$$

kur SP yra spektrinės centroidės koeficientas; $f(m)$ nusako intervalo m , kurio ilgis yra N centrinių dažnį; $X(m)$ yra DFT spektro intervalo m amplitudės svertinis vidurkis.



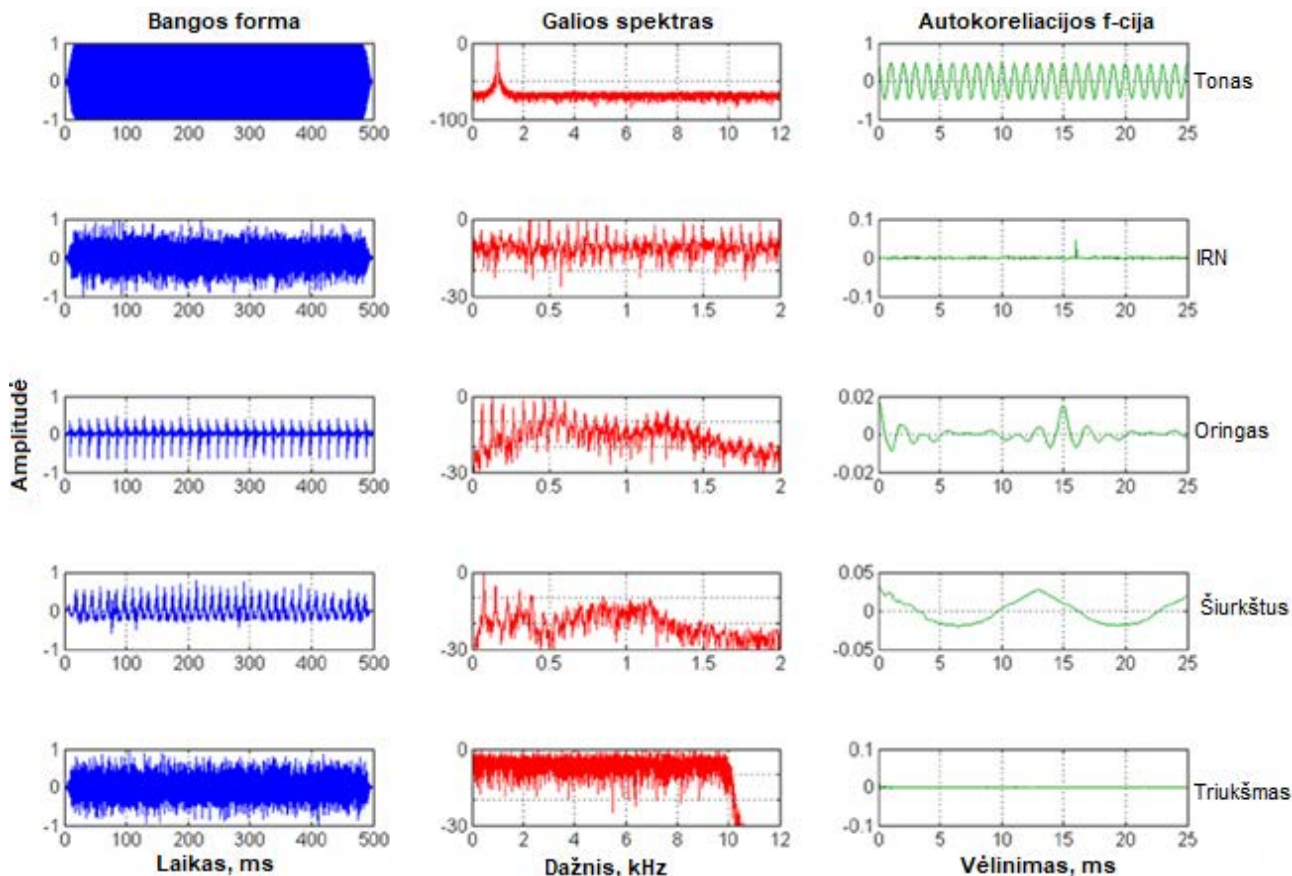
2.2 pav. Spektrogramų palyginimas tarp /i/, /u/, /a/ garsų (Ish Iswar, 2005)

Spektrinis srautas yra dydis, nusakantis, kaip greitai kinta signalo galios spektras. Dėl šios priežasties, ši charakteristika randama lyginant signalo kadro m galią su prieš tai buvusio kadro galia:

$$SF_i = \sum_{k=1}^{N/1} (|X_i(k)| - |X_i(k-1)|), \quad (2.6)$$

kur SF_i yra i -tojo kadro, kurio ilgis yra N spektrinis srautas; $X(k)$ yra DFT koeficiento vertė. Spektrinis srautas naudojamas signalo tono kitimo įvertinimui – pavyzdžiui, styginiai instrumentai pasižymi aukštais, bet greitai slopančiais tonais, todėl suformuluoti išvadą, kad pagal spektrinio srauto kitimą įmanoma atpažinti ne vien instrumentus ir pat ir šnekos signalo nukrypimus (pvz., atskirti patologinio balso segmentą nuo tokio pat ne patologinio balso).

Pasireiškus balso klostėms būdingoms patologijoms kinta balso stiprumas ir verta iširti, kokių metodų pagalba yra apskaičiuojamas signalo stiprumas. Čia naudinga tampa Parsevalio santykis, kurio esmė yra, kad laiko ir dažnio srityse pavaizduotos signalo reprezentacijos yra lygiavertės, todėl



2.3 pav. Bangos formos, spektro ir autokoreliacijos funkcijos, esant specifiniems dirgikliams (Shrivastav et al. 2012)

jos turi tą pačią energiją. Naudojantis 2.3 ir 2.4 formulėmis, Parsevalio santykis tarp signalo energijos laiko srityje ir dažnių srityje išreiškiamas formule (pagal Styveno Smito rekomendacijas):

$$E(t) = E(f) = \sum_{i=0}^{N-1} x[i]^2 = \int_0^T x(t)e^{-2\pi ift} dt, \quad (2.7)$$

kur $E(t)$ – balso energija laiko srityje; $E(f)$ – balso energija dažnių srityje; $x[i]$ – skaitmeninis signalas laiko srityje; $X[k]$ – skaitmeninis signalas dažnių srityje. Signalo energija pagal apibrėžimą yra signalo amplitudė pakelta kvadratu, todėl abi lygties pusės yra keliamos kvadratu. Suminė energija dažnių srityje lygties pusėje yra dalinama iš $2N$, kad apskaičiuoti kiekvieną signalo imtį dažnių srityje, vėliau konvertuojamą į sinusoidę, turinčią N skaitinių verčių laiko srityje.

Remiantis šnekos signalo spektru, galima įvertinti signalo savybes dažnių srityje. Signalo spektrogramose matomas formančių pasiskirstymas dažnių srityje – tai leidžia atskirti kalbos garsus, nes skirtingi garsai pasižymi tam tikromis dažnių juostomis, kuriose yra aptinkamos būdingos formantės. Taip pat spektro forma yra naudinga nusakant balso jutimines savybes (šiurkštumą, kimumą) ir spektro srauto pokytis padeda atpažinti individualius instrumentus. Dėl balso silpnumo pasireiškus balso klosčių sutrikimams, nustatyta, kad įmanoma įvertinti signalo intensyvumą, remiantis spektro energija.

2.3. Balso klosčių būklės požymiai

Pastebėta, kad tyrimuose informatyvios yra charakteristikos, susijusios su signalo trikdžiais, triukšmu ir lūžinėjimu. Šio tipo parametrinės charakteristikos yra naudojamos bandant įvertinti balso klosčių generuojamo garso kokybę, nes suteikia galimybę įvertinti signalo pagrindinio tono (f_0), signalo amplitudės, signalo harmonikų ir tylos segmentų pokyčius ir triukšmus viena skaitine išraiška. Pagal Santošo 2011 m. tyrimą, charakteristikos, apibūdinančios signalo pagrindinio tono f_0 pokytį yra: absoliutus drebulys (*Jitta*), drebulio procentinė vertė (*jitt*), tono triukšmo koeficientas (*PPQ*), santykinis vidutinio triukšmo įvertinimas (*RAP*). Amplitudės maksimalių verčių pokytis įvertinamas apskaičiavus mirguliavimą (mirguliavimas decibelais (*ShdB*), mirguliavimas procentais (*Shim*), o amplitudės verčių pokyčiai glotninant signalą įvertinami amplitudės svyravimo koeficientais (amplitudės svyravimo koeficientas (*APQ*), glotnintas amplitudės svyravimo koeficientas (*sAPQ*)). Balso turbulencijos indeksas yra santykis tarp signalo neharmoninės energijos ir harmoninės energijos, o švelnios fonacijos indeksas nurodo energijos santykį tarp žemesnio dažnio harmonikų ir aukštesnio dažnio harmonikų. Pagal Santošą balso kokybė įvertinama su triukšmo–harmonikos santykiu (*NHR*), balso turbulencijos indeksu (*VTI*), švelnios fonacijos indeksu (*SPI*), balso lūžinėjimo lygiu (*DVB*), balso lūžių kiekiu (*NVB*). Šis rinkinys neapima visos įmanomos akustinių charakteristikų visumos, bet yra naudingas, nes padeda įvertinti, kokio pobūdžio akustinės charakteristikos yra naudojamos balso patologijoms įvertinti. Santošo tyrime buvo vertinamos pateiktos charakteristikos, siekiant nustatyti, kurios yra tinkamiausios nusakyti tam tikras disfoninio balso savybes. Dėl šios priežasties, reikia įvertinti, kurios iš autoriaus minimų charakteristikų yra svarbios pasireiškiant balso klostėms būdingoms patologijoms. Rasta, kad Danšri Gundžiauvi et al. 2018 m. atliktoje apžvalgoje buvo pastebėta, kad žmogaus ausimi lengvai atpažįstamos balsinės charakteristikos yra sunkiai apibūdinamos akustinėmis. Nepaisant šios problemos, koreliacijos tarp jutiminių savybių ir akustinių charakteristikų tyrimai yra atliekami, nes yra poreikis sukurti garso tyrimų žinių bazę (Gunjawate et al. 2018).

Kaip minėta anksčiau, pasireiškus balso klosčių patologijoms, be pakitusio pagrindinio tono (f_0), balsas įgyja kimumo, šiurkštumo ir oringumo charakteristikas ir pat lūžinėja (žr. poskyrį 1.3). Vertinga yra Beno Lastožek'o et al. 2018 m. metaanalizė, kurioje nagrinėjama koreliacija tarp įvairių akustinių charakteristikų ir balso šiurkštumo ir oringumo. Toliau pateiktos akustinės charakteristikos ir autorių apskaičiuoti Pirsono koreliacijos koeficientai (toliau žymimi r) su jutiminėmis balso charakteristikomis. Autoriai teigia, kad balso šiurkštumui nusakyti yra naudingos tokios parametrinės charakteristikos, kaip antrosios signalo harmonikos amplitudė (*H2A*) ($r = 0,73$), Pirsono koeficiento vertė autokoreliacijos funkcijos maksimume (*RPK*) ($r = 0,71$). Balso oringumas koreliuoja su signalo periodo standartinio nuokrypio logaritmu (*LNPSD*) ($r = 0,76$), glotalo–triukšmo žadinimo santykis (*GNE*) ($r = 0,73$), skirtumas tarp pirmosios ir antrosios formantės amplitudės (*F0–F1*) ($r = 0,72$),

salyginis aukštų dažnių triukšmo energijos lygmuo ($Hfno$) ($r = 0,70$) parametrais. Pagal Kavitos Sachdevos ir Tanvi Šrivastavos 2019 m. tyrimą, kimumas ir kitos dėl balso klosčių būklės pakitimo pasireikšiančios savybės, koreliuoja su akustinėmis charakteristikomis, tokiomis kaip drebulys ($r = 0,73$), mirguliavimas ($r = 0,74$), harmonikos–triukšmo santykis (HNR) ($r = 0,83$). Galiausiai, balso lūžinėjimas įprastai yra nusakomas balso lūžių kiekis (NVB) charakteristika (pagal Maruti Santošo ir Belur'o Radžiašekhar'o 2018 m. tyrimą), o Brahimo Sabiro et al. 2017 m. tyrime nustatyta, kad ši charakteristika stipriai koreliuoja su disfoniskais balsais ($r = 0,97$) (įvairaus tipo disfonijos diagnozuojamos esant balso klosčių sutrikimams, žr. 1.3 poskyrį). Matyti, kad be Santošo ir Radžiašekhar'o pateiktų charakteristikų, tyrimuose naudojamos signalo spektrinės savybės (pvz.: antrosios signalo harmonikos amplitudė ($H2A$)), bet balso pagrindinio tono (f_0) analizė tyrimuose yra naudojama retai.

Įvertinus studijose pateiktus koeficientus, pastebima, kad svarbios yra tokios charakteristikos: $H2A$, $LNPSD$, HNR ir NVB . Pateiktos formulės (žr. 2.8, 2.9, 2.10 formules) nurodo šių charakteristikų matematinės išraiškas. Idealiu atveju $H2A$ apskaičiuojama:

$$H2A = 2f_0, \quad (2.8)$$

kur $H2A$ – antrosios signalo harmonikos amplitudė; f_0 – balso pagrindinis tonas.

Signalų periodo standartinio nuokrypio logaritmas $LNPSD$ diskretiesiems garso signalams *Praat* programoje yra:

$$LNPSD = \ln \left(\frac{1}{n-1} \sum_{i=m..m+n-1} (x_i - \mu)^2 \right), \quad (2.9)$$

kur $LNPSD$ – signalo periodo standartinio nuokrypio logaritmas; x_i – diskretizuotojo akustinio signalo periodo vertė; μ – signalo periodo vidutinė vertė; n – kadrų skaičius.

Harmonikos–triukšmo santykis HNR decibelais, pagal Boersmą yra vertinamas:

$$HNR = 10 \cdot \log_{10} \frac{r'_x(\tau_{max})}{1-r'_x(\tau_{max})}, \quad (2.10)$$

kur HNR – harmonikos–triukšmo santykis, dB; $r'_x(\tau_{max})$ – lokalus normalizuotos autokoreliacijos funkcijos minimumas.

Pagal Boersmą balso lūžių kiekis (NVB) *Praat* programoje vertinamas suskaičiavus, kiek balso signale yra tylos segmentų, ilgesnių nei 1,25 s. Toks segmentas yra įskaitomas, jei buvo aptiktas po bent jau 250 ms trukmės modalinės fonacijos ir 250 ms iki senakčios modalinės fonacijos. Balso lūžiai gali atsirasti nukritus kalbos dažniui arba dėl atsiranda dėl balso aparato spazmų.

Be šių charakteristikų, taip pat rasta, kaip yra vertinamos drebulio ir mirguliavimo charakteristikos, nes jos yra dažnai naudojamos balso klosčių vertinime (žr. 1.4 lentelę). Remiantis Paulo Boersmos *Praat* programa, šios charakteristikos skaičiuojamos 2.10 ir 2.11 formulėmis:

$$Drebulys = \frac{\frac{1}{N-1} \sum_{i=2}^N |T_i - T_{i-1}|}{\frac{1}{N} \sum_{i=1}^N T_i}, \quad (2.11)$$

kur *Drebulys* – signalo periodo pokytis; T_i – periodo trukmė, s; i – periodo numeris; N – periodų skaičius. Remiantis drebulio charakteristika mirguliavimo charakteristika yra:

$$Mirguliavimas = \frac{\frac{1}{N-1} \sum_{i=1}^N |A_i - A_{i-1}|}{\frac{1}{N} \sum_{i=1}^N A_i}, \quad (2.12)$$

kur *Mirguliavimas* – signalo periodo pokytis; A_i – amplitudė; i – amplitudės numeris; N – amplitudžių skaičius.

Apibendrinant galima teigti, kad vertinant patologinio balso būklę, remiantis jutiminių balso charakteristikų koreliacija su akustinėmis charakteristikomis, verta remtis laiko srities parametrais, nusakančiais signalo periodo nestabilumą (*drebulys*), signalo periodo standartinio nuokrypio logaritmas (*LNPSD*), amplitudės nestabilumą (*Mirguliavimas*). Pastebima, kad yra naudinga vertinti signalo harmonikos–triukšmo santykis (*HNR*), nes ši charakteristika koreliuoja su balso kimumu ($r = 0,83$) ir pat egzistuoja rekomendacijos, kuriose minima, kad tariamiems balsiams /a/ ir /i/, *HNR* turėtų siekti 20 dB, o kimiems balsams charakteristikos vertė yra žemesnė nei 20 dB.

2.4. Balso klosčių būklės identifikavimas

Įvertinus ankstesniuose poskyriuose pateiktą informaciją, nuspręsta, kad balso klosčių patologijų paveiktą balsą geriausia yra vertinti pasitelkus sekančias parametrines akustines charakteristikas:

- balso pagrindinis tonas (f_0);
- balso stiprumas (I);
- antrosios signalo harmonikos amplitudė ($H2A$);
- signalo periodo standartinio nuokrypio logaritmas (*LNPSD*);
- signalo ir triukšmo santykis (*HNR*);
- balso lūžių kiekis (*NVB*).

Šios charakteristikos yra svarbios, nes padeda nusakyti pacientui būdingus balso klosčių fizinius parametrus (pvz., balso klosčių ilgis), kurie lemia balso pagrindinį toną f_0 , taip pat nusako triukšmo kiekį balse (*HNR*), kas yra svarbu, nes balsai paveikti balso klosčių patologijų pasižymi kimumu. Be šių charakteristikų, vertinga yra signalo periodo standartinio nuokrypio logaritmo (*LNPSD*) charakteristika, nusakanti standartinį periodo (T) nuokrypį, nes pasireiškus oringumui balse, balso klostės nesugeba tinkamai susiglausti ir taip būna generuojamas nestabilų oro srautą. Balso stiprumo (I) ir balso lūžių kiekio (*NVB*) charakteristikos yra svarbios, nes disfoniški balsai pasižymi silpnu garsu, o pasireiškus ūmiems balso klosčių patologijų atvejams, balsas lūžinėja. Galiausiai, antrosios signalo harmonikos amplitudė ($H2A$) nusako antrosios harmonikos amplitudę.

Siekiant įvertinti, ar sudarytas charakteristikų rinkinys yra tinkamas, nuspręsta palyginti šį rinkinį su balso klosčių kokybės tyrimuose naudojamomis charakteristikomis. Kaip pastebėta anksčiau (žr. 1.4 poskyrį), yra žinoma, kad nėra apibrėžto akustinių charakteristikų rinkinio, teikiančio daugiausiai informacijos apie balso klosčių patologijų paveiktų balsų būklę (Gunjawate et al. 2018) ir nėra metodikos ar rekomendacijų, kokio turinio balsų įrašai turi būti naudojami (tiriami įvairaus tipo balsų įrašai – pvz., dainuojantis balsas, tariamos balsės ir pan. – žr. 1.4 lentelę). Nes tyrimuose naudojamos charakteristikos ir balsų įrašai yra įvairaus tipo, nuspręsta remtis akustinėmis charakteristikomis, apskaičiuojamomis remiantis ta pačia atvira *Saarbruecken* medicininių balso įrašų duomenų baze. Remiantis apžvalgoje atliktu *GRADE* vertinimu ir tyrimuose naudotų duomenų specifika, pasirinkti du tyrimai:

- balso pagrindinis tonas (f_0), drebulys, mirguliavimas, *balso stiprumas* (I) (Prakup, 2012);
- harmonikos–triukšmo santykis (HNR), drebulys, mirguliavimas (Guzman, 2013) .

