



Kauno technologijos universitetas

Informatikos fakultetas

Širdies plakimo garso įrašų (fonokardiogramų) klasifikavimui skirtų metodų tyrimas

Baigiamasis magistro studijų projektas

Mantas Gricius

Projekto autorius

prof. dr. Dalia Čalnerytė

Vadovė

Kaunas, 2025



Kauno technologijos universitetas

Informatikos fakultetas

Širdies plakimo garso įrašų (fonokardiogramų) klasifikavimui skirtų metodų tyrimas

Baigiamasis magistro studijų projektas

Dirbtinio intelekto informatika (6211BX007)

Mantas Gričius

Projekto autorius

prof. dr. Dalia Čalnerytė

Vadovė

dr. Vytenis Putys

Recenzentas

Kaunas, 2025



Kauno technologijos universitetas

Informatikos fakultetas

Mantas Gricius

Širdies plakimo garso įrašų (fonokardiogramų) klasifikavimui skirtų metodų tyrimas

Akademinio sąžiningumo deklaracija

Patvirtinu, kad:

1. baigiamąjį projektą parengiau savarankiškai ir sąžiningai, nepažeisdama(s) kitų asmenų autoriaus ar kitų teisių, laikydamasi(s) Lietuvos Respublikos autorių teisių ir gretutinių teisių įstatymo nuostatų, Kauno technologijos universiteto (toliau – Universitetas) intelektinės nuosavybės valdymo ir perdavimo nuostatų bei Universiteto akademinės etikos kodekse nustatytų etikos reikalavimų;
2. baigiamajame projekte visi pateikti duomenys ir tyrimų rezultatai yra teisingi ir gauti teisėtai, nei viena šio projekto dalis nėra plagijuota nuo jokių spausdintinių ar elektroninių šaltinių, visos baigiamojo projekto tekste pateiktos citatos ir nuorodos yra nurodytos literatūros sąrašė;
3. įstatymų nenumatytų piniginių sumų už baigiamąjį projektą ar jo dalis niekam nesu mokėjęs (-usi);
4. suprantu, kad išaiškėjus nesąžiningumo ar kitų asmenų teisių pažeidimo faktui, man bus taikomos akademinės nuobaudos pagal Universitete galiojančią tvarką ir būsiu pašalinta(s) iš Universiteto, o baigiamasis projektas gali būti pateiktas Akademinės etikos ir procedūrų kontrolieriaus tarnybai nagrinėjant galimą akademinės etikos pažeidimą.

Mantas Gricius

Patvirtinta elektroniniu būdu

Gricius, Mantas. Širdies plakimo garso įrašų (fonokardiogramų) klasifikavimui skirtų metodų tyrimas. Magistro baigiamasis projektas / vadovė prof. dr. Dalia Čalnerytė; Kauno technologijos universitetas, informatikos fakultetas.

Studijų kryptis ir sritis (studijų krypčių grupė): Informatikos mokslai, Informatika (B01).

Reikšminiai žodžiai: fonokardiogramos, dirbtinis intelektas, mašininis mokymas, klasifikavimas, giliųjų dirbtinių neuronų tinklas, konvoliucinis neuroninis tinklas.

Kaunas, 2025. 54 p.

Santrauka

Pasaulio sveikatos organizacijos duomenimis, širdies ir kraujagyslių ligos yra pagrindinė mirties priežastis pasaulyje. Dirbtinio intelekto klasifikavimo algoritmai gali padėti aptikti širdies ligas ankstyvoje stadijoje. Šio darbo tikslas – sukurti algoritmą širdies plakimo garso įrašams klasifikuoti (tipinis / netipinis signalas) ir palyginti gauto modelio balansuotą tikslumą su egzistuojančiais sprendimais. Viena pagrindinių problemų šios srities sprendimuose yra ribotas mokymo duomenų kiekis. Šiame darbe naudojamas „2016 CinC/PhysioNet Challenge“ duomenų rinkinys. Pagrindinė problema duomenų rinkinyje – duomenų disbalansas. Tipinių garso signalų yra beveik 4 kartus daugiau nei netipinių. Išbandomi 3 skirtingi klasifikavimo modeliai (atraminių vektorių klasifikatorius (SVM), atsitiktiniai miškai (RF) ir konvoliuciniais neuroniniais tinklais (CNN) paremtas giliųjų dirbtinių neuronų tinklas) ir gauti rezultatai palyginami su egzistuojančiais sprendimais. Eksperimentai atliekami naudojant 3 skirtingus įvesčių tipus: naudojant segmentaciją, kai garso įrašai yra segmentuojami į medicinoje atpažįstamus širdies garsų segmentus (S1, S2, S3 ir S4) ir iš gautų segmentų sudaromos savybės, kurios naudojamos kaip įvestis modeliams; nenaudojant segmentacijos ir įvestys gaunamos apskaičiuojant MFCC kadrų vidurkius; naudojant augmentaciją prijungiami papildomi sintetiniai netipiniai garso įrašai, kurie panaikina duomenų disbalansą ir MFCC kadrų vidurkiai naudojami kaip įvestis. Naudojant augmentaciją pašalinus duomenų disbalansą, visi modeliai klasifikavo fonokardiogramas 5-10% tiksliau. Geriausią rezultatą pavyko pasiekti naudojant CNN paremtą giliųjų dirbtinių neuronų tinklą – 0,94 MAcc, tačiau ganėtinai panašus rezultatas tikslumo prasme buvo gautas ir naudojant SVM metodą – 0,9392 MAcc. Tiek CNN, tiek SVM metodais pavyko pasiekti aukštesnę klasifikavimo tikslumą, nei geriausiai tarp literatūroje nagrinėtų egzistuojančių sprendimų.

Gricius, Mantas. A study of methods for classifying heart sound recordings (phonocardiograms). Master's Final Degree Project / supervisor prof. dr. Dalia Čalnerytė; Faculty of Informatics, Kaunas University of Technology.

Study field and area (study field group): Computer science, Informatics (B01).

Keywords: phonocardiograms, artificial intelligence, machine learning, classification, deep neural network, convolutional neural network.

Kaunas, 2025. 54 pages.

Summary

According to the World Health Organization, cardiovascular diseases are the leading cause of death in the world. Artificial intelligence classification algorithms can help detect heart disease at an early stage. The goal of this work is to develop an algorithm for classifying heartbeat audio recordings (normal / abnormal signal) and compare the balanced accuracy of the resulting model with existing solutions. One of the main problems in phonocardiogram classification field is the limited amount of training data. This work uses the “2016 CinC/PhysioNet Challenge” dataset. The main problem in the dataset is data imbalance. There are almost 4 times more normal audio signals than abnormal ones. 3 different classification models (support vector classifier (SVM), random forests (RF) and a deep neural network based on convolutional neural networks (CNN)) are used to classify heart sounds and the results obtained are compared with existing solutions. Experiments are performed using 3 different input types: using segmentation, when audio recordings are segmented into medically recognizable heart sound segments (S1, S2, S3 and S4) and features are formed from the resulting segments, which are used as input to the models; without using segmentation, where inputs are obtained by calculating the averages of MFCC frames; using augmentation, where additional synthetic abnormal audio recordings are appended to the dataset to eliminate data imbalance and the averages of MFCC frames are used as input. When using augmentation and eliminating data imbalance, all models classified phonocardiograms 5-10% more accurately. The best result was achieved using a CNN-based deep neural network – 0.94 MAcc, quite similar result in terms of accuracy was also obtained using the SVM method – 0.9392 MAcc. Both CNN and SVM methods managed to achieve higher classification accuracy than the best existing solutions examined in the literature.

Turinys

Lentelių sąrašas	8
Paveikslų sąrašas	9
Santrumpų ir terminų sąrašas	10
Įvadas.....	11
1. Širdies plakimo garso įrašų (fonokardiogramų) klasifikavimo literatūros analizė	12
1.1. Fonokardiogramų klasifikavimui naudojami duomenų rinkiniai	12
1.1.1. „2016 CinC/PhysioNet Challenge“	12
1.1.2. „Pattern Analysis, Statistical Modeling and Computational Learning (PASCAL) Heart Sounds Challenge“	12
1.1.3. „CirCor DigiScope Phonocardiogram“	13
1.1.4. Duomenų rinkinių apibendrinimas	14
1.2. Fonokardiogramų duomenų požymių išskyrimas	15
1.2.1. CNN.....	15
1.2.2. „Mel-dažnio Cepstrum“ koeficientai.....	16
1.3. Fonokardiogramų klasifikavimo metodai.....	16
1.4. Problemos esamuose sprendimuose	17
1.4.1. Mokymas su ribotu kiekiu širdies garso duomenų	18
1.4.2. Mokymo efektyvumas ir sparta	18
1.4.3. DNN reikalaujami ištekliai.....	18
2. Širdies plakimo garso įrašų eksperimentų eiga	19
2.1. Klasterizavimo eksperimentų eiga	19
2.2. Klasifikavimo eksperimentų eiga su segmentacija.....	20
2.3. Klasifikavimo eksperimentų eiga be segmentacijos.....	21
2.4. Sintetinių duomenų generavimo eiga	22
3. Širdies plakimo garso įrašų apdorojimas ir segmentacija.....	24
3.1. Duomenų rinkinys	24
3.2. Širdies plakimo garso įrašo struktūra	24
3.3. Duomenų apdorojimas atliekant segmentaciją.....	25
3.4. Segmentacijos procesas	27
3.5. Homomorfinis apvalkalas (angl. <i>homomorphic envelope</i>).....	28
3.6. Hilberto apvalkalas (angl. <i>Hilbert envelope</i>).....	29
3.7. Segmentacijos rezultatai	31
3.8. Duomenų apdorojimas nenaudojant segmentacijos ir MFCC požymių išskyrimas.....	31
3.8.1. Augmentuoti netipiniai širdies plakimo garso įrašai.....	33
4. Eksperimentai	35
4.1. Aplinka	35
4.2. Klasterizavimo eksperimentai	35
4.2.1. Calinski-Harabasz (CH) įvertis	35
4.2.2. K-vidurkių metodo aprašymas	36
4.2.3. Rezultatai.....	37
4.3. Klasifikavimo eksperimentai.....	39
4.3.1. SVM metodas	39
4.3.2. Atsitiktinių miškų metodas.....	42
4.3.3. CNN metodas	46

4.3.4. Klasifikavimo eksperimentų rezultatų apibendrinimas	49
4.4. Palyginimas su egzistuojančiais sprendimais	50
Išvados	51
Literatūros sąrašas	52

Lentelių sąrašas

1.1 lentelė. Duomenų rinkinių apibendrinimas	15
1.2 lentelė. PCG klasifikavimo tyrimų analizė.....	17
4.1 lentelė. K-vidurkių klasterių pasiskirstymas	37
4.2 lentelė. Klasterių garso įrašų savybių vidutinės reikšmės	38
4.3 lentelė. SVM rezultatai naudojant duomenų segmentaciją	40
4.4 lentelė. SVM rezultatai nenaudojant segmentacijos.....	41
4.5 lentelė. SVM rezultatai naudojant augmentaciją.....	42
4.6 lentelė. Atsitiktinių miškų rezultatai naudojant segmentaciją.....	43
4.7 lentelė. Atsitiktinių miškų rezultatai nenaudojant segmentacijos	44
4.8 lentelė. Atsitiktinių miškų rezultatai naudojant augmentaciją.....	45
4.9 lentelė. CNN rezultatai naudojant segmentaciją.....	47
4.10 lentelė. CNN rezultatai nenaudojant segmentacijos	48
4.11 lentelė. CNN rezultatai naudojant augmentaciją	49
4.12 lentelė. Bendri rezultatai.....	50
4.13 lentelė. Gautų rezultatų palyginimas su egzistuojančiais sprendimais.....	50

Paveikslų sąrašas

1 pav. Duomenų surinkimo proceso sekų diagrama.....	14
2 pav. Klasterizavimo eksperimentų veiklų diagrama.....	20
3 pav. Klasifikavimo eksperimentų veiksmų seka naudojant segmentaciją.....	21
4 pav. Klasifikavimo eksperimentų veiksmų seka nenaudojant segmentacijos	22
5 pav. Duomenų augmentacijos veiksmų seka.....	23
6 pav. Duomenų pasiskirstymo diagrama.	24
7 pav. Tipinio širdies plakimo garso įrašo struktūra	25
8 pav. Butterworth'o žemų dažnių filtravimas	26
9 pav. Butterworth'o aukštų dažnių filtravimas.	26
10 pav. Originalaus ir filtruoto garso įrašų palyginimas.	27
11 pav. Springer [24] segmentacijos proceso eiga	28
12 pav. Homomorfinis apvalkalas pritaikytas vienam iš garso įrašų.	29
13 pav. Hilberto apvalkalas pritaikytas vienam iš garso įrašų.	30
14 pav. Segmentacijos procesu atpažinti širdies garsai	31
15 pav. Tipiniai ir netipiniai širdies plakimo garso įrašai	32
16 pav. Tipinių ir netipinių įrašų MFCC	33
17 pav. Netipinių ir augmentuotų fonokardiogramų palyginimas.....	33
18 pav. Netipinių ir augmentuotų netipinių įrašų MFCC	34
19 pav. Calinski-Harabasz įverčių grafikas.	36
20 pav. Garso įrašai esantys arčiausiai kiekvieno K-vidurkių klasterio.....	38
21 pav. SVM klaidų matricos naudojant segmentaciją	40
22 pav. SVM klaidų matricos nenaudojant segmentacijos.....	41
23 pav. SVM klaidų matricos naudojant augmentaciją.....	42
24 pav. Atsitiktinių miškų klaidų matricos naudojant segmentaciją.....	44
25 pav. Atsitiktinių miškų klaidų matricos nenaudojant segmentacijos	45
26 pav. Atsitiktinių miškų klaidų matricos naudojant augmentaciją	46
27 pav. CNN klaidų matricos naudojant segmentaciją.....	47
28 pav. CNN klaidų matricos nenaudojant segmentacijos	48
29 pav. CNN klaidų matricos naudojant augmentaciją	49

Santrumpų ir terminų sąrašas

Santrumpos:

VHD – širdies vožtuvų ligos (angl. *Valvular Heart Diseases*).

PCG – fonokardiograma (angl. *Phonocardiogram*).

DNN – giliųjų dirbtinių neuronų tinklai (angl. *Deep Neural Networks*).

DI – dirbtinis intelektas.

CNN – konvoliucinis neuroninis tinklas (angl. *Convolutional Neural Network*).

ReLU – „lygintuvo“ aktyvacijos funkcija (angl. *Rectified Linear Unit*).

MFCC – „Mel-dažnio Cepstrum“ koeficientai (angl. *Mel-Frequency Cepstrum Coefficients*).

FFT – Furjė transformacija (angl. *Fast Fourier Transform*).

SVM – atraminių vektorių mašina (angl. *Support Vector Machine*).

GAN – generatyvinis priešininkų tinklas (angl. *Generative Adversarial Network*).

FHS – pagrindiniai širdies garsai (angl. *Fundamental Heart Sounds*).

HMM – paslėptasis Markovo modelis (angl. *Hidden Markov Model*).

RF – atsitiktinių miškų metodas (angl. *Random Forest*).

CV – kryžminė validacija (angl. *Cross Validation*).

MAcc – subalansuotas tikslumas (angl. *Balanced Accuracy*).

Įvadas

Darbas priklauso Dirbtinio intelekto informatikos studijų programai.

Toliau supažindinama su darbo specifika, aktualumu, išdėstomi tikslai bei uždaviniai, praktinė darbo reikšmė.

Darbo problematika ir aktualumas

Pasaulio sveikatos organizacijos duomenimis, širdies ir kraujagyslių ligos yra pagrindinė mirties priežastis pasaulyje. Apskaičiuota, kad 2019 m. nuo kraujotakos sistemos ligų mirė 17,9 mln. žmonių, t. y. 32 proc. visų mirčių pasaulyje. Iš šių mirčių 85 proc. mirė nuo širdies infarkto ir insulto [1]. Širdies vožtuvų ligos (angl. *valvular heart diseases*, VHD) yra viena iš pagrindinių širdies ir kraujagyslių ligų priežasčių, dėl kurių visame pasaulyje yra didelis mirtingumas [2]. Ankstyvas tokių problemų diagnozavimas gali padėti sumažinti mirtingumą ir sustabdyti ligos vystymąsi.

Širdies garsas yra tam tikras fiziologinis signalas, o jo matavimas vadinamas fonokardiografija. Tai viena iš veiksmingų diagnostikos priemonių. Širdies sistolės ir diastolės metu ji gali atspindėti fiziologinę informaciją apie kūno komponentus, tokius kaip prieširdžiai, skilveliai ir stambiosios kraujagyslės, bei jų funkcinę būklę [3]. Pagrindinė problema dirbtinio intelekto taikymams šioje srityje yra duomenų surinkimo sudėtingumas. Tokius duomenis surinkti sunku, todėl sudėtinga sėkmingai išmokyti klasifikavimo modelį.

Siekiant sušvelninti šią problemą, buvo pristatytas „2016 CinC/PhysioNet iššūkis“ [4]. Duomenys, pateikti šiam iššūkiui, suteikė galimybę sukurti įvairius mašininio mokymosi algoritmus, skirtus fonokardiogramų (angl. *phonocardiogram*, PCG) įrašams klasifikuoti į tipinius arba netipinius garsus. Nors sukurta daug projektų ir straipsnių, dabartinė problema yra ta, kad sėkmingiausi giliųjų dirbtinių neuronų tinklai (angl. *Deep Neural Networks*, DNN) reikalauja daug resursų. Supaprastinus ir sumažinus DNN metodus fonokardiogramų klasifikacijai, būtų galima pagerinti tokių sistemų prieinamumą mobiliuosiuose įrenginiuose, o tai atvertų galimybes atlikti fonokardiografinę stebėseną realiu laiku ir išvengti daugelio su širdimi susijusių ligų ir problemų [5]. Taip pat dauguma modelių yra priklausomi nuo segmentacijos proceso, kuris dažnu atveju reikalauja tiek pat resursų kaip ir pačių įrašų klasifikavimas ar net modelių mokymas. Paspartinus ar atsisakius šio proceso būtų galima paspartinti įrašų klasifikavimą ir pagerinti modelių efektyvumą.

Darbo tikslas ir uždaviniai

Darbo tikslas – sukurti algoritmą širdies plakimo garso įrašams klasifikuoti (tipinis / netipinis signalas) ir palyginti gauto modelio balansuotą tikslumą su egzistuojančiais sprendimais.

Darbo uždaviniai:

1. atlikti literatūros apie klasifikavimo metodus, požymių išskyrimą, duomenų rinkinius apžvalgą;
2. išanalizuoti egzistuojančių sprendimų realizacijas;
3. atlikti duomenų analizę ir apdorojimą;
4. sukurti algoritmą širdies plakimo garso įrašams klasifikuoti;
5. atlikti skaitinius eksperimentus ir siūlomo algoritmo rezultatus palyginti su egzistuojančių sprendimų rezultatais.

1. Širdies plakimo garso įrašų (fonokardiogramų) klasifikavimo literatūros analizė

Šiame skyriuje pateikiama fonokardiogramų klasifikavimo tyrimų analizė, duomenų rinkinių pasirinkimai, aprašomos naudojamos dirbtinio intelekto (DI) technologijos, analizuojamos jų silpnosios savybės, ieškoma galimybių jų tobulinimui. Taip pat analizuojami garso įrašų požymių išskyrimo metodai. Analizė atliekama atsižvelgiant į kelis išskirtus esminius kriterijus, kurių tikslas – užtikrinti, kad literatūros analizės metu būtų atrasti būdai atlikti projektą bei atrasta kaip gali būti patobulinti esami sprendimai.

Esminiai analizuojami aspektai (kriterijai):

1. duomenų rinkinių apžvalga;
2. požymių išskyrimas;
3. klasifikavimo metodai;
4. esamų sprendimų silpnosios savybės.

1.1. Fonokardiogramų klasifikavimui naudojami duomenų rinkiniai

Fonokardiogramų klasifikavimui, kaip ir daugeliui kitų mašininio mokymosi problemų, reikalingas kokybiškas ir apdorotas duomenų rinkinys. Būtent šiam uždaviniui (fonokardiogramų klasifikacijai) klasifikavimo modelio validacijai / testavimui svarbu naudoti viešai prieinamą ir žinomą duomenų rinkinį, kad būtų galima palyginti gauto modelio tikslumą esamų sprendimų spektre.

1.1.1. „2016 CinC/PhysioNet Challenge“

„2016 CinC/PhysioNet Challenge“ – tai 2016-aisiais metais įkurtas iššūkis, kurio tikslas paskatinti kurti algoritmus, skirtus širdies garso įrašams, surinktiems įvairiose klinikinėse ir neklinikinėse aplinkose, klasifikuoti [4].

Tikslui pasiekti buvo sukurta didelė duomenų bazė iš įvairių pasaulio mokslinių tyrimų įrašų, gautų įvairiose realiose klinikinėse ir neklinikinėse aplinkose (pvz., apsilankymai namuose), taip paskatinant kurti algoritmus, kurie iš vieno trumpo įrašo (10–60 s) tiksliai nustatytų, ar signalas yra normalus, nenormalus, ar prastos kokybės, iš ko būtų galima nustatyti, ar įrašo objektas turėtų būti nukreiptas ekspertinei diagnozei [6]. Iki šio iššūkio nebuvo jokios reikšmingos atviros prieigos širdies garsų duomenų bazės, pagal kurią mokslininkai galėtų mokytis ir vertinti automatinės diagnostikos algoritmus [7]. Be to, nebuvo prieinami jokie atviro kodo širdies garsų segmentavimo ir klasifikavimo algoritmai. Šis iššūkis iš esmės pakeitė šią situaciją [4].

