

KAUNO TECHNOLOGIJOS UNIVERSITETAS
VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

JONAS MINELGA

**SKAITINIO INTELEKTO METODAI
NEINVAZINEI GERKLŲ LIGŲ
DIAGNOSTIKAI**

Daktaro disertacijos santrauka
Technologijos mokslai, informatikos inžinerija (07T)

2018, Kaunas

Disertacija rengta 2012 – 2017 metais Kauno technologijos universiteto Elektros ir elektronikos fakultete Elektros energetikos sistemų katedroje.

Mokslinis vadovas:

Prof. Dr. Habil. Antanas VERIKAS (Kauno technologijos universitetas, technologijos mokslai, informatikos inžinerija, 07T).

Redagavo: Virginija Stankevičienė (Kauno technologijos universitetas, Užsienio kalbų centras)

Informatikos inžinerijos mokslo krypties disertacijos gynimo taryba:

Prof. Dr. Robertas DAMAŠEVIČIUS (Kauno technologijos universitetas, technologijos mokslai, informatikos inžinerija, 07T) – **pirmininkas**;

Doc. Dr. Nikolaj GORANIN (Vilniaus Gedimino technikos universitetas, technologijos mokslai, informatikos inžinerija, 07T);

Prof. Dr. Rytis MASKELIŪNAS (Kauno technologijos universitetas, technologijos mokslai, informatikos inžinerija, 07T);

Prof. Dr. Alfonsas VAINORAS (Lietuvos sveikatos mokslų universitetas, biomedicinos mokslai, medicina, 06B);

Dr. Marcin WOZNIAK (Silezijos technologijos universitetas, Lenkija, technologijos mokslai, informatikos inžinerija, 07T).

Disertacija bus ginama viešame informatikos inžinerijos mokslo krypties disertacijos gynimo tarybos posėdyje 2018 m. gruodžio 17 d. 10 val. Kauno technologijos universiteto disertacijų gynimo salėje.

Adresas: K. Donelaičio g. 73-403, 44249 Kaunas, Lietuva. Tel. (370) 37 300 042; faks. (370) 37 324 144; el. p. doktorantura@ktu.lt.

Disertacijos santrauka išsiųsta 2018 m. lapkričio 16 d. Su disertacija galima susipažinti internete <http://ktu.edu> bei Kauno technologijos universiteto (K. Donelaičio g. 20, Kaunas) ir Vilniaus Gedimino technikos universiteto (Saulėtekio al. 14, Vilnius) bibliotekose.

Trumpiniai

AUC Area under the curve.

C_{llr} Cost of log-likelihood ratio - the comprehensive goodness-of-detection criterion.

DET Detection error trade-off.

DT Decision Tree.

EER Equal error rate.

GMM Gaussian mixture models.

HMM Hidden Markov model.

HNR Harmonics to noise ratio.

k-NN K-nearest neighbors algorithm.

LVQ Learning vector quantization.

MFCC Mel-frequency cepstrum coefficient.

MLP Multilayer perceptron.

NNE Normalized noise energy.

OOB Out-of-bag.

PDFs Probability Density Functions.

RF Random Forest.

ROC Receiver operating characteristic.

ROCCH ROC convex hull.

SNR Signal to noise ratio.

SVM Support vector machine.

***t*-SNE** T-distributed stochastic neighbor embedding.

WAV Waveform Audio File Format.

ĮVADAS

1.1. Tyrimų sritis

Šiame darbe taikant skaitinio intelekto metodus, yra nagrinėjama gerklų patologijos aptikimo tema. Tam naudojami neinvaziniais metodais surinkti duomenys, tokie kaip balso įrašai bei klausimynai. Taip pat apžvelgiami ir panaudojami metodai, skirti grafiniam duomenų atvaizdavimui, kurie suteikia galimybę giliau nagrinėti turimą informaciją.

Žmogaus balsas, toks kokį mes girdime, yra suformuojamas gerklų sužadavimo ir filtruojamas balso trakte. Šį procesą kontroliuoja mūsų smegenys, raumenų judesiai ir klausos jutiminė sistema [12]. Jeigu bent viena iš šių dalių netinkamai atlieka savo darbą, balso signalas pakinta. Dažniausiai balso pakitimai atsiranda dėl gerklų, kurios, priklausomai nuo ligos, gali prarasti elastingumą, pastorėti ar nebesugebėti iki galo užsidaryti. Tokie balso signalo pakitimai leidžia atskirti sveiką balsą nuo patologinio.

Eksperimentams atlikti naudojamas naujesnis papildytas duomenų rinkinys, kurio pirminė versija buvo panaudota ankstesniuose darbuose [35, 19]. Duomenys buvo praplėsti papildomais duomenų rinkiniais iš naujų pacientų. Siekiant pagerinti klasifikavimo tikslumą, buvo analizuojamas klasifikavimo rezultatų apjungimas sprendimų lygyje.

Šio darbo tikslas – išanalizuoti neinvaziniais metodais surinktu duomenų, tokių kaip balso įrašai ar klausimynai, panaudojimą patologiniam balsui atskirti. Skirtingų tipų duomenys pirmiausia klasifikuojami atskirai, panaudojant skirtingus metodus, vėliau rezultatai apjungiami panaudojant aukštesnio lygio algoritmą. Pagrindinės šio darbo sritys yra duomenų apdorojimas ir mašinų mokymo metodai. Taip pat, norint palengvinti duomenų analizę, panaudojami ir vizualinio atvaizdavimo metodai. Kaip papildomas šio darbo rezultatas yra sukurta kompiuterinė programa, skirta otolaringologams.

1.2. Problemos aktualumas

Kaip teigiama skirtinguose šaltiniuose, gerklų ligos paveikia apie 5 % [4] arba 6,2 % [23] pasaulio populiacijos. Dėl vėžio, siejamo su ger-

klomis, kiekvienais metais pasaulyje miršta apie 200000 žmonių ir šis skaičius nuolatos didėja, kai tuo tarpu kitų vėžinių susirgimų kiekiai nuolatos mažėja. Taigi, prevencinės gerklų ligų diagnozės technologijos yra labai reikalingos.

Gerklų patologijos diagnozavimas yra sudėtingas procesas, reikalaujantis skirtingų tipų duomenų analizės. Pacientų nusiskundimai dažniausiai yra surenkami klausimyno forma, senoji ligų istorija saugoma tekstinėje formoje, o instrumentiniai tyrimų metodai pateikia paveikslėlius ir/arba balso įrašus.

Daugelyje tyrimų pateikiamos išvados, kad neinvaziniais metodais surinkti duomenys, tokie kaip atsakymai į klausimynus ir balso įrašai, gali būti panaudoti automatinei analizei, kuri leidžia aptikti patologinius pakitimus ankstyvoje stadijoje ir suteikia galimybę plačiau panagrinėti balso kokybę. Tokia analizė gali būti panaudota tiek diagnostikai, tiek kaip prevencinė priemonė. Nagrinėjamoje srityje balso signalo garsinė analizė yra taikoma vis dažniau, tačiau bandymų panaudoti kito tipo neinvaziniais metodais surinktus duomenis, tokius kaip atsakymai į klausimus, yra vos keletas.

1.3. Tikslai ir uždaviniai

Pagrindinis šio tyrimo tikslas – pagerinti balso patologijos klasifikavimo tikslumą pasiūlant naują metodą, kuris galėtų klasifikuoti keliuose duomenų šaltiniuose esančią informaciją. Antras tyrimo tikslas – pagerinti balso patologijos klasifikavimo tikslumą panaudojant balso ir apklausos duomenų apjungimą sprendimų lygyje.

Šiam darbui suformuluotos užduotys:

1. išanalizuoti naujausius darbus balso patologijos aptikimo srityje įvertinant naudojamų metodų trūkumus;
2. apžvelgti balso patologijos klasifikavimui iš garso ir apklausos duomenų naudojamus metodus ir įvertinti jų trūkumus;
3. pasiūlyti naują metodą duomenims, esantiems skirtinguose rinkiniuose, apjungti ir klasifikuoti;
4. sukurti „permatomą“ metodą balsui ir apklausos duomenims analizuoti;

5. eksperimentiškai patikrinti pasiūlytus metodus naudojant apmokymo imtyje nenaudotus (angl. Out Of Bag) duomenis;
6. sukurti neinvazinę sprendimų priėmimo dėl balso patologijos pagalbos sistemą, suteikiančią galimybę grafiškai analizuoti pacientų duomenis ir naudojančią keletą duomenų šaltinių;

1.4. Tyrimo metodika

Dauguma skaitinio intelekto metodų yra pripažinti kaip tinkami sėkmingam balso patologijos aptikimui. Daug tyrimų padaryta naudojant GMM, HMM, k-NN, LVQ, MLP, SVM, atsitiktinių miškų, sprendimų medžių, diskriminantinės analizės ir kitus metodus. Kai kuriuose tyrimuose naudojama kelių algoritmų kombinacija ar papildomi duomenys (kartais abu – keletas algoritmų ir keletas duomenų rinkinių), tokie kaip klausimynų atsakymai tam, kad būtų galima pasiekti didesnę klasifikavimo tikslumą. Nepriklausomai nuo pasirinkto metodo, neinvazinė gerklų ligų diagnostika iš balso įrašo ir apklausos duomenų yra sudėtinga užduotis, kuri gali būti suskirstyta į keletą žingsnių:

1. garso parametrų gavimas iš balso įrašų;
2. klausimyno informacijos surinkimas iš paciento;
3. surinktų duomenų analizavimas siekiant nustatyti sveiką arba patologinę klasę;

Šiame darbe kaip aukštesniojo lygio klasifikatorius (angl. meta-learner) buvo pasirinktas atsitiktinių miškų metodas. Siekiant padidinti klasifikavimo tikslumą, balso ir apklausos duomenys buvo apjungti sprendimų lygyje. Apklausos duomenims klasifikuoti buvo sukurtas ir panaudotas asociacijų taisyklių metodas. Balso duomenų klasifikavimas buvo atliktas panaudojant sprendimų medžio algoritmą ir duomenų rinkinį su 6 garso parametrais (gerai pažįstamais otolarinologams) bei GFI parametru iš apklausos duomenų. Šis metodas leidžia vartotojui pateikti grafinę sprendimų medžio vizualizaciją, kuri suteikia galimybę lengviau suprasti sprendimo priėmimo taktiką. Platesnieji garso parametrų rinkiniai buvo klasifikuojami panaudojant

atsitiktinių miškų metodą. Buvo naujai sukurtas ir pasiūlytas metodas nuo duomenų priklausančio atsitiktinio miško formavimui, kuris leidžia apjungti keliuose rinkiniuose esančius duomenis. Duomenų atvaizdavimui buvo panaudotas t -SNE metodas, kuris leido duomenis atvaizduoti dvimatėje erdvėje.