Matyti, kad kiekvieno paciento balso įrašą galima aprašyti penkių charakteristikų vektoriumi (balso pagrindinis tonas (f_0), drebulys, mirguliavimas, *balso stiprumas* (I), harmonikos–triukšmo santykis (HNR)). Šios savybės gali apibūdinti disfonijos metu aiškiai žmogaus ausimi girdimus balso iškrypimus ir nestabilumus balso signalo periode ir amplitudėje (drebulys, mirguliavimas), nusakyti paciento balso specifiką (balso pagrindinis tonas (f_0)) ir klaidingos informacijos kiekį garso duomenyse (harmonikos–triukšmo santykis (HNR)) (Ferrand, 2007). Verta pastebėti, kad vertinant mirguliavimo charakteristikos koreliaciją su užkimusiu balsu, gaunami įvairūs rezultatai (Wendahl, 1966; Heiberger ir Horii, 1982; Hillenbrand, 1988), nors Kavitos Sachdevos ir Tanvi Šrivastavos 2019 m. tyrime pastebima stipri koreliacija su disfonišku balsu, paveiktu balso klosčių patologijų. Galima daryti prielaidą, kad ši charakteristika nebus esminė, siekiant apibūdinti balso kimumą. Lyginant tyrimuose naudotų charakteristikų vektorių, dar kartą patvirtinta prielaida, kad harmonikos–triukšmo santykio (HNR) charakteristika yra tinkama apibūdinti balso triukšmui arba kitaip – šiurkštumui. Galiausiai, patvirtinta balso pagrindinio tona (f_0) ir balso stiprumas (I) svarba siekiant įvertinti, ar paciento balso signalas yra patologinis ar ne.

Pastebėta, kad tyrimuose susijusiuose su balso patologijomis ir jų klasifikacija, yra naudojami mašinių mokymosi algoritmai. Tyrimuose kai kurių klasifikatorių tikslumas siekia 100 %, kai bandoma atskirti sveikus ir patologinius balsus (Selvakumari ir Radha, 2017), todėl yra galimybė kurti automatinio klasifikavimo įrankius. Šilos Selvakumari apžvalgoje pastebėta, kad klasifikatoriai apmokomi charakteristikų vektoriais su daug akustinių charakteristikų. Taip daroma, nes bandoma identifikuoti platesnį patologijų spektrą ir suteikti vietos galimiems nukrypimams. Vis dėlto, imtis su dideliu kintamųjų vektoriumi gali sukelti perpildymo problemą (angl. *overfitting*). Siekiant išvengti šios problemos, charakteristikų vektoriai yra apdorojami pasitelkus principinių komponentų analizę (angl. *Principal Component Analysis*), kurios metu koreliuojantys kintamieji yra

prastinami – taip gaunamas tiesiškai nepriklausomų verčių vektorius, kurio pirmoji pozicija priklauso kintamajam pasižyminčiam didžiausia dispersija. Be matematinių transformacijų taip pat rekomenduojama naudotis ypatybių atranka rankiniu būdu – pasitelkus balso specialistų ekspertizę, įmanoma sumažinti charakteristikų vektorių, klasifikuoti duomenis (Gupta, 2016).

Siekiant įvertinti sudarytų charakteristikų rinkinių pritaikomumą klosčių funkcinai būklei vertinti ir patologijoms aptikti, verta pasinaudoti mašinų mokymosi metodais. Siekiant tinkamai apmokyti mašinų mokymosi programą, skirtą patologijos atpažinimui, netinkamos charakteristikos lems žemą programos našumą. Dėl to įvertinama, kurios charakteristikos turi daugiausiai informacijos – tai galima atlikti pasinaudojus matematiniais metodais, ar net empiriniais stebėjimais. Taip pat verta pastebėti, kad per mažas duomenų kiekis bus nepakankamas algoritmo apmokymui, o per didelis sukels perpildymo problemą. Apibendrinus galima teigti, kad duomenų imtis lems naudojamo mašinų mokymosi algoritmo pasirinkimą, našumą, bet algoritmo optimizacija gali reikalauti imties duomenų modifikacijos, dėl to kuriant charakteristikų rinkinius yra svarbu remtis įvairių sričių specialistų rekomendacijomis.

3. EKSPERIMENTINIO TYRIMO PLANAS

Sukurto akustinių charakteristikų rinkinio tinkamumas būdingų balso klosčių patologijoms aptikti ir balso klosčių funkinei būklei vertinti, turi būti patikrintas naudojantis matematiniais kiekybiniais ir kokybiniais metodais. Uždavinio įgyvendinimas pradedamas nuo eksperimentinių duomenų atrankos ir duomenų imties rūšiavimo. Tinkamų balso įrašų imtys apdorojamos akustinės analizės programomis, tokiu būdu įgyjant akustinių charakteristikų rinkinius, kurių pritaikomumo, balso klosčių funkinei būklei vertinti ir patologijoms aptikti, įvertinimą būtų galima atlikti pasinaudojus mašinų mokymosi algoritmais. Gauti rezultatai interpretuojami pasitelkus metodiką, medicinoje naudojamą diagnozės tikslumui įvertinti.

Medicininiai duomenys yra konfidencialūs, todėl patologinių balsų įrašų atranka turi būti atliekama naudojantis atviromis duomenų bazėmis, kurių populiariausia yra laikoma *Saarbruecken* atviro naudojimo medicininio balso įrašų duomenų bazė, sukurta Sarlando universiteto Fonetikos instituto pastangomis. Šioje duomenų bazėje paskelbti įvairaus amžiaus ir skirtingų lyčių pacientų patologijų paveikti balsai ir šių pacientų sveiki balsai. Šių balso įrašų akustinės charakteristikos privalo būti tikslios, svarbu tinkamai pasirinkti akustinės analizės programą, kurios pagalba būtų įmanoma apskaičiuoti charakteristikas, minimas 2.4 poskyryje. Dėl šios priežasties, apžvelgta, kokios programos yra naudojamos balso klosčių tyrimuose. Žinant akustinių charakteristikų rinkinį, toliau vertinamas šių charakteristikų pritaikomumas.

Sukurtas akustinių charakteristikų rinkinys lyginamas su sudarytu etaloniniu charakteristikų rinkiniu sprendžiant mašinų mokymosi klasifikavimo uždavinį. Gautiems rezultatams vertinti pritaikytas medicinoje dažnai naudojamas jautrumo ir specifiškumo metodas (angl. *sensitivity and specificity*), kurio pagrindas yra santykio tarp teisingos ir klaidingos diagnozės esant tam tikrai duomenų imčiai vertinimas. Gauti rezultatai lyginami su tyrimais, kurių metu vertinama balso klosčių būklė, bet akustinių charakteristikų rinkinys sudarytas remiantis automatiniais imties analizės metodais (pvz., principinių komponentų analize) arba imtis yra neribojama (bandoma klasifikaciją atlikti remiantis charakteristikų imtimi, sudaryta iš visų įmanomų akustinių charakteristikų, nevertinant jų svarbos ar koreliacijos su jutiminėmis charakteristikomis).

3.1. Eksperimentinių duomenų atranka

Siekiant patikrinti sudaryto akustinių charakteristikų rinkinio pritaikomumą balso klosčių patologijoms aptikti ir balso klosčių funkcinės būklės vertinti, akustinės charakteristikos turi būti apskaičiuojamos remiantis balsų įrašais, gautais esant toms pačioms sąlygoms. Dėl šios priežasties pasirinkta *Saarbruecken* atviro naudojimo medicininio balso įrašų duomenų bazė.

Duomenų bazėje paskelbti daugiau kaip 2000 žmonių balsų įrašai (Barry ir Pützer). Patologijos, varginančios šiuos pacientus yra įvairaus pobūdžio (pvz., laringitas, disfonija, gerklų vėžys ir pan.). Kadangi tiriamos balso klostės, jų būklė ir patologijos, duomenų bazėje atrinktos balso patologijos, pasireiškiančios esant funkciniais balso sutrikimams, nurodytiems 1.3 lentelėje (3.1 lentelėje pavaizduotos duomenų bazėje paskelbti patologijos ir jų atitikmenys iš 1.3 lentelės). Pacientai, kurie dažniausiai skundžiasi šiomis patologijomis, priklauso tokioms profesinėms grupėms, kaip mokytojai, dainininkai, prekybininkai, taip pat ir gamyklų darbininkai (Cohen et al. 2012, Fritzell, 1996; Miller ir Verdolini, 1995; Thibeault et al. 2004, Williams, 2003). Ši pacientų grupė yra aktuali, nes yra nustatyta, kad tarp 50 % ir 80 % mokytojų yra bent kartą per gyvenimą skundęsi balso sutrikimais (Cutiva et al. 2013; Martins et al. 2015).

3.1 lentelė. Funkciniai balso sutrikimai ir medicininių patologijų diagnozės atitikmenys

| Nr. | Funkcinis balso sutrikimas pagal ASHA | Medicininė patologijos diagnozė pagal <i>Saarbruecken</i> duomenų bazę |
|-----|---------------------------------------|--|
| 1 | Fontrauma | Balso klosčių polipai, mazgeliai (angl. <i>nodules</i>) |
| 2 | Raumenų įtempio disfonija | Hipofunkcinė disfonija, hiperfunkcinė disfonija |
| 3 | Ventrikuliacinė fonacija | Ūmi hiperfunkcinė disfonija |
| 4 | Balso nuovargis | Funkcinė disfonija |
| 5 | „Dainininko mazgeliai“ | Balso klosčių polipai, mazgeliai |

Lyginant 1.3 lentelės duomenis matyti, sutrikimų ir patologijų pavadinimai nesutampa. Visiškai realu, kad atrinktos patologijos neatspindės funkcinų sutrikimų, pasireiškiančių esant balso klosčių pokyčiams. Šie nesutapimai atsirado dėl to, kad funkcinų balso sutrikimų rinkinys sudarytas remiantis *American Speech-Language-Hearing Association* duomenimis, o patologijų atitikmenys *Saarbruecken* duomenų bazėje yra sudaryti Sarlando universiteto Fonetikos instituto, esančio Vokietijoje – skirtingi regionai naudoja skirtingą terminologiją įvardyti tiems patiems sutrikimams. Tačiau, įvertinus abiejų rinkinių aprašus, nuspręsta, kad pateikti ekvivalentai yra teisingi (pvz., esant ventrikuliacinei fonacijai, fonacijoms generuoti naudojamos netikros balso klostės, tas pats pastebima pasireiškus ir hiperfunkcinei disfonijai (pagal Mičigano universiteto ligoninės duomenis)). Pasireiškus balso klostėms būdingoms patologijoms, balso sutrikimai yra sukeliama kelių etiologinių faktorių (žr. 1.3 poskyrį) – automatinis individualių balso klosčių patologijų atpažinimas tampa sunkiu uždaviniu, ypač dėl to, kad nagrinėjamų patologijų imtis pasireiškia dėl tos pačios priežasties – intensyvaus balso naudojimo ir balso klosčių audinių pokyčių, atsiradusių dėl netinkamo balso klosčių naudojimo.

Apžvelgus tyrimus, kuriuose yra nagrinėjama balso klosčių būklė arba bandoma atpažinti balso klosčių patologijas remiantis akustinėmis charakteristikomis ir mašinų mokymusi, pastebėta, kad

rekomenduojama vyrų ir moterų balsus nagrinėti atskirai. Zuzana Dankovičova et al. savo 2018 m. disfoniškų balsų tyrime rado, kad tyrime naudotų klasifikatorių našumas, kai atrenkami akustinių charakteristikų požymiai, yra 80,3 %, kai lyčių duomenys neatskiriami, 80,6 % kai imtį sudaro vien moterų balsai ir 86,2 %, kai imtis yra sudaryta vien iš vyriškų balsų. Šiuos duomenis ir anksčiau minėtą prielaidą dėl vyrų ir moterų patologinių balsų įrašų imčių atskyrimo patvirtina 2.1 poskyryje nagrinėtos Paulo Boersmos ir Deivido Vyniko rekomendacijos *Praat* programai, kuriose minima, kad skirtingų lyčių balsai pasiskirsto skirtinguose dažnių intervaluose – balsus reikia vertinti atskirai, kad neįprastai žemi, disfoniški moteriški balsai nebūtų klaidingai įvertinami kaip sveiki vyriški.

Kaip minėta 1.4 poskyryje, tyrimuose kuriuose vertinama balso klosčių būklė akustinės analizės metodais, naudojami balsų įrašai, kurių turinys yra įvairus (dainuojantis balsas, tariama frazė ar tęsiama balsio fonacija). *Saarbruecken* duomenų bazėje yra paskelbti balsų įrašai, kurių turinys yra įvairiais aukštumais tariamos balsės, taip pat tariama vokiška frazė. Reikia įvertinti, kokio turinio įrašus naudojant bus gaunami geriausi klasifikacijos rezultatai. 1.2 poskyryje minėta, kad kalbos specifika ir paciento nuotaika lemia kalbos toną ir intonaciją, o tai veikia akustinių charakteristikų turinį. Pavyzdžiui, pastebima, kad liūdesio ir laimės jausmai išreikšti skirtingomis intonacijomis skirsis balso pagrindiniu tonu (f_0) (pagal Tanjos Banzinger ir Klausio Ščererio 2005 m. publikuotą straipsnį). Todėl atliekant balso akustinę analizę, reikia naudoti ilgai tariamus balsius, nes yra panaikinama paciento emocijų įtaka balso įrašui. Kadangi duomenų bazėje taip pat yra laikomi skirtingų balsių įrašai, svarbu įvertinti, kokiais balsiais naudojantis, apskaičiuojamos tiksliausios akustinės charakteristikos. Panaudota Zuzana Dankovičovos ir kitų 2018 m. tyrimas, kuriame nustatyta, kad naudojant pacientui būdingu tonu tariamą balsį /a/, vyrų patologiniai balsai yra atpažįstami 82,61 % tikslumu, kai klasifikatorius yra atraminių vektorių klasifikatorius (angl. *Support Vector Machine*), o tariant balsį /u/, tikslumas siekia tik 65,22 %, kai klasifikatorius yra tas pats. Akivaizdu, kad palaikoma balsio /a/ fonacija pacientui būdingu pagrindiniu tonu (f_0) turi teikti tiksliausias charakteristikas, kurios yra tinkamos balso klosčių funkcinės būklės vertinimui ir patologijoms aptikti, nes tariant balsį /a/, dėl liežuvio atsitraukimo į burnos ertmės galą, burnos ertmė yra atvira – įmanoma stiprus garso rezonansas ir balso klosčių generuojamas garso signalas gali kilti gerklomis aukštyn be kliūčių.

Įvertinus anksčiau pateiktas prielaidas, iš *Saarbruecken* duomenų bazės atrinktos balso klostėms būdingos patologijos ir balsų įrašų imtys suskirstytos pagal pacientų lytį. Imtis sudaryta tik iš pacientų, kurių amžius viršija 18 m. – vaikų ir paauglių organizmai yra bręstantys, jų balso klosčių fiziniai parametrai kinta ir sudaro nuokrypius charakteristikose, nes jiems yra būdingas aukštas pagrindinis tonas (f_0) (pvz., laikoma, kad vaikų balsų pagrindinis tonas (f_0) yra pasiskirstęs 250–300 Hz dažnių ruože, kai vyriškiems balsas būdingas pagrindinis tonas f_0 randamas 85 Hz–155 Hz

dažnių ruožą (Loftas, 2019). Būdingų balso klostėms patologijų įrašų kiekis, kai balsų įrašai atrinkti įvertinus pacientų amžių ir lytį, nurodyti 3.2 lentelėje.

3.2 lentelė. Patologinių balsų įrašų kiekis, įvertinus pacientų amžių ir lytį

| Eil. Num. | Funkcinis balso sutrikimas | Moterys | Vyrai |
|-----------|---|---------|-------|
| 1. | Pacientų skaičius (sveikų balsų įrašų kiekis) | 438 | 1404 |
| 2. | Balso klosčių mazgeliai | 21 | 6 |
| 3. | Balso klosčių polipai | 15 | 26 |
| 4. | Funkcinė disfonija | 76 | 38 |
| 5. | Hipofunkcinė disfonija | 5 | 13 |
| 6. | Hiperfunkcinė disfonija | 169 | 48 |

Matyti, kad skirtingoms lytims yra būdingos skirtingo tipo patologijos. Balso klostėms būdingų patologijų ir pacientų skaičiaus santykis įvertintas pagal formulę:

$$c = \frac{\sum_{i=1}^N a}{b}, \quad (3.1)$$

kur c – santykis tarp pacientų skaičiaus *Saarbruecken* duomenų bazėje ir funkcinio balso sutrikimų; a – specifinės funkcinio pobūdžio patologijos balso įrašų kiekis; b – pacientų skaičius; N – patologijos eilės numeris iš 4.1 lentelės.

Įvertinus balso klostėms būdingų patologijų ir sveikų balsų santykį pagal 4.1 formulę, gauta, kad remiantis *Saarbruecken* duomenų baze, moterys dažniausiai skundžiasi funkciniais balso sutrikimais (funkciniai sutrikimai sudaro 65,3 % visų moteriškos lyties pacientų balsų įrašų), o vyrams funkcinės prigimties balso klosčių patologijos nėra būdingos, nes sudaro 9,3 % visų vyriškos lyties pacientų skaičiaus. Šį pastebėjimą patvirtina Renatos Džiardim, Sandi Marijos Baretos ir Ados Avilos Asunkao 2007 m. tyrimas, kurio metu nagrinėti mokytojų balsai – pastebima, kad šios profesinės grupės moteriškos lyties nariams dažniau diagnozuojamos funkcinio pobūdžio patologijos, nes moterys turi mažesnes gerklas, dėl šios priežasties kad išgauti tokio stiprumo garsą, koks yra galimas turint vyriškas gerklas, balso klostės yra naudojamos intensyviau.

3.2 lentelėje matyti, kad balso įrašų kiekiai kiekvienai patologijai nėra vienodi. Esant mažam įrašų kiekiui galimas prastas algoritmo apmokymas, kai siekiama atpažinti patologijas. Kad išvengtų šios problemos, nuspręsta funkcinio balso sutrikimų patologijų įrašų imtis papildyti ir apjungti į kelias

grupės, kad būtų pasiektas 100 įrašų patologijų grupei kiekį (siekiama kiekvienai lyčiai priskirti bent 50 įrašų). Šis sprendimas priimtas, nes pastebėta, kad nagrinėjant duomenų bazę, pastebėta, kad patologiniams balsų įrašams priskirti keli etiologiniai faktoriai (pvz., pastebėta, kad balso klosčių polipai pasireiškia kartu su Reinkės edema). Patologinių balsų įrašų imčių dydžiai ir grupės nurodyti 3.3 lentelėje.