1.1.2. „Pattern Analysis, Statistical Modeling and Computational Learning (PASCAL) Heart Sounds Challenge“

„PASCAL Heart Sounds Challenge“ – tai 2011-aisiais metais paskelbtas iššūkis, kurio tikslas paskatinti širdies garso įrašų klasifikavimo algoritmų kūrimą.

Šio iššūkio duomenų rinkinys buvo surinktas iš dviejų šaltinių: A – iš visuomenės naudojant „iPhone“ įrenginio programėlę „iStethoscope Pro“; B – iš klinikinių tyrimų ligoninėse, naudojant skaitmeninį stetoskopą „DigiScope“ [8].

Taip pat būtų verta paminėti, jog anksčiau minėto „2016 CinC/PhysioNet Challenge“ iššūkio duomenys yra sudaryti iš kelių duomenų bazių ir viena iš jų yra būtent „PASCAL Heart Sounds Challenge“ iššūkio duomenų rinkinys.

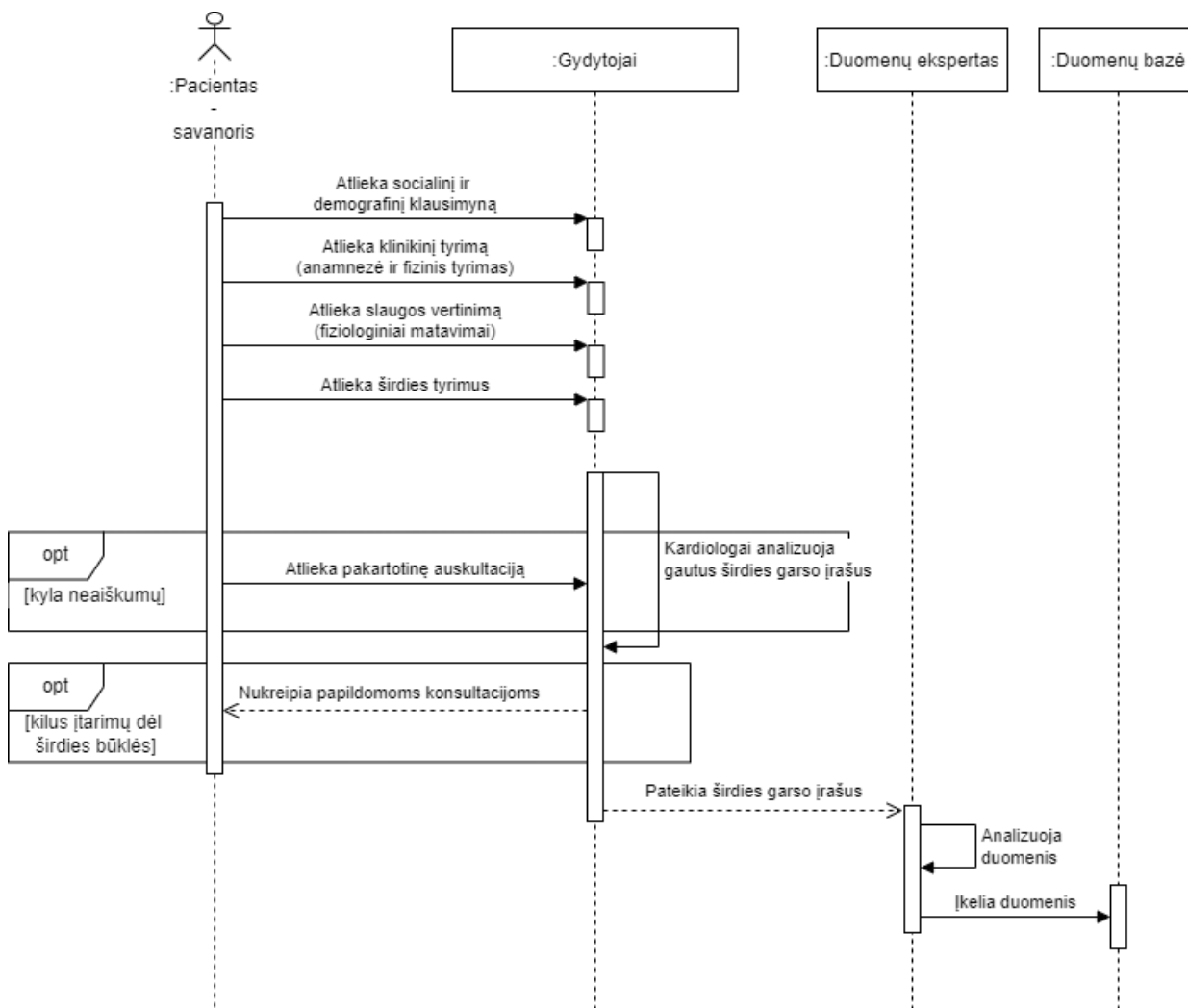
1.1.3. „CirCor DigiScope Phonocardiogram“

Duomenų rinkinys buvo surinktas per dvi masinės patikros kampanijas, vykdytas šiaurės rytų Brazilijoje 2014 ir 2015 m. [9]. Tikslinę populiaciją sudarė visi dalyviai, savanoriškai atvykę atlikti atrankinės patikros per tyrimo laikotarpį. Į tyrimą buvo įtraukti jaunesni nei 21 metų amžiaus asmenys, turintys tėvų pasirašytą sutikimą (jei reikia). Iš viso 2014 ir 2015 m. „Caravana do Coração“ (portugališkai „Širdies karavanas“) kampanijose dalyvavo 2061 dalyvis, 493 dalyviai buvo pašalinti, nes neatitiko tinkamumo kriterijų. Visi dalyviai užpildė sociodemografinį klausimyną, vėliau jiems buvo atliktas klinikinis tyrimas (anamnezė ir fizinė apžiūra), slaugos įvertinimas (fiziologiniai matavimai) ir širdies tyrimai (krūtinės ląstos rentgenograma, elektrokardiograma ir echokardiograma), kai jie buvo paskirti atliktus tyrimą.

Buvo atliktas duomenų kokybės vertinimas, visi įrašai buvo patikrinti dėl neteisingai įvestų ar išmatuotų verčių, nenuoseklių duomenų ar nukrypimų ir atitinkamai ištrinti. Po to duomenų rinkinys sudaromas priskiriant gautus įrašus prie užfiksuotų socialinių, demografinių bei klinikinių kintamųjų [10].

Vėliau buvo atlikta elektroninė auskultacija ir paprastai buvo renkami garso mėginiai iš keturių auskultacijos taškų. Visus mėginius rinko tas pats operatorius visą patikros laiką realioje klinikinėje aplinkoje. Gautuose fonokardiogramos (PCG) garso failuose buvo įvertinta signalo kokybė, o juos segmentavo širdies fiziologai. Kiekvienos kampanijos metu signalo kokybės vertinimą ir segmentavimą atliko skirtingi ekspertai, o užrašų anotaciją atliko tas pats anotatorius visam duomenų rinkiniui (įskaitant abi kampanijas). Iš viso 119 dalyvių įrašai neatitiko reikalaujamų signalo kokybės standartų, t. y. šių tiriamųjų įrašai nepadėjo patikimai apibūdinti ir aprašyti užesio. Šiuos įrašus anotatorius laikė neidentifikuojamais (nežinomais) [10].

1.1.4. Duomenų rinkinių apibendrinimas



1 pav. Duomenų surinkimo proceso sekų diagrama

Apibendrinant, buvo išanalizuoti 3 populiariausi duomenų rinkiniai (žr. 1.1 lentelė.). Daugiausiai širdies garso įrašų turi „CirCor DigiScope“ duomenų rinkinys (5272 įrašai). Mašininio mokymo modeliams mokytis dažnai yra naudojami net šimtai tūkstančių duomenų įrašų, taigi net ir 5272 įrašai nėra labai didelis kiekis duomenų. Duomenų surinkimas fonokardiogramoms klasifikuoti yra labai svarbus, tačiau ganėtinai sudėtingas procesas. Širdies plakimo garso įrašų duomenims surinkti reikalingas didelis kiekis savanorių pacientų bei gydytojų, kurie atlieka tyrimus pacientams ir padeda suskirstyti garsus į tipinius ir netipinius. Taip pat reikalingi duomenų ekspertai, kurie suformatuoja duomenis į tinkamus formatus ir įkelia duomenis į duomenų bazines (žr. 1 pav.). Todėl duomenų trūkumas vis dar lieka viena iš pagrindinių fonokardiogramų klasifikavimo uždavinio problemų. Nepaisant to, 1.3 skyriuje yra apžvelgiami egzistuojantys sprendimai, kuriems pavyko pasiekti gan aukštą klasifikavimo tikslumą.

1.1 lentelė. Duomenų rinkinių apibendrinimas

Duomenų rinkinio pavadinimas	Klasių skaičius	Įrašų skaičius	Papildomi duomenys	Pastabos
2016 „CinC/PhysioNet Challenge“	3 klasės (tipinis, netipinis ir neaiškūs).	Iš viso 3153 įrašai. 576 netipiniai, 2303 tipiniai ir 278 neaiškūs.	-	Duomenų rinkinys sudarytas iš 8 skirtingų duomenų bazių, viena iš jų – PASCAL iššūkio duomenys.
PASCAL	A dalis turi 4 klases: tipinis, užesys, papildomas širdies tonas, artefaktas. B dalis turi 3 klases: tipinis, užesys, ekstrasistolė.	Viso 832 įrašai. A: 176 įrašai (31 tipiniai, 34 užesiai, 19 papildomi širdies tonai, 40 artefaktai, 52 nežinomi). B: 656 įrašai (320 tipiniai, 95 užesiai, 46 ekstrasistolė, 195 nežinomi).	-	Šis rinkinys buvo įtrauktas į 2016 „CinC/PhysioNet Challenge“ duomenų rinkinį
„CirCor DigiScope Phonocardiogram“	2 klasės (tipinis, netipinis).	Viso 5272 įrašai.	Socialiniai-demografiniai duomenys (pacientų svoris, ūgis, lytis, amžiaus grupė).	-

1.2. Fonokardiogramų duomenų požymių išskyrimas

Vienas iš svarbių žingsnių norint atlikti širdies garso įrašų klasifikavimą yra požymių išskyrimas. Mokslininkai įvairiuose tyrimuose šį etapą atlieka skirtingai, naudojami konvoliuciniai neuroniniai tinklai, „Mel-dažnio Cepstrum“ koeficientai, Furjė transformacijos, ir nuo šių skirtingų duomenų požymių išskyrimo būdų priklauso ir gaunamas rezultatas. Šiame poskyryje aprašomi įvairūs požymių išskyrimo būdai.

1.2.1. CNN

Vienas iš būdų išskirti požymius yra panaudoti konvoliucinį neuroninį tinklą [11, 2, 12, 13]. Šis būdas yra itin patogus jei klasifikavimui yra naudojamas DNN. Būtent PCG garsų požymiams išskirti dažnai yra naudojamas vienmatis CNN (1D-CNN). Požymiams išskirti galima naudoti įvairių parametų ir struktūrų CNN tinklą.

Vienas iš pavyzdžių būtų Krishnan et al. [11] sprendimas:

Požymių išskyrimui iš įvesties laiko sekos sudaromas 1D-CNN iš atsitiktinai priskiriamų filtrų. Sudaromi trys skirtingų struktūrų tinklai. Visi trys tinklai naudoja „lygintuvo“ (angl. *rectified linear unit*, ReLU) aktyvacijos funkciją. Visas požymių išskyrimo procesas aprašomas formule:

$$c_{j,i} = \begin{cases} b_i + \sum_i^z k_i \cdot X_i, & z > 0, \\ 0, & z < 0 \end{cases}, \quad (1)$$

čia c - konvoliucinių sluoksnių požymių žemėlapiai, b – nuolydis (angl. *bias*), k – branduolys arba filtrų svoriai, X – įvesties duomenys.

1.2.2. „Mel-dažnio Cepstrum“ koeficientai

„Mel-dažnio Cepstrum“ koeficientai (angl. *Mel-Frequency Cepstrum Coefficients*, MFCC) – vienas geriausių ir populiariausių būdų parametriškai atvaizduoti garso signalus [14]. MFCC atvaizduoja garso signalus Mel dažnių skalėje, kuri padeda atskirti labiausiai žmogaus ausiai girdimus signalus. Taip pat MFCC praplečia parametrus gautus pritaikius greitąją Furjė transformaciją (FFT), nes leidžia atvaizduoti garsus tiek laiko, tiek dažnių atžvilgiu. Tai padeda sumažinti parametru skaičių (lyginant su originalaus garso parametrais) ir išsaugoti svarbiausias garso įrašo savybes, kas yra itin svarbu norint taikyti giliojo mokymosi algoritmus [15].

MFCC parametrai apskaičiuojami taip: FFT, 24 nevienalyčių trikampių filtrų (Mel skalės) energijos apskaičiavimas, logaritmo taikymas energijos vėrtėms ir atvirkštinė diskrečioji kosinuso transformacija (DCT) [16].

MFCC gali būti apskaičiuotas pagal formulę [17]:

$$c_n = \sum_{k=1}^K (\log S_k) \cos \left[n(k - 0.5) \frac{\pi}{K} \right]; n = 0, 1, 2, \dots, C - 1 \quad (2)$$

čia S_k - filtrų banko k -tojo filtro išėjimo galia, o C – MFCC koeficientų skaičius (klasikinės MFCC sistemos dažniausiai turi 8-13 koeficientų).

1.3. Fonokardiogramų klasifikavimo metodai

Atlikus požymių išskyrimą sekantis žingsnis yra išsirinkti tinkamą DI metodą ir parametrus, nuo šio žingsnio priklauso didžioji projekto dalis ir gauti rezultatai. Kad būtų lengviau palyginti rezultatus, šioje analizės dalyje formuojama lentelė, kurioje pateiki tyrimai, juose naudojamos technologijos ir pasiekti rezultatai (1.2 lentelė).

1.2 lentelė. PCG klasifikavimo tyrimų analizė.

Šaltinis	Duomenų rinkinys validacijai	Požymių išskyrimo metodai	Klasifikavimo metodai	Naudoti parametrai	Tikslumas (mAcc)
Krishnan et al., [11], 2020	„Physionet Challenge 2016“	3-jų sluoksnių vienmatis konvoliucinis neuroninis tinklas (1D-CNN)	Feedforward Neural Network (FNN)	Adam optimizatorius; learning rate (lr) = 0,001; moment, $\beta_1 = 0,9$; $\beta_2 = 0,999$	0,8574
Alkhodari et al., [2], 2021	„Physionet Challenge 2016“	3-jų sluoksnių 1D-CNN	dvikryptis Long Short-Term Memory (BiLSTM) rekurentinis neuroninis tinklas (RNN)	$lr = 0.01$; L2-regularization = 0.0001	0,8731
Zabihi et al., [18], 2016	„Physionet Challenge 2016“	Požymiai išskaičiuoti iš 5 skirtingų garso signalų apdorojimo metodų	FNN	tanh activation, Levenberg—Marquardt optimizatorius	0,8590
Shuvo et al., [12], 2021	„Physionet Challenge 2016“	4-ių sluoksnių 1D-CNN	CardioXNet (DNN)	$lr = 0.00001$; Adam optimizatorius	0,8657
Demir et al., [13], 2019	PASCAL	iš anksto išmokyti VGG16, VGG19, AlexNet	SVM	-	0,8
Chen et al., [19], 2020	„Physionet Challenge 2016“	modified frequency slice wavelet transform (MFSWT), Sample entropy (SampEn)	2 CNN modeliai	$lr = 0.001$	0,9391

Iš analizės (1.2 lentelė) matome, jog „Physionet Challenge 2016“ duomenų rinkinys yra gan dažnas pasirinkimas šios srities moksliniuose darbuose. Požymių išskyrimui daugiausiai naudojami 2 metodai:

- Automatinis išskyrimas naudojant CNN arba jo pagrindu kurtus algoritmus;
- Skaičiuojant statistines, matematinės, fizikines garsų reikšmes (Fourier transformacija, MFSWT) gaunamas pagal tam tikras formules ir jas naudojant kaip požymius.

Žiūrint į klasifikavimo metodus matome, jog ganėtinai sėkmingai užduotį atlieka DNN su įvairiomis struktūromis bei autoriai mini, jog tokiu atveju požymių išskyrimas itin paprastas ir greitas, kadangi galima tiesiog panaudoti viena ar kelis CNN sluoksnius. Jei ieškotume paprastesnio ir mažiau išteklių reikalaujančio modelio, tai dažniausiai pastebimas buvo atraminių vektorių klasifikatorius (SVM).

1.4. Problemos esamuose sprendimuose

Norint atlikti unikalų tyrimą reikia pateikti inovatyvų sprendimą ar patobulinimą esantiems sprendimams. Vienas iš svarbiausių literatūros analizės tikslų yra atrasti kuo galima patobulinti esamus sprendimus. Dauguma autorių nagrinėtuose tyrimuose prie išvadų pateikia sprendimo silpnąsias savybes ir kuo būtų galima patobulinti aprašytą sprendimą.

Chen et al., [5], 2021 straipsnyje, kuriame atliekama PCG klasifikavimo sprendimų / metodų sisteminga apžvalga išskiriamos trys pagrindinės problemos esamuose sprendimuose:

- Mokymas su ribotu kiekiu širdies garso duomenų;

- Mokymo efektyvumas ir sparta;
- Giliesiems dirbtiniams neuroniniams tinklams mokyti reikalingi atminties ir skaičiavimo ištekliai.

1.4.1. Mokymas su ribotu kiekiu širdies garso duomenų

Reikia daug laiko ir darbo sąnaudų, kad būtų gauta daug širdies garso mėginių, ypač tam tikro tipo anomalijoms nustatyti. Todėl klasifikacija dažnai atliekama tik dvejoms klasėms (normalus ir sutrikęs širdies dažnis) ir sunku sukurti stiprų sprendimą, kuris galėtų tiksliai klasifikuoti kokias anomalijas nustatytas sutrikęs širdies dažnis priklauso. Dauguma nagrinėtų straipsnių [11, 2, 12, 20, 18, 19] susiduria su šia problema ir negali detaliai klasifikuoti širdies sutrikimų anomalijų.

Galimi sprendimai:

- Duomenų augmentacija;
- Algoritmų modifikavimas taikant skirtingus svorius sąnaudų funkcijai pagal mokymo duomenų pasiskirstymą (daugiau dėmesio scenarijams su mažesniu duomenų kiekiu).
- Generative Adversary Networks (GANs) sintetinių duomenų generavimui.

1.4.2. Mokymo efektyvumas ir sparta

Dažnai neefektyvu mokyti DNN nuo nulio atsitiktinai parenkant parametrus. Vis dar dideliu iššūkiu išlieka gilus mokymosi modelio mokymas automatiškai parenkant parametrus. Vienas iš galimų sprendimų – perkeliamasis mokymas (angl. *transfer learning*). Perkeliamasis mokymas suteikia galimybę pasitelkti kitai problemai išmokytą modelį ir pritaikyti jį savo problemai, kas ženkliai sumažina mokymo laiką ir resursus.

1.4.3. DNN reikalaujami ištekliai

Pastebėta tendencija, jog gilesni DNN modeliai su daugiau sluoksnių pasižymi didesniu tikslumu. Pavyzdžiui modifikuotas AlexNet tinklas [5, 21] su 35 konvoliuciniais sluoksniais, modifikuotas VGGNet tinklas [5, 22] su 16 konvoliucinių sluoksnių ir modifikuotas InceptionResNet tinklas [5] su 138 konvoliuciniais sluoksniais. Tačiau šių modelių reikalavimai taip pat didesni. Tokių modelių valdymas ir kūrimas yra sudėtingesnis, reikalauja daugiau operatyviosios atminties ir kitų kompiuterio išteklių.

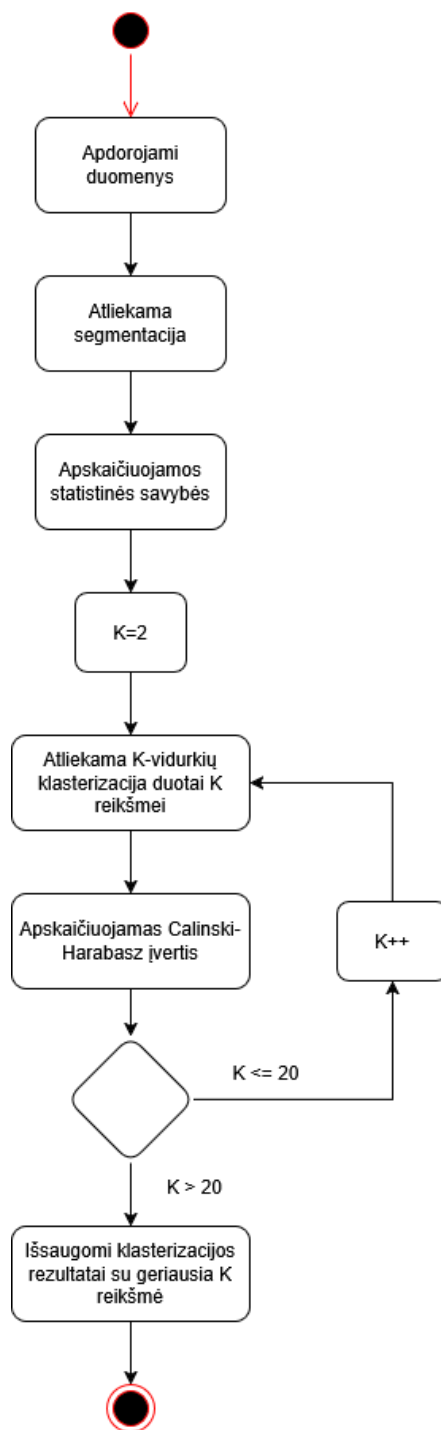
Vienas iš būdų išspręsti šią problemą yra perteklinių svorių mažinimas, taip sumažinant reikalaujamą operatyvios atminties ir kompiuterio išteklių kiekį. Pagrindiniai dabartiniai DNN modelių svorių mažinimo metodai: parametrų atpjovimas ir dalijimasis (angl. *parameter pruning and sharing*); mažo rango faktorių dekompozicija (angl. *low-rank factor decomposition*); žinių distiliavimas (angl. *knowledge distillation*); retas reguliavimas (angl. *sparse regularization*) ir kaukės pagreitinimas (angl. *mask acceleration*) [5, 23].