1.5. Mokslinis naujumas

Pagrindiniai šio tyrimo naujumo aspektai:

- Pasiūlytas naujas, nuo duomenų priklausančio atsitiktinio miško metodas, skirtas duomenims iš kelių šaltinių apjungti ir klasifikuoti. Šis metodas klasifikuoja įvairius garso parametrų rinkinius naudodamas atsitiktinio miško algoritmą, o galutinį rezultatą pateikia naujai sukonstruotas atsitiktinis miškas iš pirmame lygyje naudotų miškų medžių, kurie teisingai klasifikuoja į nagrinėjamą pacientą panašius atvejus.
- Sukurtas naujas asociacijų taisyklėmis paremtas algoritmas apklausos duomenims klasifikuoti. Šis metodas yra paremtas asociacijų taisyklėmis, kurios buvo gautos iš apklausos duomenų panaudojant panašumų analizės metodą. Šis metodas yra visiškai „permatomas“, todėl yra labai naudingas kaip prevencinė priemonė.
- Grafinis duomenų atvaizdavimas paprastesnei pacientų analizei ir geresniam supratimui apie priimtą konkretų sprendimą. Šiam tikslui pasiekti buvo naudojamas t -SNE metodas, tikimybinio tankio funkcijų grafikai ir grafinė sprendimų medžio reprezentacija. Grafiškai 2D erdvėje atvaizduota artumų matrica leidžia lengviau palyginti pacientus, tikimybinio tankio funkcijų grafikai atskleidžia apklausos duomenų pasiskirstymą, o sprendimų medžio vizualizacija parodo algoritmo logiką.

1.6. Praktinė reikšmė

Sėkmingas gerklų ligų diagnozavimas kartais gali būti itin sudėtingas net ir labai patyrusiam gydytojui. Tokio tipo algoritmai ir įrankiai, kurie buvo sukurti ir panaudoti šiame darbe, gali labai pagerinti gydytojo darbo kokybę, suteikdami jam papildomos lengvai pasiekiamos

diagnostinės informacijos. Mūsų sukurta programa gali būti naudojama tiesiog kaip atskiras diagnostinis įrankis arba kaip įrankis skirtas gauti papildomai informacijai apie pacientą: klasę pagal asociacijų taisykles, grafinę sprendimų medžio vizualizaciją, grafinį vaizdą kuriame matoma paciento vieta tarp kitų (panaudojant t -SNE). Taip pat, mūsų sukurta įrankis gali būti modifikuotas ir pateiktas vartotojams kaip internetinis puslapis ar mobilioji aplikacija, kas leistų juo pasinaudoti daug didesniai vartotojų ratui ir keliais paprastais žingsniais patikrinti savo gerklų būklę.

1.7. Ginamieji teiginiai

1. Žmogaus balso signalas ir atsakymai į tam tikrus klausimus suteikia informacijos apie gerklų būklę, taigi ji gali būti panaudota neinvazinei gerklų ligų diagnostikai.
2. Skaitinio intelekto metodai yra pajėgūs atskirti sveiką balsą nuo patologinio su didesniu nei 90 % tikslumu.
3. Mūsų pasiūlytas nuo duomenų priklausančio atsitiktinio miško metodas leidžia klasifikuoti duomenis iš keleto duomenų rinkinių ir klasifikuoja geriau nei kiti darbe nagrinėti metodai apjungiantys duomenis sprendimų arba duomenų lygyje. Mūsų metodo didžiausias pasiektas tikslumas 86,37 %.
4. Balso patologija gali būti sėkmingai diagnozuota naudojant tik apklausos duomenis.
5. Balso ir apklausos duomenų klasifikavimo apjungimas sprendimų lygyje naudojant mūsų metodus leidžia pasiekti 90,48 % tikslumą, kas pralenkia klasifikavimą naudojant šiuos duomenis atskirai.

1.8. Disertacijos struktūra

Disertacija pradedama literatūros apžvalga, kuri pateikiama 2 skyriuje. Ten apžvelgiami naujausi tyrimai ir naudojami metodai nagrinėjama tematika. 3 skyriuje aprašomi šiame tyrime naudoti duomenys. Balso parametrų gavimas iš įrašų ir apklausos duomenų surin-

kimo metodika aprašomi 4 skyriuje. Duomenų matmenų mažinimui ir atvaizdavimui 2D erdvėje panaudotas *t*-SNE metodas.

Šiame tyrime naudoti balso įrašų parametrai ir duomenų analizės metodai išsamiai aprašomi 4 skyriuje:

- 4.1 poskyryje aprašomi balso parametrai, gauti iš balso įrašų ir naudojami klasifikavimui.
- 4.2, 4.3, 4.4 ir 4.5 poskyriuose aprašomi šiame tyrime naudoti klasifikavimo metodai. Tokie metodai kaip nuo duomenų priklausantis atsitiktinis miškas, asociacijų taisyklės, atsitiktinis miškas ar sprendimų medis aprašomi išsamiai pateikiant pagrindines savybes ir naudojimo žingsnius.
- Trūkstumų duomenų apdorojimo metodai apžvelgiami 4.7 poskyryje.
- 4.9 poskyryje aprašomi duomenų atvaizdavimo metodai, pateikiami jų pavyzdiniai paveikslėliai. Sukurta kompiuterinė programa nuosekliai aprašoma pateikiant pavyzdžius, *t*-SNE paaiškinimus, tikimybinių pasiskirstymų funkcijas ir sprendimų medį.

Šio tyrimo eksperimentų rezultatai pateikiami 5 skyriuje. Patologijų aptikimas įvertinamas poskyryje 5.1, o kituose poskyriuose apžvelgiami mikrofonų palyginimo, asociacijų taisyklių, duomenų ir sprendimų atvaizdavimo bei programos panaudojamumo rezultatai. Darbo rezultatų aptarimas ir išvados yra pateikiami 6 ir 7 skyriuose.

2. LITERATŪROS APŽVALGA

Akustinių ir apklausos duomenų analizė – tai puikus įrankis balso kokybei įvertinti ir patologijoms aptikti. Kaip teigiama mokslinėje literatūroje [32], apklausos duomenų panaudojimas dažnai tikslumu aplenkia garso duomenų naudojimą, tačiau šių dviejų tipų duomenų apjungimas leidžia pasiekti dar didesnę tikslumą. Šia tematika apžvelgti tyrimai gali būti suskirstyti į 3 grupes:

1. Tyrimai, kuriuose nagrinėjami balso įrašai ir siekiama išskirti pačius svarbiausius parametrus.
2. Tyrimai, kuriuose nagrinėjami apklausos duomenys ir bandoma atrinkti pačius svarbiausius klausimus.
3. Tyrimai, kuriuose mėginama sukonstruoti kuo tikslesnę klasifikavimo algoritmą panaudojant garso duomenis ir / arba akustinius duomenis.

Šio darbo tikslas – išanalizuoti jau egzistuojančius metodus neinvaziniam gerklų patologijos diagnozavimui ir pasiūlyti naują metodą, kuris leistų pasiekti didesnę klasifikavimo tikslumą. Kadangi tai yra testinis darbas, tikslas buvo pagerinti metodus, kurie buvo aptarti [33] bei pasiūlyti naujus savo sprendimus. Ši disertacija atitinka visas 3 kryptis, nes buvo naudojami akustiniai duomenys, apklausos duomenys bei buvo bandoma sukonstruoti tikslesnę klasifikavimo algoritmą. Kaip papildomas rezultatas, buvo sukurta kompiuterinė programa, skirta otolaringologams.

2.1. Balso funkcijos ir kokybės įvertinimas

Siekiant sėkmingai įvertinti paciento būklę, reikia atlikti vizualinę analizę, aerodinامينius matavimus, akustinę analizę ir įvertinti paciento nuomonę apie jo paties būklę [16]. Medicinoje analizei naudojami tiek invaziniai, tiek neinvaziniai analizės metodai, tačiau kaip teigiama mokslinėje literatūroje [36], svarbiausi yra netiesioginė laringoskopija ir video laringostroboskopija. Nepaisant invazinių tyrimo metodų populiarumo, paciento būklė gali būti sėkmingai įvertinta ir

naudojant neinvazinius metodus, kurie tyrimuose naudojami vis dažniau [29, 1, 37, 35, 22, 18, 38].

Gerklų patologijos diagnozavimas iš balso įrašų gali būti atliktas išgavus garso parametrus ir tik tada gali būti taikomi klasifikavimo metodai. Diagnozavimo tikslumas tiesiogiai priklauso nuo naudojamų garso parametrų, nes kai kurie pakitimai gali matytis tik specifiniuose parametruose. Praktikoje dažniausiai naudojami nepertraukiami balsių įrašai, kadangi juose nelieta lingvistinių artefaktų [26, 31], jie yra efektyvesni laiko prasme ir sumažina dispersiją balsėse [40, 39]. Tyrimuose pasiekti klasifikavimo rezultatai labai priklauso nuo naudotų garso parametrų bei duomenų kiekio ir kinta nuo 78,72 % [21] iki 92,86 % [29], ar net 100 % [7]. Aukšti rezultatai gali sudaryti įspūdį, kad klasifikavimo tikslumo problema jau išspręsta, tačiau dažniausiai jie pasiekiami su itin maža (nereprezentatyvia) duomenų imtimi.

Apklauskos duomenys, surinkti per paciento apžiūrą gali tarnauti kaip dar vienas neinvazinių duomenų šaltinis. Kaip teigia [32], kai kuriuose atsakymuose gali būti naudingos informacijos, kuri nesimato balso įrašų parametruose. Rinkoje egzistuoja ne vienas įrankis automatinei balso analizei, tačiau kiek mums yra žinoma, nė vienas iš jų negali analizuoti balso parametrų ir apklauskos duomenų kartu.

2.2. Garso parametrų išgavimas

Norint panaudoti balso įrašus klasifikavimui, reikia iš jų išgauti parametrus. Norint išgauti standartinius parametrus, galima naudoti egzistuojančius įrankius, kurių populiariausi yra Dr. Speech, LibXtract, YAAFE, jAudio, Librosa, Marsyas, Aubio, Essentia, Meyda ar MIRtoolbox. Kaip minima [20], visi įrankiai gali būti suskirstyti į 3 variantus:

- savarankiška kompiuterinė programa;
- papildinys pagrindinei programai;
- programinė biblioteka;

Kiekvienais metais yra pasiūloma vis naujesnių tokio tipo įrankių, iš kurių dalis yra visiškai nauji, o kiti – patobulintos senųjų versijos.

Bėgant metams pasiūlomi ir nauji parametrų rinkiniai, kurie išryškina kitus signalo pakitimus patologiniame balse ir padeda pasiekti didesnį klasifikavimo tikslumą.

2.3. Apklausos duomenų parametrai

Apklausose gali būti pateikti įvairaus tipo klausimai, kurie priklausomai nuo tikslo, gali turėti skirtingo tipo atsakymus: skaičius, tekstą ar varneles. Skaičiai gali būti rašomi laisvai (tarkim amžius) arba pasirenkami iš duoto intervalo (pavyzdžiui savo būklės įvertinimas 10 balų sistemoje). Tekstiniai atsakymai gali suteikti daugiau informacijos, tačiau juos reikia papildomai apdoroti, todėl nėra labai dažnai naudojami [24, 13].

Šiame tyrime atsakymai į daugelį klausimų yra skaitiniai. Kai kuriuose klausimuose atsakymai gali būti tekstiniai, tačiau juose yra iš anksto apibrėžti galimi žodžiai (Vyras, Moteris, Taip, Ne), todėl tokie atsakymai gali būti nesudėtingai transformuoti į skaitines vertes (0 ir 1). Apibendrinus apklausos atsakymus, gaunamas vienas skaitinių vėrčių vektorius.