3.3 lentelė. Patologinių balsų įrašų kiekis, sugrupavus panašias patologijas ir įvertinus pacientų amžių ir lytį

| Eil. Num. | Funkcinių balso sutrikimų grupė | Pasireiškiantys sutrikimai | Moterys | Vyrai |
|------------------|---|---|----------------|--------------|
| 1 | Pacientų skaičius (sveikų balsų įrašų kiekis) | – | 50 | 50 |
| 2 | Sutrikimai pasireiškiantys per balso klosčių struktūrinius pakitimus | Balso klosčių mazgeliai Balso klosčių polipai Reinkės edema | 50 | 39 |
| 3 | Sutrikimai pasireiškiantys dėl balso klosčių raumenų neįprasto įtempimo / atpalaidavimo | Hipofunkcinė disfonija, hiperfunkcinė disfonija | 50 | 50 |

Funcinių balso sutrikimų grupės eksperimente naudojamos kaip klasės. Tai reiškia, kad akustinių charakteristikų rinkiniui patikrinti mašinų mokymosi klasifikavimo algoritmais, reikia naudoti algoritmus, pritaikytus spręsti kelių klasių klasifikavimo problemą. Pasinaudojus keliais klasifikavimo algoritmais, įmanoma palyginti sudaryto akustinių charakteristikų rinkinio universalumą.

Sudaryto akustinių charakteristikų rinkinio rezultatai taip pat palyginami su tyrimais, kurių metu balso klosčių patologijos yra vertinamos pasitelkus požymių mažinimo algoritmus. Palyginimui pasirinkti Zuzanos Dankovičovos ir kitų 2018 m. tyrimų rezultatai, kuriuose tiriama balso klosčių disfonija. Šiame tyrime iš *Saarbruecken* duomenų bazės atrinkti 100 sveikų pacientų balsų įrašai ir disfoniški balsai, iš kurių 53 yra moteriški ir 41 yra vyriškas. Studijoje remiantis automatiniais imties analizės metodais (šiuo atveju principinių komponentų analize) sudaroma charakteristikų imtis. Autoriai lygina klasifikatorių, apmokytų automatiniais imties analizės metodais, rezultatus su rezultatais, gautais pasiėmus maksimalų įmanomą akustinių charakteristikų rinkinį. Kadangi tyrime

tiriami tik disfoniški balsai, klasifikavimo algoritmai apmokomi atskirti dvi klases (viena klasė – sveiki balsai, antra klasė – disfoniški balsai). Dėl šios priežasties, iš *Saarbruecken* duomenų bazės atsisiųsti tie patys aukščiau minėtoje studijoje naudoti balsų įrašai (žr. 3.4 lentelę). Iš šių įrašų apskaičiuojamas sudarytas akustinių charakteristikų rinkinys, apmokyti klasifikatoriai ir rezultatai paruošti palyginimui su Zuzanos Dankovičovos ir kitų 2018 m. studija.

3.4 lentelė. Patologinių balsų įrašų kiekis, lyginant sveikus ir disfoniškus balsus ir įvertinus pacientų amžių ir lytį

| Eil. Num. | Įrašų apibūdinimas | Pasireiškiantys sutrikimai | Moterys | Vyrai |
|-----------|---|----------------------------|---------|-------|
| 1. | Pacientų skaičius (sveikų balsų įrašų kiekis) | – | 50 | 50 |
| 2. | Bendrai diagnozuoti disfoniški balsai (naudojama rezultatų palyginimui su Zuzanos Dankovičovos ir kitų 2018 m. studijoje) | Disfonija | 53 | 41 |

Siekiant įvertinti sudaryto akustinių charakteristikų rinkinio pritaikomumą balso klosčių patologijoms aptikti ir balso klosčių funkcinę būklę vertinti, eksperimentui atrinkti balsų įrašai, įrašyti tomis pačiomis sąlygomis. Iš *Saarbruecken* atviros medicininių balso įrašų duomenų bazės atsisiųsti patologinių balsų įrašai ir sveikų balsų įrašai. Sudarytos įrašų grupės išskirtos pagal lytis, įrašų grupėse siekta palaikyti panašų įrašų kiekį (po 50 įrašų kiekvienai lyčiai). Eksperimento metu atliekami palyginimai:

- Sudarytas akustinių charakteristikų rinkinys ir etaloninis akustinių charakteristikų rinkinys, kai klasifikatorius apmokytas atskirti tris klases;
- Sudarytu akustinių charakteristikų rinkiniu apmokytas klasifikatorius, kurio rezultatai lyginami su kitų autorių tyrimų rezultatais ir etaloninio akustinių charakteristikų rinkinio rezultatais.

Eksperimento metu palyginimai atliekami apmokius kelis klasifikatorius. Taip užtikrinama rezultatų įvairovė ir patikrinamas sudaryto akustinių charakteristikų rinkinio pritaikomumas skirtingiems klasifikavimo metodams.

3.2. Eksperimentinio tyrimo metodika

Eksperimentinio tyrimo metu turimas akustinių charakteristikų rinkinys lyginamas su sudarytu etaloniniu charakteristikų rinkiniu sprendžiant mašinų mokymosi klasifikavimo uždavinį. Siekiant patikrinti akustinių charakteristikų rinkinio pritaikomumą balso klosčių patologijoms aptikti, rezultatams įvertinti pritaikytas medicinoje dažnai naudojamas jautrumo ir specifiškumo metodas (angl. *sensitivity and specificity*), o vizualizacijai panaudotos maišaties matricos (angl. *confusion matrix*). Eksperimento eigos algoritmas pateiktas 3.1 pav.

Eksperimentui svarbu pasirinkti programinę įrangą, tinkamą akustinei analizei atlikti. Akustinėje analizėje dažnai naudojama Computerized Speech Lab sukurta *MDVP (Multi-Dimensional Voice Program)* ir Paulo Boersmos ir Deivido Vyniko sukurta *Praat*. Mehmedo Akifo Kiliko et al. 2011 m. studijoje lyginamos *MDVP* ir *Praat* programos. Autoriai pastebi, kad tarp programų apskaičiuojamų akustinių charakteristikų yra koreliacija, bet galutiniams rezultatams reikia daugiau tyrimų. Vis dėlto, *Praat* programa yra laikoma patikima didelio kiekio gydytojų ir mokslininkų. *Praat* programoje taip pat yra įgyvendintas autokoreliacijos metodas pagrindiniam tonui (f_0) apskaičiuoti. Šis metodas yra laikomas patikimiausiu apskaičiuojant šią charakteristiką (žr. 2.1 poskyrį). Kadangi balso pagrindinio tono pokyčiai pasireiškia atsiradus balso klosčių patologijoms yra itin svarbu šią charakteristiką apskaičiuoti kiek įmanoma tiksliau. Dėl aukščiau paminėtų priežasčių, akustinėms charakteristikoms apskaičiuoti pasirinkta *Praat* programa.

Eksperimentui sukurtas programos kodas, skirtas *Praat* programoje apskaičiuoti ekspertų rekomendacijomis paremta charakteristikų rinkinio vertes (balso pagrindinis tonas (f_0), drebulys, mirguliavimas, intensyvumas (I), harmonikos–triukšmo santykis (HNR) – šis rinkinys toliau vadinamas etaloniniu) ir koreliacija paremtą akustinių charakteristikų rinkinį (balso pagrindinis tonas (f_0), balso stiprumas (I) antrosios signalo harmonikos amplitudė ($H2A$), signalo periodo standartinio nuokrypio logaritmas ($LNPSD$), harmonikos–triukšmo santykis (HNR), balso lūžių kiekis (NVB) – šis rinkinys toliau vadinamas sudarytu). Programos kode įvertintos 2 skyriuje aptartos rekomendacijos *Praat* programai. Programoje balsų įrašai atskirti pagal lytį ir patologiją. Rezultatai įkelti į tekstinį failą.

Klasifikatorių mokymas ir rezultatų analizė atlikta *MATLAB* aplinkoje. Atlikus *Praat* rezultatų importą, suformuotos duomenų lentelės, kur akustinės charakteristikos yra klasės ypatybės, o patologijų grupės – klasės (žr. 3.3 lentelę, 3.4 lentelę). Klasių etiketėms duomenų lentelėse sukurti papildomi stulpeliai, į kuriuos pridėtos šios etiketės. Klasių ypatybių duomenų vektoriai normalizuoti centravimo palei nulį metodu (angl. *zero-centering*). Testui atrinkti 25 % duomenų, likę naudojami algoritmų mokymui. Duomenų lentelių pavadinimai ir paaiškinimai pateikti 3.5 lentelėje.

3.5 lentelė. Duomenų lentelių pavadinimai ir paaiškinimai

| Eil. Num. | Lentelės pavadinimas | Paskirtis |
|-----------|---|---|
| 1. | Etaloninės akustinės charakteristikos vyrams, kai patologijų klasės 3 | Funkcinių balso sutrikimų grupėms atskirti, kai pacientai yra vyrai |
| 2. | Etaloninės akustinės charakteristikos moterims, kai patologijų klasės 3 | Funkcinių balso sutrikimų grupėms atskirti, kai pacientės yra moterys |
| 3. | Sudarytos akustinės charakteristikos vyrams, kai patologijų klasės 3 | Funkcinių balso sutrikimų grupėms atskirti, kai pacientai yra vyrai |
| 4. | Sudarytos akustinės charakteristikos moterims, kai patologijų klasės 3 | Funkcinių balso sutrikimų grupėms atskirti, kai pacientės yra moterys |
| 5. | Sudarytos akustinės charakteristikos vyrams, kai patologijų klasės 2 | Disfoniškų balsų atskyrimui nuo sveikų balsų, kai pacientai yra vyrai |
| 6. | Sudarytos akustinės charakteristikos moterims, kai patologijų klasės 2 | Disfoniškų balsų atskyrimui nuo sveikų balsų, kai pacientės yra moterys |

Ekperimentui ieškant tinkamų duomenų klasifikavimo mašinų mokymosi algoritmų, pastebėta, kad tyrimuose naudojama didelė algoritmų įvairovė. Nėra patvirtinta, kad viena metodika teiktų geresnius rezultatus nei likusios visais įmanomais atvejais (Drotar ir Smekal, 2014). Toliau pateikta apžvalga dažniausiai naudojamų balso patologijų vertinime programinės įrangos priemonėmis algoritmų. Tyrimų apžvalgos metu pastebėta, kad naudojami tokie algoritmai:

- atraminių vektorių klasifikatorius (angl. *Support Vector Machine*);
- sprendimų medis (angl. *Decision tree*);
- paprastas Bajeso klasifikatorius (angl. *Naive Bayesian classification*);
- logistinio modelio medis (angl. *Logistic Model Tree*);
- atsitiktinio miško algoritmas (angl. *Random Forest Classifier*);
- k -artimiausių kaimynų metodas (angl. *k-nearest neighbour*);
- giliojo mokymo neuroniniai tinklai (angl. *Deep neural networks*);
- Gauso mišinio modelis (angl. *Gaussian mixture model*).

Nors tyrimuose naudojamos įvairios priemonės patologijoms vertinti ir nepastebima, kad viena metodika smarkiai lenktų likusias tikslumu, bet naudojant tą pačią duomenų imtį, įmanoma palyginti metodų tikslumą. Zuzanos Dankovičovos et al. 2018 m. tyrime lyginti atraminių vektorių klasifikatoriaus, k -artimiausių kaimyno ir atsitiktinio miško algoritmo metodai ir nustatyta, kad tai pačiai duomenų imčiai, atraminių vektorių klasifikatoriaus algoritmas pasižymi geresniu tikslumu

aptinkant disfonijas nei kiti naudoti algoritmai. Jajos Alemamil ir Ladžiali Almazaideh studijoje lyginant k -artimiausių kaimynų metodą ir paprastąjį Bajeso klasifikatorių, gauta, kad Bajeso klasifikatorius pasižymi geresniu tikslumu aptinkant balso patologijas. Fang Ši et al. studijoje pastebima, kad lyginant giliojo mokymo neuroninio tinklo ir Gauso mišinio modelį, giliojo mokymo neuroninis tinklas pasižymi geresniais rezultatais (gautas 99,32 % tikslumas), nei Gauso mišinio modelis. Įvertinus aukščiau pateiktą informaciją, galima daryti prielaidą, kad siekiant pasirinkti metodą patologijoms aptikti, verta rinktis iš: atraminių vektorių klasifikatoriaus, paprastojo Bajeso klasifikatoriaus ir giliojo mokymo neuroninio tinklo.

Atraminio vektoriaus mašinos algoritmas (angl. *Support Vector Machine*) yra plačiai naudojamas balso patologijų tyrimo srityje (Verde et al. 2018). Eksperimentui naudojamas daugiaklasis klasifikatorius, įgyvendintas per išvesties šifrų su klaidos pataisa modelį (angl. *Error-Correction Output Codes Model*). Išvesties šifrų su klaidos pataisa modelio algoritmas pasižymi klasifikavimo tikslumu, lyginant su kitais klasifikatoriais, skirtais daugiau nei dviejų klasių atpažinimui (Fürnkranz, 2002). Modeliui parinktas vienas prieš vieną (angl. *one-versus-one*) šifravimas, tai reiškia, kad K skaičiui klasių, reikalingas $K(K-1)/2$ skaičius vienas prieš vieną dizaino atraminių vektorių klasifikatorių. Apmokant šias mašinas, lyginamos dvi klasės (žr. 3.5 lentelę), trečioji klasė praleidžiamos (žr. 3.5 lentelę – klasės, pažymėtos 0, yra praleidžiamos).

3.5 lentelė. Klasių šifravimo matrica, kai duomenys yra suskirstyti į tris klases

| | 1 mašina | 2 mašina | 3 mašina |
|----------------|-----------------|-----------------|-----------------|
| 1 klasė | 1 | 1 | 0 |
| 2 klasė | -1 | 0 | 1 |
| 3 klasė | 0 | -1 | -1 |

Toliau išvesties šifrų su klaidos pataisa modelio algoritmas klasei k apskaičiuoja naują stebėjimo vertę \hat{k} , kur nuostolių suma yra minimali. Stebėjimo vertė \hat{k} skaičiuojama (Escalera, Pujol ir Radeva, 2010):

$$\hat{k} = \underset{k}{\operatorname{argmin}} \frac{\sum_{l=1}^L |m_{kl}| g(m_{kl} s_l)}{\sum_{l=1}^L |m_{kl}|}, \quad (3.2)$$

kur \hat{k} – stebėjimo vertė; m_{kl} – šifravimo matricos M elementas; k – klasės indeksas; l – mašinos indeksas; s_l – numatomas klasifikavimo balas; $g(\cdot)$ – nuostolių funkcija; L – mašinų skaičius.

MATLAB aplinkoje įgyvendintas išvesties šifrų su klaidos pataisa modelis atraminių vektorių klasifikatoriui. Sukurtas atraminių vektorių klasifikatoriaus šablonas. Keičiant šablono parametrus, apmokytas klaidos pataisos išvesties šiframs modelis, kiekvienas mokymas kartotas 10 kartų. Modelis pritaikytas sudaryto akustinių charakteristikų rinkinio duomenų lentelei ir etaloninio

akustinių charakteristikų rinkinio duomenų lentelei, kai pacientų akustinių charakteristikų duomenys atskirti pagal lytį.

Naudoti atraminių vektorių klasifikatoriaus šablono parametrai:

- Branduolio funkcija: tiesinė, *RFB*, polinomine;
- C : 0,1, 1, 20, 100 (*RFB* ir polinominei branduolio funkcijai), 1000 (*RFB* ir polinominei branduolio funkcijai);
- Polinominės funkcijos laipsnis: 1, 2.

Bajeso klasifikacija (angl. *Bayesian classification*) yra paprastos struktūros tikimybinis klasifikavimo algoritmas. Šis algoritmas nereikalauja daug resursų ir gerai veikia net esant duomenų trūkumui, kas svarbu analizuojant balsinį signalą. *MATLAB* aplinkoje įgyvendintas paprastasis Bajeso algoritmas, kuriam nustatyta pasiskirstymo funkcija skirta daugiamačiams, daugialypiams atsitiktiniams kintamiesiems.

Kai pasiskirstymas yra pritaikytas daugiamačiams, daugialypiams atsitiktiniams kintamiesiems, algoritmas kiekvienam numatomajam modeliui atlieka statistinį kaupimą į atskirų kategorinių lygių rūšiuotą sąrašą. Kiekvienam numatomajam modeliui j (angl. *predictor*) klasėje k skaičiuojami kategoriniai lygiai ir nustatoma tikimybė, kad numatomasis modelis j klasėje k turės kategorinį lygį L . Naudojant pridėtinį glotninimą (Manning et al., 2008) ši tikimybė skaičiuojama:

$$P(\text{predictor } j = L | \text{class } k) = \frac{1 + m_{j|k}(L)}{m_j + m_k}, \quad (3.3)$$

kur $P()$ – tikimybė, kad numatomasis modelis j klasėje k turės kategorinį lygį L ; j – numatomojo modelio indeksas; k – klasės indeksas; L – kategorinis lygis; m_j – unikalių kategorinių lygių skaičius numatomajame modelyje; m_k – stebėjimų klasėje k svertinis skaičius. Funkcija $m_{j|k}(L)$ skaičiuojama:

$$m_{j|k}(L) = n_k \frac{\sum_{i: y_i \in \text{class } k} I\{x_{ij} = L\} w_i}{\sum_{i: y_i \in \text{class } k} w_i}, \quad (3.4)$$

kur $m_{j|k}(L)$ – stebėjimų svertinis skaičius, kai numatomasis modelis j klasėje k lygus L ; n_k – stebėjimų skaičius klasėje k ; w_i – stebėjimo i svorio koeficientas; $I\{x_{ij} = L\} = 1$, kai $x_{ij} = L$, kitu atveju 0.

Paprastasis Bajeso klasifikatorius įgyvendintas *MATLAB* aplinkoje. Sukurtasis Bajeso klasifikatorius yra pritaikytas sudaryto akustinių charakteristikų rinkinio duomenų lentelei ir etaloninio akustinių charakteristikų rinkinio duomenų lentelei, kai pacientų akustinių charakteristikų duomenys atskirti pagal lytį.

Giliojo mokymo neuroninis tinklai (angl. *Deep Neural Network*) yra plačiai naudojami medicinos tyrimuose ir biometrinių signalų apdorojime. Įprastu atveju, giliojo mokymo neuroninis tinklu laikomas bet koks neuroninis tinklas, turintis kelis dirbtinių neuronų sluoksnius. Šie sluoksniai tampa svarbūs klasifikavimo uždaviniuose, nes leidžia apskaičiuoti sudėtingų duomenų požymius (Deng ir Yu, 2014). Eksperimentui pasirinktas giliojo mokymo neuroninis tinklas, paremtas *Levenberg–Marquardt* algoritmu. Algoritmas pasirinktas dėl mokymo greičio. Laikoma, kad kai neuroninis tinklas yra vidutinio dydžio (turi ne daugiau nei kelis šimtus svorio koeficientų), šis algoritmas apmokymą atlieka greičiausiai, kai lyginama su kitais neuroninio tinklo mokymo algoritmais (Hagan ir Menhaj, 1999).