2. Širdies plakimo garso įrašų eksperimentų eiga

Šiame skyriuje pateikiama fonokardiogramų eksperimentų eiga. Atliekami eksperimentai naudojant skirtingus klasifikavimo metodus bei skirtingo tipo įvestis. Pirmiausia atliekami eksperimentai naudojant Springer segmentacijos procesą [24], kur naudojamos segmentacijos procesu gautos įvestys. Naudojant segmentacijos procesu gautas įvestis atliekami ir duomenų klasterizavimo eksperimentai (žr. 4.2 skyrių). Skyriuje pateikiama eksperimentų eiga ir nenaudojant segmentacijos proceso, kur vietoje segmentacijos procese gautų įvesčių naudojamos 20 MFCC įvestys. Taip pat pateikiama sintetinių duomenų generavimo proceso eiga.

2.1. Klasterizavimo eksperimentų eiga

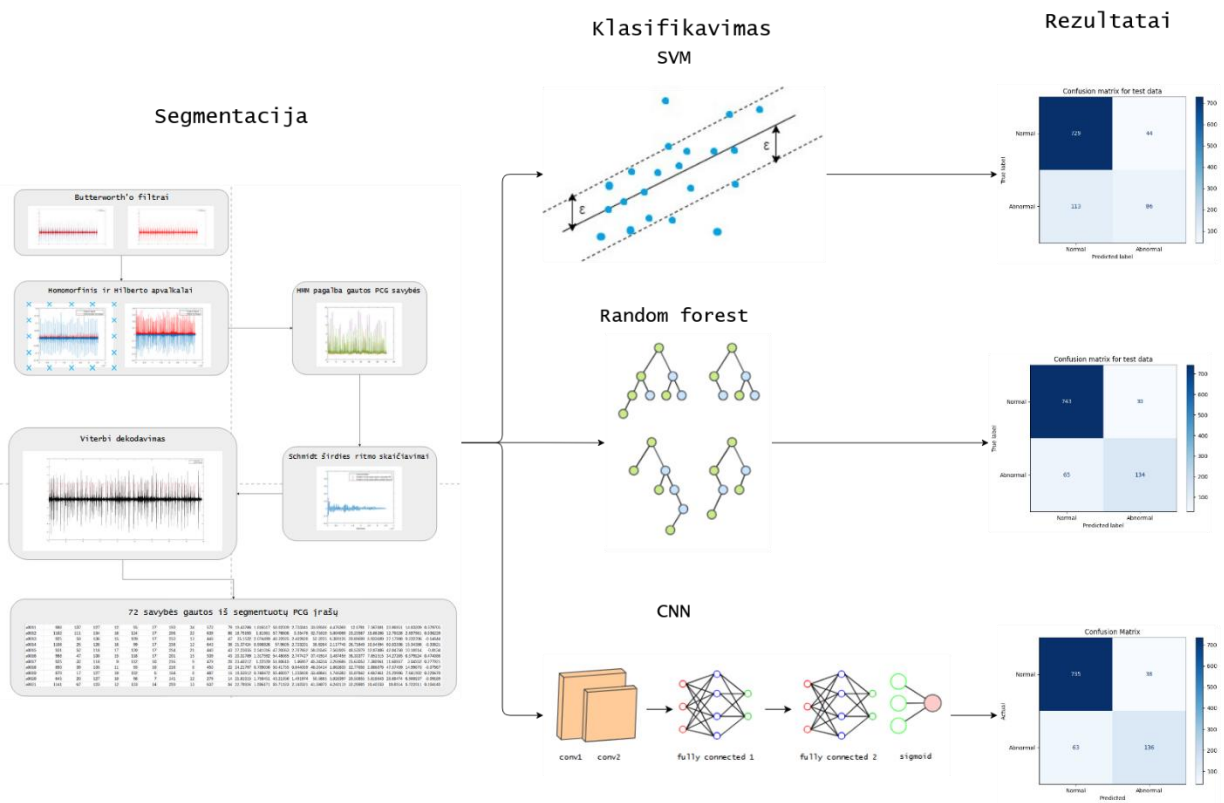
Šiame tyrime stengiamasi giliau pažvelgti į turimus duomenis ir klases. K-vidurkių metodo (plačiau žr. 4.2.2 skyriuje) pagalba randami panašumai tarp skirtingų širdies plakimo garso įrašų ir garsai priskiriami atitinkamai klasei bei atrandama kiek iš viso tokių klasių gali būti pasitelkiant Calinski-Harabasz įvertį (plačiau 4.2.1). Pirmiausia atliekamas duomenų apdorojimas ir segmentacija, gaunamos garso įrašų savybės. Tuomet atliekamas K-vidurkių klasterizavimas vis parenkant skirtingą K reikšmę ($1 < K < 21$) ir apskaičiuojamas Calinski-Harabasz įvertis. Išsaugomas rezultatas turintis aukščiausią Calinski-Harabasz įvertį (optimaliausia K reikšmė) ir gaunama K skirtingų klasterių su skirtingomis fonokardiogramomis (žr. 2 pav.).



2 pav. Klasterizavimo eksperimentų veiklų diagrama

2.2. Klasifikavimo eksperimentų eiga su segmentacija

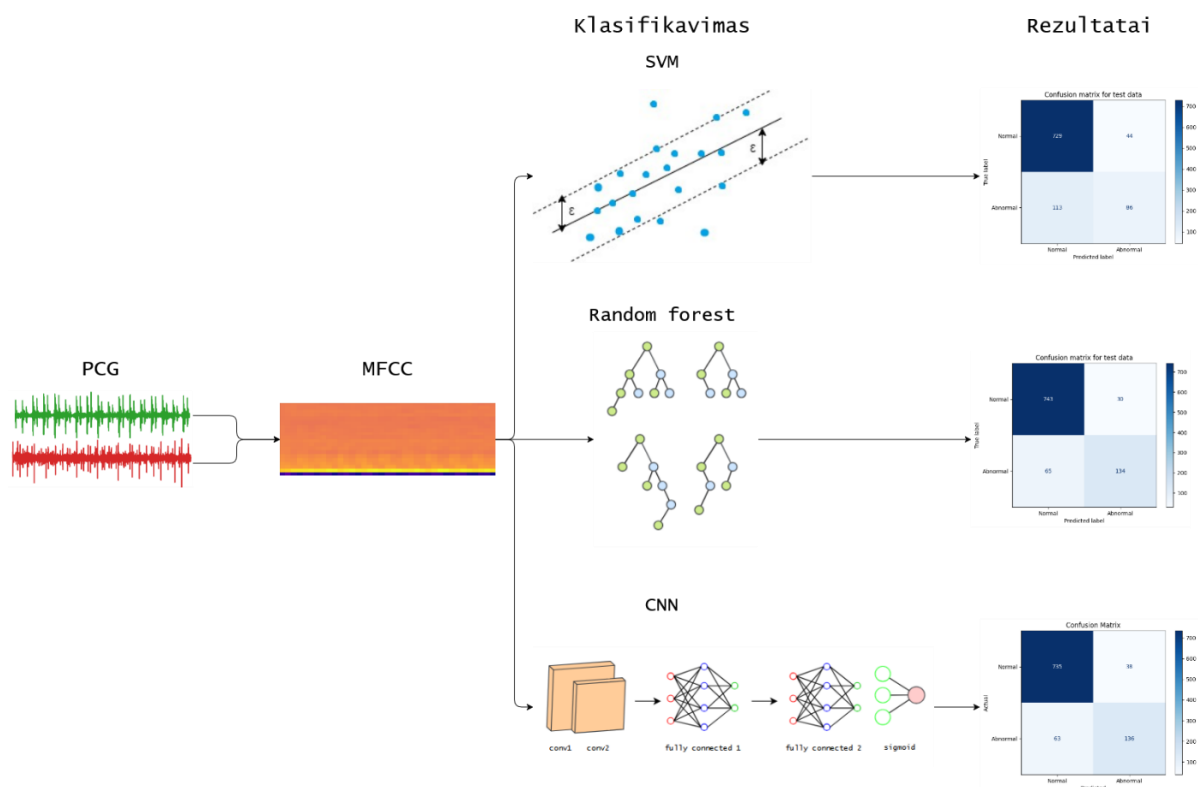
Atliekant klasifikavimo eksperimentus veiksmai taip pat prasideda nuo duomenų apdorojimo ir segmentacijos (plačiau apie segmentaciją 3.3 skyriuje). Tuomet atliekant įvairias statistines operacijas su įrašų segmentais (vidurkiai, standartiniai nuokrypiai tarp skirtingų segmentų, pasikartojimų skirtumai) gautos 72 savybės yra naudojamos kaip įvestys eksperimentams su 3 skirtingais mašininio mokymosi modeliais: SVM, Atsitiktinių miškų bei CNN. 3 pav.eikslėlyje pavaizduota bendra tyrimo schema, detaliau etapai paaiškinami tolimesniuose skyriuose.



3 pav. Klasifikavimo eksperimentų veiksmų seka naudojant segmentaciją

2.3. Klasifikavimo eksperimentų eiga be segmentacijos

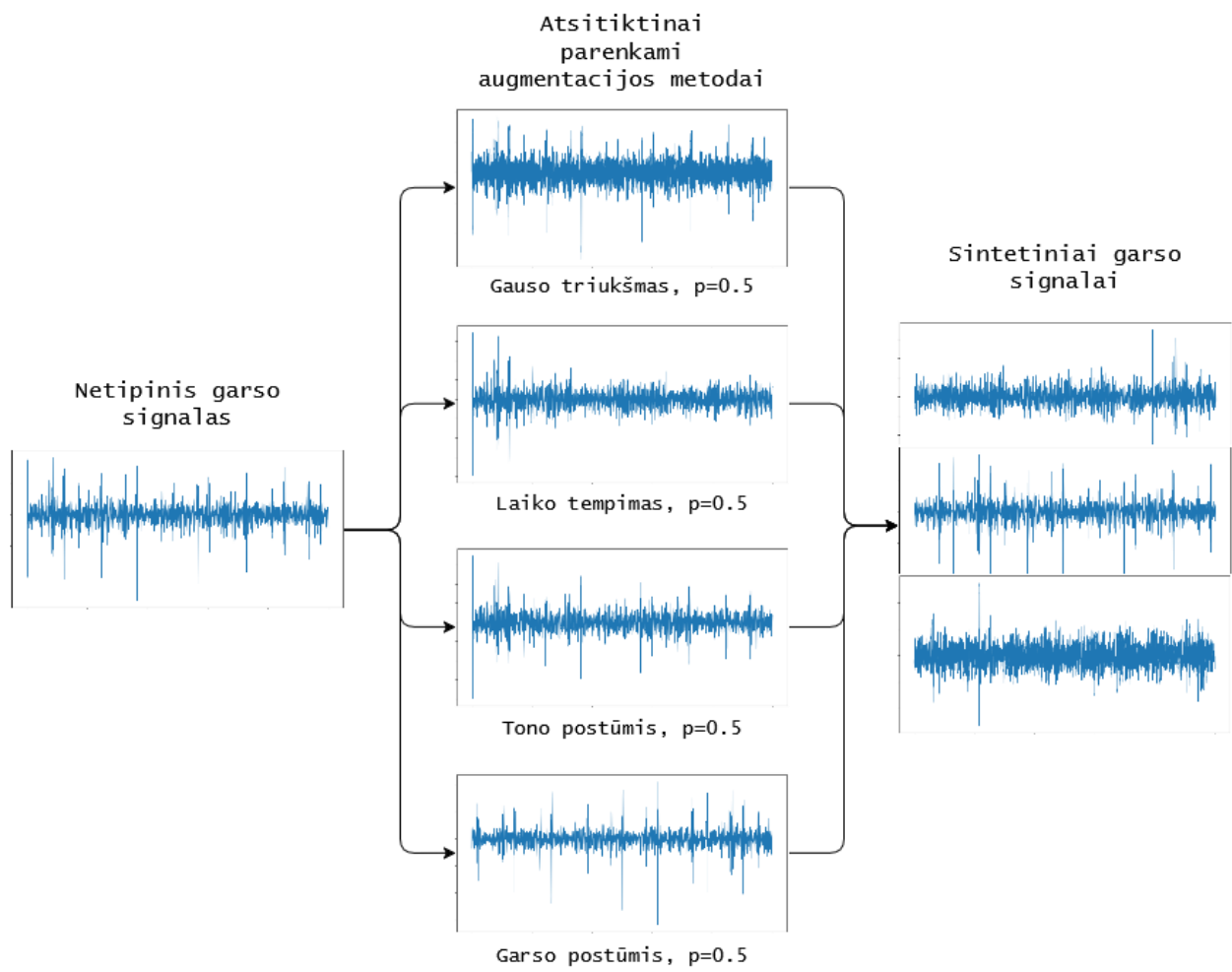
Nenaudojant širdies plakimo garsų segmentacijos proceso, klasifikavimo eksperimentų eiga šiek tiek pasikeičia (žr. 4 pav.). Pirmiausia nuskaitomi visi fonokardiogramų duomenys ir kiekvienam garso įrašui apskaičiuojamos 20 MFCC vertės. Toliau šios 20 MFCC savybių naudojamos kaip įvestys SVM, Atsitiktinių miškų ir CNN klasifikatoriams, kurie yra apmokomi klasifikuoti garso įrašus į tipinius ir netipinius. Tuomet modelio mokymo įvertinimui naudojamas testavimo duomenų rinkinys ir gaunamos klasifikavimo rezultatų klaidų matricos iš kurių galima apskaičiuoti ir visas kitas modelių įvertinimo metrikas (specifiškumą, jautrumą, tikslumą). Gavus rezultatus naudojant pirminį duomenų rinkinį, atliekama duomenų rinkinio augmentacija (plačiau 2.4 skyriuje) ir atliekama ta pati 4 pav. pavaizduota veiksmų seka naudojant naują duomenų rinkinį papildytą sintetiniais duomenimis.



4 pav. Klasifikavimo eksperimentų veiksmų seka nenaudojant segmentacijos

2.4. Sintetinių duomenų generavimo eiga

Duomenų disbalansui panaikinti naudojami duomenų augmentacijos metodai ir duomenų rinkinys papildomas sintetiniais netipiniais duomenimis. Pirmiausia netipiniai garso įrašai yra atskiriami nuo pilno duomenų rinkinio. Garso įrašų augmentacijai naudojami 4 skirtingi augmentacijos metodai: Gauso triukšmas (angl. *Gaussian noise*), laiko tempimas (angl. *time stretch*), tono postūmis (angl. *pitch shift*), garso postūmis (angl. *shift*). Kiekvieną kartą atliekant augmentaciją yra atsitiktinai parenkami augmentacijos metodai, kiekvienas augmentacijos metodas turi vienodą tikimybę būti parinktas, visada yra parenkamas bent vienas metodas, tačiau nebūtinai yra parenkami visi 4. Augmentacijos metodų parametrai taip pat parenkami atsitiktinai, tačiau nurodomi režiai. Gauso triukšmo amplitudės dydis parenkamas nuo 0.001 iki 0.005. Laiko tempimas sulėtina ar pagreitina signalą nuo 0,97 (sulėtina 3%) iki 1,03 (pagreitina 3%) pradinio greičio. Tono postūmis pakeičia toną nuo -0,5 (nuleidžia) iki 0,5 (pakelia) pustonių. Garso postūmis paslenka garsą nuo -0,3 (30% atgal) iki 0,3 (30% į priekį). Kiekvienam netipiniam garso signalui augmentacija atliekama 3 kartus ir gaunami 3 papildomi sintetiniai netipiniai garso signalai. Visa veiksmų seka pavaizduota 5 pav.

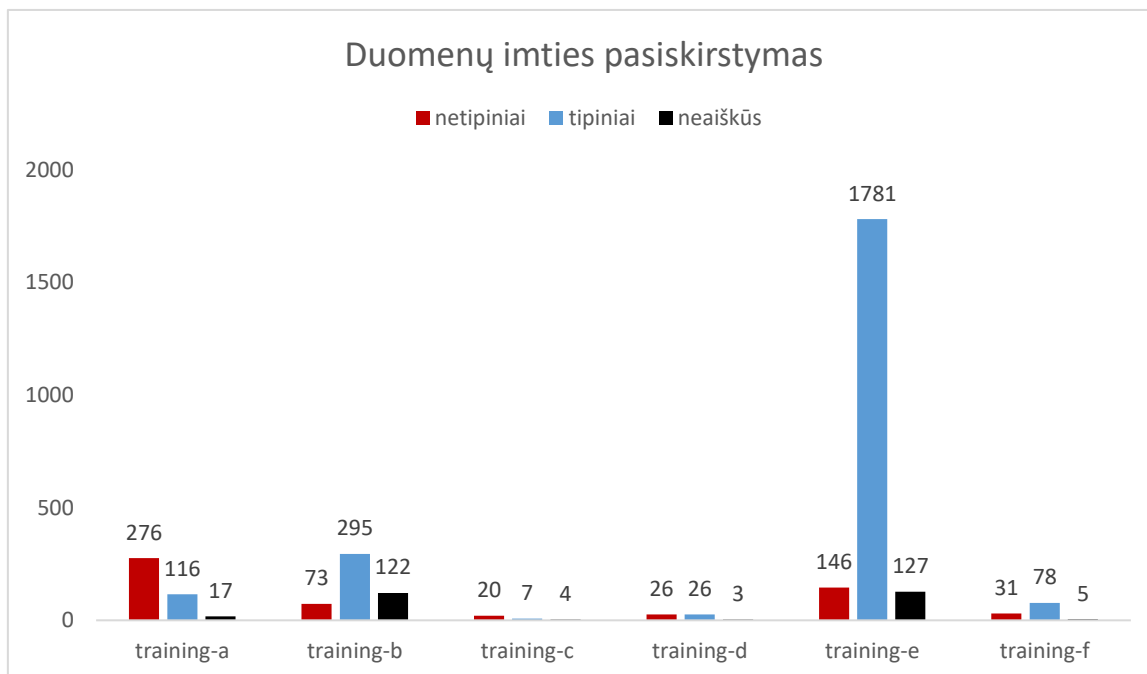


5 pav. Duomenų augmentacijos veiksmų seka

3. Širdies plakimo garso įrašų apdorojimas ir segmentacija

3.1. Duomenų rinkinys

Eksperimentams naudojamas „2016 CinC/PhysioNet Challenge“ [7] duomenų rinkinys. Duomenų rinkinys sudarytas naudojant 6 skirtingas duomenų bazines. Iš viso duomenų rinkinyje yra 3153 įrašai, iš kurių 576 yra netipiniai širdies įrašai, 2303 tipiniai širdies įrašai ir 278 neaiškūs (žr. 6 pav.).



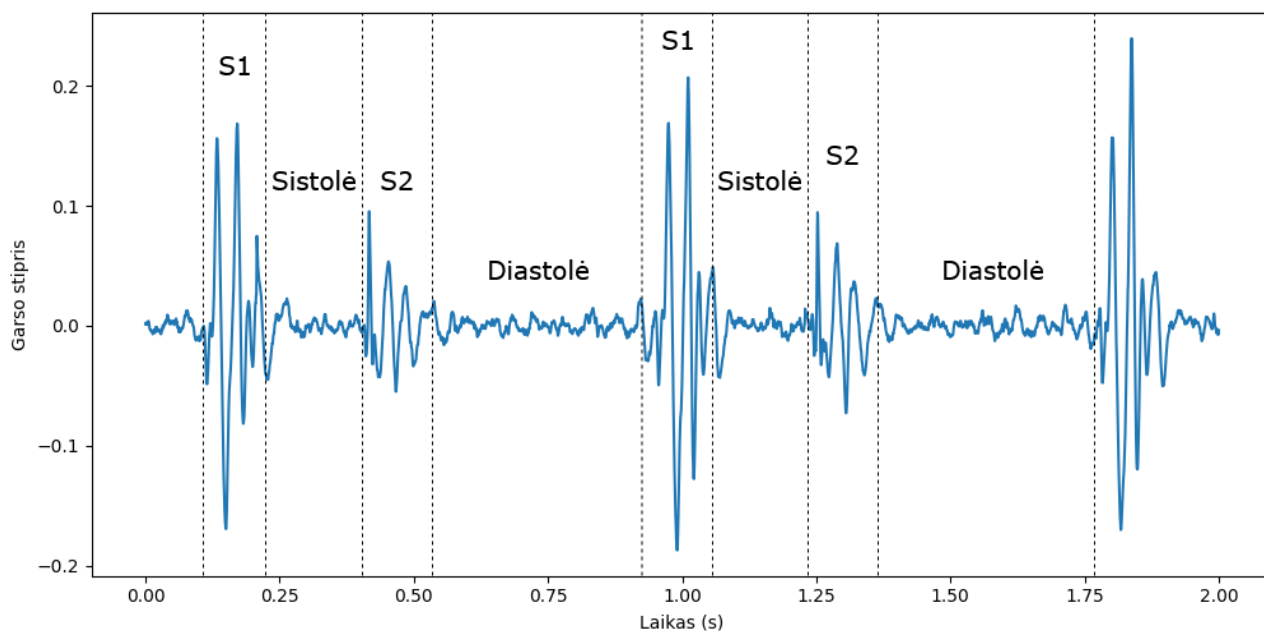
6 pav. Duomenų pasiskirstymo diagrama.

training-a – *training-f* nusako duomenis iš skirtingų duomenų bazių, kurie buvo surinkti skirtingų pasaulio universitetų. Garso įrašai duomenų rinkinyje varijuoja nuo 5 iki 120 sekundžių. Validacijai naudojamas duomenų rinkinys sudarytas iš specialiai atrinktų kokybiškų širdies plakimo garso įrašų. Validacijos duomenų rinkinį sudaro 301 įrašai iš kurių 150 tipiniai ir 151 netipiniai. „Physionet“ iššūkio pateiktas validacijos duomenų rinkinys yra sudarytas iš mokymo duomenų, visi validacijos įrašai egzistuoja ir mokymo imtyje. Kadangi „Physionet“ nepateikia testavimo duomenų rinkinio, siekiant tiksliau ištestuoti modelius „Physionet“ iššūkio pateikti mokymo duomenys buvo suskirstyti į mokymo (70%) ir testavimo (30%). Naujai gautą mokymo duomenų rinkinį sudaro 2268 įrašai, o testavimo rinkinį sudaro 972 įrašai.