2.4. Vizualinė balso ir apklausos duomenų analizė

Duomenų vizualizavimas yra svarbi duomenų analizės dalis, kuri padeda geriau juos suprasti [9]. Kai naudojami daugiamaciai duomenys, tenka pritaikyti įvairius matmenų mažinimo metodus, kad būtų įmanoma tokius duomenis atvaizduoti paprasčiau, tačiau išlaikant jų informatyvumą [17]. Tam tikrais atvejais duomenų statistinio pasiskirstymo atvaizdavimas irgi gali būti naudingas.

Šiame tyrime naudojami daugiamaciai duomenys, kurie prieš juos atvaizduojant turi būti papildomai apdoroti. Matmenų mažinimui ir duomenų atvaizdavimui panaudotas t -SNE algoritmas. Kaip teigiama literatūroje [34], tai yra našus algoritmas, galintis sėkmingai dirbti su dideliais duomenų kiekiais. 2D erdvėje atvaizduoti duomenys suteikia daug papildomos informacijos, kuri leidžia aptikti blogai suklasifikuotus, netinkamai sužymėtus ar kitai klasei priskirtus, bet šalia priešingos klasės esančius pacientus [19].

Statistiniam duomenų pasiskirstymui atvaizduoti buvo panaudotos tikimybinio tankio funkcijos. Gauti grafikai suteikia informaci-

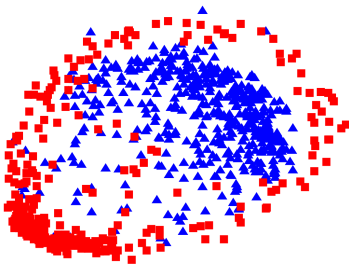
jos apie paciento vietą, lyginant su kitais pacientais, ir atskleidžia, kokios informacijos duomenų rinkinyje yra per mažai norint turėt įvairiapusišką imtį.

3. BALSŲ IR APKLAUSOS DUOMENYS

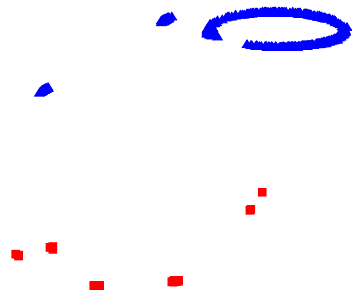
Šiame darbe buvo naudoti neinvaziniai duomenys – balsų įrašai ir apklausos rezultatai. Duomenys buvo gauti iš Lietuvos sveikatos mokslų universiteto (LSMU) Ausies, nosies ir gerklų ligų (ANG) klinikos. Iš viso buvo 3 duomenų rinkiniai – du garso duomenų ir vienas apklausos.

3.1. Balsų įrašų duomenys

Balsų informacijai gauti buvo naudoti iššęstos balsės /a/ įrašai. Toks pasirinkimas buvo paremtas tuo, kad iššęstos balsės įrašai yra paprastesni, lengviau analizuojami, jų nepaveikia kalbos greitis ar patiriamas stresas [40, 39, 18]. Visi balsų įrašai buvo įrašyti uždaroje, garso nepraleidžiančioje patalpoje. Mikrofonas buvo pastatytas 10 centimetrų atstumu nuo paciento ir pakreiptas 90 laipsnių kampu. Kiekvienam pacientui buvo daroma po 3 įrašus, kurie išsaugoti WAV formatu.



3.1 pav. Balsų įrašų duomenų bazės vizualizacija (Matmenų mažinimas ir atvaizdavimas atliktas naudojant *t*-SNE)



3.2 pav. Apklausos duomenų vizualizacija (Matmenų mažinimas ir atvaizdavimas atliktas naudojant *t*-SNE)

Sukauptoje duomenų bazėje buvo 273 įvairios lyties ir įvairaus amžiaus pacientų duomenys (163 sveiki, 110 patologiniai). Pacientai buvo patikrinti specialistų ir buvo nustatyta jų diagnozė. Patologinių

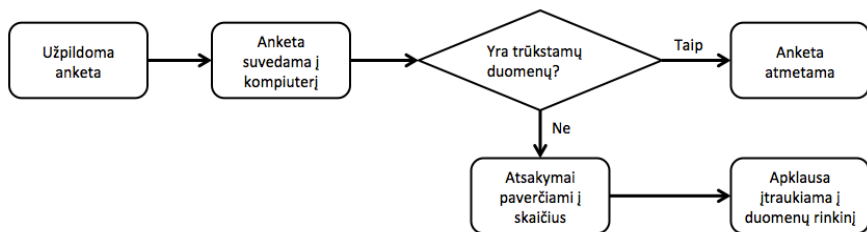
pacientų grupėje buvo mazgelinių susirgimų, polipų, cistų, papilomų, hiperplazinio laringito, keratozės, karcinomos. Pacientų patikrinimui ir tiksliam diagnozės nustatymui buvo naudojama laringostroboskopija ir teisioginė mikrolaringoskopija. Pacientų pasiskirstymas pagal jų klases pateiktas 3.1 paveikslėlyje.

3.2. Apklausos duomenys

Apklausos duomenys buvo surinkti otolaringologijos specialistų per pacientų patikrą. Klausimynui sudaryti buvo panaudoti reikšmingiausi klausimai, atrinkti ankstesniuose tyrimuose [2] ir yra pateikti 3.1 lentelėje. Klausimyno duomenų bazę sudarė 596 įvairios lyties ir įvairaus amžiaus pacientai. Jų pasiskirstymas pagal diagnozės klases matomas 3.2 paveikslėlyje. Apklausos duomenų surinkimo proceso žingsniai pavaizduoti 3.3 paveikslėlyje.

Naudojant apklausos duomenys panašumų analizės metu, atsakymai į klausimus su plačiomis skalėmis buvo padalinti į kvartilius. Nulinės vertės nebuvo koreguojamos. Pradinė duomenų analizė parodė, kad daugumai pacientų pažymėjus reikšmę 0 klausimuose 22 – 23, kiti šios grupės klausimų atsakymai sutapdavo, todėl klausimai 22 – 25 buvo pakeisti viena verte G_0 , kuri parodo, kad bent vienas iš šių klausimų buvo atsakytas kaip *problemos nėra*:

$$G_0 = (G_1 = 0 \vee G_2 = 0 \vee G_3 = 0 \vee G_4 = 0) \quad (3.1)$$



3.3 pav. Apklausos duomenų surinkimo procesas

3.1 lentelė. Klausimynas. SSV – Subjektyvus savęs vertinimas

#	Klausimas	Galima skalė arba variantai
1.	Lytis	{Vyras, Moteris}
2.	Amžius	Sveikas skaičius
3.	Vidutinė intensyvaus kalbėjimo trukmė	Valandos per dieną
4.	Vidutinė intensyvaus kalbėjimo trukmė	Dienos per savaitę
5.	Rūkymas	{Taip, Ne}
6.	Rūkymo intensyvumas	Cigaretės per dieną
7.	Rūkymo istorija	Metai
8.	Maksimalus fonavimo laikas	Sekundės
9.	Balso funkcijos kokybės SSV	Skalė nuo 0 iki 100
10.	Balso silpnumo SSV	Nuo 0 (<i>Ne</i>) iki 100 (<i>Labai nusilpęs</i>)
11.	Balso trūkinėjimo progresavimas	Įvertinimas nuo 1 iki 4
12.	Per dieną patiriamo streso SSV	Nuo 0 (<i>Nėra</i>) iki 100 (<i>Labai didelis stresas</i>)
13.	Dainavimo dažnumas	Įvertinimas nuo 1 iki 5
14.	Dainavimo / kalbėjimo dažnumas pri rūkytoje patalpoje	Įvertinimas nuo 1 iki 5
15.	Dėl ligos patiriamo diskomforto SSV	Nuo 0 (<i>Jokio</i>) iki 100 (<i>Didelis diskomfortas</i>)
16.	„Per daug silpno balso“ SSV	Nuo 0 (<i>Nėra</i>) iki 100 (<i>Labai aiškus</i>)
17.	Pasikartojančio balso „dingimo“ SSV	Nuo 0 (<i>Nėra</i>) iki 100 (<i>Labai aiškus</i>)
18.	Nusilpusio balso SSV	Nuo 0 (<i>Nėra</i>) iki 100 (<i>Labai aiškus</i>)
19.	Sumažėjusios galimybės dainuoti SSV	Nuo 0 (<i>Nėra</i>) iki 100 (<i>Labai aiškus</i>)
20.	Balso „lūžinėjimo“ dažnis	Nuo 0 (<i>Nėra</i>) iki 100 (<i>Labai dažnai</i>)
21.	Balso naudojimo lygis	Įvertinimas nuo 1 iki 4
22.	Kalbėjimas reikalauja papildomų pastangų (G_1)	Nuo 0 (<i>Ne</i>) iki 5 (<i>Didelė problema</i>)
23.	Gerklės skausmas arba diskomfortas po kalbėjimo (G_2)	Nuo 0 (<i>Ne</i>) iki 5 (<i>Didelė problema</i>)
24.	Kalbant balsas nusilpsta (G_3)	Nuo 0 (<i>Ne</i>) iki 5 (<i>Didelė problema</i>)
25.	Balsas trūkčioja ar skamba kitaip (G_4)	Nuo 0 (<i>Ne</i>) iki 5 (<i>Didelė problema</i>)
26.	Gerklinės funkcijos indeksas [3, 27] ($GFI=G_1+G_2+G_3+G_4$)	Įvertinimas nuo 0 iki 20

4. METODOLOGIJA

Paciento būklei nustatyti – sveikas ar patologinis – buvo panaudota keletas klasifikavimo metodų. Apmokymui skirtus duomenų rinkinius sudarė vektoriai, kuriuose buvo duomenų parametrai ir paciento klasė. Eksperimentams buvo panaudoti 3 skirtingi audio parametru duomenų rinkiniai.

Kuriams ir naudojamiems metodams buvo iškelta sąlyga, kad jie būtų kiek įmanoma labiau „permatomi“ – leistų suprasti kaip ir kodėl buvo priimtas sprendimas dėl paciento klasės. Šiam reikalavimui patenkinti buvo sukurtas asociacijų taisyklių metodas ir panaudotas sprendimų medis.

Kaip pagrindinis klasifikatorius balso įrašo parametrui buvo panaudotas atsitiktinių miškų metodas. Siekiant pagerinti tikslumą, šiame darbe buvo pasiūlytas naujas, nuo duomenų priklausančio atsitiktinio miško algoritmas, kuris taip pat leidžia klasifikuoti duomenis iš kelių duomenų šaltinių. Naudojant šį algoritmą buvo palyginti akustinis ir kontaktinis mikrofonai. Rezultatai parodė, kad kontaktinis mikrofonas nesuteikia papildomos naudos, kad klasifikavimo tikslumo pokytis būtų statistiškai reikšmingas.

Duomenų ir klasifikavimo rezultatų atvaizdavimui buvo panaudotas *t*-SNE metodas. Jo pagalba artumų matrica, gauta iš atsitiktinio miško, buvo atvaizduota 2D erdvėje.