Levenberg–Marquardt algoritmas aproksimuoja *Hessian* matricą per Jakobo matricas, apskaičiuotas standartiniu atgalinio skleidimo metodu (angl. *backpropagation*). Naudojant šį algoritmą, parametrai perskaičiuojami:

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e, \quad (3.5)$$

kur x – parametras; k – parametro indeksas; J – Jakobo matrica; e – visų klaidų išraiškų vektorius; I – tapatybės matrica; μ – slopinimo faktorius. Giliojo mokymo neuroninio tinklo klasifikatoriaus mokymo pradžioje inicijuojama didelė μ vertė. Kiekvienos iteracijos metu, μ vertė mažinama gradientinio nusileidimo kryptimi, kol mažinant μ mažėja nuostoliai. Iteracijos kartojamos, kol gradientinis nusileidimas nekonverguoja palei 0.

Eksperimentui naudoti giliojo mokymo neuroninio tinklo nustatymai:

- neuronų sluoksnių skaičius: 2, 3, 4, 5;
- neuronų skaičius sluoksnyje: 10, 5, 20.

Giliojo mokymo neuroninio tinklo klasifikatorius, paremtas *Levenberg–Marquardt* algoritmu įgyvendintas *MATLAB* aplinkoje. Sukurtasis giliojo mokymo tinklo klasifikatorius pritaikytas sudaryto akustinių charakteristikų rinkinio duomenų lentelei ir etaloninio akustinių charakteristikų rinkinio duomenų lentelei, kai pacientų akustinių charakteristikų duomenys atskirti pagal lytį.

Klasifikatorių rezultatai įvertinti kryžminės patikros metodu. Naudoti du kryžminės patikros metodai:

- K -kryžminio skaidymo (angl. *k-fold*) – duomenys padalinti į k skaičių imčių. Viena imtis naudojama tikrinimui, likusios mokymui. Patikra vykdoma k kartų, kol kiekviena imtis panaudojama vieną kartą tikrinimui. Eksperimento metu duomenys dalinti į $k = 4$ imtis.
- Sulaikymo metodas (angl. *holdout*) – duomenys padalinami nustatytu santykiu. Santykiu padalintos dalys naudojamos algoritmo apmokymui ir tikrinimui. Eksperimento metu

duomenys padalinti santykiu 1:3, kur 3 dalys duomenų naudojamos mokymui ir 1 dalis tikrinimui.

Iš kryžminės patikros rezultatų lentelės atrinkti klasifikatoriai, kur po patikros gauti žemiausi tikslumo įvertinimai. Klasifikatoriai yra apmokyti atskiromis normalizuotomis testinių duomenų imtinimis (akustinių charakteristikų rinkiniu ir su sudarytu etaloninių charakteristikų rinkiniu). Galutiniai rezultatai įvertinti remiantis jautrumo ir specifiškumo skaičiavimu. Jautrumo ir specifiškumo vertinimas atliktas remiantis Ven'o Zu, Nensės Zeng ir Ning'o Vango 2010 m. tyrimu. Remiantis tyrimu pateikta metodika, jautrumas skaičiuojamas:

$$\text{Jautrumas} = \frac{TT}{TT+NN} \cdot 100\%, \quad (3.6)$$

kur *Jautrumas* – jautrumo įvertinimas, %; *TT* – tikrų teigiamų diagnozių kiekis; *NN* – netikrų neigiamų reikšmių kiekis. Jautrumo įvertinimas parodo kaip gerai algoritmas aptinka patologijas. Jautrumo skaičiavimas atliktas kiekvieno klasifikatoriaus rezultatams.

Siekiant įvertinti, kaip gerai algoritmas atskiria sveikus balsus, įvertinamas specifiškumas:

$$\text{Specifiškumas} = \frac{TN}{TN+NT} \cdot 100\%, \quad (3.7)$$

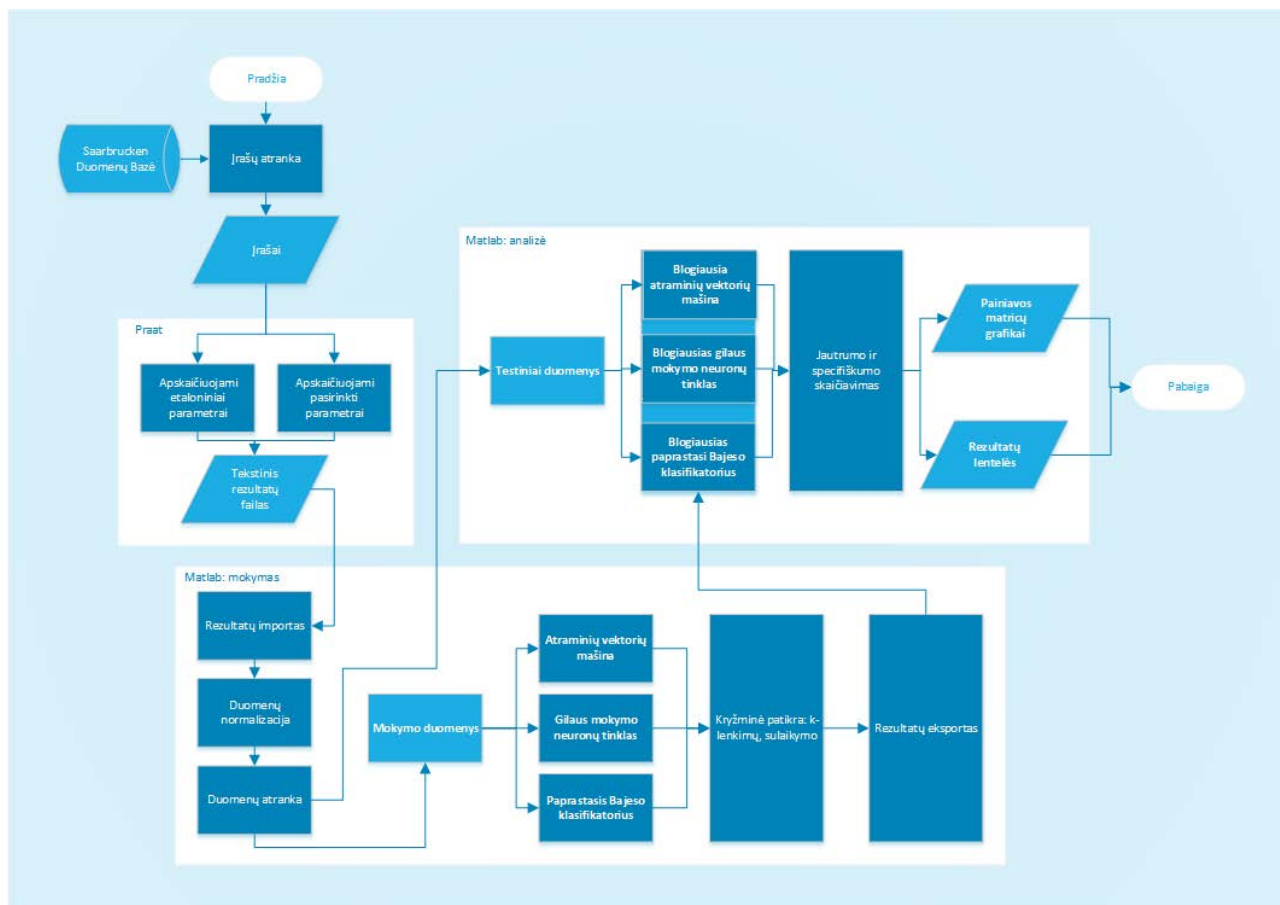
kur *Specifiškumas* – specifiškumo įvertinimas, %; *TN* – tikrų neigiamų diagnozių kiekis; *NT* – netikrų teigiamų diagnozių kiekis. Specifiškumo skaičiavimas atliktas kiekvieno klasifikatoriaus rezultatams.

Bendras klasifikatoriaus pritaikomumas patologijoms aptikti, įvertinamas remiantis taiklumo skaičiavimu pagal formulę:

$$\text{Taiklumas} = \frac{TT+TN}{TT+TN+NT+NN} \cdot 100\%, \quad (3.8)$$

kur *Taiklumas* – taiklumo įvertinimas, %; *TT* – tikrų teigiamų diagnozių kiekis; *TN* – tikrų neigiamų diagnozių kiekis; *NT* – netikrų teigiamų diagnozių kiekis; *NN* – netikrų neigiamų reikšmių kiekis. Tikslumo skaičiavimas atliktas kiekvieno klasifikatoriaus rezultatams. Tikslumo skaičiavimų rezultatai vizualizuoti su maišaties matricomis, siekiant parodyti klaidas klasifikuojant kiekvieną klasę.

Eksperto metu atliktas sudaryto akustinių charakteristikų rinkinio pritaikomumo balso klosčių patologijoms aptikti patikrinimas. Sudarytas akustinių charakteristikų rinkinys patikrintas pasinaudojus atraminių vektorių klasifikatoriumi, paprastuoju Bajeso klasifikatoriumi ir giliojo mokymo neuroniniu tinklu. Klasifikatoriai apmokyti sudarytomis normalizuotų duomenų lentelėmis, klasifikatoriams atlikta kryžminė patikra (naudoti *k*-kryžminio skaidymo ir išlaikymo metodai). Atrinkti mažiausio tikslumo klasifikatoriai (po vieną kiekvieno tipo klasifikatorių), taip siekiant įvertinti charakteristikų rinkinių našumą neoptimaliomis sąlygomis. Normalizuoti testiniai duomenys



3.1 pav. Eksperimentinio tyrimo eigos algoritmas

apmokyti atrinktaisiais klasifikatoriais, atliktas jautrumo ir specifiškumo įvertinimas. Rezultatai vizualizuoti maišaties matricose. Palyginimui eksperimentas pakartotas su nenormalizuotais duomenimis. Eksperimento rezultatai pateikti A priede, analizė 4.2 poskyryje.

4. BALSŲ KLOSČIŲ FUNKCINĖS BŪKLĖS TYRIMO REZULTATAI

Darbo metu uždaviniams įgyvendinti atlikta mokslinės literatūros kokybinė analizė. Analizės metu pagilintos žinios apie balsų klostes, jų patologijas ir akustinės analizės metodus. Remiantis tyrimų analizės duomenimis, sudaryta akustinių charakteristikų imtis, kurios pritaikomumas balsų klosčių funkcinėi būklei vertinti ir patologijoms aptikti įvertintas atlikus eksperimentinį tyrimą. Darbo metu suformuluotos išvados pateiktos šiame skyriuje. Skyriuje taip pat pateikiamos rekomendacijos ir siūlymai tolimesniems tyrimams ir sukurto akustinių charakteristikų rinkinio panaudojimui.

4.1. Analitiniai rezultatai

Literatūros analitinės analizės metu nustatyta, kad dauguma su balsų patologijomis susijusių sutrikimų veikia balsų klostes. Balsų klostės yra viena iš pagrindinių kalbos aparato dalių, o jų fiziniai parametrai lemia paciento balsą, šių organų būklės vertinimas yra daugelio tyrimų, susijusių su balsų patologijomis objektas ir darytina prielaida, kad šių organų pokyčiai atsispindės ir apskaičiuojamose balsų akustinėse charakteristikose.

Balsas pasižymi subjektyviai suvokiamomis jutiminėmis charakteristikomis (pvz., balsų šiurkštumas, oringumas ar kimumas), kurios yra lengvai įvertinamos garsiniu suvokimu, bet sunkiai aprašomos matematiniais metodais. Svarbios akustinės charakteristikos yra balsų pagrindinis tonas ir stipris ir rasta, kad vertinant balsų kokybę, reikia įvertinti balsų intonaciją ir toną, nes šios charakteristikos priklauso nuo paciento kalbos ir emocinės būklės – šios charakteristikos gali iškreipti tyrimo rezultatus.

Nustatyta, kad balsų klostes labiausiai veikia funkciniai balsų klosčių sutrikimai, nes šio tipo sutrikimai atsiranda dėl įtempto balsų naudojimo. Pasireiškus šiems sutrikimams, pakinta balsų klosčių elastingumas, forma – trikdomas jų gebėjimas vibruoti būdingu pacientui dažniu. Pasireiškus šiems sutrikimų grupėms, kinta balsų pagrindinis tonas (f_0) ir geba išlaikyti tam tikrą toną (balsas lūžinėja) ir pacientų balsai įgyja savybes, tokias kaip: kimumas, šiurkštumas, oringumas.

Galima teigti, kad tyrimuose vyrauja parametrinė akustinė analizė, bet nepasiūlytas universalus charakteristikų rinkinys, tinkamas nusakyti balsų klosčių būklę. Suformuluota išvada, kad atliekant patologinių balsų akustinę analizę, nėra bendros metodikos, tinkamos visoms duomenų imtims ir atvejams (Gupta, 2016). Šią prielaidą patvirtina 1.4 lentelės informacija, kurioje matyti, kad kai tyrimo objektas yra tas pats (pacientų balsų kokybė) yra naudojama didelė įvairovė akustinių charakteristikų. Siekiant surasti akustinių charakteristikų ir balsų patologijų sąryšius tyrimuose remiamasi įvairaus tipo metodikomis, pvz., akustiniai parametrai yra analizuojami pasitelkus įvairius savybių atrankos algoritmus arba yra atrenkami pasitelkus balsų ekspertų nuomonę (Gupta, 2016).

Pastebima, kad bandant klasifikuoti balso klosčių patologijas remiantis įvairaus pobūdžio klasifikavimo algoritmais, nebandoma įvertinti jutiminių savybių (pvz., oringumo, šiurkštumo ar kimumo) ir akustinių charakteristikų koreliacijos, tokiu būdu sukuriant optimalų charakteristikų rinkinį – tai yra galimybė tyrimams.

Atlikus balso pagrindinio tono (f_0) tyrimų analizę, suformuluota išvada, kad šiai charakteristikai apskaičiuoti yra sukurta įvairių metodų, nes tenka įvertinti balso signalo periodo t_s nepastovumą ir formančių energijos kitimą. Nustatyta, kad dažniausiai yra naudojamas Paulo Boersmos 1994 m. pasiūlytas autokoreliacijos metodas, kuris yra pranašesnis už kitų autorių siūlomus, nes yra patvaresnis (veikia taip pat gerai tiek žemiems, tiek aukštiems dažniams (Boersma, 1994)), atsparesnis signalo triukšmams ir taip pat gerai veikia esant tylos segmentams įrašė. Įvertinus Boersmos metodo specifiką, suformuluota išvada, kad bandant nustatyti balso tono pagrindinį dažnį f_0 , privaloma įvertinti balsų dažnių ruožų skirtumus, atsiradusius dėl vyrų ir moterų balso klosčių fizinių parametrų. Autoriaus rekomendacijose Praat programai (Boersma ir Weenik, 2005), minima, kad vyrų balsams geriausia nustatyti 75–300 Hz dažnių ribas, o moterų 100–500 Hz. Matyti, kad balsų pagrindinio tono dažnių ruožai persidengia, dėl šios priežasties, vertinant paciento balsą, tenka atkreipti dėmesį į lytį, nes neįprastai žemi, disfoniški moteriški balsai gali būti vertinami kaip sveiki vyriški balsai, jei vertinant duomenų imtis nėra atskiriamos. Vis dėlto, teisinga yra prielaida, kad siekiant optimizuoti f_0 aptikimą Boersmos metodu, svarbu įvertinti pacientų imties *Fhi* ir *Flo* charakteristikas, nusakančias minimalų ir maksimalų pasiekiamą balso dažnį, taip nustatant dažnių ribas pacientų balsų imčiai.

Remiantis šnekos signalo spektru, galima įvertinti signalo savybes dažnių srityje. Signalų spektrogramose matomas formančių pasiskirstymas dažnių srityje – tai leidžia atskirti kalbos garsus, nes skirtingi garsai pasižymi tam tikromis dažnių juostomis, kuriose yra aptinkamos būdingos formantės. Taip pat spektro forma yra naudojama nusakyti balso jutimines savybes (šiurkštumą, kimumą), o spektro srauto pokytis padeda atpažinti individualius instrumentus. Rasta, kad pasireiškus balso klosčių sutrikimams, balsas pasižymi silpnumu ir nustatyta, kad įmanoma įvertinti signalo intensyvumą, remiantis spektro energija.

Apibendrinant galima teigti, kad vertinant patologinio balso būklę remiantis jutiminių balso charakteristikų koreliacija su akustinėmis charakteristikomis, verta naudoti laiko srities charakteristikomis, nusakančiomis signalo periodo nestabilumą (*drebulys*, *LNPSD*) ir amplitudės nestabilumą (*Mirguliavimas*). Pastebima, kad yra naudinga vertinti signalo harmonikos–triukšmo santykį (*HNR*), nes ši charakteristika koreliuoja su balso kimumu (koreliacijos laipsnis (r) lygus 0,83). Rastos rekomendacijos, kuriose minima, kad tariamiems balsams /a/ ir /i/, *HNR* turėtų siekti 20 dB, o kimiems balsams charakteristikos vertė yra žemesnė nei 20 dB.

Siekiant įvertinti sudarytų charakteristikų rinkinių pritaikomumą klosčių funkcinei būklei vertinti ir patologijoms aptikti, verta pasinaudoti mašinų mokymosi metodais. Siekiant tinkamai apmokyti patologijų atpažinimui skirtą mašinų mokymosi programą, prastos charakteristikos lems žemą programos našumą. Dėl to įvertinama, kurios charakteristikos turi daugiausiai informacijos apie patologišią balsą – tai galima atlikti pasinaudojus matematiniais metodais, ar net empiriniais stebėjimais. Taip pat pastebėta, kad per mažas duomenų kiekis bus nepakankamas algoritmo apmokymui, o per didelis sukels perpildymo problemą. Apibendrinus galima teigti, kad duomenų imtis lems naudojamo mašinų mokymosi algoritmo pasirinkimą, našumą, bet algoritmo optimizacija gali reikalauti imties duomenų modifikacijos, dėl to kuriant charakteristikų rinkinius yra svarbu remtis įvairių sričių profesionalų rekomendacijomis.

4.2. Eksperimentinio tyrimo rezultatai

Remiantis analitinių rezultatų gairėmis, sukurta eksperimentinio tyrimo metodika, skirta įvertinti sudaryto akustinių charakteristikų rinkinio pritaikomumą balso klosčių patologijoms aptikti. Eksperimentui duomenys atrinkti iš *Saarbruecken* duomenų bazės, kurioje paskelbti pacientų balsų įrašai, suskirstyti pagal nustatytas balso patologijas. Pacientai atskirti pagal lytis ir atrinkti tik pacientai, kurių amžius viršija 18 m.. Įvertinus duomenų bazėje esančių balso klosčių patologijų skaičių, nustatyta, kad moterys dažniausiai skundžiasi funkciniais balso sutrikimais (funkciniai sutrikimai sudaro 65,3 % visų moteriškos lyties pacientų balsų įrašų), o vyrams funkcinės prigimties balso klosčių patologijos nėra būdingos, nes sudaro 9,3 % visų vyriškos lyties pacientų skaičiaus.