3.2. Širdies plakimo garso įrašo struktūra

Fonokardiogramose yra atpažįstami 2 pagrindiniai garsai (angl. *fundamental heart sounds*, FHS) S1 ir S2 (žr. 7 pav.). S1 įvyksta izovolumetrinio skilvelių susitraukimo (sistolės) pradžioje, kai mitralinis ir triburis vožtuvai užsidaro ir kraujas yra išstumiamas. Tokie struktūriniai pasikeitimai širdyje sukelia vibracijas, kurios girdimos krūtinės ląstoje. S2 garsas įvyksta širdies atsipalaidavimo (diastolės) metu kai aortos ir plaučių vožtuvai užsidaro ir širdis prisipildo krauju. Be šių įprastai aptinkamų širdies garso signalų, kurie neįtakoja širdies patologijų, aptinkami ir kiti širdies garsai sistolės ir diastolės metu. Su širdies ligomis dažniausiai yra asocijuojami papildomi garsai diastolės metu, kurie yra išskiriami kaip S3 ir S4. Širdies garsas S3 susijęs su padidėjusiu kairiojo prieširdžio

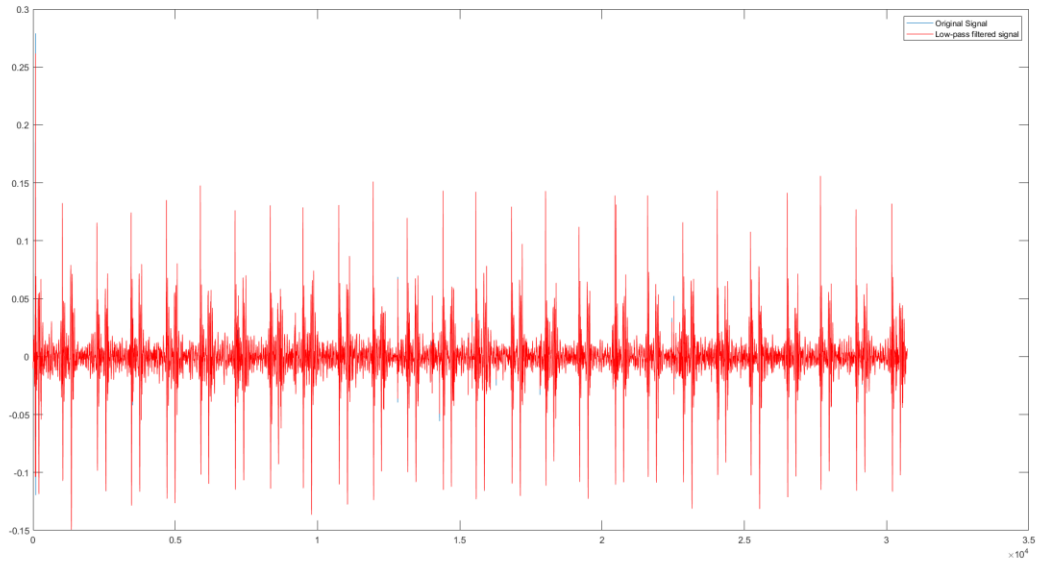
tūriu ir (arba) padidėjusiu skilvelio prisipildymo slėgiu. Kai kuriems vaikams ir sportininkams šis garsas gali būti fiziologinis. S4 garsas susidaro, kai žmogaus skilvelis tampa mažiau elastingas. Kai prieširdžiai susitraukia vėlyvos diastolės metu ir spaudžia sustandėjusį skilvelį, jiems tenka generuoti didesnę jėgą, todėl kraujo tėkmė tampa sūkurinė. Tai būdinga ligoms, dėl kurių sumažėja skilvelių paslankumas, pavyzdžiui, kairiojo skilvelio hipertrofijai. [25]



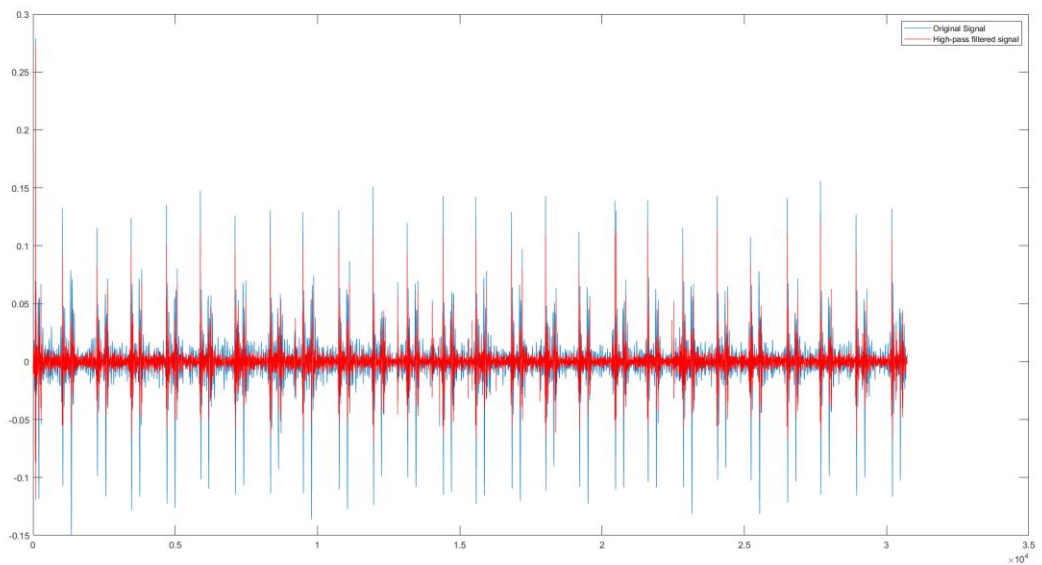
7 pav. Tipinio širdies plakimo garso įrašo struktūra

3.3. Duomenų apdorojimas atliekant segmentaciją

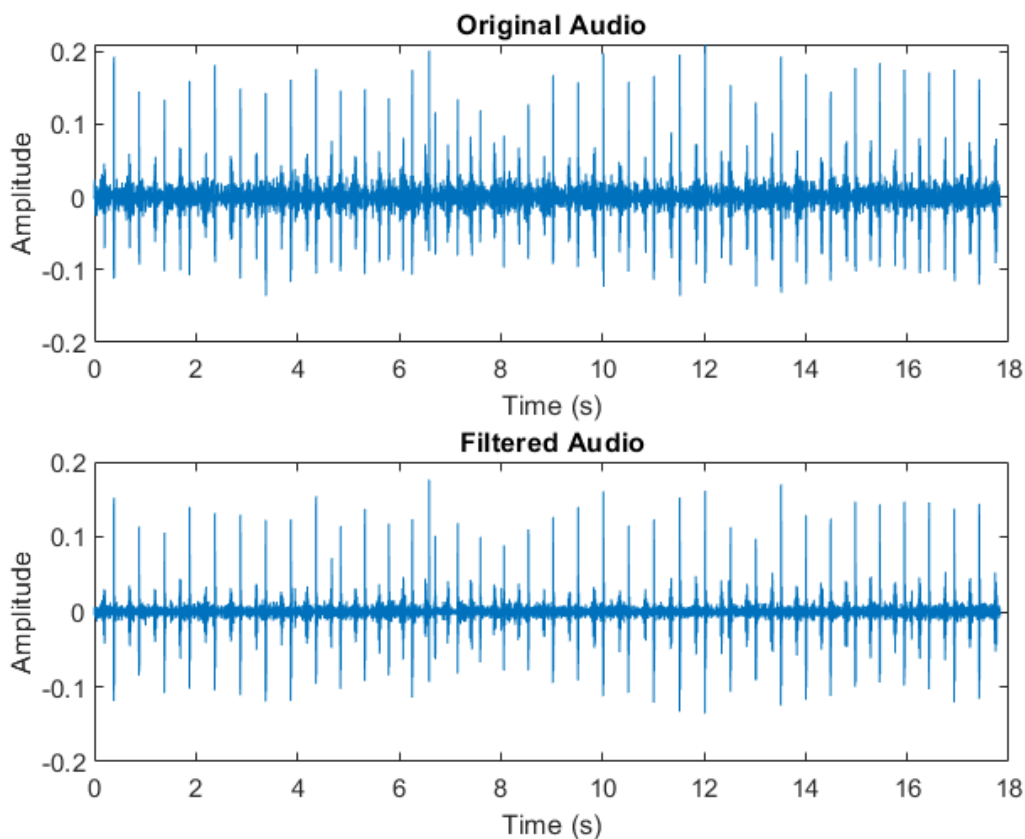
Nuskaičius garso įrašą yra gaunami 2 kintamieji: duomenų masyvas, kuriame laikomos garso dažnio reikšmės kiekvienam laiko vienetui; mėginių ėmimo dažnis (angl. *sampling frequency*), kuris nusako dažnių kiekį vienoje sekundėje. Duomenų rinkinyje garso įrašai yra skirtingos trukmės. Garso įrašų trukmė varijuoja nuo 5 iki 120 sekundžių. Naudojant segmentaciją klasifikavimas atliekamas iš trumpo garso įrašo (iki 60 sekundžių). Pirmas žingsnis yra sutraukti įrašus, kurie yra ilgesni nei 60 sekundžių. Surinkti garso įrašai taip pat neretai turi triukšmo, kuris trukdo tikslią širdies plakimo garso įrašų klasifikaciją. Įprasti triukšmo šaltiniai širdies plakimo garso įrašuose yra aplinkos kalba, judesiai ir fiziologiniai garsai, pavyzdžiui, žarnyno ir kvėpavimo garsai [24]. Šiems triukšmo garsams naikinti ar bent sumažinti yra naudojami Butterworth'o aukštų ir žemų dažnių filtrai (žr. 8 pav., 9 pav. ir 10 pav.), Schmidt'o šuolių naikinimas [26]. Filtrams naudojami Springer segmentacijos sprendime [24] nurodyti parametrai: Butterworth'o žemų dažnių filtrui naudojamas 400 Hz ribinis dažnis, Butterworth'o aukštų dažnių filtrui naudojamas 25 Hz ribinis dažnis, abiejų Butterworth'o filtrų filtro eilė, kuri nusako kaip staigiai filtras nukerta signalą už ribinio dažnio – 2.



8 pav. Butterworth'o žemų dažnių filtravimas



9 pav. Butterworth'o aukštų dažnių filtravimas.



10 pav. Originalaus ir filtruoto garso įrašų palyginimas.

3.4. Segmentacijos procesas

Apdorojus duomenis dar vienas svarbus etapas prieš klasifikuojant širdies plakimo garso įrašus yra segmentacija. Segmentacijos proceso eiga (žr. 11 pav.) pradedama nuo duomenų nuskaitymo ir apdorojimo. Pritaikomi jau minėti Butterworth'o aukštų ir žemų dažnių filtrai triukšmui sumažinti. Toliau naudojant Homomorfinį (plačiau 3.5 skyriuje) ir Hilberto (plačiau 3.6 skyriuje) apvalkalus išskiriamos savybės. Homomorfinis apvalkalas parodo kaip garsas stiprėja ir silpnėja bėgant laikui bei ignoroja staigias ir smulkias garso bangas, kas leidžia lengviau identifikuoti ir izoliuoti širdies plakimo garsus įrašė. Hilberto apvalkalas išskiria garso stiprumą / energijos lygį laike, kas leidžia lengviau atpažinti garso šuolius (širdies garse tai gali padėti identifikuoti S1 ir S2 garsus). Tuomet naudojamas Paslėptasis Markovo modelis (angl. *Hidden Markov Model*, HMM), kurio pagalba gaunamos fonokardiogramų savybės. Gautos savybės naudojamos Schmidt širdies ritmo skaičiavimams. Atlikus skaičiavimus atliekamas Viterbi dekodavimas ir širdies plakimo garsas suskirstomas į S1, S2, S3 ir S4 garsus. Iš šių segmentuotų garso įrašų gaunamos 72 savybės kiekvienam garso įrašui, kurios toliau gali būti naudojamos garsų klasifikavimo uždaviniui. Visas segmentacijos procesas buvo atliekamas pasitelkiant atvirojo kodo sprendimą, kuris nurodomas kaip rekomendacija 2016 metų „Physionet“ iššūkyje. Detalesnė informacija apie segmentacijos eigą, modelių parametrus, gautas savybes pateikta Springer straipsnyje [24].



11 pav. Springer [24] segmentacijos proceso eiga

3.5. Homomorfinis apvalkalas (angl. *homomorphic envelope*)

Homomorfinis apvalkalas atliekant duomenų segmentaciją yra vienas iš filtrų leidžiantis išskirti svarbiausias garso įrašo savybes (žr. 12 pav.). Homomorfinis apvalkalas išgaunamas naudojant homomorfinį filtrą. Į signalą žiūrima kaip į lėtai kintančio komponento, kontroliuojančio signalo amplitudę, ir greitai kintančio komponento, vaizduojančio svyruojančią signalo dalį, sandaugą [26]:

$$x(t) = a(t) \cdot o(t) \quad a(t) > 0, \quad (3)$$

čia $a(t)$ yra amplitudės komponentė, o $o(t)$ - osciliacinė komponentė. Logaritminiu būdu transformuotas signalas yra amplitudės ir svyruojančios komponentės suma:

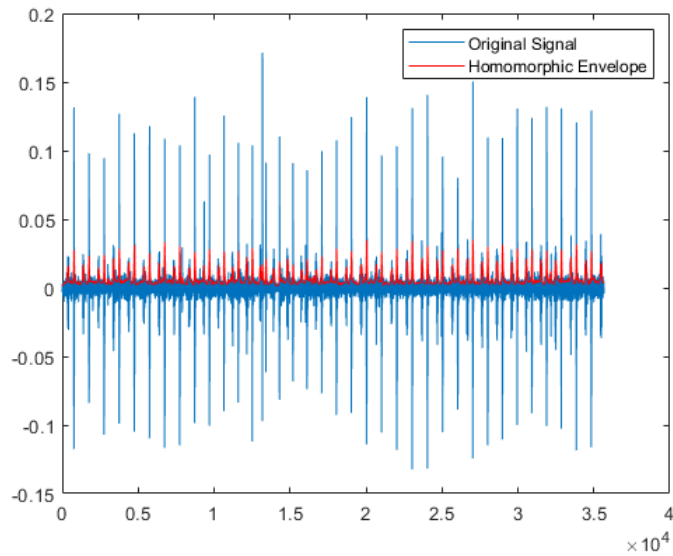
$$\ln|x(t)| = \ln|a(t)| + \ln|o(t)|, \quad (4)$$

Tokiu būdu daugyba pakeičiama sudėtimi ir galima atmesti aukšto dažnio tiesiniu žemųjų dažnių filtru:

$$L(\ln|x(t)|) = L(\ln|a(t)|) + L(\ln|o(t)|) \approx L(\ln|a(t)|), \quad (5)$$

čia L yra žemųjų dažnių filtras, šiuo atveju - pirmos eilės Butterworth'o filtras, kurio empirinis nustatytas ribinis dažnis yra 25 Hz. Homomorfinis apvalkalas sukuriamas taikant atvirkštinę transformaciją:

$$HEnv(t) = \exp(L(\ln|x(t)|)) \approx a(t), \quad (6)$$



12 pav. Homomorfinis apvalkalas pritaikytas vienam iš garso įrašų.

3.6. Hilberto apvalkalas (angl. *Hilbert envelope*)

Hilberto apvalkalas taip pat naudojamas kaip vienas iš filtrų garso įrašo savybėms išskirti (žr. 13 pav.). Hilberto transformacija grindžiamą apvalkalo išskyrimą dažniausiai galima suskirstyti į dvi šakas, t. y. (1) apvalkalas E yra kompleksinio analitinio signalo realiosios dalies decimuotas signalas ir (2) momentinis dažnis IF yra kompleksinio analitinio signalo vaizduojamosios dalies išvestinė. Tegul stetoskopu užfiksuotas širdies plakimo garso signalas $x(t)$ yra realiosios vertės signalas. Tuomet Hilberto transformacija [27]:

$$H[x(t)] = 1\pi \int_{-\infty}^{\infty} \frac{x(\tau)}{\tau-t} d\tau = x(t) \cdot \frac{1}{\pi t}, \quad (7)$$

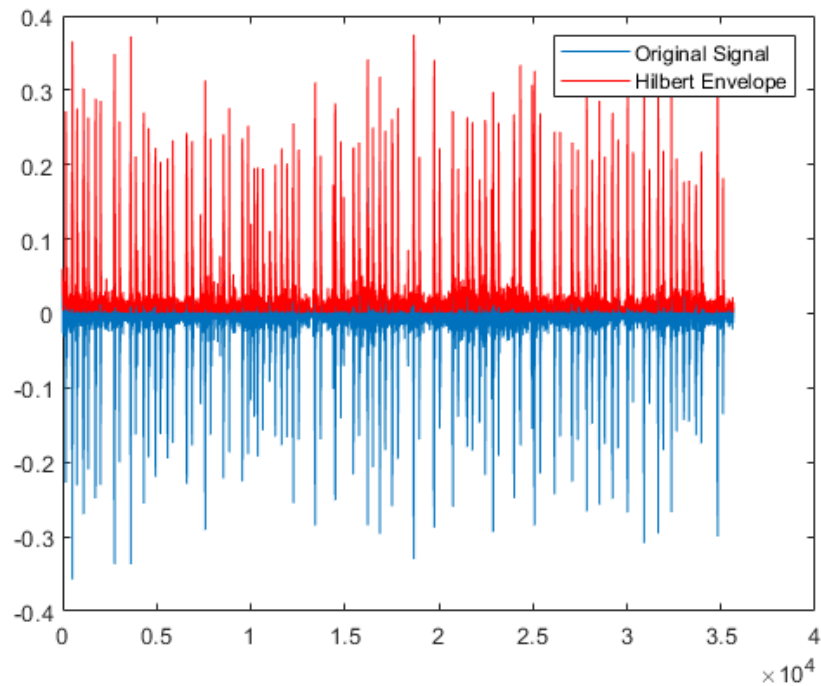
čia naudojama principinė integralo vertė, o „*“ reiškia konvoliucijos operatorių. Hilberto transformacija dažnai interpretuojama kaip 90° fazės keitiklis: $[H[x(t)]] = -x(t)$. Be to, iš duoto signalo $x(t)$ kompleksinį analitinį signalą $A[x(t)]$ galima išreikšti taip:

$$A[x(t)] = x(t) + jH[x(t)] = E(t)e^{j\varphi(t)}, \quad (8)$$

Apvalkas $E(t)$ apskaičiuojamas pagal (5), (6) lygtis. Tačiau apvalkas vis tiek gali greitai virpėti, nes jam būdingas labai aukštas imties ėmimo dažnis. Tada $E(t)$ perskaičiuojamas taikant slenkančio vidurkio filtrą, kurio lango ilgis yra 0,0145 s ir 0,013 s persidengimas. Taigi galutinis signalas apskaičiuojamas:

$$G(t) = \sum_{i=1}^W E(i) \cdot w(i), \quad (9)$$

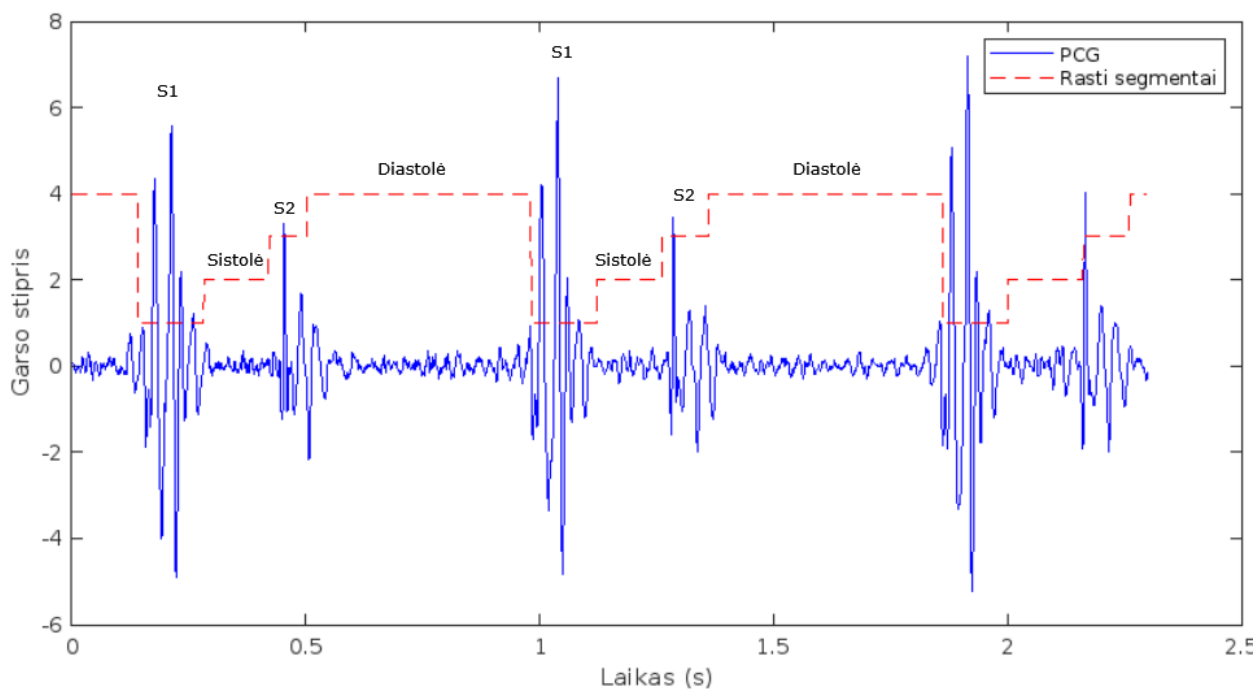
čia W - lango pavyzdžių skaičius, o w - hammingo langas. Signalas $G(t)$ vadinamas Hilberto apvalkalu.