4.1. Balso parametrai

Balso įrašų parametrai buvo išgauti panaudojant akustinės ir signalų analizės metodus. Duomenų bazę sudaro 14 parametru rinkinių, kurie pateikti 4.1 lentelėje. Ši parametru kolekcija buvo panaudota, nes [11, 35] buvo parodyta, kad su ja galima pasiekti aukštą klasifikavimo tikslumą. Garso signalai buvo sudalinti į laiko intervalus (langus) ir parametrai buvo išgaunami kiekvienam „langui“ atskirai.

Antrajame garso parametru rinkinyje sukaupti duomenys, kuriuos pateikė LSMU ANG specialistai. Šiame rinkinyje buvo tik 6 parametrai, kurie yra pateikti 4.2 lentelėje. Minėtiems parametrui gauti buvo panaudotas „Dr. Speech“ (<http://www.doctorspeech.com>) įrankis.

4.1 lentelė. Garso signalų parametrai naudojami šiame darbe (iš viso 927 parametrai)

#	Požymių tipas	Kiekis
1.	Tono ir amplitudės perturbacijos	24
2.	Dažnių spektras (0 – 5000 Hz)	100
3.	Mel skalės dažnių juostos	35
4.	Kepstrai (kepstrinė energija)	100
5.	Mel dažnių skalės koeficientai	35
6.	Autokoreliacijos koeficientai	80
7.	Harmonikų-triukšmo santykis (spektrinis)	11
8.	Harmonikų-triukšmo santykis (kepstrinis)	11
9.	Tiesioginio prognozavimo (TP) koeficientai	77
10.	Tiesioginio prognozavimo koeficientų kosinuso transformacija	77
11.	Signalų dažnių gaubtinės forma	128
12.	Levinsono-Durbino atspindžio koeficientai	24
13.	Balstro trakto nereguliarumai	71
14.	Percepcinio tiesioginio prognozavimo kepstriniai koeficientai (PTPK)	154

Kadangi otolaringologai naudoja šiuos parametrus kasdieniame darbe, jie jiems yra lengvai suprantami ir aiškūs.

4.2 lentelė. Garso parametrai, gauti iš LSMU ANG

#	Parametras
1.	Fundamentalus dažnis (F0)
2.	F_0 neperiodiškumas (Jitter)
3.	Amplitudės neperiodiškumas (Shimmer)
4.	Normalizuota ūžesio energija (NNE)
5.	Harmonikų-triukšmo santykis (HNR)
6.	Signalų-triukšmo santykis (SNR)

4.2. Sprendimų medis

Šiame darbe panaudotas CART tipo sprendimų medis. Jis atvaizduojamas kaip medžio tipo hierarchinė struktūra. Yra keletas metodų

sprendimo medžiui sudaryti, tačiau šiame tyrime buvo naudojami tokie žingsniai:

1. Pirmajame žingsnyje panaudojami visi mokymui skirti duomenys išbandant visus įmanomus padalinimo variantus.
2. Antrame žingsnyje parenkamas optimaliausias padalinimas.
3. Pritaikomas pasirinktas padalinimas.
4. Ankstesni žingsniai rekursiškai pritaikomi abiemis vaikiniais elementams.

Rekursinis padalinimas vyksta tol, kol nagrinėjame mazge lieka vienos klasės elementai. Pagrindiniai sprendimų medžio privalumai tokie, kad jį lengva sudaryti ir jis yra lengvai interpretuojamas. Kaip minima mokslinėje literatūroje [14], šis klasifikatorius gana plačiai naudojamas įvairių ligų diagnozavimui.

4.3. Asociacijų taisyklės

Šio tyrimo metu buvo sukurtas ir panaudotas asociacijų taisyklėmis paremtas apklausos duomenų klasifikavimo metodas. Taisyklėmis išskirti buvo panaudota panašumų analizė, kuria buvo atrinkta po 11 jeigu < sąlyga > tada < rezultatas > tipo taisyklių kiekvienai klasei (patologinei ir sveikai). Šis metodas yra visiškai „permatomas“ ir leidžia matyti kaip priimamas sprendimas, todėl yra labai naudingas mokymo procese. Taip pat jis leido atrinkti 17 reikšmingiausių klausimų taip sumažinant mūsų pradinį klausimyną. Kiekvienos taisyklės svarbumui įvertinti buvo paskaičiuotos tokios reikšmės:

$$\text{support}(A \rightarrow C) = P(A \wedge C), \quad (4.1)$$

$$\text{confidence}(A \rightarrow C) = \frac{P(A \wedge C)}{P(A)}, \quad (4.2)$$

$$\text{lift}(A \rightarrow C) = \frac{\text{confidence}(A \rightarrow C)}{P(C)}, \quad (4.3)$$

kur: $P(\text{condition})$ yra sąlygos tikimybė (duomenų dalis tenkinanti šią sąlygą), \wedge yra loginis IR, *support* yra taisyklės populiarumas (duomenų dalis turinti A ir C), *confidence* yra taisyklės tikrumas (kaip dažnai A eina kartu su C), *lift* yra netikėtumo įvertinimas.

4.3 lentelė. Taisyklių aprėptis anksčiau paminėtiems klausimams. Klausimų sutrumpinimai su klausimais siejami taip: 1 (Gender), 4 (U), 6 (C), 7 (Y), 8 (MFT), 9 (VAS), 11 (H), 15 (D), 16 (W), 18 (R), 19 (S), 20 (X), 21 (L), 22–25 (G_0). „Sveiki“ ir „Patologiniai“ stulpeliuose pateikiamas kiekis, keliose taisyklėse klausimas yra naudojamas

Taisyklė	Sveiki	Patologiniai
$G_0 = 0$	11	0
$H = 2$	11	0
$C = 0$	6	1
$Y = 0$	6	0
$U = 7$	3	2
$MFT = [2, 12]$	0	4
$L = 3$	4	0
$VAS = (63, 100]$	0	2
$G_0 = 0$	0	1
$X = (73, 100]$	0	1
$S = (75,8, 100]$	0	1
$R = (60, 100]$	0	1
$W = (60, 100]$	0	1
$D = (65, 100]$	0	1
Gender = F	1	0

Siekiant panaudoti gautas taisykles klasifikavimui, buvo panaudotas paprastas metodas, išvestas iš spėjimo pagal svorinę daugumą [25]. Klasifikavimo rezultato tikrumui gauti buvo naudojama formulė:

$$Certainty = 100 \cdot \frac{\sum_{i=1}^J P_i - \sum_{i=1}^K H_i}{\sum_{i=1}^J P_i + \sum_{i=1}^K H_i}, \quad (4.4)$$

kur: J yra tenkinamų „patologinių“ taisyklių kiekis, K yra tenkinamų „sveikų“ taisyklių kiekis, P yra tenkinamos „patologinės“ taisyklės tikrumas ir H yra tenkinamos „sveikos“ taisyklės tikrumas. Gautas teigiamas rezultatas nurodo patologinę klasę, neigiamas – sveiką. Su-

generuotos taisyklės buvo atrinktos paliekant tik reikšmingas, kur $support > 0,28$ ir $confidence > 0,9$ sveikai klasei ir $support > 0,16$ ir $confidence > 0,9$ patologinei klasei. Sugeneruotų taisyklių sąrašai pateikiami lentelėse 5.1 ir 5.2. Atsakymai į klausimus # 22 – 25 buvo apjungti į vieną kintamąjį G_0 (pagal formulę 4.5).

$$G_0 = (G_1 = 0 \vee G_2 = 0 \vee G_3 = 0 \vee G_4 = 0). \quad (4.5)$$

4.4. Atsitiktinis miškas

Atsitiktinis miškas (RF) – tai klasifikatorius, sudarytas iš grupės sprendimų medžių [5], kur rezultatas gaunamas daugumos balsavimu. Kaip minima teorinėje literatūroje [15, 8], RF yra greitai apmokomas, išskaidomas, atsparus triukšmams ir produktyvus su daugiadimensiais duomenimis. Atsitiktinio miško tikslumui įvertinti yra naudojama OOB imtis, kur balsuoja tik tie medžiai, kurių sudarymui nebuvo naudoti paciento duomenys. Mažas šališkumas ir maža koreliacija yra labai svarbūs klasifikavimo tikslumui, kurie pasiekiami auginant medžius iki maksimalaus dydžio naudojant atsitiktines imtis:

- Kiekvienas atskiras RF medis yra sukuriamas panaudojant atsitiktinę imtį iš mokymui skirtų duomenų.
- Sudarant medžius kiekviename mazge yra atsitiktinai parenkama n kintamųjų iš N galimų ir tik vienas kintamasis, lemiantis geriausią padalinimą, yra naudojamas iš n atrinktų.

Konstruojant RF n yra vienintelis eksperimentiškai parenkamas parametras. Medžio konstravimui nepanaudoti duomenys gali būti panaudoti jo testavimui. OOB taip pat galima panaudoti parametų svarbumo įvertinimui. Iš atsitiktinio miško galima išgauti artumų matricą Φ . Tai atliekama einant per kiekvieną medį ir padidinant ϕ_{ij} vienetu, jeigu x_i ir x_j užima tą patį mazgą. Tokia matrica šiame tyrime buvo panaudota duomenims ir sprendimams atvaizduoti 2D erdvėje panaudojant t -SNE metodą. Atsitiktinio miško artumų matrica Φ buvo gauta taip:

$$\Phi = \frac{\sum_{i=1}^{N_F} w_i \Phi_i}{\sum_{i=1}^{N_F} w_i}, \quad (4.6)$$

kur: N_F yra atskirų miškų skaičius ($N_F=14$) ir w_i yra svoris, proporcingas vidutiniam i -tojo RF tikslumui [35].

4.5. Nuo duomenų priklausantis atsitiktinis miškas

Šiame tyrime buvo pasiūlytas ir panaudotas naujas, atsitiktiniais miškais paremtas klasifikavimo algoritmas, leidžiantis panaudoti duomenis iš keleto šaltinių. Jis paremtas tuo, kad kiekvienas duomenų rinkinys yra klasifikuojamas naudojant atsitiktinių miškų metodą, o galutinis rezultatas gaunamas atlikus rezultatų apjungimą sprendimų lygyje. Formuojant atsitiktinius miškus medžių skaičius B buvo nustatytas 1000, o n buvo paskaičiuotas taip, kaip rekomenduojama mokslinėje literatūroje [5] – $n = \sqrt{N}$. Naudojant pasiūlytą metodą, nagrinėjamo paciento x duomenys yra klasifikuojami sukonstruotų atsitiktinių miškų, iš kurių vėliau atrenkami kaimynai $\mathcal{N}_i(x)$ iš m galimų, kurie yra panašiausi į x . Norint gauti galutinį sprendimą dėl x klasės, yra formuojamas naujas RF, kuris sudaromas iš medžių, atrinktų iš $\mathcal{N}_i(x)$ kaimynų, kurie teisingai klasifikuoja testavimui skirtus duomenis (OOB). Panašumui tarp x ir x_j nustatyti yra matuojamas atstumas tarp mazgų, kuriuos jie užima sprendimų medyje:

$$p_j = \frac{1}{K} \sum_k^K 1/(e^{w \cdot g_{jk}}), \quad (4.7)$$

kur: k ėjimai per K medžių, kur x_j yra tarp OOB elementų, w yra parametras ir g yra medžio šakų skaičius tarp x ir x_j . Kai x ir x_j yra tame pačiame mazge, $g = 0$ ir p_j bus padidinta vienetu. Parametro g daroma įtaka yra kontroliuojama parametru w .