Balsų įrašai apdoroti *Praat* programa. Sukurtas kodas, skirtas apskaičiuoti sudaryto akustinių charakteristikų rinkinio vertes. Siekiant įvertinti šių akustinių charakteristikų pritaikomumą, charakteristikos lyginamos su etaloniniu charakteristikų rinkiniu, kurio vertės apskaičiuotos *Praat* programa. Duomenys normalizuoti *MATLAB* aplinkoje, rinkinio vertinimas atliktas *MATLAB* aplinkoje.

Sudaryto akustinių charakteristikų rinkinio vertinimas *MATLAB* aplinkoje atliktas remiantis trijų tipų klasifikatoriumis:

- Atraminių vektorių klasifikatorius;
- paprastasis Bajeso klasifikatorius;
- giliojo mokymo neuroninis tinklas.

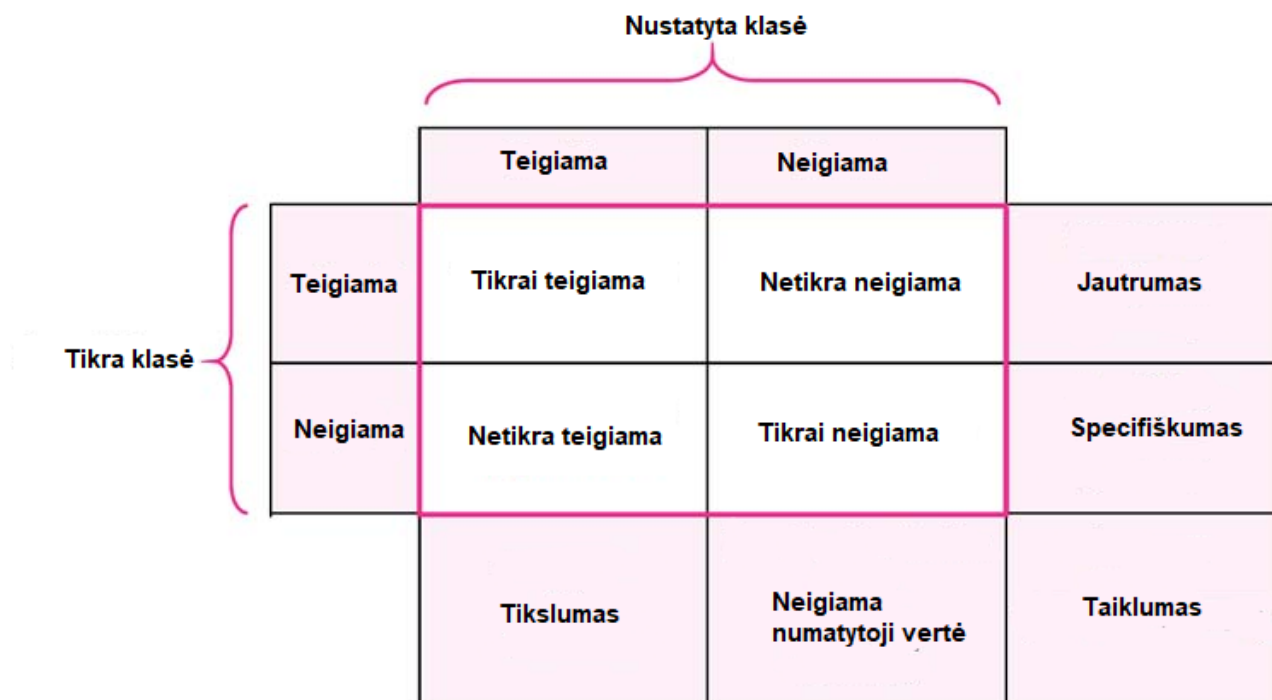
Klasifikatoriai atskirai apmokyti sudaryto akustinių charakteristikų rinkinio vertėmis ir etaloninio akustinių charakteristikų rinkinio vertėmis. Mokymui duomenų imtys atskirtos pagal lytį. Klasifikatoriai įvertinti k -kryžminio skaidymo ir sulaikymo kryžminės patikros metodais. Atrinkti blogiausi klasifikatoriai, apmokyti testiniais duomenimis. Blogiausi klasifikatoriai atrinkti, kad

įvertinti akustinių charakteristikų imties našumą esant neoptimalioms sąlygoms. Klasifikavimo rezultatai pateikti 4.1 lentelėje, 4.2 lentelėje, 4.3 lentelėje ir 4.4 lentelėje.

4.1 lentelė. Eksperimento rezultatai, klasifikuojant 2 klases, kai mokymo duomenys normalizuoti

| | Rinkinio tipas | Atraminių vektorių klasifikatorius | Bajeso klasifikatorius | Gilias mokymo neuroninis tinklas |
|----------------|-----------------------|---|-------------------------------|---|
| Vyrai | Sudarytas rinkinys | 100 % | 100 % | 100 % |
| | Etaloninis rinkinys | 100 % | 100 % | 83,05 % |
| Moterys | Sudarytas rinkinys | 100 % | 100 % | 100 % |
| | Etaloninis rinkinys | 100 % | 100 % | 79,03 % |

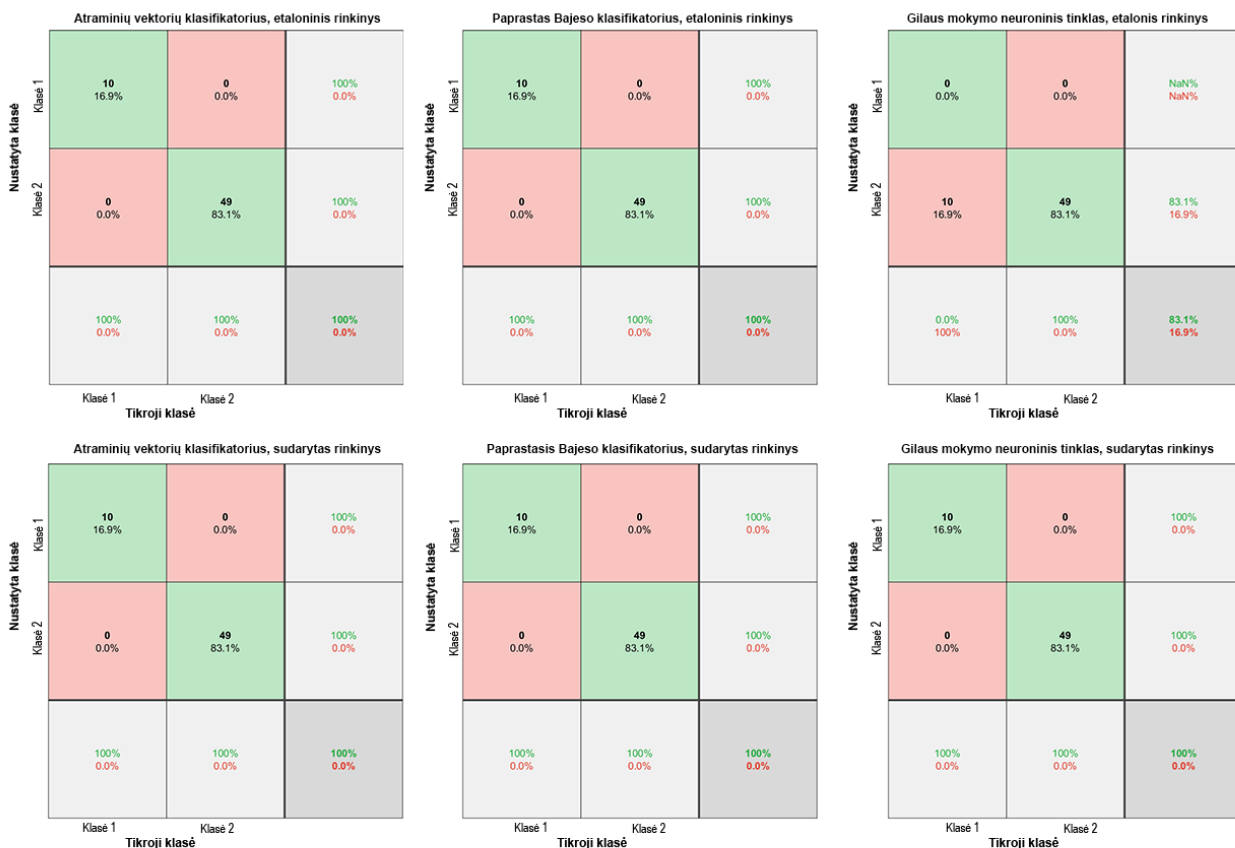
Normalizavus duomenų imtį, nustatyta, kad atskirai tiriant moterų ir vyrų patologinius balsus eksperimento sąlygoms, tiek sudarytas akustinių charakteristikų rinkinys, tiek etaloninis akustinių charakteristikų rinkinys yra tinkami atskirti sveikus ir disfoniškus balsus. Šie rinkiniai gali būti naudojami balso klosčių funkcinės būklės įvertinimui. Dėl *Saarbruecken* duomenų bazėje esančio įrašų kiekio (pacientų, su diagnozuota disfonija yra mažiau nei 100), rinkinio pritaikomumas reikalauja daugiau tyrimų. Papildomam rinkinio įvertinimui rinkinių akustinės charakteristikos turėtų būti apskaičiuojamos iš didesnės įrašų imties.



4.1 pav. Maišaties matricos verčių paaiškinimas

Klasifikatorių *Jautrumas* ir *Specifiškumas* pateikti painiavų matricių paveiksluose. Painiavų matricių verčių interpretavimas pateiktas 4.1 pav. Paveiksle matyti, kad procentinės vertės, pateiktos dešinėje maišaties matricos paveikslo pusėje reiškia jautrumą ir specifiškumą. Klasifikatoriaus *Taiklumas* kiekvienai klasei pateiktas ant matricos įstrižainės.

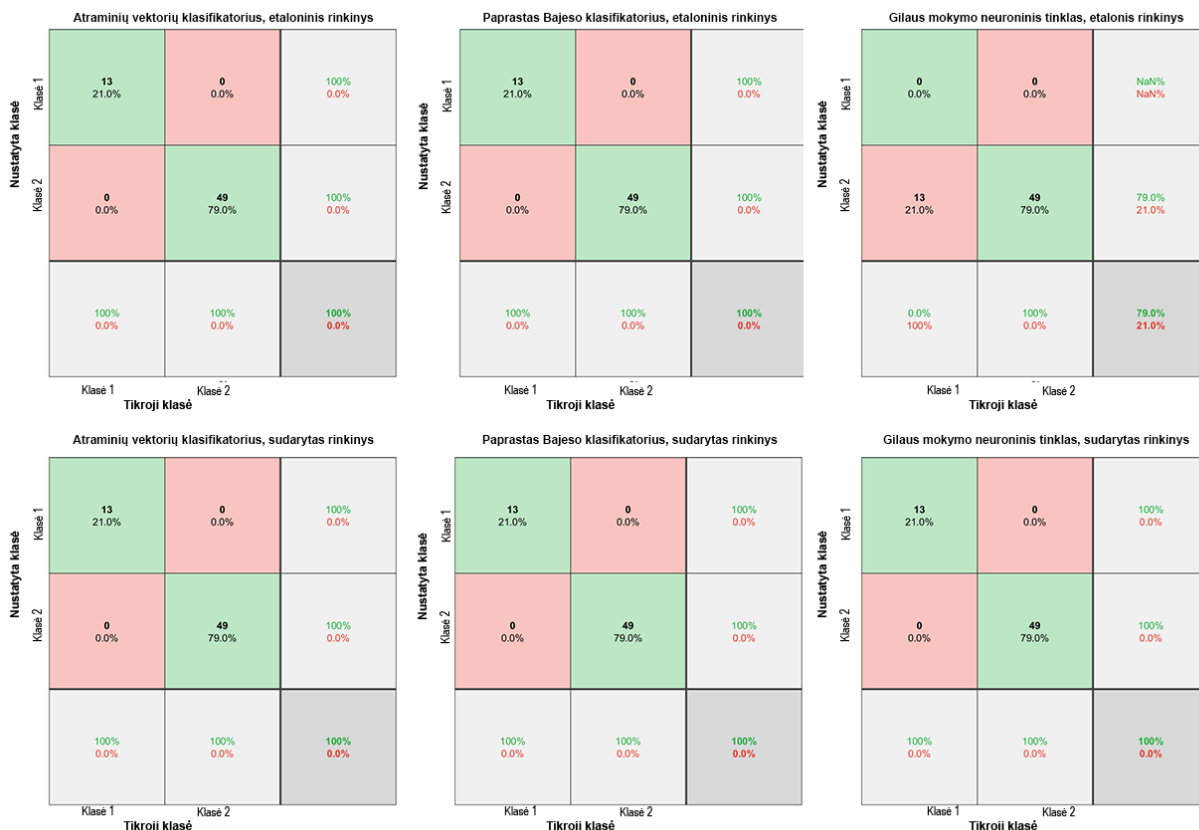
Maišaties matricioje (žr. 4.2 pav.) grafiškai pavaizduoti klasifikatorių duomenys, kai tiriami vyrų balsai. Paveiksle pateiktoji *Klasė 1* yra disfoniški balsai, o *Klasė 2* – sveiki. Maišaties matricioje pateikti jautrumo ir specifiškumo įvertinimai. Matyti, kad giliojo mokymo neuroninis tinklas, apmokytas etaloniniu charakteristikų rinkiniu, kai pacientai yra vyrai, nesugebėjo aptikti vyriškų disfoniškų balsų. Galima suformuluoti išvadą, kad nepaisant charakteristikų rinkinio kokybės, siekiant geriausių įmanomų rezultatų svarbu naudoti tinkamus klasifikatorius arba klasifikatorius optimizuoti. Šiuo atveju, neuroninio tinklo struktūra buvo sudaryta iš 5 sluoksnių, kiekviename sluoksnyje po 5 neuronus. Šio klasifikatoriaus *Taiklumas* po kryžminės patikros – 82,78 %. Patikros metu šis klasifikatorius neaptiko disfoniškų balsų.



4.2 pav. 2 klasių klasifikatorių maišaties matrica, kai pacientai yra vyrai, mokymo duomenys normalizuoti

Maišaties matricioje (žr. 4.3 pav.) grafiškai pavaizduoti klasifikatorių duomenys, kai tiriami moterų balsai. Paveiksle pateiktoji *Klasė 1* yra disfoniški balsai, o *Klasė 2* – sveiki. Maišaties matricioje pavaizduoti jautrumo ir specifiškumo įvertinimai, taip pat klasifikatoriaus bendras taiklumo

įvertinimas. Taip pat, kaip 4.2 pav., matyti, kad giliojo mokymo neuroninis tinklas, apmokytas etaloniniu charakteristikų rinkiniu, kai pacientai yra moterys, nesugebėjo aptikti disfoniškų balsų. Giliojo mokymo neuroninis tinklo struktūra buvo sudaryta iš 3 sluoksnių, kiekviename sluoksnyje po 5 neuronus. Šio klasifikatoriaus *Taiklumas* po patikros – 79,26 %. Šis klasifikatorius patikros metu neaptiko disfoniškų balsų.



4.3 pav. 2 klasių klasifikatorių maišaties matrica, kai pacientai yra moterys, mokymo duomenys normalizuoti

Eksperimento metu įvertintas sudaryto charakteristikų rinkinio pritaikomumas daugiau nei dviejų klasių atskirymui – bandoma atskirti kelias patologijas vieną nuo kitos. Kiekvienam įvertinimui gauti išrinkti apmokyti klasifikatoriai, kur po kryžminės patikros gauti žemiausi taiklumo įvertinimai. Pasirinktų klasifikatorių kryžminės patikros įvertinimai pateikti 4.2 lentelėje.

4.2 lentelė. Pasirinktų klasifikatorių kryžminės patikros įvertinimai

| | | | | |
|--|------------------------------------|---|---|--|
| | Klasifikatorius pritaikytas | Atraminių vektorių klasifikatoriaus kryžminės patikros metodas | Bajeso klasifikatoriaus kryžminės patikros metodas | Gilaus mokymo neuroninio tinklo tinklo patikros metodas |
|--|------------------------------------|---|---|--|

| | | K – kryžminio skaidymo | Išlaikymo | K – kryžminio skaidymo | Išlaikymo | Taiklumo įvertinimas |
|----------------|--------------------------|---------------------------------------|------------------|---------------------------------------|------------------|---------------------------------|
| Vyrai | Sudarytam rinkiniui | 86,98 % | 88,27 % | 98,6 % | 100 % | 98,6 % |
| | Etaloniniam rinkiniui | 87,44 % | 88,27 % | 98,14 % | 100 % | 89,3 % |
| Moterys | Sudarytam rinkiniui | 85,09 % | 84,21 % | 82,02 % | 100 % | 83,77 % |
| | Etaloniniam rinkiniui | 84,65 % | 84,21 % | 98,68 % | 98,35 % | 81,58 % |

Išrinkti apmokyti atraminių vektorių klasifikatoriaus klasifikatoriai įgyvendinti su linijine branduolio funkcija. Akivaizdu, kad šie klasifikatoriai turimiems duomenims nėra optimalūs, nes šie klasifikatoriai iš visų apmokytų pasižymi žemiausiais patikros rezultatais. Akivaizdu, kad svarbu optimizuoti pasirinktą klasifikatorių, kad pasiekti kiek įmanoma tikslesnius taiklumo rezultatus. Tai reiškia, kad esant prastos kokybės duomenims, rezultatus galima gerinti pasirinkus optimalesnį klasifikavimo algoritmą.

Pritaikius pasirinktus trijų klasių klasifikatorius testiniams duomenims, apskaičiuotas *Taiklumas*. 4.3 lentelėje pateikti klasifikatorių taiklumo įvertinimai. Matyti, kad sudarytas akustinių charakteristikų rinkinys pasižymi geresniais klasifikavimo rezultatais nei etaloninis. Etaloninis charakteristikų rinkinys pasižymėjo geresniais tikslumo rezultatais, kai klasifikuojami moteriški balsai.