13 pav. Hilberto apvalkalas pritaikytas vienam iš garso įrašų.

3.7. Segmentacijos rezultatai

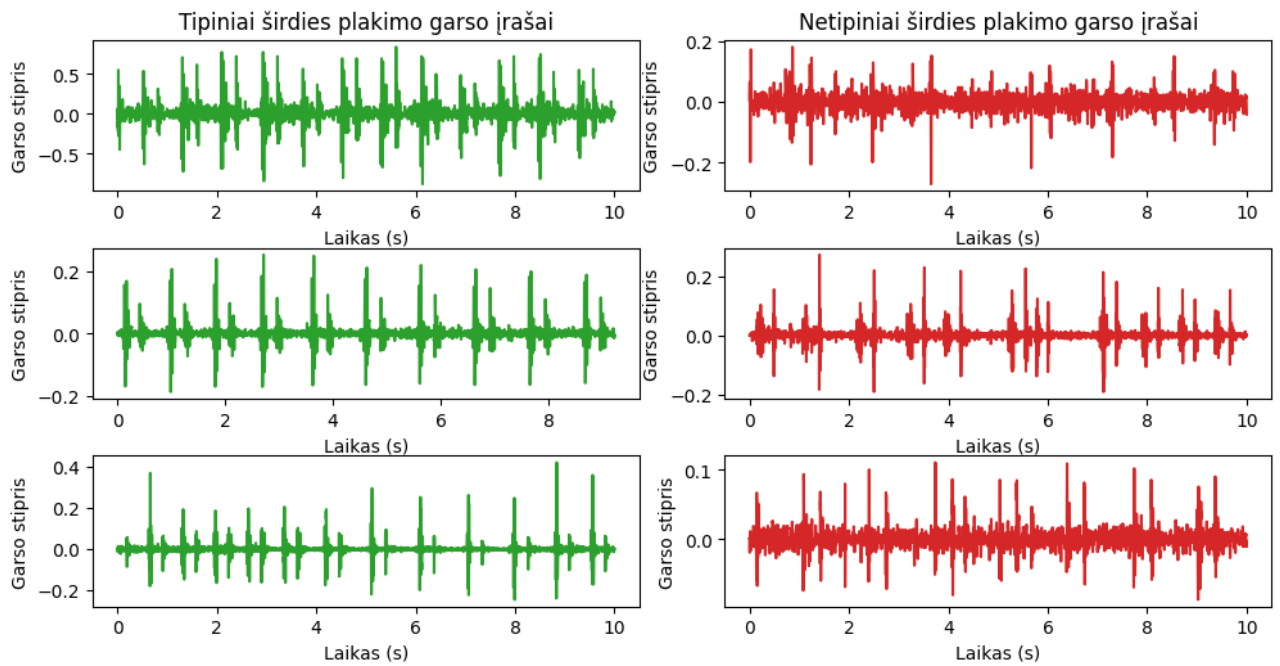
Atlikus segmentaciją ir atvaizdavus vieną pasirinktą garso įrašą, grafike galime pamatyti segmentacijos procese atpažintus skirtingus širdies plakimo garsus. 14 pav. paveikslėlyje pavaizduotas susegmentuotas širdies plakimo garso įrašas. Čia pasirinktas tas pats garso įrašas kaip ir 3.2 skyriuje, kur analizuojami fonokardiogramose atpažįstami garsai. Lyginant abu paveikslėlius galime matyti, kad išskirti segmentai beveik sutampa ir Springer [24] segmentacijos metodas gali gan tiksliai atpažinti skirtingus garsus tipinėje fonokardiogramoje.



14 pav. Segmentacijos procesu atpažinti širdies garsai

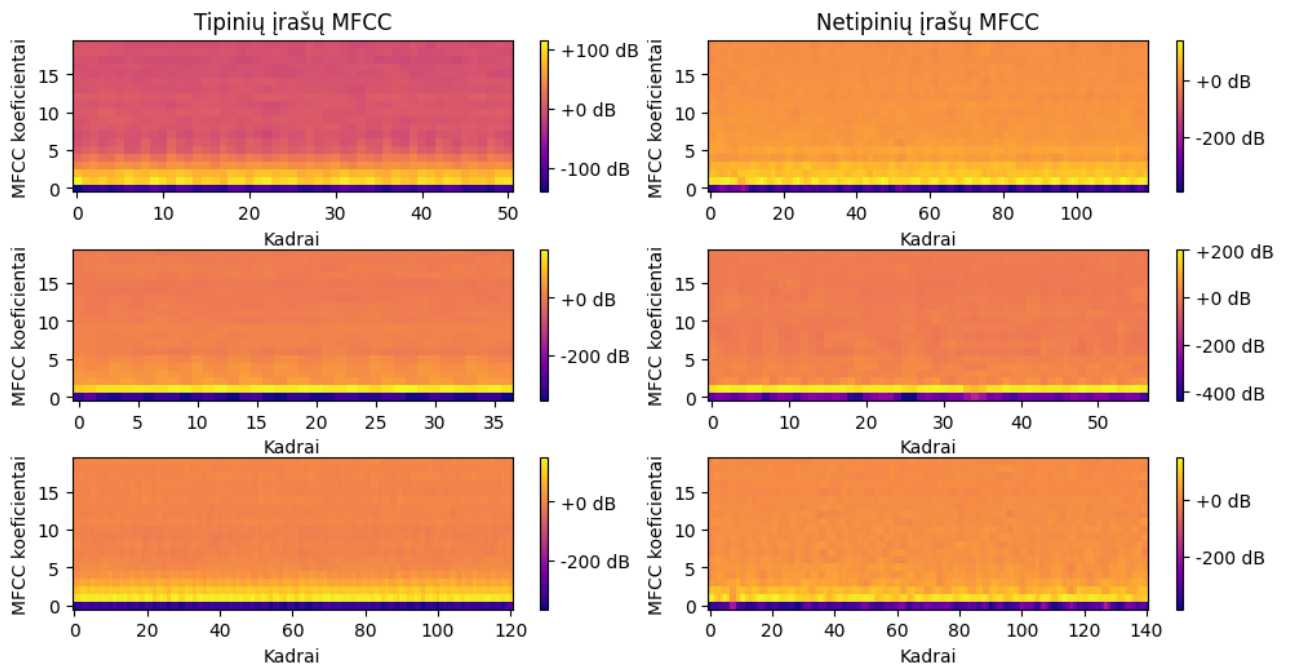
3.8. Duomenų apdorojimas nenaudojant segmentacijos ir MFCC požymių išskyrimas

Neatliekant garso įrašų segmentacijos, duomenų apdorojimo procesas gerokai paprastesnis. Čia siekiant išsaugoti kuo daugiau informacijos širdies garsai nėra karpomi ar filtruojami ir išlaikomi originalūs duomenų rinkinyje pateikti garsai. 15 pav. paveikslėlyje pavaizduoti 3-jų atsitiktinai parinktų tipinių ir 3-jų netipinių širdies garsų iš mokymo imties pirmosios 10 sekundžių.



15 pav. Tipiniai ir netipiniai širdies plakimo garso įrašai

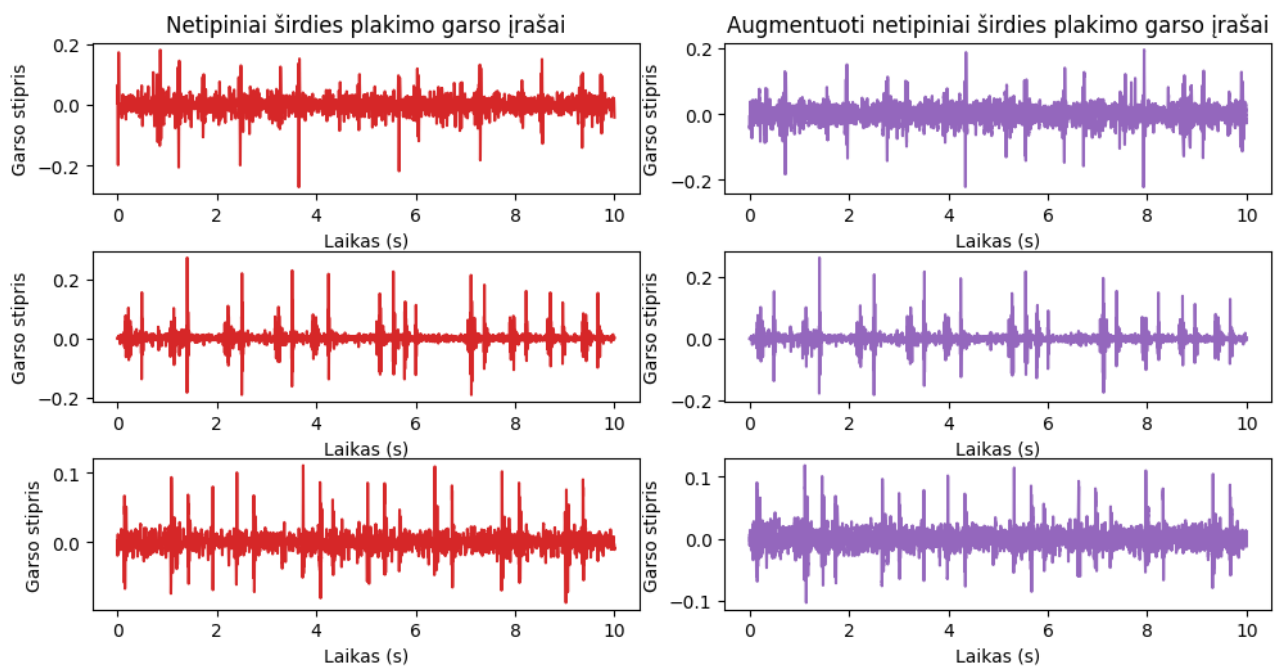
Mašininio mokymosi modeliams vis tiek reikia pateikti duomenis priimtiniu formatu ir savybių išskyrimas vis dar lieka svarbus žingsnis atliekant iš šių eksperimentus. Savybės čia išskiriamos pasitelkiant „Mel-dažnio Cepstrum“ koeficientus. Kiekvienam įrašui yra apskaičiuojama 20 MFCC koeficientų. Pasirinkta naudoti pirmuosius 20, nes tolimesni koeficientai gali lemti modelio persimokymą. Tų pačių atsitiktinai parinktų širdies plakimo garso įrašų MFCC koeficientai atvaizduojami 16 pav. paveikslėlyje. Kadangi garso įrašai nėra apkarpmi, o įrašų trukmė skiriasi, matome, kad ir MFCC kadrų skaičius įrašuose yra skirtingas. Daugumai mašininio mokymosi modelių skirtingų matmenų duomenų įvestis nėra priimtina. Šiai problemai pašalinti apskaičiuojami kadrų vidurkiai kiekvienam MFCC koeficientui. Galiausiai po savybių išskyrimo kiekvienas duomenų įrašas turi 20 savybių, kurios yra standartizuojamos ir pateikiamos kaip įvestis mašininio mokymosi modelių mokymui ir testavimui.



16 pav. Tipinių ir netipinių įrašų MFCC

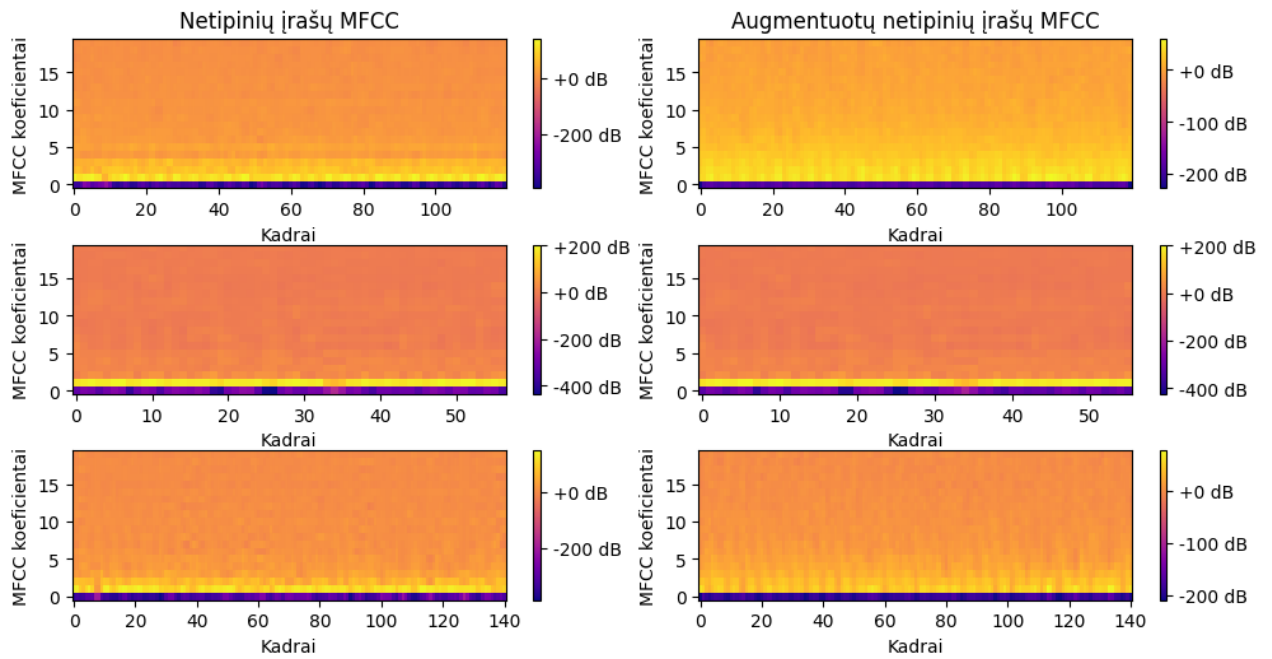
3.8.1. Augmentuoti netipiniai širdies plakimo garso įrašai

Atlikus duomenų augmentaciją ir gavus sintetinius duomenis galime juos palyginti su originaliais duomenimis. Žvelgiant į tų pačių, jau anksčiau apžvelgtų, netipinių garsų signalų ir atitinkamai po vieną iš jų sugeneruotų sintetinių signalų pirmas 10 sekundžių (žr. 17 pav.) matome, jog pirmasis signalas įgyjo šiek tiek triukšmo bei amplitudės šuoliai pasistūmėjo laiko prasme. Žvelgiant į antrąjį signalą plika akimi sunku pasakyti kuo garsas pasikeitė, tačiau galima išžvelgti minimalų amplitudės pokytį tarp 6 ir 7 sekundžių. O trečiajame signale galime pastebėti pakitusį triukšmingumą.



17 pav. Netipinių ir augmentuotų fonokardiogramų palyginimas

Žvelgiant į tų pačių signalų MFCC reikšmes (žr. 18 pav.) pirmajame signale matome aukštesnes 5-20 koeficientų reikšmes. Antrajame signale taip pat sunku išžvelgti skirtumų. O trečiajame matome žemesnes pirmojo koeficiento reikšmes.



18 pav. Netipinių ir augmentuotų netipinių įrašų MFCC

4. Eksperimentai

Atliekami skirtingi eksperimentai naudojant duomenų segmentaciją ir nenaudojant segmentacijos. Eksperimentams su segmentacija naudojamos iš atskirtų širdies plakimo garsų segmentų gautos statistinės savybės. Iš viso gaunamos 72 savybės, kurios yra naudojamos kaip įvestis mašininio mokymosi modeliams. Tuo tarpu nenaudojant segmentacijos įvestis gerokai skiriasi. Kiekvienam garso įrašui yra apskaičiuojami 20 MFCC ir MFCC kadru vidurkiai yra naudojami kaip įvestis.

4.1. Aplinka

Segmentacijos procesui ir klasterizavimo eksperimentams buvo naudojama MATLAB Online aplinka, versija R2023b. Visiems kitiems eksperimentams naudojama Google Colab Python 3 T4 GPU aplinka.

4.2. Klasterizavimo eksperimentai

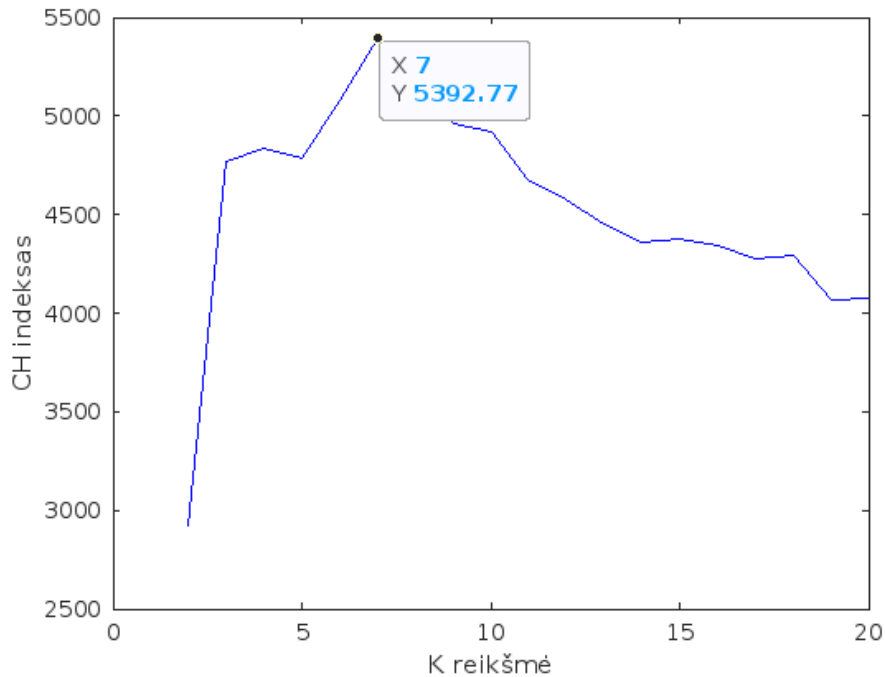
Atliekant klasterizavimą pirmas žingsnis yra rasti optimalų klasterių skaičių. Atlikus klasterizaciją galima analizuoti pagrindinius skirtumus tarp įrašų esančių skirtinguose klasteriuose ir pastebėti pagrindines tendencijas duomenų rinkinyje.

4.2.1. Calinski-Harabasz (CH) įvertis

Calinski-Harabasz (CH) metodas yra vienas sėkmingiausių būdų nustatyti optimalų klasterių skaičių klasterizavimo tyrimuose. Metodas pagrįstas „tarp klasterių sklaidos matricos“ (angl. *Between-Cluster Sum of Squares*, BCSM) ir „klasterio viduje sklaidos matricos“ (angl. *Within-Cluster Sum of Squares*, WCSM) santykiu [28]:

$$CH = \frac{\text{pėdsakas}(BCSM)}{\text{pėdsakas}(WCSM)} \cdot \frac{N-k}{k-1}, \quad (10)$$

BCSM pėdsakas yra atstumų tarp kiekvieno klasterio centro ir duomenų rinkinio centroido kvadratų suma, pasverta pagal klasterio dydį. WCSM pėdsakas yra atstumų tarp kiekvieno klasterio centro ir kiekvieno klasterio taško kvadratų suma. CH atskyrimui apskaičiuoti naudoja klasterių centrus, tačiau atskyrimas matuojamas pagal duomenų rinkinio centrą, o ne pagal konkrečius klasterius. Normalizavimo koeficientas $(N - k)/(k - 1)$ mažina rezultatą, kai k didėja.



19 pav. Calinski-Harabasz įverčių grafikas.

Kuo didesnė Calinski-Harabasz įverčio reikšmė, tuo labiau tinkama yra K reikšmė. Šiuo atveju nustatyta, jog geriausia K reikšmė yra 7 su 5392,77 CH indeksu (žr. 19 pav.).

4.2.2. K-vidurkių metodo aprašymas

Buvo naudojama MATLAB $idx = kmeans(X,k)$ funkcija [29]. MATLAB *kmeans* atlieka k-vidurkių klasterizavimą, siekiant suskirstyti $n \cdot p$ duomenų matricos X stebėjimus į k klasterių, ir grąžinamas $n \cdot 1$ vektorius (idx), kuriame yra kiekvieno stebėjimo klasterio indeksai. X eilutės atitinka taškus, o stulpeliai – kintamuosius. Pagal numatytuosius nustatymus *kmeans* naudoja kvadratinio Euklido atstumo metriką ir *k-means++* algoritmą klasterių centrams inicializuoti.

k-means++

k-means++ algoritmas naudoja euristinį metodą centroidų pradinėms reikšmėms rasti k-vidurkių klasterizacijai. Arthur ir Vassilvitskii [30] teigimu, *k-means++* pagerina Lloyd algoritmo veikimo laiką ir galutinio sprendinio kokybę.