Atsitiktiniame miške priimant sprendimą yra įvertinami du variantai: balsavimas ir svorinis balsavimas, kur k^{th} medis iš i^{th} atsitiktinio miško gauna svorį, paskaičiuotą pagal šią išraišką:

$$w_{ik} = \frac{1}{|\mathcal{N}_i^c(x)|} \sum_{x_j \in \mathcal{N}_i^c(x)} 1/(e^{w \cdot g_{jk}}), \quad (4.8)$$

kur: $|\mathcal{N}_i^c(x)|$ yra teisingai suklasifikuotų OOB duomenų kiekis iš $\mathcal{N}_i(x)$ kaimynų. Kitų kintamųjų reikšmės yra tokios pačios kaip 4.7 formulėje.

4.6. Klasifikatoriaus tikslumo įvertinimas

Šiame darbe klasifikatorių tikslumui įvertinti buvo naudojami tokie įverčiai kaip aptikimo klaidos nuokrypio (DET) kreivė, lygios klaidos įvertis EER, imtuvo veikimo charakteristikos (ROC) kreivė, loginio panašumo koeficiento kaina (C_{lrr}) ir plotas po kreive (AUC). Šios vertės buvo paskaičiuotos iš interpoliuotos ROC, kuri buvo apdorota panaudojant BOSARIS įrankį [6].

Greitas būdas palyginti klasifikatorių tikslumą turint jų DET (ROC) kreives yra pusiausvyros taškas, dar žinomas kaip lygios klaidos koeficientas (EER). EER yra taškas, kuriame DET (ROC) kreivė kerta diagonalę ir:

- a) klaidinga teigiama norma (nepataikymo dažnis) = klaidinga neigiama norma (klaidingo aliarmo dažnis), DET kreivei.
- b) teisinga teigiama norma (jautrumas) = teisinga neigiama norma (specifiškumas), ROC kreivei.

DET ir EER gavimui buvo panaudota interpoliuota ROC versija, panaudojant šalia esančių pažeidėjų algoritmą, kuris vadinamas ROC korpuso išgaubtumo (ROCCH) metodu (pateikiamas [6]). Kaip teigia [28], DET kreivė yra labiau linijinė dėl logaritminių ašių, taigi leidžia lengviau palyginti algoritmus nei ROC. Kai kurie autoriai teigia [10], kad DET kreivė gali būti išreikšta 4.9 funkcija, kur p_i^F yra klaidingo pranešimo tikimybė:

$$p_i^D = f_i(p_i^F). \quad (4.9)$$

4.7. Sprendimo tikrumo įvertinimas

Tyrimo metu sukurta programinė įranga ne tik pateikia paciento klasę, kaip klasifikavimo rezultata, bet ir procentą kiek algoritmas tikras dėl šios klasės. Tikrumas paskaičiuojamas naudojant informaciją iš RF sudarytų iš $\{t_1, \dots, t_L\}$ medžių. Sprendimo priėmimui suformuluotos dvi atmetimo taisyklės, kurios paremtos tikimybių vertėmis:

1. Panašumu paremtas atmetimas, kur x sudarytas jei:

$$p^{max1}(\{t_1, \dots, t_L\}, x, q) - p^{max2}(\{t_1, \dots, t_L\}, x, q) < \varepsilon_s, \quad (4.10)$$

kur: ε_s yra vartotojo nustatytas panašumo slenkstis ir p^{max1} and p^{max2} yra pirma ir antra didžiausios klasės tikimybės:

$$p^{max1}(\{t_1, \dots, t_L\}, x, q) = \max_{q=1, \dots, Q} p(\{t_1, \dots, t_L\}, x, q). \quad (4.11)$$

Šio tipo atmetimas reiškia, kad klasifikatorius negali priimti patikimo sprendimo dėl klasės.

2. Skirtumais paremtas atmetimas, kur x sudarytas jei:

$$p^{max1}(\{t_1, \dots, t_L\}, x, q) < \varepsilon_d, \quad (4.12)$$

kur: ε_d yra vartotojo nustatytas panašumo slenkstis. Šio tipo atmetimas parodo, kad nė viena klasė nėra pakankamai panaši į x .

4.8. Panašumo įvertinimas

Tyrimo metu dviejų pacientų (\mathbf{x}_i ir \mathbf{x}_j) panašumas buvo skaičiuojamas išmatuojant atstumą tarp jų užimamų mazgų. Kaip buvo pasiūlyta [19], tai buvo atliekama pagal 4.13 formulę, kuri pateikta toliau:

$$p_{ij}^t = 1/(e^{w \cdot g_{ij}}), \quad (4.13)$$

kur w yra parametras, g_{ij} medžio šakų kiekis tarp \mathbf{x}_i ir \mathbf{x}_j užimamų mazgų. Jeigu užimamas tas pats mazgas, $g = 0$ ir $p_{ij} = 1$. Parametras w naudojamas atstumo tarp dviejų pacientų įtakai reguliuoti.

Naudojant asociacijų taisyklių metodą, panašumas tarp \mathbf{x}_i ir \mathbf{x}_j skaičiuojamas pagal formulę:

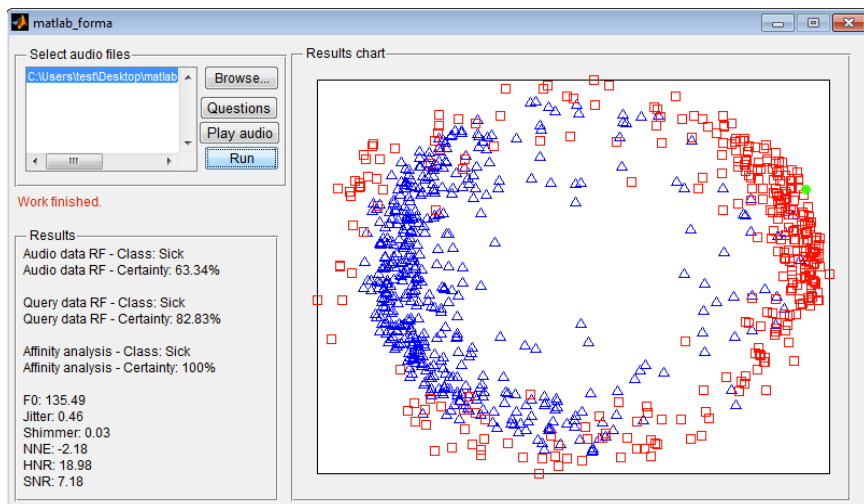
$$p_{ij}^r = \frac{2 \sum_{n \in \mathcal{U} \cap \mathcal{V}} p_{in}}{\sum_{n=1}^U p_{in} + \sum_{n=1}^V p_{jn}}, \quad (4.14)$$

kur \mathcal{U} ir \mathcal{V} yra tenkinamų taisyklių rinkiniai. Atitinkamai, U ir V yra taisyklių kiekiai tuose rinkiniuose, o p_{in} yra n -tosios tenkinamos taisyklės tikrumas.

4.9. Duomenų vizualizavimas

Kaip vienas iš šio tyrimo rezultatų – buvo sukurta kompiuterinė programa. Jos ekrano vaizdas pateiktas 4.1 paveikslėlyje, kur matosi, kad yra padalintas į 4 dalis. Atlikus paciento analizę, programa pateikia vartotojui paciento klasę, tikrumą ir 6 garso parametrus. Taip pat vartotojas gali peržiūrėti ir statistinius duomenų grafikus, kurie gauti panaudojus statistinio pasiskirstymo funkcijas (PDFs).

Programos kodas yra padalintas į dvi dalis. Pirmoji dalis yra atsakinga už garso parametrų gavimą iš įrašo failo, klasifikavimą ir artumų matricos atvaizdavimą panaudojant t -SNE. Antroji dalis atsakinga už naujai pateiktų duomenų išsaugojimą bendroje duomenų bazėje.



4.1 pav. Pagrindinio programos lango ekranvaizdis su trimis grafinės sąsajos dalimis: balso įrašo failo parinkimas, 2D grafikas ir tekstinių rezultatų pateikimo laukas

Pradiniame programos lange pateikiamas esamų duomenų atvaizdavimas dvimatėje erdvėje. Po klasifikavimo, panaudojant t -SNE, nagrinėjamas paciento pažymimas tame pačiame grafike. Tai yra viena iš unikalų t -SNE funkcijų, kuri leidžia naują elementą įterpti į anksčiau sugeneruotą grafiką. Toks pacientų atvaizdavimas leidžia juos pa-

4.4 lentelė. Vartotojui pateikiami paciento parametrai

Parametrai	Aprašymas
Standartinė paciento informacija	Paciento ID, Lytis, Amžius, Diagnozė
Balso parametrai	F_0 , F_0 neperiodiškumas, amplitudės neperiodiškumas, NNE, HNR, SNR
Apklauskos duomenys	Atsakymai į klausimus (žr. 3.1 lentelę)

lyginti tarpusavyje, įvertinant arčiau kurios klasės yra nagrinėjamas pacientas. 4.1 paveikslėlyje nagrinėjamas pacientas yra pažymėtas žalia spalva.

Pradinis paciento būklės nustatymas yra labai svarbus paciento diagnozėje. Tokie garso parametrai kaip F_0 , Jitter, Shimmer, NNE, HNR, SNR ir apklausos duomenys yra puikiai suprantami otolaringologams, todėl jiems yra pateikiami šių duomenų pasiskirstymo grafikai. Kaip rekomenduota literatūroje [30], skaičiuojant PDFs buvo panaudotas Epanechnikov branduolio išlyginimo metodas. Su kiekvienu, naujai įtrauktu pacientu, šie grafikai atsinaujina taip leisdami matyti turimų duomenų įvairovę ir aptikti tam tikrų duomenų stoką.

Sukurta programinė įranga taip pat leidžia palyginti ir keletą pacientų balso signalų parametrus bei apklausos duomenis. Tą padaryti galima paspaudus ant paciento žymės programos lange esančiame 2D grafike. Kadangi kiekvienas taškas žymi atskirą pacientą, kiekvienam iš jų bus atidarytas naujas langas su parametrais, kurie pateikti 4.4 lentelėje. Pateikiami parametrai buvo atrinkti su gydytojais, kuriuos jie įvertino kaip pačius svarbiausius. Toks peržiūros būdas leidžia palyginti greta esančius pacientus ir tiksliau įvertinti nagrinėjamo paciento diagnozę.

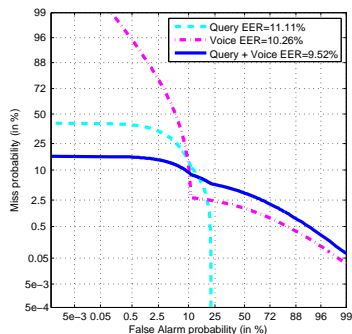
Kad būtų lengviau suprasti paciento klasės nustatymo procesą, vartotojui pateikiamas sprendimų medžio vaizdas. Jis perskaičiuojamas kiekvieną kartą kai įtraukiami naujo paciento duomenys. Sprendimų medis sudarytas naudojant 6 garso parametrus ir GFI parametras iš apklausos duomenų. Pateiktame paveikslėlyje ryškesne linija pavaizduotas pavyzdinis sprendimo būdas.