4.3 lentelė. Eksperimento rezultatai, klasifikuojant 3 klases, kai mokymo duomenys normalizuoti

| | Rinkinio tipas | Atraminių vektorių klasifikatorius | Bajeso klasifikatorius | Gilaus mokymo neuroninis tinklas |
|----------------|-----------------------|---|-----------------------------------|---|
| Vyrai | Sudarytas rinkinys | 89,85 % | 100 % | 98,55 % |
| | Etaloninis rinkinys | 80,34 % | 91,35 % | 76,54 % |
| Moterys | Sudarytas rinkinys | 80,45 % | 83,91 % | 100 % |
| | Etaloninis rinkinys | 83,91 % | 95,40 % | 82,27 % |

Maišaties matricoje (žr. 4.4 pav.) grafiškai pavaizduoti klasifikatorių duomenys, kai tiriami vyrų balsai. Paveiksle pateiktoji *Klasė 1* yra sveiki balsai, *Klasė 2* – pokyčiai balse dėl raumenų nuovargio, o *Klasė 3* – balso klosčių struktūriniai pakitimai (klases sudarančios patologijos pateiktos 3.3 lentelėje). Matricose matyti, kad klasifikatoriai yra jautrūs sveikiems balsams (minimalus jautrumas

buvo 98 %, kai naudojamas atraminių vektorių klasifikatorius, o charakteristikų rinkinys yra etaloninis). Vertinant specifiškumą, algoritmai atskiria patologijų klases įvairiu tikslumu. Geriausiai pasirodė paprastas Bajeso klasifikatorius ir giliojo mokymo neuroninis tinklas, kai charakteristikų rinkinys yra sudarytas. Įvertinus kitų algoritmų ir rinkinių jautrumo ir specifiškumo vertes, matyti, kad algoritmai sunkiai atskiria patologijų grupių klases (*Klasė 2*, *Klasė 3*). Įvertinus šias vertes, rasta, kad paprastas Bajeso klasifikatorius buvo tinkamiausias klasifikatorius abiem charakteristikų rinkiniams.



4.4 pav. 3 klasių klasifikatorių maišaties matrica, kai pacientai yra vyrai, mokymo duomenys normalizuoti

Maišaties matricoje (žr. 4.7 pav.) grafiškai pavaizduoti klasifikatorių duomenys, kai tiriami moterų balsai. Paveiksle pateiktoji *Klasė 1* yra sveiki balsai, *Klasė 2* – pokyčiai balse dėl raumenų nuovargio, o *Klasė 3* – balso klosčių struktūriniai pakitimai (klases sudarančios patologijos pateiktos 3.3 lentelėje). Taip pat kaip ir vyrų balsų eksperimento atveju, pastebėta, kad algoritmai geriausiai atskiria moteriškus sveikus balsus nuo patologinių (5 iš 6 klasifikatorių atskyrė sveikus balsus nuo patologinių 100 % tikslumu). Bendras taiklumas vertinant moteriškus balsus yra žemesnis nei vyriškų. Šie rezultatai gauti dėl didesnės duomenų dispersijos (žr. 4.5 pav. ir 4.6. pav. – matyti, kad sudaryto charakteristikų rinkinio duomenys yra labiau išsisklaidę ir grupės susikerta).

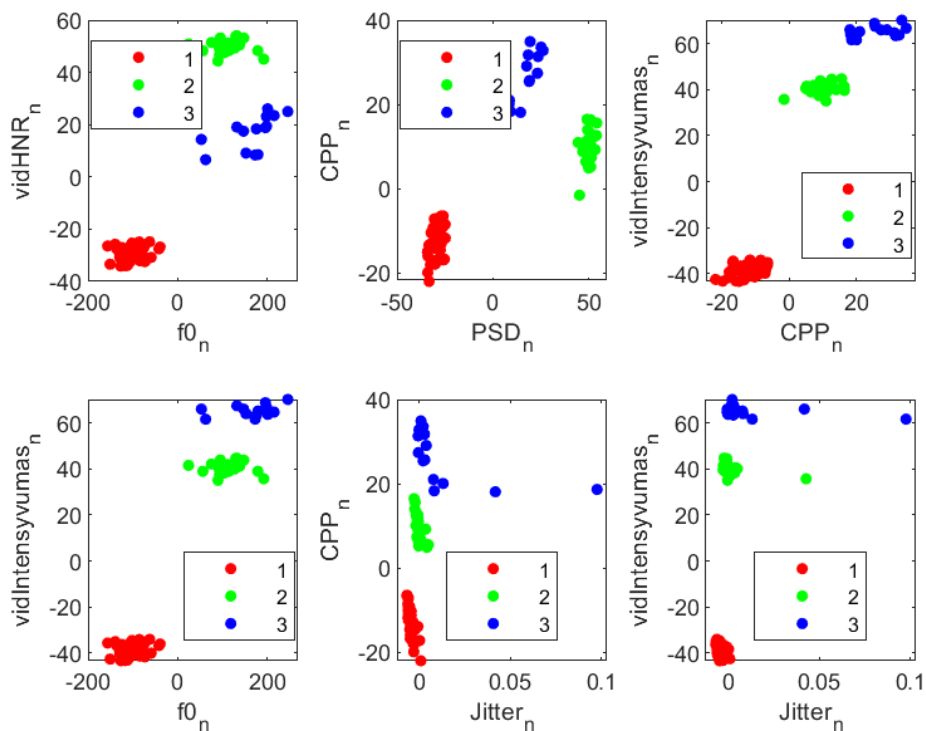
Eksperimento metu atlikta analizė su nenormalizuotais duomenimis. Rezultatai pateikti 4.3 lentelėje ir 4.4 lentelėje. Matyti, kad klasifikatorių *Taiklumas* nukrito. Pastebėta, kad nenormalizavus duomenų, Atraminų vektorių klasifikatoriaus algoritmas pasižymėjo didesniu taiklumu, nei kiti naudoti algoritmai. Pastebėta, kad Atraminų vektorių klasifikatoriaus klasifikavimas prastesnis, nei klasifikavimas kitais algoritmais. 4.4 lentelėje matyti, kad klasifikuojant 3 klases, paprastas Bajeso ir giliojo mokymo neuroninio tinklo algoritmų *Taiklumas* mažesnis, nei algoritmų, apmokytų atpažinti normalizuotus duomenis.

4.3 lentelė. Eksperimento rezultatai, klasifikuojant 2 klases, duomenys nenormalizuoti

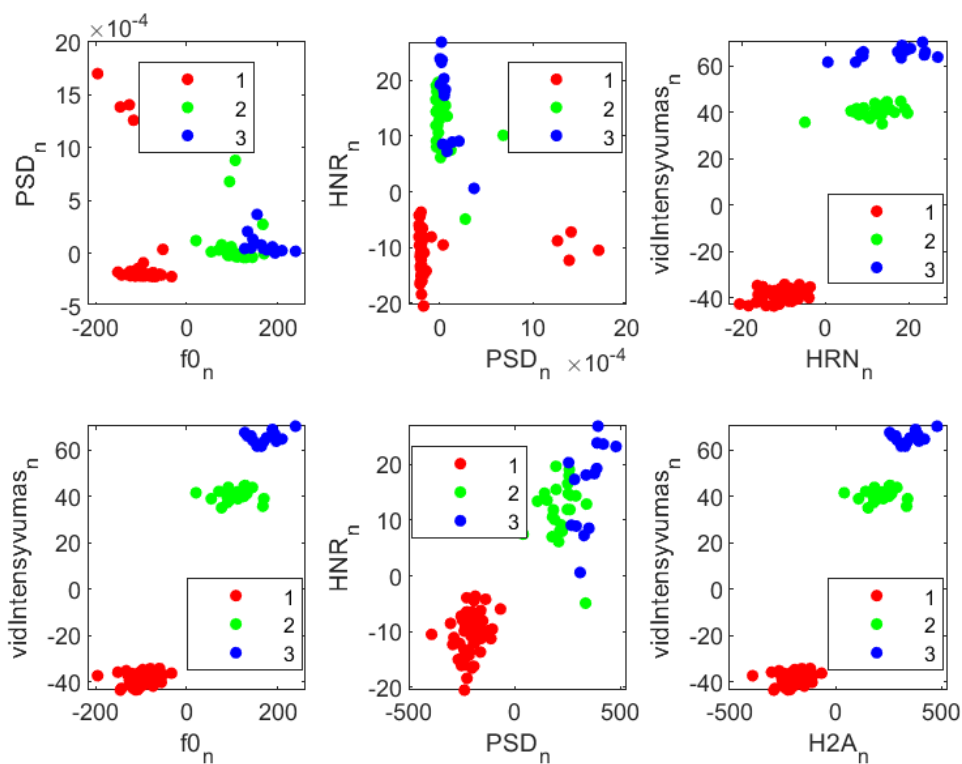
| | Rinkinio tipas | Atraminų vektorių klasifikatorius | Bajeso klasifikatorius | Gilaus mokymo neuroninis tinklas |
|----------------|---------------------|-----------------------------------|------------------------|----------------------------------|
| Vyrai | Sudarytas rinkinys | 98,31 % | 86,44 % | 77,96 % |
| | Etaloninis rinkinys | 98,31 % | 84,74 % | 81,35 % |
| Moterys | Sudarytas rinkinys | 79,03 % | 72,58 % | 79,03 % |
| | Etaloninis rinkinys | 100 % | 100 % | 100 % |



4.5 pav. 3 klasių klasifikatorių maišaties matrica, kai pacientai yra moterys, mokymo duomenys s normalizuoti



4.5 pav. Etalominio charakteristikų rinkinio duomenų dispersija, kai pacientai moteriškos lyties



4.6 pav. Sudaryto charakteristikų rinkinio duomenų dispersija, kai pacientai moteriškos lyties

4.4 lentelė. Eksperimento rezultatai, klasifikuojant 3 klases, duomenys nenormalizuoti

| | Rinkinio tipas | Atramiųjų vektorių klasifikatorius | Bajeso klasifikatorius | Gilias mokymo neuroninis tinklas |
|----------------|-----------------------|---|-------------------------------|---|
| Vyrai | Sudarytas rinkinys | 92,75 % | 60,87 % | 71,01 % |
| | Etaloninis rinkinys | 72,83 % | 75,31 % | 60,49 % |
| Moterys | Sudarytas rinkinys | 93,82 % | 62,34 % | 66,04 % |
| | Etaloninis rinkinys | 75,92 % | 69,13 % | 67,28 % |

Eksperimento metu gauti rezultatai palyginti su tyrimais, kurių metu balso klosčių patologijos yra vertinamos pasitelkus akustinių charakteristikų mažinimo algoritmus. Palyginimui pasirinktas Zuzanos Dankovičovos et al. 2018 m. tyrimas. Tyrime akustinių charakteristikų imtis mažinama principinių komponentų analizės (angl. *Principal Component Analysis*) metodu. Tyrimo metu autoriai gavo bendrą klasifikavimo taiklumą (*Taiklumas*) 80,6 % moteriškos lyties pacientams ir 86,2 % vyriškos lyties pacientams, kai klasifikuojami sveiki ir disfoniški balsai, o klasifikatoriai yra atramiųjų vektorių klasifikatorius, k -artimiausias kaimynas ir atsitiktinio miško klasifikatorius. Studijoje naudoti iš *Saarbruecken* duomenų bazės atsisųsti sveikų ir su diagnozuota disfonija pacientų balsų įrašai. Eksperimento metu naudojant tos pačios duomenų bazės įrašus ir atlikus klasifikavimo į sveikus ir disfoniškus balsus eksperimentą, gauta, kad eksperimento sąlygomis, sudarytas akustinių charakteristikų rinkinys yra tinkamas disfonijai nustatyti.

4.3. Rekomendacijos balso klosčių būklei vertinti

Įvertinus analitinius ir eksperimentinius darbo rezultatus suformuotos rekomendacijos balso klosčių būklei vertinti. Balso klosčių būklės vertinimas akustinės analizės metodais yra sudėtingas uždavinys. Dėl balso daugialypės ir netiesinės prigimties (kalbos aparatą sudaro sudėtinga organų sistema, valdoma nervų sistemos), balso klostės yra veikiamos įvairių tipų patologijų, iš kurių dažniausiai diagnozuojamos yra funkcinio pobūdžio patologijos, tokios kaip hiperfunkcinė disfonija, hipofunkcinė disfonija, balso klosčių polipai ir pan.

Balso klosčių patologijos dažnai yra diagnozuojamos kelios vienu metu. Dėl šios priežasties, patologijų klasifikavimo uždavinys tampa sudėtingu, nes tą patį organą (balso klostes) veikiančios patologijos turės panašią įtaką balso klostėms ir tuo pačiu metu lems akustinių charakteristikų pokyčius. Siūloma įprastų balso klosčių patologijų tyrimo problemą laikyti dvejetainio klasifikavimo uždaviniu. Tokiu atveju klasifikatoriai apmokomi didelėmis sveikų ir patologinių balsų imtimis ir klasifikatorius tampa pagalbine priemone pacientams įvertinti savo balso būklę.

Dėl balso akustinių charakteristikų kiekio ir įvairios prigimties, balso vertinimas su akustinėmis charakteristikomis yra sunkus. Žmogaus ausis gerai atskiria patologinius ir sveikus balsus arba bendrą

balso kokybę, bet šį uždavinį spręsti automatizuotais metodais yra sunku. Rekomenduojama balso klosčių būklės vertinimui kurti specifinius akustinių charakteristikų rinkinius, nusakančius konkrečias patologijas (pvz., mikčiojimas pasižymės aukštu balso lūžių indeksu (*NVB*)). Tokie rinkiniai įprastai yra pritaikomi paprastiems klasifikatoriams (žr. 4.2 poskyrio rezultatus), kurie lengvai ir greitai įgyvendinami programinėje įrangoje.

Akustinių charakteristikų ir jų pokyčių dėl balso klosčių patologijų rinkiniai yra jautrūs pasirinkto klasifikatoriaus tipams (žr. 4.2 poskyrio rezultatus). Jei nėra galimybės patologijoms sukurti specifinių akustinių charakteristikų rinkinių, geri klasifikavimo rezultatai gali būti pasiekiami optimizavus klasifikavimo algoritmus ar pritaikius matematinės transformacijos duomenis (pvz., normalizaciją). Šiuo atveju reikia įvertinti ar sistema turi pakankamai skaičiavimo resursų.

Bendru atveju balso klosčių ir balso vertinimas akustinės analizės metodais yra dėmesio verta tyrimų sritis. Akustinės charakteristikos ir jų pokyčiai gali būti vertinami įvairiais metodais, kurių palyginimas ir vertinimas suteikia įžvalgą į balso, balso kokybės ir balso estetikos prigimtį.

APIBENDRINIMAS. IŠVADOS

Darbo metu išnagrinėti ir ištirti akustinės analizės metodai balso klosčių funkciniai būklei vertinti ir patologijoms aptikti. Remiantis gautais analitiniais ir eksperimentiniais rezultatais, sukurti funkcinę būklę atspindintys akustinių charakteristikų rinkiniai ir rekomendacijos.

Atsižvelgus į darbo uždavinius, analitinius ir eksperimentinius rezultatus, suformuluotos šios išvados:

1. Dažniausiai balso klostes veikia funkcinio pobūdžio sutrikimai, nes šių sutrikimų priežastis yra įtemptas balso klosčių naudojimas. Šio pobūdžio balso sutrikimai yra būdingi pacientams, priklausantiems profesinėms grupėms, tokioms kaip mokytojai, dainininkai, prekybininkai, taip pat ir gamyklų darbininkai (tarp 50 % ir 80 % mokytojų yra bent kartą per gyvenimą skundęsi funkciniais balso sutrikimais).
2. Dėl mažesnių gerklų fizinių matmenų, moterims funkcinio pobūdžio sutrikimai yra diagnozuojami dažniau (65,3 % moterų *Saarbruecken* duomenų bazėje diagnozuoti funkcinio pobūdžio sutrikimai, kai tik 9,3 % vyrų buvo priskirta tokia pati diagnozė).
3. Pasireiškus funkciniam balso sutrikimams, balsas tampa kimumus, oringas, pakinta balso pagrindinis tonas, stiprumas, balsas lūžinėja. Šios balso savybės yra lengvai suprantamos pasitelkus garsinį suvokimą, bet yra sunkiai apibūdinamos pasitelkus akustines charakteristikas.
4. Balso patologijoms vertinti medicinoje naudojamos garsinio suvokimo skalės. Populiariausios yra *GRBAS* ir *CAPE-V* skalės, kuriomis įvertinamas balso disfoniškumas remiantis jutiminėmis balso charakteristikomis (pvz.: balso kimumu, šiurkštumu ir pan.). Šių skalių vertinimo tikslumas priklauso nuo personalo kompetencijos.
5. Vienodos metodikos balso klosčių būklei įvertinti akustinės analizės metodais nėra. Egzistuoja rekomendacijos, teigiančios, kad akustinių charakteristikų rinkinys privalo būti sudarytas remiantis balso ekspertų rekomendacijomis, jei siekiama apmokyti įvairaus pobūdžio klasifikatorius. Remiantis ekspertų rekomendacijomis, sudarytas akustinių charakteristikų iš balso pagrindinio tono (f_0), drebulio, mirguliavimo, balso stiprumo (I), harmonikos–triukšmo santykio (HNR).
6. Remiantis akustinių charakteristikų ir jutiminių charakteristikų koreliacijos vertinimu, nustatyta, kad funkcinis balso sutrikimus įmanoma apibūdinti šiomis charakteristikomis: balso pagrindiniu tonu (f_0), balso stiprumu (I), antrąja signalo harmonikos amplitude ($H2A$), signalo periodo standartinio nuokrypio logaritmu ($LNPSD$), harmonikos–triukšmo santykiu (HNR), balso lūžių kiekiu (NVB).

7. Apmokius atraminių vektorių klasifikatorių, paprastą Bajeso klasifikatorių ir giliojo mokymo neuroninio tinklo klasifikatorių sudarytais akustinių charakteristikų rinkiniais gauta, kad naudojant ekspertų rekomendacijomis paremtą charakteristikų rinkinį (etaloninį rinkinį), pasiekiamas vidutinis 82,74 % klasifikavimo *Taiklumas* vyrams ir 87,19 % moterims, kai duomenys klasifikuojami į tris klases (sveikų balsų klasė ir dvi patologijų grupių klasės). Šie rezultatai gauti naudojant normalizuotus mokymo duomenis. Klasifikavimo *Taiklumas* pagerintas klasifikavimo algoritmus apmokius koreliacija paremtu charakteristikų rinkiniu (sudarytu rinkiniu). Gauta, kad įmanomas vidutiniškas 96,13 % vyrų ir 88,12 % moterų pataloginių balsų atpažinimo tikslumas.
8. Klasifikuojant balsą į dvi klases (sveiki ir disfoniški balsai), gauta, kad naudojant ekspertų rekomendacijomis paremtą charakteristikų rinkinį (etaloninį rinkinį), pasiekiamas vidutinis 94,35 % klasifikavimo *Taiklumas* vyrams ir 93,01 % moterims, kai duomenys klasifikuojami į dvi klases (sveikų balsų klasė ir disfoniško balso klasės). Šie rezultatai gauti naudojant normalizuotus mokymo duomenis. Klasifikavimo *Taiklumas* pagerintas klasifikavimo algoritmus apmokius koreliacija paremtu charakteristikų rinkiniu (sudarytu rinkiniu). Gauta, kad eksperimento sąlygomis įmanoma pasiekti 100 % taiklumą (*Taiklumas*) vyrų ir moterų balsų imtims. Šie rezultatai reikalauja papildomų eksperimentų, kad būtų patvirtintas šių duomenų tikslumas.

Įvertinus suformuotas išvadas, galima teigti, kad darbo tikslas yra įgyvendintas. Sudarytas akustinių charakteristikų rinkinys yra tinkamas balso klosčių funkcinėi būklei vertinti ir funkcinio pobūdžio patologijoms atpažinti. Pastebėta, kad šio charakteristikų rinkinio pritaikomumas ne balso klosčių patologijoms atpažinti reikalauja papildomų tyrimų. Pastebėta galimybė papildomiems tyrimams, paremtiems akustinės analizės metodais, kurių metu įvertinamas balso klosčių patologijų ūmumas.