Kai klasterių skaičius yra k, algoritmas *k-means++* parenka sėklas taip:

1. Iš duomenų aibės X tolygiai atsitiktinai parenkamas įrašas. Pasirinktas įrašas yra pirmasis centroidas ir žymimas c_1 .
2. Apskaičiuojamas atstumus nuo kiekvieno stebėjimo iki c_1 . Atstumas tarp c_j ir įrašo m žymimas $d(x_m, c_j)$.
3. Iš duomenų imties X atsitiktinai parenkamas sekantis centroidas, c_2 su tikimybe:

$$c_2 = \frac{d^2(x_m, c_1)}{\sum_{j=1}^n d^2(x_j, c_j)}, \quad (11)$$

4. Parenkamas centras j :

- a. Apskaičiuojamas atstumas nuo kiekvieno stebinio iki kiekvieno centroido ir kiekvienas įrašas priskiriamas kuriam nors centroidui.
- b. Iš duomenų imties X atsitiktinai parenkama reikšmė j su tikimybe:

$$j = \frac{d^2(x_m, c_p)}{\sum_{\{h; x_h \in c_p\}} d^2(x_h, c_p)}, \quad (12)$$

Čia C_p yra visų stebėjimų, artimiausių centroidui C_p , aibė, o x_m priklauso C_p .

Tai reiškia, kad kiekvienas sekantis centras parenkamas su tikimybe, proporcinga atstumui nuo savęs iki jau pasirinkto artimiausio centro.

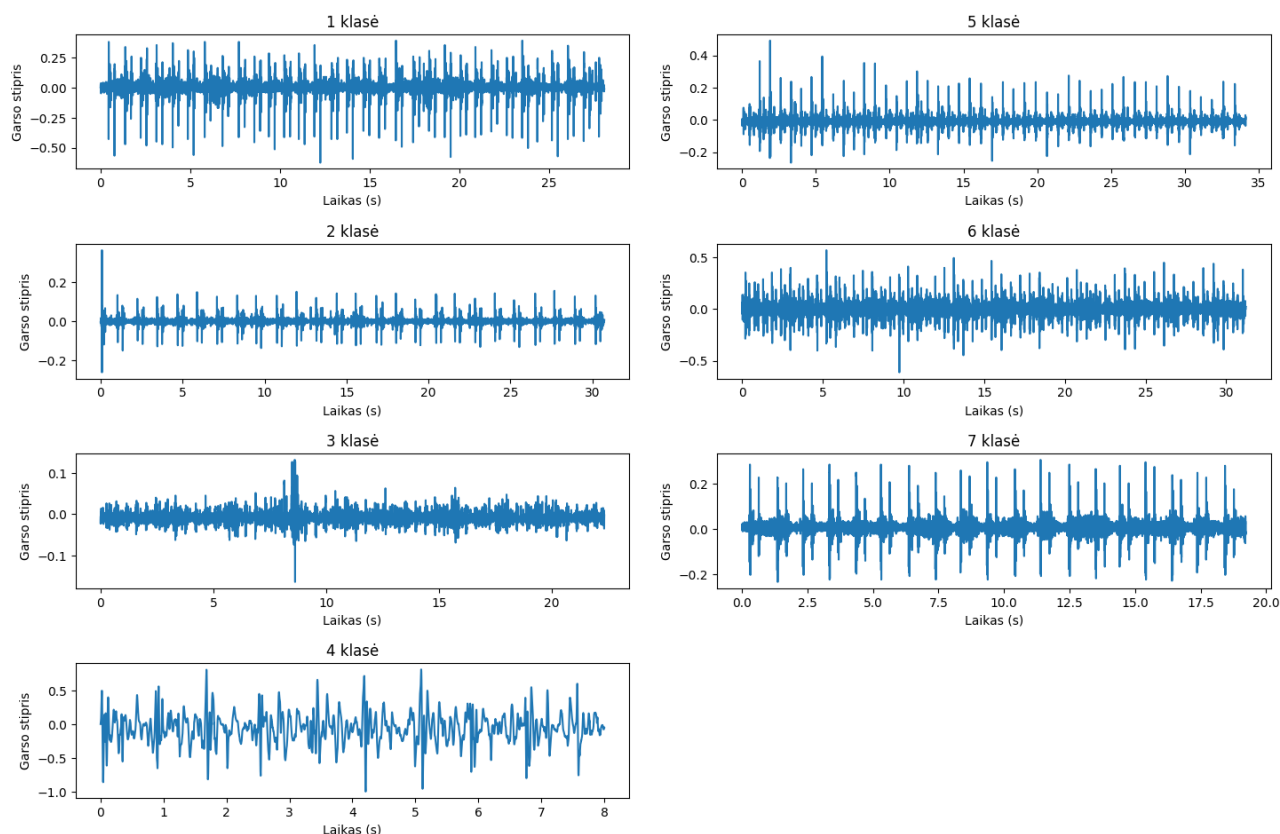
5. Kartojamas 4 žingsnis kol pasirenkama k centroidų.

4.2.3. Rezultatai

Atlikus K -vidurkių klasterizaciją 4.1 lentelėje matome klasterių pasiskirstymą. Daugiausiai įrašų priklauso 1, 5, 6 ir 7 klasteriams. Visi klasteriai turi daugiau tipinių įrašų nei netipinių. Galima priežastis tokiam šališkam duomenų pasiskirstymui klasteriuose yra klasių disbalansas duomenų rinkinyje. 20 pav.eikslėlyje vaizduojami garso įrašai esantys arčiausiai kiekvieno K -vidurkių klasterio.

4.1 lentelė. K -vidurkių klasterių pasiskirstymas

Klasterio Nr.	Elementų skaičius	Tipinių signalų skaičius	Tipinių signalų dalis, %	Netipinių signalų skaičius	Netipinių signalų dalis, %
1	894	685	77	209	23
2	155	88	57	67	43
3	54	34	63	20	37
4	42	28	67	14	33
5	767	625	81	142	19
6	628	561	89	67	11
7	695	554	80	141	20



20 pav. Garso įrašai esantys arčiausiai kiekvieno K-vidurkių klasterio.

Žvelgiant į įrašus esančius arčiausiai klasterių, matome, kad įrašai ženkliai skiriasi, tačiau sunku pasakyti ką šie skirtumai lemia ir ar kaip nors padeda atskirti tipinius signalus nuo netipinių. Siekiant aiškiau suprasti K-vidurkių klasterių skirtumus pasitelkiamos 4 garso įrašų savybės: įrašo trukmė, amplitudės maksimali reikšmė, amplitudės vidurkis, amplitudės standartinis nuokrypis ir 2 savybės gautos segmentacijos metu: RR intervalų vidurkis, RR intervalų standartinis nuokrypis. 4.2 lentelėje pateikiamos šių savybių vidutinės reikšmės kiekviename klasteryje.

4.2 lentelė. Klasterių garso įrašų savybių vidutinės reikšmės

Klasterio Nr.	Įrašo trukmė (s)	Amplitudės maksimumas	Amplitudės vidurkis	Amplitudės standartinis nuokrypis	RR intervalų vidurkis	RR intervalų standartinis nuokrypis
1	21.21	0.65004	-0.03564	0.1061	872.24	41.401
2	20.55	0.60351	-0.03847	0.1045	1219.92	60.477
3	20.26	0.61112	-0.00166	0.1019	1871.27	160.537
4	16.43	0.71492	-0.03278	0.1633	1689.11	70.357
5	22.51	0.57494	-0.01737	0.0960	726.62	35.762
6	25.53	0.54385	-0.01238	0.0785	576.93	27.853
7	22.07	0.63883	-0.03398	0.0972	1008.91	44.604

Rezultatuose matome, jog trumpiausi įrašai buvo atskirti į 4 klasterį ir šie įrašai taip pat pasižymėjo aukštesniu amplitudės standartiniu nuokrypiu. 3-iasis klasteris turi aukščiausią amplitudės vidurkį bei aukščiausias RR intervalų vidurkio ir standartinio nuokrypio reikšmes. Klasterizavimo eksperimentai

rodo, jog duomenų rinkinyje esantys duomenys gali išsiskirti ne tik kaip tipiniai ar netipiniai širdies garsai, tačiau ir pagal kitas savybes. Tokia analizė galėtų būti labai naudinga norint atfiltruoti tam tikrus duomenis pagal garso savybes, pavyzdžiui, turint tam tikrus trukmės reikalavimus, būtų galima atskirti 4 klasterio duomenis ir jų nenaudoti, kadangi šiam klasteriui priklauso šiek tiek trumpesni garsai. Žemo amplitudės standartinio nuokrypio klasterį (šiam tyrime gautas 6 klasteris) būtų galima interpretuoti kaip triukšmingų garsų klasterį ir taip pat atskirti, norint gauti kokybiškesnių garsų duomenų imtį.

4.3. Klasifikavimo eksperimentai

Klasifikavimo eksperimentai atlikti naudojant 3 skirtingus mašininio mokymo klasifikavimo metodus: SVM klasifikatorius (angl. *Support Vector Machine*), atsitiktinių miškų metodas (angl. *Random Forest*, RF) ir 2-jų sluoksnių CNN (angl. *Convolutional Neural Network*) gilusis neuroninis tinklas. Atskiri eksperimentai buvo atliekami naudojant segmentacijos metu gautas savybes, originalių duomenų MFCC savybes bei augmentedų duomenų MFCC savybes. Kiekvienam metodui buvo atlikti eksperimentai išbandant įvairias modelių parametrų kombinacijas, siekiant atrasti optimalius modelio parametrus duomenų rinkiniui ir klasifikavimo uždaviniui. Gauti modeliai testuojami pasitelkiant TEST duomenų rinkinį, kuris sudarytas iš modelio mokymui nenaudojamų garso įrašų ir VALID – „Physionet“ pateiktas rinkinys, kuris sudarytas iš kokybiškų ir aiškių garsų mokymo duomenų imtyje. Taip pat modelių ištestuojami ir 5-ių grupių kryžminės validacijos (angl. *5-Fold Cross Validation*, CV) būdu. Visų eksperimentų rezultatams pateikiamos klaidų matricos (angl. *confusion matrix*), tikslumas, F1 įvertis, specifiškumas, jautrumas ir subalansuotas tikslumas (angl. *balanced accuracy*, MAcc). Išskiriant geriausius rezultatus, VALID rezultatai nebus svarstomi, kadangi VALID imtyje egzistuoja mokymo imtyje esantys duomenys, todėl modelis visuomet rodys geresnius rezultatus. Geriausi rezultatai visuomet bus išskiriami naudojant kryžminės validacijos (CV) testavimo būdą, nes tai išsamiausias testavimo būdas bei tokiu pačiu būdu yra pateikiami rezultatai ir panašiuose sprendimuose literatūroje.

4.3.1. SVM metodas

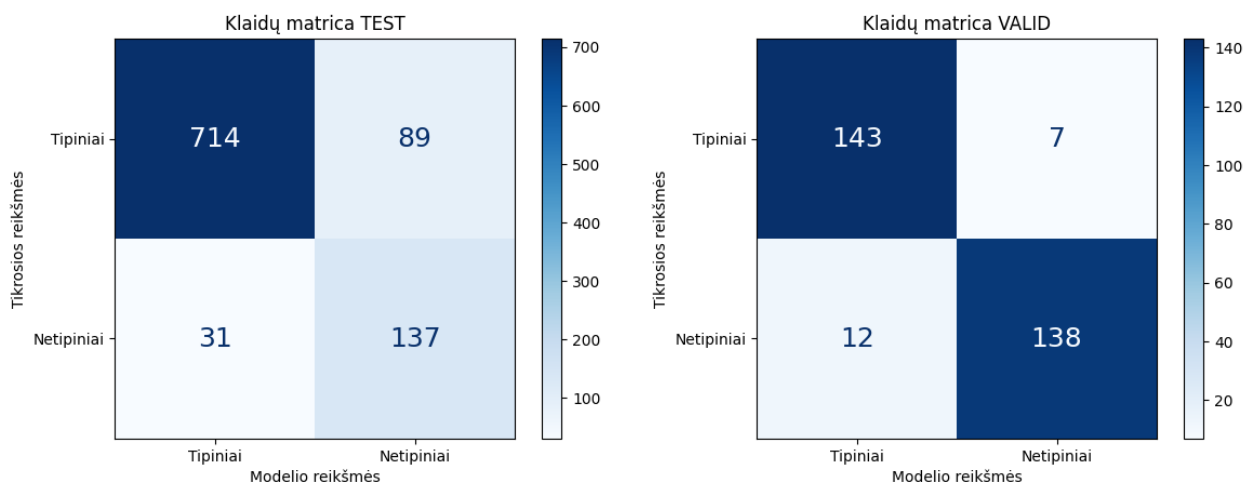
Naudojant atraminių vektorių klasifikatorių eksperimentai atliekami naudojant 4-ių parametrų kombinacijas. Branduolio funkcija (*kernel*): tiesinė (*linear*), polinominė (*poly*), radialinės bazinės funkcijos (*rbf*) ir sigmoidinė (*sigmoid*). C reikšmė: 0,1, 1, 10 ir 50. Gama reikšmė: mastelinė (*scale*), automatinė (*auto*), 0,01 ir 0,1. Laipsnio reikšmė (taikoma tik polinominei branduolio funkcijai): 2, 3, 4. Iš viso gaunama 84 skirtingos parametrų kombinacijos.

Rezultatai naudojant duomenų segmentaciją

4.3 lentelėje matomi SVM rezultatai naudojant duomenų segmentaciją. Geriausi rezultatai pasiekti naudojant radialinės bazinės funkcijos branduolio funkciją, C reikšmę 1 bei mastelinę gama reikšmę. Gautas rezultatas – 0,8547 MAcc

4.3 lentelė. SVM rezultatai naudojant duomenų segmentaciją

Testavimo būdas	Parametrai (branduolio funkcija, C, gama, laipsnis)	Tikslumas	F1 įvertis	Specifiškumas	Jautrumas	MAcc
TEST	rbf, 10, 0.01	0.8764	0.6954	0.8892	0.8155	0.8523
	rbf, 10, scale/auto	0.8929	0.7174	0.9153	0.7857	0.8505
	rbf, 1, scale	0.8311	0.6419	0.8219	0.8750	0.8485
VALID	poly, 1, 0.1, 4	0.9367	0.9356	0.9533	0.9200	0.9367
	poly, 10, 0.1, 3	0.9333	0.9338	0.9267	0.9400	0.9333
	poly, 0, 0.1, 4	0.9300	0.9278	0.9600	0.9000	0.9300
CV	rbf, 1, scale	0.8343	0.6858	0.8209	0.8884	0.8547
	rbf, 1, auto	0.8334	0.6845	0.8198	0.8884	0.8541
	rbf, 1, 0.01	0.8238	0.6725	0.8073	0.8892	0.8483



21 pav. SVM klaidų matricos naudojant segmentaciją

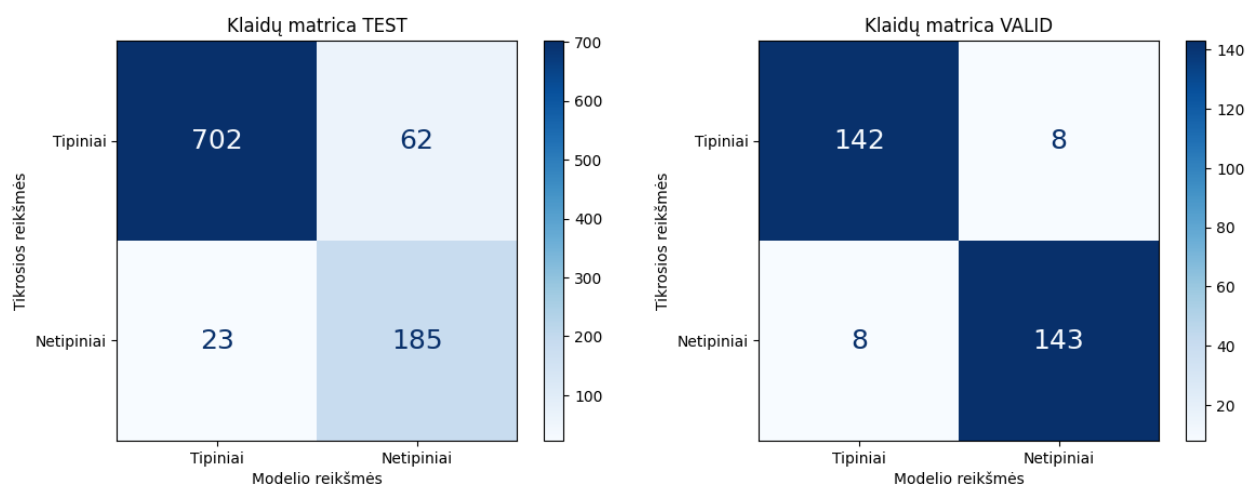
Klaidų matricose (žr. 21 pav.) matome, jog modeliui tiek TEST, tiek VALID duomenų rinkiniuose geriau sekėsi klasifikuoti tipinius garso signalus.

Rezultatai nenaudojant segmentacijos

4.4 lentelėje matome, jog nenaudojant segmentacijos SVM pavyko pasiekti šiek tiek geresnių rezultatų. Geriausi rezultatai gauti naudojant taip pat radialinės bazinės funkcijos branduolio funkciją, ta pačia C reikšme (1) ir gama reikšme 0,1. Geriausias rezultatas nenaudojant segmentacijos – 0,8935 MAcc.

4.4 lentelė. SVM rezultatai nenaudojant segmentacijos

Testavimo būdas	Parametrai (branduolio funkcija, C, gama, laipsnis)	Tikslumas	F1 įvertis	Specifiškumas	Jautrumas	MAcc
TEST	rbf, 10, scale/auto	0.9126	0.8132	0.9188	0.8894	0.9041
	rbf, 1, 0.1	0.9002	0.7949	0.8992	0.9038	0.9015
	rbf, 10, 0.1	0.9300	0.8373	0.9542	0.8413	0.8978
VALID	rbf, 50, 0.1	0.9468	0.9470	0.9467	0.9470	0.9468
	rbf, 10, 0.1	0.9203	0.9236	0.8800	0.9603	0.9201
	poly, 50, 0.1, 3	0.9136	0.9150	0.9000	0.9272	0.9136
CV	rbf, 1.0, 0.1	0.8883	0.7678	0.8843	0.9027	0.8935
	rbf, 10, auto	0.8963	0.7743	0.9035	0.8690	0.8862
	rbf, 10, scale	0.8966	0.7745	0.9042	0.8674	0.8858



22 pav. SVM klaidų matricos nenaudojant segmentacijos

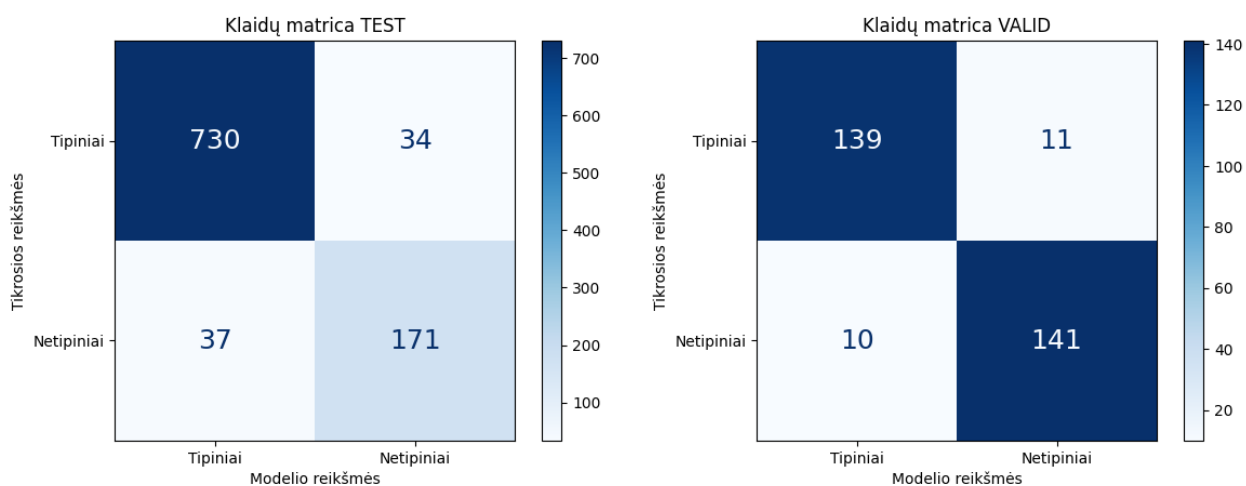
Klaidų matricose (žr. 22 pav.) taip pat matome geresnius rezultatus. TEST duomenų imtyje neteisingai suklasifikuoti 23 netipiniai signalai (kai naudojant segmentaciją buvo 31) ir 62 tipiniai signalai (kai naudojant segmentaciją buvo 89). VALID duomenų imtyje nenaudojant segmentacijos kiekviena klasė klasifikuojama ganėtinai neprastai (~0,947 tikslumu).

Rezultatai naudojant duomenų augmentaciją

Pasitelkus duomenų augmentaciją gauti dar geresni rezultatai (žr. 4.5 lentelė) Geriausi parametrai: branduolio funkcija – taip pat radialinė bazinė funkcija; C – 10, gama – 0,1. Geriausias rezultatas – 0,9392 MAcc.