5. EKSPERIMENTŲ REZULTATAI

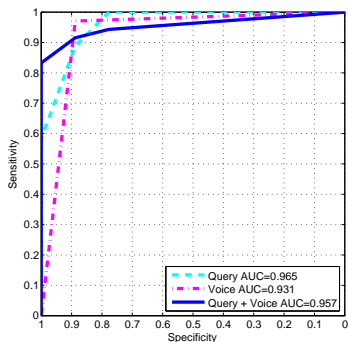
5.1. Klasifikavimo tikslumo įvertinimas

Klasifikavimo tikslumas buvo vertinamas naudojant 48 apmokymams nenaudotų pacientų duomenis. Tikslumui įvertinti, kai naudojami garso duomenys, kiekvienam pacientui buvo suteiktas įvertinimas pagal klasės tikimybę tam mazgui, kuriame jis yra. Naudojant apklausos duomenis klasifikavimo tikslumas buvo skaičiuojamas pagal 4.4 formulę. Klasifikavimo tikslumų rezultatai matomi 5.1 ir 5.2 paveikslėliuose, kur taip pat matosi EER ir AUC vertės.

Nepaisant panaudotų algoritmų paprastumo, pasiekti rezultatai yra aukšti ir daug žadantys. Naudojant balso ir apklausos duomenis atskirai, buvo pasiektos EER vertės: 11,11 % ir 10,26 %. Apjungus klasifikatorius sprendimų lygyje buvo pasiektas dar didesnis tikslumas, kur EER tik 9,52 %. Kaip matyti DET ir ROC kreivėse, apjungtas klasifikatorius turi didžiausią specifiškumą aukščiausio jautrumo taške. Gauti rezultatai parodo, kad apklausos duomenų panaudojimas dažnai leidžia pasiekti geresnius klasifikavimo rezultatus nei balso parametrai. Nepaisant to, duomenų apjungimas sprendimų lygyje duoda didžiausią tikslumą.



5.1 pav. DET kreivės ir EER įvertiniai nematytiems pacientams



5.2 pav. ROC kreivės ir AUC įvertiniai nematytiems pacientams

5.2. Duomenų apjungimas

Kuriant nuo duomenų priklausančio atsitiktinio miško algoritmą buvo analizuojama keletas duomenų apjungimo variantų ir kartu lyginami akustinis bei kontaktinis mikrofonai. Klasifikavimo tikslumui palyginti buvo naudojamos DET kreivės. Kaip matyti iš rezultatų, akustinis mikrofonas lenkia kontaktinį naudojant visas tyrinėtas duomenų apjungimo schemas ir lenkia kontaktinį 13-oje duomenų rinkinių iš 14, o tikslumo skirtumas yra statistiškai reikšmingas. Reikia paminėti, kad kiekvienam parametrų rinkiniui buvo kuriamas atskiras atsitiktinis miškas. Kai kuriais atvejais tikslumas išaugo po parametrų atrinkimo palikus mažiau kaip 15 % rinkinio parametrų. Lyginant skirtingus duomenų apjungimo metodus, didžiausias tikslumas (86,37 %) buvo pasiektas naudojant nuo duomenų priklausantį atsitiktinį mišką, kai miškai atrenkami, o medžiai turi svorius.

5.3. Asociacijų taisyklės

Asociacijų taisyklės buvo atrinktos ir paliktos tik pačios reikšmingiausios. Taisyklės skirtos sveikai klasei išėjo sudėtingesnės, o skirtos patologinei klasei yra paprastesnės, jų komponentai įvairesni. Taisyklių sąrašai sveikai ir patologinei klasėms pateikti 5.1 ir 5.2 lentelėse, kur žymekliai su klausimyno klausimais siejami taip (žymeklis – klausimo numeris): Lytis – 1, U – 4, C – 6, Y – 7, MFT – 8, VAS – 9, H – 11, D – 15, W – 16, R – 18, S – 19, X – 20, L – 21, G_0 – 22–25.

Šiame darbe atlikta panašumų analizė asociacijų taisyklių išskyrimui išryškino 17 pačių svarbiausių klausimų. Tai parodo, kad kai kurie klausimai suteikia labai mažai naudos, todėl gali būti pašalinti iš klausimyno. Kaip rodo rezultatai, svarbiausias apklausos duomenų parametras yra GFI (22–25 klausimų kombinacija), kadangi bet kokia jo vertė nelygi 0 kartu su trumpa fonacija aiškiai indikuoja patologiją.

5.3.1. Parametrų svarba

Šiame tyrime be klausimų svarbos pagal asociacijų taisykles buvo įvertinta ir klausimų bei garso parametrų svarba naudojant RF. Gauti rezultatai pateikiami 5.3 ir 5.4 paveikslėliuose. Kaip matyti, 10 klausimų iš 26 gali padidinti klasifikavimo tikslumą daugiau kaip 1 %, o tarp garso parametrų tokių yra net 16 iš 26. Apklausos duomenyse

5.1 lentelė. Asociacijų taisyklės sveikai klasei

#	Taisyklės sąlyga	Svarbumas	Tikrumas	Tikimybė
1	$H=2, G_0=1, \text{Lytis} = F$	0,302	0,928	1,691
2	$H=2, G_0=1, Y=0, L=3$	0,292	0,926	1,687
3	$H=2, G_0=1, C=0, Y=0, L=3$	0,289	0,925	1,685
4	$H=2, G_0=1, C=0, L=3$	0,297	0,922	1,680
5	$H=2, G_0=1, Y=0$	0,341	0,919	1,674
6	$H=2, G_0=1, C=0, Y=0$	0,337	0,918	1,673
7	$H=2, G_0=1, C=0$	0,351	0,917	1,671
8	$H=2, G_0=1, Y=0, U=7$	0,285	0,909	1,657
9	$H=2, G_0=1, C=0, Y=0, U=7$	0,284	0,909	1,656
10	$H=2, G_0=1, L=3$	0,381	0,908	1,655
11	$H=2, G_0=1, C=0, U=7$	0,294	0,907	1,653

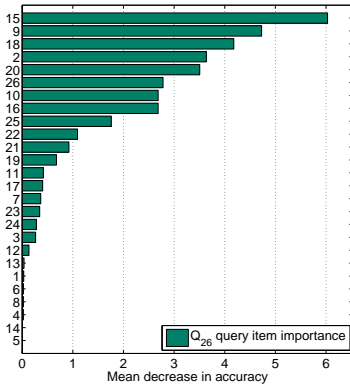
5.2 lentelė. Asociacijų taisyklės patologinei klasei

#	Taisyklės sąlyga	Svarbumas	Tikrumas	Tikimybė
1	$\text{MFT}=[2,12], G_0=0$	0,168	1	2,216
2	$R=(60,100]$	0,191	0,983	2,177
3	$\text{MFT}=[2,12], C=0$	0,164	0,980	2,171
4	$\text{MFT}=[2,12]$	0,267	0,975	2,161
5	$D=(65,100]$	0,190	0,974	2,158
6	$\text{MFT}=[2,12], U=7$	0,211	0,969	2,147
7	$X=(73,100]$	0,191	0,966	2,141
8	$\text{VAS}=(63,100]$	0,210	0,947	2,098
9	$\text{VAS}=(63,100], U=7$	0,169	0,944	2,091
10	$S=(75,8, 100]$	0,185	0,940	2,083
11	$W=(60,100]$	0,188	0,933	2,068

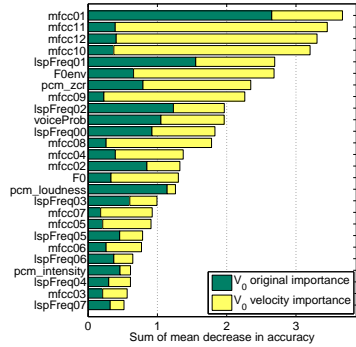
atsitiktiniam miškui reikšmingiausi yra subjektyvūs savęs įvertinimo klausimai (15, 9, 18). Balso parametrų rinkinyje didžiausią svarbą turi MFCC parametrai.

5.4. Duomenų ir sprendimų analizė

Duomenims atvaizduoti 2D erdvėje buvo panaudota artumų (panašumo) matrica, paskaičiuota pagal 4.13 ir 4.14 formules. Duomenų matmenų mažinimui ir atvaizdavimui buvo naudojamas t -SNE metodas su nu-



5.3 pav. RF permutacija pagrįsta kintamųjų svarba apklausos duomenims

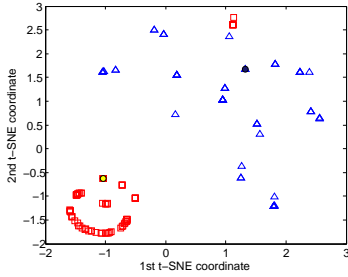


5.4 pav. RF permutacija pagrįsta kintamųjų svarba balso duomenims

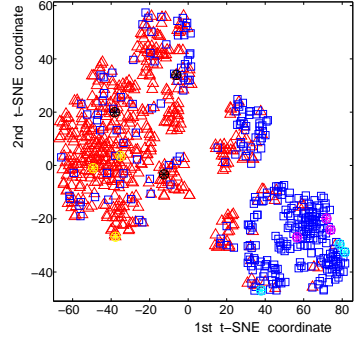
statyta perplexity parametro verte 50. Sekimo tikslais 2 pacientų duomenys iš „sveikos“ klasės buvo pažymėti skirtingomis spalvomis. Kaip matyti 5.5 paveikslėlyje, kartais pacientas gali atsidurti prie priešingos klasės pacientų, kas parodo, kad reikia išsamesnės analizės.

Balso duomenys atvaizduoti 5.6 paveikslėlyje panaudojant visus įrašus (kiekvienam pacientui po 3). Būtų galima tikėtis, kad vieno paciento žymekliai bus šalia, tačiau kaip tai iliustruoja skirtingų spalvų apskritimai 5.6 paveikslėlyje, tai ne visada tiesa. Klasifikavimo patikimumui vertinti buvo atvaizduoti duomenys su išklasifikuota klase, didinant žymeklį priklausomai nuo patikimumo (žr. 5.7 pav.). Palyginus gautą paveikslėlį su atvaizduotais pradiniais duomenimis buvo pastebėta, jog net ir klaidinga diagnozė gali būti pateikta su dideliu patikimumu.

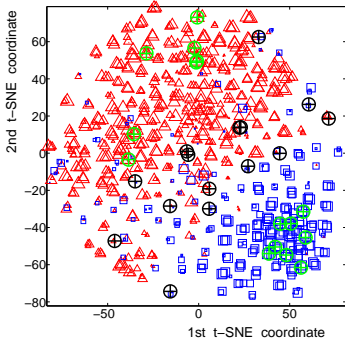
Tais atvejais, kai reikia papildomos analizės, galima pasinaudoti tikimybinio dažnio funkcijų grafikai. Tokių grafikų pavyzdys pateiktas 5.8 paveikslėlyje. Atliekant analizę, galima palyginti pažymėtas paciento vietas 5.5 ir 5.8 grafikuose, taip patikrinant ar abiejuose mato ta pati tendencija.