Monika Danilovaitė



LITERATŪRA

1. Ramig, L. O.; Verdolini, K. (1998). Treatment efficacy: Voice disorders. *Journal of Speech, Language, and Hearing Research* 41, 101–116.
2. Roy, N.; Merrill, R. M.; Thibeault, S.; Parsa, R. A.; Gray, S. D.; Smith, E. M. (2004). Prevalence of voice disorders in teachers and the general population. *Journal of Speech, Language, and Hearing Research* 47, 281–293.
3. Americal Speech-Language-Hearing Association. *Voice Disorders*.(2018).
<https://www.asha.org/PRPSpecificTopic.aspx?folderid=8589942600§ion=Assessment>
4. Mourao, J.; Moreira, J.; Barbosa, J.; Carvalho, J.; Tavares, J. 2015. Soft tissue injuries after direct laryngoscopy. *Journal of Clinical Anesthesia* 27(8), 668–671.
5. Gupta, R.; Chaspari, T.; Kim, J.; Kumar, N.; Bone, D.; Naryanan, Sh. (2016). Pathological speech processing: State-of-the-art, current challenges, and future directions. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6470–6474.
6. Halawa, W. E.; Rodriguez Fernandez Freire, A.; Munoz, I.V.; Perez, S. S. (2014). Assessment of effectiveness of acoustic analysis of voice for monitoring the evolution of vocal nodules after vocal treatment. *Eur Arch Otorhinolaryngol* 271(4), 749–56.
7. Vizza, P.; Mirarchi, D.; Tradigo, G.; Redavide, M.; Bossio, R. B.; Veltri, P. (2017). Vocal signal analysis in patients affected by Multiple Sclerosis. *Procedia Computer Science* (108), 1205–1214.
8. Braga, D.; Madureira, A. M.; Coelho, L.; Ajith, R. (2018). Automatic detection of Parkinson’s disease based on acoustic analysis of speech. *Engineering Applications of Artificial Intelligence* (77), 148–158.
9. Titze, I. R. (1994). Toward standards in acoustic analysis of voice. *J. Voice* 8(1), 1–7.
10. Vegienė, A. (2014). Balso daugiaparametrio tyrimo sisteminės analizės reikšmė pirminei gerklų ligų atrankai. *Institutional Repository of Lithuanian University of Health Sciences*, 146–147.
11. Dankovičová, Z.; Sovák, D.; Drotar, P.; Vokorokos, L. (2018). Machine Learning Approach to Dysphonia Detection. *Applied Sciences*, 8, 1927.
12. Iqbal, M.; Manzoor, T.; Ibrahim, M.; Tahira, S.; Azmat, R. (2015). *VOCAL HYGIENE PRACTICE AMONG PROFESSIONAL VOICE USERS*.
https://www.researchgate.net/publication/301549160_VOCAL_HYGIENE_PRACTICE_AMONG_PROFESSIONAL_VOICE_USERS
13. Ward, D. (2006). *Stuttering and Cluttering: Frameworks for understanding treatment: vadovėlis*. Psychology Press.
14. National Center for Voice and Speech. (2018). *Voice Qualities*.
<http://www.ncvs.org/ncvs/tutorials/voiceprod/tutorial/quality.html>
15. Titze, I.R. and Martín, D.W. (1994). Principles of voice production. *The Journal of the Acoustical Society of America* 104(3), 1148–1148.
16. Yip, M. (2002). *Tone: vadovėlis*. Cambridge University Press.

17. National Center for Voice and Speech. (2018). *3 Levels of Intensity Control: Lungs, Glottis & Vocal Tract*. <http://www.ncvs.org/ncvs/tutorials/voiceprod/tutorial/three.html>
18. GPNotebook. (2006). *GRBAS (grade, roughness, breathiness, asthenia, strain)*. <https://www.gpnotebook.co.uk/simplepage.cfm?ID=x20161228165732191130>
19. ASHA Special Interest Division 3. (2006). *Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V)*. <https://www.asha.org/uploadedfiles/members/divs/d3capevprocedures.pdf>
20. Martins, R. H., do Amaral, H. A.; Tavares, E. L.; Martins, M. G.; Gonçalves, T. M.; Dias, N. H. (2015). Voice disorders: Etiology and diagnosis. *Journal of Voice*. 30(6), 761.e1-761.e9.
21. Hagen, P.; Lyons, G. D. (1996). Dysphonia in the elderly: Diagnosis and management of age-related voice changes. *Southern Medical Journal*, 89, 204–207.
22. Martins, R. H.; Hidalgo Ribeiro, C. B.; Fernandes de Mello, B. M.; Branco, A.; Tavares, E. L. (2012). Dysphonia in children. *Journal of Voice*, 26, e17–e20.
23. Stemple, J. C.; Roy, N.; Klaben, B. G. (2014). *Clinical voice pathology: Theory and management*. CA: Plural.
24. Verdolini, K.; Rosen, C.; Branski, R. C. (Eds.). (2006). *Classification manual for voice disorders-I*, NJ: Erlbaum.
25. Watson, S. (2018). *Caring for and Preventing Vocal Nodules*. <https://www.healthline.com/health/vocal-nodules>
26. Galindo, G. E.; Peterson S. D.; Erath B. D.; Castro, C.; Hillman, R. E.; Zanatu, M. (2017). Modeling the Pathophysiology of Phonotraumatic Vocal Hyperfunction With a Triangular Glottal Model of the Vocal Folds. *J Speech Lang Hear Res.*, 60(9), 2452–2471.
27. University of Pittsburgh, Department of Otolaryngology. *Muscle Tension Dysphonia*. (2020). <http://www.otolaryngology.pitt.edu/centers-excellence/voice-center/conditions-we-treat/muscle-tension-dysphonia>
28. Columbia University Department of Otolaryngology Head and Neck Surgery. (2019). *Ventricular Phonation (Plica Ventricularis)*. <https://www.entcolumbia.org/health-library-0/ventricular-phonation-plica-ventricularis>
29. Kerkar, P. (2018). *What is Vocal Fatigue, Know its Causes, Symptoms, Treatment*. <https://www.epainassist.com/face-mouth-throat/what-is-vocal-fatigue>
30. Kostek, B.; Wierzchowska, A. (1997). Parametric representation of musical sounds. *Archives of Acoustics* 22, 1, 3–26.
31. Keisuke, I.; Nobutaka, O. (2016). *Online Acoustic Scene Analysis Based on Nonparametric Bayesian Model*. Pranešimas konferencijoje: European Signal Processing Conference (EUSIPCO), Budapeštas, Vengrija.
32. Chen, M. Y. (1995) Acoustic parameters of nasalized vowels in hearing-impaired and normal-hearing speakers. *The Journal of Acoustical Society of America*, 98(5), 2443–2453.
33. Kostek, B. (1995). Computer Based Recognition of Musical Phases Using the Rough Set Approach. *2nd Annual Joint Conference on Inform. Sciences*. 401–404.

34. Gunjawate, D. R.; Ravi, R.; Bellur, R. (2018). Acoustic Analysis of Voice in Singers: A Systematic Review. *Journal of Speech, Language, and Hearing Research*, 40–51.
35. Gerratt, B. R.; Kreiman, J.; Garellek, M. (2016). Comparing Measures of Voice Quality from Sustained Phonation and Continuous Speech. *Journal of Speech, Language, and Hearing Research, Volume 59 Issue 5*, 994–1001.
36. Selvakumari, S.; Radha, V. (2017). Voice Pathology Identification: A Survey on Voice Disorder. *I.J. Engineering and Manufacturing 2*, 39–49.
37. Gupta, R.; Chaspari, T.; Kim J.; Kumar, N.; Bone, D.; Narayanan, Sh. (2016). Pathological speech processing: State-of-the-art, current challenges, and future directions. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6470 – 6474.
38. Rizzo, G.; Copetti, M.; Arcuti, S.; Martino, D.; Fontana, A.; Logroscino, G. (2016). Accuracy of clinical diagnosis of Parkinson disease. A systematic review and meta-analysis. *Neurology 86 (6)*, 566–576.
39. Cheveigné, A.; Kawahara, H. (2002). YIN, A fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America, 111(4)*, 1917-193.
40. Strömbergsson, S. (2016). Today's Most Frequently Used F0 Estimation Methods, and Their Accuracy in Estimating Male and Female Pitch in Clean Speech. *Conference: Interspeech*, 525–529.
41. Chu, W., Alwan, A. (2009). Reducing F0 Frame Error of F0 tracking algorithms under noisy conditions with an unvoiced/voiced classification frontend. *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 3969–3972.
42. Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *IFA Proceedings 17*, 97–110.
43. Boersma, P. (2003). *Sound: To Pitch (ac)*. http://www.fon.hum.uva.nl/praat/manual/Sound_To_Pitch_ac.html
44. Boersma, P.; Weenink, D. (2003). *Intro 4.2. Configuring the pitch contour*. http://www.fon.hum.uva.nl/PRAAT/manual/Intro_4_2__Configuring_the_pitch_contour.html
45. Boersma, P.; Weenink P. (2003). *Praat: doing phonetics by computer*. <http://www.fon.hum.uva.nl/praat/>
46. Ishwar, I. (2005). *Spectrogram -iua-.png*. https://commons.wikimedia.org/wiki/File:Spectrogram_-iua-.png
47. Shrivastav, R.; Eddins, D.; Anand, S. (2012). Pitch strength of normal and dysphonic voices. *The Journal of the Acoustical Society of America, 131(3)*, 2261–2269.
48. Smith, S. W. *Digital signal procesing*. <http://www.dspguide.com/CH10.PDF>
49. Santosh, M.; Rajashekhar, B. (2011). Perceptual and acoustic analysis of voice in individuals with total thyroidectomy: pre-post surgery comparison [corrected]. *Indian journal of otolaryngology and head and neck surgery: official publication of the Association of Otolaryngologists of India, 63(1)*, 32–39.
50. Gunjawate, D.; Ravi, R.; Rajashekhar, B. (2018). Acoustic Analysis of Voice in Singers: A Systematic Review. *Journal of Speech Language and Hearing Research, 61(1)*, 40-51.
51. Barsties v. L. B.; Maryn, Y.; Gerrits, E.; Bodt, M. (2018). A Meta-Analysis: Acoustic Measurement of Roughness and Breathiness. *Journal of Speech Language and Hearing Research, 61(2)*, 298-323.

52. Sachdeva, K.; Shrivastava, T. (2019). Dysphonia and its Correlation with Acoustic Voice Parameters. *International Journal of Phonosurgery & Laryngology*, 8.
53. Santosh, M.; Rajashekhar, B. (2011). Perceptual and acoustic analysis of voice in individuals with total thyroidectomy: pre-post surgery comparison [corrected]. *Indian journal of otolaryngology and head and neck surgery: official publication of the Association of Otolaryngologists of India*, 63(1), 32–39.
54. Sabir, B.; Touri, B.; Moussetad, M. (2017). Correlation between acoustic measures, voice handicap index and GRBAS scales scores among Moroccan students. *Current Pediatric Research*, 21, 343–353.
55. Ruzs, J.; Klempíř, J.; Baborová, E.; Tykalová, T.; Majerová, V.; Cmejla, R.; Růžička, E.; Roth, J. (2013). Objective acoustic quantification of phonatory dysfunction in Huntington's disease. *PloS one*, 8(6), e65881.
56. Gunjawate, D. R.; Ravi, R.; Bellur, R. (2018). Acoustic analysis of voice in singers: A systematic review. *Journal of Speech, Language, and Hearing Research*, 61(1), 40–51.
57. Prakup B. (2012). Acoustic measures of the voices of older singers and nonsingers. *J Voice*, 26(3) 341-350.
58. Guzman, M.; Angulo, M.; Muñoz, D.; Mayerhoff, R. Effect on long-term average spectrum of pop singers' vocal warm-up with vocal function exercises. *Int J Speech Lang Pathol.*, 15(2), 127-135.
59. Ferrand, C. (2001). Speech Science: An Integrated Approach to Theory and Clinical Practice. *Ear and Hearing*, 22, 549.
60. Wendahl, R.W. (1966). Laryngeal analog synthesis of jitter and shimmer auditory parameters of harshness. *Folia Phoniatica*, 18, 98–108.
61. Heiberger, V.L.; Horii, Y. (1982). Jitter and shimmer in sustained phonation. Iš N.J. Lass (Ed.), *Speech and Language: Advances in Basic Research and Practice*, Vol. 7. (pp. 299–332). New York: Academic Press.
62. Hillenbrand, J.M. (1988). Perception of aperiodicities in synthetically generated voices. *Journal of the Acoustical Society of America*, 83, 2361–2371.
63. Barry, W. J.; Pützer, M. *Saarbrücken Voice Database, Institute of Phonetics, Univ. of Saarland*. <http://www.stimmdatenbank.coli.uni-saarland.de/>
64. Cohen, S. M.; Kim, J.; Roy, N.; Asche, C.; Courey, M. (2012). Prevalence and causes of dysphonia in a large treatment-seeking population. *The Laryngoscope*, 122, 343–348.
65. Fritzell, B. (1996). Voice disorders and occupations. *Logopedics Phoniatics Vocology*, 21, 7–12.
66. Miller, M.; Verdolini, K. (1995). Frequency of voice problems reported by teachers of singing and control subjects and risk factors. *Journal of Voice*, 8, 348–362.
67. Thibeault, S. L.; Merrill, R. M.; Roy, N.; Gray, S. D.; Smith, E. M. (2004). Occupational risk factors associated with voice disorders among teachers. *Annals of Epidemiology*, 14, 786–792.
68. Williams, N. R. (2003). Occupational groups at risk of voice disorders: A review of the literature. *Occupational Medicine*, 53, 456–460.
69. Cantor Cutiva, L. C.; Vogel, I.; Burdorf, A. (2013). Voice disorders in teachers and their associations with work-related factors: A systematic review. *Journal of Communication Disorders*, 46, 143–155.

70. Martins, R. H.; do Amaral, H. A.; Tavares, E. L.; Martins, M. G.; Gonçalves, T. M.; Dias, N. H. (2015). Voice disorders: Etiology and diagnosis. *Journal of Voice*, 30(6), 761.e1-761.e9.
71. Michigan Medicine. (2020). *Functional Dysphonia*. <https://www.uofmhealth.org/conditions-treatments/ear-nose-throat/functional-dysphonia>
72. Bänziger, T.; Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication*, 46(3–4), 252–267.
73. Jardim R.; Barreto S. M.; Assunção A. A. (2007). Condições de trabalho, qualidade de vida e disfonia entre docentes [Work conditions, quality of life, and voice disorders in teachers]. *Cad Saude Publica*, 23(10), 2439-2461.
74. KILIÇ, M.; Oguz, H.; şafak, M. A. (2011). Comparison of results in two acoustic analysis programs: Praat and MDVP. *Turkish Journal of Medical Sciences*. 41, 835–841.
75. Drotar, P.; Smekal, Z. Comparative study of machine learning techniques for supervised classification of biomedical data. *Acta Electrotech. Inform.*, 14, 5–11.
76. Alemami, Y.; Almazaydeh, L. (2018). Pathological Voice Signal Analysis Using Machine Learning Based Approaches. *Computer and Information Science*, 11, 8-13.
77. Fang, S. H.; Tsao, Y.; Hsiao, M. J.; Chen, J. Y.; Lai, Y. H.; Lin, F. C.; Wang, C. T. (2018). Detection of Pathological Voice Using Cepstrum Vectors: A Deep Learning Approach. *Journal of Voice*, 33(5), 634-641
78. Verde, L.; De Pietro, G.; Sannino, G. (2018). Voice Disorder Identification by using Machine Learning Techniques. *IEEE Access*. 1-1.
79. Fürnkranz, J. (2002). Round Robin Classification. *J. Mach. Learn. Res.*, Vol. 2, 721–747.
80. Escalera, S.; Pujol O.; Radeva P. (2010). On the decoding process in ternary error-correcting output codes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 32, Issue 7, 120–134.
81. Manning, C. D.; Raghavan, P.; Schütze, M. (2008). *Introduction to Information Retrieval*. Cambridge University Press.
82. Deng, L.; Yu, D. (2013). Deep Learning: Methods and Applications. *Foundations and Trends® in Signal Processing: Vol. 7: No. 3–4*, 197-387.
83. Hagan, M.T.; Menhaj, M. (1999). Training feed-forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, Vol. 5, No. 6, 989–993.
84. Zhu, W.; Zeng, N.; Wang, N. (2010). *Sensitivity, Specificity, Accuracy, Associated Confidence Interval and ROC Analysis with Practical SAS® Implementations*. NorthEast SAS users group, health care and life sciences. <https://www.lexjansen.com/nesug/nesug10/hl/hl07.pdf>
85. Boersma, P.; Weenink, D. (2003). *Praat: doing phonetics by computer*. <http://www.fon.hum.uva.nl/praat/>