4.5 lentelė. SVM rezultatai naudojant augmentaciją

Testavimo būdas	Parametrai (branduolio funkcija, C, gama, laipsnis)	Tikslumas	F1 įvertis	Specifiškumas	Jautrumas	MAcc
TEST	rbf, 10, 0.1	0.9270	0.8281	0.9555	0.8221	0.8888
	poly, 10, 0.1, 4	0.9084	0.7963	0.9280	0.8365	0.8823
	poly, 50, scale, 4	0.9043	0.7891	0.9228	0.8365	0.8797
VALID	rbf, 50, 0.1	0.9302	0.9307	0.9267	0.9338	0.9302
	poly, 50, 0.1, 4	0.9103	0.9126	0.8867	0.9338	0.9102
	poly, 50, 0.1, 3	0.9070	0.9085	0.8933	0.9205	0.9069
CV	rbf, 10, 0.1	0.9391	0.9313	0.9389	0.9395	0.9392
	rbf, 50, 0.1	0.9371	0.9291	0.9369	0.9375	0.9372
	rbf, 50, scale	0.9332	0.9249	0.9325	0.9340	0.9332



23 pav. SVM klaidų matricos naudojant augmentaciją

Klaidų matricose (žr. 23 pav.) matome geresnius rezultatus. Nors ir VALID duomenų imtyje modelis padarė daugiau klaidų nei naudojant originalų duomenų rinkinį ir segmentaciją, TEST imtyje klasifikuojant tipinius signalus buvo padaryta beveik per pus mažiau klaidų (34 neteisingai suklasifikuoti, kai prieš tai buvo 62 ir 89).

4.3.2. Atsitiktinių miškų metodas

Atsitiktinių miškų metodui taip pat svarbu atrasti geriausius parametrus ir išbandyti skirtingas parametrų kombinacijas. Čia optimizuojamos 6-ių skirtingų parametrų reikšmės: medžių skaičius miške (*n_estimators*), maksimalus medžio gylis (*max_depth*), minimalus įrašų kiekis reikalingas išskirstymui (*min_samples_split*), minimalus kiekis įrašų medžių galuose (*min_samples_leaf*), savybių kiekis geriausiai medžiui nustatyti (*max_features*), klasių balansavimas (*class_weight*). Medžių skaičius miške nusako kiek skirtingų sprendimų medžių bus sudaryta geriausiai rezultatui gauti, pasirinktos išbandyti reikšmės: 10, 50, 100. Kuo didesnis medžių skaičius, tuo ilgiau užtrunka modelio mokymas ir klasifikavimas, tačiau naudojant daugiau medžių didesnė tikimybė gauti tikslesnį rezultatą. Čia maksimalus medžių skaičius pasirinktas 100, nes naudojant daugiau medžių

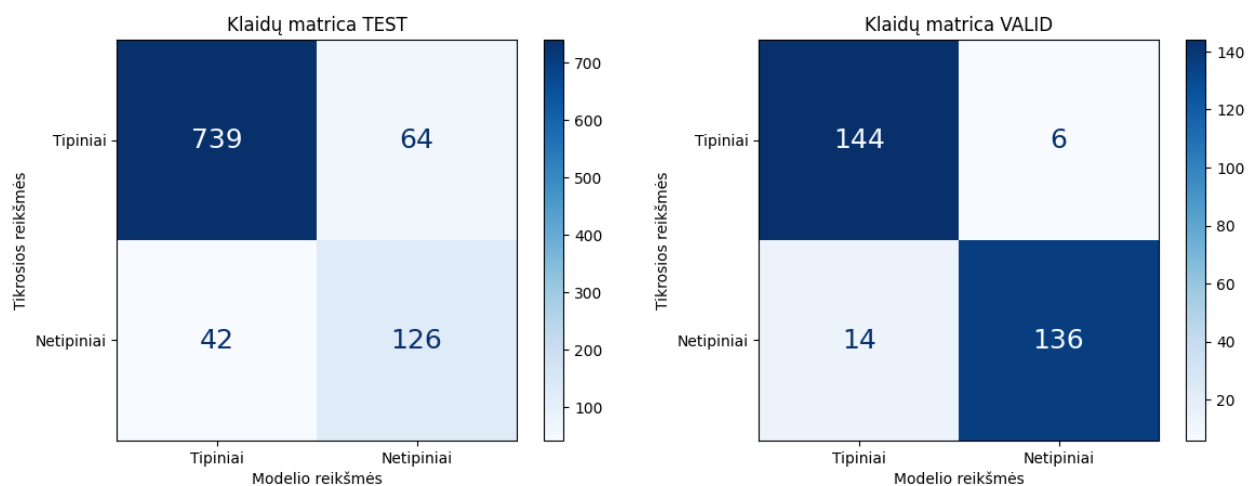
pastebėta tendencija modelio persimokymui (angl. *overfitting*). Maksimalus medžio gylis nusako, kiek kartų maksimaliai gali būti medis skirstomas į šakas, kuo šis skaičius didesnis, tuo modelis potencialiai gali būti sudėtingesnis (ilgiau trunka mokymas ir klasifikavimas). Išbandomos šios *max_depth* reikšmės: be limitu (-), 10, 20, 30. Minimalus įrašų kiekis reikalingas išskirstymui nusako kiek mažiausiai įrašų reikia, kad medis būtų toliau skirstomas į šakas, kuo šis skaičius mažesnis, tuo sudėtingesnis modelis. Pasirinkta išbandyti šias *min_samples_split* reikšmes: 2, 5, 10. Minimalus kiekis įrašų medžių galuose nusako kiek mažiausiai įrašų turi būti mazge po paskutinio medžio skirstymo, kuo mažesnė reikšmė, tuo sudėtingesnis modelis. Čia išbandomos šios *min_samples_leaf* reikšmės: 1, 2, 4. Savybių kiekis geriausiam medžiui nustatyti nusako kiek savybių yra naudojama geriausiam medžiui nustatyti. Pasirinktos išbandyti *max_features*: *sqrt* (savybių kiekio šaknis), *log2* (savybių kiekio logaritmas pagrindu 2). Klasių balansavimas nusako kuriuo būdu *scikit-learn* [31] Python biblioteka balansuoja klases. Išbandomos šios *class_weight* reikšmės: *balanced* (naudojama įrašų klasių reikšmių (tipinis, netipinis) aibė ir proporciškai priskiriami svoriai įrašams), *balanced_subsample* (veikia taip pat kaip ir *balanced*, tačiau taikomas atskirai kiekvienam savirankos (angl. *bootstrapping*) metu atrinktą poaibiui). Su tikslu išbandyti visus parametrus atliekami 432 eksperimentai naudojant kiekvieną duomenų tipą (su segmentacija, be segmentacijos, su augmentacija). Toliau pateikti eksperimentų rezultatai. 4.6 lentelė, 4.7 lentelė ir 4.8 lentelėse pateikiami parametrai tokia tvarka: *n_estimators*, *max_depth*, *min_samples_split*, *min_samples_leaf*, *max_features*, *class_weight* (B – *balanced*, BS – *balanced_subsample*).

Rezultatai naudojant duomenų segmentaciją

Naudojant atsitiktinių miškų metodą su duomenų segmentacija geriausi rezultatai (žr. 4.6 lentelė.) gauti su šiais parametrais: medžių skaičius – 100, maksimalus medžio gylis – 10, minimalus įrašų kiekis išskirstymui – 10, minimalus įrašų kiekis medžių galuose – 4, savybių kiekis geriausiam medžiui nustatyti – *sqrt*, klasių balansavimas – *balanced*. Gautas rezultatas – 0,8295 MAcc.

4.6 lentelė. Atsitiktinių miškų rezultatai naudojant segmentaciją

Testavimo būdas	Parametrai	Tikslumas	F1 įvertis	Specifiškumas	Jautrumas	MAcc
TEST	100, -, 5, 4, log2, B	0.8908	0.7039	0.9203	0.7500	0.8351
	100, 20, 2, 4, log2, B	0.8888	0.7000	0.9178	0.7500	0.8339
	100, 10, 10, 4, sqrt, BS	0.8785	0.6845	0.9029	0.7619	0.8324
VALID	50, -, 5, 1, sqrt, BS	0.9333	0.9315	0.9600	0.9067	0.9333
	50, 20, 2, 1, sqrt, B	0.9300	0.9273	0.9667	0.8933	0.9300
	50, -, 5, 1, log2, B	0.9267	0.9241	0.9600	0.8933	0.9267
CV	100, 10, 10, 4, sqrt, B	0.8816	0.7195	0.9182	0.7407	0.8295
	50, 10, 10, 4, sqrt, B	0.8810	0.7167	0.9187	0.7356	0.8271
	50, -, 10, 4, sqrt, BS	0.8915	0.7291	0.9384	0.7108	0.8246



24 pav. Atsitiktinių miškų klaidų matricos naudojant segmentaciją

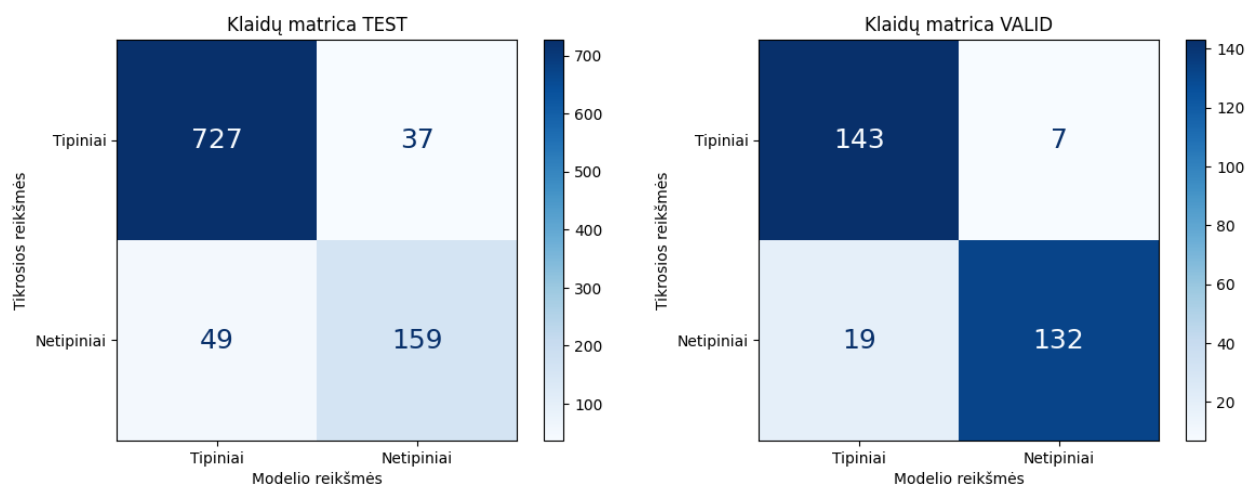
Klaidų matricose (žr. 24 pav.) matome, jog atsitiktiniams miškams naudojant segmentaciją klasifikuoti netipinius signalus sekėsi šiek tiek prasčiau nei SVM. Tačiau VALID duomenų rinkinyje tipinius signalus atsitiktiniai miškai klasifikavo aukštu tikslumu ir padarė tik 6 klaidas.

Rezultatai nenaudojant segmentacijos

Nenaudojant segmentacijos rezultatai (žr. 4.7 lentelė.) ženkliai nesiskiria, tačiau modelis šiek tiek tiksliau klasifikuoja tiek tipinius, tiek netipinius signalus. Geriausi rezultatai gauti naudojant tokius pat parametrus kaip ir naudojant segmentaciją. Geriausias rezultatas – 0,8427 MAcc.

4.7 lentelė. Atsitiktinių miškų rezultatai nenaudojant segmentacijos

Testavimo būdas	Parametrai	Tikslumas	F1 įvertis	Specifiškumas	Jautrumas	MAcc
TEST	50, -, 10, 2, sqrt, B	0.9115	0.7871	0.9516	0.7644	0.8580
	10, 10, 10, 4, sqrt, B	0.8930	0.7570	0.9241	0.7788	0.8515
	50, 10, 10, 4, sqrt, B	0.9012	0.7670	0.9398	0.7596	0.8497
VALID	50, 20, 5, 1, sqrt, B	0.9136	0.9103	0.9533	0.8742	0.9138
	100, -, 5, 2, sqrt, B	0.9070	0.9041	0.9400	0.8742	0.9071
	50, -, 5, 1, sqrt, B	0.9037	0.8997	0.9467	0.8609	0.9038
CV	100, 10, 10, 4, sqrt, BS	0.8935	0.7442	0.9285	0.7569	0.8427
	50, 10, 2, 4, sqrt, BS	0.8904	0.7379	0.9254	0.7549	0.8402
	100, 10, 2, 4, sqrt, BS	0.8929	0.7408	0.9300	0.7488	0.8394



25 pav. Atsitiktinių miškų klaidų matricos nenaudojant segmentacijos

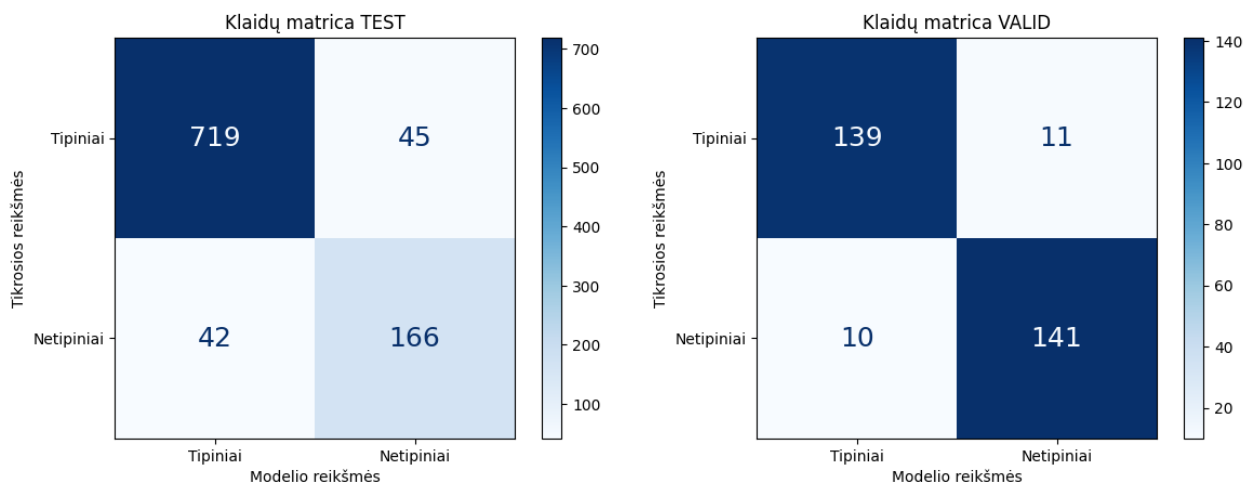
25 pav. esančiose klaidų matricose matome, kad nenaudojant segmentacijos rezultatai stipriai nepasikeičia. TEST duomenų rinkinyje geriau sekėsi klasifikuoti tipinius garso įrašus. VALID duomenų rinkinyje tiek tipiniai, tiek netipiniai signalai buvo klasifikuojami tiksliau naudojant segmentaciją.

Rezultatai naudojant duomenų augmentaciją

4.8 lentelėje matome rezultatus naudojant duomenų augmentaciją. Naudojant augmentaciją pavyko gauti gerokai geresnius rezultatus. Geriausi rezultatai gauti naudojant šiuos parametrus: medžių skaičius – 100, maksimalus medžio gylis – 20, minimalus įrašų kiekis išskirstymui – 2, minimalus įrašų kiekis medžių galuose – 1, savybių kiekis geriausiam medžiui nustatyti – *sqrt*, klasių balansavimas – *balanced_subsample*. Gautas rezultatas – 0,9255 MAcc. Naudojant augmentaciją atsitiktinių miškų modelis klasifikuoja netipinius širdies plakimo garsus net 20% tiksliau ir gautas beveik 10% geresnis balansuotas tikslumas.

4.8 lentelė. Atsitiktinių miškų rezultatai naudojant augmentaciją

Testavimo būdas	Parametrai	Tikslumas	F1 įvertis	Specifiškumas	Jautrumas	MAcc
TEST	100, -, 5, 2, sqrt, B	0.9105	0.7924	0.9411	0.7981	0.8696
	100, -, 10, 4, sqrt, BS	0.9043	0.7822	0.9319	0.8029	0.8674
	100, 10, 10, 2, sqrt, BS	0.9033	0.7804	0.9306	0.8029	0.8668
VALID	50, 20, 2, 1, sqrt, B	0.9302	0.9307	0.9267	0.9338	0.9302
	50, -, 2, 1, sqrt, B	0.9269	0.9272	0.9267	0.9272	0.9269
	10, -, 2, 1, sqrt, BS	0.9236	0.9231	0.9333	0.9139	0.9236
CV	100, 20, 2, 1, sqrt, BS	0.9271	0.9166	0.9409	0.9102	0.9255
	100, 20, 5, 2, sqrt, BS	0.9267	0.9165	0.9373	0.9135	0.9254
	100, -, 2, 1, sqrt, BS	0.9265	0.9158	0.9420	0.9073	0.9247



26 pav. Atsitiktinių miškų klaidų matricos naudojant augmentaciją

Nors ir kryžminės validacijos rezultatas naudojant augmentaciją ženkiai tikslesnis, klaidų matricose (žr. 26 pav.) tokių stiprių pokyčių nesimato. TEST duomenų imtyje naudojant augmentaciją šiek tiek tiksliau klasifikuojami netipiniai širdies garsai, o VALID duomenų imtyje tiek tipiniai, tiek netipiniai signalai klasifikuojami panašiu tikslumu kaip ir naudojant augmentacijos.

4.3.3. CNN metodas

Eksperimentams naudojamo giliojo dirbtinio neuroninio tinklo struktūrą sudaro 2 dvimačiai konvoliuciniai sluoksniai požymių išskyrimui, 2 pilnai sujungti sluoksniai duomenų šablonams atpažinti ir sudaryti bei sigmoidinis sluoksnis rezultato transformacijai į binarinį tipą (tik tipinis arba netipinis garso signalas). Konvoliuciniams ir pilnai sujungtiems sluoksniams naudojama ReLU aktyvacijos funkcija.

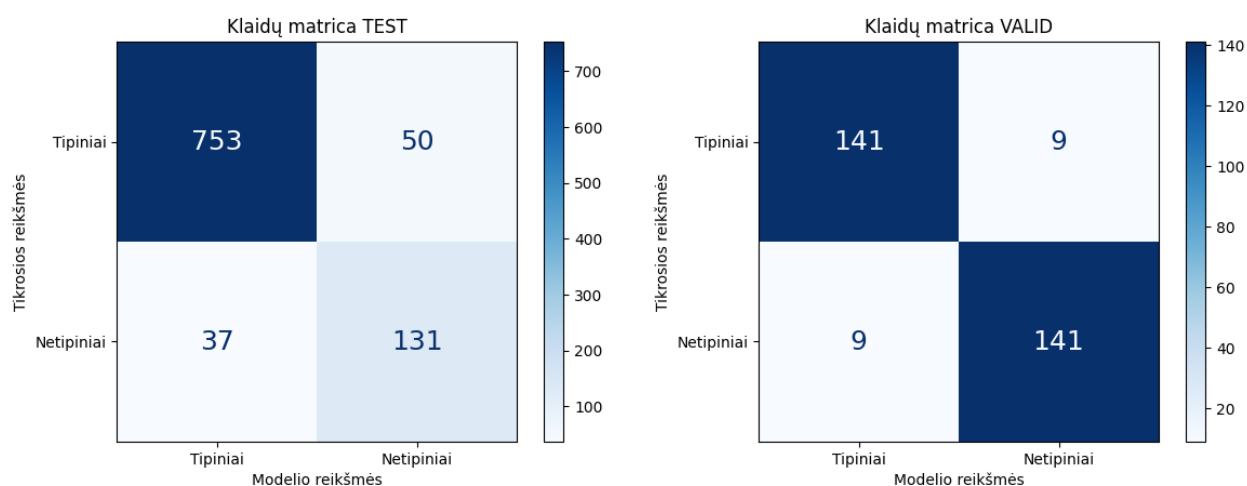
Atliekant CNN eksperimentus optimizuojami 4 skirtingi parametrai: iteracijų kiekis (*epochs*), mokymosi greitis (*learning rate*), optimizatorius (*optimizer*) ir nuostolių funkcija (*criterion*). Išbandomi 3 skirtingi iteracijų kiekiai: 25, 50 ir 100. Išbandomos mokymosi spartos reikšmės: 0,1, 0,01, 0,001. Pasirinkti išbandyti šie optimizatoriai: *Adafactor*, *Adam*, *SGD* (*stochastic gradient descent*). Pasirinktos nuostolių funkcijos: *BCELoss* (*binary cross entropy loss*), *MSELoss* (*mean square error loss*), *HuberLoss*. Iš viso gaunama 81 kombinacija.

Rezultatai naudojant duomenų segmentaciją

Naudojant duomenų segmentaciją geriausi rezultatai (žr. 4.9 lentelė.) gauti naudojant 100 iteracijų, 0,1 mokymosi spartą, SGD optimizatorių ir *HuberLoss* nuostolių funkciją. Gautas rezultatas – 0,8465. CNN metodu pavyko gauti geresnius rezultatus nei atsitiktinių miškų metodu, tačiau prastesnius nei SVM metodu.