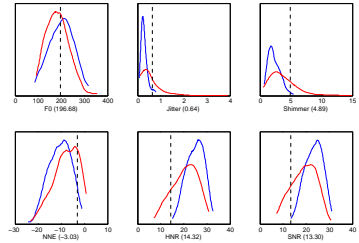
5.5 pav. Panašumų matricos 2D vaizdas gautas panaudojant *t*-SNE algoritimą (Balso ir apklausos duomenų apjungimas)



5.6 pav. Panašumų matricos vaizdas iš meta RF (Balso duomenys)



5.7 pav. Apibendrinta panašumų matrica. Žymeklio dydis atspindi sprendimo tikrumą (didesnis žymeklis – didesnis tikrumas). Juodi žymekliai atspindi mažo tikrumo sprendimus, o žali – didelio



5.8 pav. Tikimybinio pasiskirstymo funkcijų grafikai sveikos (*Mėlyna*) ir *patologinės* (*Raudona*) klasių garso parametrams

IŠVADOS

1. Neinvazinių balso patologijos metodų apžvalga atskleidė, kad tokių metodų egzistuoja labai daug, tačiau pagrindine jų problema išlieka per mažas tikslumas, kuris neleidžia jų naudoti kasdieniame gydytojų darbe. Tai parodo jog reikia atlikti daugiau tyrimų šioje srityje siekiant pagerinti balso patologijos klasifikavimo tikslumą.
2. Pasiūlytas nuo duomenų priklausančio atsitiktinio miško metodas pagerino klasifikavimo tikslumą 1,6 % – 3,19 %, lyginant su kitais šiame darbe bandytais metodais. Didžiausias pasiektas šio metodo tikslumas 86,37 % parodo didelį jo potencialą, tačiau reikia tolimesnių tyrimų siekiant jį patobulinti.
3. Naujai sukurtas, asociacijų taisyklėmis paremtas, metodas apklausos duomenų klasifikavimui pasiekė 88,89 % tikslumą. Tai sutampa su kitų autorių atradimais, kad balso patologija gali būti sėkmingai diagnozuojama panaudojant vien tik apklausos duomenis.
4. Balso ir apklausos duomenų klasifikavimas apjungiant rezultatus meta klasifikatoriumi leido pasiekti 90,48 % tikslumą. Tai yra 1,5 % daugiau, lyginant su geriausiais pasiektais rezultatais kai skirtingo tipo duomenys naudojami atskirai. Kadangi baziniams klasifikatoriams panaudoti visiškai „permatomi“ metodai, jie leidžia aiškiai matyti sprendimo priėmimo logiką ir yra naudingi kaip prevencinė ar mokomoji priemonė.
5. Dimensionalumo sumažinimas naudojant t-SNE metodą leido atvaizduoti pradinius duomenis ir klasifikatoriaus rezultataus grafiškai viename 2D paveikslėlyje, kas suteikia galimybę patogiau juos analizuoti. Statistinis duomenų atvaizdavimas panaudojant tikimybinio tankio funkcijas yra naudingas išsamesnei paciento analizei, parodo duomenų trūkumus ir naudojama kaip mokymo medžiaga.

Literatūra

- [1] Meisam Khalil Arjmandi, Mohammad Pooyan, Mohammad Mikaili, Mansour Vali, and Alireza Moqarehzadeh. Identification of voice disorders using long-time features and support vector machine with different feature reduction methods. *Journal of Voice*, 25(6):e275–e289, nov 2011.
- [2] M. Bacauskiene, A. Verikas, A. Gelzinis, and A. Vegiene. Random forests based monitoring of human larynx using questionnaire data. *Expert Systems with Applications*, 39(5):5506–5512, apr 2012.
- [3] Kevin K. Bach, Peter C. Belafsky, Kathleen Wasylik, Gregory N. Postma, and Jamie A. Koufman. Validity and reliability of the glottal function index. *Archives of Otolaryngology–Head & Neck Surgery*, 131(11):961, nov 2005.
- [4] Hans Behrbohm, Oliver Kaschke, Tadeus Nawka, and Andrew Swift. *Ear, Nose and Throat Diseases: With Head and Neck Surgery*. Thieme Medical, 3rd edition, August 2009.
- [5] Leo Breiman. *Random Forests*, volume 45. Kluwer Academic Publishers, Boston, October 2001.
- [6] Niko Brummer and Edward de Villiers. The BOSARIS toolkit: Theory, algorithms and code for surviving the new DCF. *arXiv*, 1304(2865v1):1–23, April 10 2013. Presented at the NIST SRE’11 Analysis Workshop, Atlanta, December 2011. Available at <http://sites.google.com/site/bosaristoolkit/>.
- [7] Hugo Tito Cordeiro, José Manuel Fonseca, and Carlos Meneses Ribeiro. LPC spectrum first peak analysis for voice pathology detection. *Procedia Technology*, 9:1104–1111, 2013.
- [8] Peijun Du, Alim Samat, Björn Waske, Sicong Liu, and Zhenhong Li. Random forest and rotation forest for fully polarized SAR image classification using polarimetric and spatial features. *ISPRS Journal of Photogrammetry and Remote Sensing*, 105:38–53, jul 2015.
- [9] Daniel Engel, Lars Huttenberger, and Bernd Hamann. A survey of dimension reduction methods for high-dimensional data analysis and visualization, 2012.

- [10] Fangwen Fu, Deepak S. Turaga, Olivier Verscheure, Mihaela van der Schaar, and Lisa Amini. Configuring competing classifier chains in distributed stream mining systems. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):548–563, dec 2007.
- [11] A. Gelzinis, A. Verikas, and M. Bacauskiene. Automated speech analysis applied to laryngeal disease categorization. *Computer Methods and Programs in Biomedicine*, 91(1):36–47, jul 2008.
- [12] Jorge Andrés Gómez-García, Laureano Moro-Velázquez, Juan Ignacio Godino-Llorente, and César Germán Castellanos-Domínguez. An insight to the automatic categorization of speakers according to sex and its application to the detection of voice pathologies: A comparative study. *Revista Facultad de Ingeniería Universidad de Antioquia*, (79), jun 2016.
- [13] Adidah Lajis and Normaziah Abdul Aziz. Competency assessment of short free text answers. In *2014 4th International Conference on Engineering Technology and Technopreneuship (ICE2T)*. Institute of Electrical and Electronics Engineers (IEEE), aug 2014.
- [14] D Lavanya. Ensemble decision tree classifier for breast cancer data. *International Journal of Information Technology Convergence and Services*, 2(1):17–24, feb 2012.
- [15] Claudia Lindner, Paul A. Bromiley, Mircea C. Ionita, and Tim F. Cootes. Robust and accurate shape model matching using random forest regression-voting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9):1862–1874, sep 2015.
- [16] Leonardo Wanderley Lopes, Jocélio Delfino da Silva, Layssa Batista Simões, Deyverson da Silva Evangelista, Priscila Oliveira Costa Silva, Anna Alice Almeida, and Maria Fabiana Bonfim de Lima-Silva. Relationship between acoustic measurements and self-evaluation in patients with voice disorders. *Journal of Voice*, 31(1):119.e1–119.e10, jan 2017.
- [17] Dalton Lunga, Saurabh Prasad, Melba M. Crawford, and Okan Ersoy. Manifold-learning-based feature extraction for classification of hyper-spectral data: A review of advances in manifold learning. *IEEE Signal Processing Magazine*, 31(1):55–66, jan 2014.
- [18] Youri Maryn, Marc De Bodt, Ben Barsties, and Nelson Roy. The value of the acoustic voice quality index as a measure of dysphonia severity in subjects speaking different languages. *Eur Arch Otorhinolaryngol*, oct 2013.

- [19] Jonas Minelga, Antanas Verikas, Evaldas Vaiciukynas, Adas Gelzinis, and Marija Bacauskiene. A transparent decision support tool in screening for laryngeal disorders using voice and query data. *Applied Sciences*, 2017.
- [20] David Moffat, David Ronan, and Joshua Reiss. An evaluation of audio feature extraction toolboxes, 2015.
- [21] Ghulam Muhammad, Mansour Alsulaiman, Awais Mahmood, and Zulfiqar Ali. Automatic voice disorder classification using vowel formants. In *2011 IEEE International Conference on Multimedia and Expo*. Institute of Electrical & Electronics Engineers (IEEE), jul 2011.
- [22] Ghulam Muhammad, Tamer A. Mesallam, Khalid H. Malki, Mohamed Farahat, Awais Mahmood, and Mansour Alsulaiman. Multidirectional regression (MDR)-based features for automatic voice disorder detection. *Journal of Voice*, 26(6):817.e19–817.e27, nov 2012.
- [23] R.M.M. Nelson Roy, S. Thibeault, R.A. Parsa, S.D. Gray, and E.M. Smith. Prevalence of voice disorders in teachers and the general population. *Journal of Speech, Language, and Hearing Research*, 47(2):281–293, 2004.
- [24] Ifeyinwa Okoye, Steven Bethard, and Tamara Sumner. Cu : Computational assessment of short free text answers - a tool for evaluating students understanding. In *Second Joint Conference on Lexical and Computational Semantics (SEM)*. Association for Computational Linguistics, jun 2013.
- [25] Senthil Kumar Palanisamy. Association rule based classification. Computer science, Worcester Polytechnic Institute, Worcester, Massachusetts, USA, May 2006. Prof. Carolina Ruiz, Advisor.
- [26] Daria Panek, Andrzej Skalski, Janusz Gajda, and Ryszard Tadeusiewicz. Acoustic analysis assessment in speech pathology detection. *International Journal of Applied Mathematics and Computer Science*, 25(3), jan 2015.
- [27] Rūta Pribušienė, Migle Baceviciene, Virgilijus Uloza, Aurelija Vegiene, and Jelena Antuseva. Validation of the lithuanian version of the glottal function index. *Journal of Voice*, 26(2):e73–e78, mar 2012.
- [28] Nicolás Sáenz-Lechón, Juan I. Godino-Llorente, Víctor Osmar-Ruiz, and Pedro Gómez-Vilda. Methodological issues in the development of automatic systems for voice pathology detection. *Biomedical Signal Processing and Control*, 1(2):120–128, apr 2006.