PRIEDAI

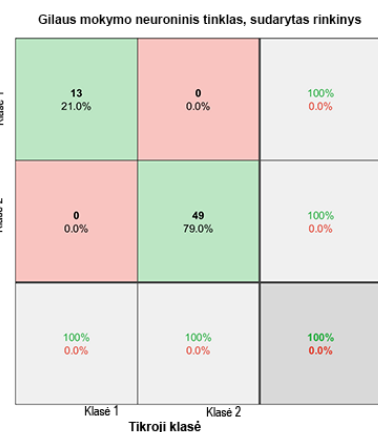
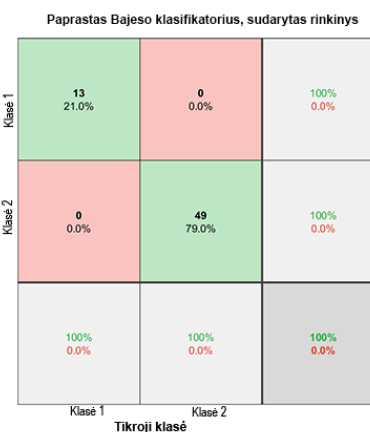
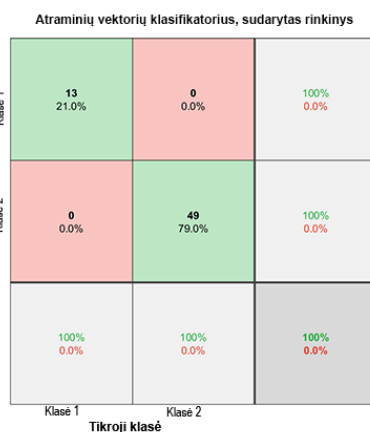
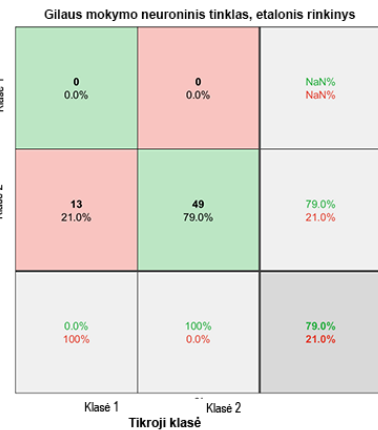
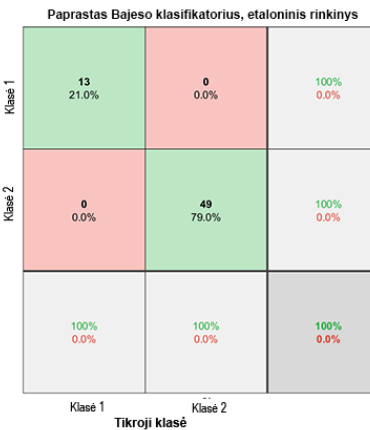
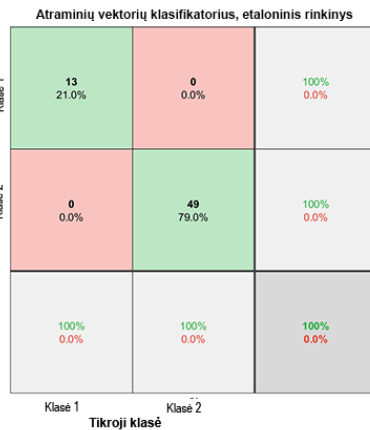
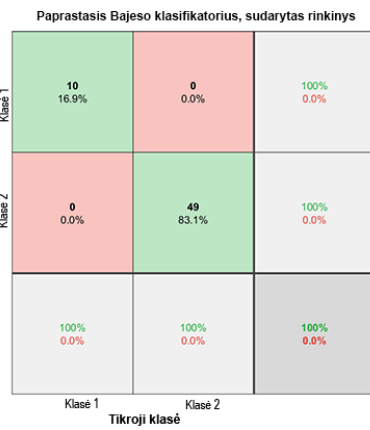
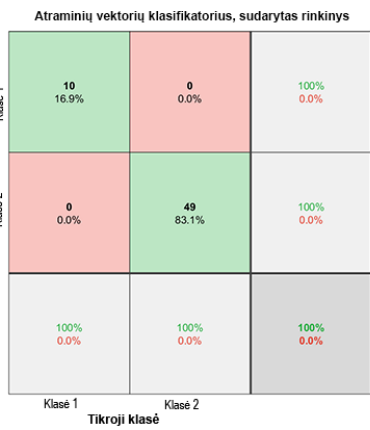
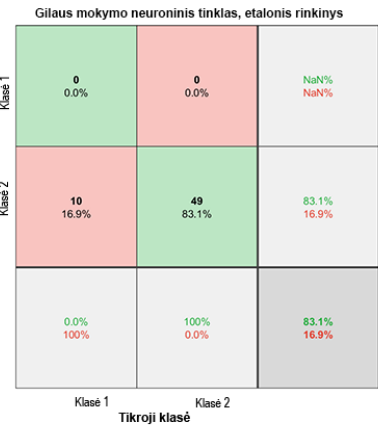
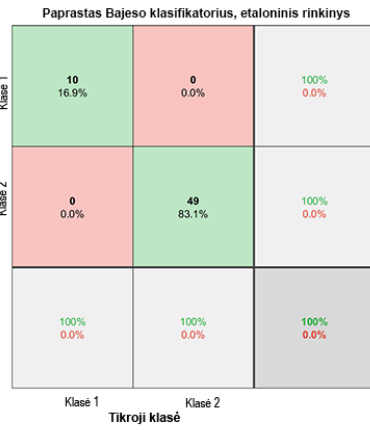
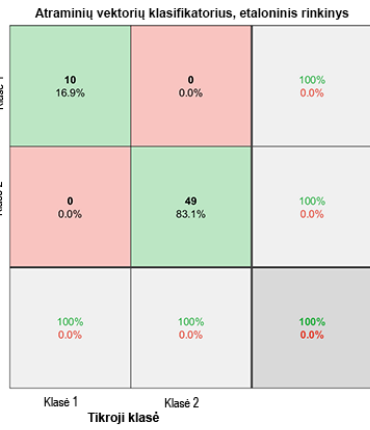
A priedas. Balso klosčių būklės vertinimo analitiniai ir eksperimentiniai rezultatai

| | Rinkinio tipas | Atraminų vektorių klasifikatorius | Bajeso klasifikatorius | Gilaus mokymo neuroninis tinklas |
|----------------|-----------------------|--|-------------------------------|---|
| Vyrai | Sudarytas rinkinys | 100 % | 100 % | 100 % |
| | Etaloninis rinkinys | 100 % | 100 % | 83,05 % |
| Moterys | Sudarytas rinkinys | 100 % | 100 % | 100 % |
| | Etaloninis rinkinys | 100 % | 100 % | 79,03 % |

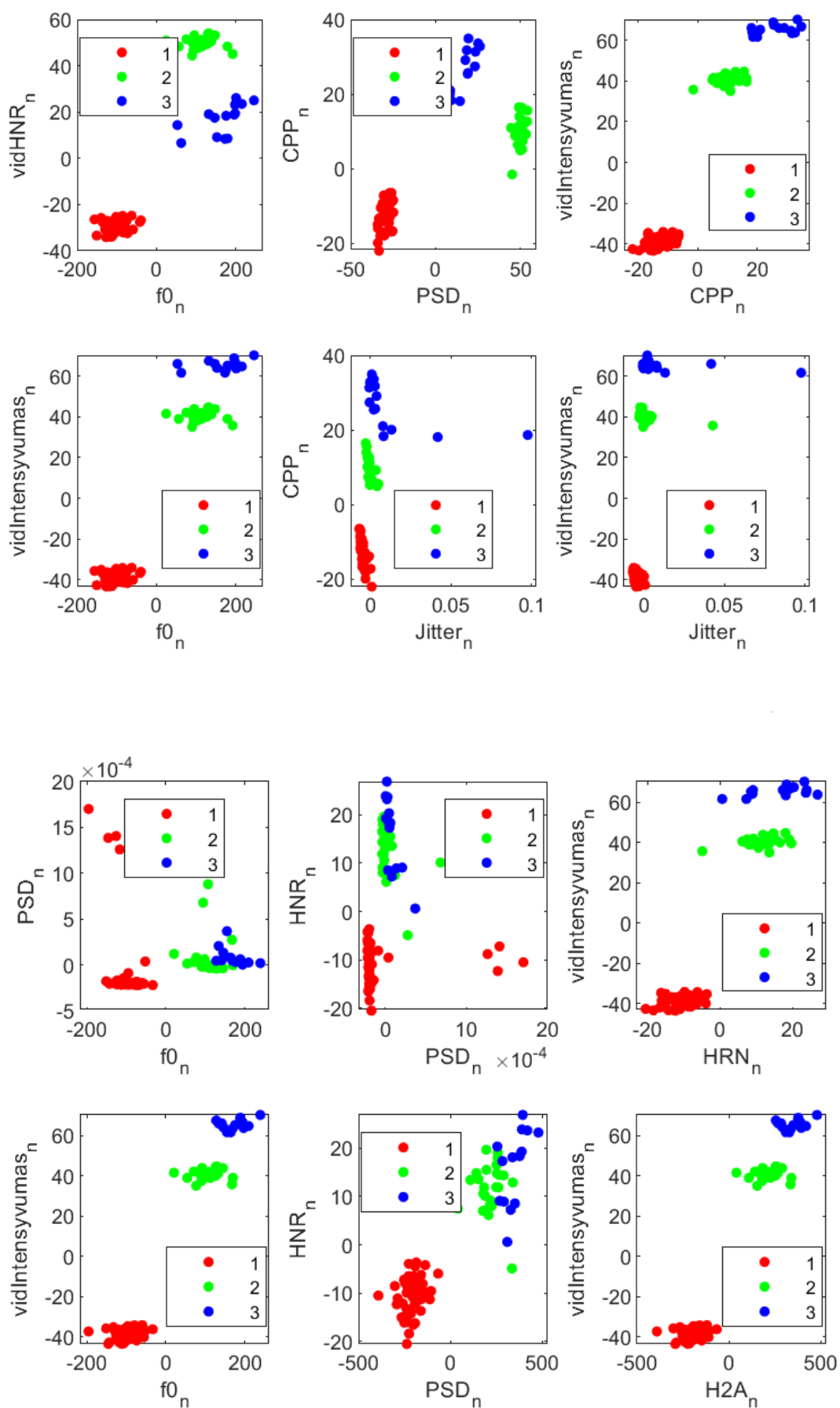
| | Rinkinio tipas | Atraminų vektorių klasifikatorius | Bajeso klasifikatorius | Gilaus mokymo neuroninis tinklas |
|----------------|-----------------------|--|-------------------------------|---|
| Vyrai | Sudarytas rinkinys | 89,85 % | 100 % | 98,55 % |
| | Etaloninis rinkinys | 80,34 % | 91,35 % | 76,54 % |
| Moterys | Sudarytas rinkinys | 80,45 % | 83,91 % | 100 % |
| | Etaloninis rinkinys | 83,91 % | 95,40 % | 82,27 % |

| | Rinkinio tipas | Atraminų vektorių klasifikatorius | Bajeso klasifikatorius | Gilaus mokymo neuroninis tinklas |
|----------------|-----------------------|--|-------------------------------|---|
| Vyrai | Sudarytas rinkinys | 98,31 % | 86,44 % | 77,96 % |
| | Etaloninis rinkinys | 98,31 % | 84,74 % | 81,35 % |
| Moterys | Sudarytas rinkinys | 79,03 % | 72,58 % | 79,03 % |
| | Etaloninis rinkinys | 100 % | 100 % | 100 % |

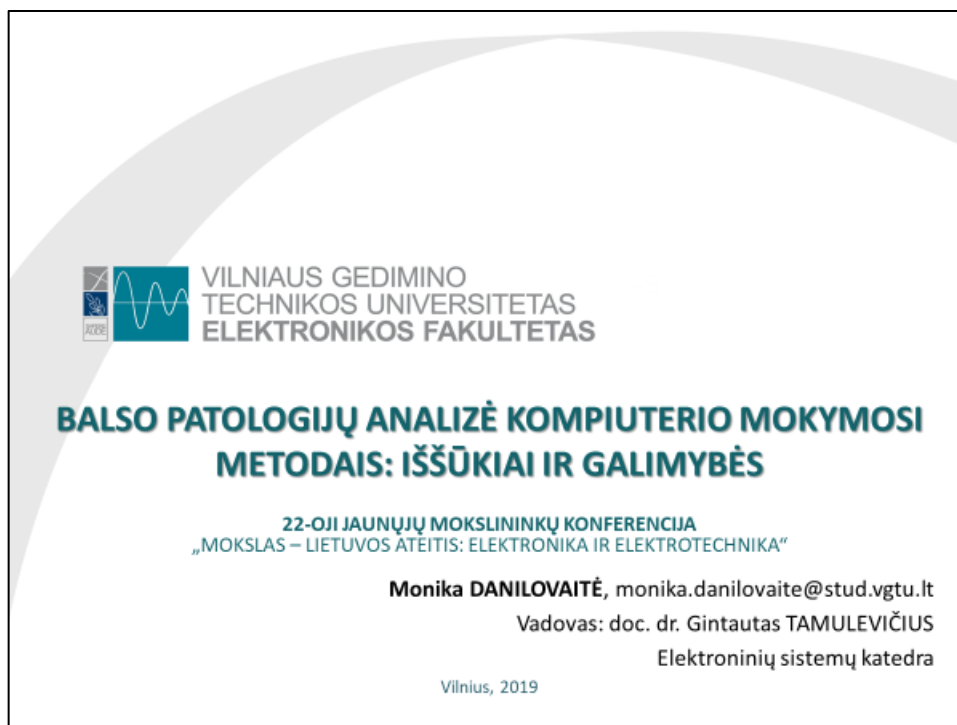
| | Rinkinio tipas | Atraminų vektorių klasifikatorius | Bajeso klasifikatorius | Gilaus mokymo neuroninis tinklas |
|----------------|-----------------------|--|-------------------------------|---|
| Vyrai | Sudarytas rinkinys | 92,75 % | 60,87 % | 71,01 % |
| | Etaloninis rinkinys | 72,83 % | 75,31 % | 60,49 % |
| Moterys | Sudarytas rinkinys | 93,82 % | 62,34 % | 66,04 % |
| | Etaloninis rinkinys | 75,92 % | 69,13 % | 67,28 % |








B priedas. Pranešimo 22-oje Lietuvos jaunųjų mokslininkų konferencijoje medžiaga



 VILNIAUS GEDIMINO
TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS

BALSO PATOLOGIJŲ ANALIZĖ KOMPIUTERIO MOKYMOŠI METODAIS: IŠŠŪKIAI IR GALIMYBĖS

22-OJI JAUNŲJŲ MOKSLININKŲ KONFERENCIJA
„MOKSLAS – LIETUVOS ATĖITIS: ELEKTRONIKA IR ELEKTROTECHNIKA“

Monika **DANILOVAITĖ**, monika.danilovaite@stud.vgtu.lt
Vadovas: doc. dr. Gintautas TAMULEVIČIUS
Elektroninių sistemų katedra

Vilnius, 2019



 VILNIAUS GEDIMINO
TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS

Monika **DANILOVAITĖ** 2019-03-15
Jaunųjų mokslininkų konferencija
„Mokslas – Lietuvos ateitis: Elektronika ir elektrotechnika“

Turinys

Įvadas

Problema ir tikslai

- Kalbos padargai
- Balso patologijos
- Galimybės
- Iššūkiai
- Studijose naudojami metodai

Apibendrinimas

2 11

VILNIAUS GEDIMINO
TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS

Monika DANILOVAITĖ 2019-03-15

Jaunųjų mokslininkų konferencija
„Mokslas – Lietuvos ateitis: Elektronika ir elektrotechnika“

Įvadas

Balso sutrikimai vargina tarp 3% ir 9% JAV populiacijos, bet 0,98% iš visų registruotų ligonių gydomi dėl balso sutrikimų (Ramig 1998; Roy 2005).

Balso sutrikimai diagnozuojami gydytojo, bet pastebima, jog laringoskopijos metu atsiranda didelė minkštųjų audinių traumos tikimybė (net 52,1%, pagal Mourao 2015) ir atsiranda žmogiškos klaidos pavojus

3 11

VILNIAUS GEDIMINO
TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS

Monika DANILOVAITĖ 2019-03-15

Jaunųjų mokslininkų konferencija
„Mokslas – Lietuvos ateitis: Elektronika ir elektrotechnika“

Problema ir tikslai

Problema

Vieningos metodikos trūkumas vertinant balso patologijas

Pranešimo tikslas

Kompiuterio mokymosi ir signalų apdorojimo metodų taikymo balso patologijų apdorojimui iššūkių ir galimybių apžvalga

4 11

VILNIAUS GEDIMINO
TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS

Monika DANILOVAITĖ 2019-03-15

Jaunųjų mokslininkų konferencija
„Mokslas – Lietuvos ateitis: Elektronika ir elektrotechnika“

Kalbos padargai



- Balso aparatas
- Artikuliacijos aparatas
- Oro aparatas

5 11

VILNIAUS GEDIMINO
TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS

Monika DANILOVAITĖ 2019-03-15

Jaunųjų mokslininkų konferencija
„Mokslas – Lietuvos ateitis: Elektronika ir elektrotechnika“

Balso patologijos

Trys grupės (pagal American Speech-Language-Hearing association):

- **Organiniai (struktūriniai arba neurologiniai)**
Atsiranda dėl struktūrinių pakitimų balse arba nervų sistemos pakitimų
- **Funkciniai**
Atsiranda dėl įtempto balso naudojimo
- **Psichogeniniai**
Stresas, nerimas, depresija

6 11

VILNIAUS GEDIMINO
TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS

Monika DANILOVAITĖ 2019-03-15

Jaunųjų mokslininkų konferencija
„Mokslas – Lietuvos ateitis: Elektronika ir elektrotechnika“

Galimybės

- **Hormoniniai pokyčiai**
- **Parkinsono liga**
- **Išsėtinė sklerozė**
- **Sutrikimai atsiradę dėl balso nuovalgio**

7 11

VILNIAUS GEDIMINO
TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS

Monika DANILOVAITĖ 2019-03-15

Jaunųjų mokslininkų konferencija
„Mokslas – Lietuvos ateitis: Elektronika ir elektrotechnika“

Iššūkiai

Nėra nustatyta vieninga metodika balso patologijų vertinimui kompiuterio mokymosi metodais

- **Pacientų imties sudarymas**
Paciento savitos balsinės charakteristikos
- **Balso patologijos yra linkusios pasireikšti kelios vienu metu**
- **Įrašo turinys**
Esant SVM algoritmui /a/ - 77.55%, o /i/ - 67.35%) (Dankovičová, 2018)
Fonacijos atsiradusios dėl skirtumų tarp kalbų
- **Sunkumai sudarant požymių imtį**
Persipildymo problema

8 11

VILNIAUS GEDIMINO
TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS

Monika DANILOVAITĖ 2019-03-15

Jaunųjų mokslininkų konferencija
„Mokslas – Lietuvos ateitis: Elektronika ir elektrotechnika“

Studijose naudojami metodai

- Lyginant **SVM** (angl. Support Vector Machine), **KNN** (angl. K-Nearest Neighbour) ir **RFN** (angl. Random Forest Classifier), SVM geriausias (Dankovičová, 2018)
- Lyginant **KNN** ir **Naive Bayess**, Bayess pasižymi geresniu patologijų aptikimu (93,3 % prieš 80 %) (Alemami, 2017)
- Lyginant **DNN** (angl. Deep Neural Networks) ir **Gaussian Mixture Model**, DNN pasižymi geresniais rezultatais (99.32% aptikimas) (Fang, 2018)
- Hemmerling 2015 atliktas tyrimas gavo 100 % tikslumą naudojant **RFN**
- Siekiant aptikti Parkinsono ligą, gautas 85 % tikslumas naudojant **ANN** (angl. Artificial Neural Networks), kai ekspertų pirminės diagnozės tikslumas 79.6% (Rizzo, 2016)

9 11

VILNIAUS GEDIMINO
TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS

Monika DANILOVAITĖ 2019-03-15

Jaunųjų mokslininkų konferencija
„Mokslas – Lietuvos ateitis: Elektronika ir elektrotechnika“

Apibendrinimas

- **Iš balso pokyčių įmanoma nustatyti įvairaus tipo sutrikimus**
- **Yra galimybių nagrinėjant senatinių ligų poveikį balsui ankstyvosiose ligų stadijose**
- **Dėl didelio informacijos kiekio balse, sunku sudaryti požymių rinkinius**
- **Aptinkant tam tikras balso patologijas, jau yra pasiektas 100% tikslumas**

10 11



VILNIAUS GEDIMINO
TECHNIKOS UNIVERSITETAS
ELEKTRONIKOS FAKULTETAS

Jaunųjų mokslininkų konferencija
„Mokslas – Lietuvos ateitis: Elektronika ir elektrotechnika“

Ačiū už dėmesį

11