4.9 lentelė. CNN rezultatai naudojant segmentaciją

Testavimo būdas	Parametrai (<i>epochs, learning rate, optimizer, criterion</i>)	Tikslumas	F1 įvertis	Specifiškumas	Jautrumas	MAcc
TEST	25, 0.1, Adafactor, BCELoss	0.9197	0.7706	0.9489	0.7798	0.8644
	25, 0.01, Adafactor, MSELoss	0.9145	0.7608	0.9415	0.7857	0.8636
	100, 0.1, SGD, BCELoss	0.9104	0.7535	0.9352	0.7917	0.8635
VALID	25, 0.1, Adafactor, BCELoss	0.9600	0.9605	0.9467	0.9733	0.9600
	25, 0.1, Adafactor, MSELoss	0.9367	0.9360	0.9467	0.9267	0.9367
	25, 0.1, Adafactor, HuberLoss	0.9500	0.9498	0.9533	0.9467	0.9500
CV	100, 0.1, SGD, HuberLoss	0.8964	0.7493	0.9321	0.7608	0.8465
	100, 0.1, Adafactor, HuberLoss	0.8983	0.7523	0.9347	0.7582	0.8464
	50, 0.1, Adafactor, MSELoss	0.9005	0.7529	0.9406	0.7453	0.8429



27 pav. CNN klaidų matricos naudojant segmentaciją

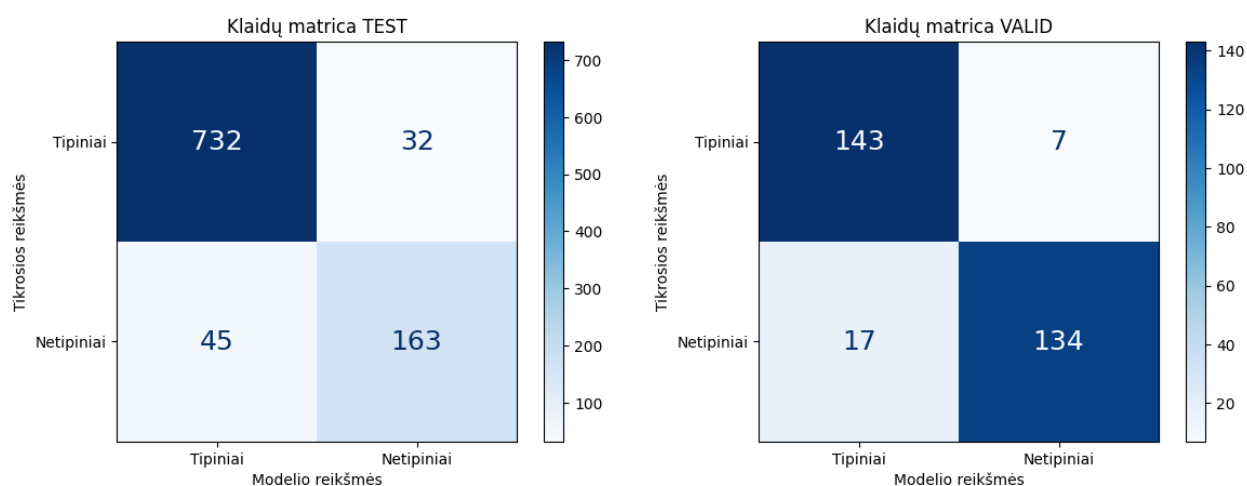
Žiūrint į klaidų matricas (27 pav.) matome, kad CNN šiek tiek geriau susitvarko su nesubalansuotais duomenimis nei SVM ar RF ir VALID duomenų rinkinyje tik 9 įrašai kiekvienoje klasėje buvo suklasifikuoti neteisingai.

Rezultatai nenaudojant segmentacijos

Naudojant originalius duomenis ir MFCC reikšmes kaip įvestis CNN geriausius rezultatus (žr. 4.10 lentelė.) pasiekė naudojant šiuos parametrus: 50 iteracijų, 0,01 mokymosi spartą, *Adafactor* optimizatorių ir *BCELoss* nuostolių funkciją. Nenaudojant segmentacijos modelis įrašus klasifikavo šiek tiek tiksliau – 0,87 MAcc.

4.10 lentelė. CNN rezultatai nenaudojant segmentacijos

Testavimo būdas	Parametrai (<i>epochs, learning rate, optimizer, criterion</i>)	Tikslumas	F1 įvertis	Specifiškumas	Jautrumas	MAcc
TEST	50, 0.01, Adam, MSELoss	0.9187	0.8132	0.9437	0.8269	0.8853
	100, 0.001, Adam, HuberLoss	0.9239	0.8177	0.9581	0.7981	0.8781
	100, 0.01, Adafactor, MSELoss	0.9239	0.8159	0.9607	0.7885	0.8746
VALID	25, 0.1, Adafactor, BCELoss	0.9369	0.9352	0.9667	0.9073	0.9370
	100, 0.001, Adam, BCELoss	0.9369	0.9360	0.9533	0.9205	0.9369
	100, 0.1, Adafactor, HuberLoss	0.9302	0.9288	0.9533	0.9073	0.9303
CV	50, 0.01, Adafactor, BCELoss	0.9173	0.7963	0.9502	0.7897	0.8700
	100, 0.01, Adafactor, BCELoss	0.9077	0.7810	0.9346	0.8018	0.8682
	100, 0.01, Adam, BCELoss	0.9117	0.7847	0.9440	0.7858	0.8649



28 pav. CNN klaidų matricos nenaudojant segmentacijos

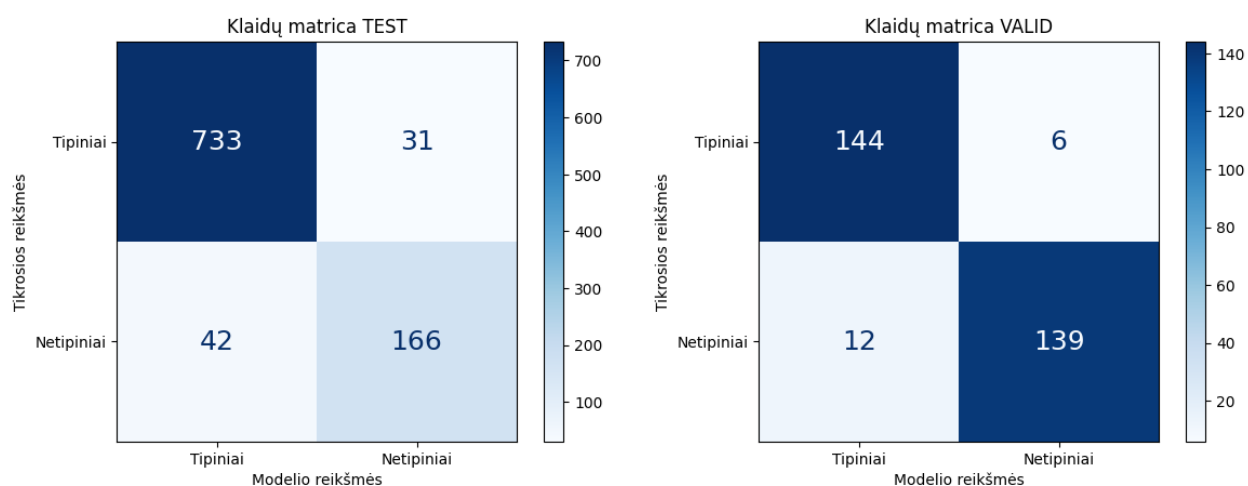
Klaidų matricose (žr. 28 pav.) matome, jog nors ir bendrame rezultate įrašai buvo klasifikuojami tiksliau nei naudojant segmentaciją, tiek TEST, tiek VALID duomenų rinkiniuose modelis buvo labiau paveiktas duomenų disbalanso ir tiksliau klasifikavo tipinius širdies plakimo garso signalus, o netipinių signalų klasifikavimo tikslumas buvo prastesnis.

Rezultatai naudojant duomenų augmentaciją

Pašalinus duomenų disbalansą naudojant duomenų augmentaciją gauti ženkliai geresni CNN klasifikavimo rezultatai (žr. 4.11 lentelė.). Geriausius rezultatus pavyko pasiekti naudojant šiuos parametrus: 100 iteracijų, 0,1 mokymosi sparta, SGD optimizatorius ir *BCELoss* nuostolių funkcija. Čia gautas geriausias rezultatas visų eksperimentų metu – 0,94 MAcc.

4.11 lentelė. CNN rezultatai naudojant augmentaciją

Testavimo būdas	Parametrai (<i>epochs, learning rate, optimizer, criterion</i>)	Tikslumas	F1 įvertis	Specifiškumas	Jautrumas	MAcc
TEST	100, 0.01, Adafactor, BCELoss	0.9259	0.8269	0.9529	0.8269	0.8899
	100, 0.001, Adam, HuberLoss	0.9177	0.8104	0.9437	0.8221	0.8829
	50, 0.01, Adafactor, MSELoss	0.9228	0.8175	0.9542	0.8077	0.8809
VALID	25, 0.1, Adafactor, BCELoss	0.9502	0.9498	0.9600	0.9404	0.9502
	100, 0.1, Adafactor, BCELoss	0.9435	0.9424	0.9667	0.9205	0.9436
	50, 0.1, Adafactor, BCELoss	0.9402	0.9388	0.9667	0.9139	0.9403
CV	100, 0.1, SGD, BCELoss	0.9401	0.9325	0.9415	0.9385	0.9400
	50, 0.1, Adafactor, MSELoss	0.9401	0.9325	0.9417	0.9382	0.9400
	50, 0.01, Adafactor, BCELoss	0.9371	0.9294	0.9357	0.9392	0.9375



29 pav. CNN klaidų matricos naudojant augmentaciją

Pažvelgus į klaidų matricas (žr. 29 pav.) matoma, kad naudojant augmentaciją tiek TEST, tiek VALID duomenų rinkinyje įrašai klasifikuojami šiek tiek tiksliau, o VALID rinkinyje itin gerai klasifikuojami tipiniai garso įrašai.

4.3.4. Klasifikavimo eksperimentų rezultatų apibendrinimas

4.12 lentelėje apžvelgiami geriausi rezultatai gauti naudojant skirtingus mašininio mokymosi modelius ir įvestis (AUG – naudojant augmentaciją, ORG – nenaudojant segmentacijos, SEG – naudojant segmentaciją). Lentelėje pateikiami tik kryžminės validacijos rezultatai. Geriausią rezultatą pavyko pasiekti naudojant 2-jų sluoksnių CNN gilųjų dirbtinį neuroninį tinklą (0,94 MAcc). Geriausias rezultatas gautas naudojant augmentacijos pagalba subalansuotą duomenų rinkinį. Visi trys modeliai geriau klasifikavimo širdies plakimo garso įrašus naudojant subalansuotą duomenų rinkinį su sintetiniais duomenimis. Naudojant augmenteduotą duomenų rinkinį CNN klasifikuoja širdies plakimo garso įrašus maždaug 8% tiksliau, atsitiktinių miškų metodas ~10% tiksliau, o SVM apie 5% tiksliau.

4.12 lentelė. Bendri rezultatai

Metodas	Įvestis	Tikslumas	F1 įvertis	Specifiškumas	Jautrumas	MAcc
CNN	AUG	0.9401	0.9325	0.9415	0.9385	0.9400
	ORG	0.9173	0.7963	0.9502	0.7897	0.8700
	SEG	0.8964	0.7493	0.9321	0.7608	0.8465
RF	AUG	0.9271	0.9166	0.9409	0.9102	0.9255
	ORG	0.8935	0.7442	0.9285	0.7569	0.8427
	SEG	0.8816	0.7195	0.9182	0.7407	0.8295
SVM	AUG	0.9391	0.9313	0.9389	0.9395	0.9392
	ORG	0.8883	0.7678	0.8843	0.9027	0.8935
	SEG	0.8343	0.6858	0.8209	0.8884	0.8547

4.4. Palyginimas su egzistuojančiais sprendimais

Lyginant gautus rezultatus su egzistuojančiais sprendimais (4.13 lentelė), pavyko gauti tiksliau klasifikuojančius modelius nei Chen et al. [19] (0,9391 MAcc) tiek naudojant CNN metodą (0,94 MAcc), tiek naudojant SVM (0,9392 MAcc).

4.13 lentelė. Gautų rezultatų palyginimas su egzistuojančiais sprendimais

Šaltinis	Duomenų rinkinys	Klasifikavimo metodai	Naudoti parametrai	Tikslumas (mAcc)
Alkhodari et al., [2], 2021	„Physionet Challenge 2016“	dvikryptis Long Short-Term Memory (BiLSTM) rekurentinis neuroninis tinklas (RNN)	lr = 0.01; L2-regularization = 0.0001	0,8731
Shuvo et al., [12], 2021	„Physionet Challenge 2016“	CardioXNet (DNN)	lr = 0.00001; Adam optimizatorius	0,8657
Chen et al., [19], 2020	„Physionet Challenge 2016“	2 CNN modeliai	lr = 0.001	0,9391
Šis darbas	„Physionet Challenge 2016“ praplėstas augmentacijomis	2-jų sluoksnių CNN	100 iteracijų, 0,1 mokymo sparta, SGD optimizatorius, BCELoss nuostolių funkcija	0,9400
Šis darbas	„Physionet Challenge 2016“ praplėstas augmentacijomis	SVM klasifikatorius	branduolio funkcija – radialinė bazinė funkcija (rbf); C – 10, gama – 0,1	0,9392

Išvados

1. Literatūroje dažniausiai pasirenkami metodai fonokardiogramų klasifikavimui yra įvairių struktūrų CNN paremti gilieji dirbtiniai neuroniniai tinklai ir atraminių vektorių klasifikatorius, duomenų išskyrimui naudojami MFCC koeficientai ir CNN dirbtinių neuronų sluoksniai, o duomenims dažniausiai naudojamas „2016 CinC/PhysioNet Challenge“ duomenų rinkinys.
2. Atlikus egzistuojančių sprendimų analizę atrastos pagrindinės esamų sprendimų problemos: mokymas su ribotu kiekiu širdies garso duomenų, mokymo efektyvumas ir sparta, giliesiems dirbtiniams neuroniniams tinklams mokyti reikalingi atminties ir skaičiavimo išteklių.
3. Atliekant duomenų analizę klasterizavimo tyrimas parodė, jog duomenyse yra daugiau nei dviejų tipų būdinguosius požymius turinčios grupės (klasteriai) ir kai kurie klasteriai gali padėti geriau atskirti kokybiškai įrašytus širdies garsus nuo nekokybiškų garsų su daug triukšmo ar trikdžių. Taip pat atrastas duomenų disbalansas (tipinių širdies garsų duomenų imtyje 3,9 karto daugiau nei netipinių). Duomenų apdorojimui, skirtingai nuo eksperimento tipo, naudojamas širdies garsų segmentacijos procesas arba apskaičiuojamos vidutinės MFCC kadru reikšmės.
4. Sukurtas algoritmas fonokardiogramų klasifikavimui naudoja 3 skirtingus modelius: atraminių vektorių (SVM), atsitiktinių miškų (RF) bei 2 sluoksnių CNN paremtą giliųjų dirbtinių neuronų tinklą. Algoritmas priima segmentacijos būdu gautas garso įrašo savybes arba MFCC vidurkius kaip įvestį modeliams mokyti. Taip pat atliekama duomenų augmentacija ir sukuriama sintetiniai netipiniai duomenys klasių disbalansui duomenų imtyje pašalinti.
5. Atlikus klasifikavimo eksperimentus, geriausi 5 grupių kryžminės validacijos rezultatai gauti naudojant CNN metodą (0,94 MAcc). Panaikinus duomenų disbalansą modeliai klasifikavo garso įrašus 5-10% tiksliau. O lyginant gautus rezultatus su literatūroje egzistuojančiais sprendimais tiek CNN, tiek SVM modeliams pavyko pranokti geriausią apžvelgtoje literatūroje egzistuojantį sprendimą.

Literatūros sąrašas

1. Cardiovascular diseases (CVDs). URL: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
2. ALKHODARI, M., FRAIWAN, L. Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings. In *Computer Methods and Programs in Biomedicine*. 2021. Vol. 200, p. 105940. doi: 10.1016/j.cmpb.2021.105940
3. LIU, C., MURRAY, A. Applications of Complexity Analysis in Clinical Heart Failure. BARBIERI, R. ir kt. Sud. In *Complexity and Nonlinearity in Cardiovascular Signals*. Cham: Springer International Publishing, 2017. p. 301–325. ISBN 978-3-319-58709-7. doi: 10.1007/978-3-319-58709-7_11
4. CLIFFORD, G.D. ir kt. Recent advances in heart sound analysis. In *Physiological Measurement*. 2017. Vol. 38, no. 8, p. E10. doi: 10.1088/1361-6579/aa7ec8
5. CHEN, W. ir kt. Deep Learning Methods for Heart Sounds Classification: A Systematic Review. In *Entropy*. 2021. Vol. 23, no. 6, p. 667. doi: 10.3390/e23060667
6. LIU, C. ir kt. An open access database for the evaluation of heart sound algorithms. In *Physiological Measurement*. 2016. Vol. 37, no. 12, p. 2181. doi: 10.1088/0967-3334/37/12/2181
7. CLIFFORD, G.D. ir kt. Classification of normal/abnormal heart sound recordings: The PhysioNet/Computing in Cardiology Challenge 2016. In *2016 Computing in Cardiology Conference (CinC)*. 2016. p. 609–612.
8. Classifying Heart Sounds Challenge. URL: <http://www.peterjbentley.com/heartchallenge/>
9. OLIVEIRA, J. ir kt. The CirCor DigiScope Dataset: From Murmur Detection to Murmur Classification. In *IEEE Journal of Biomedical and Health Informatics*. 2022. Vol. 26, no. 6, p. 2524–2535. doi: 10.1109/JBHI.2021.3137048
10. OLIVEIRA, JORGE ir kt. [s.l.]: PhysioNet. doi: 10.13026/TSHS-MW03
11. KRISHNAN, P.T. ir kt. Automated heart sound classification system from unsegmented phonocardiogram (PCG) using deep neural network. In *Physical and Engineering Sciences in Medicine*. 2020. Vol. 43, no. 2, p. 505–515. doi: 10.1007/s13246-020-00851-w
12. SHUVO, S.B. ir kt. CardioXNet: A Novel Lightweight Deep Learning Framework for Cardiovascular Disease Classification Using Heart Sound Recordings. In *IEEE Access*. 2021. Vol. 9, p. 36955–36967. doi: 10.1109/ACCESS.2021.3063129
13. DEMIR, F. ir kt. Towards the classification of heart sounds based on convolutional deep neural network. In *Health Information Science and Systems*. 2019. Vol. 7, no. 1, p. 16. doi: 10.1007/s13755-019-0078-0
14. HASAN, M. ir kt. Speaker Identification Using Mel Frequency Cepstral Coefficients. In *Proceedings of the 3rd International Conference on Electrical and Computer Engineering (ICECE 2004)*. 2004.
15. MOURAD, T. *The Stationary Bionic Wavelet Transform and its Applications for ECG and Speech Processing*. Cham: Springer International Publishing, 2022. ISBN 978-3-030-93404-0

16. ISTRATE, D. ir kt. Information extraction from sound for medical telemonitoring. In *IEEE Transactions on Information Technology in Biomedicine*. 2006. Vol. 10, no. 2, p. 264–274. doi: 10.1109/TITB.2005.859889
17. HAN, W. ir kt. An efficient MFCC extraction method in speech recognition. In *2006 IEEE International Symposium on Circuits and Systems (ISCAS)*. 2006. p. 4 pp.-. doi: 10.1109/ISCAS.2006.1692543
18. ZABIHI, M. ir kt. Heart sound anomaly and quality detection using ensemble of neural networks without segmentation. In *2016 Computing in Cardiology Conference (CinC)*. 2016. p. 613–616.
19. CHEN, Y. ir kt. Classification of heart sounds based on the combination of the modified frequency wavelet transform and convolutional neural network. In *Medical & Biological Engineering & Computing*. 2020. Vol. 58, no. 9, p. 2039–2047. doi: 10.1007/s11517-020-02218-5
20. YADAV, A. ir kt. Machine learning-based classification of cardiac diseases from PCG recorded heart sounds. In *Neural Computing and Applications*. 2020. Vol. 32, no. 24, p. 17843–17856. doi: 10.1007/s00521-019-04547-5
21. DOMINGUEZ-MORALES, J.P. ir kt. Deep Neural Networks for the Recognition and Classification of Heart Murmurs Using Neuromorphic Auditory Sensors. In *IEEE Transactions on Biomedical Circuits and Systems*. 2018. Vol. 12, no. 1, p. 24–34. doi: 10.1109/TBCAS.2017.2751545
22. WU, J.M.-T. ir kt. Applying an ensemble convolutional neural network with Savitzky–Golay filter to construct a phonocardiogram prediction model. In *Applied Soft Computing*. 2019. Vol. 78, p. 29–40. doi: 10.1016/j.asoc.2019.01.019
23. LECUN, Y. ir kt. Deep learning. In *Nature*. 2015. Vol. 521, no. 7553, p. 436–444. doi: 10.1038/nature14539
24. SPRINGER, DAVID [s.l.]: PhysioNet. doi: 10.13026/VNT9-KF93
25. DORNBUSH, S., TURNQUEST, A.E. Physiology, Heart Sounds. In *StatPearls*. Treasure Island (FL): StatPearls Publishing, 2025. URL: <http://www.ncbi.nlm.nih.gov/books/NBK541010/>
26. SCHMIDT, S.E. ir kt. Segmentation of heart sound recordings by a duration-dependent hidden Markov model. In *Physiological Measurement*. 2010. Vol. 31, no. 4, p. 513. doi: 10.1088/0967-3334/31/4/004
27. CHOI, S., JIANG, Z. Comparison of envelope extraction algorithms for cardiac sound signal segmentation. In *Expert Systems with Applications*. 2008. Vol. 34, no. 2, p. 1056–1069. doi: 10.1016/j.eswa.2006.12.015
28. BAARSCH, J., CELEBI, M.E. Investigation of Internal Validity Measures for K-Means Clustering. In *Hong Kong*. 2012.
29. k-means clustering - MATLAB kmeans - MathWorks Nordic. URL: <https://se.mathworks.com/help/stats/kmeans.html#bueftl4-1>
30. ARTHUR, D., VASSILVITSKII, S. K-means++ the advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. 2007. p. 1027–1035.

31. scikit-learn: machine learning in Python — scikit-learn 1.6.1 documentation. URL: <https://scikit-learn.org/stable/>