- [29] Ali Salih Mahmoud Saudi, Aliaa A. A. Youssif, and Atef Z. Ghalwash. Computer aided recognition of vocal folds disorders by means of RASTA-PLP. *CIS*, 5(2), feb 2012.
- [30] Bernard W. Silverman. *Density estimation for statistics and data analysis*. Monographs on statistics and applied probability. Chapman and Hall, London, UK, 1st edition, 1986.
- [31] Athanasios Tsanas. Acoustic analysis toolkit for biomedical speech signal processing: concepts and algorithms. 8th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications (MAVEBA), dec 2013.
- [32] E. Vaiciukynas, A. Verikas, A. Gelzinis, M. Bacauskiene, J. Minelga, M. Hällander, E. Padervinskis, and V. Uloza. Fusing voice and query data for non-invasive detection of laryngeal disorders. *Expert Systems with Applications*, 42(22):8445–8453, dec 2015.
- [33] Evaldas Vaičiukynas. *Computational Intelligence Methods for Voice Function Assessment and Laryngeal Pathology Detection*. PhD thesis, Kaunas University of Technology, Kaunas, 2013.
- [34] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, November 2008.
- [35] A. Verikas, A. Gelzinis, E. Vaiciukynas, M. Bacauskiene, J. Minelga, M. Hällander, V. Uloza, and E. Padervinskis. Data dependent random forest applied to screening for laryngeal disorders through analysis of sustained phonation: Acoustic versus contact microphone. *Medical Engineering & Physics*, 37(2):210–218, feb 2015.
- [36] Antanas Verikas, Adas Gelzinis, Marija Bacauskiene, Magnus Hällander, Virgilijus Uloza, and Marius Kaseta. Combining image, voice, and the patient’s questionnaire data to categorize laryngeal disorders. *Artificial Intelligence in Medicine*, 49(1):43–50, may 2010.
- [37] Xiang Wang, Jianping Zhang, and Yonghong Yan. Discrimination between pathological and normal voices using GMM-SVM approach. *Journal of Voice*, 25(1):38–43, jan 2011.
- [38] Kai-Pun Wong, Jung-Woo Woo, Jason Yu-Yin Li, Kyu Eun Lee, Yeo Kyu Youn, and Brian Hung-Hin Lang. Using transcutaneous laryngeal ultrasonography (TLUSG) to assess post-thyroidectomy patients’ vocal cords: Which maneuver best optimizes visualization and assessment accuracy? *World J Surg*, 40(3):652–658, nov 2015.

- [39] R. N. Wormald, R. J. Moran, R. B. Reilly, and P. D. Lacy. Performance of an automated, remote system to detect vocal fold paralysis. *Annals of Otolology, Rhinology & Laryngology*, 117(11):834–838, nov 2008.
- [40] Yu Zhang and Jack J. Jiang. Acoustic analyses of sustained and running voices from patients with laryngeal pathologies. *Journal of Voice*, 22(1):1–9, jan 2008.

Mokslinių publikacijų disertacijos tema sąrašas

Mokslinės informacijos instituto (ISI) referuojamuose leidiniuose:

1. A. VERIKAS, A. GELZINIS, E. VAICIUKYNAS, M. BACAUSKIENE, J. MINELGA, M. HALLANDER, V. ULOZA, E. PADERVINSKIS, Data dependent random forest applied to screening for laryngeal disorders through analysis of sustained phonation: acoustic versus contact microphone, *Medical Engineering & Physics*, 37, 2015, 210-218.
2. E. VAICIUKYNAS, A. VERIKAS, A. GELZINIS, M. BACAUSKIENE, J. MINELGA, M. HALLANDER, E. PADERVINSKIS, V. ULOZA, Fusing voice and query data for non-invasive detection of laryngeal disorders, *Expert Systems with Applications*, 42(22), 2015, doi:10.1016/j.eswa.2015.07.001.
3. J. MINELGA, A. VERIKAS, E. VAICIUKYNAS, A. GELZINIS, M. BACAUSKIENE, A transparent decision support tool in screening for laryngeal disorders using voice and query data, *Applied Sciences*, 2017, doi:10.3390/app7101096.

Pranešimai mokslinėse konferencijose:

1. Comparing throat and acoustic microphones for laryngeal pathology detection from human voice, *9th International Conference on Electrical and Control Technologies ECT-2014*, 2014 May 8-9, Kaunas, Lithuania.
2. Exploring sustained phonation recorded with acoustic and contact microphones to screen for laryngeal disorders, *2014 IEEE Symposium Series on Computational Intelligence: IEEE Symposium on Computational Intelligence in Healthcare and e-health*, 2014 December 9-12, Orlando, Florida, USA.

3. Towards Voice and Query Data-based Non-invasive Screening for Laryngeal Disorders, *14th International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases (AI-KED '15)*, 2015 January 10-12, Tenerife, Canary Islands, Spain.

Trumpai apie autorių

Jonas Minelga gimė 1987 metų balandžio 14 dieną Kaune. 2006 metais baigė Raudondvario gimnaziją. 2010 metais baigė KTU programų inžinerijos bakalauro studijas. 2012 metais tame pačiame fakultete įgijo programų sistemų inžinerijos magistro laipsnį. Nuo 2012 metų iki 2017 metų - KTU doktorantas. Turi 8 metų patirtį dirbant programuotoju ir 4 metų patirtį vedant Taikomosios elektronikos bei Mikroprocesorinių įtaisų ir valdiklių laboratorinius darbus. Mokslinių interesų sritys: skaitinis intelektas, signalų analizė, atpažinimų teorija. Autoriaus kontaktinis elektroninis paštas jonas.minelga@ktu.lt.

Computational intelligence methods for non-invasive larynx pathology diagnosis

Abstract

Laryngeal disorders affect roughly 5-6% of the general human population and larynx-related cancer alone causes around 200,000 annual deaths worldwide. This being one of only a few areas where annual deaths are increasing, requires effort in targeting easy, effective and accessible preventive laryngeal health care. This research analyzes computational intelligence techniques for larynx pathology detection, using non-invasive measurements, such as human voice recordings and answers to specific questionnaires. The intention was to develop a technique for voice and query data analysis, which would be capable of detecting voice pathology and providing support in screening for laryngeal disorder.

This study is performed by using a subject's voice recordings and answers to specific questionnaire, obtained by otolaryngology specialists. The collected data is gathered into 3 databases. Voice database contains sustained phonation (/a/ as in word „large“) recordings of 273 subjects (163 healthy and 110 pathological voices) varying in sex and age (from 19 to 85 years old), where each recording is labeled by a clinical diagnosis, obtained from clinical voice specialists. There are up to three recordings done for each patient, where part of them is also recorded with contact microphone. Query database contains data from 596 subjects (327 healthy and 269 pathological) also varying in sex and age. All subjects from the voice database are present in the query database. Patient classification discerns patients into *healthy* and *pathological* classes.

The methodological approach towards the analysis of voice and query data consisted of several steps. Firstly, characterization of audio recordings was obtained by using various techniques to extract diverse features from voice recordings. 14 different feature sets of varying size were extracted from each recording, resulting in 927 features per recording. Secondly, 6 audio parameters (extracted using „Dr. Speech“ software) were provided by otolaryngology specialists, together with

query data (answers to 25 questions). The three mentioned data sets were used for data classification either individually or in fusion. For data fusion task, new data dependent random forest-based way of available information from multiple data sets combination was introduced.

Random forest (RF) algorithm was used as base classifier for audio data classification and as a meta-learner in decision-level fusion cases. Both variations were used for classification of audio recordings from two types of microphones, which allowed assessment if contact microphones can provide useful information for classification accuracy improvement compared to acoustic ones or used together. Data dependent random forest-based data combination and classification technique was proposed and applied for voice data classification. In this work, affinity analysis of the query data was used to extract rules for each class (*healthy* and *pathological*), which allows identification if patient has larynx pathology. Classification of query data was performed by extracted association rules and Decision tree (DT) algorithm.

Two distinct techniques were used to represent voice and query data visually for inspection and analysis. Voice data was mapped to 2D space using *t*-SNE algorithm, which also allows analysis of subject's similarity. Meanwhile, graphs of probability density functions (PDFs) were obtained for analysis of query data.

Our introduced data-dependent random forest-based technique of data combination and classification helped to achieve the highest classification accuracy of 86.37%, compared to the results achieved by using only one single data set. The used affinity analysis highlighted 17 important questions which allows reduction of the questionnaire. Our developed association rules technique for query data classification together with DT are completely transparent, which allows deeper exploration of decision-making process and is very useful for teaching/learning purposes, early preventive health care. Results of acoustic and contact microphone comparison revealed that the acoustic microphone is superior to the contact one. However, contact microphones may be more useful in the noisy environments, but additional research is required to determine the noise level when contact microphone becomes superior to the acoustic one.

Applied visual data analysis techniques is another useful contribution of this research, which allows detection of incorrectly labeled or more thorough examination requiring subjects. PDFs provide additional information (statistical) about a patient and serve as a learning material, as well as show which type of data is not present in the used data set and might affect classification accuracy. Accurate pathology detection was observed for unseen subjects with equal error rate (EER) of 11.11% by using association rules and EER of 10.26% by using the decision tree. When using decision-level fusion, an even lower EER of only 9.52% was achieved.

In conclusion, results of this research indicate that the developed techniques can be very useful for diagnostics, education and exploratory tasks in Otolaryngology departments.

Research Object

The object of this research is voice pathology detection by using voice and query data analysis in separate and fusion matter. Non-invasiveness is the main advantage of this data.

Objectives and Tasks

The main objective of this research is to improve voice pathology classification accuracy by developing new classification technique, which would be capable of classifying data from multiple sources. Second objective is to improve classification accuracy by the use of voice and questionnaire data decision-level fusion.

Tasks formulated for this study:

1. Analyze state-of-the-art work of laryngeal pathology detection identifying the drawbacks of techniques used.
2. Review classification techniques using voice and query data fusion and identify their limitations.
3. Propose a new technique for combination and classification of data available in different feature sets.

4. Develop a transparent technique for voice and query data analysis.
5. Experimentally validate the proposed techniques with out of bag (OOB) data.
6. Develop a non-invasive decision support system for voice pathology detection, providing multiple approaches for graphical patient data analysis and employing several data sources.

Conclusions

1. The review of non-invasive techniques used for voice pathology detection showed that many techniques exist for this task, however classification accuracy remains an obstacle while trying to apply these techniques in everyday work of medical specialists. This indicates that more research is required to improve the accuracy of these techniques.
2. Proposed data dependent Random Forest based technique for combination and classification of data from multiple data sets increased classification accuracy by 1.6% - 3.19%, compared to other our tested techniques. The maximum achieved accuracy of 86.37% indicates high potential of our method, but further improvements are necessary for even higher accuracy improvement.
3. The newly constructed algorithm from our extracted association rules achieved Equal Error Rate of only 11.11% . This confirmed the finding of related work analysis that successful voice pathology detection can be performed using only query data which represents patient's voice quality and function evaluation.
4. Voice and query data classification by combination of meta-learner achieved Equal Error Rate of only 9.52%. This was a 1.5% improvement, compared to the best achieved result when classifying both types of data separately. Base classifiers, Association rules and Decision tree, being completely transparent, provide the required transparency and helps in preventive health care and for learning purposes.

5. Dimensionality reduction by T-distributed stochastic neighbor embedding allowed to visualize initial data and classifier decision in a single two-dimensional image, which helps to analyze data and decisions. Statistical representation of data by Probability Density Functions is useful for deeper patient analysis and allows to indicate data deficiency, serves as a learning material.

UDK 004.89 + 616.22-008-07](043.3)

SL344. 2018-10-26, 3 leidyb. apsk. l. Tiražas 50 egz.

Išleido Kauno technologijos universitetas, K. Donelaičio g. 73, 44249
Kaunas

Spausdino leidyklos "Technologija" spaustuvė, Studentų g. 54, 51424
Kaunas